



**PODER EXECUTIVO  
MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO AMAZONAS  
INSTITUTO DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA**



**ANTECIPANDO A USABILIDADE NAS FASES INICIAIS DO  
PROCESSO DE DESENVOLVIMENTO DE SOFTWARE**

**NATASHA MALVEIRA COSTA VALENTIM**

Manaus, Março de 2017

**NATASHA MALVEIRA COSTA VALENTIM**

**ANTECIPANDO A USABILIDADE NAS FASES INICIAIS DO  
PROCESSO DE DESENVOLVIMENTO DE SOFTWARE**

Tese de Doutorado submetida ao corpo docente do Programa de Pós-Graduação em Informática da Universidade Federal do Amazonas (PPGI-UFAM).

Orientadora: Prof<sup>a</sup>. Tayana Uchôa Conte, D.Sc.

Manaus, Março de 2017

## Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

V155a Valentim, Natasha Malveira Costa  
Antecipando a usabilidade nas fases iniciais do processo de desenvolvimento de software / Natasha Malveira Costa Valentim. 2017  
249 f.: il. color; 31 cm.

Orientadora: Tayana Uchôa Conte  
Tese (Doutorado em Informática) - Universidade Federal do Amazonas.

1. Usabilidade cedo. 2. Usabilidade nos estágios iniciais do processo de d. 3. Modelo de Projeto. 4. Modelo de Análise. 5. Integração de IHC e ES. I. Conte, Tayana Uchôa II. Universidade Federal do Amazonas III. Título



PODER EXECUTIVO  
MINISTÉRIO DA EDUCAÇÃO  
INSTITUTO DE COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA



## FOLHA DE APROVAÇÃO

**"Antecipando a Usabilidade nas Fases Iniciais do Processo de Desenvolvimento de Software"**

**NATASHA MALVEIRA COSTA VALENTIM**

Tese de Doutorado defendida e aprovada pela banca examinadora constituída pelos Professores:

Profa. Tayana Uchôa Conte - PRESIDENTE

Prof. Altigran Soares da Silva - MEMBRO INTERNO

Prof. Raimundo da Silva Barreto - MEMBRO INTERNO

Profa. Simone Diniz Junqueira Barbosa - MEMBRO EXTERNO

Prof. Gleison dos Santos Souza - MEMBRO EXTERNO

Manaus, 24 de Março de 2017

À Deus,  
que tem me dado sabedoria e amor à pesquisa e ao ensino ao longo dessa jornada.

## AGRADECIMENTOS

A Deus, o dono da Ciência e Autor da minha vida. Em cada momento sentia a Sua presença me orientando a não me contentar com o básico e me impulsionando a ir além. Meu deu forças em um dos momentos mais difíceis (falecimento da vó). Grata sou à Deus!

Ao meu esposo Diego Valentim, que sempre me deu uma palavra de Deus diante das minhas dificuldades e medos. Cuidou de mim em cada momento, me dando todo o suporte necessário para que eu pudesse me dedicar ao máximo. Eu te amarei para sempre!

À toda a minha família e amigos, que sempre acreditaram em mim. Aos meus avós Lenir (*in memoriam*) e Jorge que me deram todo incentivo. Ao meu pai Jonir, que sempre investiu em meus estudos e sonhou comigo cada passo dessa jornada. À minha mãe Débora, por torcer e se alegrar com minhas vitórias. Aos meus tios por sempre orarem por mim, em especial à minha tia Jolene, por todo seu amor e incentivo. Aos meus irmãos, primos e madrastra por todo o afeto e auxílio nesta caminhada. Aos meus sogros, por terem orado para que Deus abençoasse a minha vida acadêmica. Ao meu pastor Aroldo e sua família, por me acompanharem espiritualmente e por me levarem pra mais perto de Deus. À minha amada igreja IBAPE, que sempre me apoiou através de suas orações.

À minha querida orientadora Tayana Conte que me orientou de forma brilhante em cada passo dessa caminhada! Como aprendi com você! Minhas vitórias e conquistas no meio acadêmico são suas também, pois tudo que aprendi devo a você! Muito obrigada por toda dedicação a este trabalho, sua confiança e seu apoio constante!

À Dra. Simone Barbosa, Dr. Gleison Santos, Dr. Altigran Soares, e Dr. Raimundo Barreto por aceitarem participar de minha banca. Eu me sinto muitíssimo honrada!

Ao grupo USES, pela cooperação e compartilhamento de conhecimento em nossas reuniões. Em especial, ao Williamson Silva por todas as discussões em prol da pesquisa.

A todos os que participaram da realização dos estudos apresentados nesta tese.

À todos os pesquisadores que me ajudaram a formar a pesquisadora que hoje sou, em especial, ao José Carlos Maldonado, Alessandro Garcia, Clarisse de Souza, Simone Barbosa, Rafael Prikladnick, Sabrina Marczak, Milene Silveira, Tiago Silva, Guilherme Travassos, Gleison Santos, Cristiano Maciel, Bernardo José, Roberto Oliveira e Leonardo Silva. Ao Rafael Prikladnick da TECNOPUC, e Márcio Lira, Daniel Tadeu e Stones Junior da PRODAM por terem aberto as portas da indústria de *software* para mim.

À UFAM, ao Icomp, à CAPES e FAPEAM por todo apoio financeiro.

## SUMÁRIO

Lista de figuras .....	IX
Lista de tabelas .....	XII
Resumo .....	XIV
Abstract .....	XV
<b>CAPÍTULO 1 - INTRODUÇÃO .....</b>	<b>1</b>
1.1. Contexto .....	1
1.2. Motivações .....	3
1.3. Definição do problema .....	4
1.4. Objetivo de pesquisa .....	5
1.5. Metodologia baseada em evidências .....	6
1.6. Principais contribuições .....	8
1.7. Outros resultados .....	10
1.8. Organização .....	13
<b>CAPÍTULO 2 - AVALIAÇÃO DE USABILIDADE NO PROCESSO DE DESENVOLVIMENTO DE APLICAÇÕES .....</b>	<b>15</b>
2.1. Introdução.....	15
2.2. Tipos de Avaliação de Usabilidade .....	15
2.3. Avaliação de Usabilidade nas Fases Iniciais do Processo de Desenvolvimento .....	20
2.4. Síntese do Capítulo.....	24
<b>CAPÍTULO 3 - MAPEAMENTO SISTEMÁTICO SOBRE TECNOLOGIAS QUE APOIAM A INTEGRAÇÃO ENTRE AS ÁREAS DE ENGENHARIA DE SOFTWARE E INTERAÇÃO HUMANO-COMPUTADOR.....</b>	<b>26</b>
3.1. Introdução.....	26
3.2. Protocolo do Mapeamento Sistemático .....	27
3.3. Resultados obtidos.....	33
3.4. Síntese do Capítulo.....	40
<b>CAPÍTULO 4 - MAPEAMENTO SISTEMÁTICO SOBRE TECNOLOGIAS QUE AUXILIAM NO PROJETO/CRIAÇÃO OU AVALIAÇÃO DE MODELOS VISANDO à USABILIDADE DA APLICAÇÃO .....</b>	<b>41</b>
4.1. Introdução.....	41
4.2. Protocolo do Mapeamento Sistemático .....	42
4.3. Resultados obtidos.....	51
4.4. Combinação dos Resultados das Subquestões .....	64
4.5. Ameaças à Validade .....	67
4.6. Síntese do Capítulo.....	68
<b>CAPÍTULO 5 - CONJUNTO DE TECNOLOGIAS QUE APOIA A ANTECIPAÇÃO DA USABILIDADE NAS FASES INICIAIS DO PROCESSO DE DESENVOLVIMENTO DE SOFTWARE.....</b>	<b>70</b>
5.1. Introdução.....	70
5.2. Proposta de Tecnologias que auxiliam na Avaliação de Modelos utilizados nas Fases Iniciais do Processo de Desenvolvimento, visando à Usabilidade da Aplicação.....	72

5.3	Proposta de Tecnologias que auxiliam no Projeto/Criação de Modelos utilizados nas Fases Iniciais do Processo de Desenvolvimento de <i>Software</i> , visando à Usabilidade da Aplicação .....	77
5.4	Síntese do Capítulo.....	85
<b>CAPÍTULO 6 - ESTUDOS EXPERIMENTAIS DA MIT 1 .....</b>		<b>86</b>
6.1.	Introdução.....	86
6.2.	1º Estudo de Viabilidade da MIT 1 .....	88
6.3.	2º Estudo de Viabilidade da MIT 1 .....	98
6.4.	Ameaças à Validade do 1º e 2º Estudo de Viabilidade da MIT 1.....	108
6.5.	Síntese do Capítulo.....	109
<b>CAPÍTULO 7 - ESTUDOS EXPERIMENTAIS DA MIT 2 .....</b>		<b>110</b>
7.1.	Introdução.....	110
7.2.	1º Estudo de Viabilidade da MIT 2 .....	111
7.3.	Estudo de Observação da MIT 2 .....	123
7.4.	Estudo de Caso na Indústria da MIT 2.....	134
7.5.	Síntese do Capítulo.....	137
<b>CAPÍTULO 8 - ESTUDOS EXPERIMENTAIS DA MIT 3 E UDRT-AD.....</b>		<b>138</b>
8.1.	Introdução.....	138
8.2.	1º Estudo de Viabilidade da MIT 3 .....	138
8.3.	2º Estudo de Viabilidade da MIT 3 e UDRT-AD .....	145
8.4.	Síntese do Capítulo.....	156
<b>CAPÍTULO 9 - ESTUDOS EXPERIMENTAIS DA UDG .....</b>		<b>157</b>
9.1.	Introdução.....	157
9.2.	Estudo piloto com a UDG .....	157
9.3.	1º Estudo de viabilidade da UDG .....	159
9.4.	2º Estudo de viabilidade da UDG .....	173
9.5.	Ameaças à Validade do 1º e 2º Estudo de Viabilidade da UDG .....	187
9.6.	Síntese do Capítulo.....	188
<b>CAPÍTULO 10 - CONCLUSÕES E PERSPECTIVAS FUTURAS .....</b>		<b>189</b>
10.1.	Conclusões.....	189
10.2.	Contribuições.....	191
10.3.	Perspectivas Futuras .....	192
<b>REFERÊNCIAS .....</b>		<b>195</b>
<b>APÊNDICE A – TÉCNICA MIT 1.....</b>		<b>205</b>
<b>APÊNDICE B – TÉCNICA MIT 2 .....</b>		<b>208</b>
<b>APÊNDICE C – TÉCNICA MIT 3 .....</b>		<b>210</b>
<b>APÊNDICE D – TÉCNICA UDRT-AD .....</b>		<b>212</b>
<b>APÊNDICE E – EXEMPLOS DE APLICAÇÃO DA TÉCNICA UDRT-AD .....</b>		<b>216</b>
<b>APÊNDICE F – TÉCNICA UDG .....</b>		<b>219</b>

## LISTA DE FIGURAS

<b>Figura 1. Visão geral da metodologia .....</b>	<b>8</b>
<b>Figura 2. Família de experimentos do conjunto de tecnologias proposto .....</b>	<b>9</b>
<b>Figura 3. Total de artigos no 1° filtro e 2° filtro no 1° MSL .....</b>	<b>32</b>
<b>Figura 4. Visão Temporal dos artigos selecionados no 1° MSL .....</b>	<b>34</b>
<b>Figura 5. Total de artigos no 1° filtro e 2° filtro no 2° MSL .....</b>	<b>50</b>
<b>Figura 6. Visão temporal das publicações identificadas no 2° MSL.....</b>	<b>52</b>
<b>Figura 7. Distribuição de artigos por periódico no 2° MSL.....</b>	<b>53</b>
<b>Figura 8. Distribuição de artigos por conferência no 2° MSL.....</b>	<b>53</b>
<b>Figura 9. Distribuição de artigos por workshop no 2° MSL.....</b>	<b>54</b>
<b>Figura 10. Modelos em que as tecnologias são aplicadas - 2° MSL .....</b>	<b>59</b>
<b>Figura 11. Tipo de contribuição dos artigos do 2° MSL .....</b>	<b>61</b>
<b>Figura 12. Combinação 1 de subquestões de pesquisa .....</b>	<b>65</b>
<b>Figura 13. Combinação 2 de subquestões de pesquisa .....</b>	<b>65</b>
<b>Figura 14. Combinação 3 de subquestões de pesquisa .....</b>	<b>66</b>
<b>Figura 15. Combinação 4 de subquestões de pesquisa .....</b>	<b>66</b>
<b>Figura 16. Componentes do Conjunto de Tecnologias para antecipação da Usabilidade.....</b>	<b>71</b>
<b>Figura 17. Visão geral do Conjunto de Técnicas MIT .....</b>	<b>73</b>
<b>Figura 18. Itens da heurística Visibilidade do status do Sistema (MIT 1) .....</b>	<b>74</b>
<b>Figura 19. Exemplo do processo de inspeção com a MIT 1.....</b>	<b>75</b>
<b>Figura 20. Itens da heurística Controle e liberdade ao usuário (MIT 2) .....</b>	<b>76</b>
<b>Figura 21 - Itens da heurística Concordância entre o sistema e o mundo real.....</b>	<b>76</b>
<b>Figura 22. Processo de aplicação das heurísticas da técnica UDRT-AD.....</b>	<b>78</b>
<b>Figura 23. Exemplo de um dos elementos da técnica UDRT-AD.....</b>	<b>79</b>

<b>Figura 24. Processo de aplicação da UDG.....</b>	<b>79</b>
<b>Figura 25. Elementos da técnica UDG.....</b>	<b>84</b>
<b>Figura 26. <i>Guidelines</i> da Heurística 2 da UDG.....</b>	<b>84</b>
<b>Figura 27. Dicas da Heurística 2 da UDG .....</b>	<b>85</b>
<b>Figura 28. Processo de Inspeção baseado em Sauer <i>et al.</i> (2000) .....</b>	<b>86</b>
<b>Figura 29. Estudos experimentais realizados com a MIT 1.....</b>	<b>87</b>
<b>Figura 30. Boxplots da eficiência do 1° estudo de viabilidade da MIT 1 .....</b>	<b>92</b>
<b>Figura 31. Boxplots da eficácia do 1° estudo de viabilidade da MIT 1.....</b>	<b>93</b>
<b>Figura 32. Percepção sobre Facilidade de Uso - 1° estudo de viabilidade MIT 1... 94</b>	
<b>Figura 33. Percepção sobre a Utilidade - 1° estudo de viabilidade da MIT 1.....</b>	<b>95</b>
<b>Figura 34. Item de Verificação MIT-1AD1 .....</b>	<b>97</b>
<b>Figura 35. Item de Verificação MIT-1AI1 .....</b>	<b>97</b>
<b>Figura 36. Boxplots da eficiência do 2° estudo de viabilidade da MIT 1 .....</b>	<b>103</b>
<b>Figura 37. Boxplots da eficácia do 2° estudo de viabilidade da MIT 1.....</b>	<b>104</b>
<b>Figura 38. Percepção sobre Facilidade de Uso - 2° estudo de viabilidade MIT 1. 105</b>	
<b>Figura 39. Percepção sobre Utilidade - 2° estudo de viabilidade da MIT 1.....</b>	<b>106</b>
<b>Figura 40. Estudos experimentais realizados com a MIT 2.....</b>	<b>110</b>
<b>Figura 41. Boxplots da eficiência e eficácia - 1° estudo de viabilidade da MIT 2. 115</b>	
<b>Figura 42. Percepção sobre Facilidade de Uso - 1° Estudo de Viabilidade MIT 2 116</b>	
<b>Figura 43. Percepção sobre a Utilidade da MIT 2 no 1° Estudo de Viabilidade.. 117</b>	
<b>Figura 44. Categoria Estrutura da MIT 2.....</b>	<b>119</b>
<b>Figura 45. Categorias Feedback do uso da MIT 2 e Sugestões de Melhoria .....</b>	<b>120</b>
<b>Figura 46. Item de verificação 2BD1 da MIT 2 .....</b>	<b>121</b>
<b>Figura 47. Item de verificação 2BA2 da MIT 2 – Baixo Detalhamento .....</b>	<b>121</b>
<b>Figura 48. Item de verificação 2BG6 da MIT 2 – Baixo e Alto Detalhamento .....</b>	<b>121</b>
<b>Figura 49. Item de verificação 2BF1 da MIT 2.....</b>	<b>121</b>

<b>Figura 50. Estudos experimentais realizados com a MIT 3 e UDRT-AD .....</b>	<b>138</b>
<b>Figura 51. Boxplots de eficiência do 1º Estudo de Viabilidade da MIT 3.....</b>	<b>141</b>
<b>Figura 52. Boxplots de eficácia do 1º Estudo de Viabilidade da MIT 3.....</b>	<b>142</b>
<b>Figura 53. Percepção sobre Facilidade de Uso - 1º Estudo de Viabilidade MIT 3</b>	<b>143</b>
<b>Figura 54. Percepção sobre a Utilidade - 1º Estudo de Viabilidade da MIT 3.....</b>	<b>144</b>
<b>Figura 55. Boxplots da eficácia - Estudo de Viabilidade da UDRT-AD e MIT 3.</b>	<b>150</b>
<b>Figura 56. Boxplots do tempo de modelagem do Estudo de Viabilidade da UDRT-AD e MIT 3.....</b>	<b>151</b>
<b>Figura 57. Boxplots com o comparativo da corretude .....</b>	<b>151</b>
<b>Figura 58. Estudos experimentais realizados com a UDG.....</b>	<b>157</b>
<b>Figura 59. Boxplots da corretude - 1º estudo de viabilidade UDG .....</b>	<b>167</b>
<b>Figura 60. Boxplots da completude - 1º estudo de viabilidade UDG .....</b>	<b>167</b>
<b>Figura 61. Boxplots da eficácia - 1º estudo de viabilidade UDG .....</b>	<b>168</b>
<b>Figura 62. Boxplots do tempo de construção - 1º estudo de viabilidade UDG.....</b>	<b>169</b>
<b>Figura 63. Percepção sobre Facilidade de Uso - 1º estudo de viabilidade UDG...</b>	<b>171</b>
<b>Figura 64. Percepção sobre Utilidade - 1º estudo de viabilidade UDG .....</b>	<b>171</b>
<b>Figura 65. Percepção sobre Intenção de Uso - 1º estudo de viabilidade UDG.....</b>	<b>172</b>
<b>Figura 66. Boxplots da corretude - 2º estudo de viabilidade UDG .....</b>	<b>178</b>
<b>Figura 67. Boxplots da completude - 2º estudo de viabilidade UDG .....</b>	<b>179</b>
<b>Figura 68. Boxplots da eficácia - 2º estudo de viabilidade UDG .....</b>	<b>180</b>
<b>Figura 69. Boxplots do tempo de construção - 2º estudo de viabilidade UDG.....</b>	<b>180</b>
<b>Figura 70. Percepção sobre Facilidade de Uso - 2º estudo de viabilidade UDG...</b>	<b>183</b>
<b>Figura 71. Percepção sobre Utilidade - 2º estudo de viabilidade UDG .....</b>	<b>184</b>
<b>Figura 72. Percepção sobre Intenção de Uso - 2º estudo de viabilidade UDG.....</b>	<b>184</b>

## LISTA DE TABELAS

Tabela 1. Requisitos para o Conjunto de Tecnologias .....	5
Tabela 2. Heurísticas de Nielsen (1994a) .....	19
Tabela 3. Heurísticas do Método UCE propostas por Hornbæk <i>et al.</i> (2007).....	21
Tabela 4. Questões do Método DU (DA SILVA e SILVEIRA, 2010) .....	22
Tabela 5. Objetivo do 1º MSL segundo Paradigma GQM .....	27
Tabela 6. Subquestões de Pesquisa do 1º MSL .....	28
Tabela 7. Termos e <i>String</i> de busca utilizados no 1º MSL.....	29
Tabela 8. Resultados do 1º MSL para cada uma das Subquestões .....	33
Tabela 9. Objetivo do 2º MSL segundo Paradigma GQM .....	42
Tabela 10. Subquestões de pesquisa do 2º MSL .....	42
Tabela 11. Fontes utilizadas no 2º MSL .....	43
Tabela 12. Termos e <i>String</i> de busca em inglês utilizados no 2º MSL.....	44
Tabela 13. Termos e <i>String</i> de busca em português utilizados no 2º MSL.....	44
Tabela 14. Resultados do 2º MSL para cada uma das Subquestões .....	51
Tabela 15. Heurísticas da técnica UDG e referências bases .....	82
Tabela 16. Resultado da inspeção - 1º Estudo de Viabilidade da MIT 1 .....	91
Tabela 17. Eficácia e Eficiência no 1º Estudo de Viabilidade da MIT 1 .....	92
Tabela 18. Resultado da inspeção - 2º Estudo de Viabilidade da MIT 1 .....	102
Tabela 19. Eficácia e Eficiência - 2º Estudo de Viabilidade da MIT 1.....	103
Tabela 20. Resultado da inspeção no 1º Estudo de Viabilidade da MIT 2.....	114
Tabela 21. Eficácia e Eficiência no 1º Estudo de Viabilidade da MIT 2 .....	114
Tabela 22. Pontos de investigação e suas definições .....	124
Tabela 23. Pontos de investigação e questões usadas na entrevista .....	125
Tabela 24. Resultados da inspeção - Estudo de Observação da MIT 2 .....	126

<b>Tabela 25. Resultado da inspeção no 1º estudo de viabilidade da MIT 3 .....</b>	<b>140</b>
<b>Tabela 26. Eficácia e Eficiência no 1º Estudo de Viabilidade da MIT 3.....</b>	<b>141</b>
<b>Tabela 27. Caracterização dos participantes do Estudo de Viabilidade da UDRT-AD e MIT 3.....</b>	<b>146</b>
<b>Tabela 28. Categorias de Defeitos (adaptada de Travassos <i>et al.</i> (1999)).....</b>	<b>149</b>
<b>Tabela 29. Dados quantitativos - Estudo de Viabilidade da UDRT-AD e MIT 3.</b>	<b>150</b>
<b>Tabela 30. Número de defeitos – Estudo de viabilidade da UDRT-AD e MIT 3..</b>	<b>152</b>
<b>Tabela 31. Quantidade de problemas de usabilidade do Grupo 1 .....</b>	<b>154</b>
<b>Tabela 32. Categorias de Defeitos em <i>mockups</i> (adaptadas de Travassos <i>et al.</i> (1999)) .....</b>	<b>160</b>
<b>Tabela 33. Critérios de avaliação da estética visual dos <i>mockups</i> (adaptados de Moshagen e Thielsch (2013)) .....</b>	<b>161</b>
<b>Tabela 34. Categorização dos participantes - 1º estudo de viabilidade da UDG..</b>	<b>163</b>
<b>Tabela 35. Dados quantitativos do 1º Estudo de Viabilidade da UDG.....</b>	<b>165</b>
<b>Tabela 36. Resultados dos indicadores do 1º estudo de viabilidade da UDG .....</b>	<b>166</b>
<b>Tabela 37. Resultados do indicador Estética - 1º estudo de viabilidade da UDG.</b>	<b>170</b>
<b>Tabela 38. Categorização dos participantes - 2º estudo de viabilidade da UDG..</b>	<b>174</b>
<b>Tabela 39. Dados quantitativos do 2º Estudo de Viabilidade da UDG.....</b>	<b>176</b>
<b>Tabela 40. Resultados dos indicadores do 2º estudo de viabilidade da UDG .....</b>	<b>177</b>
<b>Tabela 41. Resultados do indicador Estética do 2º estudo de viabilidade da UDG do Avaliador 1 .....</b>	<b>181</b>
<b>Tabela 42. Resultados do indicador Estética do 2º estudo de viabilidade da UDG do Avaliador 2.....</b>	<b>182</b>

## RESUMO

Pesquisas recentes têm objetivado melhorar o grau de usabilidade nas fases iniciais do processo de desenvolvimento de aplicações, a chamada "*Early Usability*". Parte das tecnologias propostas destina-se a avaliar a usabilidade através de modelos das fases iniciais do processo de desenvolvimento das aplicações. Este tipo de tecnologia visa apresentar bom nível de eficiência e eficácia, além de oferecer uma satisfação do usuário e redução do tempo de desenvolvimento. Outra parte destas tecnologias cria/projeta estes modelos já visando à usabilidade da aplicação final. No entanto, estas tecnologias têm sido desenvolvidas de forma independente e para modelos específicos, além de exigir um maior conhecimento em usabilidade por parte de quem as aplica. Deste modo, nesta pesquisa, propôs-se um conjunto de tecnologias (métodos, técnicas, ferramentas e processos, dentre outros) que auxilia na criação e avaliação de modelos visando à usabilidade da aplicação com ênfase nos estágios iniciais do processo de desenvolvimento. Este conjunto de tecnologias permite a identificação de um maior número de problemas de usabilidade já nas fases iniciais, se comparado às abordagens convencionais. Nas abordagens convencionais a identificação de problemas de usabilidade é realizada em modelos independentes, ou seja, ou é realizada em modelos específicos para um determinado contexto, como modelos *web*, ou é realizada quando o *software* está pronto. Portanto, o conjunto de tecnologias proposto abrange uma maior quantidade de modelos, não é específico para uma contexto e apoia profissionais de *software* com pouco conhecimento em usabilidade a construir/avaliar modelos das fases iniciais. Além disso, o conjunto de tecnologias orienta tanto engenheiros de *software* quanto engenheiros de usabilidade a evitar problemas de usabilidade de maneira mais proativa em seus projetos de desenvolvimento, pois eles poderão construir seus modelos já visando à usabilidade da aplicação. Com isso, espera-se contribuir para a melhoria da qualidade do *software* desenvolvido e diminuição do custo de correção de problemas de usabilidade antes que o código fonte seja escrito.

**Palavras-chave:** Usabilidade cedo, usabilidade nos estágios iniciais do processo de desenvolvimento, Modelo de Projeto, Modelo de Análise, Integração de IHC (Interação Humano-Computador) e ES (Engenharia de *Software*).

## ABSTRACT

Recent researches tried to improve the degree of usability in the early stages of the application's process development, the so called "Early Usability". Part of the proposed technologies is intended to evaluate usability through models elaborated in the initial phases of the application development process. This type of technology aims to present a good level of efficiency and effectiveness, besides offering an user satisfaction and reducing development time. Other technologies create/design these models already targeting the usability of the final application. However, these technologies have been developed independently and for specific models, besides requiring a greater knowledge in usability on the part of those who apply them. Thus, in this research, we proposed a set of technologies (methods, techniques, tools, processes, among others) which assists in creating and evaluating models addressing the usability of the application with emphasis in the early stages of development process. This set of technologies allows the identification of a higher number of usability problems already in the early stages, compared to conventional approaches. In conventional approaches, the identification of usability problems is performed in independent models, that is, or it is performed in models specific to a particular context, such as web models, or it is performed when the software is ready. Therefore, the proposed set of technologies encompasses a higher number of models, is not context-specific, and supports software professionals with little usability knowledge to build/evaluate early-stage models. In addition, the set of technologies guides both software engineers and usability engineers to avoid usability problems more proactively in its development projects, because this professionals can build their models already aiming at the usability of the application. With this, we hope to improve the software quality and to decrease the cost of correcting usability problems before the source code is written.

**Keywords:** Early usability, usability in early stage of development process, design model, analysis model, integration of HCI (Human-Computer Interaction) and SE (Software Engineering).

# CAPÍTULO 1 - INTRODUÇÃO

*Este capítulo apresenta a contextualização desta pesquisa de doutorado. Além disso, apresenta as motivações, a definição do problema, os objetivos de pesquisa e a metodologia seguida.*

## 1.1. Contexto

A usabilidade é conhecida como um aspecto significantemente da qualidade global de sistemas interativos (DE ANGELI *et al.*, 2003). A norma ISO 9240-11 (ISO/IEC, 1998), define usabilidade como a “medida que um produto pode ser usado por um usuário específico para alcançar objetivos específicos com eficácia, eficiência e satisfação em um contexto de uso específico”. Nielsen (1993) a define como “a capacidade de aprender e memorizar um sistema de *software*, a sua eficiência de uso, a sua capacidade de evitar e gerenciar erros de usuário e satisfação do usuário” e a norma ISO 25010 (ISO/IEC, 2011) como “a capacidade do produto de *software* de ser compreendido, aprendido, operado e atraente ao usuário, quando usado sob condições específicas”. Estas múltiplas definições implicam que a usabilidade tem sido interpretada de diferentes perspectivas e que diferentes *stakeholders* a observam através de diferentes caminhos (SEFFAH e METZKER, 2004).

A usabilidade possibilita benefícios como: a melhoria da produtividade dos usuários, a redução de treinamentos e custos com documentação (MOLINA e TOVAL, 2009). Diante disso, um grande número de pesquisadores tem investigado sobre a usabilidade no desenvolvimento de *software* (HORNBAEK *et al.*, 2007; JURISTO *et al.*, 2007). Segundo Donahue (2001), os investimentos em usabilidade têm possibilitado benefícios também para empresas de desenvolvimento de *software* como retorno econômico. Isto tem motivado mais organizações a considerar a usabilidade como um fator relevante em seus produtos de *software* (FERRÉ *et al.*, 2001).

No entanto, Seffah e Metzker (2004) destacam alguns obstáculos que as empresas de *software* enfrentam para não incluir a usabilidade durante o processo de desenvolvimento: (a) as atividades relacionadas a usabilidade são geralmente dissociadas do processo de desenvolvimento de *software* utilizado, e (b) os métodos de usabilidade ainda são difíceis de entender e pouco utilizados pelas equipes e organizações de desenvolvimento de *software*. Lauesen (1997), afirma que desenvolvedores: (a) possuem uma má compreensão das

técnicas de usabilidade e; (b) acreditam que sejam necessários laboratórios caros ao invés de uma abordagem de baixo custo que possa ser aprendida e aplicada por eles mesmo em seus projetos. Além disso, Ardito *et al.* (2011) realizaram um survey com organizações de desenvolvimento de *software* e verificaram que: (a) falta métodos de usabilidade disponíveis para as empresas de *software*; (b) os métodos de usabilidade disponíveis requerem muitos recursos em termos de custo, tempo e pessoas envolvidas; e (c) alguns desenvolvedores têm o principal interesse na funcionalidade e eficiência do código, o que gera pouco impacto para os usuários finais.

Modelos são produtos (artefatos) do processo de desenvolvimento de *software*. Em outras palavras, descrevem processos de negócio, o problema a ser resolvido e o sistema a ser desenvolvido. Alguns modelos, por utilizarem representações gráficas, são geralmente mais compreensíveis do que descrições detalhadas em linguagem natural (SOMMERVILLE, 2007). Nas fases iniciais do processo de desenvolvimento há modelos de análise e modelos de projeto.

Os modelos de análise são elaborados para se obter uma compreensão maior acerca do sistema a ser desenvolvido e para especificá-lo. Os modelos criados durante a atividade de análise cumprem os seguintes papéis (PRESSMAN, 2006):

- Ajudam o analista a entender a informação, a função e o comportamento do sistema, tornando a tarefa de análise mais fácil e sistemática;
- Tornam-se o ponto focal para a revisão e, portanto, a chave para a determinação da consistência da especificação.

Modelos de projeto também devem ser construídos com o objetivo de prover uma visão geral do sistema a ser construído, bem como uma variedade de visões mais específicas de seus elementos, de modo a guiar a implementação. Os modelos de projeto devem ser facilmente compreensíveis, tendo em vista que seu propósito é comunicar informações para profissionais responsáveis pela codificação, testes e manutenção (PRESSMAN, 2006). Um modelo de projeto deve:

- Prover uma visão da totalidade do que vai ser construído;
- Decompor o todo em partes e prover diferentes visões do sistema;
- Refinar e descrever com mais detalhes cada parte ou visão do sistema, de modo a prover orientação para a construção de cada detalhe.

## 1.2. Motivações

As comunidades de IHC (Interação Humano-Computador) e ES (Engenharia de Software) evoluíram separadamente e cada uma desenvolveu seus próprios métodos para atender às necessidades dos seus clientes e usuários de *software* (JURISTO *et al.*, 2007). Porém, nos últimos vinte anos, as fronteiras entre IHC e ES têm diminuído e as tentativas de minimizar a distância entre elas têm aumentado (JURISTO *et al.*, 2007). Seffah *et al.* (2001) propõem que os engenheiros de *software* e engenheiros de usabilidade aprendam uns com os outros para facilitar e incentivar a convergência das práticas em ambas as comunidades. Portanto, é de grande importância promover a mútua compreensão das atividades e responsabilidades das duas comunidades na construção de *software* com boa usabilidade, garantindo assim que os problemas de usabilidade serão adequadamente tratados no projeto de *software* durante o ciclo de desenvolvimento (JURISTO *et al.*, 2007).

Várias empresas de desenvolvimento de *software* estão começando a integrar práticas de IHC em seus processos de ES (FERRÉ *et al.*, 2004). Um indício da importância da integração de questões de usabilidade com o processo de desenvolvimento está no Processo de Usabilidade criado na primeira alteração à Norma ISO/IEC 12207 para Processos de Ciclo de Vida do *software* (ISO/IEC 2002). A primeira atividade deste processo específica como atividades centradas no usuário se encaixam em todo o processo do ciclo de vida do *software* e seleciona métodos de avaliação de usabilidade. Assim a integração de ambas as áreas não deve ser uma área de interesse somente de alguns especialistas (FERRÉ *et al.*, 2005), já que esta é essencial para desenvolvimento de *software*.

Tecnologias (métodos, técnicas, processos, abordagens, dentre outras) que avaliam a usabilidade têm sido utilizadas nas últimas fases do desenvolvimento de *software* (CONTE *et al.*, 2009a; MORAGA *et al.*, 2007; OLSINA e ROSSI 2002), com a avaliação da aplicação final ou mesmo de uma versão preliminar da interface final. No entanto, estudos mostram que se as avaliações de usabilidade forem realizadas durante as fases iniciais do desenvolvimento de uma aplicação, elas auxiliam a revelar problemas que podem ser corrigidos a um custo menor (TRAVASSOS *et al.*, 1999).

Por esta razão, torna-se importante considerar a usabilidade nos estágios iniciais do processo de desenvolvimento, a chamada “*Early Usability*” (JURISTO *et al.*, 2007). O objetivo da “*Early Usability*” é antecipar a usabilidade através de modelos de análise e projeto construídos durante o desenvolvimento de *software* que influenciarão na qualidade do *software* desenvolvido. A “*Early Usability*” pode auxiliar na redução do número de falhas detectadas em projetos de desenvolvimento de *software*. Ela também apresenta benefícios

como o aumento da qualidade do *software* desenvolvido e maior satisfação dos usuários (MOLINA e TOVAL, 2009). Segundo Fernandez *et al.* (2011a), se problemas de usabilidade são detectados e consertados cedo, a qualidade da aplicação final pode ser melhorada, poupando assim recursos na fase de desenvolvimento e contribuindo assim para a redução do custo do processo de desenvolvimento.

### 1.3. Definição do problema

Tecnologias têm sido propostas para garantir um bom nível de usabilidade das aplicações nas fases iniciais do processo de desenvolvimento (FERNANDEZ 2009; LUNA *et al.*, 2010; MOLINA e TOVAL 2009; PANACH *et al.*, 2007). Dentre estas tecnologias, uma parte delas objetiva a melhoria da usabilidade através da avaliação de modelos construídos durante as fases iniciais do processo de desenvolvimento das aplicações. Outra parte destas tecnologias cria estes modelos já visando à usabilidade. No entanto, estas tecnologias têm sido independentemente desenvolvidas para modelos específicos.

Com isso, o problema abordado neste trabalho está relacionado com a dificuldade de se obter modelos utilizados nas fases iniciais do processo de desenvolvimento que visem a usabilidade da aplicação. Para solucionar esse problema, criou-se um conjunto de tecnologias que auxilia na criação e avaliação desses modelos. Este conjunto de tecnologias permite a identificação de um maior número de defeitos de usabilidade já nas fases iniciais, se comparado às abordagens convencionais. Nas abordagens convencionais a identificação de problemas de usabilidade é realizada em modelos independentes, ou seja, em um único tipo de modelo, é realizada em modelos específicos para um determinado contexto, como modelos *web* ou é realizada quando o *software* já está pronto. Portanto, o conjunto de tecnologias proposto abrange uma maior quantidade de modelos, não é para um contexto específico e apoia profissionais de *software* com pouco conhecimento em usabilidade a construir/avaliar modelos das fases iniciais. Além disso, o conjunto de tecnologias guia tanto engenheiros de *software* quanto engenheiros de usabilidade a evitar problemas de usabilidade de maneira mais proativa em seus projetos de desenvolvimento, pois eles poderão já construir seus modelos visando à usabilidade da aplicação.

Portanto, a questão de pesquisa desse trabalho consiste em: “*Como antecipar a usabilidade das aplicações finais através dos modelos utilizados nas fases iniciais do processo de desenvolvimento de software a um baixo custo?*”. Para isso, será realizada a antecipação de usabilidade durante o projeto e avaliação desses modelos com um esforço reduzido, se comparado a uma abordagem convencional.

## 1.4. Objetivo de pesquisa

O objetivo principal deste trabalho consiste em propor um conjunto de tecnologias que apoie o projeto e avaliação de modelos utilizados nas fases iniciais do processo de desenvolvimento visando à usabilidade da aplicação final, de acordo com a questão de pesquisa apresentada na Subseção 1.3.

Os objetivos específicos são:

1. Obtenção de um corpo de conhecimento acerca das tecnologias existentes nas comunidades de ES e IHC que criam e avaliam modelos utilizados nas fases iniciais do processo de desenvolvimento, visando à usabilidade da aplicação final;
2. Definição de um conjunto de tecnologias que auxiliem no projeto e avaliação de modelos utilizados nas fases iniciais do processo de desenvolvimento, visando à usabilidade;
3. Avaliação experimental do conjunto de tecnologias proposto a fim de avaliar seu desempenho na prática através de experimentos controlados;
4. Definição de um conjunto de recursos que dê suporte ao uso das tecnologias por organizações de *software*, como: capacitação (treinamento) para uso das tecnologias e recursos de apoio à aplicação.

O propósito final é que o conjunto de tecnologias, apoiado pelo conjunto de recursos, possa ser empregado pelos próprios envolvidos nos projetos de desenvolvimento de *software*, como, por exemplo, engenheiros de *software* e profissionais de IHC. Para alcançar isto, o conjunto de tecnologias deve atender aos requisitos mostrados na Tabela 1.

**Tabela 1. Requisitos para o Conjunto de Tecnologias**

Requisitos	Justificativa
Ser fácil de aprender e de utilizar	As pessoas que realizarão a antecipação da usabilidade devem se tornar <b>aptas a aplicar as tecnologias</b> em um curto período de tempo, mesmo que essas pessoas não sejam especialistas em usabilidade. Idealmente, elas devem ser capazes de aplicar as tecnologias sem treinamento ou após um treinamento que dure poucas horas.
Apresentar bom nível de eficácia	As tecnologias devem <b>apoiar</b> as pessoas (utilizadores das tecnologias) a <b>identificar problemas de usabilidade nas fases iniciais e/ou projetar soluções de design com boa usabilidade</b> . Comparativamente a uma abordagem convencional, as tecnologias devem apresentar um nível de eficácia superior.
Apresentar bom nível de eficiência	As tecnologias também devem apoiar as pessoas (utilizadores das tecnologias) a antecipar a usabilidade com o <b>menor esforço</b> (homens-hora) possível. Comparativamente a uma abordagem convencional, as tecnologias devem apresentar um nível de eficiência equivalente ou superior.
Oferecer uma boa relação	O benefício resultante do uso das tecnologias deve superar seus custos. O maior componente do custo de uso de uma tecnologia é o <b>esforço</b> (homens-

Requisitos	Justificativa
custo-benefício na sua aplicação	hora). Também devem ser considerados os custos de <b>capacitação</b> das pessoas que usarão as tecnologias. O esforço (homens-hora) empregado no uso da tecnologia deve representar um percentual baixo do esforço total de desenvolvimento. Os custos de capacitação estão diretamente relacionados ao tempo de treinamento necessário para estar apto a aplicar as tecnologias.

Estes requisitos foram baseados em atributos que, segundo BOLCHINI e GARZOTTO (2007), podem contribuir para a aceitação e adoção das tecnologias pela indústria de *software*.

## 1.5. Metodologia baseada em evidências

Para atingir os objetivos relacionados a esta pesquisa será utilizada uma metodologia baseada em experimentação, que utiliza estudos experimentais para avaliar e evoluir as tecnologias de *software*. Para isso, será utilizada como base a metodologia proposta em Shull *et al.* (2001) que faz uso de Estudos Primários.

Mafra (2006) identificou a necessidade de executar Estudos Secundários antes dos Estudos Primários sugeridos em Shull *et al.* (2001), para que a definição da nova tecnologia pudesse estar baseada em evidências da literatura. Desta maneira, Mafra (2006) propõe a adoção de atividades iniciais para condução de Estudos Secundários, como Revisões Sistemáticas e Mapeamentos Sistemáticos (KITCHENHAM e CHARTERS, 2007), conforme será explicado no Capítulo 3 e 4 desta tese de doutorado.

Os **Estudos Secundários** visam identificar, avaliar, e interpretar todos os resultados relevantes a um determinado tópico de pesquisa, fenômeno de interesse ou questão de pesquisa. Um dos tipos de estudo secundário é o Mapeamento Sistemático da Literatura (MSL) (KITCHENHAM e CHARTERS, 2007). MSLs são realizados para fornecer uma visão geral de uma área de pesquisa e para estabelecer se existe evidência de pesquisa em um tópico (BUDGEN *et al.*, 2008).

Já os **Estudos Primários** são estudos que o pesquisador realiza para gerar novo conhecimento, isto é, levantar hipóteses e testá-las. A metodologia definida em Shull *et al.* (2001) é composta por quatro etapas, que contemplam a avaliação da tecnologia desde sua definição até sua transferência para a indústria, descritas a seguir:

- **Estudo de Viabilidade:** o objetivo principal deste tipo de estudo não é encontrar uma resposta definitiva, mas sim criar um corpo de conhecimento sobre a aplicação da tecnologia. Desta forma, é possível ao pesquisador avaliar se a aplicação da tecnologia é viável, ou seja, se atende de forma razoável aos objetivos inicialmente definidos.

- **Estudo de Observação:** geralmente é realizado em ambiente no qual a aplicação prática da tecnologia é observada por pesquisadores. Através dessa observação, é possível coletar dados sobre como a tecnologia é aplicada. Nesse sentido, os pesquisadores podem adquirir uma compreensão refinada sobre a tecnologia, ao presenciarem eventuais dificuldades que os participantes possam apresentar.
- **Estudo de Caso em Ciclo de Vida Real:** para que a tecnologia seja realmente útil, tem que ser capaz de ser usada em um ciclo de vida de desenvolvimento real. Este tipo de estudo deve ser conduzido com o propósito de caracterizar a aplicação da tecnologia no contexto de um ciclo de vida de desenvolvimento.
- **Estudo de Caso na Indústria:** tem como principal objetivo identificar se a aplicação da tecnologia possui alguma interação negativa com o ambiente industrial. Este tipo de estudo é realizado na última etapa da metodologia, pois é necessário ter certeza de que a nova tecnologia é boa antes de pedir para a indústria investir tempo e dinheiro na sua utilização.

A Figura 1 apresenta a metodologia adotada nesta pesquisa. A seguir são detalhadas as atividades realizadas em cada etapa da metodologia:

1. **Realização de Estudos Secundários:** esta etapa consiste na realização de dois estudos secundários (dois Mapeamentos Sistemáticos da Literatura). O objetivo dos mapeamentos sistemáticos é identificar tecnologias que auxiliem no projeto e avaliação de modelos, visando à usabilidade da aplicação final. Os resultados destes mapeamentos servirão para criar uma base de conhecimento sólida para esta pesquisa, ajudando a categorizar as tecnologias existentes.
2. **Criação do Conjunto de Tecnologias:** nesta etapa, com base no conhecimento adquirido e nas evidências coletadas através dos estudos secundários, propõe-se um conjunto de tecnologias que auxilie na avaliação e projeto de modelos utilizados nas fases iniciais do processo de desenvolvimento, visando à usabilidade
3. **Realização de Estudos Primários:** para avaliar o conjunto de tecnologias proposto diferentes tipos de estudos experimentais serão realizados:
  - **Estudos de viabilidade:** para determinar a viabilidade prática de aplicação das tecnologias propostas;

- **Estudos de observação:** para aprimorar o entendimento dos pesquisadores em relação à aplicação das tecnologias propostas e possibilitar seu refinamento;
- **Estudo de caso em um ciclo de vida:** para caracterizar a aplicação das tecnologias propostas em um ciclo de vida de desenvolvimento;
- **Estudo de caso na indústria:** para identificar se existem problemas de integração na aplicação das tecnologias propostas na indústria.

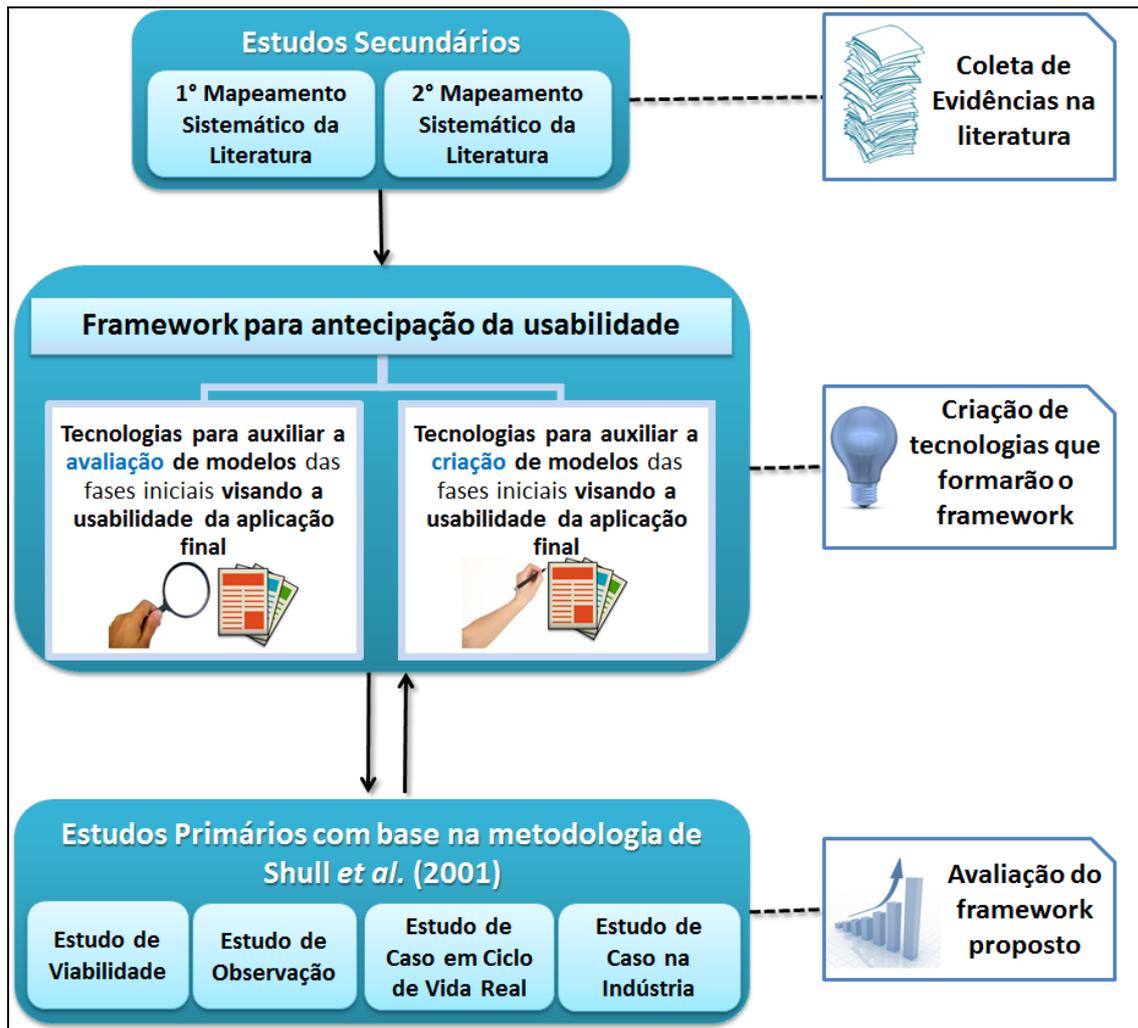


Figura 1. Visão geral da metodologia

## 1.6. Principais contribuições

As duas principais contribuições desta tese de doutorado são descritas abaixo:

- Identificação, categorização e análise das tecnologias identificadas na literatura científica que apoiam a integração entre as áreas de Engenharia de Software e Interação Humano-Computador e que auxiliam no projeto ou avaliação de modelos

das fases iniciais do processo de desenvolvimento de software visando à usabilidade da aplicação final. Foram conduzidos estudos secundários para investigar as lacunas de pesquisa existentes neste tópico de pesquisa.

- Conjunto de tecnologias para apoiar o projeto e avaliação de modelos utilizados nas fases iniciais do processo de desenvolvimento, que podem ser empregadas pelos próprios envolvidos nos projetos de desenvolvimento de *software*. As tecnologias propostas são: *Model Inspection Technique for Usability Evaluation* (MIT), sendo que a MIT 1 apoia a avaliação de usabilidade em especificações de casos de uso, a MIT 2 apoia a avaliação de usabilidade em mockups e a MIT 3 apoia a avaliação de usabilidade em diagramas de atividades; *Usability Design Reading Technique for Activity Diagrams* (UDRT-AD); e *Usability Design Guidelines* (UDG). Estas tecnologias foram avaliadas através de uma família de experimentos representada na Figura 2. Pode-se notar que alguns estudos iniciais foram realizados com a versão 1 das técnicas MIT 1, MIT 2 e MIT 3. Estes estudos iniciais foram realizados no âmbito da minha pesquisa de mestrado e; portanto, não serão descritos nesta proposta de doutorado. Como o objetivo da família de experimentos de uma tecnologia é extrair conclusões significativas e também evoluir estas tecnologias, novos estudos experimentais foram realizados, já no âmbito desta pesquisa de doutorado.

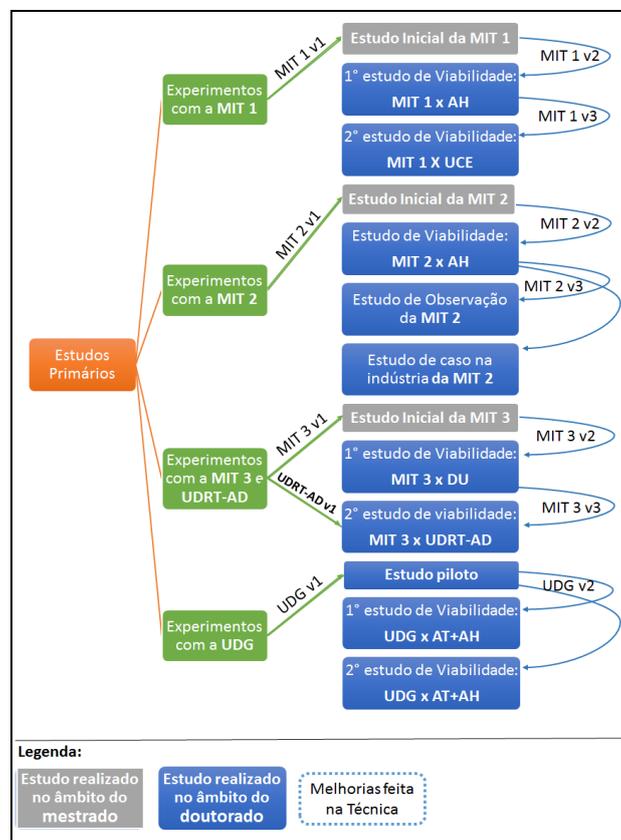


Figura 2. Família de experimentos do conjunto de tecnologias proposto

## 1.7. Outros resultados

Esta pesquisa de doutorado resultou em publicações científicas, prêmios e um projeto financiado por uma agência regional. Estes resultados são apresentados nas Subsubseções a seguir.

### 1.7.1 Artigos publicados

Durante esta pesquisa de doutorado, alguns artigos foram publicados relacionados ao tópico de pesquisa desta tese. Os principais resultados foram publicados no International Conference on Model Driven Engineering Languages and Systems (MODELS), International Conference on Software Engineering and Knowledge Engineering (SEKE), International Conference on Enterprise Information Systems (ICEIS), Internacional Conference on Human-Computer Interaction (HCI), Simpósio Brasileiro de Engenharia de Software (SBES) e Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC). As referências resumidas dos artigos originados desta pesquisa são apresentados a seguir em ordem cronológica de publicação:

- **Artigo 1:** VALENTIM, N. M. C.; DA SILVA, T. S.; SILVEIRA, M. S.; CONTE, T. Estudo Comparativo entre Técnicas de Inspeção de Usabilidade sobre Diagramas de Atividades. In: XII Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais (IHC 2013), 2013.
- **Artigo 2:** VALENTIM, N. M. C.; CONTE, T. Improving a Usability Inspection Technique Based on Quantitative and Qualitative Analysis. In: Simpósio Brasileiro de Engenharia de Software (SBES 2014), 2014.
- **Artigo 3:** SILVA, W.; VALENTIM, N. M. C.; CONTE, T. U. Projetando Diagramas de Atividade visando a Usabilidade de Aplicações Interativas. In: XIII Simpósio Brasileiro Sobre Fatores Humanos em Sistemas Computacionais (IHC 2014), 2014.
- **Artigo 4:** VALENTIM, N. M. C.; MALDONADO, J. C.; CONTE, T. Evaluating an Inspection Technique for Use Case Specifications: Quantitative and Qualitative Analysis. In: 17th International Conference on Enterprise Information Systems (ICEIS 2015), 2015a.
- **Artigo 5:** VALENTIM, N. M. C.; RABELO, J.; ORAN, A.; CONTE, T. A Controlled Experiment with Usability Inspection Techniques Applied to Use Case Specifications: Comparing the MIT 1 and the UCE Techniques. In: ACM/IEEE

18th International Conference on Model Driven Engineering Languages and Systems (MODELS 2015), 2015b.

- **Artigo 6:** VALENTIM, N. M. C.; CONTE, T.; ESTACIO, B. J. S.; PRIKLADNICKI, R. How do software engineers apply an early usability inspection technique? A qualitative study. In: 27th International Conference on Software Engineering and Knowledge Engineering (SEKE 2015), USA, 2015c.
- **Artigo 7:** VALENTIM, N. M. C.; CONTE, T. U. Framework que apoia a Antecipação da Usabilidade nas Fases Iniciais do Processo de Desenvolvimento de Software. In: Workshop de Teses e Dissertações em IHC (WTDIHC 2015), 2015.
- **Artigo 8:** SILVA, W.; VALENTIM, N. M. C.; CONTE, T. Integrating the Usability into the Software Development Process: A Systematic Mapping Study. In: 17th International Conference on Enterprise Information Systems (ICEIS 2015), Barcelona, 2015. p. 105-113.
- **Artigo 9:** SILVA, W.; VALENTIM, N. M. C.; CONTE, T. Designing Activity Diagrams Aiming at Achieving Usability in Interactive Applications: An Empirical Study.. In: Conference on Human-Computer Interaction (HCII 2016), Toronto, 2016, v. 1. p. 208-219.

O artigo 8 apresenta os resultados do 1º Mapeamento Sistemático da Literatura realizado no âmbito desta pesquisa de doutorado. O artigo 7 descreve a ideia central desta pesquisa. O artigo 3 apresenta uma das tecnologias proposta nesta pesquisa, a UDRT-AD. E por fim, os artigos 1, 2, 4, 5, 6 e 9 descrevem estudos realizados com as tecnologias propostas nesta pesquisa de doutorado.

Além disso, outros artigos foram publicados e são listados a seguir:

- **Artigo 10:** VALENTIM, N. M. C.; Rabelo, J.; Silva, W.; Coutinho, W.; CONTE, T. Avaliando a qualidade de um aplicativo web móvel através de um teste de usabilidade: um relato de experiência. In: Simpósio Brasileiro de Qualidade de Software (SBQS 2014), 2014.
- **Artigo 11:** SILVA, W.; VALENTIM, N. M. C.; CONTE, T. Estudo Comparativo entre Diagrama de Atividade e Análise Hierárquica de Tarefas: uma Análise da Facilidade de Modelagem e Compreensão. In: Simpósio Brasileiro Sobre Fatores Humanos em Sistemas Computacionais (IHC 2014), 2014.
- **Artigo 12:** VALENTIM, N. M. C.; SILVA, W.; CONTE, T. Avaliando a Experiência do Usuário e a Usabilidade de um Aplicativo Web Móvel: um relato de

- experiência. In: XVIII Congresso Ibero-Americano em Engenharia de Software (CIbSE 2015), Lima, 2015.
- **Artigo 13:** PINHEIRO, S. F.; VALENTIM, N. M. C.; VINCENZI, A. M. R. Um Comparativo na Execução de Testes Manuais e Testes de Aceitação Automatizados em uma Aplicação Web. In: XIV Simpósio Brasileiro de Qualidade de Software (SBQS 2015), 2015.
  - **Artigo 14:** ESTACIO, B. J. S.; VALENTIM, N. M. C.; RIVERO, L.; CONTE, T.; PRIKLADNICKI, R. Evaluating the Use of Pair Programming and Coding Dojo in Teaching Mockups Development: An Empirical Study. In: 48th Annual Hawaii International Conference on System Sciences (HICSS 2015), 2015.
  - **Artigo 15:** LOPES, A.; VALENTIM, N. M. C.; FERREIRA, B. M.; ZILSE, R.; CONTE, T. Utilizando Modelagem de Interação para Projetar uma Aplicação Móvel Multiusuário: Um Relato de Experiência. In: Simpósio Brasileiro de Qualidade de Software (SBQS 2016), 2016.
  - **Artigo 16:** FERREIRA, B. M.; RIVERO, L.; VALENTIM, N. M. C.; ZILSE, R.; KOSTER, A.; CONTE, T. Evaluation Of UX Methods: Lessons Learned When Evaluating a Multi-user Mobile Application. In: 18th International Conference on Human-Computer Interaction (HCI 2016), 2016.
  - **Artigo 17:** VALENTIM, N. M. C.; SILVA, W.; CONTE, T. The Students' Perspectives on Applying Design Thinking for the Design of Mobile Applications. In: 39th International Conference on Software Engineering (ICSE 2017) - Software Engineering Education and Training (SEET), 2017.
  - **Artigo 18:** VALENTIM, N. M. C.; LOPES, A.; OLIVEIRA, E.; CONTE, T.; VINCENZI, A. M. R.; MALDONADO, J. C. An Acceptance Empirical Assessment of Open Source Test Tools. In: 19th International Conference on Enterprise Information Systems (ICEIS 2017), 2017.
  - **Artigo 19:** OLIVEIRA, R.; SOUSA, L.; DE MELLO, R.; VALENTIM, N. M. C.; LOPES, A.; CONTE, T.; GARCIA, A.; OLIVEIRA, E.; LUCENA, C. Collaborative Identification of Code Smells: A Multi-case Study. In: 39th International Conference on Software Engineering (ICSE 2017) - Software Engineering in Practice (SEIP), 2017.

### 1.7.2 Prêmios

O artigo 4 foi premiado como melhor artigo (1º lugar) na área de Interação Humano-Computador no 17th International Conference on Enterprise Information Systems (ICEIS 2015).

Os artigos 10, 13 e 15 foram premiados como melhores artigos (1º lugar) na categoria relato de experiência no Simpósio Brasileiro de Qualidade de Software (SBQS) nos anos de 2014, 2015 e 2016, respectivamente.

### 1.7.3 Financiamento

Edital N° 016/2013 PROTI – PESQUISA, processo N° 062.00578/2014. "Melhorando a qualidade no desenvolvimento de software no estado do Amazonas por meio da integração de usabilidade no processo de desenvolvimento". Financiadora: Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM). Coordenadora: Tayana Uchôa Conte.

## 1.8. Organização

Esta tese de doutorado está organizada em mais nove capítulos, além deste primeiro capítulo de introdução, que apresentou o contexto, motivação, o problema, os objetivos e a metodologia. A organização do texto deste trabalho segue a estrutura abaixo:

- **Capítulo 2 – Avaliação de Usabilidade no Processo de Desenvolvimento de Aplicações:** descreve o referencial teórico desta pesquisa, mostrando os tipos existentes de avaliações de usabilidade. Este capítulo apresenta também avaliações de usabilidade utilizadas nas etapas iniciais do processo de desenvolvimento.
- **Capítulo 3 – Mapeamento Sistemático sobre Tecnologias que Apoiam a Integração entre as Áreas de Engenharia de *Software* e Interação Humano-Computador:** apresenta o mapeamento sistemático realizado com o objetivo de identificar as tecnologias que melhoram a usabilidade no processo de desenvolvimento de *software*.
- **Capítulo 4 – Mapeamento Sistemático sobre Tecnologias que Auxiliam no Projeto ou Avaliação de Modelos visando à Usabilidade da Aplicação:** apresenta o mapeamento sistemático realizado com o objetivo de encontrar as tecnologias que auxiliam no projeto ou avaliação de modelos visando à usabilidade.
- **Capítulo 5 – Conjunto de Tecnologias que apoia a Antecipação da Usabilidade nas Fases Iniciais do Processo de Desenvolvimento de**

**Software:** apresenta a proposta de tecnologias para auxiliar no projeto e avaliação de modelos utilizados nas fases iniciais do processo de desenvolvimento visando à usabilidade da aplicação final.

- **Capítulo 6 – Estudos Experimentais da MIT 1:** descreve os experimentos realizados com o propósito de avaliar e aperfeiçoar a técnica MIT 1 que compõe o conjunto de tecnologias proposto.
- **Capítulo 7 – Estudos Experimentais da MIT 2:** descreve os experimentos realizados com o propósito de avaliar e aperfeiçoar a técnica MIT 2 que compõe o conjunto de tecnologias proposto.
- **Capítulo 8 – Estudos Experimentais da MIT 3 e UDRT-AD:** descreve os experimentos realizados com o propósito de avaliar e aperfeiçoar as técnicas MIT 3 e UDRT-AD que compõem o conjunto de tecnologias proposto.
- **Capítulo 9 – Estudos Experimentais da UDG:** descreve os experimentos realizados com o propósito de avaliar e aperfeiçoar a técnica UDG que compõem o conjunto de tecnologias proposto.
- **Capítulo 10 – Conclusão e Perspectivas Futuras:** contém as conclusões e contribuições do trabalho, além de indicar possibilidades de trabalhos futuros.

## CAPÍTULO 2 - AVALIAÇÃO DE USABILIDADE NO PROCESSO DE DESENVOLVIMENTO DE APLICAÇÕES

*Este capítulo apresenta os tipos existentes de avaliações de usabilidade utilizadas no processo de desenvolvimento de aplicações. Além disso, são apresentadas avaliações de usabilidade utilizadas nas fases iniciais do processo de desenvolvimento.*

### 2.1. Introdução

Segundo Rocha e Baranauskas (2003), uma avaliação de usabilidade tem três grandes objetivos: avaliar a usabilidade das funções do sistema (se está adequada aos requisitos da tarefa do usuário), avaliar o efeito da interface junto ao usuário (avaliar sua usabilidade) e identificar problemas específicos do sistema (aspectos do design os quais quando usados no contexto alvo, causam resultados inesperados ou confusão entre os usuários). As tecnologias comumente adotadas para avaliação de usabilidade podem ser divididas em duas grandes categorias: (1) Inspeções de Usabilidade, nas quais inspetores examinam aspectos da aplicação para detectar violações de princípios de usabilidade estabelecidos; e (2) Testes de Usabilidade, que são tecnologias de avaliação baseados na participação direta de usuários (PRATES e BARBOSA, 2003). Além disso, Kieras (2009) aborda sobre uma terceira categoria chamada Avaliação baseada em Modelos, também conhecida como Modelagem Analítica, no qual avaliadores predizem a usabilidade por empregar diferentes tipos de modelos.

A seguir são apresentadas as principais tecnologias para Avaliação e Projeto de Usabilidade, divididos nas categorias citadas acima: Inspeção de Usabilidade, Teste de usabilidade e Avaliação baseada em Modelos.

### 2.2. Tipos de Avaliação de Usabilidade

A **Inspeção de Usabilidade** é uma avaliação onde não há a participação do usuário final e pode ser usada em qualquer fase do desenvolvimento do *software*. Segundo Rocha e Baranaukas (2003), esta objetiva encontrar problemas de usabilidade em um design

de uma interface de usuário e com base nesses problemas fazer recomendações no sentido de eliminar os problemas e melhorar a usabilidade do design. Isso significa que inspeções de usabilidade são feitas em um estágio onde a interface está sendo gerada e necessita ser avaliada. Além disto, Prates e Barbosa (2003) ressaltam que este tipo de avaliação possibilita a formação e capacitação da equipe com relação a projetos de interface centrados no usuário. Entre os principais métodos de inspeção existentes estão:

- **Avaliação Heurística:** proposta por Nielsen (1994a), envolve um pequeno conjunto de avaliadores examinando a interface e julgando suas características tomando como base princípios de usabilidade, denominados heurísticas (PRATES e BARBOSA, 2003). Este método será apresentado na Subseção 2.2.1.
- **Percorso Cognitivo (*Walkthrough Cognitivo*):** é um método que tem como foco principal avaliar o design quanto à sua facilidade de aprendizagem, onde o autor de um aspecto do design apresenta uma proposta para um grupo de pares e estes avaliam a solução usando critérios apropriados ao design específico (POLSON *et al.*, 1992).
- **Guidelines e Checklists:** segundo Winckler e Pimenta (2002) e Cybis *et al.* (2007), são recomendações que podem ser usadas com o duplo propósito de auxiliar o processo de concepção ou guiar a avaliação. Durante o processo de concepção, designers devem consultar tais recomendações que auxiliam a evitar problemas de usabilidade. E pode-se utilizar o conjunto de recomendações ergonômicas como apoio a inspeção da interface; neste caso, a interface é inspecionada minuciosamente por um avaliador que verifica se todas as recomendações ergonômicas são respeitadas.

Os **Testes de Usabilidade** são métodos de avaliação centrados no usuário que incluem métodos experimentais, métodos observacionais e técnicas de questionamento. Para usar esses métodos é necessária uma implementação real do sistema em algum formato, que pode variar desde uma simulação da capacidade interativa do sistema, um protótipo básico implementado, um cenário ou a implementação completa do sistema (ROCHA e BARANAUSKAS, 2003). Entre as técnicas utilizadas nos Testes de Usabilidade, pode-se destacar:

- **Técnicas baseadas em Observação:** são técnicas que têm por fundamento a observação do usuário interagindo com o sistema. Uma delas é *Thinking Aloud*, na qual se solicita ao usuário que verbalize tudo que pensa enquanto usa um sistema e

a expectativa é que seus pensamentos mostrem como o usuário interpreta cada item da interface (LEWIS, 1982).

- **Técnicas baseadas em Perguntas:** são técnicas que coletam a opinião dos usuários sobre a interface. Entrevista é uma destas técnicas onde um entrevistador faz perguntas ao usuário, podendo gravar ou registrar a entrevista.
- **Avaliação através da Monitoração de Respostas Fisiológicas:** um exemplo deste tipo de avaliação é o uso de aparelhos de *Eye Tracking*, que determinam o movimento e a fixação dos olhos de uma pessoa (NAMAHN, 2001) e isto reflete na quantidade de processamento cognitivo que um display da interface requer.
- **Avaliação Experimental:** avaliação de aspectos específicos através de Estudos Controlados e Estudos de Observações. Um Estudo Controlado é um estudo direcionado por hipóteses e permite a análise quantitativa. Já um Estudo de Observação é um estudo onde é possível coletar dados sobre como a tecnologia é aplicada (MAFRA, 2006) e permite uma análise qualitativa.

Embora o teste de usabilidade seja considerado o método mais eficaz para avaliar sistemas e protótipos do ponto de vista do usuário das aplicações, seu custo é alto, pois envolve o tempo dos usuários e muitas vezes o uso de laboratórios específicos de usabilidade (MATERA *et al.*, 2002). Os métodos de inspeção foram propostos como uma alternativa com bom custo-benefício em comparação com os testes de usabilidade.

A **Avaliação baseada em Modelos (Modelagem Analítica)** tem o objetivo de obter alguns resultados de usabilidade antes de ser realizado um protótipo ou testes com usuários (KIERAS, 2009). A abordagem utiliza um modelo para representar o design da interface e produzir medições previstas da usabilidade da interface. O modelo é baseado na descrição detalhada do design proposto e uma análise da tarefa detalhada. O modelo explica como os usuários realizarão as tarefas através da interação com a interface proposta, e usa a teoria psicológica e dados paramétricos para gerar as métricas previstas de usabilidade. Uma vez que o modelo é construído, as previsões de usabilidade podem ser rapidamente e facilmente obtidas através de cálculos ou executando uma simulação. Além disso, as implicações de variações no design podem ser rapidamente exploradas através de alterações no modelo (KIERAS, 2009). Entre as mais relevantes abordagens estão:

- **Modelos de Redes de Tarefas (*Task network models*):** o desempenho da tarefa é modelado em termos de uma rede PERT (*Program Evaluation Review Technique*) como gráfico de processos. Cada processo começa quando seus processos pré-requisitos foram concluídos, e tem uma distribuição de tempo de conclusão. Os

processos são geralmente chamados de "tarefas", mas eles não precisam ser desempenhados somente por pessoas, pois também há processos de máquina. Além disso, outras informações, tal como os parâmetros de carga de trabalho ou de recursos, podem ser anexadas a cada processo (LAUGHERY, 1989);

- **Modelos de Arquitetura Cognitiva (*Cognitive architecture models*):** sistemas de arquitetura cognitiva consistem em um conjunto de interações hipotéticas perceptivas, cognitivas e motoras assumidas por um ser humano. O funcionamento dos componentes e as suas interações são simuladas com um programa de computador. Este programa fornece entradas (estímulos) para o humano simulado, e fornece saídas (respostas) produzidas pelo humano simulado. Tarefas são modeladas principalmente programando o componente cognitivo de acordo com uma análise de tarefas, e, em seguida, as previsões de desempenho são obtidas por executar a simulação usando cenários selecionados para gerar os eventos de entrada na tarefa (ANDERSON, 1983) e;
- **Modelos GOMS:** descrevem o conhecimento dos procedimentos que um usuário deve ter para operar um sistema. O acrônimo e a abordagem podem ser resumidos da seguinte forma: o usuário pode realizar certos Objetivos (G) com o sistema; Operadores (O) são as ações básicas que podem ser realizadas no sistema, por exemplo, encontrar um ícone na tela; Métodos (M) são sequências de operadores que, quando executados, realizam um objetivo; Regras de Seleção (S) descrevem que métodos devem ser utilizados em que situação para realizar um objetivo, se houver mais do que um disponível. Construir um modelo GOMS envolve escrever os métodos para realizar os objetivos da tarefa de interesse, e depois calcular métricas de usabilidade previstas da representação do método (CARD *et al.*, 1983).

### 2.2.1. Avaliação Heurística

Nielsen (1994b) propôs um método de inspeção de *software*, chamado Avaliação Heurística (AH) e este tem por objetivo encontrar problemas de usabilidade através de uma análise de conformidade do sistema frente a padrões de qualidade ou heurísticas definidas por especialistas. Heurísticas são princípios que podem guiar uma decisão de projeto ou ser utilizados para uma análise crítica de uma decisão tomada (NIELSEN, 1994b).

Matera *et al.* (2006) cita a AH como um dos principais métodos de inspeção de aplicações interativas. Além disso, estudos mostram que este é um método muito eficiente, com uma boa relação custo-benefício (NIELSEN, 1994a).

A Tabela 2 apresenta as 10 Heurísticas de Nielsen desenvolvidas a partir do estudo de 249 problemas de usabilidade (NIELSEN, 1994a):

**Tabela 2. Heurísticas de Nielsen (1994a)**

Nº	Heurística	Explicação
1	Visibilidade do estado do sistema	O sistema deve sempre manter os usuários informados sobre o que está acontecendo, através de um <i>feedback</i> apropriado dentro de um tempo razoável.
2	Concordância entre o sistema e o mundo real	O sistema deve utilizar a linguagem do usuário, com palavras, frases e conceitos familiares ao usuário. Seguir as convenções do mundo real e fazer a informação aparecer na ordem natural e lógica.
3	Controle e liberdade ao usuário	O sistema deve dar apoio a funções como <i>Undo</i> e <i>Redo</i> ou funções que permitam ao usuário utilizar “saídas de emergência” em caso de escolhas de funções erradas ou para sair de um estado não esperado.
4	Consistência e padrões	Devem ser seguidas convenções da plataforma de desenvolvimento e padrões de interface normalmente aceitos. Usuários não devem ter que adivinhar se palavras, situações ou ações diferentes significam a mesma coisa.
5	Prevenção de Erros	O sistema deve evitar a ocorrência de erros na sua utilização. Melhor do que apresentar boas mensagens de erro, é ter um projeto cuidadoso que evite a ocorrência de um problema.
6	Reconhecer ao invés de lembrar	Tornar objetos, ações e opções visíveis, para que o usuário não tenha que se lembrar de informações de uma parte do diálogo para outra. Instruções para uso do sistema devem estar visíveis, ou facilmente recuperáveis, quando necessário.
7	Flexibilidade e eficiência de uso	Aceleradores (abreviações, teclas de função...) podem tornar mais rápida a interação com o usuário. Permitir aos usuários customizar ações frequentes.
8	Projeto minimalista e estético	Diálogos não devem conter informação irrelevante ou raramente necessária. Todas as unidades extras de informações em um diálogo competem com aquelas que são realmente relevantes, e diminuem sua visibilidade relativa
9	Reconhecimento, diagnóstico e recuperação de erros	Mensagens de erro devem ser expressas em linguagem simples (sem códigos), indicando precisamente o problema, e sugerindo construtivamente uma solução.
10	Ajuda e Documentação	As informações de ajuda e documentação devem ser fáceis de procurar, com foco na tarefa do usuário, listando passos concretos que devem ser seguidos e não serem grandes demais.

Segundo Rocha e Baranauskas (2003), a Avaliação Heurística deve ser vista como parte do processo de design iterativo de uma interface. Esta envolve um pequeno conjunto de avaliadores examinando a interface e julgando suas características usando os princípios de usabilidade, descritos acima.

Na primeira fase, cada avaliador percorre a interface diversas vezes inspecionando os diferentes componentes do diálogo e ao encontrar problemas os descreve associando-os às heurísticas de usabilidade que foram violadas, julgando-os também quanto à gravidade de acordo com a escala de severidade abaixo:

- 0 – Não representa um problema de usabilidade;
- 1 – Somente um problema cosmético. Não precisa ser consertado a não ser que haja tempo disponível;
- 2 – Problema leve. Baixa prioridade para consertá-lo;
- 3 – Problema grave. Alta prioridade para consertá-lo;
- 4 – Problema catastrófico. É obrigatório consertá-lo.

A segunda fase consiste de uma reunião onde é feita a consolidação das listas individuais de problemas dos avaliadores. Cada avaliador tem acesso aos relatórios individuais de todos os avaliadores, e podem julgar sobre os problemas encontrados pelos outros avaliadores. Ao final dessa fase um relatório unificado é obtido com todos os problemas de usabilidade encontrados.

E na última fase, selecionam-se os problemas que devem ser corrigidos de acordo com a severidade, prazos e orçamentos. Esta etapa deve ser realizada junto ao cliente ou ao gerente de projeto.

### **2.3. Avaliação de Usabilidade nas Fases Iniciais do Processo de Desenvolvimento**

Diversas pesquisas têm buscado propor tecnologias que auxiliem na antecipação da usabilidade nas fases iniciais do processo de desenvolvimento, a chamada “*Early Usability*” (ABRAHÃO e INSFRAN, 2006; HONRBÆK *et al.*, 2007; MOLINA e TOVAL, 2009; PANACH *et al.*, 2007). Como citado anteriormente, o objetivo da “*Early Usability*” é antecipar a usabilidade através de modelos de análise e projeto construídos durante o desenvolvimento de *software* que influenciarão na qualidade do *software* desenvolvido. Suas vantagens são: redução do número de falhas detectadas em projetos de desenvolvimento de *software*, aumento da qualidade do *software* desenvolvido, maior satisfação dos usuários e redução do custo do processo de desenvolvimento.

Entre as tecnologias que avaliam a usabilidade nas fases iniciais e que utilizam como objeto de avaliação modelos que são comumente utilizados em um processo de desenvolvimento pode-se citar os seguintes trabalhos: Hornbæk *et al.* (2007), Da Silva e Silveira (2010) e Luna *et al.* (2010).

#### **2.3.1. Método *Use Case Evaluation* (UCE)**

Hornbæk *et al.* (2007) propuseram um método de avaliação chamado *Use Case Evaluation* (Avaliação de Casos de Uso) adaptado para avaliação de usabilidade baseada em

casos de uso. Este método tem o objetivo de facilitar a identificação dos problemas de usabilidade no início do processo de desenvolvimento, onde os primeiros casos de uso essenciais são descritos.

O método de Avaliação de Casos de Uso (UCE) é composto por três atividades: (1) Inspeção de casos de uso, (2) Avaliação de casos de uso, e (3) Documentação de avaliação.

Na primeira atividade o objetivo principal é identificar problemas de usabilidade que o avaliador está convencido que um usuário pode vir a experimentar. Esta atividade possui duas etapas:

- *Brainstorm*: o avaliador percorre os casos de uso, procurando por problemas de usabilidade, sem qualquer procedimento sistemático.
- *Inspeção Sistemática*: o avaliador tenta identificar os problemas de usabilidade empregando as Heurísticas de Nielsen adaptadas (Tabela 3) buscando encontrar semelhanças em que um caso de uso ultrapasse uma heurística.

**Tabela 3. Heurísticas do Método UCE propostas por Hornbæk *et al.* (2007)**

Nº	Heurística	Explicação	Baseado em
1	Visibilidade do estado do sistema	Mesma descrição da Avaliação Heurística	Heurísticas de Nielsen
2	Concordância entre o sistema e o mundo real	Mesma descrição da Avaliação Heurística	Heurísticas de Nielsen
3	Controle e liberdade ao usuário	Mesma descrição da Avaliação Heurística	Heurísticas de Nielsen
4	Consistência e padrões	Mesma descrição da Avaliação Heurística	Heurísticas de Nielsen
5	Prevenção de Erros	Mesma descrição da Avaliação Heurística	Heurísticas de Nielsen
6	Reconhecer ao invés de lembrar	Mesma descrição da Avaliação Heurística	Heurísticas de Nielsen
7	Flexibilidade e eficiência de uso	Mesma descrição da Avaliação Heurística	Heurísticas de Nielsen
8	Reconhecimento, diagnóstico e recuperação de erros	Mesma descrição da Avaliação Heurística	Heurísticas de Nielsen
9	Evitar operações mentais pesadas e menor carga de trabalho	Não forçar o usuário a operações mentais pesadas e manter a carga de trabalho do usuário no mínimo.	Heurísticas de Nielsen e Princípios de Dimensões Cognitivas
10	Evite forçar o usuário a compromissos prematuros	Não forçar o usuário a executar uma determinada tarefa ou decisão até que seja necessário. Será que o usuário sabe por que algo deve ser feito?	Heurísticas de Nielsen e Princípios de Dimensões Cognitivas
11	Proporcionar funções que são úteis para o usuário	Considerar se a funcionalidade descrita é útil para ser usada pelos usuários e se funções/dados estão faltando.	Ideia de que as primeiras avaliações em casos de uso devem ajudar a estabelecer o uso das funcionalidades.

A segunda atividade tem o objetivo de avaliar a qualidade dos casos de uso, onde o avaliador expressa o porquê de algo não poder ser adequadamente analisado. Isso produz como resultado uma avaliação de cada caso de uso, que enfatiza o quão útil é o caso de uso para a inspeção.

E, por fim, na terceira atividade os resultados são compilados em um produto de avaliação coerente. O conteúdo principal da documentação é a lista de problemas de usabilidade, que descreve os problemas que os avaliadores esperam que um usuário vá enfrentar ao usar o sistema, sendo que cada um desses problemas de usabilidade deve incluir um motivo claro do porquê é percebido como tal.

Pode-se observar que duas Heurísticas de Nielsen foram excluídas da proposta de Hornbæk *et al.* (2007): “Projeto minimalista e estético” e “Ajuda e Documentação”. Isto se deve ao fato de que, segundo os autores, na especificação dos casos de uso não se têm a preocupação com detalhes do design da interface do usuário.

### 2.3.2. Método Diretrizes de Usabilidade (DU)

Da Silva e Silveira (2010) propuseram um método informal de avaliação em IHC, chamado de Diretrizes de Usabilidade (DU). Este método possui diretrizes (questões de verificação) que são utilizadas nas etapas iniciais do processo de desenvolvimento de sistemas interativos. As questões do método DU são apresentados na Tabela 4.

**Tabela 4. Questões do Método DU (DA SILVA e SILVEIRA, 2010)**

<b>Heurísticas e Questões</b>	
<b>Nº</b>	<b>Correspondência entre o sistema e o mundo real</b>
1	Em etapas onde é necessária uma entrada de dados, as tarefas são descritas em uma terminologia familiar ao usuário? (Por exemplo, solicitar login e senha)
2	A terminologia usada nas Atividades é familiar ao usuário?
3	A terminologia usada nas atividades a serem executadas pelo usuário são familiares?
4	As abreviaturas são significativas?
5	O sistema reflete o fluxo de trabalho do usuário?
<b>Nº</b>	<b>Visibilidade do estado do sistema</b>
6	Há algum feedback do sistema para cada ação do usuário?
7	Quando o usuário possui mais de uma opção (caminho a ser percorrido), isto fica claro?
8	As mensagens do sistema são sempre afirmativas e na voz ativa?
9	O usuário possui a opção de cancelar suas ações?
10	O usuário pode sair do sistema a qualquer momento?
<b>Nº</b>	<b>Reconhecimento em vez de lembrança</b>
11	São mostrados todos os dados que o usuário necessita em cada etapa da sequência de uma transação?
12	O sistema apresenta correspondência, ou seja, relações entre os controles e as ações são óbvias para o usuário?

A proposta é que o próprio projetista possa testar suas ideias, utilizando essas diretrizes para avaliar um sistema na etapa de modelagem e sem a necessidade de um especialista da área (DA SILVA e SILVEIRA, 2010). As diretrizes avaliam a usabilidade através de Diagrama de Atividades e possuem 21 questões, agrupadas nas seguintes heurísticas (Da SILVA e SILVEIRA, 2010): Correspondência entre o sistema e o mundo real, Visibilidade do estado do sistema, Reconhecimento em vez de lembrança e Prevenção de erros. Da Silva e Silveira (2010) realizaram estudos com este método e constataram que o diagrama de atividades possibilitou obter o fluxo de atividades que são relevantes para o desenvolvimento de *software*, assim como ajudou a encontrar problemas de usabilidade.

### 2.3.3. Abordagem dirigida por Modelos

Luna *et al.* (2010) apresentam uma abordagem de desenvolvimento que visa ser ágil, *model-based development approach* (MDSD) e apoia a especificação de requisitos e testes de usabilidade. A abordagem começa construindo testes de interação e navegação derivados de *mockups* de apresentação e *User Interaction Diagrams* (UIDs). Estes testes são depois executados na aplicação gerada pela ferramenta de desenvolvimento dirigida por modelos para verificar se são aceitos ou não.

O ciclo de desenvolvimento é dividido em *sprints*. No início do *sprint*, a equipe de desenvolvimento tem apenas um conjunto de especificações do que eles têm que fazer. Estas especificações foram definidas por meio de entrevistas com o usuário. Desenvolvedores começam a trabalhar, escolhendo uma delas de cada vez. As etapas deste ciclo são descritas a seguir:

- Etapa 1: Um pequeno ciclo começa com a captura de uma análise mais detalhada usando artefatos (modelos) de requisitos informal. Uma variedade de artefatos pode ser utilizada dependendo do tipo de requisito que se está captando:
  - Para requisitos envolvendo interações, usa-se UIDs que servem como uma especificação parcial da navegação da aplicação. *Mockups* são usados para aspectos da *User Interface* (UI), e *Use Case* (UC) ou *User Stories* (US) para aspectos do negócio ou de domínio.
  - Para os requisitos de usabilidade, usa-se um conjunto de propriedades de usabilidade derivadas de orientações de requisitos de usabilidade definidas na literatura. Estas propriedades de usabilidade são representadas por: UIDs para aspectos de navegação e *mockups* para aspectos UI.

- Etapa 2: Os artefatos que foram usados para capturar os requisitos são descritos em linguagem informal, portanto, os desenvolvedores transformam esses requisitos em testes. Antes que um requisito seja implementado, os testes correspondentes devem ser executados para verificar se a aplicação cumpre ou não com os requisitos. A falta de testes mostra que requisitos ainda não são apoiados pelo sistema em desenvolvimento. Se nesta fase a aplicação passar em todos os testes, então, ou eles não expressam o novo requisito de forma adequada, ou o novo requisito não é novo na realidade, porque a aplicação já o apoia. No primeiro caso, deve-se detalhar mais os testes voltando para a etapa 2 e no segundo, deve-se descartar o requisito e retornar a etapa 1.
- Etapa 3: Uma vez que o requisito foi especificado em um conjunto de testes, a fase de desenvolvimento pode começar. Usando uma abordagem MDSD, o desenvolvedor cria ou amplia os modelos existentes gerando uma maior versão da aplicação. A geração de código é realizada automaticamente por meio da transformação que toma como entrada os modelos.
- Etapa 4: A fim de verificar se o requisito foi implementado com sucesso e nenhuma funcionalidade anterior está corrompida, o programador executa todo o conjunto de testes para verificar ambas as coisas. Se um ou mais testes falharem, ele deve voltar para a etapa 3, fazer algum retrabalho nos modelos, gerar o código novamente, e repetir a etapa 4 até que todos os testes passem.
- Etapa 5: Finalmente, tem-se uma nova aplicação com um requisito adicional. O ciclo continua com a escolha de um novo requisito (Etapa 1) e seguindo as etapas de 2 a 5 até que terminem os requisitos do *sprint*.

## 2.4. Síntese do Capítulo

Este capítulo teve como objetivo apresentar os tipos de avaliações de usabilidade: Inspeção, Teste e Modelagem Analítica. Foram apresentadas tecnologias para serem utilizadas nas etapas iniciais do processo de desenvolvimento. As tecnologias abordadas avaliam modelos como especificações de caso de uso, diagrama de atividades e *mockups*.

Hornbæk *et al.* (2007) avaliou a usabilidade através de especificações de casos de uso. Estes modelos têm sido sugeridos como um meio valioso para a integração de engenharia de usabilidade diretamente no processo de desenvolvimento de *software* (FERRÉ *et al.*, 2001). Os casos de uso são frequentemente disponíveis no início do desenvolvimento de um sistema interativo, e são relevantes tanto para o desenvolvimento de *software* quanto

para o design de interface de usuário. Portanto, este modelo foi considerado importante no contexto desta pesquisa.

Os Diagramas de Atividades foram utilizados como modelo pelo método DU para avaliação de usabilidade (DA SILVA e SILVEIRA, 2010). Estes modelos visam mostrar o fluxo de atividades em um único processo. Além disso, este diagrama da UML mostra como as atividades dependem umas das outras. Portanto, têm-se evidências que este é um importante modelo para a descoberta de problemas de usabilidade.

A abordagem proposta por Luna *et al.* (2010) faz teste de interação a partir de *mockups* de apresentação. Estes *mockups* são modelos usados para aspectos da Interface do Usuário que funcionam como esboços das aplicações. Destinam-se a ser desenvolvidos rapidamente para refletir os desejos dos clientes em termos de apresentação de forma mais substancial do que os requisitos expressos em linguagem escrita (LUNA *et al.* 2010). Por este motivo, este modelo desempenha um papel importante para a avaliação de usabilidade.

Existem outros modelos importantes que precisam ser considerados em avaliações de usabilidade nas fases iniciais do processo de desenvolvimento de *software*. No entanto, para saber da existência de tecnologias que avaliam a usabilidade através de outros modelos e que integrem conceitos das áreas de IHC e ES, faz-se necessário caracterizar as tecnologias existentes. Para isso, foi realizado um mapeamento sistemático da literatura, descrito no próximo capítulo.

# CAPÍTULO 3 - MAPEAMENTO SISTEMÁTICO SOBRE TECNOLOGIAS QUE APOIAM A INTEGRAÇÃO ENTRE AS ÁREAS DE ENGENHARIA DE SOFTWARE E INTERAÇÃO HUMANO- COMPUTADOR

*Este capítulo apresenta um estudo secundário (Mapeamento Sistemático da Literatura) realizado com o propósito de caracterizar as tecnologias que apoiam a integração entre as áreas de IHC e ES no processo de desenvolvimento de software visando à melhoria da usabilidade. Neste capítulo são discutidos os principais resultados deste estudo secundário.*

## 3.1. Introdução

Para desenvolver tecnologias que apoiem projetistas de IHC ou engenheiros de *software* no processo de desenvolvimento de *software* e que visam a melhoria da qualidade das aplicações é necessário caracterizar as tecnologias já existentes, tendo como principal finalidade conhecer o seu funcionamento e em que modelos estas tecnologias podem ser utilizadas. Por este motivo, decidiu-se executar um levantamento na literatura das tecnologias que apoiam a inclusão da usabilidade dentro do processo de *software*. Por este motivo, optou-se por realizar um Mapeamento Sistemático da Literatura.

Um Mapeamento Sistemático da Literatura (MSL) é um tipo de Revisão Sistemática e, segundo Kitchenham e Charters (2007), é utilizado para prover uma ampla visão de uma área de pesquisa e estabelecer se existem evidências de pesquisas em um determinado tópico. Ao contrário das revisões informais da literatura, onde o pesquisador não segue um processo definido para sua condução, um Mapeamento Sistemático é executado de maneira formal obedecendo um protocolo pré-estabelecido. Em comparação com revisões informais da literatura, os mapeamentos sistemáticos requerem maior rigor na sua realização. Em compensação, seus resultados tendem a ser mais confiáveis, visto que eles fazem uso de uma metodologia rigorosa e passível de auditoria e repetição.

Sendo assim, um MSL foi realizado em conjunto com Williamson Alison Freitas Silva (SILVA *et al.*, 2015) com o objetivo de identificar (a) quais tecnologias<sup>1</sup> são utilizadas pelos projetistas de IHC ou engenheiros de *software*, (b) em que fase do processo do desenvolvimento elas podem ser aplicadas e (c) se elas são utilizadas dentro de um ciclo de vida, além de outras verificações. Este capítulo mostra como foi realizado o mapeamento, discutindo os resultados alcançados, apresentando um resumo das tecnologias encontradas. Ao apresentar este mapeamento, pretende-se também oferecer um levantamento sistemático que sirva como base de pesquisa para trabalhos futuros nas áreas de Engenharia de *Software* e Interação Humano-Computador.

### 3.2. Protocolo do Mapeamento Sistemático

O objetivo deste Mapeamento Sistemático segundo o Paradigma GQM (*Goal-Question-Metric*) (BASILI e ROMBACH, 1988) está definido na Tabela 5:

Tabela 5. Objetivo do 1º MSL segundo Paradigma GQM

<b>Analisar</b>	publicações científicas
<b>Com o propósito de</b>	identificar e analisar
<b>Em relação a</b>	tecnologias de IHC (Interação Humano-Computador) ou Engenharia de <i>Software</i> (ES) que promovem a usabilidade no processo de desenvolvimento de <i>software</i>
<b>Do ponto de vista dos</b>	Pesquisadores
<b>No contexto</b>	acadêmico e industrial com foco na integração das áreas de Interação Humano-Computador (IHC) e Engenharia de <i>Software</i> (ES) no processo de desenvolvimento de <i>software</i>

Com base neste objetivo, pretende-se responder a seguinte questão de pesquisa: **“Quais tecnologias melhoram a usabilidade no processo de desenvolvimento de *software*?”**. Através desta questão de pesquisa é possível classificar e categorizar o conhecimento atual sobre as tecnologias que possibilitam a integração das áreas de IHC e ES e que melhoram a usabilidade das aplicações interativas. Além disso, permite identificar novas áreas de investigações sobre o tema de pesquisa e proporcionar um conhecimento útil para os profissionais e pesquisadores de ambas as áreas.

Devido à questão de pesquisa ser bastante ampla, foram definidas subquestões para responder questionamentos específicos sobre a aplicabilidade de cada tecnologia (ver Tabela 6). Cada subquestão será explicada em detalhes na Subseção 3.2.3:

<sup>1</sup> O termo “tecnologia” é utilizado como generalização para procedimentos, ferramentas, técnicas, metodologias e outras propostas realizadas na área de Engenharia de *Software* e Interação Humano-Computador (SANTOS *et al.*, 2012).

Tabela 6. Subquestões de Pesquisa do 1º MSL

Nº SQ's	Descrição da Subquestão
<b>SQ1</b>	Tipo de tecnologia (método, técnica, modelo, ferramenta ou abordagem)
<b>SQ2</b>	Origem da tecnologia
<b>SQ3</b>	Contexto de uso
<b>SQ4</b>	Fase do processo de desenvolvimento em que a tecnologia é empregada
<b>SQ5</b>	Ciclo de vida específico
<b>SQ6</b>	Objeto Projetado/Avaliado
<b>SQ7</b>	Avaliação Empírica
<b>SQ8</b>	Apoio ferramental

### 3.2.1. Estratégia utilizada para pesquisa dos Estudos Primários/Secundários

A estratégia para pesquisa deve tornar explícito o escopo da pesquisa e os termos a serem utilizados nela. Os termos compõem as sequências de palavras-chave utilizados na *string* de busca (*search strings*).

- **Escopo da Pesquisa:** a pesquisa foi realizada nas seguintes bibliotecas digitais: IEEEExplore<sup>2</sup> e Scopus<sup>3</sup>. Estas bibliotecas foram escolhidas, pois:
  - Permitem o uso de expressões lógicas para as buscas ou mecanismo similar para localizar as publicações através do título e resumo das publicações;
  - Possuem um bom funcionamento e abrangência de suas máquinas de busca;
  - Contêm em suas bases, publicações de variadas áreas de conhecimento;
  - A Scopus é uma das maiores bases de dados de indexação de resumos e citações (KITCHENHAM e CHARTERS, 2007);
- **Idioma dos artigos:** os idiomas escolhidos foram o Inglês e o Português. O inglês por ser adotado pela grande maioria das conferências e periódicos internacionais relacionados ao tema da pesquisa. Além disso, é o idioma utilizado pela maioria das editoras listadas no Portal de Periódicos da CAPES. O português foi escolhido por ser utilizado em conferências nacionais, como o IHC (Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais) e o SBES (Simpósio Brasileiro de Engenharia de *Software*), nos quais o tema pesquisado também é abordado.
- **Termos utilizados:** com a finalidade de melhorar e estruturar a busca nas bibliotecas digitais selecionadas utilizou-se o PICOC (*Population, Intervention, Comparison, Outcome e Context*), conforme Kitchenham e Charters (2007) sugerem:
  - **(P) Population:** Processo de desenvolvimento de *software*;

<sup>2</sup> <http://ieeexplore.ieee.org/>.

<sup>3</sup> <http://www.scopus.com/home.url>.

- **(I) *Intervention***: tecnologias de IHC e ES que são utilizadas no processo de desenvolvimento de *software*;
- **(C) *Comparison***: Não se aplica, pois o objetivo não é realizar uma comparação entre as tecnologias, e sim, caracterizá-las.
- **(O) *Outcome***: A melhoria da usabilidade na aplicação através dos modelos construídos utilizando as tecnologias que projetam/avaliam a usabilidade;
- **(C) *Context***: Não se aplica, pois como não se deseja fazer comparação entre as tecnologias, não é necessário determinar um contexto.

Logo em seguida, foram identificados termos que instanciassem a População (***Population***), Intervenção (***Intervention***) e Resultado (***Outcome***) e foi elaborada uma *string* de busca. Os termos estão descritos na língua inglesa, por este ser o idioma utilizado pelas máquinas de buscas. Os termos que foram utilizados neste mapeamento foram agrupados em três grupos que combinados entre si formam a *string* de busca. A Tabela 7 mostra os termos e a *string* de busca utilizados para esta pesquisa. O caractere booleano OR tem sido usado para juntar outros termos, enquanto o caractere booleano AND tem sido usado para unir os três conceitos (População, Intervenção e Resultados).

**Tabela 7. Termos e *String* de busca utilizados no 1º MSL**

<b>Termos e <i>String</i> de busca utilizados</b>		
População	(software development OR software project OR software engineering OR software process)	AND
Intervenção	(technique OR method OR methodology OR tool)	AND
Resultados	(usability inspection OR usability evaluation OR usability design OR usability testing)	

O período de busca incluiu publicações produzidas até o ano de 2013. Como a pesquisa foi realizada em janeiro de 2014, as publicações relativas a esse ano não foram consideradas neste mapeamento sistemático.

### 3.2.2. Critérios de Seleção de Artigos e Procedimentos

Kitchenham e Charters (2007) dizem que devem ser seguidos critérios de inclusão e exclusão para os artigos identificados. Sendo assim, foram definidos os seguintes critérios:

Os critérios de Inclusão são:

- **CI1**. Podem ser selecionadas publicações que descrevam quais tecnologias que são utilizadas para promover a usabilidade no processo de desenvolvimento de *software*;
- **CI2**. Podem ser selecionadas publicações que apresentam apoio ferramental para antecipar a usabilidade no processo de *software*;

- **CI3.** Podem ser selecionadas publicações que discutam aspectos relacionados à inclusão da usabilidade no processo de *software*;
- **CI4.** Podem ser selecionadas publicações que apresentam a melhoria da usabilidade em uma das fases do processo de *software* aplicável em alguma organização;

Os critérios de Exclusão são:

- **CE1.** Não serão selecionadas publicações em que o idioma não seja o inglês e português;
- **CE2.** Não serão selecionados artigos que não estão disponíveis para a leitura e coleta dos dados (artigos pagos ou não disponibilizados pela máquina de busca);
- **CE3.** Não serão selecionadas publicações duplicadas, ou seja, publicações já selecionadas a partir de outra biblioteca digital;
- **CE4.** Não serão selecionadas publicações que não atendam aos critérios de inclusão

Quanto ao procedimento de seleção preliminar (1º Filtro), decidiu-se que seriam selecionados artigos que apresentassem informações no título e no abstract sobre tecnologias de IHC e ES que são utilizadas com o intuito de promover a usabilidade no processo de desenvolvimento de *software*, utilizando os critérios de inclusão e exclusão. Em caso de dúvida, o artigo seria incluído. Para cada estudo foi apresentado um dos critérios.

Em relação ao procedimento de seleção final (2º Filtro), realizou-se a leitura completa dos artigos que restaram da seleção preliminar. Dessa forma, esta fase tem como objetivo fazer uma análise mais detalhada dos estudos, identificando e extraindo dados também de acordo com os critérios de inclusão e exclusão.

### 3.2.3. Estratégia para extração de dados

Após a definição da pesquisa e os processos de seleção dos artigos, o processo de extração de dados foi realizado através da leitura completa de cada um dos trabalhos selecionados. Para a extração adotou-se a estratégia de fornecer um conjunto de possíveis respostas. Fernandez *et al.* (2011a) afirmam que esta estratégia garante a aplicação dos mesmos critérios de extração de dados para todos os trabalhos selecionados, além de facilitar a classificação. De cada artigo aprovado pelo processo de seleção final (2º filtro) foram extraídas informações de acordo com as Subquestões:

Com respeito à **SQ1** (Tipo de tecnologia), a tecnologia encontrada no artigo pode ser um método, técnica, modelo, ferramenta, abordagem ou outro procedimento adotado.

No que diz respeito à **SQ2** (Origem da tecnologia), o artigo pode ser classificado de acordo com as seguintes respostas:

- a) **Nova:** se o artigo apresenta uma tecnologia, mas esta não se baseia em outras tecnologias das áreas de IHC e ES;
- b) **Existente:** se o artigo apresenta uma tecnologia, mas esta foi proposta com base em outras tecnologias das áreas de IHC e ES.

No que diz respeito à **SQ3** (Contexto de uso), o artigo pode ser classificado de acordo com as seguintes respostas:

- a) **Indústria:** se a tecnologia apresentada no artigo foi utilizada ou avaliada no contexto industrial;
- b) **Academia:** se a tecnologia apresentada no artigo foi utilizada ou avaliada no contexto acadêmico;
- c) **Ambas:** se a tecnologia apresentada no artigo foi utilizada ou avaliada tanto no contexto industrial quanto acadêmico.

No que diz respeito à **SQ4** (Fase do processo de desenvolvimento em que a tecnologia é empregada), a tecnologia encontrada no artigo pode ser adotada em uma ou mais fases descritas pelo SWEBOK (2004):

- a) **Requisitos:** quando a tecnologia é utilizada na fase onde são desenvolvidos modelos que visam a identificação das necessidades dos usuários. Modelos como: casos de uso, modelos de tarefas e cenários de uso;
- b) **Projeto:** quando a tecnologia é utilizada na fase onde são desenvolvidos modelos que projetam a aplicação. Estes modelos são construídos antes da codificação (por exemplo: diagramas de atividades e modelos de navegação);
- c) **Construção (Codificação):** quando a tecnologia é utilizada durante a codificação da aplicação;
- d) **Verificação, Validação & Teste:** a tecnologia auxilia a: (a) verificar se o produto corresponde aos requisitos do usuário (Verificação), (b) assegurar a consistência, completude e corretude da aplicação (Validação); e (c) examinar o comportamento da aplicação por meio de sua execução (Teste);
- e) **Manutenção:** a tecnologia é utilizada durante a manutenção da aplicação.

Com respeito à **SQ5** (Ciclo de vida específico), é verificado se a tecnologia encontrada no artigo é abordada em algum ciclo de vida específico e qual é o ciclo de vida:

- a) **Sim:** a tecnologia é utilizada em um ciclo de vida específico;
- b) **Não:** a tecnologia não é utilizada em um ciclo de vida específico.

Em relação à **SQ6** (Objeto Projetado/Avaliado), a tecnologia pode ser classificada de acordo com o objeto em que a tecnologia pode ser empregada. Por exemplo, protótipos, aplicações *Web*, aplicações *Mobile*, dentre outros.

De acordo com a **SQ7** (Avaliação Empírica), a tecnologia apresentada no artigo pode ser classificada de acordo com as seguintes respostas:

- a) **Não:** não é descrito nenhum tipo de avaliação empírica (Estudo de Caso, Experimento Controlado ou outros) com a tecnologia;
- b) **Sim:** é descrito algum tipo de avaliação empírica (Estudo de Caso, Experimento Controlado ou outros) com a tecnologia.

Na **SQ8** (*Apoio ferramental*), o artigo pode ser classificado como:

- a) **Sim:** a tecnologia apresentada no artigo requer algum apoio ferramental;
- b) **Não:** a tecnologia apresentada no artigo não requer um apoio ferramental.

O pacote contendo mais informações sobre este MSL, assim como os artigos selecionados em cada fase e os campos de extração estão disponíveis no Relatório Técnico disponível em Silva *et al.* (2014a).

### 3.2.4. Artigos Selecionados após a Condução do Mapeamento Sistemático

A Figura 3 apresenta a quantidade de artigos retornados por cada máquina de busca das bibliotecas digitais selecionadas, o total de artigos que foram selecionados no 1º Filtro e o total de artigos selecionados após o 2º Filtro. Como resultado quantitativo este MSL obteve um total de 124 artigos selecionados. Estes artigos foram selecionados com base nos critérios de inclusão (ver Subseção 3.2.2).

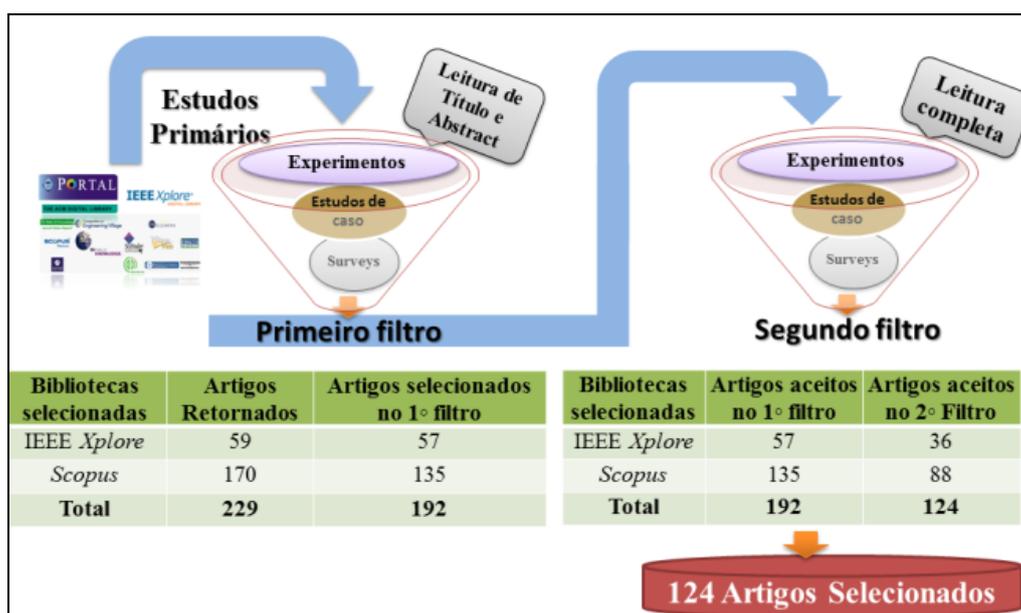


Figura 3. Total de artigos no 1º filtro e 2º filtro no 1º MSL

Um ponto importante que precisa ser destacado é que a biblioteca Scopus retornou um alto número de artigos sem acesso gratuito ao seu conteúdo através do portal de periódicos CAPES, impossibilitando assim a sua leitura. Outro ponto é que houve artigos duplicados nas bibliotecas. Porém, estes artigos foram contabilizados apenas uma vez e de acordo com a ordem de busca realizada, que foi a seguinte: IEEEExplore e Scopus.

Todas as informações mostradas neste Capítulo podem ser encontradas em maior detalhe no Relatório Técnico criado por Silva *et al.* (2014a). Neste relatório são apresentadas as informações relacionadas ao planejamento, execução e análise dos dados encontrados a partir do mapeamento sistemático executado. Também está disponível a tabela com os dados de extração dos artigos selecionados neste mapeamento.

### 3.3. Resultados obtidos

#### 3.3.1. Visão Geral do Resultados

Uma visão geral dos resultados baseado no número de tecnologias classificadas em cada resposta das Subquestões é apresentada na Tabela 8.

Tabela 8. Resultados do 1º MSL para cada uma das Subquestões

Subquestão de Pesquisa	Possíveis Respostas	Resultados	
		Artigos	Porcentagem (%)
SQ1. Tipo de Tecnologia	Métodos	50	40,65
	Ferramentas	25	20,33
	<i>Frameworks</i> /Abordagens	23	18,70
	Técnicas	16	13,01
	Modelos	6	4,88
	Metodologias	3	2,44
SQ2. Origem da Tecnologia	Existente	111	90,24
	Nova	12	9,76
SQ3. Contexto de uso	Academia	87	70,73
	Indústria	27	21,95
	Ambos	9	7,32
SQ4. Fase do Processo de desenvolvimento em que a tecnologia é empregada	Requisitos	22	13,58
	Projeto	27	16,67
	Construção	4	2,47
	Verificação, Validação & Teste	109	67,28
	Manutenção	0	0,00
SQ5. Ciclo de Vida Específico	Não	106	86,18
	Sim	17	13,82
SQ6. Objeto Projetado/Avaliado	Aplicações ( <i>Web, Mobile, Desktops</i> e aplicações não especificadas)	76	50,67
	Interfaces, <i>Mockups</i> ou Protótipos	43	28,67
	Modelos	24	16,00
	Outros objetos	7	4,66

Subquestão de Pesquisa	Possíveis Respostas	Resultados	
		Artigos	Porcentagem (%)
SQ7. Avaliação Empírica	Sim	90	73,17
	Não	3	26,83
SQ8. Apoio Ferramental	Não	95	77,24
	Sim	28	22,76

Apesar de 124 artigos serem selecionados, a maioria das Subquestões da Tabela 8 contabilizam 123 artigos. Isto acontece porque um dos artigos selecionados é um MSL que investiga que métodos de avaliação de usabilidade têm sido empregados pelos pesquisadores para avaliar artefatos *Web* e como esses métodos foram usados (FERNANDEZ *et al*, 2011a). Por este motivo, não foi possível extrair os dados deste artigo e agrupá-los nesta tabela de classificação juntamente com as tecnologias que eram apresentadas nos outros artigos. Portanto, decidiu-se mantê-lo fora desta tabela de resultados. Para este tipo de artigo identificado (MSL), elaborou-se outro tipo de extração de dados, que se encontra disponível em Silva *et al*. (2014a).

As seguintes subseções apresentam a análise dos resultados obtidos a partir de cada subquestão de pesquisa. Nota-se que a SQ4 e a SQ6 possuem dados que não são exclusivos. Nestas subquestões as tecnologias podem ser classificadas em uma ou mais respostas, logo, a soma das percentagens é maior que 100%. Por exemplo, na SQ4 algumas tecnologias podem ser utilizadas em mais de uma fase no processo de desenvolvimento.

### 3.3.2. Ano de Publicação

Os artigos selecionados foram publicados entre 1988 e 2013. Do ponto de vista temporal (Figura 4), houve um aumento do número de publicações entre os anos de 2005 e 2007. Percebe-se também, de acordo com as publicações coletadas neste MSL, que nos anos de 2008 e 2011 houve um decréscimo nas publicações.

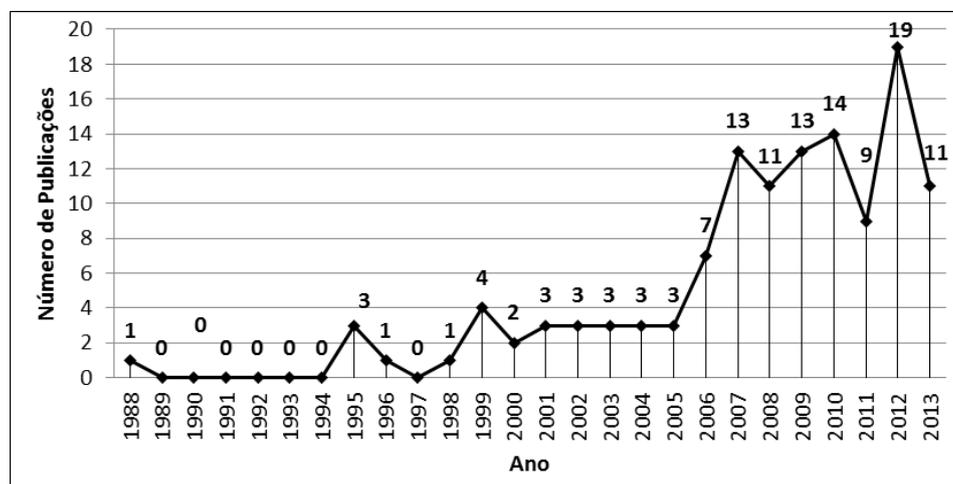


Figura 4. Visão Temporal dos artigos selecionados no 1º MSL

O ano de 2012 (15,32%) é o ano com a maior quantidade de publicações, seguido de 2010 (11,29%), 2009 (10,48%) e 2007 (10,48%). Como este MSL foi realizado em janeiro de 2014, nem todas as conferências do ano de 2013 tinham indexado suas publicações nas bibliotecas digitais pesquisadas. Isto pode ser um dos motivos para o baixo número de artigos neste ano.

### 3.3.3. Tipo de Tecnologia

Os resultados para a Subquestão **SQ1** (Tipo de Tecnologia) mostraram que 40,65% dos artigos selecionados apresentam métodos que promovem a usabilidade no processo de desenvolvimento. Como exemplo destes métodos pode-se citar o trabalho de Hornbæk *et al.* (2007), que propuseram um método para avaliação de usabilidade de casos de uso, chamado UCE (*Use Case Evaluation*), já descrito no Capítulo 2 (Subsubseção 2.3.1).

Cerca de 20,33% dos artigos apresentam alguma ferramenta que auxilia a melhorar a usabilidade no processo de desenvolvimento. Um exemplo destas ferramentas pode ser encontrada em Vaz *et al.* (2012), que apresentam a WDT Tool, que é uma ferramenta que auxilia na identificação de problemas de usabilidade em aplicações *Web*.

Além disso, 18,70% dos artigos apresentam algum *framework*, 13,01% abordam técnicas, 4,88% apresentam modelos e 2,44% dos artigos apresentam metodologias.

### 3.3.4. Origem da Tecnologia

Os resultados para a Subquestão **SQ2** (Origem das tecnologias) mostraram que cerca de 9,76% das tecnologias apresentadas não utilizaram como base outras tecnologias. Por exemplo, Pankratius (2011) apresenta uma ferramenta chamada Linsight Suite. A ferramenta possui um plug-in que integra questionários na IDE Eclipse, com o objetivo de coletar informações subjetivas do programador com relação à aplicação que está sendo desenvolvida, enquanto este realiza a codificação da mesma.

Cerca de 90,24% das tecnologias selecionadas basearam-se em outras tecnologias já existentes na literatura, seja da área de IHC ou ES. Por exemplo, Conte *et al.* (2007a), propuseram uma técnica que avalia a usabilidade de aplicações *Web*, a WDP (*Web Design Perspectives*). Esta técnica combina as perspectivas de projeto *Web* (Apresentação, Conceituação e Navegação) com as heurísticas propostas por Nielsen (1994b), como forma de guiar a inspeção de usabilidade em aplicações *Web*.

Os resultados desta Subquestão (**SQ2**) e da Subquestão anterior (**SQ1**) indicam que diversas tecnologias estão sendo propostas na literatura. Estas tecnologias têm como objetivo principal auxiliar, tanto os designers de IHC quanto os engenheiros de *software* na

melhoria da usabilidade no processo de desenvolvimento de aplicações interativas, seja avaliando ou projetando a usabilidade em aplicações interativas.

### 3.3.5. Contexto de uso

Os resultados desta Subquestão **SQ3** (Contexto de Uso) mostraram que cerca de 70,73% das tecnologias selecionadas foram utilizadas no contexto acadêmico. Por exemplo, Fernandes *et al.* (2012) realizaram dois estudos (viabilidade e observacional) da tecnologia que propuseram. Os participantes deste estudo foram alunos de graduação de uma universidade. Cerca de 21,95% das tecnologias foram aplicadas no contexto industrial. Por exemplo, Sivaji *et al.* (2013) realizaram dois estudos de caso, em dois projetos de *software*, onde especialistas em usabilidade realizavam a avaliação em ambos os projetos utilizando a tecnologia proposta pelos autores. Além disso, 7,32% das tecnologias foram utilizadas tanto na academia quanto indústria.

Os resultados desta Subquestão mostram que a maioria das tecnologias encontradas neste MSL estão sendo avaliadas na academia. Este fato pode estar relacionado ao número reduzido de profissionais que se dispõem a participar das avaliações. Além disso, disponibilizar parte do tempo dos profissionais para realizar estas avaliações é mais custoso para a indústria. Uma das soluções encontrada pelos pesquisadores é realizar avaliações com estudantes, uma vez que muitos deles estão atuando ou já atuaram na indústria. Além disso, estudantes que não possuem experiência na indústria podem apresentar habilidades similares a profissionais menos experientes (CARVER *et al.*, 2003).

### 3.3.6. Fase do processo de desenvolvimento em que a tecnologia é empregada

Os resultados desta Subquestão **SQ4** (Fase do processo de desenvolvimento em que a tecnologia é empregada) revelaram que 67,28% das tecnologias são utilizadas na fase de V, V&T. Nesta fase, os profissionais utilizam as tecnologias após a aplicação ter sido desenvolvida, ou, possivelmente, quando uma versão do protótipo funcional da aplicação está pronta. Na fase de Verificação, Validação & Teste (V, V & T), as tecnologias foram divididas em duas categorias: (1) Inspeção de Usabilidade e (2) Teste de Usabilidade. Estas duas categorias já foram descritas no Capítulo 2 desta tese (Subseção 2.2).

Do total de tecnologias utilizadas na fase de V, V&T, 44,95% são tecnologias para Inspeção de Usabilidade. Por exemplo, Fernandes *et al.* (2012) apresentam uma técnica de inspeção, chamada WE-QT (*Web Evaluation-Question Technique*), que auxilia inspetores novatos a identificar problemas de usabilidade de aplicações *Web*, por meio de perguntas e respostas. Do total das tecnologias utilizadas na fase de V, V&T., 55,06% são tecnologias

para Teste de Usabilidade. Por exemplo, a ferramenta proposta por Fabo e Durikovic (2012) identifica problemas de usabilidade através da captura automática de dados (tempo de sucesso, distância percorrida, cliques do mouse e tempo de execução das atividades).

Cerca de 16,67% dos artigos desta Subquestão apresentam tecnologias que podem ser utilizadas na fase de Projeto, ou seja, utilizam os modelos produzidos durante as fases iniciais do processo de desenvolvimento (como diagramas de atividades, dentre outros). Por exemplo, Rivero e Conte (2012) apresentam a técnica *Web DUE (Web Design Usability Evaluation)*, que tem por finalidade avaliar *mockups* de aplicações *Web*.

Cerca de 13,58% das tecnologias podem ser utilizados na etapa Requisitos. Por exemplo, Ormeno *et al.* (2013) apresentaram um novo método para capturar os requisitos de usabilidade nas fases iniciais do processo de desenvolvimento de *software*, de tal maneira que profissionais não especialistas em usabilidade possam utilizá-lo. Após realizar a captura desses requisitos, a abordagem permite transformar esses requisitos de usabilidade em um modelo conceitual de qualquer método MDD (*Model-Driven Development*).

Cerca de 2,47% das tecnologias podem ser utilizadas enquanto os desenvolvedores realizam a codificação da aplicação. Por exemplo, Pankratius (2011) apresenta uma ferramenta que visa automatizar avaliação de usabilidade através de um plugin (*Linsight Suite*) que integra questionários na IDE Eclipse. Quando o desenvolvedor seleciona um pedaço do código, um conjunto de perguntas associadas àquele código é aberto, auxiliando o programador a relatar problemas de usabilidade do código.

Neste mapeamento, nenhuma tecnologia que é utilizada na fase de Manutenção foi identificada. No geral, os resultados desta Subquestão indicam que há poucas tecnologias que possam ser utilizadas nas fases iniciais do processo de desenvolvimento (Requisitos e Projeto, isto é, antes da aplicação ser codificada). Muitas tecnologias focam em avaliações nas fases finais (Inspeções e Testes de Usabilidade – 68,28%). No entanto, quanto mais tarde questões de usabilidade são tratadas na aplicação, mais tarde também serão identificados os problemas de usabilidade. E isto pode acarretar maior custo na correção dos problemas de usabilidade ou aumento do tempo dos profissionais para desenvolvimento e manutenção da aplicação interativa.

### 3.3.7. Ciclo de vida específico

Os resultados para a Subquestão **SQ5** (Ciclo de vida específico) revelou que 86,18% das tecnologias não são utilizadas em um ciclo de vida específico. Desta forma, tais tecnologias podem ser adequadas para serem utilizadas em qualquer ciclo de vida de desenvolvimento adotados pela indústria.

Por outro lado, 13,82% dos artigos apresentam tecnologias que são utilizadas em um ciclo de vida específico. Por exemplo, Sivaji *et al.* (2013) apresentam uma abordagem híbrida que integra a Avaliação Heurística e o Teste de Usabilidade em um ciclo de vida de desenvolvimento específico. Este ciclo de vida é iterativo e possui as seguintes fases: Requisitos, Projeto, Teste para Prototipação Inicial, Reprojeto & Desenvolvimento e Solução de Defeitos, Teste para Prototipação Intermediária, Reprojeto & Desenvolvimento e Solução de Defeitos, Teste para Prototipação Avançada e Implantação.

Os resultados para esta subquestão indicam que 86,18% das tecnologias são usadas independentemente do ciclo de vida adotado pela equipe de desenvolvimento. No entanto, 13,82% das tecnologias selecionadas estão usando um ciclo de vida específico (por exemplo, o *Model Driven Development - MDD*).

### 3.3.8. Objeto Projetado/Avaliado

Os resultados para a Subquestão **SQ6** (Objeto Projetado/Avaliado) revelou que 50,67% das tecnologias avaliam/projetam a usabilidade de aplicações já desenvolvidas. Do total de tecnologias que utilizam aplicações como objeto, 52,65% destas tecnologias utilizam aplicações *Web*, 26,30% aplicações Desktop, 3,94% aplicações Móveis e 17,12% não especificavam em que tipo de objeto a tecnologia foi aplicada. Como exemplos de trabalhos que apresentam tecnologias que utilizam as Aplicações *Web* como objeto, pode-se citar: Sivaji *et al.* (2013), Vaz *et al.* (2012), Ramli e Jaafar (2010) e Conte *et al.* (2007a). E, como exemplo de trabalho que utiliza uma tecnologia que é empregada a aplicações desktop tem-se o trabalho de Fabo e Durikovic (2012).

Cerca de 28,67% dos artigos utilizaram como objeto as interfaces, *mockups* ou protótipos das aplicações. Um exemplo é o método proposto por Ormeno *et al.* (2013) que realiza a coleta de requisitos de usabilidade através de interfaces de aplicações móveis.

Cerca de 16% das tecnologias empregavam modelos como objetos, tais como: modelos de tarefas e modelos de navegação. E, 4,66% dos artigos utilizaram alguns outros tipos de objetos. Dentre esses, 2% dos artigos utilizavam linhas de código como objeto as linhas de código, 2% dos artigos utilizavam arquivos de log e, 0,66% utilizavam tarefas dos usuários.

Os resultados para esta Subquestão indicam que muitas tecnologias estão sendo desenvolvidas para melhorar a usabilidade de aplicações, especialmente voltadas para a *Web*. No entanto, um ponto a ser levado em consideração é o baixo número de tecnologias que auxiliam a melhorar a usabilidade em aplicações móveis. Com o crescimento do uso de dispositivos móveis, as aplicações móveis têm se tornado cada vez mais presentes entre os

usuários. Com base nisso, faz-se necessárias novas pesquisas nesta área, a fim de investigar e propor tecnologias que visam melhorar a usabilidade e que possam ser utilizadas no processo de desenvolvimento das aplicações móveis.

### 3.3.9. Avaliação Empírica

Os resultados para esta Subquestão **SQ7** (Avaliação Empírica) revelaram que em 26,83% das tecnologias selecionadas não foi realizada nenhum tipo de avaliação empírica. Os artigos apenas faziam uma descrição da tecnologia ou indicava como utilizá-la. Cerca de 73,17% das tecnologias foram avaliadas empiricamente. Por exemplo, Santos e Conte (2011) apresentam um assistente que apoia a inspeção de usabilidade chamado APIU (Apoio ao Processo de Inspeção de Usabilidade) e descrevem a evolução do assistente através de estudos empíricos, bem como as melhorias realizadas no assistente.

Os resultados para esta Subquestão mostram que os autores estão realizando estudos empíricos nas tecnologias que estão propondo. Realizar estudos empíricos é uma prática comum nas áreas de Interação Humano-Computador e Engenharia de *Software* (FERNANDEZ *et al.*, 2012a). Estas duas áreas têm se preocupado em melhorar as tecnologias propostas para que estas auxiliem os profissionais a projetar/avaliar a usabilidade das aplicações.

### 3.3.10. Apoio Ferramental

Os resultados para a Subquestão **SQ8** (Apoio ferramental) revelou que 22,76% dos artigos apresentam tecnologias que necessitam de uma ferramenta ou *framework* para auxiliar no uso pelos profissionais. Conforme foi citado anteriormente, têm-se os trabalhos de Vaz *et al.* (2012) e Santos e Conte (2011). No entanto, cerca de 77,24% dos artigos apresentam tecnologias que não necessitam de um apoio ferramental.

Como resultados, percebeu-se que as tecnologias encontradas neste mapeamento em sua minoria (22,76%) necessitam de apoio ferramental. Tal característica se traduz em um aumento do emprego de tecnologias que não requerem apoio ferramental. Por outro lado, as ferramentas podem aumentar o desempenho, diminuir a sobrecarga e facilitar o trabalho dos profissionais no processo de desenvolvimento. Portanto, tecnologias com apoio ferramental que estejam disponíveis para o uso podem diminuir o esforço dos profissionais e, conseqüentemente, trazer alguns benefícios para a indústria.

### 3.4. Síntese do Capítulo

Este capítulo descreveu os resultados de um estudo secundário (MSL). Este estudo discutiu as evidências existentes sobre as tecnologias propostas pelas áreas de Interação Humano-Computador (IHC) e Engenharia de *Software* (ES) que podem ser utilizadas dentro do processo de desenvolvimento de *software* para promover a usabilidade da aplicação. A partir de um conjunto inicial de 229 artigos, um total de 124 trabalhos de pesquisa foram selecionados neste mapeamento, após o 1º e o 2º filtro.

Este mapeamento contribuiu com resultados que são úteis para a promoção e melhoria da prática atual de pesquisa sobre a integração entre as áreas de IHC e ES, além de fornecer um corpo de conhecimento de tecnologias que auxiliam na melhoria da usabilidade através de diversos modelos gerados no processo de desenvolvimento.

Os resultados obtidos neste MSL identificaram várias tecnologias de IHC que têm como foco principal apoiar os designers de IHC e engenheiros de *software* a melhorar a usabilidade das aplicações finais. Foram identificadas as seguintes lacunas: (i) necessidade de novas tecnologias para apoiar a usabilidade nas fases iniciais do processo de desenvolvimento, pois corrigir problemas de usabilidade nestas fases é menos custoso e evita esforço dos profissionais com retrabalho; (ii) necessidade de novas tecnologias que auxiliem na antecipação de usabilidade de aplicações móveis, uma vez que o desenvolvimento deste tipo de aplicação está crescendo consideravelmente; e (iii) necessidade de novas tecnologias que auxiliem a projetar as aplicações já visando à sua usabilidade, uma vez que esta é uma maneira proativa de considerar a usabilidade.

# CAPÍTULO 4 - MAPEAMENTO SISTEMÁTICO SOBRE TECNOLOGIAS QUE AUXILIAM NO PROJETO/CRIAÇÃO OU AVALIAÇÃO DE MODELOS VISANDO À USABILIDADE DA APLICAÇÃO

*Este capítulo apresenta um mapeamento sistemático de tecnologias que auxiliam na criação ou avaliação de modelos das fases iniciais do processo de desenvolvimento visando à usabilidade, abordando a seguinte questão de pesquisa: “Quais tecnologias tem sido usadas durante a criação e/ou avaliação de modelos nos estágios iniciais do processo de desenvolvimento visando à usabilidade da aplicação final?” Além disso, foi descrito como o mapeamento sistemático foi realizado e quais são os resultados obtidos. Em seguida, discute-se as implicações para pesquisadores e profissionais, bem como as limitações do estudo.*

## 4.1. Introdução

Como dito anteriormente, a avaliação de usabilidade geralmente ocorre no final do processo de desenvolvimento de *software*. Por este motivo, possíveis problemas de usabilidade deixam de ser encontrados e tratados logo no início do ciclo de desenvolvimento, especificamente antes da codificação da aplicação, onde são mais baratos de resolver. Identificar problemas de usabilidade cedo, no entanto, é difícil com a prática atual de desenvolvimento de *software*, porque o trabalho de usabilidade é geralmente separado das atividades do núcleo de desenvolvimento de *software* (HORNBAEK *et al.*, 2007).

Os resultados do Mapeamento Sistemático da Literatura (MSL) descritos no Capítulo 3 apontaram a necessidade de propor novas tecnologias que apoiem a usabilidade das aplicações a partir das fases iniciais do processo de desenvolvimento. Portanto, a fim de identificar tecnologias de usabilidade mais específicas para as fases iniciais do processo de *software*, foi realizado um 2º MSL com o seguinte objetivo descrito na próxima seção.

## 4.2. Protocolo do Mapeamento Sistemático

### 4.2.1. Objetivo

Tabela 9 mostra o objetivo segundo o paradigma GQM (BASILI e ROMBACH, 1988).

**Tabela 9. Objetivo do 2º MSL segundo Paradigma GQM**

<b>Analisar</b>	estágios iniciais de projetos de desenvolvimento de <i>software</i>
<b>Com o propósito de</b>	caracterizar
<b>Em relação a</b>	tecnologias de IHC e ES que apoiam a criação/avaliação de usabilidade
<b>Do ponto de vista dos</b>	pesquisadores de IHC e ES
<b>No contexto</b>	fontes primárias disponíveis no mecanismo de busca da ACM e SCOPUS

### 4.2.2. Questão de Pesquisa

A questão de pesquisa principal investigada neste mapeamento é a seguinte: **“Quais tecnologias têm sido usadas durante a criação e/ou avaliação de modelos nos estágios iniciais do processo de desenvolvimento visando à usabilidade da aplicação final?”**.

Além desta questão de pesquisa, foram definidas subquestões, para responder questionamentos específicos sobre a aplicabilidade de cada tecnologia, apresentadas na Tabela 10.

**Tabela 10. Subquestões de pesquisa do 2º MSL**

<b>Subquestões de Pesquisa</b>	<b>Objetivo</b>
SQ1. Estágio do processo de desenvolvimento	Investigar os estágios iniciais do processo de desenvolvimento em que as tecnologias podem ser usadas
SQ2. Área de pesquisa	Investigar se as tecnologias propostas são da área de IHC e/ou ES
SQ3. Tipo de tecnologia	Descobrir se a tecnologia de usabilidade avalia ou cria um modelo
SQ4.1 Tipo de tecnologia de avaliação de usabilidade	Descobrir quais são os tipos de Tecnologias de Avaliação de Usabilidade
SQ4. Modelos em que as tecnologias são aplicadas	Investigar os modelos em que as tecnologias são aplicadas
SQ5. Tipo de contribuição	Descobrir a principal contribuição das fontes primárias da tecnologia
SQ6. Apoio Ferramental	Investigar quais tecnologias necessitam de apoio ferramental
SQ7. Estudos experimentais	Descobrir quais tecnologias têm sido experimentalmente avaliadas
SQ7.1 Tipo de estudo experimental	Descobrir quais são os tipos de estudos experimentais realizados com as tecnologias
SQ7.2 Ambiente de avaliação	Investigar quais tecnologias têm sido validadas em ambientes acadêmicos e/ou industrial
SQ7.3 Tipo de análise	Descobrir quais tecnologias têm sido analisadas de maneira quantitativa e/ou qualitativa
SQ8. Contexto de aplicação	Descobrir quais tecnologias identificadas são específicas ou genéricas

Estas Subquestões de pesquisa permitiram: categorizar e resumir os conhecimentos atuais sobre “*Early Usability*”; identificar lacunas na pesquisa atual para sugerir áreas para investigação futura e fornecer conhecimentos úteis para os profissionais de usabilidade.

#### 4.2.3. Estratégia utilizada para pesquisa dos estudos primários

- Escopo da Pesquisa

A pesquisa foi feita a partir de duas bibliotecas digitais através de seus mecanismos de busca avançada. A Tabela 11 mostra as fontes consideradas no escopo desta pesquisa.

Tabela 11. Fontes utilizadas no 2º MSL

Nome da Fonte	Link	Tipo de Pesquisa
ACM Digital Library	<a href="http://portal.acm.org/dl.cfm">http://portal.acm.org/dl.cfm</a>	Máquina de Busca
Scopus	<a href="http://www.scopus.com/home.url">http://www.scopus.com/home.url</a>	Máquina de Busca

Estas bibliotecas foram escolhidas porque: (1) elas permitem uma boa operação e escopo de seus mecanismos de busca; (2) ACM também indexa algumas publicações da Springer Link, Science Direct, e muitas publicações relacionadas à área de IHC; e (3) Scopus é uma das maiores bases de dados que indexa resumos e citações (KITCHENHAM e CHARTERS, 2007). A biblioteca IEEEExplore não foi utilizada neste MSL, pois ela retornou um número extremamente grande de artigos, o que inviabilizou o seu uso.

- Idioma dos artigos

Os idiomas escolhidos foram o Inglês e Português. Inglês, pois é o idioma adotado pela grande maioria das conferências e periódicos internacionais relacionados com tema de pesquisa e por ser o idioma utilizado pela maioria das editoras relacionadas com o tema, listadas no Portal de Periódicos da CAPES. E português por ser a língua nativa da pesquisadora.

- Termos utilizados na pesquisa (palavras-chave)

Com a finalidade de melhorar e estruturar a busca nas bibliotecas digitais selecionadas também foi utilizado neste MSL o PICOC (KITCHENHAM e CHARTERS, 2007).

Neste Mapeamento Sistemático, o PICOC foi aplicado da seguinte forma:

- **Population (P):** estágios iniciais do processo de desenvolvimento.
- **Intervention (I):** tecnologias de IHC e ES empregadas em estágios iniciais do processo de desenvolvimento.

- **Comparison (C):** Não se aplica, pois, o objetivo não é fazer uma comparação entre tecnologias, mas caracterizá-las.
- **Outcome (O):** melhorias da aplicação através dos modelos desenvolvidos através do uso de tecnologias que criam/avaliam usabilidade.
- **Context (C):** Não se aplica, pois, como não há comparação, não é necessário determinar um contexto.

Na Tabela 12 são mostrados os termos e a *string* de busca em inglês e na Tabela 13 são mostrados os termos e a *string* de busca em português utilizados na pesquisa. Os termos estão agrupados em três partes: a primeira parte representa a população, ou seja, a(s) fase(s) do processo de desenvolvimento em que a tecnologia será abordada; a segunda representa a intervenção: o que se planeja encontrar; e a terceira representa os resultados: o que se deseja melhorar, avaliar ou projetar.

**Tabela 12. Termos e *String* de busca em inglês utilizados no 2º MSL**

<b><i>String</i> de busca em inglês</b>		
População	("design phase" OR "analysis phase" OR "early stage*" OR "early phase*" OR "early usability")	AND
Intervenção	("tool" OR "framework" OR "technique" OR "method" OR "model" OR "process" OR "guideline" OR "pattern" OR "metric" OR "approach*" OR "inspection" OR "principle" OR "aspect" OR "requirement" OR "feature elicitation" OR "feature specification" OR "feature requirements" OR "scenario" OR "engineering perspective" OR "oriented engineering process" OR "requirements engineering" OR "heuristic" OR "functionalit*")	AND
Resultados	("usability design" OR "usability evaluation" OR "usability assessment" OR "usability improvement" OR "usability assurance")	

**Tabela 13. Termos e *String* de busca em português utilizados no 2º MSL**

<b><i>String</i> de busca em português</b>		
População	("fase de projeto" OR "fase de análise" OR "estágio* inicia*" OR "fase* inicia*" OR "usabilidade cedo")	AND
Intervenção	("ferramenta" OR "framework" OR "técnica" OR "método" OR "modelo" OR "processo" OR "diretrizes" OR "padrão" OR "métrica" OR "abordagem" OR "inspeção" OR "princípio" OR "aspectos" OR "requisitos" OR "elicitação de características" OR "especificações de características" OR "requisitos de características" OR "cenário" OR "perspectiva de engenharia" OR "processo de engenharia" OR "engenharia de requisitos" OR "heurísticas" OR "funcionalidades")	AND
Resultados	("projeto de usabilidade" OR "avaliação de usabilidade" OR "melhoria de usabilidade" OR "garantia de usabilidade")	

#### 4.2.4. Critérios de seleção de artigos e Procedimentos

Cada artigo identificado neste MSL foi avaliado independentemente por dois pesquisadores que decidiram se este deveria ou não ser incluído considerando seu título e resumo. As discrepâncias na seleção foram resolvidas por consenso entre os dois pesquisadores após a análise de todo o artigo. Antes do processo de seleção, os pesquisadores discutiram e obtiveram um entendimento consistente dos critérios de inclusão e exclusão, listados a seguir.

- Critérios para a inclusão e exclusão de artigos

Como critérios de Inclusão, têm-se:

- **CI1.** Publicações que apresentam tecnologias que projetam e/ou avaliam a usabilidade nos estágios iniciais do processo de desenvolvimento de *software*;
- **CI2.** Publicações onde são apresentadas ferramentas que apoiam tecnologias que projetam e/ou avaliam a usabilidade nos estágios iniciais;
- **CI3.** Publicações onde são descritos estudos experimentais de tecnologias que projetam e/ou avaliam a usabilidade nos estágios iniciais;
- **CI4.** Publicações que discutam aspectos relacionados a tecnologias que projetam e/ou avaliam a usabilidade nos estágios iniciais.

Como critérios de Exclusão, têm-se:

- **CE1.** Não foram selecionadas publicações que não atendam aos critérios acima;
- **CE2.** Não foram selecionadas publicações que não têm disponibilidade de conteúdo para leitura e análise dos dados (especialmente em casos, onde os estudos são pagos ou não disponibilizados pelas máquinas de buscas).
- **CE3.** Não foram selecionadas publicações que possuem linguagem diferente de Inglês e Português.

- Processo de seleção

Os resultados do processo de seleção foram discutidos entre dois pesquisadores e todos os conflitos entre os resultados foram resolvidos, para mitigar viés pessoal. Um terceiro pesquisador foi consultado em reuniões específicas que foram realizadas para resolver desacordos ou incerteza durante a seleção dos artigos.

- Processo de seleção preliminar (1º filtro): Dois pesquisadores avaliaram o título e o resumo de cada artigo de acordo com os critérios de inclusão e exclusão e os artigos

selecionados que estariam dentro do escopo da questão de pesquisa. Os pesquisadores mantinham qualquer artigo para a próxima etapa do processo de seleção se não conseguiam decidir se incluí-lo ou descartá-lo com base apenas em seu título e resumo. A confiabilidade da inclusão de um artigo candidato neste estudo de mapeamento sistemático foi avaliada pela aplicação do Fleiss' Kappa (FLEISS, 1981). Kappa é uma medida estatística para avaliar a confiabilidade de concordância entre um número fixo de avaliadores na classificação de itens. Esta medida é pontuada como: não existe concordância ( $<0$ ), concordância mínima (0 – 0,20), concordância razoável (0,21 – 0,40), concordância moderada (0,41 – 0,60), concordância substancial (0,61 – 0,80) e concordância quase perfeita (0,81 - 1). Os dois pesquisadores envolvidos no mapeamento classificaram de forma independente uma amostra aleatória de 255 artigos no primeiro filtro. O Kappa obtido foi 0,614. Ele indica um nível substancial de concordância entre avaliadores (FLEISS, 1981).

- o Processo de seleção final (2º filtro): Como a estratégia de leitura de somente duas informações (título e abstract) não é suficiente para identificar se o estudo é realmente relevante, torna-se necessário realizar a leitura completa dos artigos que restaram do 1º filtro. Dessa forma, no segundo filtro, os pesquisadores realizaram uma leitura completa dos artigos selecionados do primeiro filtro. Os pesquisadores usaram os critérios de seleção para julgar se os artigos devem ser finalmente incluídos ou não. Os resultados foram verificados, e qualquer desacordo entre os pesquisadores foi discutido e resolvido.

#### **4.2.5. Definição de estratégia de extração de dados**

A estratégia de extração de dados empregada neste mapeamento foi baseada em fornecer o conjunto de possíveis respostas para cada subquestão de pesquisa definida anteriormente. Esta estratégia assegura a aplicação dos mesmos critérios de extração de dados para todos os artigos selecionados e facilita a sua classificação. Os dados extraídos foram registrados em um documento para posterior análise e síntese. As possíveis respostas a cada subquestão de pesquisa são explicadas em mais detalhes a seguir.

Sobre a SQ1 (Estágios do processo de desenvolvimento), a tecnologia pode ser usada em um dos seguintes estágios do processo de desenvolvimento:

- a. Análise: a tecnologia é usada para avaliar ou projetar modelos que são especificações de alto nível do *software* (por exemplo, modelos de tarefas).
- b. Projeto: a tecnologia é usada para avaliar ou projetar modelos criados em tempo de projeto e antes da codificação (por exemplo, modelos de navegação).
- c. Ambos: se a tecnologia é utilizada em modelos de análise e de projeto.

Sobre a SQ2 (Área de Pesquisa), a tecnologia pode ser classificada em uma das seguintes áreas de pesquisa:

- a. IHC: se a tecnologia de usabilidade é do campo de IHC.
- b. ES: se a tecnologia de usabilidade é do campo de ES.
- c. Ambas: se a tecnologia de usabilidade integra interesses das áreas de IHC e ES.

Na SQ3 (Tipo de tecnologia), a tecnologia pode ser categorizada nos seguintes tipos:

- a. Avaliação de Usabilidade: se a tecnologia avalia um modelo do estágio inicial do processo de desenvolvimento visando à usabilidade da aplicação final.
- b. Projeto de Usabilidade: se a tecnologia projeta um modelo do estágio inicial visando à usabilidade da aplicação final.

Dentro da SQ3, há uma subquestão chamada SQ3.1 (Tipo da tecnologia de avaliação de usabilidade). Sobre a SQ3.1, as tecnologias de avaliação da usabilidade foram classificadas com base na taxonomia proposta por Ivory e Hearst (2001). A tecnologia pode ser categorizada em um ou mais dos seguintes tipos:

- a. Teste: se a tecnologia necessita de um avaliador observando os participantes interagindo com um modelo para determinar problemas de usabilidade (por exemplo, protocolo de *Think Aloud*, teste de usuário, método *Wizard of Oz*).
- b. Inspeção: se envolve um avaliador especialista usando um conjunto de critérios para identificar possíveis problemas de usabilidade representados em modelos (por exemplo, Método Inspector, *walkthroughs*, checklist).
- c. Investigação: se a tecnologia necessita de um método de coleta de dados pessoais dos participantes, tais como seus sentimentos ou suas preferências (por exemplo, Focus Group, questionários ou entrevistas).
- d. Modelagem analítica: se a tecnologia envolve uma abordagem de engenharia que permita aos avaliadores prever a usabilidade empregando diferentes tipos de modelos (por exemplo, *Usability-Supporting Architectural Patterns* - USAPs).

- e. Simulação: se a tecnologia permite simular a interação do usuário através de qualquer algoritmo de simulação ou a análise de dados de uso (por exemplo, modelos de rede de Petri ou Algoritmo de Árvore de Tarefa).

Sobre a SQ4 (Modelos em que a tecnologia é aplicada), seu objetivo é identificar o(s) modelo(s) em que a tecnologia pode ser empregada. Por exemplo, protótipos, especificações de casos de uso, diagramas de atividades, cenários, storyboards, entre outros.

O objetivo da SQ5 (Tipo de Contribuição) é identificar a principal contribuição do artigo. O tipo de contribuição refere-se à determinação do tipo de intervenção sendo estudada (PETERSEN *et al.*, 2015), que pode ser uma ferramenta, processo, métricas, entre outros.

Sobre a SQ6 (Apoio ferramental), a tecnologia pode ser classificada em uma das seguintes respostas:

- a. Sim: a tecnologia requer algum suporte de ferramenta específico.
- b. Não: a tecnologia não requer suporte de ferramentas específico.

Sobre a SQ7 (Estudos empíricos), a tecnologia pode ser categorizada em uma das seguintes respostas:

- a. Sim: existe uma avaliação empírica da tecnologia proposta descrita no artigo.
- b. Não: não há avaliação empírica da tecnologia proposta.

Dentro da SQ7, existem subquestões chamadas SQ7.1 (Tipos de estudos empíricos), SQ7.2 (Ambiente de avaliação) e SQ7.3 (Tipos de análise). Sobre a SQ7.1, os estudos foram classificados com base na taxonomia proposta por Shull *et al.* (2001). Além disso, alguns estudos foram classificados como *surveys* (Wöhlin *et al.*, 2000) se apenas coletam informações sobre a tecnologia. Provas de conceitos e exemplos ilustrativos não foram consideradas nesta subquestão, uma vez que não apresentam evidência empírica. Para SQ7.1, um estudo empírico avaliando a tecnologia pode ser categorizado em um ou mais dos seguintes tipos:

- a. Estudo de Viabilidade: estudo para determinar a possibilidade de uso da tecnologia, comparando-a com outra tecnologia;
- b. Estudo Observacional: estudo para melhorar a compreensão ou a relação custo-eficácia da tecnologia;
- c. Estudo de caso: estudo para caracterizar a aplicação da tecnologia durante um ciclo de vida real ou para identificar se a aplicação da tecnologia se encaixa em um contexto industrial foi realizado, ou;

- d. *Survey*: estudo para coletar informações de pessoas para descrever, comparar ou explicar seus conhecimentos, atitudes e comportamento (Wöhlin *et al.*, 2000).

Em SQ7.2, o ambiente onde a tecnologia foi avaliada pode ser categorizado em um dos seguintes tipos:

- a. Ambiente industrial com profissionais: se a tecnologia foi avaliada em um ambiente industrial com profissionais;
- b. Ambiente acadêmico com alunos: se a tecnologia foi utilizada ou avaliada em um ambiente acadêmico com os alunos;
- c. Ambiente de laboratório com profissionais: se a tecnologia foi utilizada ou avaliada em laboratório com profissionais (não estudantes), ou;
- d. Misto: se a tecnologia foi utilizada ou avaliada em contextos industriais e acadêmicos, contextos industriais e em laboratório ou ambientes acadêmicos e em laboratórios.

Sobre a SQ7.3, uma análise de estudo pode ser classificada em um dos seguintes tipos:

- a. Quantitativa: a análise do estudo da tecnologia foi realizada de forma quantitativa;
- b. Qualitativa: a análise do estudo da tecnologia foi realizada de forma qualitativa;
- c. Ambas: a análise do estudo da tecnologia foi conduzida de forma quantitativa e qualitativa.

Na SQ8 (Contexto de aplicação), a tecnologia pode ser categorizada em uma das seguintes respostas:

- a. Específico: a tecnologia foi utilizada ou avaliada em um contexto específico, isto é, limitada a um tipo específico de aplicação ou processo de desenvolvimento (por exemplo, aplicação móvel, ciclo de vida ágil, entre outros), ou;
- b. Genérico: se a tecnologia foi utilizada ou avaliada em um contexto geral, isto é, não limitado a um tipo específico de aplicação ou processo de desenvolvimento.

#### 4.2.6. Artigos Seleccionados após a Condução do 2º Mapeamento Sistemático

Conforme ilustrado na Figura 5, 1013 artigos foram inicialmente retornados aplicando a *string* de busca antes do primeiro filtro. Um total de 352 artigos foram seleccionados após a aplicação do primeiro filtro, com base nos critérios de inclusão (ver Subsecção 4.2.4). Um total de

136 artigos foram selecionados após a aplicação do segundo filtro. Alguns artigos apareceram em mais de uma biblioteca digital. Neste caso, um artigo repetido foi considerado apenas uma vez de acordo com a nossa ordem de busca realizada (primeiro ACM e, em seguida, Scopus).

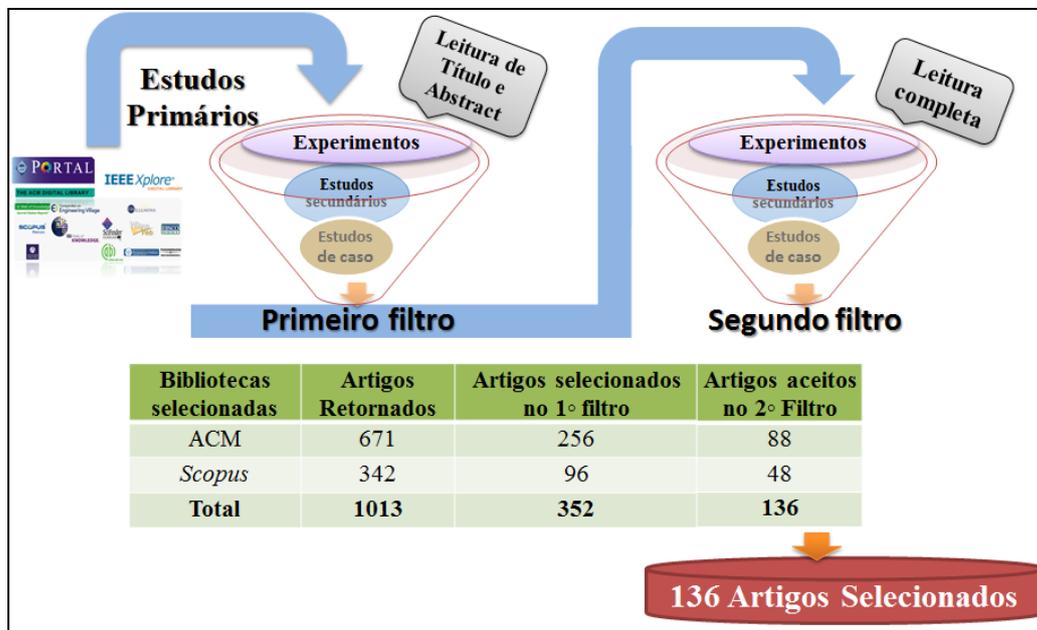


Figura 5. Total de artigos no 1º filtro e 2º filtro no 2º MSL

Os resultados da pesquisa revelaram que os artigos sobre “*Early Usability*” foram publicados em várias conferências/periódicos relacionados a diferentes áreas como ES, IHC e outras áreas relacionadas.

Dos 136 artigos selecionados neste MSL, os artigos de Vermeeren *et al.* (2010), Fernandez *et al.* (2011a), Ivory e Hearst (2001), Salvador *et al.* (2014), Silva *et al.* (2015) e Yen e Bakken (2012) apresentam estudos secundários (revisão da literatura ou mapeamentos sistemáticos ou revisões sistemáticas da literatura). Não foi realizada uma análise desses estudos secundários porque a intenção, neste momento, não é realizar uma revisão terciária. Em trabalhos futuros, pretende-se fazer uma análise destes seis artigos que contêm estudos secundários identificados neste MSL. Portanto, na apresentação dos resultados deste MSL foram considerados apenas 130 artigos.

Todas as informações mostradas neste Capítulo podem ser encontradas em maior detalhe no Relatório Técnico criado por Valentim *et al.* (2017a). Neste relatório são apresentadas as informações relacionadas ao planejamento, execução e análise dos dados encontrados a partir

do mapeamento sistemático executado. Também está disponível a tabela com os dados de extração dos artigos selecionados.

### 4.3. Resultados obtidos

#### 4.3.1. Visão geral dos resultados

Os resultados gerais, que são baseados na contagem das tecnologias que são classificadas em cada uma das respostas às subquestões de pesquisa, são apresentados na Tabela 14.

Tabela 14. Resultados do 2º MSL para cada uma das Subquestões

Subquestão de Pesquisa	Possíveis Respostas	Resultados	
		Tecnologias	Porcentagem (%)
SQ1. Estágio do processo de desenvolvimento	Análise	54	26,60
	Projeto	108	53,20
	Ambas	41	20,20
SQ2. Área de pesquisa	IHC	92	45,32
	ES	29	14,29
	Ambas	82	40,39
SQ3. Tipo de tecnologia	Avaliação de Usabilidade	189	93,10
	Projeto de Usabilidade	14	6,90
SQ3.1 Tipo de tecnologia de avaliação de usabilidade	Teste	67	33,00
	Inspeção	107	52,71
	Investigação	25	12,32
	Modelagem Analítica	23	11,33
	Simulação	8	3,94
SQ6. Apoio ferramental	Sim	42	20,69
	Não	161	79,31
SQ7. Avaliação experimental	Sim	95	46,80
	Não	108	53,20
SQ7.1 Tipo de estudo experimental	Estudo de Viabilidade	32	32,99
	Estudo de Observação	2	2,06
	Estudo de Caso	64	65,98
	Survey	1	1,03
SQ7.2 Ambiente de avaliação	Indústria	40	41,24
	Academia	32	32,99
	Laboratório	24	24,74
	Misto	1	1,03
SQ7.3 Tipo de Análise	Quantitativa	9	9,28
	Qualitativa	45	46,39
	Ambas	43	44,33
SQ8. Contexto de aplicação	Específico	108	53,20
	Genérico	95	46,80
<b>Nota:</b> SQ3.1 e SQ7.1 não são exclusivos, e uma tecnologia pode ser classificada em uma ou mais das possíveis respostas. A soma das porcentagens pode ser superior a 100%.			

Este MSL identificou 203 tecnologias relacionadas à usabilidade de *software*. Pode-se observar que as subquestões SQ4 e SQ5 foram omitidas da Tabela 14 porque elas têm muitas respostas. Além disso, as subquestões SQ3.1 e SQ7.1 não são exclusivas; o que significa que a tecnologia pode ser classificada em uma ou mais das possíveis respostas. Portanto, nestas subquestões, a soma das percentagens pode ser superior a 100%.

#### 4.3.2. Ano de Publicação de artigos

Os artigos selecionados foram publicados entre 1995 e setembro de 2015. Do ponto de vista temporal (Figura 6), houve um aumento do número de publicações nos anos de 2006 e 2007. Percebe-se também, de acordo com as publicações coletadas neste MSL, que nos anos de 2008, 2010, 2011 e 2014, o número de publicações diminuiu. Entre os anos de 1996 e 1998 não houve nenhuma publicação. O ano de 2015 não pôde ser analisado integralmente pois este MSL considerou artigos publicados até setembro de 2015. Isto pode ser um dos motivos para o baixo número de artigos neste ano. O ano de 2007 é o ano com a maior quantidade de publicações (20 artigos), seguido de 2009 (17 artigos) e 2013 (16 artigos).

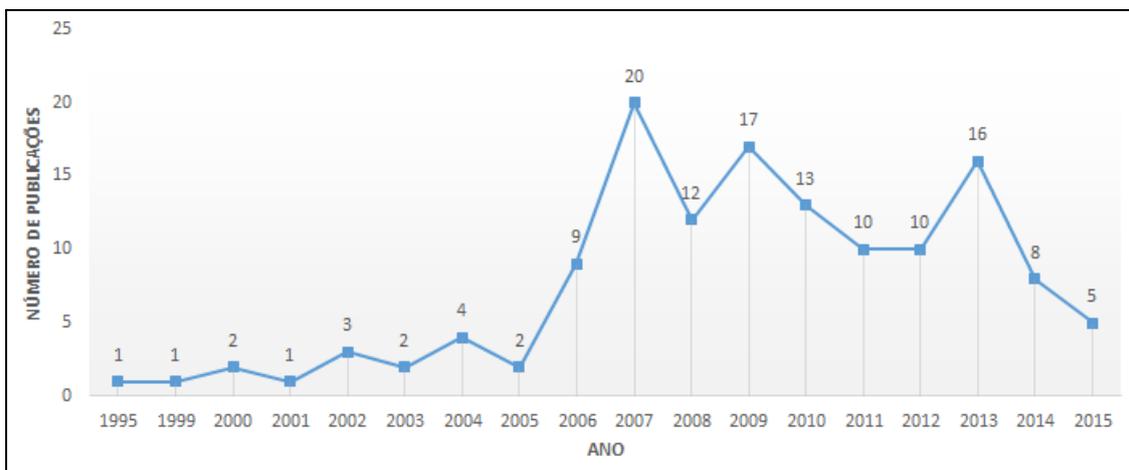


Figura 6. Visão temporal das publicações identificadas no 2º MSL

#### 4.3.3. Locais de Publicação

Neste MSL, foi considerado locais de publicação avaliados por pares (incluindo periódicos, bem como conferências e workshops com revisão por pares). A Figura 7 fornece uma visão geral da distribuição de artigos por periódico. O principal local é o *Journal of Systems and Software* (JSS), com quatro artigos publicados. Em seguida, tem-se o *Software Quality Journal*

(SQJ) com três artigos publicados, seguido por *Information and Software Technology* (IST), *Personal and Ubiquitous Computing* (PUC), e *Audio Engineering Society* (AES), com dois artigos cada.

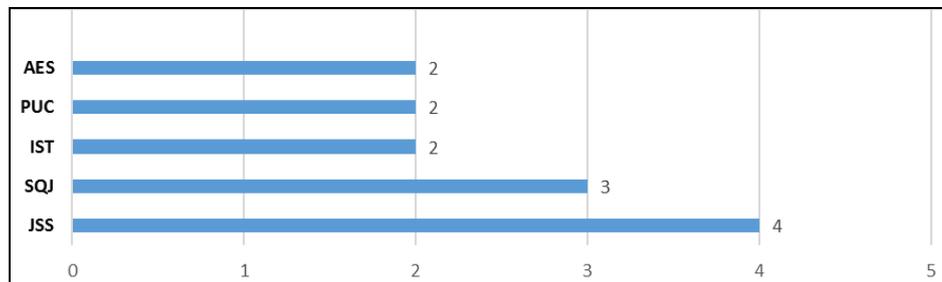


Figura 7. Distribuição de artigos por periódico no 2º MSL

Sobre a distribuição de artigos por conferência (ver Figura 8), o principal local de publicação é o *Nordic Conference on Human-Computer Interaction* (NordiCHI), com seis artigos publicados. *Conference on Human Factors in Computing Systems* (CHI), *International Conference on Human-Computer Interaction* (INTERACT), *International Conference on Human-Computer Interaction* (HCII), e *International Conference on Web Information Systems Engineering* (WISE), possui quatro artigos em cada um. Além disso, há conferências, como *Symposium on Engineering Interactive Computing Systems* (EICS) e *International Symposium on Empirical Software Engineering and Measurement* (ESEM), com três artigos cada.

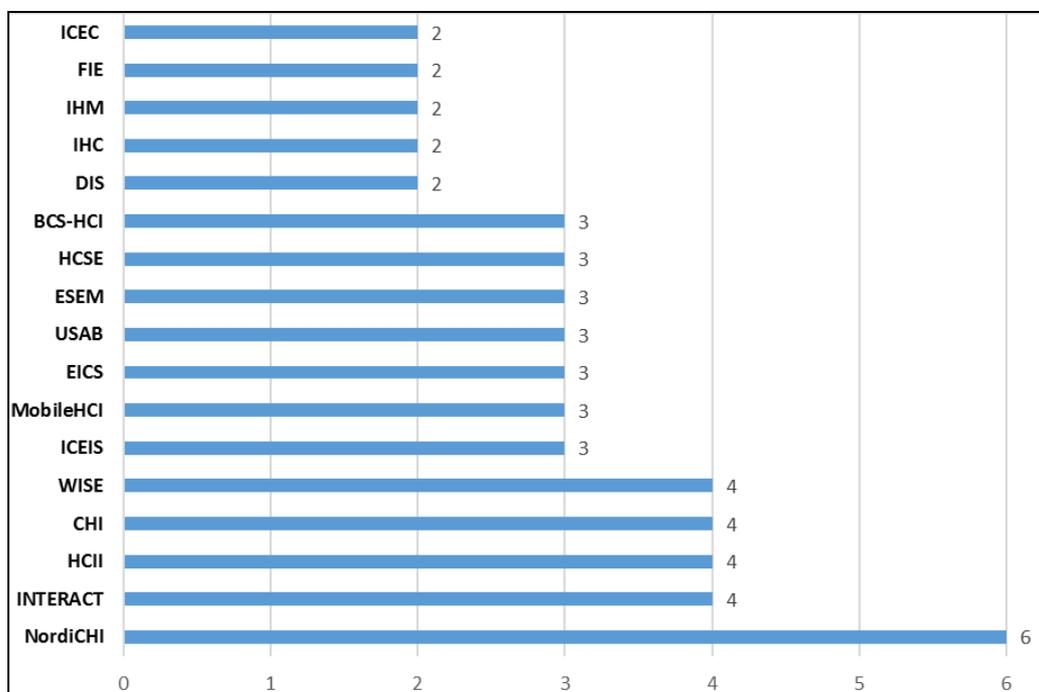


Figura 8. Distribuição de artigos por conferência no 2º MSL

A Figura 9 fornece uma visão geral da alocação de artigos por workshops. O principal workshop é o *International Workshop - Task Models and Diagrams for User Interface Design* (TAMODIA), com três artigos publicados.

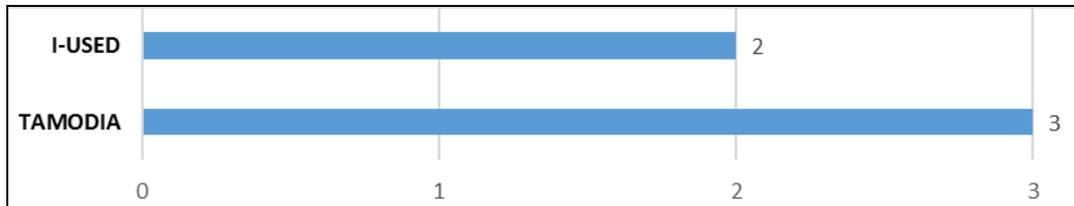


Figura 9. Distribuição de artigos por workshop no 2º MSL

Houve outras conferências, periódicos e workshops com apenas um artigo publicado, mas não foram representadas nas figuras acima.

#### 4.3.4. Estágio do processo de desenvolvimento (SQ1)

Os resultados desta subquestão revelaram que 53,20% das tecnologias podem ser empregadas no estágio de Projeto (ver Tabela 14). Estas tecnologias apoiam a avaliação ou o projeto de artefatos intermediários durante o processo de desenvolvimento (por exemplo, modelos de navegação). Por exemplo, Peischl *et al.* (2013) apresentam uma pesquisa que trata da integração de métodos de engenharia centrados no usuário no desenvolvimento de aplicativos móveis em uma fase inicial do desenvolvimento de protótipos de papel. Nesta tecnologia, questões de usabilidade são consideradas.

Cerca de 26,60% das tecnologias podem ser aplicadas no estágio de Análise. Estas tecnologias apoiam a avaliação ou o projeto de especificações de alto nível do *software* (por exemplo, cenários). Um exemplo de uma dessas tecnologias é o processo de Integração de Requisitos e Segurança de Informação (*Integrating Requirements and Information Security - IRIS*) para o desenvolvimento e aplicação de personas para apoiar a elicitação e especificação de requisitos para sistemas seguros e usáveis, proposto por Faily e Fléchais (2010). IRIS é dividido em três fases: (1) um workshop onde o escopo do sistema é definido, (2) observação e entrevista com usuários para identificar os contextos de uso e (3) workshop de requisitos e análise de risco.

Por fim, 20,20% das tecnologias podem ser utilizadas durante os estágios de análise e projeto. Por exemplo, Velmourougan *et al.* (2014) apresentam uma nova tecnologia chamada Ciclo de Vida de Desenvolvimento de *Software*-Usabilidade (*Usability-Software Development Life Cycle - U-SDLC*). U-SDLC introduz tarefas de desenvolvimento de usabilidade e atividades a

serem seguidas durante o ciclo de vida de desenvolvimento de *software*. Especificamente, na fase de análise, existem três atividades: (a) entrevista contextual, (b) documentação de requisitos de usabilidade e (c) documentação de checklist de usabilidade. A fase de projeto é composta de (e) desenvolvimento de protótipo, (f) avaliação de protótipo, (g) análise de impacto, (h) documentação de diretrizes de usabilidade e a tarefa e (i) verificação de requisitos de usabilidade.

Os resultados desta subquestão indicam que há necessidade de desenvolver mais tecnologias que possam ser empregadas na fase de análise do processo de desenvolvimento. De acordo com Fernandez *et al.* (2011a), se problemas de usabilidade forem identificados cedo, a qualidade das aplicações finais pode ser melhorada antes da codificação, economizando recursos. Além disso, é importante propor novas tecnologias que possam ser utilizadas em ambas as fases. Ao fazer isso, as tecnologias ajudariam a antecipar mais a usabilidade.

#### 4.3.5. Área de Pesquisa (SQ2)

Nesta subquestão, classificou-se uma tecnologia da área de ES ou IHC se os autores citam isso claramente no artigo. Se os autores não citam a área de pesquisa no artigo, a tecnologia foi classificada de acordo com o local da publicação. Os resultados desta subquestão mostraram que 45,32% das tecnologias são propostas na área de IHC (ver Tabela 14). Por exemplo, Bacim *et al.* (2010) propõe uma adição ao processo de engenharia de usabilidade ao projetar um sistema de aprendizagem: utilizar as informações fornecidas pelos experts de domínio e mapeá-los aos processos de percepção e cognição após a análise de requisitos e durante a fase de projeto de informação.

Cerca de 14,29% das tecnologias são propostas na área de ES. A comunidade de ES tem trabalhado vários anos no paradigma do Desenvolvimento Orientado a Modelos (*Model-Driven Development* - MDD), que afirma que todo o esforço dos analistas deve ser focado em um modelo conceitual e os meios do modelo devem implementar o sistema para transformação do código. Um exemplo de uma dessas tecnologias é o Método para incorporar Características de Usabilidade Funcional em Desenvolvimento dirigido a Modelos (*Method to Incorporate Functional Usability Features into Model-Driven Development* - MIFUM) proposto por Panach *et al.* (2013). MIFUM é dividido em duas etapas: (1) identificar as propriedades a serem modeladas e (2) propor mudanças ao método MDD para incluir essas propriedades. O primeiro estágio é composto de dois passos: (1) definição de modos de uso e (2) a identificação de propriedades.

Um especialista em usabilidade, que sabe como trabalhar com diretrizes de usabilidade, executa essas etapas. A segunda etapa é composta de duas etapas: (1) definição das mudanças no modelo conceitual e (2) descrição das variações no compilador do modelo. O designer MDD, que deve aprimorar o método MDD para oferecer suporte a recursos de usabilidade, executa essas etapas.

Cerca de 40,39% das tecnologias são propostas em ambas as áreas. Uma tecnologia foi classificada em ambas as áreas se os autores citam claramente no artigo que a tecnologia integra as áreas IHC e ES. Por exemplo, Karppinen e Liinasuo (2008) descrevem como a usabilidade das funcionalidades de *software* é estimulada e avaliada durante a fase de projeto de um projeto de *software* desenvolvendo funcionalidades relacionadas à segurança em middleware. Neste artigo, os autores desta tecnologia estavam preocupados com a compreensão de termos que têm significados diferentes em IHC e ES. Portanto, algumas tecnologias têm sido propostas para a inclusão da usabilidade no processo de desenvolvimento, envolvendo ambas as áreas.

Os resultados desta subquestão mostram que a maioria das tecnologias encontradas neste MSL estão sendo propostas na área de IHC. Isto pode ter acontecido porque essa área surgiu na década de 1980 e introduziu a capacidade técnica da computação pessoal com conceitos de usabilidade e design centrado no usuário (ver os trabalhos seminais de Shackel (1991), Nielsen (1989), e Norman e Draper (1986)). No geral, a usabilidade é um conceito essencial em IHC (ISSA E ISAIAS, 2015). No entanto, de acordo com Boehm (2006), a ênfase na usabilidade de produtos de *software* por não-programadores só aumentou dentro da área de ES na década de 1990. Conseqüentemente, era necessário reinterpretar a Regra de Ouro, "Faça aos outros o que os outros fariam com você". Para programadores e estudantes de Ciência da Computação, isso significava desenvolver interfaces de usuário amigáveis (BOEHM, 2006). Além disso, muitas tecnologias estão sendo propostas para ambas as áreas. Nos últimos vinte anos foram feitas tentativas crescentes para preencher a lacuna entre as comunidades ES e IHC (JURISTO *et al.*, 2007). Seffah *et al.* (2011) sugerem maneiras pelas quais os engenheiros de *software* e engenheiros de usabilidade podem aprender uns com os outros para facilitar e encorajar a convergência de práticas em ambas as comunidades. De acordo com Juristo *et al.* (2007), é necessário garantir que os problemas de usabilidade sejam adequadamente tratados ao longo do ciclo de desenvolvimento de um produto de *software*.

#### 4.3.6. Tipo de Tecnologia (SQ3)

Os resultados desta subquestão mostraram que 93,10% das tecnologias avaliam um modelo desde os primeiros estágios do processo de desenvolvimento visando à usabilidade da aplicação final (ver Tabela 14). Por exemplo, Fernandez *et al.* (2011b) propõem o Processo de Avaliação de Usabilidade *Web* (*Web Usability Evaluation Process* - WUEP) como um método de inspeção de usabilidade que integra avaliações de usabilidade durante várias etapas do processo de Desenvolvimento *Web* dirigido a Modelos (*Model-Driven Web Development* - MDWD). O WUEP oferece amplo suporte ao conceito de usabilidade, já que seu Modelo de Usabilidade *Web* (*Web Usability Model* - WUM) subjacente foi estendido e adaptado ao domínio da *Web*, considerando a nova série de padrões ISO 25010 (SQuaRE), juntamente com várias diretrizes de usabilidade. As métricas deste modelo fornecem uma definição genérica, que deve ser operacionalizada para se aplicar a artefatos de diferentes níveis de abstração (PIMs, PSMs e Modelos de Código) em vários processos MDWD (por exemplo, OO-H, WebML).

Cerca de 6,90% das tecnologias projetam um modelo das primeiras etapas do processo de desenvolvimento visando à usabilidade da aplicação final. Por exemplo, Panach *et al.* (2015) propõem incluir Características de Usabilidade Funcional (*Functional Usability Features* - FUFs) em um modelo conceitual. FUFs fornecem uma descrição detalhada de como características de usabilidade afetam a arquitetura do sistema (JURISTO *et al.*, 2007). A proposta de Panach *et al.* (2015) é um passo adiante para incorporar características de sistemas de *software* que não são combinadas em métodos MDD. A abordagem para a incorporação de FUFs em um método MDD é dividida em quatro etapas: (1) identificar os possíveis modos de uso de cada funcionalidade de usabilidade, (2) determinar as propriedades que configuram cada modo de uso em relação aos requisitos de usabilidade, (3) definir primitivas conceituais para representar abstratamente a maneira de usar as propriedades e (4) descrever as mudanças que devem ser feitas ao compilador do modelo para implementar as propriedades identificadas. Essa proposta nos aproxima dos modelos conceituais onde os modelos representam não apenas funcionalidade, comportamento ou persistência, mas também características de usabilidade.

Os resultados desta subquestão indicam que há poucas tecnologias que projetam modelos desde as primeiras fases do processo de desenvolvimento visando à usabilidade da aplicação final. Com base nos seus benefícios (JURISTO *et al.*, 2007), há uma necessidade de novas tecnologias neste contexto, porque é uma forma proativa para apoiar o projeto de

aplicação visando à sua usabilidade. Uma das vantagens deste tipo de tecnologia é o tempo reduzido e esforço, uma vez que o *stakeholder* antecipa a usabilidade ao construir o modelo, ou seja, a equipe de desenvolvimento não tem que esperar até que o modelo seja construído para se preocupar com essas questões. Este tipo de tecnologia também contribui para a redução de custos no processo de desenvolvimento de *software*.

Sobre a subquestão SQ3, há uma subquestão chamada SQ3.1 (Tipo de tecnologia de avaliação de usabilidade). Classificamos as tecnologias para a avaliação da usabilidade com base na taxonomia proposta por Ivory e Hearst (2011). Os resultados do SQ3.1 mostraram que o tipo mais usado de tecnologia de avaliação de usabilidade é a inspeção de usabilidade, cerca de 52,71% das tecnologias. Tecnologias de inspeção de usabilidade precisam de um avaliador especialista usando critérios para identificar potenciais problemas de usabilidade em modelos. Por exemplo, Pinelle e Gutwin (2002), e Rivero e Conte (2013). Cerca de 33% das tecnologias são testes de usabilidade, como Mendes e Furtado (2014) e De Sá e Churchill (2012). Os testes de usabilidade precisam de um avaliador observando os participantes interagindo com um modelo para identificar problemas de usabilidade. Os métodos de investigação representam cerca de 12,32% das tecnologias, como Ryu *et al.* (2006), que coleta informações pessoais dos participantes. A modelagem analítica corresponde a cerca de 11,33% das tecnologias, como Molina e Toval (2009) e John *et al.* (2009), onde a abordagem permite aos avaliadores prever a usabilidade empregando diferentes tipos de modelos. Os métodos de simulação representam apenas cerca de 3,94% das tecnologias, como o trabalho de Palanque *et al.* (2007), onde a interação do usuário é simulada.

Os resultados da subquestão SQ3.1 mostram que a Inspeção de Usabilidade foi o tipo mais usado de tecnologia de avaliação de usabilidade. As inspeções são executadas por avaliadores ou projetistas especializados (não requerem a participação de usuários finais) e baseiam-se na revisão dos aspectos de usabilidade dentro dos artefatos avaliados, sobre sua conformidade com um conjunto de diretrizes (FERNANDEZ *et al.*, 2011a). Poucas tecnologias foram de Simulação. Isto pode ter acontecido porque essas tecnologias apresentam características de outros tipos de tecnologias e geralmente requerem suporte de ferramentas.

#### 4.3.7. Modelos em que as tecnologias são aplicadas (SQ4)

A Figura 10 fornece uma visão geral dos modelos nos quais as tecnologias são aplicadas.

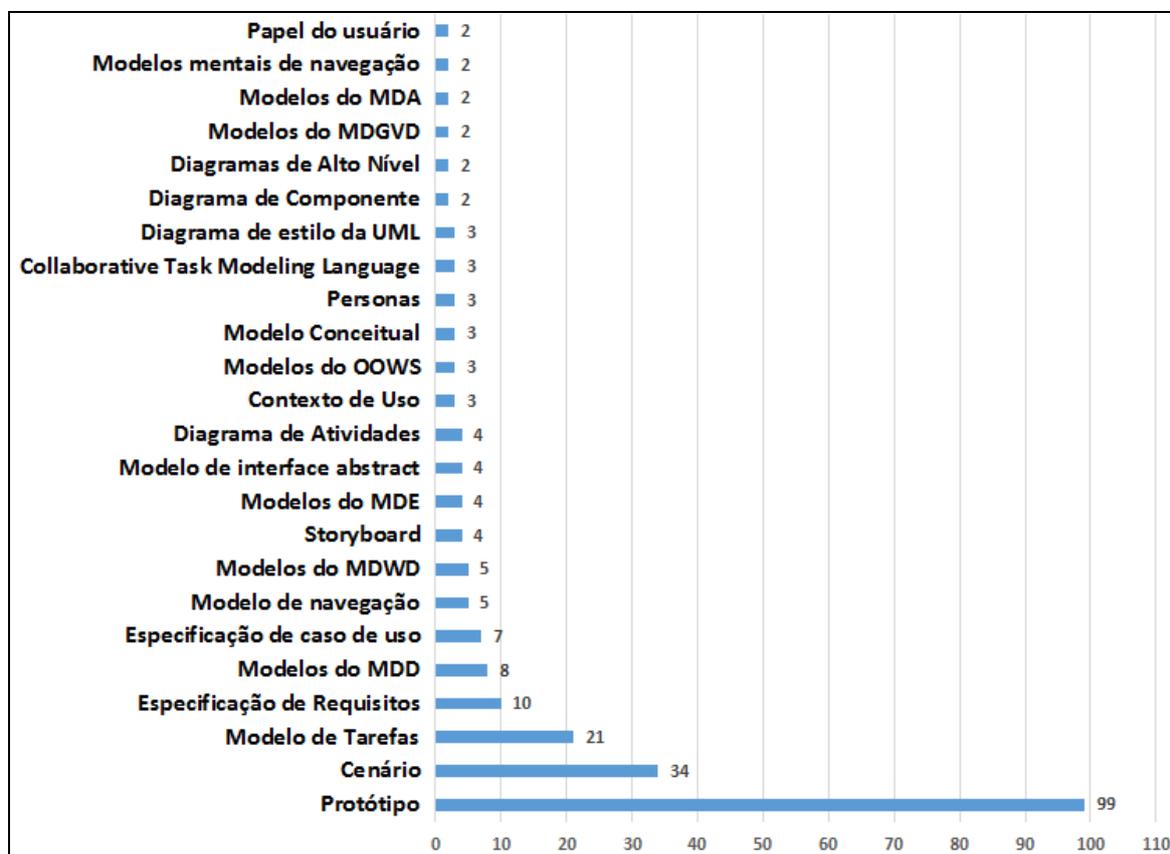


Figura 10. Modelos em que as tecnologias são aplicadas - 2º MSL

Os resultados desta subquestão mostraram que 99 tecnologias utilizam um protótipo de baixa fidelidade objetivando a usabilidade da aplicação final. Nesta tese, considera-se protótipos como esboços que são construídos principalmente para projetar layouts. Eles podem ser criados em papel ou com *software* específico. Considera-se *wireframes* e *mockups* como protótipos. *Wireframes* são esquemas que exibem elementos básicos sem tantos detalhes. *Mockups* são exibições estáticas de como o visual do produto final deve parecer, oferecendo a oportunidade de tomar as decisões necessárias sobre esquemas de cores, tipografia e estilo. Ambos representam elementos centrais da aplicação e apresentam um comportamento estático da aplicação. Um exemplo de uma tecnologia que usa protótipo é apresentado por Paelke e Nebe (2008). Os autores propuseram a realização de iterações iniciais de projeto na interface, seu comportamento e componentes usando artefatos ainda mais leves nas iterações iniciais, ou seja,

esboços de projeto, protótipos de papel e *mockups*. Esses modelos podem ser empregados tanto para revisar o projeto quanto para realizar avaliações de usabilidade.

O cenário é o segundo modelo mais utilizado (com 34 tecnologias identificadas). O uso de cenários é uma abordagem estabelecida para pensar sobre os usuários dentro do projeto. Rafla *et al.* (2007) propôs o método Workshop Atributo de Qualidade (*Quality Attribute Workshop - QAW*) que envolve *stakeholders* no início do ciclo de vida do desenvolvimento do sistema na descoberta dos atributos de qualidade das arquiteturas de *software*. O processo QAW consiste nos seguintes sete passos e em seu último passo, os casos de uso definidos são analisados para compreender as razões por trás dos requisitos de usabilidade e representam uma compreensão de alto nível da interface do usuário. Ter que fornecer pré e pós-condições para cada caso de uso força os analistas de usabilidade a pensar mais logicamente sobre os requisitos de usabilidade e os objetivos do usuário.

Os resultados desta subquestão indicam que muitas tecnologias estão sendo desenvolvidas para melhorar a usabilidade através de protótipos. Isso mostra que os pesquisadores investiram nesse tipo de modelo, talvez porque ele demonstra mais características visuais da aplicação final. Outra razão é que as empresas de *software* desejam desenvolver seu *software* mais rapidamente. Para fazer isso, eles querem um feedback rápido de seus clientes e usuários através de um formato visual de seu projeto do sistema.

#### 4.3.8. Tipo de Contribuição (SQ5)

A Figura 11 apresenta os resultados da subquestão SQ5. Pode-se notar que cinquenta e quatro das tecnologias são técnicas (tipo de contribuição). Um exemplo de uma técnica é apresentado por Pinelle e Gutwin (2002). A técnica é chamada de Groupware Walkthrough (GW), uma técnica de inspeção de usabilidade para groupware. A técnica é uma modificação do walkthrough cognitivo para incluir a consideração das complexidades do trabalho em equipe. O GW tem dois componentes: um modelo de tarefa para identificar e analisar tarefas colaborativas do mundo real e um processo para avaliar o apoio de um sistema para essas tarefas.

Cinquenta tecnologias abordam métodos. Por exemplo, Wyk e De Villiers (2008) propuseram a Análise de Contexto de Usabilidade (Usability Context Analysis - UCA), que é um método estruturado para obter informações detalhadas sobre um produto e como ele será usado, e para derivar um plano para uma avaliação baseada em usuário de um produto.

Os resultados desta subquestão indicam que os maiores tipos de contribuição identificados nos artigos dentro deste MSL são técnicas e métodos.

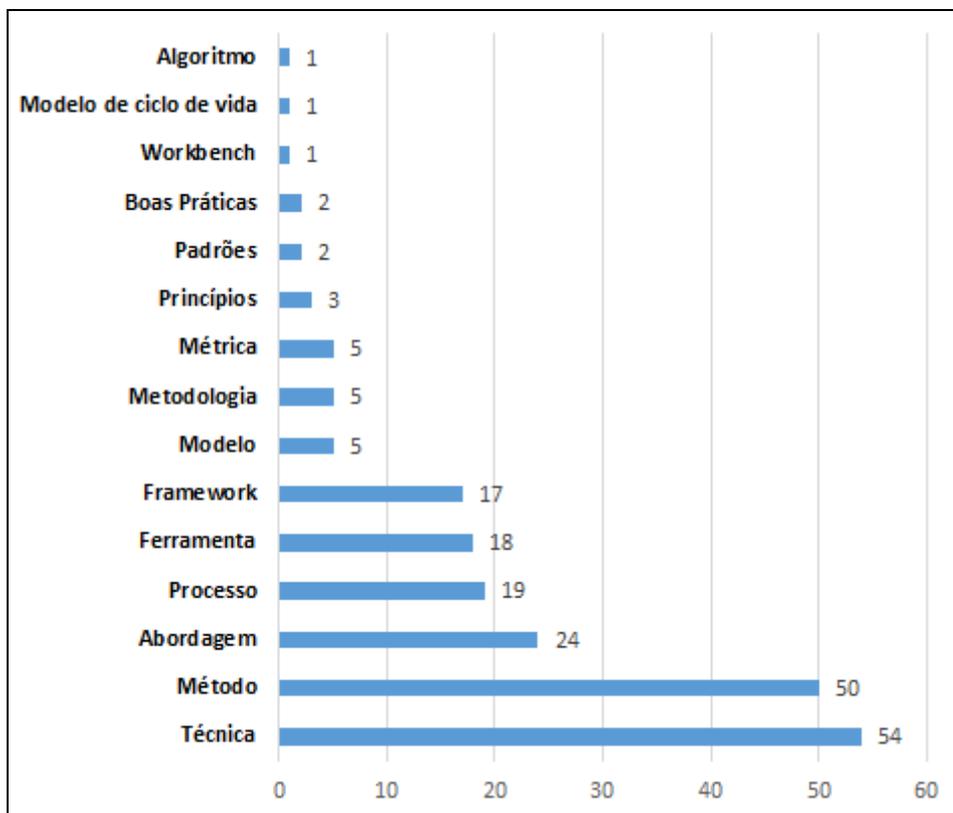


Figura 11. Tipo de contribuição dos artigos do 2º MSL

#### 4.3.9. Apoio Ferramental (SQ6)

Os resultados desta subquestão indicaram que 79,31% das tecnologias não necessitam de suporte de ferramentas para auxiliar os profissionais (ver Tabela 14). Alguns exemplos são apresentados por Siek *et al.* (2011) e Blandford *et al.* (2007). No entanto, cerca de 20,69% das tecnologias requerem apoio de ferramentas. Por exemplo, Memmel e Reiterer (2008) apresentam um método mais um apoio ferramental empírico para gerar especificações visuais de interface do usuário, conhecido como Inspector (*Interdisciplinary UI Specification Tool*). Um dos requisitos do Inspector é permitir uma avaliação de usabilidade da aparência e sensação da interface. Nos estágios iniciais, por exemplo, um template de usuário pode ser ligado e ser parte de papéis de usuário, personas e diagramas de casos de uso.

É importante observar que a avaliação de usabilidade automatizada tem algumas deficiências. Algumas delas não são orientadas a coleta de dados subjetivos. Portanto, as

percepções dos usuários e o contexto do usuário não podem ser considerados. No entanto, o emprego de tecnologias automatizadas pode reduzir os esforços dos profissionais e o consumo de recursos, proporcionando muitos benefícios para a indústria.

#### 4.3.10. Estudos Experimentais (SQ7)

Os resultados desta subquestão revelaram que 53,20% das tecnologias não foram avaliadas experimentalmente. As tecnologias restantes (46,80%) foram avaliadas experimentalmente. Os resultados desta subquestão mostram que menos da metade das tecnologias foram avaliadas experimentalmente. No entanto, um número significativo de tecnologias não foi avaliado. Isto demonstra que há necessidade de mais estudos experimentais no campo da “*Early Usability*”.

Dentro do SQ7, também foram definidas algumas subquestões: SQ7.1 (Tipos de estudos experimentais), SQ7.2 (Ambiente de avaliação) e SQ7.3 (Tipos de análise). Sobre a SQ7.1, os estudos foram classificados com base na taxonomia proposta por Shull *et al.* (2001). Além disso, alguns estudos foram classificados como *surveys* (Wöhlin *et al.*, 2000). *Surveys* coletam informações sobre as tecnologias.

Os resultados para SQ7.1 revelaram que o tipo de estudo experimental mais empregado foi o Estudo de Caso. Cerca de 65,98% das tecnologias foram avaliadas por Estudo de Caso, como por exemplo De Sá e Carriço (2008) e Riihaho (2002). Cerca de 32,99% das tecnologias foram avaliadas por Estudos de Viabilidade, como Wania e Atwood (2009), Hornbæk *et al.* (2007). Além disso, cerca de 2,06% das tecnologias foram avaliadas por Estudos de Observação, como os descritos por Golden (2009) e Olsen *et al.* (2011). Somente Gordillo *et al.* (2014) apresentou um *Survey* em seu artigo.

O estudo de caso foi o tipo de estudo experimental mais empregado. Contudo, foram realizados poucos estudos de observação e *surveys*. Isto demonstra que há necessidade de mais estudos de observação e *surveys* nas áreas IHC e ES. De acordo com Shull *et al.* (2001), um estudo de observação pode ser um pouco mais demorado para o pesquisador e menos confortável para os participantes do que questionários ou entrevistas. No entanto, oferece resultados qualitativos mais precisos. *Surveys* exigem um investimento de tempo relativamente pequeno do pesquisador, mas às vezes ele pode não coletar as informações no nível de detalhes desejado (SHULL *et al.*, 2001).

Quanto à SQ7.2 (Ambiente de avaliação), cerca de 32,99% das tecnologias apresentadas foram utilizadas em ambiente acadêmico com alunos, como Swierenga *et al.* (2011) e Tohidi *et al.* (2006a). 41,24% das tecnologias foram aplicadas em ambiente industrial com profissionais, como Span *et al.* (2014) e Haynes (2009). Cerca de 24,74% das tecnologias foram aplicadas em ambiente laboratorial com profissionais, como Or e Tao (2012) e Panach *et al.* (2008). Apenas Chen *et al.* (2007) apresentou uma tecnologia que foi aplicada em um ambiente misto, onde o projeto colaborativo tinha oito parceiros da academia e da indústria. Os resultados da SQ7.2 mostram que mais de metade das tecnologias encontradas neste MSL não foram avaliadas em ambiente industrial. Isto pode ter acontecido devido aos custos altos para as empresas realizar avaliações com profissionais. Uma vez que alguns alunos trabalham na indústria e têm o perfil profissional esperado para participar das avaliações, eles têm sido considerados nos estudos.

Quanto à SQ7.3 (Tipos de análise), 46,39% das tecnologias apresentam a análise do estudo de forma qualitativa, como a apresentada por Paelke e Nebe (2008) e Stoll *et al.* (2009). Cerca de 9,28% das tecnologias apresentam a análise dos resultados do estudo de forma quantitativa, como a de Ben Ammar *et al.* (2013) e Panach *et al.* (2008). Cerca de 44,33% das tecnologias apresentam a análise do estudo de forma quantitativa e qualitativa, como Følstad e Knutsen (2010) e Tohidi *et al.* (2006b). Os resultados da SQ7.3 mostram que a maioria dos estudos experimentais encontrados neste MSL estão sendo analisados tanto de forma quantitativa quanto qualitativa. Uma análise quantitativa enfatiza a coleção de números. Neste tipo de análise, os pesquisadores testam hipóteses através de estatísticas e tendências de dados para procurar padrões que permitam a previsão da ocorrência e comportamento dos fenômenos. Por outro lado, a análise qualitativa compreende atividades cognitivas intensas, tais como interpretação, associação e correlação. Há uma subjetividade inerente nesse tipo de análise. Ambas as análises são relevantes para evoluir uma tecnologia.

#### **4.3.11. Contexto de Aplicação (SQ8)**

Os resultados desta subquestão indicaram que 53,20% das tecnologias são específicas, isto é, limitadas a um tipo específico de aplicação ou processo de desenvolvimento (por exemplo, aplicação móvel, ciclo de vida ágil, entre outros) (ver Tabela 14). Por exemplo, Furtado *et al.* (2007) propôs um *framework* conceitual para associar os conceitos de usabilidade, educação de computadores e qualidade efetiva. Os autores analisaram a interação professor-

aluno, à luz das estratégias de aprendizagem utilizadas em geo-simuladores educacionais para a definição dos principais construtos emocionais envolvidos nesse processo. O *framework* conceitual é específico para os Sistemas Educativos de Geo-simulação.

Cerca de 46,80% das tecnologias identificadas foram utilizadas em um contexto genérico, isto é, não limitado a um tipo específico de aplicação ou processo de desenvolvimento. Por exemplo, Seffah *et al.* (2006) propôs a Qualidade de Uso Integrada à Medição (*Quality in Use Integrated Measurement* - QUIM) como um modelo consolidado para a medição da usabilidade. O QUIM é hierárquico em que decompõe a usabilidade em fatores, critérios e, finalmente, em métricas específicas. O principal objetivo do QUIM é fornecer um *framework* consistente e repositório de fatores de usabilidade, critérios e métricas para fins educacionais e de pesquisa.

Os resultados desta subquestão revelaram que a maioria das tecnologias são utilizadas em um contexto específico. Isto não é tão positivo, porque a uso dessa tecnologia só é possível se for aplicada dentro do contexto para o qual se destina. Estes resultados indicam que existe uma necessidade de mais tecnologias que possam ser empregadas em configurações gerais, isto é, não limitadas a um processo de desenvolvimento específico ou a um tipo de aplicação.

#### **4.4. Combinação dos Resultados das Subquestões**

As subquestões de pesquisa deste MSL foram combinadas com o objetivo de fornecer uma visão geral do campo Avaliação/Projeto de “*Early Usability*”.

Subquestões SQ4 e SQ5 não foram utilizadas nesta combinação, porque elas têm várias possíveis respostas. Além disso, a subquestão SQ7 não foi usada porque foi representada neste mapeamento por SQ7.1, onde a alternativa “sim” foi representada pelos tipos de estudos experimentais e a alternativa “não” foi representada pela resposta “nenhum”. Este tipo de análise permitiu obter mais informações sobre como os resultados de cada subquestão estão relacionados com os resultados de outras subquestões e quais as possíveis lacunas na pesquisa.

A Figura 12 mostra a combinação dos resultados obtidos a partir da subquestão de pesquisa SQ3.1 (Tipo de tecnologia de avaliação de usabilidade) em comparação com as subquestões SQ1 (Estágio do processo de desenvolvimento) e SQ7.1 (Tipos de estudos experimentais). Estes resultados indicam que: há uma escassez de tecnologias de simulação que são aplicadas nas fases de análise e projeto do processo de desenvolvimento. Isso pode ter acontecido porque poucos modelos permitem a simulação. Além disso, nenhuma tecnologia de

simulação foi avaliada através de um estudo de viabilidade, estudo de observação, estudo de caso ou *survey*. Portanto, é necessário realizar experimentos com tecnologias de simulação para verificar como esse tipo de tecnologia se comporta.

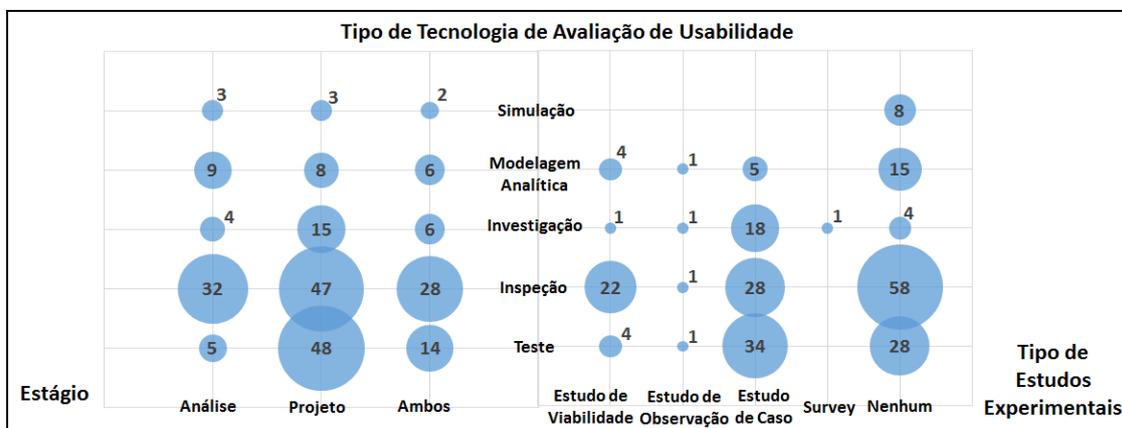


Figura 12. Combinação 1 de subquestões de pesquisa

A Figura 13 mostra a combinação dos resultados das subquestões de pesquisa SQ3 (Tipo de tecnologia) e SQ2 (Área de pesquisa) em comparação com as subquestões de pesquisa SQ8 (Contexto de aplicação) e SQ6 (Apoio ferramental). Os resultados indicam que: (a) há uma escassez de tecnologias de projeto de usabilidade que são aplicadas em contextos gerais e que apresentam apoio ferramental, e (b) poucas tecnologias de ES são usadas em um contexto genérico e precisam de apoio de ferramentas.

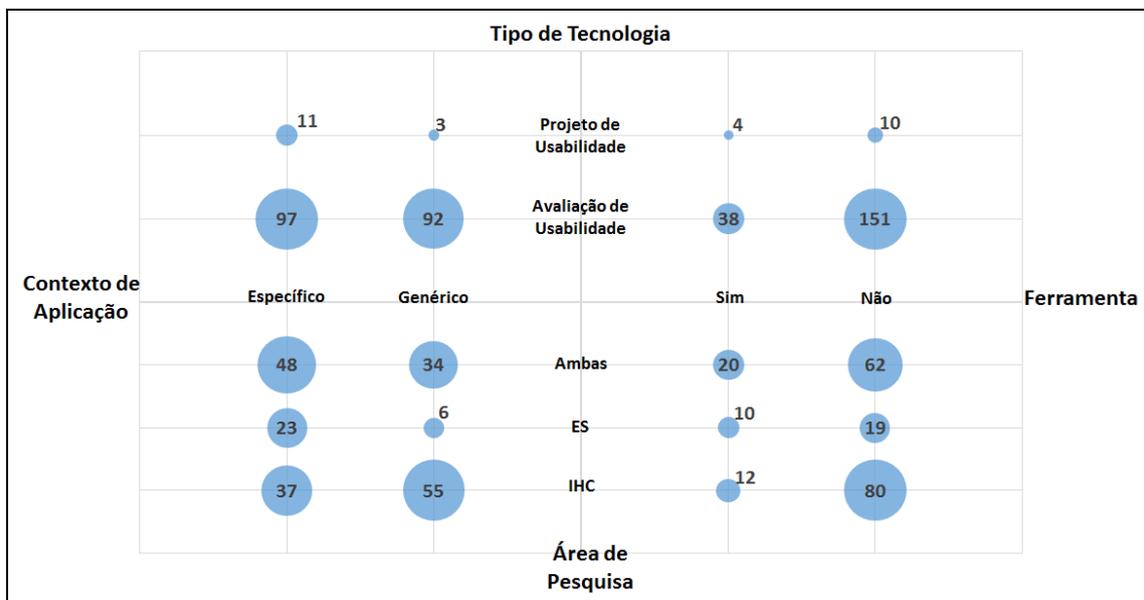


Figura 13. Combinação 2 de subquestões de pesquisa

A Figura 14 mostra a combinação dos resultados da subquestão de pesquisa SQ7.1 (Tipos de estudos experimentais) em comparação com as subquestões de pesquisa SQ7.2 (Ambiente de avaliação) e SQ7.3 (Tipos de análise). Estes resultados indicam que: (a) a maioria dos estudos de caso são realizados em um ambiente industrial com profissionais e apresentam a análise de estudos de forma qualitativa; e (b) há uma escassez de estudos de viabilidade que apresentam a análise de forma qualitativa. A análise qualitativa fornece informações relevantes para a evolução das tecnologias propostas.

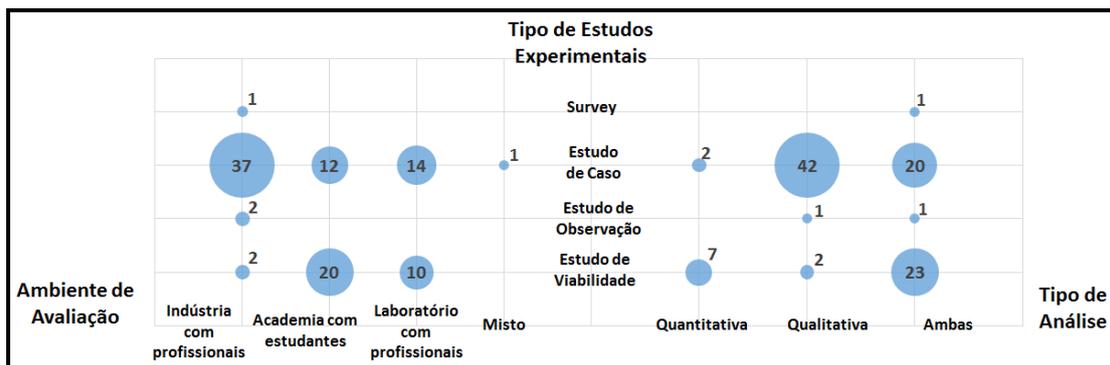


Figura 14. Combinação 3 de subquestões de pesquisa

A Figura 15 mostra a combinação dos resultados da subquestão de pesquisa SQ3 (Tipo de tecnologia) em comparação com as subquestões de pesquisa SQ1 (Estágio do processo de desenvolvimento) e SQ2 (Área de pesquisa). Os resultados mostram que: (a) poucas tecnologias de projeto de usabilidade são aplicadas nas fases de análise e projeto do processo *software*; (b) apenas algumas tecnologias de avaliação da usabilidade são propostas pela área de ES.

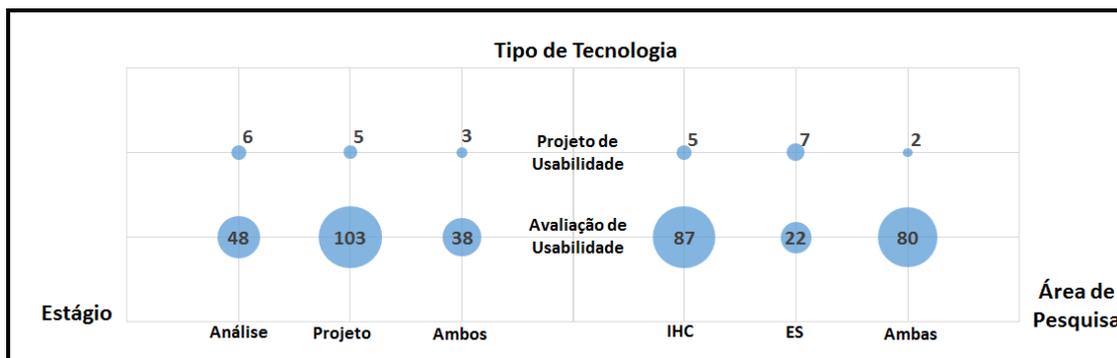


Figura 15. Combinação 4 de subquestões de pesquisa

Uma das maiores lacunas de pesquisa identificadas neste mapeamento sistemático foi a escassez de tecnologias de usabilidade que apoiam a criação de modelos desde as primeiras fases

do processo de desenvolvimento. São necessárias novas tecnologias de projeto de usabilidade, principalmente para apoiar a criação de modelos que visem a usabilidade das aplicações finais. Além disso, se um processo de *software* emprega tecnologias de projeto de usabilidade, necessitará de menos atividades para desenvolver aplicações com boa usabilidade.

#### 4.5. Ameaças à Validade

A seguir algumas ameaças à validade deste MSL são descritas.

Estratégia de Pesquisa: pode ter ocorrido de que alguns artigos relevantes não tenham sido coletados, o que pode afetar os resultados deste MSL. Para mitigar essa ameaça, duas grandes bases de dados digitais da área de ES e IHC (ACM e Scopus) (FERNANDEZ *et al.*, 2011a; PETERSEN *et al.*, 2015) foram investigadas e por fim obteve-se 136 artigos com 203 tecnologias. Além disso, considerou-se artigos de controle, isto é, artigos relevantes que foram identificados no primeiro mapeamento sistemático que foi realizado, descrito no Capítulo 3 desta tese, que abordam sobre “*Early Usability*”. Depois de verificar os artigos retornados, todos os artigos de controle foram retornados neste MSL. Após a aplicação dessa medida, a probabilidade de ausência de artigos relevantes foi reduzida. Além disso, vale ressaltar que a *string* de busca não cobriu o termo “User eXperience (UX)” pois não se desejava identificar tecnologias de UX voltadas para aspectos hedônicos e sim para aspectos pragmáticos. Aspectos pragmáticos se referem às percepções de um produto com relação ao cumprimento da sua finalidade (HASSENZAHN *et al.*, 2010). Estes aspectos se concentram na utilidade e usabilidade em relação às tarefas a serem realizadas na aplicação. Já os aspectos hedônicos focam mais nas necessidades e expectativas humanas que alguém tem para usar uma aplicação (HASSENZAHN *et al.*, 2010), isto é, como uma pessoa se sente ao usá-la. Portanto, neste MSL focou-se em identificar tecnologias que abordam um dos aspectos pragmáticos, a usabilidade.

O processo de seleção depende, em grande parte, do conhecimento pessoal e da experiência dos pesquisadores que conduziram este MSL, o que poderia ter introduzido viés nos resultados da seleção. As seguintes estratégias foram utilizadas para mitigar esses vieses: (a) um conjunto de critérios de inclusão e exclusão para seleção de artigo foram fornecidos como base de um processo de seleção objetivo; (b) as possíveis interpretações e entendimentos diferentes dos critérios de seleção pelos pesquisadores foram considerados. Para isso, realizamos reuniões de discussão entre os pesquisadores antes da seleção formal para garantir que os pesquisadores

alcançaram um entendimento claro e consistente dos critérios de seleção; (c) dois pesquisadores realizaram a seleção de artigos na primeira e segunda etapas do processo de seleção e outro pesquisador resolveu quaisquer conflitos entre os resultados dos dois pesquisadores a fim de mitigar o viés pessoal. Além disso, os pesquisadores já haviam tido experiência na execução de outro MSL (SILVA *et al.*, 2015), descrito no Capítulo 3 desta tese.

Os resultados da extração de dados podem ter sido negativamente afetados devido ao viés dos pesquisadores que extraíram os dados. Viés na extração de dados pode resultar em uma imprecisão dos itens de dados extraídos. Esse viés foi mitigado por três medidas: (a) uma lista de itens de dados extraídos foi especificada em detalhes para reduzir possíveis mal-entendidos sobre os itens de dados a serem extraídos; (b) discussões foram realizadas sobre os dados extraídos dos pesquisadores ao longo de todo o processo de extração de dados para melhorar a consistência e correção dos resultados da extração de dados; e (c) os resultados de todas as extrações de dados também foram verificados pelos pesquisadores para reduzir o viés pessoal.

#### **4.6. Síntese do Capítulo**

Este capítulo descreveu os resultados do 2º estudo secundário (Mapeamento Sistemático). Este estudo tem apresentado evidências sobre as tecnologias propostas que apoiam a usabilidade das aplicações nas fases iniciais do processo de desenvolvimento. Um total de 130 trabalhos de pesquisa foram selecionados neste MSL.

Os resultados obtidos neste MSL apresentam várias tecnologias que se concentram em apoiar os profissionais na melhoria da usabilidade das aplicações através de modelos de análise e projeto. Os resultados também mostram que há uma necessidade para a criação de novas tecnologias para apoiar o projeto de usabilidade. Com este tipo de tecnologia, a usabilidade pode ser considerada de forma proativa no processo de desenvolvimento de *software*, evitando retrabalho e esforço dos profissionais. Este MSL revelou evidências de lacunas de pesquisa para pesquisadores de ambas as áreas, tais como: (a) criação de novas tecnologias de usabilidade para a área de ES, (b) execução de mais estudos de viabilidade que apresentem a análise de forma qualitativa e, (c) proposta de mais tecnologias de simulação aplicadas nas fases de análise e projeto do processo de desenvolvimento.

Os resultados deste MSL têm implicações para os pesquisadores que estão planejando propor novas tecnologias de avaliação/projeto de usabilidade ou estão dispostos a realizar

novos estudos; e para os profissionais que estão trabalhando em empresas de desenvolvimento de *software* e gostariam de integrar tecnologias de usabilidade em seus processos de *software* de forma eficaz. Para os pesquisadores, seria interessante propor novas tecnologias de acordo com as lacunas de pesquisa identificadas acima, pois há necessidade de mais tecnologias de simulação e tecnologias de modelagem analítica aplicadas nas fases de análise e projeto do processo de desenvolvimento. Para os profissionais, é relevante saber em quais fases do processo de desenvolvimento as tecnologias são aplicadas. Através dos resultados deste MSI, pode-se ver que a maioria das tecnologias são usadas na fase de projeto usando protótipos para teste de usabilidade e inspeção. Além disso, foi identificado que poucos estudos observacionais foram realizados com as tecnologias propostas. Em relação aos métodos propostos, é relevante conduzir estudos de observação porque proporcionam resultados qualitativos mais precisos.

## CAPÍTULO 5 - CONJUNTO DE TECNOLOGIAS QUE APOIA A ANTECIPAÇÃO DA USABILIDADE NAS FASES INICIAIS DO PROCESSO DE DESENVOLVIMENTO DE SOFTWARE

*Este capítulo apresenta a proposta de um conjunto de tecnologias para auxiliar na antecipação da usabilidade. Este conjunto de tecnologias objetiva fornecer uma maneira proativa de se considerar a usabilidade em projetos de desenvolvimento de software se comparada a uma abordagem tradicional.*

### 5.1 Introdução

O objetivo principal desta pesquisa consiste em fornecer um conjunto de tecnologias que apoie a avaliação e criação de modelos utilizados nas fases iniciais do processo de desenvolvimento, visando à usabilidade da aplicação final. Este conjunto de tecnologias é apoiado por um conjunto de recursos que podem ser empregados pelos próprios envolvidos no projeto de desenvolvimento. Para alcançar isto, o conjunto de tecnologias deve atender aos seguintes requisitos: ser fácil de aprender e utilizar, apresentar bom nível de eficiência e eficácia, além de oferecer uma boa relação custo-benefício em sua aplicação.

O objetivo é que este conjunto de tecnologias seja:

1. Independente de ferramenta: não se limita à disponibilidade de uma ferramenta, além de ser distribuído de forma gratuita;
2. Independente de processo de desenvolvimento: não se limita a um determinado processo de desenvolvimento;
3. Abrangente: compreende vários modelos das fases iniciais do processo de desenvolvimento;
4. Usado por profissionais de *software* com pouco conhecimento em usabilidade: auxilie profissionais novatos na melhoria da usabilidade, ou seja, profissionais não especialistas em usabilidade;

5. Independente do tipo de aplicação: não se limita a um determinado tipo de aplicação ou ambiente.

Este conjunto de tecnologias conterà dois tipos de tecnologias conforme ilustrado na Figura 16: (1) para avaliação de modelos das fases iniciais e; (2) para a criação de modelos das fases iniciais; cujo objetivo é a antecipação da usabilidade das aplicações finais.



Figura 16. Componentes do Conjunto de Tecnologias para antecipação da Usabilidade

Para apoiar a avaliação de usabilidade através de modelos utilizados nas fases iniciais do processo de desenvolvimento foi proposto um conjunto de técnicas de leitura chamada MIT (*Model Inspection Technique for Usability Evaluation*). Estas técnicas serão apresentadas na Subseção 5.2.

Para apoiar o projeto de modelos utilizados nas fases iniciais do processo de desenvolvimento (o Diagrama de Atividades e o *Mockup*) visando à usabilidade da aplicação foram propostas as técnicas UDRT-AD (*Usability Design Reading Technique for Activity Diagrams*) e UDG (*Usability Design Guidelines*). Ambas serão apresentadas na Subseção 5.3.

Vale ressaltar que as tecnologias propostas podem ser utilizadas de forma independente. Ou seja, se o projeto ou a empresa de *software* só utiliza um modelo específico, como, por exemplo, o *diagrama de atividades*, ele/ela só precisa utilizar a tecnologia que cria ou avalia este modelo. No entanto, se a empresa de *software* utiliza todos os modelos, ela pode utilizar todas as tecnologias para apoiar o projeto e/ou avaliação destes modelos durante o seu processo de desenvolvimento de *software*.

Caso o projeto de *software* esteja no início e nenhum modelo ainda tenha sido criado, recomenda-se utilizar as tecnologias que apoiam o projeto destes modelos, como a UDG para apoiar a construção de *mockups* e a UDRT-AD para apoiar a construção de diagramas de atividades. Caso a projeto de *software* já tenha construído estes modelos, recomenda-se utilizar somente as tecnologias para avaliação, que são: MIT 1 para especificações de casos de uso, MIT 2 para *mockups* e MIT 3 para diagramas de atividades.

## **5.2 Proposta de Tecnologias que auxiliam na Avaliação de Modelos utilizados nas Fases Iniciais do Processo de Desenvolvimento, visando à Usabilidade da Aplicação**

Esta Subseção apresenta o conjunto de técnicas MIT (*Model Inspection Technique for Usability Evaluation*) cuja proposta inicial foi feita no âmbito de minha pesquisa de mestrado. Em minha pesquisa de doutorado buscou-se evoluir estas técnicas através dos resultados dos estudos experimentais. Estes experimentos serão descritos em mais detalhes nos próximos capítulos desta tese de doutorado.

As MITs são técnicas de leitura para inspeção de usabilidade em modelos de análise e projeto (VALENTIM *et al.*, 2013; VALENTIM e Conte, 2014; VALENTIM *et al.*, 2015a, VALENTIM *et al.*, 2015b, VALENTIM *et al.*, 2015c). Segundo Travassos *et al.* (2002), 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. Estas têm como objetivo aumentar a eficácia das inspeções, fornecendo diretrizes que podem ser utilizados pelos revisores para analisar um determinado artefato de *software* e identificar defeitos (TRAVASSOS *et al.* 2002). Desta forma, as técnicas propostas possuem um conjunto de itens de verificação que devem guiar o inspetor durante a avaliação de usabilidade em Caso de Uso, em *Mockups* e em Diagrama de Atividades (Figura 17).

Para a realização da inspeção de usabilidade através destas técnicas é sugerido haver mais que um avaliador, pois, segundo Rocha e Baranauskas (2003), uma única pessoa nunca é capaz de encontrar todos os problemas de usabilidade existentes em um modelo.

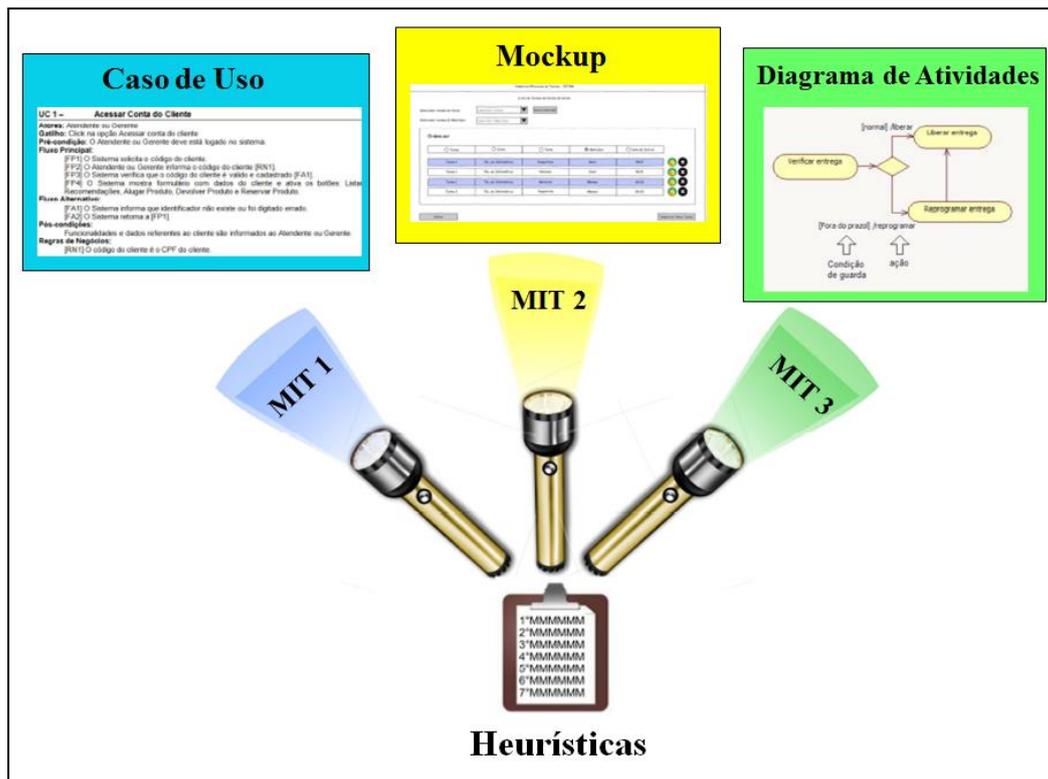


Figura 17. Visão geral do Conjunto de Técnicas MIT

### 5.2.1 Técnica MIT 1

Um dos modelos construídos durante os estágios iniciais do processo de desenvolvimento é o Modelo de Casos de Uso, que é composto por um conjunto de casos de uso que descrevem as funcionalidades dos requisitos através de especificações de casos de usuário. Da Cruz (2014) argumenta que um modelo de caso de uso de um sistema de *software* pode ser usado como um modelo das funcionalidades requeridas do sistema e as restrições necessárias sobre a interação entre o usuário, identificado como ator, e o próprio sistema. A especificação de um caso de uso é normalmente feita através de uma descrição textual (DA CRUZ, 2014). As especificações de caso de uso são um recurso importante para orientar os desenvolvedores a projetar as interações do sistema com os usuários, bem como desenvolver o próprio sistema. Este tipo de especificação tem sido sugerido como um meio valioso para integrar a engenharia de usabilidade diretamente no processo de desenvolvimento de *software* (FERRÉ *et al.*, 2011). Existem diversos estilos de especificações de caso de uso, porém o que será considerado nesta pesquisa é o que contém fluxo de eventos que descrevem como problemas são solucionados pelo sistema.

A MIT 1 tem como objetivo avaliar a usabilidade através de especificações de Casos de Uso. A MIT 1 possui 34 itens de verificação, agrupados nas seguintes heurísticas: Visibilidade do status do sistema (Figura 18); Concordância entre o sistema e o mundo real;

Controle e liberdade do usuário; Consistência e padrões; Prevenção de erros; Reconhecer ao invés de lembrar; Flexibilidade e eficiência de uso; Reconhecimento, diagnóstico e recuperação de erros; e Proporcionar funções que são úteis para o usuário. Pode-se observar que somente algumas das heurísticas Nielsen (1994b) foram consideradas como base para a MIT 1. Isto deve-se ao fato de apenas algumas delas poderem ser adaptadas para avaliar especificações de caso de uso. A técnica MIT 1 completa é apresentada no APÊNDICE A.

<b>MIT-1A. Visibilidade do Status do Sistema</b>	
<b>MIT-1A1</b>	Verifique se há algum texto no FP, FA e FE que informa em que parte do sistema o usuário se encontra;
<b>MIT-1A2</b>	Verifique se há algum texto no FP, FA e FE que informa ao usuário o que foi realizado após uma persistência de dados. Por exemplo: quando há alteração ou exclusão de algo, uma mensagem de texto é apresentada.

**Figura 18. Itens da heurística Visibilidade do status do Sistema (MIT 1)**

Os passos do processo de inspeção utilizando a técnica MIT 1 são mostrados na Figura 19. Estes passos são: (1) avaliar o caso de uso e (2) identificar problemas de usabilidade. A fim de ilustrar o processo de inspeção com a MIT 1, utilizou-se uma especificação de caso de uso. Esta especificação descreve a interação entre o usuário e a funcionalidade Gerenciar um Curso de um sistema online que disponibiliza indicadores de pesquisa e desenvolvimento no Brasil.

O primeiro passo para a identificação de problemas de usabilidade é seguir com a avaliação dos itens de verificação de usabilidade. Em outras palavras, os inspetores podem checar se a especificação de caso de uso atende cada item de verificação de usabilidade.

A fim de identificar problemas de usabilidade (segundo passo), inspetores podem apontar na aplicação que parte não atendem aos itens de verificação de usabilidade. Se olharmos a Figura 19 e a técnica MIT 1, pode-se relatar as não conformidades dos itens de verificação de usabilidade com o elemento A destacado na Figura 19.

O item de verificação MIT 1AB1 orienta que nome de campos, telas, botões, links, mensagens de erro e textos informativos tenham conceitos familiares para usuários, ou seja, seguem as convenções do mundo real. No entanto, o nome da tela “Registro de curso – Centro de Treinamento” não tem um nome que segue as convenções do mundo real (ver Figura 19 elemento A). Ou seja, este fluxo alternativo especifica a funcionalidade “Editar” e o nome da tela não representa esta funcionalidade. Este é um exemplo de problema de usabilidade identificado com a MIT 1.

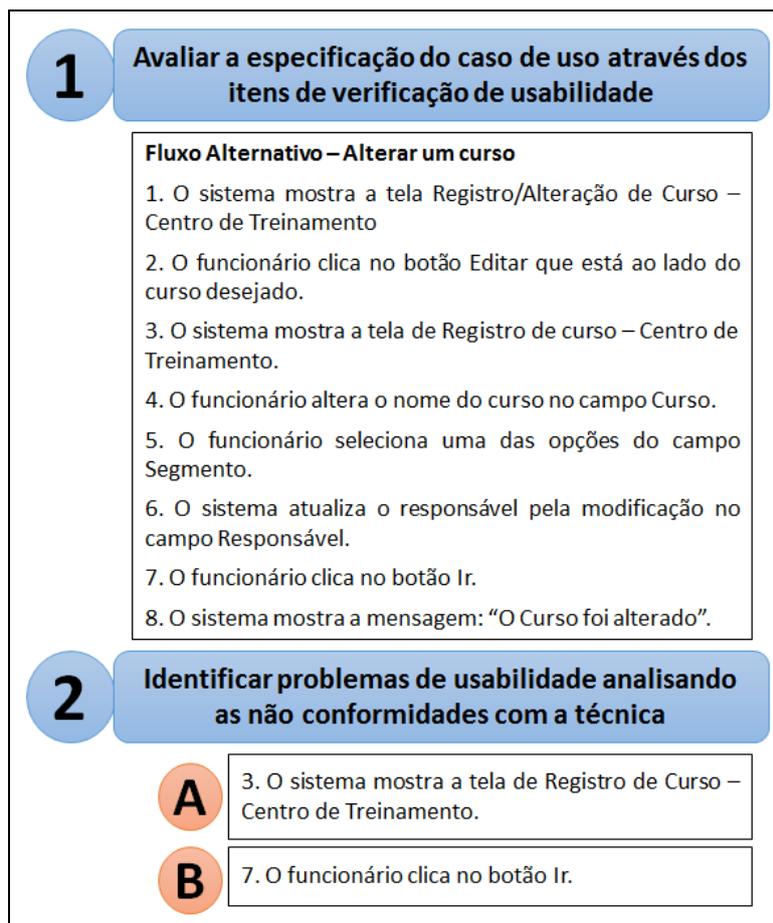


Figura 19. Exemplo do processo de inspeção com a MIT 1

Já o item de verificação MIT 1AE2 orienta que todas as opções, botões e links apresentados na aplicação devem ter nomes que definem claramente que resultados ou estados serão alcançados. No entanto, o botão Ir não tem um nome que claramente indica qual estado será alcançado ao selecioná-lo (ver Figura 19 elemento B). Ou seja, o nome do botão Ir não é representativo para o usuário, caracterizando outro problema de usabilidade.

### 5.2.2 Técnica MIT 2

A técnica MIT 2 considera a avaliação de usabilidade em *mockups*. A MIT 2 utiliza procedimentos de leitura para os inspetores avaliarem os *mockups* através de 29 itens de verificação, agrupados nas seguintes heurísticas: Visibilidade do status do sistema; Concordância entre o sistema e o mundo real; Controle e liberdade do usuário (Figura 20); Consistência e padrões; Prevenção de erros; Reconhecer ao invés de lembrar; Flexibilidade e eficiência de uso; Reconhecimento, diagnóstico e recuperação de erros; e proporcionar funções que são úteis para o usuário. Algumas das heurísticas de Nielsen (1994b) não foram adaptadas para avaliar *mockups*, pois neste tipo de modelo não se tem o nível de detalhes e informações que a heurística de Nielsen (1994b) orienta avaliar. A técnica MIT 2

completa é apresentada APÊNDICE B. O processo de inspeção utilizando a técnica MIT 2 é similar ao processo utilizando a MIT 1 apresentado anteriormente.

<b>MIT-2C. Controle e liberdade ao usuário</b>	
<b>MIT-2C1</b>	Verifique se o usuário tem as opções de desfazer ou refazer algo que ele tenha escolhido;
<b>MIT-2C2</b>	Verifique se o usuário tem a opção de cancelar o que está realizando, ou se há opções similares que permitam ao usuário utilizar saídas em caso de escolhas erradas ou para sair de um estado ou local não esperado. Por exemplo, ter a possibilidade de clicar na opção “Cancelar” ou “Sair”;

Figura 20. Itens da heurística Controle e liberdade ao usuário (MIT 2)

### 5.2.3 Técnica MIT 3

Por fim, a MIT 3 faz avaliação de usabilidade em diagrama de atividades. O estilo de diagrama de atividades considerado nesta pesquisa ilustra o fluxo de atividades de um caso de uso. Este tipo de diagrama de atividades explora a ordem das atividades. A MIT 3 possui 14 itens de verificação, agrupados nas seguintes heurísticas: Concordância entre o sistema e o mundo real (Figura 21); Controle e liberdade do usuário; Consistência e padrões; Prevenção de erros; Reconhecer ao invés de lembrar; Flexibilidade e eficiência de uso; Reconhecimento, diagnóstico e recuperação de erros; e proporcionar funções que são úteis para o usuário. Da mesma forma que nas técnicas MIT 1 e MIT 2, algumas das heurísticas de Nielsen (1994b) não foram adaptadas para avaliar diagrama de atividades na MIT 3, pois neste tipo de modelo não se tem o nível de detalhes e informações que a heurística de Nielsen (1994b) orienta avaliar. A técnica MIT 3 completa é apresentada APÊNDICE C. O processo de inspeção utilizando a técnica MIT 3 também é similar ao processo utilizando a MIT 1 apresentado anteriormente.

<b>MIT-3A. Concordância entre o sistema e o mundo real</b>	
<b>MIT-3A1</b>	Verifique se as atividades estão apresentadas em uma ordem natural e lógica segundo os conceitos do domínio do problema;
<b>MIT-3A2</b>	Verifique se os nomes das atividades utilizam termos (palavras) que seguem as convenções do mundo real, ou seja, que englobam tanto as convenções do domínio do problema quanto às convenções de terminologia de aplicações semelhantes.

Figura 21 - Itens da heurística Concordância entre o sistema e o mundo real

O propósito deste conjunto de técnicas é que seja facilmente adotável pela indústria de desenvolvimento de *software*. Além disso, outros objetivos são diminuir o esforço de retrabalho e reduzir o custo das correções dos problemas de usabilidade, já que é mais barato resolver estes problemas em estágios iniciais de desenvolvimento do que quando o sistema se encontra finalizado (FERNANDEZ *et al.*, 2011a).

Uma das características das MITs é que elas não necessitam de apoio ferramental. Outra característica relevante é que as técnicas podem ser usadas de forma independente. No entanto, melhores resultados são encontrados se utilizadas em conjunto pelo fato de integrar avaliação de artefatos de ES e IHC com o mesmo propósito de avaliação de usabilidade. Os Capítulos 6, 7 e 8 apresentam resultados dos estudos experimentais que avaliaram as MITs.

### **5.3 Proposta de Tecnologias que auxiliam no Projeto/Criação de Modelos utilizados nas Fases Iniciais do Processo de Desenvolvimento de *Software*, visando à Usabilidade da Aplicação**

Esta Subseção apresenta as duas técnicas propostas que apoiam a construção de modelos utilizados nas fases iniciais do processo de desenvolvimento de *software*.

#### **5.3.1 Técnica UDRT-AD**

A técnica UDRT-AD (*Usability Design Reading Technique for Activity Diagrams*) consiste em uma técnica de leitura para projetar Diagramas de Atividades visando à usabilidade durante o projeto de aplicações interativas (SILVA *et al.*, 2014a). A técnica proposta foi separada por heurísticas que auxiliam na construção dos diagramas de atividades, extraindo os elementos para este diagrama a partir de uma descrição textual e apoiando uma melhor usabilidade na aplicação projetada. A UDRT-AD tem como objetivos: (a) auxiliar os projetistas das aplicações a identificar os elementos do diagrama na descrição textual, (b) construir o diagrama já apoiando uma melhor usabilidade na aplicação e (c) verificar a consistência do diagrama modelado.

Ao utilizar a técnica, conforme mostra a Figura 22, o projetista (Figura 22 elemento a) deve primeiramente buscar compreender a descrição textual (Figura 22 elemento b) e entender o domínio do problema ao qual se deseja solucionar. Em seguida, o projetista utilizará um conjunto de heurísticas (Figura 22 elemento c), que servem como passos, para extrair os elementos da descrição e construir o diagrama de atividade. Este diagrama poderá ser usado depois para apoiar a construção da aplicação final. Nesta técnica, o diagrama é criado pensando na usabilidade da aplicação que está sendo projetada. A técnica permite ainda ao projetista analisar a consistência do diagrama através de Itens de Verificação. Desta maneira, utilizando esta técnica, o projetista obterá um apoio (diagrama de atividades – Figura 22 elemento d) para desenvolver uma aplicação em que a usabilidade é planejada desde as fases iniciais do processo de desenvolvimento.

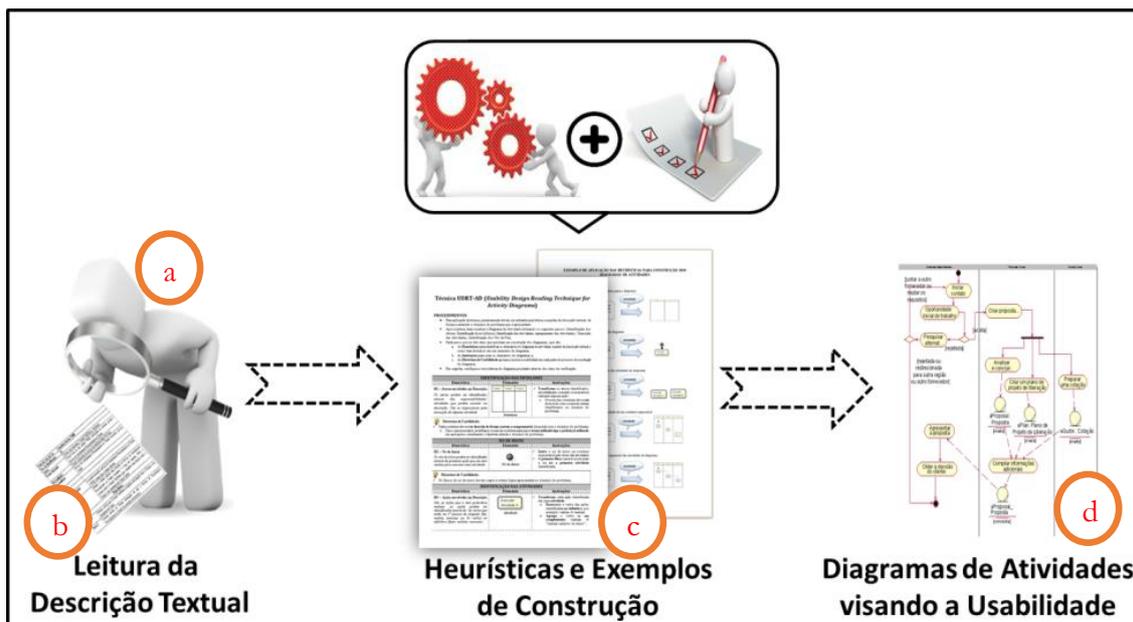


Figura 22. Processo de aplicação das heurísticas da técnica UDRT-AD

Esta técnica é composta por seis passos que auxiliam os projetistas na construção: Identificação dos Atores, Identificação do nó de Início, Identificação das Atividades, Agrupamento das Atividades, Transição das Atividades, Identificação dos Nós de Fim. Cada passo é composto de:

- a) uma Heurística de Construção que ajuda a identificar os elementos do diagrama na descrição textual;
- b) Instruções de como construir/inserir esses elementos no diagrama; e,
- c) uma Diretriz de Usabilidade que guia os projetistas, a partir do diagrama, na inclusão da usabilidade no diagrama.

Além disso, para auxiliar no uso da técnica, cada Heurística de Construção e Instrução de Construção possui um exemplo de aplicação, mostrando como extrair o elemento do diagrama de atividade a partir da descrição textual. A Figura 23 apresenta um exemplo de uma heurística e de um exemplo de construção dos elementos.

O texto completo da técnica UDRT-AD e os exemplos de construção são apresentados no APÊNDICE D e APÊNDICE E, respectivamente. O propósito desta técnica também é que seja facilmente adotável pela indústria de desenvolvimento de *software*. E uma de suas características é que também não necessita de apoio ferramental específico.

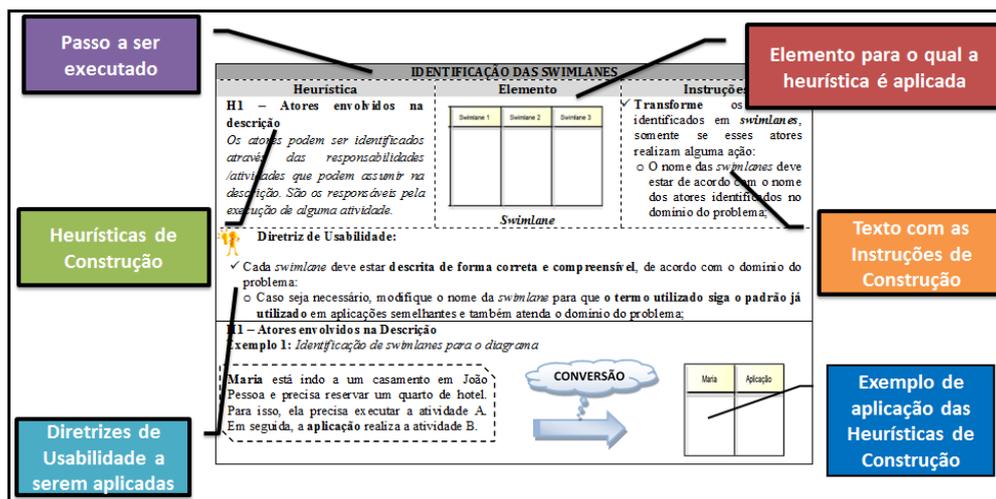


Figura 23. Exemplo de um dos elementos da técnica UDRT-AD

### 5.3.2 Técnica UDG

*Usability Design Guidelines* (UDG) é uma técnica para apoiar a construção de *mockups* visando à usabilidade durante o projeto de aplicações. As informações e funcionalidades básicas de uma aplicação são extraídas de uma descrição textual, que pode ser um documento de especificação de requisitos, especificação de caso de uso, cenário, dentre outros. A partir daí, utiliza-se a UDG para apoiar o projeto do *mockup* já pensando na usabilidade da aplicação.

A Figura 24 apresenta, de maneira similar à UDRT-AD, como utilizar a UDG. O projetista (Figura 24 elemento a) deve compreender a descrição textual e o domínio do problema ao qual se deseja solucionar (Figura 24 elemento b). Em seguida, o projetista utilizará um conjunto de *guidelines* separados por heurísticas (Figura 24 elemento c), que servem como orientações para a construção dos *mockups* refletindo sobre a usabilidade. Desta maneira, utilizando esta técnica, o projetista obterá *mockups* (Figura 24 elemento d) que podem servir de apoio para outras etapas do processo de desenvolvimento, por exemplo, apoio para validar requisitos com *stakeholders* e apoio nas discussões sobre a usabilidade da aplicação nas fases iniciais do processo de desenvolvimento.

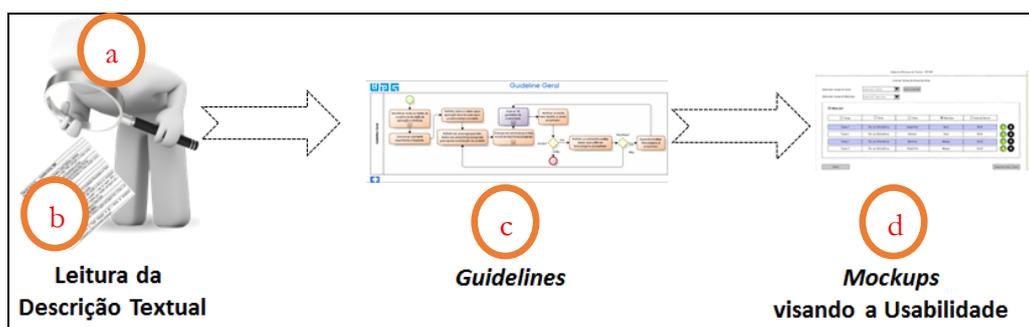


Figura 24. Processo de aplicação da UDG

A técnica proposta é separada por heurísticas e possuem *guidelines* que auxiliam na construção dos *mockups*. Para a definição das *guidelines* da UDG foram investigadas e utilizadas 8 fontes bases que abordam sobre usabilidade, apresentadas a seguir:

- Nielsen (1994b): apresenta um método de inspeção chamado Avaliação Heurística (AH), já descrito no Capítulo 2 desta tese (Subseção 2.2.1). Esta fonte foi escolhida, pois é um dos métodos mais conhecidos e amplamente utilizado pela indústria de *software* (FERNANDEZ *et al.*, 2012a).
- ISO/IEC 25010 (2011): apresenta um modelo de qualidade em que categoriza propriedades de qualidade de produto de *software* em oito características, sendo usabilidade uma delas. Nesta ISO, a usabilidade possui um conjunto de subcaracterísticas, tais como Reconhecimento de adequação, Aprendizagem, Operabilidade, Proteção de erro do usuário, Estética da interface do usuário e Acessibilidade. Esta fonte foi escolhida pois trata-se de uma norma que define e apresenta algumas subcaracterísticas de usabilidade e que tem sido utilizada por diversas instituições de pesquisa e desenvolvimento de *software* como padrão.
- Fernandez (2012b): propôs o Modelo de Usabilidade *Web* (MUW), que considera as subcaracterísticas de usabilidade propostas na ISO/IEC 25010 (2011). No entanto, como estas subcaracterísticas são altamente genéricas e também são definidas em um alto nível de abstração, o MUW quebra estas subcaracterísticas em outras subcaracterísticas e atributos, a fim de cobrir um conjunto maior de aspectos da usabilidade *Web*. Esta fonte foi escolhida pois, embora seja específica para *Web*, trata-se de uma proposta que aprofunda o nível de detalhamento das subcaracterísticas de usabilidade definidas na ISO/IEC 25010 (2011).
- Carvajal (2012): propôs um processo para apoiar desenvolvedores de *software* a incluir usabilidade em suas aplicações através do uso de *guidelines* de usabilidade. As *guidelines* de usabilidade incorporam algumas características de usabilidade em aplicações de *software*. Os principais artefatos deste processo especificam as responsabilidades que o sistema e suas partes devem cumprir para estar em conformidade com essas características de usabilidade, tornando-os diretamente implementáveis a partir do design. Esta fonte foi escolhida por apresentar *guidelines* de usabilidade que orientam desenvolvedores a construir artefatos da sua aplicação incluindo a usabilidade. Mesmo que estas *guidelines* de usabilidade

sejam específicas para o processo proposto por Carvajal (2012), elas podem servir de base para a UDG.

- Valentim *et al.*, (2015b): apresenta uma das técnicas propostas nesta pesquisa, chamada MIT 2. Esta técnica está descrita em mais detalhes na Subseção 5.2. Esta fonte foi escolhida pois apresenta itens de verificação de usabilidade que podem ser adaptados para se tornarem *guidelines* que orientam a construção de *mockups* com boa usabilidade.
- Shneiderman e Plaisant (2010): apresenta as “Regras de Ouro” que são aplicáveis à maioria dos sistemas interativos. As “Regras de Ouro” são princípios e práticas de design de boa qualidade. Esta fonte foi escolhida, pois, estes princípios e práticas de design foram derivados da experiência e aperfeiçoados ao longo de três décadas, tanto na academia quanto na indústria.
- Tognazzini (1992; 2014): apresenta alguns princípios fundamentais para a projeto e implementação de interfaces eficazes. Recentemente, o autor fez uma revisão e expansão destes princípios, acrescentando novos subprincípios, novos exemplos, estudos de caso e discussões realizadas (TOGNAZZINI, 2014). Esta fonte foi escolhida pois, estes princípios de design também foram obtidos ao longo de várias pesquisas tanto na academia como na indústria.
- *Usability Guidelines* (USABILITYGUIDELINES, 2016): apresenta algumas *guidelines* definidas pelo governo americano. *Usability Guidelines* destinam-se a fornecer melhores práticas sobre uma ampla gama de projetos *web*, além de melhorar a usabilidade dos sites projetados. Esta fonte foi escolhida, pois, estas *guidelines* foram obtidas e compiladas pelo governo americano após extensa pesquisa e revisão. As referências correspondentes a cada *guideline* estão disponíveis no site do governo americano (USABILITYGUIDELINES, 2016).

Como pode ser observado, algumas tecnologias bases são específicas, por exemplo, são voltadas para um contexto *Web* ou são usadas em um processo específico. Nestes casos, as adaptações necessárias foram feitas para que as *guidelines* da UDG fossem genéricas, ou seja, pudessem ser utilizadas em qualquer contexto ou processo de *software*.

A Tabela 15 a seguir apresenta: (i) as heurísticas da UDG, na ordem em que são abordadas na técnica, (ii) uma breve explicação das heurísticas, e (iii) as fontes que serviram de base para a construção das *guidelines* referente a cada heurística.

Tabela 15. Heurísticas da técnica UDG e referências bases

Nº	Heurística	Explicação	Fonte
1	Design minimalista e estético	<i>Guidelines</i> de construção relacionados ao layout e estética do <i>mockup</i> da aplicação e de seus elementos.	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Valentim <i>et al.</i> , (2015b) Tognazzini (1992; 2014) USABILITYGUIDELINES (2016)
2	Concordância entre o sistema e o mundo real	<i>Guidelines</i> de construção relacionados à concordância dos elementos do <i>mockup</i> da aplicação com o mundo real.	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Valentim <i>et al.</i> , (2015b) Tognazzini (1992; 2014) USABILITYGUIDELINES (2016)
3	Reconhecimento e diagnóstico de erros	<i>Guidelines</i> de construção relacionados às ações que requerem que o usuário seja alertado/avisado através de mensagens.	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Carvajal (2012) Valentim <i>et al.</i> , (2015b) Shneiderman e Plaisant (2010) Tognazzini (1992; 2014)
4	Consistência e padrões	<i>Guidelines</i> de construção relacionados à padronização e consistência do <i>mockup</i> da aplicação e das mensagens.	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Valentim <i>et al.</i> , (2015b) Shneiderman e Plaisant (2010) Tognazzini (1992; 2014) USABILITYGUIDELINES (2016)
5	Feedback do Status da Aplicação	<i>Guidelines</i> de construção relacionados ao feedback do status da aplicação, representado no <i>mockup</i> .	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Carvajal (2012) Valentim <i>et al.</i> , (2015b) Shneiderman e Plaisant (2010) Tognazzini (1992; 2014) USABILITYGUIDELINES (2016)
6	Controle e liberdade ao usuário	<i>Guidelines</i> de construção relacionados ao apoio de voltar/sair/cancelar e apoio à desfazer/refazer, com representação no <i>mockup</i> da aplicação.	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Carvajal (2012) Valentim <i>et al.</i> , (2015b) Shneiderman e Plaisant (2010) Tognazzini (1992; 2014) USABILITYGUIDELINES (2016)
7	Diminuição da sobrecarga cognitiva	<i>Guidelines</i> de construção relacionados à densidade de informação, reconhecimento ao invés de relembração e facilidade de entendimento, representados no <i>mockup</i> .	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Valentim <i>et al.</i> , (2015b) Shneiderman e Plaisant (2010) Tognazzini (1992; 2014) USABILITYGUIDELINES (2016)
8	Flexibilidade e eficiência de uso	<i>Guidelines</i> de construção relacionados à flexibilidade e eficiência de uso da aplicação, representados no <i>mockup</i> .	Nielsen (1994b) Fernandez (2012b) Valentim <i>et al.</i> , (2015b) Shneiderman e Plaisant (2010) Tognazzini (1992; 2014) USABILITYGUIDELINES (2016)

Nº	Heurística	Explicação	Fonte
9	Prevenção de erros	<i>Guidelines</i> de construção relacionadas à mecanismos de prevenção de erros, antes que eles possam acontecer na aplicação, representados no <i>mockup</i> .	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Valentim <i>et al.</i> , (2015b) Shneiderman e Plaisant (2010) USABILITYGUIDELINES (2016)
10	Ajuda e documentação	<i>Guidelines</i> de construção relacionados aos elementos de ajuda e documentação necessários para aplicação, representados no <i>mockup</i> .	Nielsen (1994b) ISO/IEC 25010 (2011) Fernandez (2012b) Carvajal (2012) USABILITYGUIDELINES (2016)

A UDG foi criada para ser utilizada através de qualquer leitor de pdf. A versão completa da técnica UDG é apresentada no APÊNDICE F. A técnica UDG é composta por oito tipos de elementos. A Figura 25 apresenta, a tela inicial da UDG, descrevendo o significado de cada elemento. Os elementos da técnica UDG foram construídos com base na notação BPMN (***Business Process Model and Notation***). O elemento “Atividade Inicial” (Figura 25 elemento 1) marca o ponto onde a *guideline* inicia e é representada por um círculo verde de linha simples. O elemento “Atividade Final” (Figura 25 elemento 2) marca o ponto onde a *guideline* termina e é representado por um círculo rosa de linha grossa. O elemento “Atividade que apoia o projeto do *mockup*” (Figura 25 elemento 3) representa as atividades que orientam a construção dos *mockups* e é representado por um retângulo com pontas arredondadas na cor laranja. O elemento “Dica de projeto do *mockup*” (Figura 25 elemento 4) dá acesso às dicas de construção dos *mockups* e é representado por um retângulo com pontas arredondadas na cor laranja contendo uma pequena cruz no centro inferior. O elemento “Atividade que apoia o projeto da usabilidade” (Figura 25 elemento 5) representa as atividades que abordam orientações de usabilidade durante a construção dos *mockups* e é representado por um retângulo com pontas arredondadas na cor azul. O elemento “Dica de projeto de usabilidade” (Figura 25 elemento 6) dá acesso às dicas de usabilidade durante a construção dos *mockups* e é representado por um retângulo com pontas arredondadas na cor azul contendo uma pequena cruz no centro inferior. O elemento “Atividade que dá acesso às 10 *guidelines* de usabilidade” (Figura 25 elemento 7) dá acesso às 10 *guidelines* de usabilidade que orientam a construção de *mockups* com boa usabilidade e é representado por um retângulo com pontas arredondadas na cor roxa com uma pequena cruz no centro inferior. E por fim, o elemento “Decisão” (Figura 25 elemento 8) representa o ponto onde uma decisão deve ser tomada e é representado por um losango na cor amarela. Os elementos da UDG que contém a cruz são elementos clicáveis.

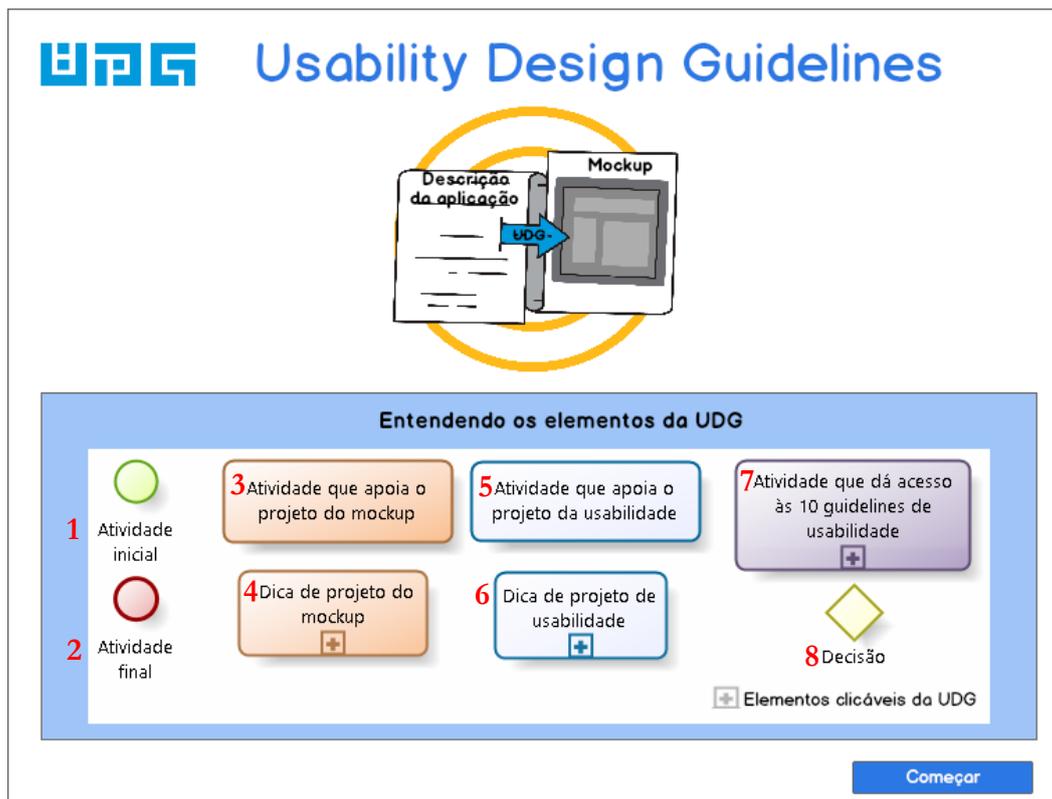


Figura 25. Elementos da técnica UDG

O projetista, entendendo o objetivo de cada elemento da UDG, pode começar a utilizar a técnica, clicando no botão “Começar” da tela inicial (Figura 25). A partir daí, ele segue as orientações (*guidelines*) que a UDG vai lhe dando. A Figura 26 apresenta um exemplo de *guidelines* da UDG. Estas *guidelines* apoiam o projeto do *mockup* e o projeto da usabilidade referente à heurística 2 “Concordância entre o sistema e o mundo real”. Através dos botões “1”, “2”, “3”, “4”, “5”, “6”, “7”, “8”, “9”, “10” e “Próxima *guideline*” representados na Figura 26 é possível navegar entre as 10 heurísticas da UDG. Além disso, através do botão “*Guideline* Geral” pode-se ir para a *Guideline* Geral da técnica.

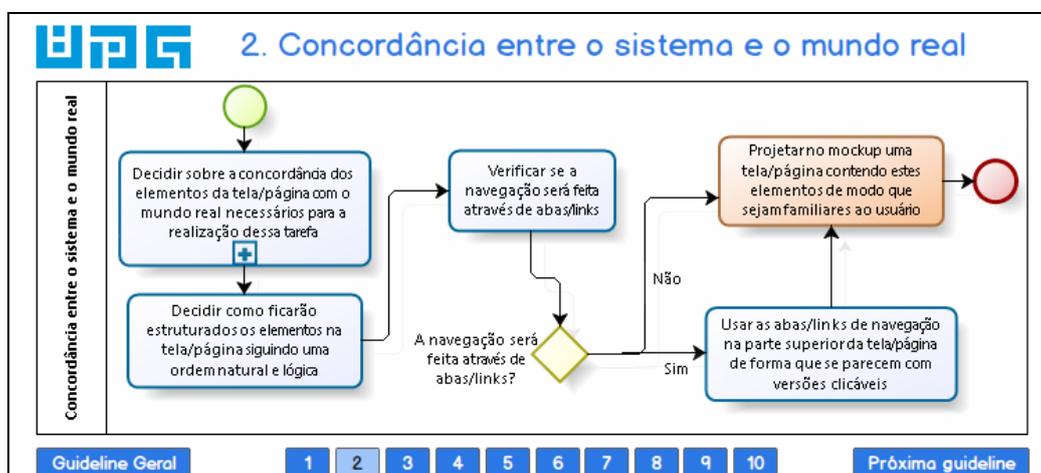


Figura 26. *Guidelines* da Heurística 2 da UDG

Para ilustrar as dicas fornecidas pela UDG, a Figura 27 apresenta um exemplo de dica de usabilidade referente à heurística 2 “Concordância entre o sistema e o mundo real”. Esta dica é acessada a partir da atividade “Decidir sobre a concordância dos elementos da tela/página com o mundo real necessários para a realização dessa tarefa” mostrada na Figura 26 (Atividade azul contendo uma cruz). Para sair da dica e voltar para a tela contendo as *guidelines* da heurística 2 é só clicar no botão chamado “Voltar para a *guideline*”.

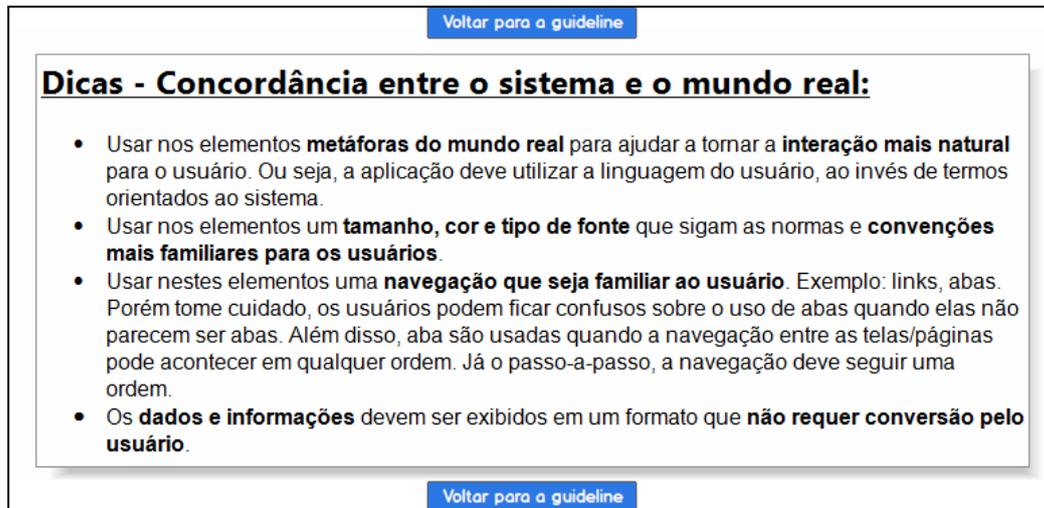


Figura 27. Dicas da Heurística 2 da UDG

## 5.4 Síntese do Capítulo

Este capítulo apresentou as tecnologias MIT 1, MIT 2, MIT 3, UDRT-AD e UDG. As tecnologias propostas possuem o objetivo de auxiliar no projeto e avaliação de modelos usados nas fases iniciais, visando à usabilidade da aplicação final. O próximo capítulo descreve a família de experimentos executados de acordo com a metodologia baseada em experimentação proposta por Shull *et al.*, (2001).

## CAPÍTULO 6 - ESTUDOS EXPERIMENTAIS DA MIT 1

*Este capítulo apresenta os estudos experimentais que foram realizados para avaliar empiricamente e aperfeiçoar a tecnologia MIT 1.*

### 6.1. Introdução

Estudos primários são necessários para criar, melhorar e avaliar processos, métodos e ferramentas para o desenvolvimento de *software* (BASILI *et al.*, 1996) e avaliação da qualidade do *software* (BOLCHINI e GARZOTTO, 2007). Realizar repetições destes estudos é necessário para conseguir uma maior validade dos resultados (SHULL *et al.*, 2008; KITCHENHAM, 2008). Diante disso, uma "família de experimentos" tem surgido como uma metodologia de pesquisa empírica com o objetivo de extrair conclusões significativas a partir de vários experimentos semelhantes que possuem o mesmo objetivo (FERNANDEZ *et al.*, 2013).

Em todos os estudos experimentais realizados no âmbito desta pesquisa foi seguido um processo de inspeção baseado no processo sugerido por Sauer *et al.* (2000). Este processo é dividido em quatro atividades que são apresentadas na Figura 28. As atividades e papéis do processo de inspeção serão descritas a seguir.

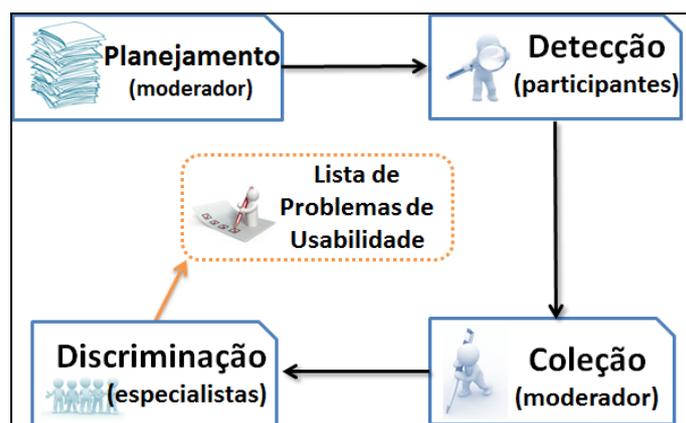


Figura 28. Processo de Inspeção baseado em Sauer *et al.* (2000)

A primeira atividade é o Planejamento da Inspeção. Nesta atividade é feita a definição do objetivo, contexto e hipóteses do experimento, seleção de variáveis, preparação do material, seleção de participantes, treinamento dos participantes nas técnicas e preparação de tarefas para os participantes. O líder da inspeção (ou moderador) realiza

esta atividade. Em nossos experimentos, uma pessoa que tem conhecimento e experiência em avaliação de usabilidade assumiu o papel de líder ou moderador da inspeção.

A segunda atividade do processo de inspeção baseado em Sauer *et al.* (2000) é a Detecção de Defeitos, onde cada inspetor procura por defeitos de usabilidade nos modelos, individualmente.

Após a detecção, o moderador realiza a atividade de Coleção, onde as listas de discrepâncias individuais são integradas em uma única lista, removendo-se o nome do inspetor que encontrou a discrepância e a técnica que ele aplicou. O nome de quem encontrou a discrepância e a técnica aplicada foram removidas para que os especialistas, na atividade de Discriminação (próxima atividade), não tendessem a classificar uma discrepância como defeito de usabilidade por ela ter sido encontrada por uma técnica específica ou por ter sido encontrada por um inspetor específico.

A última atividade do processo de inspeção é a Discriminação. Nesta atividade, o time formado por especialistas em usabilidade revisa a lista de discrepâncias, todos juntos em uma reunião de discriminação. Este time decide quais discrepâncias são únicas e quais são duplicadas. Discrepâncias duplicadas são discrepâncias apontadas por mais de um inspetor. Além disso, o time decide quais discrepâncias são defeitos reais e quais são falso-positivo (defeitos de usabilidade identificados considerados como “não real”).

Estudos experimentais foram realizados com a técnica MIT 1 a fim de avaliá-la e aperfeiçoá-la. Como pode ser visto na Figura 29, dois estudos de viabilidade foram realizados com a técnica MIT 1. O 1º estudo foi realizado com o intuito de comparar a MIT 1 com a Avaliação Heurística. E o 2º estudo foi realizado com o objetivo de comparar a MIT 1 com o método UCE. Estes dois estudos experimentais serão descritos em mais detalhes nas subseções a seguir, assim como as melhorias realizadas na MIT 1.

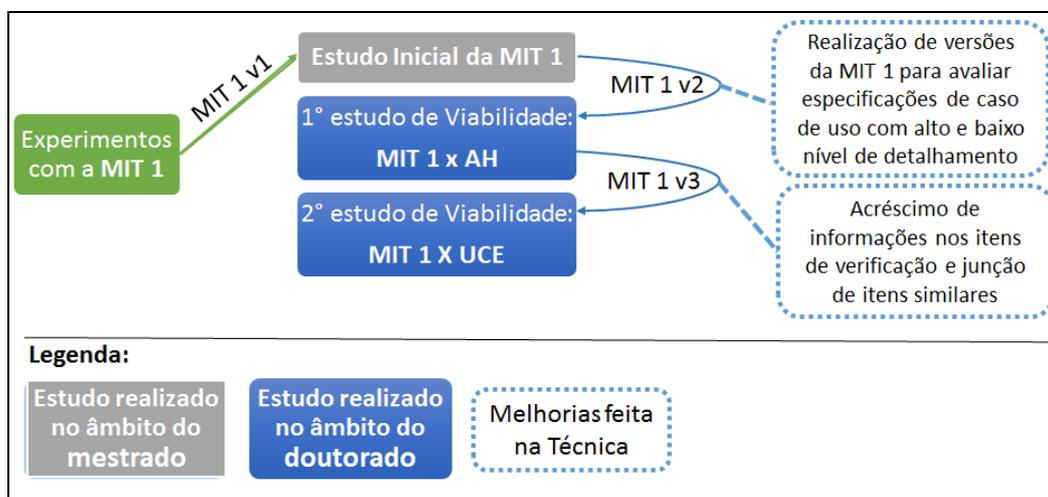


Figura 29. Estudos experimentais realizados com a MIT 1

## 6.2. 1º Estudo de Viabilidade da MIT 1

Segundo a metodologia experimental apresentada na Subseção 1.5, a primeira questão de pesquisa a ser respondida é “Os resultados são viáveis?”. Esta questão avalia se a nova tecnologia satisfaz o objetivo geral para o qual foi criada. Neste caso, o objetivo da MIT 1 é apoiar a detecção de defeitos de usabilidade em especificações de caso de uso. Para tecnologias que detectam defeitos de usabilidade, SHULL *et al.* (2001) recomendam a execução de um estudo de viabilidade, visto que este tipo de estudo permite testar a eficácia e eficiência da tecnologia.

Diante disso, a técnica MIT 1 foi avaliada através de um estudo de viabilidade (VALENTIM *et al.*, 2015a). As Subsubseções a seguir apresentam o detalhamento deste estudo de viabilidade, incluindo a execução das atividades do processo de inspeção, os resultados alcançados e as ameaças à validade.

### 6.2.1. Planejamento do 1º Estudo de Viabilidade da MIT 1

Com o objetivo de avaliar a MIT 1 realizou-se um estudo de viabilidade, usando somente a versão da técnica com alto nível de detalhamento. MIT 1 foi avaliada em comparação com o método Avaliação Heurística (AH) porque: (a) a construção da MIT 1 foi baseada na AH e assim comparar estas tecnologias é importante para verificar se a tecnologia derivada (MIT 1) é melhor que a tecnologia original (AH); e (b) os inspetores têm uma base de conhecimento sólida em princípios de usabilidade, o que lhes permite usar a AH.

#### a. Hipótese

O estudo foi planejado e conduzido a fim de testar as seguintes hipóteses (nula e alternativa, respectivamente):

- H01: Não há diferença entre as tecnologias MIT 1 e AH em relação ao indicador de eficiência.
- HA1: Há diferença entre as tecnologias MIT 1 e AH em relação ao indicador de eficiência.
- H02: Não há diferença entre as tecnologias MIT 1 e AH em relação ao indicador de eficácia.
- HA2: Há diferença entre as tecnologias MIT 1 e AH em relação ao indicador de eficácia.

## **b. Contexto**

Este estudo foi executado com um dos casos de uso de um sistema online que disponibiliza indicadores de pesquisa e desenvolvimento no Brasil (ver a especificação do caso de uso utilizado neste experimento em Valentim *et al.* (2017c)). O experimento foi conduzido com estudantes de graduação do curso de Ciência da Computação da Universidade Federal do Amazonas (UFAM). Os estudantes já tinham tido duas disciplinas introdutórias sobre “Engenharia de *Software*” e “Interação Humano-Computador” e estavam cursando a disciplina “Análise e Projeto de Sistemas” (2º Semestre de 2013).

## **c. Seleção de Variáveis**

As variáveis independentes foram as tecnologias de avaliação de usabilidade (MIT 1 e AH) e as variáveis dependentes foram os indicadores de eficiência e eficácia das tecnologias. Eficiência e eficácia foram calculadas para cada participante como: (a) a razão entre o número de defeitos detectados e o tempo gasto no processo de inspeção; e (b) a razão entre o número de defeitos detectados e o número total de defeitos existentes (conhecidos), respectivamente.

## **d. Seleção de Participantes**

Dezoito participantes assinaram um Termo de Consentimento Livre e Esclarecido (TCLE) e preencheram um Formulário de Caracterização que mediu sua experiência em Avaliação de Usabilidade (AU) e Desenvolvimento de *Software* (DS). O formulário de caracterização foi aplicado para categorizar os participantes como tendo: nenhuma, baixa, média ou alta experiência em AU e DS. Com relação à classificação da experiência em Avaliação de Usabilidade, considerou-se com:

- Alta experiência (A): participantes que tinham participado em mais de 5 projetos ou avaliação de usabilidade na indústria;
- Média Experiência (M): participantes que tinham participado entre 1 e 4 projetos ou avaliação de usabilidade na indústria;
- Baixa Experiência (B): participantes que participaram em pelo menos um projeto ou avaliação de usabilidade em sala de aula;
- Nenhuma Experiência (N): participantes que tinham nenhum conhecimento sobre usabilidade ou que conheciam alguns conceitos de usabilidade adquiridos em leituras/palestras mas nenhuma experiência prática.

De forma similar, a experiência dos participantes em Desenvolvimento de *Software* seguiu os mesmos padrões. Tabela 16 (segunda e terceira coluna) mostra a categorização de cada participante.

#### **e. Projeto Experimental**

Participantes foram divididos em dois grupos: o grupo da MIT 1 e o grupo da AH. Cada participante inspecionou o mesmo caso de uso. Os participantes foram designados para cada técnica de forma aleatória e balanceada, de acordo com o nível de experiência de cada participante. Cada grupo foi composto por 9 participantes.

#### **f. Instrumentação**

Muitos artefatos foram definidos para apoiar o experimento: formulário de categorização, Termo de Consentimento Livre e Esclarecido (TCLE), especificação das tecnologias MIT 1 e AH, instruções para a inspeção, uma planilha para a anotação das discrepâncias identificadas e um questionário pós-inspeção, que estão disponíveis em Valentim *et al.* (2017c). Além disso, foi usado um caso de uso que é parte da especificação de sistema real.

#### **g. Preparação**

Todos os participantes receberam treinamento de 2 horas em avaliação de usabilidade. Além disso, para cada grupo foi feita uma apresentação de 15 minutos sobre a tecnologia que o grupo aplicaria. Exemplos similares foram mostrados para os participantes sobre como usar ambas as tecnologias (MIT 1 e AH).

### **6.2.2. Execução do 1º Estudo de Viabilidade da MIT 1**

No começo do estudo, um pesquisador agiu como moderador sendo responsável por passar as informações da avaliação para os participantes. Em seguida, os participantes foram divididos em dois grupos que ficaram em salas diferentes (um grupo usou a AH e o outro grupo usou a MIT 1). Cada participante recebeu os artefatos descritos na Subsubseção 6.2.1 - f. Durante a inspeção, cada participante preencheu uma planilha com os defeitos encontrados. Todos os participantes devolveram uma planilha contendo os possíveis defeitos e o tempo total gasto na inspeção. Eles também entregaram o questionário pós-inspeção preenchido. Cada inspetor executou a atividade de detecção de defeitos individualmente. Durante a atividade de detecção, inspetores não receberam qualquer ajuda dos pesquisadores envolvidos no estudo e não conversaram entre si.

### 6.2.3. Coleção e Discriminação do 1º Estudo de Viabilidade da MIT 1

Após a execução, as listas de discrepâncias individuais foram integradas em uma única lista, removendo o nome do inspetor que encontrou a discrepância e a tecnologia que ele/ela tinha utilizado.

Um time formado por um engenheiro de *software* (o autor do caso de uso) e dois especialistas em usabilidade revisaram esta lista. Este time decidiu quais discrepâncias foram únicas e quais eram duplicatas (discrepâncias equivalentes apontadas por mais de um inspetor). Além disso, o time decidiu quais discrepâncias eram defeitos ou falso-positivo.

### 6.2.4. Resultados Quantitativos do 1º Estudo de Viabilidade da MIT 1

A Tabela 16 apresenta o resultado geral da avaliação de usabilidade em casos de uso do 1º Estudo de Viabilidade da MIT 1. O rótulo 'P' e um número identifica cada participante, por exemplo, P01 identifica o participante 01. Os participantes P01 a P09 aplicaram a MIT 1. E os participantes P10 a P18 aplicaram a AH. Pode-se observar que os inspetores que usaram a MIT 1 encontraram entre 9 e 19 defeitos utilizando cerca de 0,58 e 1,52 horas. Por outro lado, os inspetores que usaram a AH levaram entre 0,38 e 1 hora e encontraram entre 4 e 14 defeitos.

Tabela 16. Resultado da inspeção - 1º Estudo de Viabilidade da MIT 1

Part.	Experiência em AU	Experiência em DS	Discrepâncias	Falso-Positivo	Defeitos	Tempo (Hora)	Defeitos /Hora
P01	Baixa	Nenhuma	26	7	19	1,40	13,57
P02	Baixa	Nenhuma	12	1	11	0,97	11,38
P03	Média	Baixa	14	2	12	0,63	18,95
P04	Baixa	Nenhuma	14	1	13	1,02	12,79
P05	Baixa	Nenhuma	18	2	16	1,52	10,55
P06	Baixa	Baixa	12	0	12	0,58	20,57
P07	Baixa	Nenhuma	20	5	15	0,83	18,00
P08	Baixa	Nenhuma	14	2	12	0,92	13,90
P09	Nenhuma	Nenhuma	11	2	9	1,12	8,06
P10	Baixa	Nenhuma	12	2	10	0,42	24,00
P11	Baixa	Nenhuma	17	3	14	0,62	22,70
P12	Baixa	Nenhuma	11	3	8	1,00	8,00
P13	Baixa	Nenhuma	9	3	6	0,83	7,20
P14	Baixa	Nenhuma	4	0	4	0,50	8,00
P15	Baixa	Baixa	13	5	8	0,38	20,87
P16	Média	Nenhuma	8	0	8	0,45	17,78
P17	Baixa	Nenhuma	9	4	5	0,87	5,77
P18	Baixa	Nenhuma	9	2	7	0,50	14,00

**Legenda:** Part. – Participante. Experiência em AU. - Experiência em Avaliação de Usabilidade; Experiência em DS. - Experiência em Desenvolvimento de *Software*.

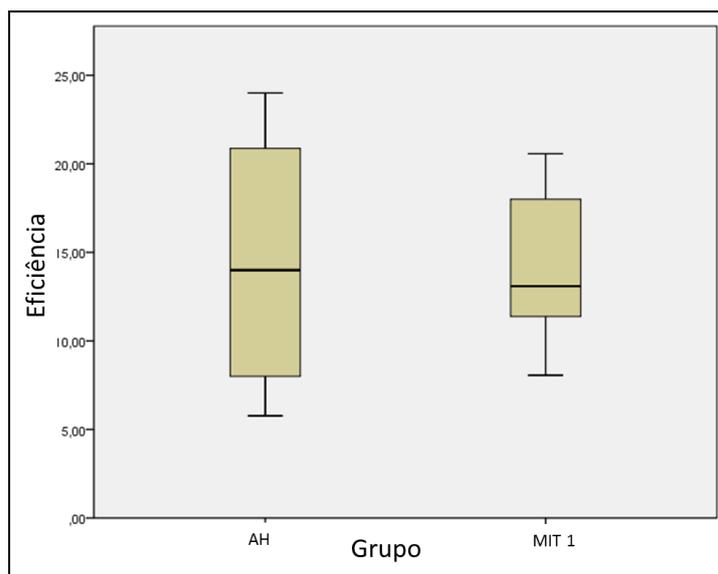
As inspeções resultaram em um conjunto de 40 defeitos diferentes de usabilidade, incluindo os 11 defeitos inseridos nos casos de uso. Defeitos foram incluídos porque havia a necessidade de ter mais defeitos encontrados, já que o caso de uso original possuía poucos defeitos de usabilidade. Tabela 17 apresenta a média de eficácia e eficiência.

**Tabela 17. Eficácia e Eficiência no 1º Estudo de Viabilidade da MIT 1**

Tecnologia	MIT 1	AH
Total de Defeitos	119	70
Média de Defeitos	13,22	7,78
Eficácia	33,06%	19,44%
Tempo Médio (min)	59,89	37,11
Eficiência (defeitos/hora)	13,25	12,57

Uma análise foi realizada usando o teste não-paramétrico de Mann-Whitney (MANN e WHITNEY, 1947), dado o tamanho limitado da amostra. Um resumo dos resultados é apresentado usando o gráfico de boxplot. A análise estatística foi realizada usando a ferramenta estatística SPSS V. 19, e  $\alpha = 0,10$ . A escolha desta significância estatística foi motivada pelo pequeno tamanho da amostra usada neste experimento. Figura 30 mostra o gráfico de boxplot com a distribuição da eficiência por técnica.

Na Figura 30 pode ser observado que o grupo da MIT 1 tem quase a mesma eficiência do grupo da AH. Quando se compara as duas amostras usando o teste de Mann-Whitney, não se encontra nenhuma diferença significativa entre os dois grupos ( $p = 0,895$ ). Estes resultados apoiam a hipótese nula  $H_0$  que afirma que não há diferença no indicador de eficiência entre a MIT 1 e AH.



**Figura 30. Boxplots da eficiência do 1º estudo de viabilidade da MIT 1**

A mesma análise foi aplicada para determinar se houve diferença estatística comparando o indicador de eficácia das duas técnicas na identificação de defeitos de usabilidade. O gráfico de boxplots com a distribuição da eficácia por técnica (ver Figura 31) mostra que o grupo da MIT 1 foi muito mais eficaz que o grupo da AH ao inspecionar a usabilidade do caso de uso. Além disso, a mediana do grupo da MIT 1 está muito mais alta que a mediana do grupo da AH, e todo boxplot do grupo da MIT 1 está acima do terceiro quartil do grupo da AH. O teste de Mann-Whitney confirmou que a eficácia da MIT 1 foi significativamente mais alta que a eficácia da AH ( $p = 0,002$ ). Estes resultados sugerem que a técnica MIT 1 foi mais eficaz que a AH quando usada para inspecionar a especificação do caso de uso deste estudo (apoia a hipótese HA2).

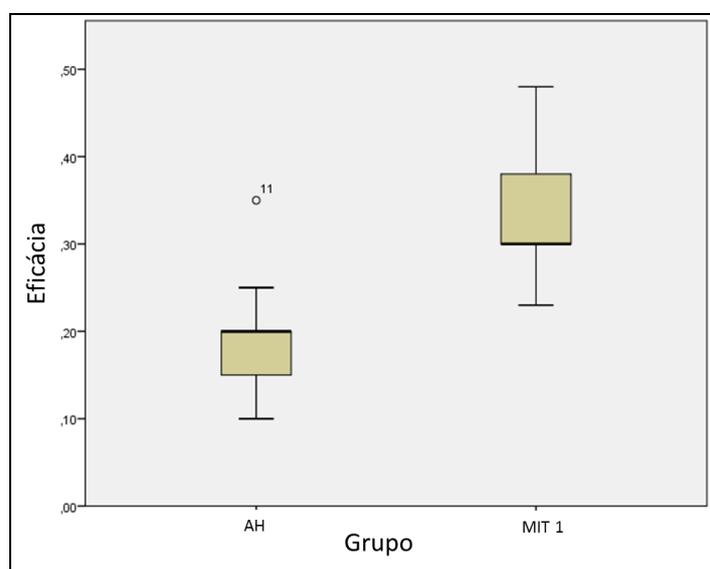


Figura 31. Boxplots da eficácia do 1º estudo de viabilidade da MIT 1

### 6.2.5. Análise da Percepção dos Participantes do 1º Estudo de Viabilidade da MIT 1

Após a análise quantitativa, os questionários pós-inspeção foram analisados. Foram definidos os seguintes indicadores: (i) Facilidade de Uso Percebida, que define o grau em que uma pessoa acredita que usar uma tecnologia específica seria livre de esforço, e (ii) Utilidade Percebida, que define o grau em que uma pessoa acredita que a tecnologia poderia melhorar seu desempenho. Estes indicadores são baseados no Modelo de Aceitação de Tecnologia (*Technology Acceptance Model - TAM*) (DAVIS, 1989).

Participantes forneceram suas respostas em uma escala de seis pontos, baseado no questionário aplicado por Lanubile *et al.* (2003). As possíveis respostas foram: concordo totalmente, concordo amplamente, concordo parcialmente, discordo parcialmente, discordo amplamente e discordo totalmente. Esta escala de respostas foi considerada apropriada porque não há valor intermediário, ou seja, ela ajuda a evitar o viés da tendência

central em classificações, forçando os participantes a julgar o resultado como adequado ou não adequado (GARLAND, 1991; JOHNS, 2005; CALEFATO, *et al.*, 2010). Neste questionário, os inspetores responderam seu grau de aceitação sobre as sentenças relacionadas a Facilidade de Uso e Utilizada Percebida.

### a. Facilidade de Uso Percebida

A Figura 32 apresenta as percepções dos participantes com relação à facilidade de uso da MIT 1. Não coletamos e analisamos as percepções dos participantes do grupo da AH, pois nosso intuito nesse estudo era apenas analisar as percepções dos participantes do grupo da MIT 1 para podermos evoluir a técnica. O eixo X do gráfico da Figura 32 se refere às possíveis respostas do questionário pós-inspeção e o eixo Y se refere ao número de participantes. P01 e outros códigos representam os participantes na Tabela 16.

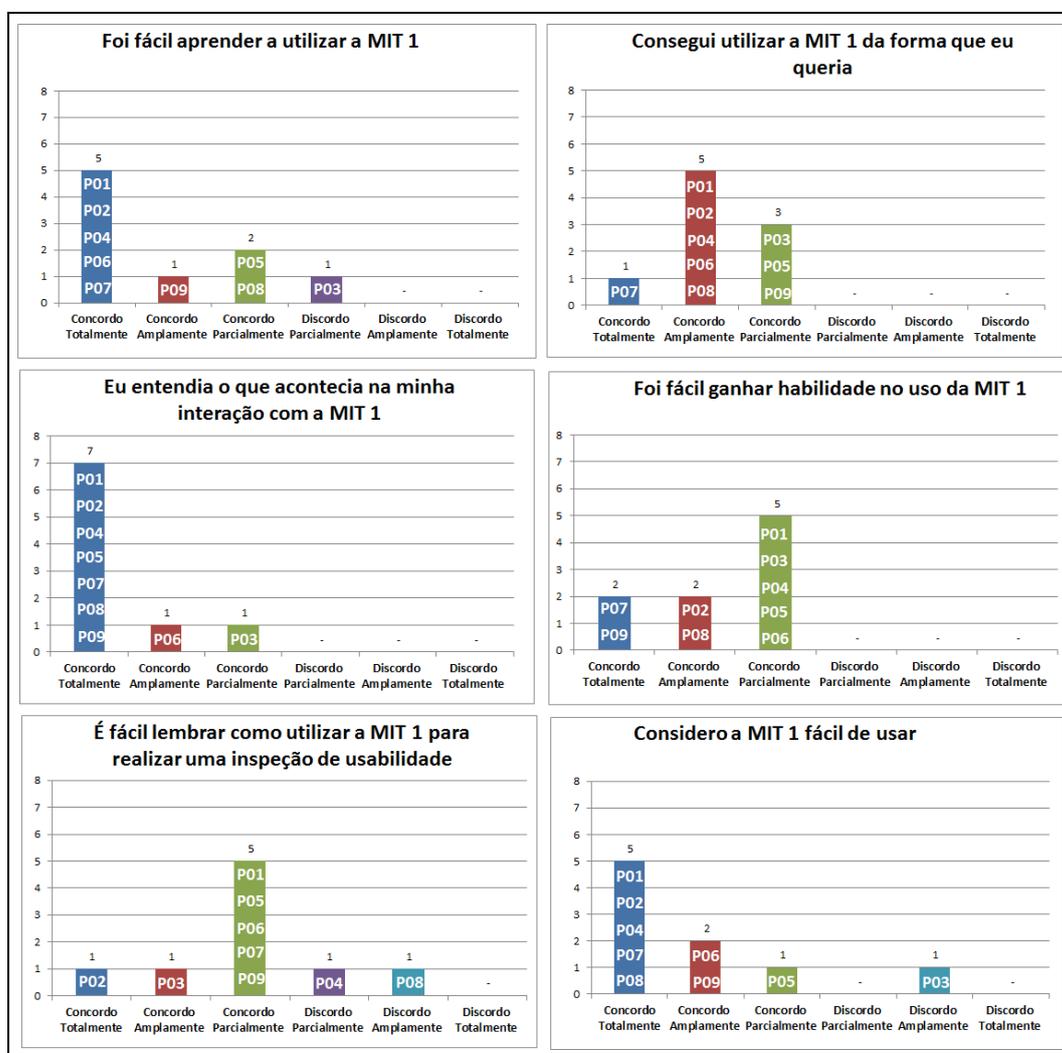


Figura 32. Percepção sobre Facilidade de Uso - 1º estudo de viabilidade MIT 1

A seguir serão analisadas somente as respostas referente à MIT 1, pois deseja-se avaliar a aceitabilidade desta técnica. Pode-se observar na Figura 32 que 7 dos 9

participantes concordaram totalmente com a sentença “*Eu entendia o que acontecia em minha interação com a MIT 1*”, uma vez que os participantes entendiam o que estava acontecendo quando eles usavam a MIT 1.

Outra sentença onde mais da metade dos participantes concordaram totalmente foi “*Foi fácil aprender a utilizar a MIT 1*”, mostrando que os participantes não tiveram muita dificuldade em aprender a usar a MIT 1. Pode-se notar que os participantes P05 e P08 concordaram parcialmente com a sentença “*Foi fácil aprender a utilizar a MIT 1*”, mostrando que a MIT 1 não é tão fácil de aprender.

Dois dos nove participantes (P04 e P08) discordaram com a sentença “*É fácil lembrar como utilizar a MIT 1 para realizar uma inspeção de usabilidade*”, mostrando que a MIT 1 não é tão fácil de lembrar. O participante P03 discordou amplamente da sentença “*Considero a MIT 1 fácil de usar*”, destacando a dificuldade que ele teve quando usou a MIT 1. No entanto, todos os inspetores concordaram com as outras sentenças, mostrando sua aceitação com relação à técnica MIT 1.

## b. Utilidade Percebida

A Figura 33 apresenta a percepção dos participantes sobre a utilidade da MIT 1.

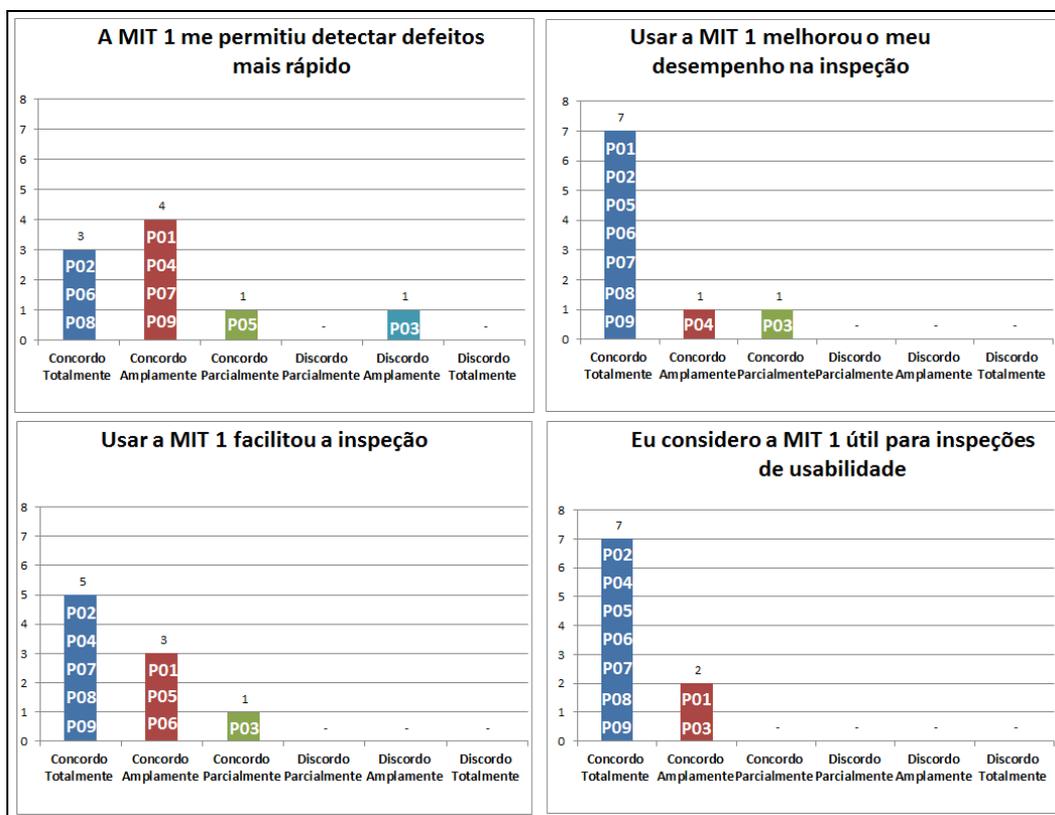


Figura 33. Percepção sobre a Utilidade - 1º estudo de viabilidade da MIT 1

Verificou-se que somente o participante P03 discordou parcialmente da sentença “A MIT 1 me permitiu detectar defeitos mais rápido”, indicando que em algum ponto este participante achou que o uso da técnica tornou o processo de inspeção mais demorado. No entanto, todos os inspetores concordaram com as demais sentenças, reforçando que a MIT 1 ajuda no processo de inspeção.

#### **6.2.6. Resultados Qualitativos do 1º Estudo de Viabilidade da MIT 1 e Melhorias**

Uma análise específica dos dados qualitativos (comentários adicionais dos inspetores) contido nos questionários foi realizada. Esta análise foi feita usando os procedimentos do método *Grounded Theory* (GT) (CORBIN e STRAUSS, 2008).

Os dados qualitativos que foram extraídos dos questionários pós-inspeção foram analisados usando um subconjunto de fases do processo de codificação sugerido por Corbin e Strauss (2008) para o método GT: codificação aberta (primeira fase) e axial (segunda fase). Para analisar os dados qualitativos, foram criados códigos (conceitos importantes para entendimento da percepção sobre a técnica e seu processo de aplicação) relacionados aos comentários dos participantes – codificação aberta (primeira fase). Após isso, os códigos foram agrupados de acordo com suas propriedades, formando conceitos que representam categorias e subcategorias. Por fim, os códigos foram relacionados entre si – codificação axial (segunda fase). Os procedimentos de GT objetivaram alcançar uma análise mais profunda, comparando e analisando o relacionamento entre estes conceitos. O propósito da análise neste estudo foi entender a percepção dos inspetores através de sua experiência usando a MIT 1. Uma vez que a intenção deste estudo não é criar uma teoria, não foi realizada a codificação seletiva (terceira fase do método GT). As fases de codificação aberta e axial foram suficientes para entender as causas de alguns problemas na aplicação da MIT 1. Os conceitos relacionados ao método GT são apresentados em detalhes em Conte *et al.* (2009b).

A seguir serão discutidas as implicações sobre os resultados qualitativos da MIT 1.

##### **a. Comentários sobre a Facilidade de Uso da MIT 1 e Melhorias**

Algumas das dificuldades que foram coletadas neste experimento quando se usou a MIT 1 foram: há heurísticas que foram aplicadas somente em algumas partes da interface representada no caso de uso (ver citação de P01 abaixo); há um grande número de itens de verificação para avaliar a especificação de caso de uso (ver citação de P02 abaixo); é difícil aplicar a MIT 1 (ver citação de P03 abaixo); e é difícil lembrar os itens da técnica (ver citação de P06 abaixo).

“(…) há heurísticas que se aplicam somente a algumas [partes da interface representada no caso de uso] (...)”. (Participante 1).

“São muitos subitens para avaliar, [o] que pode deixar meio confuso (...)”. (Participante 2).

“MIT 1 é uma técnica útil porém difícil de aplicar”. (Participante 3).

“São muitos os tipos de [itens] para serem lembrados, mas com a prática talvez dê para lembrar mais” (Participante 6).

Pode-se observar que houve algumas dificuldades e uma delas precisa ser destacada: há heurísticas que só são aplicadas a algumas partes da interface representadas no caso de uso. Isto aconteceu no item de verificação 1AD1, pois o mesmo pede para verificar se há um problema somente no nome dos botões ou links. Para que este item de verificação levasse em consideração outras partes da interface, foram acrescentados outros dois termos: “campos” e “telas” (ver Figura 34).

<b>Antes</b>	<b>Item de Verificação MIT-1AD1</b>	Verifique se todos os botões e links possuem nomes que indicam funcionalidades diferentes no Fluxo Principal, Fluxo Alternativo, Fluxo de Exceção e Regras de Negócio, ou seja, não tenham nomes diferentes, mas com a mesma funcionalidade.
<b>Depois</b>	<b>Item de Verificação MIT-1AD1</b>	Verifique se todos os <b>campos, telas</b> , botões e links possuem nomes que indicam funcionalidades diferentes no Fluxo Principal, Fluxo Alternativo, Fluxo de Exceção e Regras de Negócio, ou seja, não tenham nomes diferentes, mas com a mesma funcionalidade.

Figura 34. Item de Verificação MIT-1AD1

Da mesma forma foram acrescentados os termos: “campo”, “opção” e “tela” no item de verificação 1AI1 (ver Figura 35). Com estas melhorias, buscou-se deixar a técnica mais completa, considerando outras partes da interface que podem ser descritas na especificação de caso de uso.

<b>Antes</b>	<b>Item de Verificação MIT-1AI1</b>	Verifique se todo botão ou link que seja essencial está presente no Fluxo Principal, Fluxo Alternativo, Fluxo de Exceção e Regras de Negócio.
<b>Depois</b>	<b>Item de Verificação MIT-1AI1</b>	Verifique se todo <b>campo, opção, tela</b> , botão ou link que seja essencial está presente no Fluxo Principal, Fluxo Alternativo, Fluxo de Exceção e Regras de Negócio.

Figura 35. Item de Verificação MIT-1AI1

Os participantes também tiveram dificuldades em aplicar a técnica. Isso foi visto tanto através do indicador Facilidade de Uso quanto nas respostas ao questionário pós-inspeção. O objetivo da MIT 1 é que o inspetor utilize a técnica para o auxiliar a encontrar os problemas de usabilidade no caso de uso. No entanto, se o inspetor tem dificuldade de utilizá-la, a técnica ainda precisa ser melhorada neste sentido para que alcance seu objetivo.

#### **b. Comentários sobre a Estrutura da MIT 1 e Melhorias**

Durante este experimento algumas inadequações na estrutura da técnica foram coletadas, como: há itens de verificação semelhantes (ver citação de P09 abaixo) e há muitas heurísticas que não são utilizadas (ver citação de P05 abaixo). Além disso, foi dada uma sugestão de agrupar algumas heurísticas (ver citação de P01 abaixo).

*“(...) há algumas confusões ao compreender quais afirmações realmente especificar (...), por conter especificações semelhantes (...)”* (Participante 9).

*“(...) muitas heurísticas para tratar de diferentes tipos de discrepâncias e algumas vezes nem todas são utilizadas. (...)”* (Participante 5).

*“[As] heurísticas poderiam ser agrupados de forma a ajudar na localização utilizando o campo”.* (Participante 1).

Nota-se que houve citações de algumas inadequações na estrutura da técnica, como por exemplo: alguns itens forem identificados como similares na técnica. Isto provavelmente gerou dúvidas durante a utilização da MIT 1. Melhorias neste sentido foram feitas na MIT 1 e uma nova versão da técnica foi criada.

### **6.3. 2º Estudo de Viabilidade da MIT 1**

Com o objetivo de verificar se a versão 2 da técnica MIT 1 (MIT 1 v2) é viável e se esta técnica possui um desempenho melhor em comparação ao UCE, um método também proposto para avaliar a usabilidade através de especificações de caso de uso, realizou-se um novo estudo de viabilidade (VALENTIM *et al.*, 2015c). As Subsubseções a seguir apresentam o detalhamento deste estudo de viabilidade.

#### **6.3.1. Planejamento do 2º Estudo de Viabilidade da MIT 1**

Este experimento objetivou validar empiricamente a Eficácia, Eficiência, Facilidade de Uso Percebida e Utilidade Percebida da MIT 1 quando comparada ao método *Use Case Evaluation* (UCE) (HORNBAEK *et al.*, 2007). O método UCE foi escolhido para ser comparado com a MIT 1, pois este foi o único método encontrado na literatura que tem o mesmo objetivo da MIT 1, avaliar a usabilidade através das especificações de caso de uso.

Como mostrado anteriormente na Tabela 3, o UCE possui uma lista de *guidelines* para apoiar a inspeção de usabilidade em especificações de caso de uso.

#### **a. Hipótese**

O experimento foi planejado e conduzido a fim de testar as seguintes hipóteses (nula e alternativa, respectivamente):

- H01: Não há diferença entre as técnicas MIT 1 e UCE com relação ao indicador de eficiência.
- HA1: Há diferença entre as técnicas MIT 1 e UCE com relação ao indicador de eficiência.
- H02: Não há diferença entre as técnicas MIT 1 e UCE com relação ao indicador de eficácia.
- HA2: Há diferença entre as técnicas MIT 1 e UCE com relação ao indicador de eficácia.

#### **b. Contexto**

O experimento foi executado com o mesmo caso de uso utilizado no experimento anterior. O experimento foi conduzido com estudantes de graduação do quarto ano de Ciência da Computação e estudantes de Mestrado e Doutorado da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).

#### **c. Seleção de Variáveis**

As variáveis independentes foram as técnicas de avaliação de usabilidade (MIT 1 e UCE) e as variáveis dependentes foram indicadores de eficiência e eficácia das técnicas, calculadas da mesma forma que no estudo anterior.

#### **d. Seleção de Participantes**

Quarenta e oito estudantes (de 54) consentiram em participar do estudo. Eles assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) e preencheram um formulário de caracterização que mediu sua experiência com Avaliação de Usabilidade (AU) e Desenvolvimento de *Software* (DS). O formulário de caracterização foi aplicado para categorizar os participantes como tendo: nenhuma, baixa, média ou alta experiência em AU e DS, seguindo a mesma categorização dos estudos anteriores (ver Subsubseção 6.2.2 - f). De forma análoga, a experiência dos participantes em desenvolvimento de *software* foi classificada seguindo o mesmo padrão. Na Tabela 18, a segunda coluna (AU) e a terceira

coluna (DS) apresentam a categorização de cada participante com relação às suas experiências em AU e DS.

Em relação à participação de estudantes de graduação, todos receberam quatro pontos em uma das avaliações (equivalente a 0,5 ponto da nota final do curso) independente de seu desempenho, uma vez que todos os estudantes são de tempo parcial. Com relação à participação de estudantes de pós-graduação, a participação foi voluntária e o experimento foi realizado durante o horário de aula.

#### **e. Projeto Experimental**

Os participantes foram divididos em dois grupos, onde inspecionariam o mesmo caso de uso: grupo da MIT 1 e grupo do UCE. Os participantes foram designados para cada grupo de maneira aleatória e balanceada, de acordo com o nível de experiência de cada participante. Cada grupo foi composto por 24 participantes. No entanto, 3 participantes alocados para o grupo UCE não compareceram no dia do experimento.

#### **f. Instrumentação**

Alguns artefatos também foram definidos para apoiar o experimento: formulário de caracterização e consentimento, especificação das técnicas UCE e MIT 1, instruções para o experimento, uma planilha de anotação das discrepâncias identificadas e um questionário pós-inspeção, disponíveis em Valentim *et al.* (2017c). Além disso, foi utilizado o mesmo caso de uso do experimento anterior. Todos os artefatos foram validados antes do estudo ser executado por um pesquisador especialista em IHC e ES, para que não houvesse nenhum problema que inviabilizasse a execução do estudo.

#### **g. Preparação**

Todos os participantes receberam uma hora de treinamento em avaliação de usabilidade. O material do treinamento incluiu um conjunto de slides contendo uma introdução sobre usabilidade a fim de apresentar os conceitos básicos. Além disso, para cada grupo, foi feita uma apresentação de 15 minutos sobre a técnica que o grupo aplicaria. Exemplos similares foram apresentados a fim de mostrar como usar ambas as técnicas.

### **6.3.2. Detecção do 2º Estudo de Viabilidade da MIT 1**

No início do experimento, um pesquisador agiu como moderador da inspeção, sendo responsável por passar as informações sobre a avaliação aos inspetores, os estudantes. Após isso, dividiu-se os participantes em dois grupos para usar cada uma das

técnicas. Cada grupo fez a avaliação em salas diferentes supervisionados por outros dois pesquisadores. Cada participante recebeu os artefatos do experimento. Durante a inspeção, cada participante preencheu uma planilha com as discrepâncias (possíveis defeitos) identificadas. Todos os participantes entregaram no final da inspeção uma planilha contendo as discrepâncias identificadas e a anotação do tempo total da inspeção. Eles também preencheram um questionário pós-inspeção. Cada inspetor realizou a atividade de detecção de defeitos individualmente. Durante a atividade de detecção, os estudantes não receberam ajuda dos pesquisadores envolvidos no experimento.

### **6.3.3. Coleção e Discriminação do 2º Estudo de Viabilidade da MIT 1**

As atividades de Coleção e Discriminação foram realizadas conforme apresentado na Subseção 6.1. Na atividade de Discriminação o time que verificou a lista de discrepâncias foi formado por três especialistas em usabilidade.

### **6.3.4. Resultados Quantitativos do 2º Estudo de Viabilidade da MIT 1**

A Tabela 18 mostra o resultado geral da avaliação de usabilidade em casos de uso por participante do 2º Estudo de Viabilidade da MIT 1. Os participantes P01 a P24 aplicaram a MIT 1. E os participantes P25 a P44 aplicaram o UCE. Pode-se observar que os inspetores que usaram a MIT 1 encontraram entre 7 e 21 defeitos entre 0,32 e 1,33 horas. Por outro lado, os inspetores que usaram o UCE utilizaram a técnica entre 0,25 e 0,78 horas, no entanto eles identificaram de 4 a 16 defeitos. Pode-se notar que a MIT 1 ajudou a identificar mais defeitos que o UCE. No entanto, os inspetores levaram mais tempo usando a MIT 1 do que os inspetores que usaram o UCE.

No geral, as inspeções resultaram em um conjunto de 113 defeitos (conhecidos), incluindo os 11 defeitos inseridos pelo moderador. Tabela 19 apresenta a média de eficácia e eficiência. A eficácia da MIT 1 neste experimento foi de 11,47%. Comparando esta medida com a eficácia do grupo do UCE (7,12%), pode-se notar que esta medida foi mais alta. Além disso, pode-se observar na Tabela 19 que o UCE tende a fornecer um baixo número de falso positivos, total de 31 falso positivos comparado a 53 falso positivos da MIT 1. O baixo número de falso positivos pode ser explicado pelo fato que o UCE tem um procedimento mais simples (conteúdo mais claro) para detectar defeitos de usabilidade. No entanto, pode-se observar que a MIT 1 apoiou na identificação de mais defeitos de usabilidade (311 defeitos incluindo as duplicatas) que o UCE (169 defeitos incluindo as duplicatas). O alto grau de defeitos de usabilidade identificados com a MIT 1 pode indicar que a MIT 1 guia mais os participantes na atividade de detecção de defeitos de usabilidade.

Um resumo dos resultados quantitativos por indicador foi apresentado usando um gráfico de boxplot. Neste estudo, a análise estatística foi realizada usando a ferramenta estatística SPSS V. 22, e  $\alpha = 0,05$ . A Figura 36 mostra o gráfico de boxplot com a distribuição da eficiência por técnica.

Tabela 18. Resultado da inspeção - 2º Estudo de Viabilidade da MIT 1

Part.	Experiência em AU	Experiência em DS	Nº de Discrepância	Nº falso Positivo	Nº de Defeitos	Tempo (Hora)	Defeitos /Hora
P01	Baixa	Nenhuma	12	0	12	0,78	15,32
P02	Baixa	Baixa	14	0	14	0,92	15,27
P03	Baixa	Alta	9	0	9	1,17	7,71
P04	Baixa	Baixa	14	0	14	0,78	17,87
P05	Baixa	Baixa	22	2	20	0,75	26,67
P06	Média	Média	19	4	15	0,70	21,43
P07	Nenhuma	Alta	24	17	7	0,88	7,92
P08	Nenhuma	Nenhuma	24	9	15	1,33	11,25
P09	Baixa	Média	10	2	8	0,50	16,00
P10	Baixa	Alta	17	3	14	0,83	16,80
P11	Baixa	Baixa	25	9	16	0,50	32,00
P12	Baixa	Média	21	1	20	0,52	38,71
P13	Nenhuma	Baixa	12	1	11	0,63	17,37
P14	Baixa	Média	12	0	12	0,67	18,00
P15	Nenhuma	Média	10	0	10	0,67	15,00
P16	Baixa	Média	8	0	8	0,58	13,71
P17	Baixa	Alta	8	0	8	0,32	25,26
P18	Nenhuma	Média	19	1	18	0,58	30,86
P19	Nenhuma	Alta	7	3	4	0,67	6,00
P20	Baixa	Baixa	15	0	15	0,75	20,00
P21	Baixa	Média	11	0	11	0,50	22,00
P22	Baixa	Média	21	0	21	0,75	28,00
P23	Baixa	Média	16	0	16	0,73	21,82
P24	Nenhuma	Média	14	1	13	0,57	22,94
P25	Média	Média	9	3	6	0,67	9,00
P26	Baixa	Média	14	0	14	0,75	18,67
P27	Nenhuma	Alta	6	0	6	0,35	17,14
P28	Nenhuma	Alta	15	1	14	0,50	28,00
P29	Nenhuma	Média	10	1	9	0,48	18,62
P30	Alta	Alta	17	1	16	0,78	20,43
P31	Baixa	Média	5	0	5	0,67	7,50
P32	Baixa	Média	11	2	9	0,75	12,00
P33	Nenhuma	Média	11	1	10	0,50	20,00
P34	Nenhuma	Nenhuma	7	1	6	0,50	12,00
P35	Baixa	Média	12	2	10	0,48	20,69
P36	Nenhuma	Média	8	2	6	0,52	11,61
P37	Nenhuma	Baixa	8	3	5	0,78	6,38
P38	Baixa	Baixa	7	2	5	0,63	7,89
P39	Média	Média	5	1	4	0,27	15,00
P40	Baixa	Baixa	12	3	9	0,58	15,43
P41	Baixa	Média	11	5	6	0,60	10,00
P42	Nenhuma	Média	10	0	10	0,75	13,33
P43	Baixa	Nenhuma	8	3	5	0,50	10,00
P44	Baixa	Média	8	0	8	0,67	12,00
P45	Baixa	Média	8	0	6	0,25	24,00

Legenda: Part. – Participante; Experiência em AU. - Experiência em Avaliação de Usabilidade; Experiência em DS. - Experiência em Desenvolvimento de *Software*.

Tabela 19. Eficácia e Eficiência - 2º Estudo de Viabilidade da MIT 1

Tecnologia	MIT 1	UCE
Total de Defeitos	311	169
Média de Defeitos	12,96	8,05
Total de Falso Positivos	53	31
Eficácia	11,47%	7,12%
Tempo Médio (min)	17,08	11,98
Eficiência (defeitos/hora)	18,20	14,10

Na Figura 36, pode-se observar que o grupo da MIT 1 tem quase a mesma eficiência que o grupo do UCE. O número 12 na Figura 36 representa o participante que tem o melhor desempenho neste indicador relacionado ao grupo da MIT 1. A fim de determinar se este resultado é ou não significativo, aplicou-se o teste paramétrico t-test (JURISTO e MORENO, 2001) para amostras independentes para verificar a hipótese H01, uma vez que a Eficiência foi normalmente distribuída ( $p\text{-value} = 0,147$ ). Quando se compara as duas amostras usando o *t-test*, encontra-se diferença significativa entre os dois grupos ( $p\text{-value} = 0,030$ ). Estes resultados apoiam a rejeição da hipótese nula H01 ( $p\text{-value} < 0,05$ ), e a aceitação de sua hipótese alternativa HA1, sugerindo que a técnica MIT 1 foi mais eficiente que o UCE quando usado para inspecionar a especificação do caso de uso.

Vale lembrar que no 1º estudo de viabilidade, a MIT 1 apresentou uma eficiência melhor que a AH. No entanto, não houve diferença significativa neste indicador entre os dois grupos. Já neste 2º estudo de viabilidade houve uma diferença significativa com relação a eficiência, mostrando que a MIT 1 foi mais eficiente que o UCE. Isto pode ter ocorrido, pois buscou-se melhorar a MIT 1 a fim de que o processo de inspeção com esta técnica se torne mais rápido.

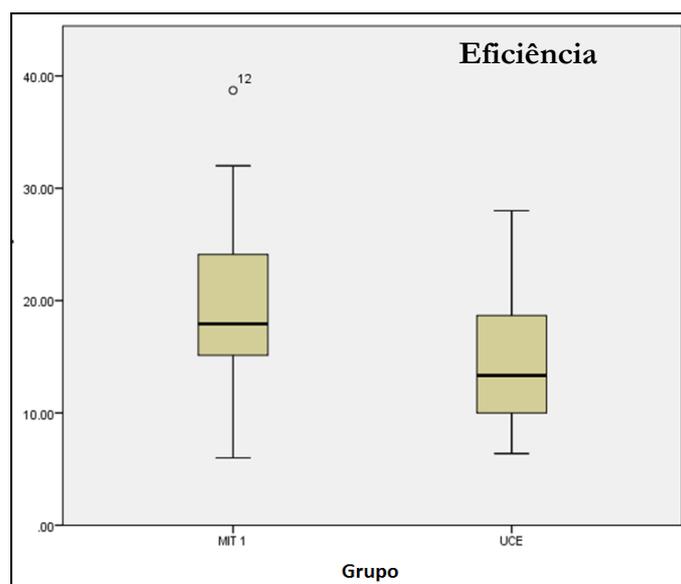


Figura 36. Boxplots da eficiência do 2º estudo de viabilidade da MIT 1

Outra análise foi realizada para determinar se houve diferença significativa comparando o indicador de eficácia das duas técnicas na detecção de defeitos de usabilidade. O gráfico de boxplots com a distribuição da eficácia por técnica (ver Figura 37) sugere que o grupo da MIT 1 foi muito mais eficaz que o grupo do UCE durante a inspeção de usabilidade do caso de uso. Além disso, a mediana do grupo da MIT 1 está muito mais alta que a mediana do grupo do UCE, e todo boxplot do grupo da MIT 1 está acima do terceiro quartil do boxplot do grupo do UCE. O número 30 na Figura 37 representa os participantes que tiveram o melhor desempenho neste indicador relacionado ao grupo do UCE. A fim de determinar se o resultado é ou não significativo, aplicou-se o teste de Mann-Whitney (MANN e WHITNEY, 1947) para verificar a H02, uma vez que a Eficácia não foi normalmente distribuída ( $p\text{-value} = 0,034$ ). O  $p\text{-value}$  obtido no teste de Mann-Whitney foi menor que 0,001. Este resultado, portanto, apoia a rejeição da hipótese nula H02 ( $p\text{-value} < 0,05$ ), e a aceitação de sua hipótese alternativa HA2, sugerindo que a técnica MIT 1 foi mais eficaz que o UCE quando usado para inspecionar a especificação de caso de uso neste experimento.

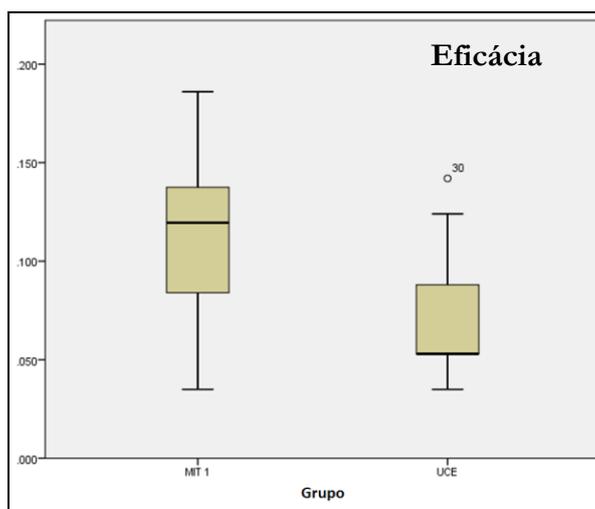


Figura 37. Boxplots da eficácia do 2º estudo de viabilidade da MIT 1

Uma análise de correlação foi realizada com as variáveis eficácia e eficiência para determinar o relacionamento entre elas. O coeficiente de correlação de Spearman entre eficácia e eficiência foi 0,726, com  $p\text{-value} < 0,001$ , mostrando correlação positiva. Portanto, participantes mais eficazes também foram mais eficientes.

### 6.3.5. Análise da Percepção dos Participantes do 2º Estudo de Viabilidade - MIT 1.

Como indicador por Hornbæk (2010), para avaliar a qualidade dos métodos de avaliação de usabilidade, não somente a contagem dos defeitos de usabilidade deve ser considerada, mas também a satisfação do participante em relação às tecnologias utilizadas.

Portanto, após a análise quantitativa, os questionários pós-inspeção sobre a aceitação das técnicas MIT 1 e UCE foram analisados. Os indicadores usados foram: (i) Facilidade de Uso Percebida; e (ii) Utilidade Percebida. Os resultados destes indicadores serão apresentados a seguir.

**a. Facilidade de Uso Percebida**

A Figura 38 apresenta as percepções dos participantes com relação à Facilidade de Uso do UCE e MIT 1. O eixo X do gráfico da Figura 38 se refere às possíveis respostas do questionário pós-inspeção e o eixo Y se refere ao número de participantes. P01, P02 e outros códigos representam os participantes apresentados na Tabela 18.

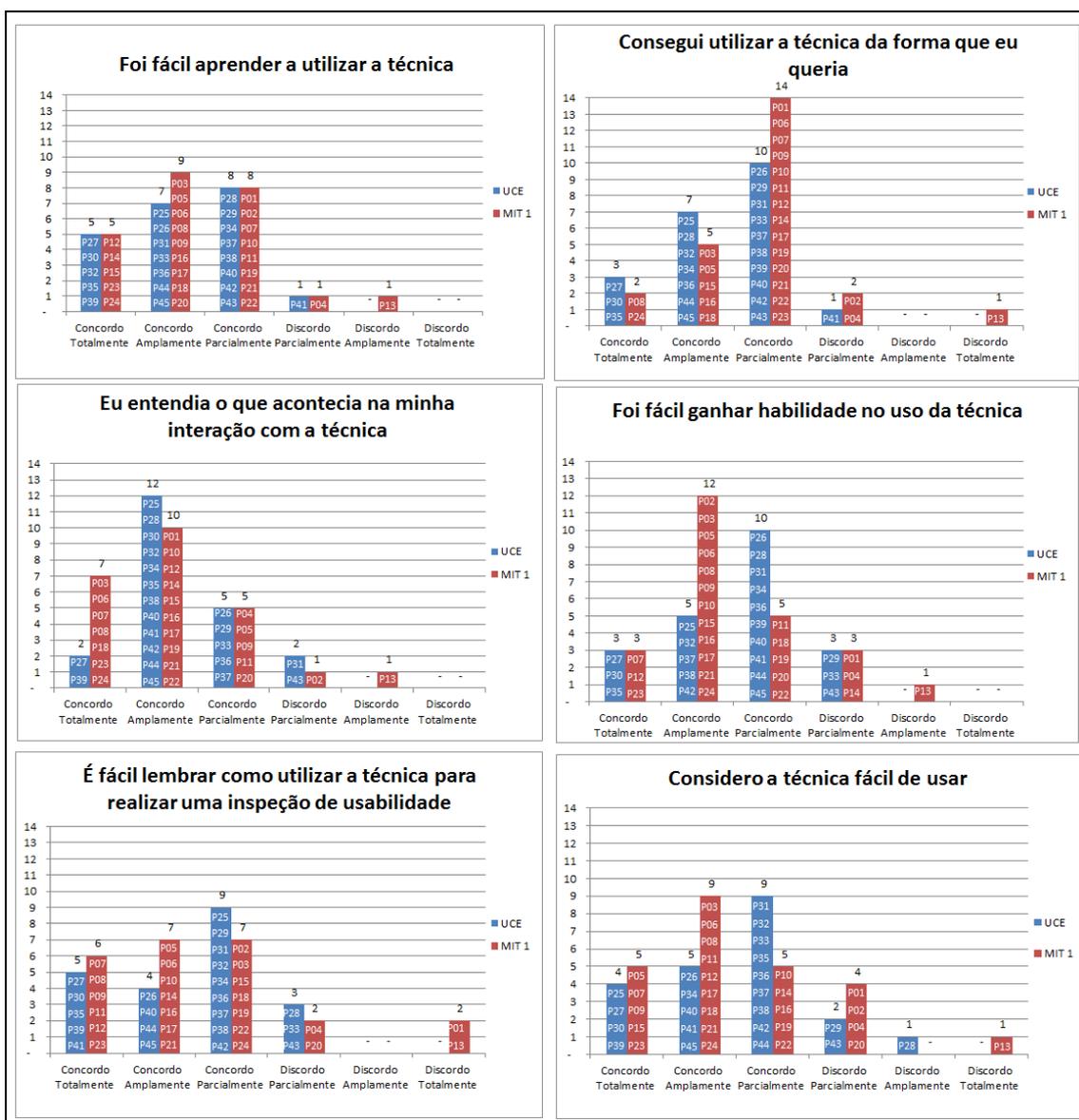


Figura 38. Percepção sobre Facilidade de Uso - 2º estudo de viabilidade MIT 1

Pode-se notar que o participante P13 concordou totalmente com a sentença “*Consegui utilizar a técnica da forma que eu queria*”. O participante P13 foi classificado com nenhuma experiência em avaliação de usabilidade (ver Tabela 18). Além disso, o mesmo participante discordou totalmente na sentença “*Considero a técnica fácil de usar*”. Isto pode indicar que a técnica MIT 1 deve ser melhorada a fim de que se torne mais fácil de usar.

A sentença “*É fácil lembrar como utilizar a técnica para realizar uma inspeção de usabilidade*” obteve discordâncias tanto por parte dos participantes da MIT 1 (P01, P04, P13, P20) quanto dos participantes do UCE (P28, P33, P43), mostrando que não é tão fácil lembrar de ambas as técnicas. Vale ressaltar que a maioria dos participantes tem baixa experiência em usabilidade; portanto, é necessário que as técnicas minimizem a carga de memória dos participantes.

### b. Utilidade Percebida

A Figura 39 apresenta as percepções dos participantes com relação à Utilidade das técnicas MIT 1 e UCE. Pode-se verificar que o participante P02 discordou totalmente das sentenças “*A técnica me permitiu detectar defeitos mais rápido*” e “*Usar a técnica melhorou meu desempenho na inspeção*” relativas à MIT 1. O participante P02 tinha baixa experiência em usabilidade. Ele levou cerca de 0,92 horas na inspeção, encontrando cerca de 14 defeitos de usabilidade. O desempenho do participante P02 bem como o tempo gasto na inspeção foi relativamente razoável, mas para o participante, a técnica deve ajudar a encontrar defeitos mais rápido.

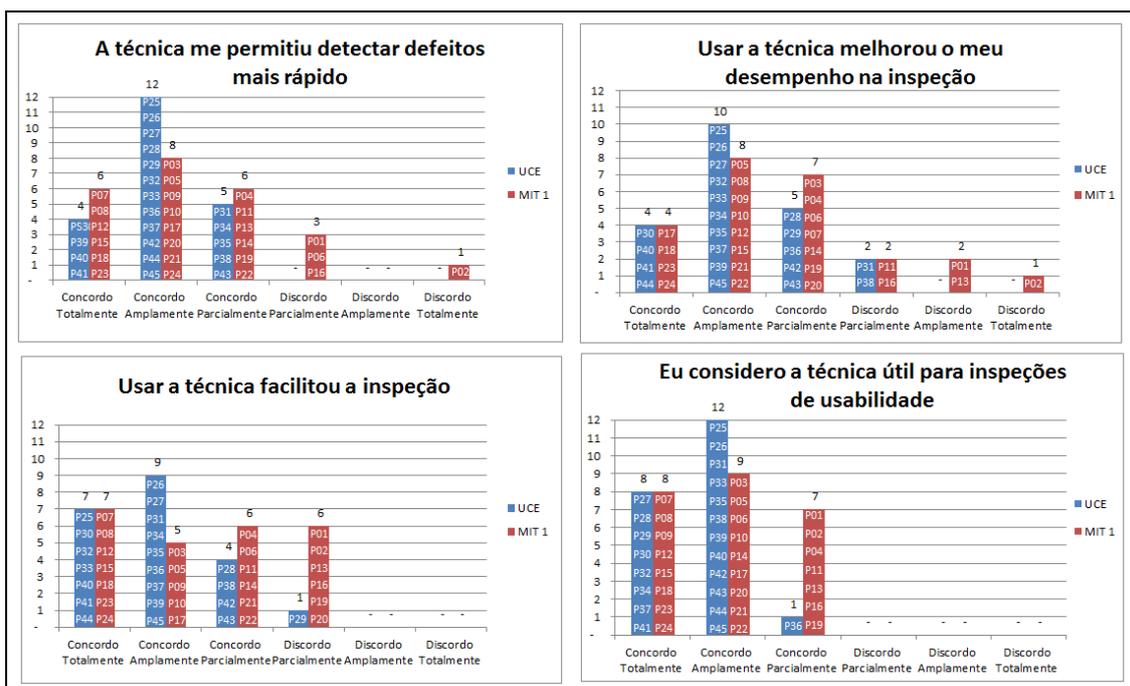


Figura 39. Percepção sobre Utilidade - 2º estudo de viabilidade da MIT 1

### c. Discussão

Foi realizada uma análise dos defeitos que foram encontrados pelo grupo da MIT 1 e que não foi encontrado pelo grupo do UCE e vice-versa. A seguir serão apresentados alguns desses defeitos.

O participante P20 (grupo da MIT 1) relatou o seguinte defeito “*O sistema não permite que o usuário refaça qualquer mudança nos dados*”. Este participante relatou na planilha de discrepância que ele identificou este defeito através do item de verificação 1AC1 da MIT 1 que solicita o seguinte “*Verifique se o usuário, através do Fluxo Alternativo e Fluxo de Exceção, pode desfazer ou refazer algo que envolva persistência de dados no sistema. Por exemplo: pode excluir ou alterar dados inseridos*”. Este defeito não foi relatado por nenhum dos participantes do grupo do UCE. Isto pode indicar que o UCE precisa orientar mais com relação à identificação de defeitos relacionados à heurística “Controle e Liberdade do Usuário”.

Outro participante do grupo da MIT 1 (P04) relatou o seguinte defeito “*A tela "Cadastro/Alteração de Curso – Centro de Treinamento" traz duas possibilidades, dando uma possível ambiguidade ao usuário*”. Este defeito foi identificado pelo item de verificação 1AD3 que solicita o seguinte “*Verifique se os nomes de campos, telas, botões e links no Fluxo Principal, Fluxo Alternativo, Fluxo de Exceção e Regras de Negócio apresentam um único sentido, sem ambiguidades*”. Nenhum dos participantes do grupo do UCE identificou este defeito. Isto pode sugerir que ter itens de verificação pode guiar mais os inspetores a encontrar defeitos de usabilidade (como a MIT 1). Note que o item de verificação 1AD3 apresentado acima também é relacionado a heurística “Controle e Liberdade do Usuário”. Portanto, o método UCE deve guiar mais inspetores a identificar defeitos de usabilidade relacionados a esta heurística.

O participante P41, que usou o método UCE, identificou o seguinte defeito “*Ausência de instruções para o usuário sobre o significado do campo Segmento*”. Ele identificou este defeito usando a heurística “Prevenção de Erros” que orienta o seguinte “*Melhor do que apresentar boas mensagens de erros, é ter um projeto cuidadoso que previne a ocorrência de um problema, em primeiro lugar*”. A falta do significado das opções do campo Segmento pode levar o usuário ao erro. Este tipo de defeito não foi relatado por nenhum participante que usou a MIT 1. Isto pode indicar que é necessário incluir itens de verificação que guiem os usuários a identificar estes tipos de defeitos nas especificações de caso de uso na técnica MIT 1.

No geral, observa-se que a MIT 1 obteve uma alta aceitação dos participantes neste estudo. Isto pode indicar que a técnica MIT 1 está adequada para ser utilizada por profissionais da indústria de *software*.

## 6.4. Ameaças à Validade do 1º e 2º Estudo de Viabilidade da MIT 1

Nestes estudos houve algumas ameaças que podem afetar a validade destes resultados. Nesta Subsubseção as ameaças serão discutidas e apresentadas conforme a categorização de Wohlin *et al.*, (2000): interna, externa, de conclusão e de constructo.

### 6.4.1. Validade Interna

Nestes experimentos, considerou-se quatro principais ameaças que representam um risco à interpretação dos resultados: (1) efeitos do treinamento, (2) classificação da experiência, (3) medição do tempo e; (4) influência do moderador. Poderia ter havido efeito do treinamento se o treinamento na AH (1º estudo de viabilidade) ou no UCE (2º estudo de viabilidade) tivesse qualidade menor que o treinamento na MIT 1. Controlou-se o efeito do treinamento preparando treinamentos equivalentes com os mesmos exemplos de detecção de discrepâncias. Além disso, com relação à classificação da experiência dos participantes, esta foi feita com base em dados factuais, ou seja, uma autoclassificação dos participantes em relação à sua experiência. Os participantes se classificaram de acordo com o número e tipo de experiências anteriores (em avaliação de usabilidade e desenvolvimento de *software*). Considerando a medição do tempo, foi solicitado aos participantes que fossem o mais preciso possível na anotação do tempo inicial e tempo final da inspeção, e o moderador também checkou o tempo anotado por cada participante quando ele/ela entregava a planilha de discrepância. E por fim, para reduzir a ameaça com relação à influência do moderador nos resultados do estudo, um time de especialistas analisaram as discrepâncias identificadas. Este time jogou se as discrepâncias eram defeitos de usabilidade ou não, sem a interferência do moderador.

### 6.4.2. Validade Externa

Cinco ameaças foram consideradas: (1) participantes foram estudantes de graduação (1º estudo de viabilidade) e estudantes de graduação, mestrado e doutorado (2º estudo de viabilidade); (2) o estudo foi conduzido em um ambiente acadêmico; (3) a validade do modelo avaliado como um modelo representativo; (4) os pesquisadores inseriram alguns defeitos no modelo; e (5) participantes necessitavam de treinamento. Com relação à ameaça 1, tanto no 1º estudo de viabilidade quanto no 2º estudo de viabilidade, poucos participantes tinham experiência na indústria. De acordo com Carver *et al.* (2003), estudantes que não têm experiência na indústria têm habilidade similares a inspetores menos experientes. Com relação à ameaça 2, o modelo inspecionado (caso de uso) é parte

da especificação de um sistema real. No entanto, não é possível afirmar que o modelo usado na inspeção represente todos os tipos de casos de uso (Ameaça 3). Com relação à ameaça 4, todos os problemas de usabilidade inseridos foram encontrados por todos os participantes de ambos os grupos. Portanto, o número de defeitos encontrados pelos inspetores de ambos os grupos foi muito maior que o número de defeitos inseridos pelo moderador. E por fim, com relação à ameaça 5, seria ideal se não houvesse necessidade de treinamento para aplicar a técnica. No entanto, o pequeno tempo gasto no treinamento permite que profissionais sem tanta experiência em avaliação de usabilidade usem a técnica.

#### **6.4.3. Validade de Conclusão**

No 1º estudo de viabilidade, o principal problema é o tamanho da amostra. O pequeno número de participantes não é o ideal do ponto de vista estatístico. Tamanho da amostra é um problema conhecido em estudos IHC e ES (CONTE *et al.*, 2007b; FERNANDEZ *et al.*, 2012a). No 2º estudo de viabilidade, o tamanho da amostra não foi considerado um problema, uma vez que houve uma quantidade alta de participantes. Outra ameaça é a homogeneidade da amostra, pois os participantes de ambos os estudos são todos estudantes de uma mesma instituição. Devido a estes fatos, há limitação dos resultados, sendo estes considerados indícios e não conclusivos. E por fim, a ameaça com relação à coleção dos dados foi tratada aplicando-se o mesmo procedimento em cada experimento individual a fim de extrair os dados e garantir que cada medida fosse calculada usando a mesma fórmula.

#### **6.4.4. Validade de Constructo**

A validade de constructo pode ter sido influenciada pelas medidas que foram aplicadas na análise quantitativa. As medidas de eficácia e eficiência adotadas são frequentemente usadas em estudos que investigam técnicas de detecção de defeitos (FERNANDEZ *et al.*, 2012a; FERNANDEZ *et al.*, 2013).

### **6.5. Síntese do Capítulo**

Este capítulo relatou os resultados da validação empírica cujo objetivo foi avaliar a eficácia e eficiência dos participantes ao utilizar a tecnologia MIT 1 em comparação com outros métodos de inspeção de usabilidade (Avaliação Heurística e UCE). Devido à alta aceitação da técnica MIT 1 pelos participantes do estudo e devido às melhorias realizadas na técnica, ela foi considerada adequada para ser utilizada em projetos de *software*.

## CAPÍTULO 7 - ESTUDOS EXPERIMENTAIS DA MIT 2

*Este capítulo apresenta os estudos experimentais que foram realizados para avaliar empiricamente e aperfeiçoar a tecnologia MIT 2.*

### 7.1. Introdução

Estudos experimentais também foram realizados com a técnica MIT 2 a fim de avaliá-la e aperfeiçoá-la. Como pode ser visto na Figura 40, foi realizado um estudo de viabilidade, um estudo de observação e um estudo de caso com a MIT 2. O estudo de observação foi realizado logo após o estudo de viabilidade porque desejava-se investigar por que a técnica MIT 2 não obteve um desempenho tão bom no estudo de viabilidade e também porque desejava-se saber como as pessoas utilizam a MIT 2. Em seguida um estudo de caso na indústria foi realizado com a técnica MIT 2 a fim de analisar seu comportamento em ambiente industrial.

Estes três estudos experimentais são descritos em mais detalhes nas subseções a seguir, assim como as melhorias realizadas na MIT 2.

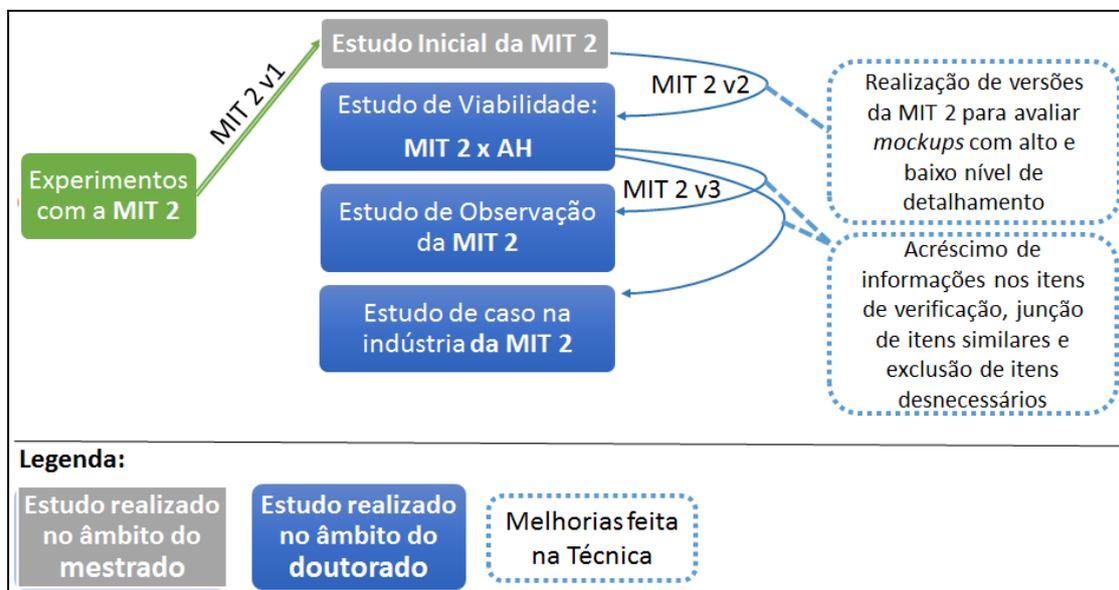


Figura 40. Estudos experimentais realizados com a MIT 2

## 7.2. 1º Estudo de Viabilidade da MIT 2

O objetivo da técnica MIT 2 é apoiar a detecção de defeitos de usabilidade de *mockups*. A técnica MIT 2 foi avaliada através de um estudo de viabilidade em comparação com a Avaliação Heurística (VALENTIM e CONTE, 2014). As Subsubseções a seguir apresentam o detalhamento deste estudo de viabilidade, incluindo a execução das atividades do processo de inspeção, os resultados alcançados e as ameaças à validade.

### 7.2.1. Planejamento do 1º Estudo de Viabilidade da MIT 2

A fim de avaliar a MIT 2 foi executado um estudo de viabilidade, usando somente o tipo da técnica com baixo nível de detalhamento. A escolha deste tipo da técnica é devida ao fato de os *mockups* utilizados não apresentarem mensagens de erro, textos informativos e advertências do sistema.

A MIT 2 foi avaliada em comparação com a AH (NIELSEN, 1994b), que, como dito anteriormente, é um método de inspeção amplamente utilizado pela indústria (FERNANDEZ *et al.*, 2012a) e é o método base da MIT 2. O objetivo deste experimento é analisar a MIT 2 com o propósito de caracterizá-la com respeito à sua eficácia e eficiência comparada à AH.

#### a. Hipóteses

O estudo foi planejado e conduzido a fim de testar as seguintes hipóteses (nula e alternativa, respectivamente):

- H01: Não há diferença entre as técnicas MIT 2 e AH com relação ao indicador de eficiência.
- HA1: Há uma diferença entre as técnicas MIT 2 e AH com relação ao indicador de eficiência.
- H02: Não há diferença entre as técnicas MIT 2 e AH com relação ao indicador de eficácia.
- HA2: Há uma diferença entre as técnicas MIT 2 e AH com relação ao indicador de eficácia.

#### b. Contexto

O estudo foi executado com dois *mockups* de um sistema real de um Centro de Treinamento, disponíveis em Valentim *et al.* (2017c). O experimento foi conduzido com estudantes de graduação do curso de Ciência da Computação (segundo semestre de 2013)

da Universidade Federal do Amazonas (UFAM). Os alunos já tinham frequentado, antes do estudo, uma disciplina introdutória de “Engenharia de *Software*” e estavam cursando uma disciplina de “Análise e Projeto de Sistema”.

### **c. Seleção de Variáveis**

As variáveis independentes foram as técnicas de avaliação de usabilidade (MIT 2 e AH) e as variáveis dependentes foram os indicadores de eficiência e eficácia das técnicas. Eficiência e eficácia foram calculadas para cada participante como: (a) a razão entre o número de defeitos detectados e o tempo gasto no processo de inspeção; e (b) a razão entre o número de defeitos detectados e o número de defeitos existentes (conhecidos), respectivamente.

### **d. Seleção de Participantes**

Dezoito participantes assinaram o TCLE e preencheram o Formulário de Caracterização. O formulário de caracterização foi aplicado para categorizar os participantes como tendo: nenhuma, baixa, média ou alta experiência em Avaliação de Usabilidade (AU) e Desenvolvimento de *Software* (DS), da mesma forma que em estudos anteriores. Na Tabela 20, a segunda coluna (AU) e a terceira coluna (DS) apresentam a categorização de cada participante com relação às suas experiências em AU e DS.

### **e. Projeto Experimental**

Participantes foram divididos em dois grupos: o grupo da MIT 2 e AH. Ambos os grupos inspecionariam os mesmos *mockups*. Os participantes foram atribuídos para cada grupo usando projeto randomizado. Cada grupo foi composto por nove participantes.

### **f. Instrumentação**

Muitos artefatos foram definidos para apoiar o experimento, tais como: formulário de consentimento e de caracterização, especificação das técnicas MIT 2 e AH, instruções para a inspeção, uma planilha para a anotação das discrepâncias identificadas, um questionário pós-inspeção e dois *mockups*, disponíveis em Valentim *et al.* (2017c). O primeiro *mockup* apresenta uma lista com as possíveis turmas cadastradas, dando a opção de refinar a busca de turmas por curso ou município, além da opção de selecionar uma turma para alteração ou exclusão. Já o segundo *mockup* representa a tela de cadastro dessas turmas. Todos os artefatos deste estudo foram validados por outro pesquisador da área.

### **g. Preparação**

Todos os participantes receberam treinamento de 2 horas sobre princípios de usabilidade. Então, para cada grupo, foi feita uma apresentação de 15 minutos para esclarecer sobre o uso de ambas as técnicas. Exemplos similares de como usar as técnicas foram mostrados para ambos os grupos. Este tempo foi suficiente para apresentar as técnicas e exemplificar o seu uso.

#### **7.2.2. Execução do 1º Estudo de Viabilidade da MIT 2**

No início do estudo, um pesquisador agiu como moderador, sendo responsável por passar as informações da avaliação para os inspetores. Então, os participantes foram divididos em dois grupos, que ficaram alocados em salas diferentes (um grupo utilizou a MIT 2 e o outro utilizou a AH). Cada participante recebeu os artefatos do estudo descritos na Subsubseção 7.2.1.f. Durante a inspeção, cada participante preencheu uma planilha com as discrepâncias encontradas. Todos os participantes devolveram a planilha contendo os possíveis defeitos e o tempo total gasto na inspeção. Eles também entregaram o questionário pós-inspeção preenchido. A atividade de detecção de defeitos foi executada individualmente por cada inspetor. Durante esta atividade, os inspetores não receberam qualquer auxílio dos pesquisadores envolvidos no estudo. Ao todo, 9 inspetores utilizaram a técnica MIT 2 e 9 inspetores utilizaram a técnica AH.

#### **7.2.3. Coleção e Discriminação do 1º Estudo de Viabilidade da MIT 2**

As atividades de Coleção e Discriminação foram realizadas de forma similar aos estudos anteriores descritos. Na discriminação, um grupo formado pelo engenheiro de *software* autor dos *mockups* e dois pesquisadores especialistas em usabilidade avaliaram as discrepâncias identificadas pelos participantes do estudo.

#### **7.2.4. Resultados Quantitativos 1º Estudo de Viabilidade da MIT 2**

A Tabela 20 mostra o resultado geral da avaliação de usabilidade em *mockups*. Os participantes P01 a P09 aplicaram a MIT 2. E os participantes P10 a P18 aplicaram a AH. Pode-se observar que os inspetores que usaram a MIT 2 encontraram entre 3 a 9 defeitos utilizando cerca de 0,08 e 0,50 horas. Por outro lado, os inspetores que usaram a AH levaram entre 0,23 e 0,50 horas e encontraram entre 4 e 8 defeitos.

Tabela 20. Resultado da inspeção no 1º Estudo de Viabilidade da MIT 2

Part.	Experiência em AU	Experiência em DS	Discrepâncias	Falso-Positivo	Defeitos	Tempo (Hora)	Defeitos /Hora
P01	Baixa	Nenhuma	7	1	6	0,42	14,40
P02	Baixa	Nenhuma	5	1	4	0,42	9,60
P03	Média	Baixa	9	2	7	0,40	17,50
P04	Baixa	Nenhuma	9	3	6	0,37	16,36
P05	Baixa	Nenhuma	4	0	4	0,08	48,00
P06	Baixa	Baixa	13	4	9	0,50	18,00
P07	Baixa	Nenhuma	8	4	4	0,43	9,23
P08	Baixa	Nenhuma	8	1	7	0,45	15,56
P09	Nenhuma	Nenhuma	4	1	3	0,50	6,00
P10	Baixa	Nenhuma	9	4	5	0,32	15,79
P11	Baixa	Nenhuma	11	5	6	0,33	18,00
P12	Baixa	Nenhuma	9	4	5	0,50	10,00
P13	Baixa	Nenhuma	6	1	5	0,38	13,04
P14	Baixa	Nenhuma	4	0	4	0,35	11,43
P15	Baixa	Baixa	10	2	8	0,35	22,86
P16	Média	Nenhuma	5	0	5	0,23	21,43
P17	Baixa	Nenhuma	8	4	4	0,42	9,60
P18	Baixa	Nenhuma	7	1	6	0,30	20,00

**Legenda:** Part. – Participante. **Experiência em AU.** - Experiência em Avaliação de Usabilidade; **Experiência em DS.** - Experiência em Desenvolvimento de Software.

No geral, a inspeção resultou em um conjunto de 35 defeitos de usabilidade, incluindo os 6 defeitos inseridos nos *mockups*. A Tabela 21 apresenta a eficácia e eficiência.

Tabela 21. Eficácia e Eficiência no 1º Estudo de Viabilidade da MIT 2

Tecnologia	MIT 2	AH
<b>Total de Defeitos</b>	50	48
<b>Média de Defeitos</b>	5,56	5,33
<b>Eficácia</b>	15,87%	15,24%
<b>Tempo Médio (min)</b>	23,78	21,22
<b>Eficiência (defeitos/hora)</b>	14,02	15,08

Uma análise foi realizada usando o teste não paramétrico de Mann-Whitney (MANN e WHITNEY, 1947), dado o tamanho limitado da amostra (DYBA *et al.*, 2006). Um resumo dos resultados foi apresentado usando o gráfico de boxplot. As análises estatísticas foram executadas usando a ferramenta SPSS V. 19, com  $\alpha = 0,10$ . A escolha desta significância estatística foi motivada pelo pequeno tamanho da amostra usada neste experimento (RIVERO e CONTE, 2012). A Figura 41 elemento A mostra o gráfico de boxplot com a distribuição da eficiência por técnica. Através deste gráfico é possível observar que o grupo da MIT 2 teve quase a mesma eficiência que o grupo da AH. Quando as duas amostras são comparadas usando o teste de Mann-Whitney, não se encontra nenhuma diferença significativa entre os dois grupos ( $p = 0,507$ ). Estes resultados apoiam a

hipótese nula H01 que afirma que não há nenhuma diferença no indicador de eficiência de ambas as técnicas.

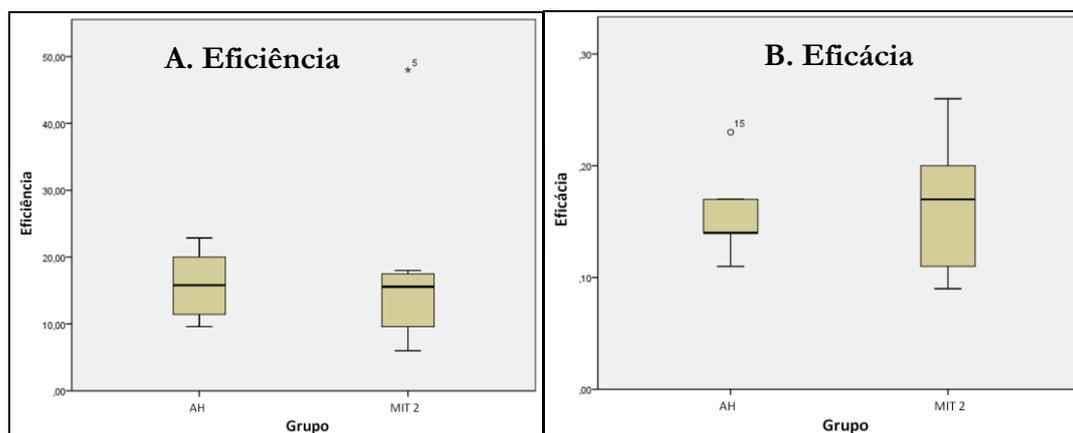


Figura 41. Boxplots da eficiência e eficácia - 1º estudo de viabilidade da MIT 2

A mesma análise foi aplicada para determinar se havia alguma diferença significativa comparando o indicador de eficácia das duas técnicas na detecção de defeitos de usabilidade. O gráfico de boxplots com a distribuição da eficácia por técnica (ver Figura 41 elemento B) mostra que o grupo da MIT 2 foi mais eficaz que o grupo da AH quando inspecionaram a usabilidade em *mockups*.

Além disso, a mediana do grupo da MIT 2 está mais alta que a mediana do grupo da AH. No entanto, quando se compara as duas amostras usando o teste de Mann-Whitney, não se encontra diferença estatística significativa entre os dois grupos ( $p = 0,892$ ). Estes resultados apoiam a hipótese nula H02 que afirma que não há nenhuma diferença significativa no indicador de eficácia da MIT 2 e AH.

Uma análise de correlação foi realizada com as variáveis eficácia e eficiência, para determinar o relacionamento entre elas. O coeficiente de correlação de Spearman entre eficácia e eficiência foi de 0,597, com  $p = 0,009$ , mostrando correlação positiva. Portanto, participantes mais eficazes também foram mais eficientes.

### 7.2.5. Análise da Percepção dos Participantes - 1º Estudo de Viabilidade da MIT 2.

Os questionários sobre a aceitação da tecnologia foram analisados somente com respeito à MIT 2, pois desejava-se avaliar a aceitabilidade somente desta técnica. Os indicadores definidos foram: (i) utilidade percebida, e (ii) facilidade de uso percebida.

#### a. Facilidade de Uso Percebida

Os dados coletados foram apresentados em formato de gráfico para fins de análise. A Figura 42 apresenta as percepções dos participantes sobre a facilidade de uso da MIT 2.

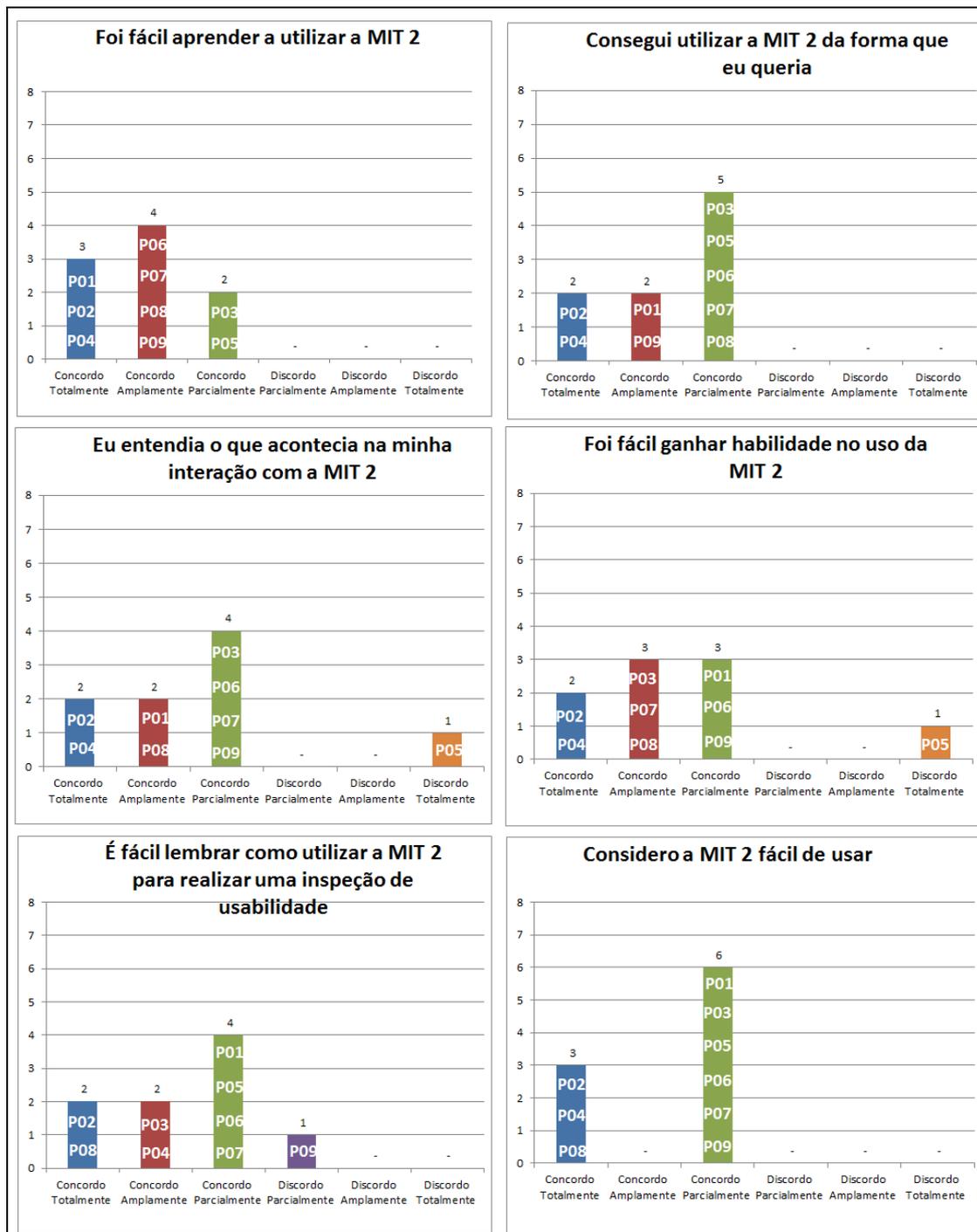


Figura 42. Percepção sobre Facilidade de Uso - 1º Estudo de Viabilidade MIT 2

O eixo X dos gráficos da Figura 42 refere-se às possíveis respostas do questionário pós-inspeção e o eixo Y refere-se à quantidade de participantes. Os códigos P01, P02 e os demais representam os participantes apresentados na Tabela 20. Pode-se notar que o participante P05, que possui baixa experiência em avaliação de usabilidade, discordou totalmente da afirmativa “Eu entendia o que acontecia na minha interação com a MIT 2”, mostrando que a MIT 2 em algum momento não foi fácil de entender. O participante P05 também discordou totalmente da afirmativa “Foi fácil ganhar habilidade no uso da MIT 2”.

Já o participante P09, que não possuía experiência em avaliação de usabilidade, discordou parcialmente da afirmativa “É fácil lembrar como usar a MIT 2 para realizar uma inspeção de usabilidade”, destacando a dificuldade de lembrar da MIT 2. No entanto, todos os inspetores concordaram com as outras afirmativas, mostrando que em geral, a técnica MIT 2 foi considerada fácil de usar.

### b. Utilidade Percebida

A Figura 43 apresenta as percepções dos participantes sobre a utilidade da MIT 2. Verificou-se que somente o participante P01, que possui baixa experiência em avaliação de usabilidade, discordou parcialmente na questão “A MIT 2 me permitiu detectar defeitos mais rápido”, indicando que em algum ponto o uso da técnica consumiu muito tempo. Porém, todos os inspetores concordaram com as outras afirmativas, reforçando que a MIT 2 é útil no processo de inspeção.

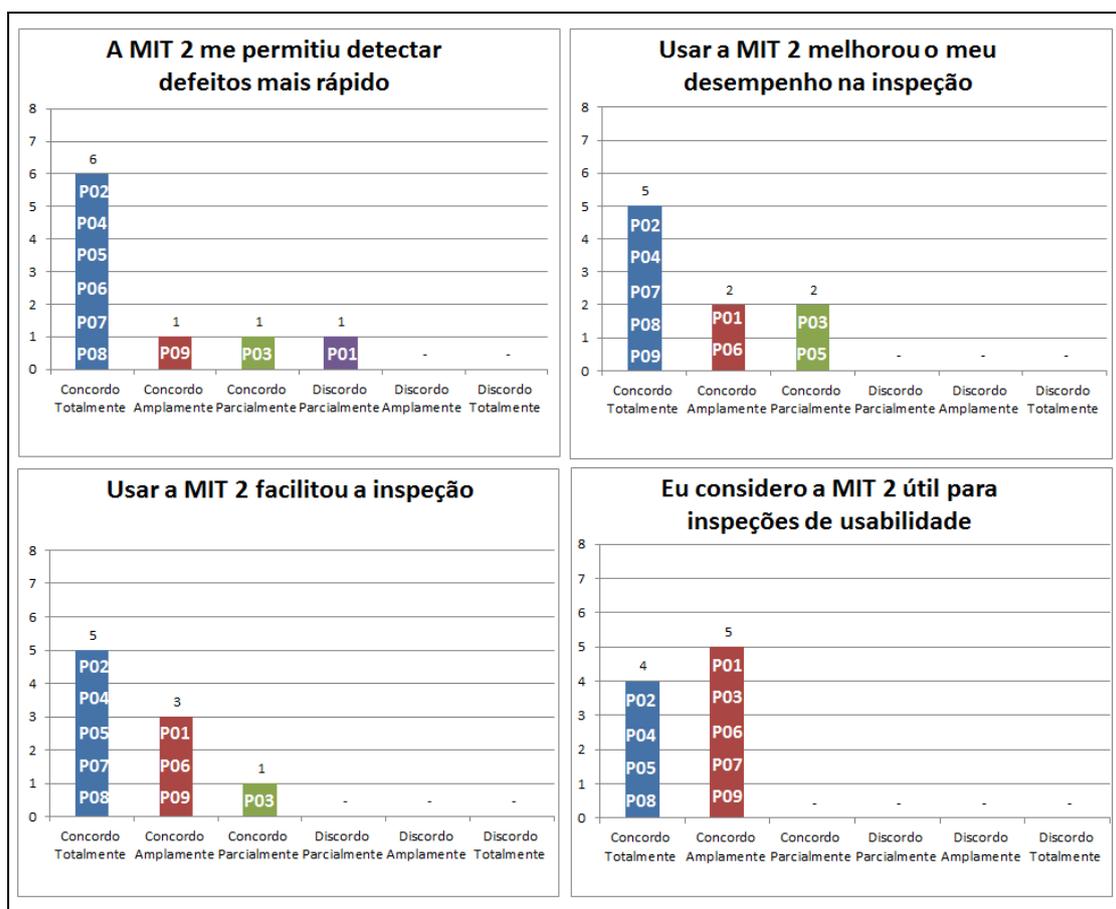


Figura 43. Percepção sobre a Utilidade da MIT 2 no 1º Estudo de Viabilidade

### 7.2.6. Análise Qualitativa do 1º Estudo de Viabilidade da MIT 2

Uma análise específica dos dados qualitativos (comentários adicionais dos inspetores) contidos nos questionários foi feita, tendo por base o método *Grounded Theory* (GT)

(CORBIN e STRAUSS, 1998). Como o objetivo da análise qualitativa era melhorar a técnica proposta (MIT 2), só foi feita a análise dos dados da MIT 2.

Os dados qualitativos extraídos dos questionários foram analisados utilizando um subconjunto das fases do processo de codificação sugerido por Corbin e Strauss (1998) para o método GT – as codificações aberta (1ª fase) e axial (2ª fase). Ao analisar os dados destes questionários, criou-se códigos (conceitos relevantes para entendimento da percepção sobre a técnica) relacionados a trechos (citações dos participantes) - codificação aberta. Após isso, os códigos foram agrupados de acordo com as suas propriedades, formando conceitos que representam categorias e subcategorias. Finalmente, estas foram relacionadas entre si - codificação axial. Os procedimentos de GT visam a uma análise mais aprofundada, através da comparação e da análise da relação entre estes conceitos. O objetivo dessa análise neste estudo foi entender qual a percepção dos inspetores sobre a experiência de uso da MIT 2. Como não se pretendia criar uma teoria a esse respeito, não foi realizada a codificação seletiva (3ª fase do método GT). As etapas de codificação aberta e axial foram suficientes para entender o motivo de alguns problemas. Os conceitos referentes ao método GT são apresentados detalhadamente em (CONTE *et al.*, 2009b).

Verificou-se que o conteúdo dos questionários teve maior direcionamento em relação à estrutura, ao uso e às possíveis melhorias na MIT 2. O processo de codificação produziu no total 17 códigos que foram associados a 2 categorias. O questionário pós-inspeção possuía afirmativas sobre a adequação e facilidade de uso da técnica, o que levou à identificação dessas duas categorias: Estrutura da MIT 2 e Feedback do Uso da MIT 2. Além disso, dentro da categoria Feedback de Uso da MIT 2 foi criada uma subcategoria chamada Sugestões de Melhorias – onde são apresentadas sugestões de modificações na MIT 2 com o objetivo de melhorar o ganho com a sua utilização.

Entre os códigos criados na categoria “Estrutura da MIT 2” (Figura 44), três deles apontam evidências de inadequação como os códigos “O item 2BA2 é inútil se não houver mensagens de erros nos *mockups*” e “2BF1 e 2BF2 têm quase a mesma função para campos de análise diferentes (um para nomes e outro para figuras)”.

O primeiro código foi citado pelo participante P08, que possui baixa experiência em avaliação de usabilidade. Ele menciona diretamente um item de verificação que não convém à MIT 2 - Baixo Nível de Detalhamento – uma vez que os *mockups* com pouco detalhamento não apresentam as mensagens do que foi realizado após uma persistência de dados (alteração, exclusão, entre outros). Pode-se notar que mesmo com pouca experiência, um participante da graduação conseguiu identificar um problema na técnica. Além disso,

cinco códigos da categoria “Estrutura da MIT 2” apontam dificuldades, como os códigos: “Não está claro o sentido da palavra “terminologia” no item 2BD1” citado pelo participante P07 e “É necessário deixar mais claro o que é consistente (com a realidade ou com o sistema)” citado pelo participante P01. Estes códigos demonstram indícios das dificuldades que inspetores tiveram com alguns termos usados pela técnica.

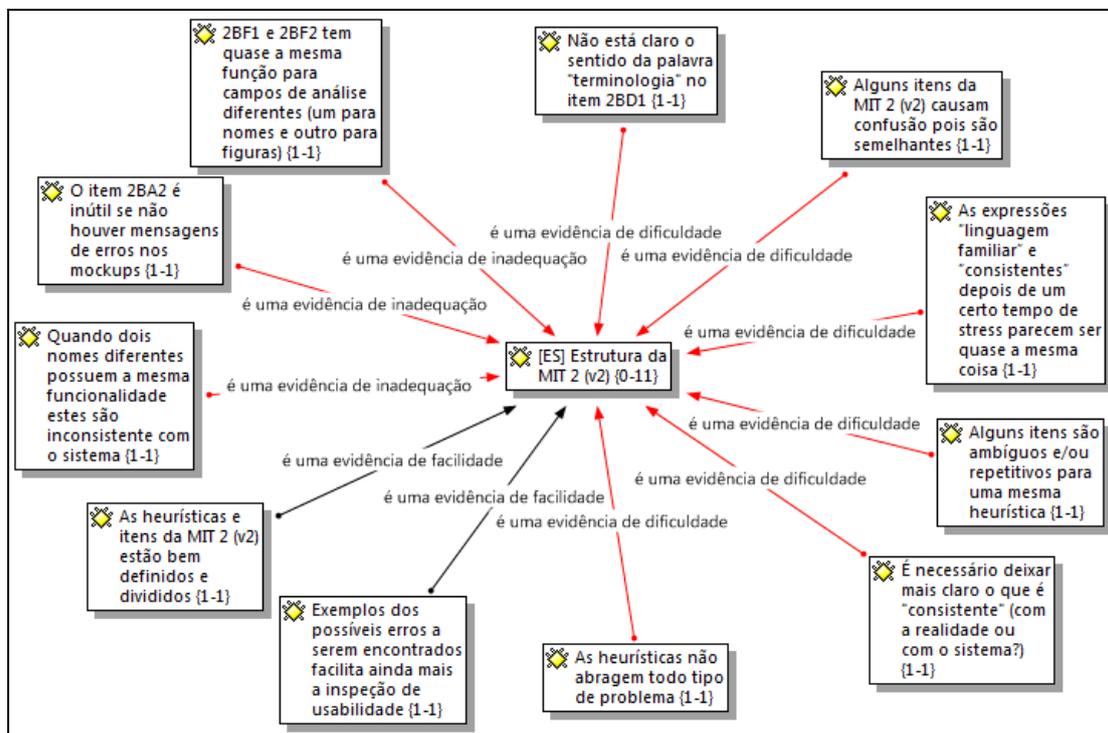


Figura 44. Categoria Estrutura da MIT 2

A Figura 45 apresenta a categoria “Feedback do uso da MIT 2”. Esta categoria possui seis códigos que apontam dificuldades, tais como: “Dificuldade de entender alguns itens da MIT 2 para procurar problemas” e “Dificuldade de imaginar links observados por pessoas experientes”. Através do primeiro código, citado pelo participante P07, é possível observar que este inspetor teve dificuldades em tentar entender alguns itens de verificação da MIT 2, levando a supor que estes itens ainda não estão claros o suficiente. O participante P07 possui baixa experiência em usabilidade e isto nos leva a notar que a técnica precisa estar mais clara para inspetores menos experientes. Além disso, no segundo código (citado pelo participante P01) é afirmado que se tem dificuldades de aplicar um determinado item, pois o mesmo também dificulta o entendimento. Quando se percebe que a técnica ainda possui itens difíceis de entender para inspetores menos experientes, novos esforços precisam ser empenhados a fim de que a melhore, pois, uma técnica difícil de entender, possivelmente não será usada.

A subcategoria “Sugestões de Melhoria” também contém 6 códigos associados que descrevem melhorias na MIT 2 e sugerem novas características (ver Figura 45). Os códigos mencionam questões referentes à diminuição do tempo de inspeção, tais como: “Sugestão de juntar alguns itens de verificação para diminuir o tempo de leitura da MIT 2” e “Sugestão de juntar os itens de verificação 2BF1 e 2BF2”, citados pelo participante P01. Além disso, os códigos sugerem inspecionar a interação entre *mockups* (“Sugestão de inspecionar a interação entre *mockups*” e “Sugestão de a MIT 2 verificar caso deixe de preencher um campo num formulário, o sistema deve avisar sobre o erro”). Ambos os códigos foram citados pelo participante P08. Isto mostra que mesmo participantes com baixa experiência em avaliação de usabilidade sugerem melhorias interessantes.

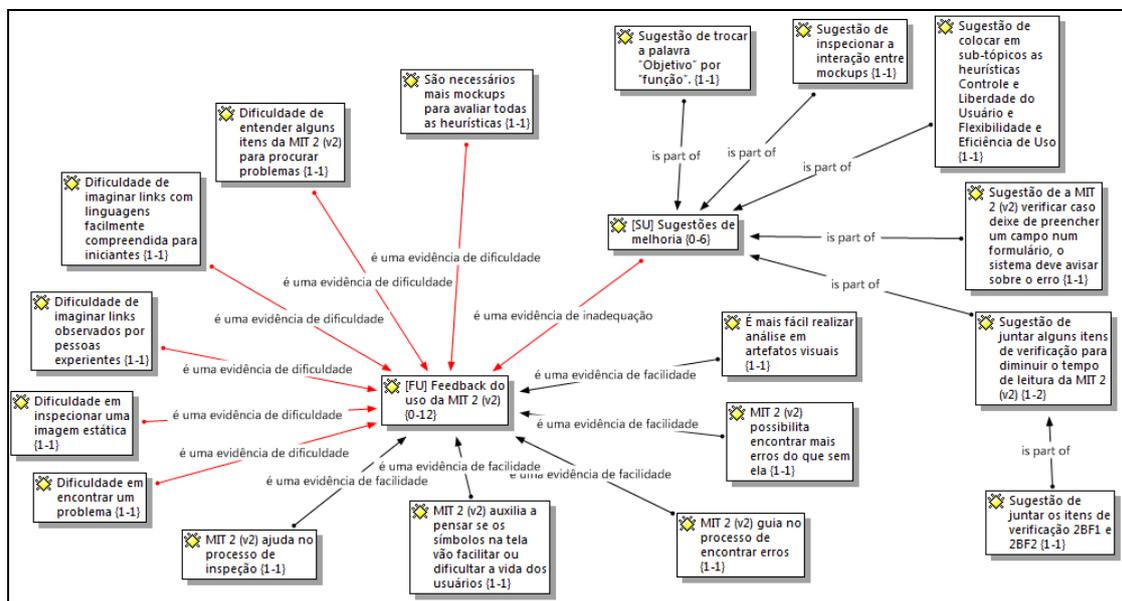


Figura 45. Categorias Feedback do uso da MIT 2 e Sugestões de Melhoria

### 7.2.7. Melhorias na MIT 2

A análise qualitativa permitiu a identificação das causas para o resultado negativo obtido com a análise quantitativa. Através da análise desses dados, alterações foram feitas e uma terceira versão (v3) da MIT 2 foi elaborada.

Um dos inspetores afirmou que “É necessário deixar mais claro o que é “consistente” (com a realidade ou com o sistema?)” no item de verificação 2BD1. Além disso, um inspetor comentou que “Não está claro o sentido da palavra “terminologia” no item 2BD1”. Para solucionar o primeiro problema, acrescentaram-se os termos “com o sistema representado”, para deixar mais claro com que a terminologia deve estar com consistente. Para resolver o segundo problema, adicionou-se o termo “nomenclatura” ao lado da palavra “terminologia”. Estas

alterações foram realizadas no item de verificação 2BD1 (ver Figura 46) que pertence à heurística 2BD - Consistência e Padrões.

<b>2BD1</b>	Verifique se a terminologia (nomenclatura), gráficos e símbolos estão consistentes com o sistema representado;
-------------	--

Figura 46. Item de verificação 2BD1 da MIT 2

Outro inspetor observou que “O item 2BA2 é inútil se não houver mensagens de erros nos mockups”. Através deste comentário foi constatado que de fato este item de verificação (ver Figura 47) não era para estar contido na MIT 2 – Baixo Detalhamento, pois esta versão da técnica não leva em consideração mensagens de erro, textos informativos e advertências. Portanto, este item de verificação foi retirado da MIT 2 – Baixo Detalhamento.

<b>2BA2</b>	Verifique se há algum texto que informa ao usuário o que foi realizado após uma persistência de dados (alteração, exclusão, etc).
-------------	---

Figura 47. Item de verificação 2BA2 da MIT 2 – Baixo Detalhamento

Através do código “Dificuldade de imaginar links observados por pessoas experientes”, um inspetor expressou sua dificuldade em tentar entender como seriam links abreviados para usuários experientes presente no item de verificação 2BG6 (ver Figura 48). Levando em consideração que este item causa dificuldades de entendimento e que o mesmo já está contido no item de verificação 2BG7 (ver Figura 48), o item de verificação 2BG6 foi retirado tanto da MIT 2 – Baixo Detalhamento quanto da MIT 2 – Alto Detalhamento.

<b>2BG6</b>	Verifique se há nomes de links abreviados para usuários mais experientes;
<b>2BG7</b>	Verifique se há como as funcionalidades do sistema serem realizadas de forma mais rápida e eficiente para usuários mais experientes.

Figura 48. Item de verificação 2BG6 da MIT 2 – Baixo e Alto Detalhamento

Um dos inspetores comentou que os itens de verificação “2BF1 e 2BF2 tem quase a mesma função para campos de análise diferentes (um para nomes e outro para figuras)” e deu a “Sugestão de juntar os itens de verificação 2BF1 e 2BF2”. Para melhorar a técnica no ponto citado, uniram-se os itens de verificação 2BF1 e 2BF2 em um único item (ver Figura 49), reduzindo o tempo de leitura da técnica. A mesma alteração foi feita na MIT 2 – Alto Detalhamento.

<b>2BF1</b>	Verifique se os nomes das opções, campos, botões e links são informados e se os ícones/figuras minimizam o esforço físico e cognitivo do usuário de se lembrar;
-------------	---

Figura 49. Item de verificação 2BF1 da MIT 2

Alguns códigos como “Dificuldade de entender alguns itens da MIT 2 para procurar problemas”, “Alguns itens são ambíguos e/ou repetitivos para uma mesma heurística”, e “Sugestão de juntar alguns itens de verificação para diminuir o tempo de leitura da MIT 2” foram atendidos com as

melhorias já relatadas. Códigos como “*Sugestão de inspecionar a interação entre mockups*” e “*Sugestão de a MIT 2 verificar caso deixe de preencher um campo num formulário, o sistema deve avisar sobre o erro*”, ainda serão melhor analisados e possivelmente serão atendidos em versões futuras da técnica.

Outros problemas foram relatados, como “*Dificuldade em encontrar um problema*”, “*Dificuldade em inspecionar uma imagem estática*”, “*Alguns itens da MIT 2 causam confusão pois são semelhantes*”, “*As heurísticas não abrangem todo tipo de problema*”. Buscou-se fazer uma verificação abrangente na técnica para atender a estes pontos. Melhorias foram realizadas na MIT 2 e uma versão da técnica foi criada.

### **7.2.8. Ameaças à Validade do 1º Estudo de Viabilidade da MIT 2**

#### **a. Validade Interna**

Neste experimento, consideraram-se quatro principais ameaças que representam um risco para uma interpretação inapropriada dos resultados: (1) efeitos do treinamento, (2) classificação da experiência, (3) medição de tempo e (4) influência do moderador. Estas mesmas ameaças já foram explicadas em estudos anteriores (ver Subseção 6.4).

#### **b. Validade Externa**

Quatro ameaças à validade externa foram consideradas: (1) participantes foram estudantes de graduação; (2) o estudo foi conduzido em um ambiente acadêmico; (3) a validade do artefato avaliado como um artefato representativo; (4) o pesquisador inseriu alguns defeitos nos *mockups*; e (5) participantes com necessidade de treinamento. Da mesma forma estas ameaças já foram abordadas nos estudos anteriores.

#### **c. Validade de Conclusão**

Neste estudo, o principal problema é o tamanho e homogeneidade da amostra. A quantidade de participantes não é a ideal do ponto de vista estatístico e, além disso, os participantes são todos estudantes de uma mesma instituição. Futuramente, pretende-se realizar novos estudos com mais participantes e que sejam da indústria. Devido a estes fatos, há limitação nos resultados, sendo estes considerados indícios e não conclusivos.

#### d. Validade de Constructo

As medidas de eficiência e eficácia adotadas neste estudo são frequentemente usadas em estudos que investigam técnicas de detecção de defeitos (FERNANDES *et al.*, 2012; FERNANDEZ *et al.*, 2012a).

### 7.3. Estudo de Observação da MIT 2

Para apoiar o desenvolvimento e validação da técnica MIT 2 adotou-se a metodologia experimental apresentada em Shull *et al.* (2001). A metodologia compreende quatro fases: (1) estudos de viabilidade: para determinar a possibilidade de utilização da tecnologia; (2) estudos observacionais: para melhorar a compreensão e a relação custo-eficácia da tecnologia; (3) estudos de caso do ciclo de vida real: para caracterizar a aplicação da tecnologia durante um ciclo de vida real; e (4) estudos de caso na indústria: identificar se a aplicação da tecnologia se encaixa nas configurações industriais.

A fim de verificar a possibilidade de uso da MIT 2 foi realizado o estudo de viabilidade descrito anteriormente. O resultado do teste estatístico deste estudo apontou que a MIT 2 obteve eficácia e eficiência similar à Avaliação Heurística. Isto indica que novos estudos necessitam ser realizados para identificar que parte do processo de inspeção com a MIT 2 necessita de melhorias. Espera-se que as pessoas envolvidas no projeto possam usar a MIT 2 para garantir a qualidade de seus *mockups*. Para alcançar este objetivo, realizou-se a segunda fase da metodologia apresentada em Shull *et al.* (2001) (estudo de observação) para obtenção de um entendimento detalhado de como a MIT 2 é aplicada (VALENTIM *et al.*, 2015b).

O objetivo de um estudo observacional é coletar dados sobre como uma tarefa particular é realizada. Realizou-se um estudo observacional com o propósito de elicitare o processo usado por desenvolvedores quando aplicam a técnica MIT 2 - Alto Detalhamento durante uma avaliação de usabilidade. O objetivo foi entender profundamente o processo de uso da MIT 2, portanto, a MIT 2 não foi comparada a nenhuma outra técnica. O estudo observacional deve responder à seguinte questão: “Que passos são adotados para aplicação da MIT 2?”.

Técnicas observacionais podem ser usadas para entender práticas de trabalho atuais (MAGUIRE, 2001). Neste estudo, obteve-se dois tipos de dados qualitativos: dados observacionais e inquisitivos. Os dados observacionais foram coletados durante o processo de inspeção. Para obter dados observacionais, utilizou-se o método de Avaliação Cooperativa (MÜLLER, 1997), uma vez que desenvolvedores descrevem o que eles(as)

estão fazendo e o observador é livre para fazer perguntas e pedir explicações sobre as decisões ou ações dos desenvolvedores (MÜLLER, 1997). Dados inquisitivos foram coletados no final das inspeções usando entrevistas.

### 7.3.1. Planejamento do Estudo de Observação da MIT 2

Para realizar este estudo, os seguintes pontos de investigação qualitativos e quantitativos foram definidos (ver Tabela 22). Os dados quantitativos foram medidos com a finalidade de comparar os resultados quantitativos deste estudo com o outro estudo realizado com a MIT 2.

Tabela 22. Pontos de investigação e suas definições

Tipo de dados	Ponto de investigação	Definição
Dados qualitativos	Processo de Aplicação	Relato de como os desenvolvedores aplicam a MIT 2 e como a aplicariam se utilizassem a técnica novamente;
	Facilidade de Uso da Técnica	Relato das dificuldades e/ou facilidades que os desenvolvedores possuem ao utilizar a técnica MIT 2;
	Intenção de Uso da Técnica	Relato de como desenvolvedores acreditam que utilizariam a MIT 2 em seus projetos de desenvolvimento ou avaliações no geral;
	Sugestão de melhoria	Sugestões de modificações na técnica MIT 2.
Dados Quantitativos	Eficiência	A taxa entre o número de defeitos encontrados e o tempo gasto;
	Eficácia	A taxa entre o número de defeitos detectados e o número total de defeitos existentes (conhecidos).

O estudo foi executado com dois *mockups* que fazem parte da especificação de um sistema real de um Centro de Treinamento, o mesmo do 2º estudo de viabilidade da MIT 2. Os *mockups* que foram avaliados nesse estudo são: *mockup* de cadastro de curso e *mockup* de listagem de curso (onde é possível selecionar um curso para excluí-lo ou editá-lo), e algumas mensagens que o sistema apresenta após a persistência de dados. Os *mockups* possuem problemas de usabilidade reais que influenciariam no uso do sistema projetado.

As entrevistas e observações foram realizadas em um Centro de Formação de uma grande empresa de TI, onde o foco é a inovação e desenvolvimento de *software*. O Centro de Formação apoia muitos desenvolvedores em projetos reais, principalmente no desenvolvimento de sistemas *Web* e aplicações *Mobile*, adotando metodologias ágeis como Scrum e XP. A fim de lidar com as questões éticas do estudo, criou-se um termo de consentimento livre para informar sobre os procedimentos de pesquisa e confidencialidade.

Quinze desenvolvedores assinaram o termo de consentimento. Todos os participantes receberam treinamento em *mockups* e princípios de usabilidade. Exemplos foram mostrados em como usar a técnica MIT 2.

Os resultados coletados durante o estudo foram analisados juntos com os dados obtidos das entrevistas. Para as entrevistas, um questionário semiestruturado foi usado com questões abertas. As questões relacionadas aos pontos de investigação Processo de aplicação, Facilidade de uso, Intenção de uso da técnica e Sugestões de melhoria são apresentadas na Tabela 23.

**Tabela 23. Pontos de investigação e questões usadas na entrevista**

Pontos de Investigação	Questões
Processo de Aplicação	Como você aplicou a técnica com relação à ordem de leitura e procura de problemas de usabilidade? Por que você acha que essa é a melhor forma de aplicar a técnica?
	Como você aplicou a técnica com relação à ordem de uso das heurísticas? Por que você acha que essa é a melhor forma de aplicar a técnica?
	Como você aplicaria a técnica se fosse realizar uma nova avaliação de usabilidade?
Facilidade de Uso da Técnica	Quais aspectos da técnica tornam sua aplicação fácil/difícil de usar? Por quê?
	Quais itens de verificação você encontrou dificuldade em entender? Por quê?
Intenção de Uso da Técnica	Você utilizaria esta técnica em um projeto de desenvolvimento de <i>software</i> no seu ambiente de trabalho? Como?
Sugestão de melhoria	Na sua opinião, como a técnica poderia ser melhorada?

### 7.3.2. Execução do Estudo de Observação da MIT 2

Cada desenvolvedor aplicou a técnica MIT 2, avaliando o *mockup* e identificando problemas de usabilidade. Quando cada desenvolvedor encontrou um problema de usabilidade, ele(a) descreveu o problema em uma planilha de discrepância. Após isso, um pesquisador entrevistou os desenvolvedores e eles forneceram suas impressões com relação à técnica MIT 2. O observador realizou algumas anotações em um formulário. É importante notar que o observador podia questionar as ações dos desenvolvedores a qualquer hora, mas não era permitido ajudar os desenvolvedores na atividade de descoberta de problemas de usabilidade.

### 7.3.3. Coleção do Estudo de Observação da MIT 2

Um dos pesquisadores agiu como moderador da inspeção e foi responsável por conduzir o estudo observacional. Depois da inspeção individual por cada desenvolvedor, o

moderador checou todas as planilhas de discrepâncias e obteve uma lista única de discrepâncias. Durante esta atividade, o moderador destacou discrepâncias duplicadas.

### 7.3.4. Discriminação do Estudo de Observação da MIT 2

Uma reunião de discriminação foi realizada com dois outros pesquisadores especialistas em usabilidade (não envolvidos com o estudo). O propósito desta reunião foi analisar todas as discrepâncias identificadas por cada desenvolvedor, da mesma forma que foi realizada nos estudos anteriores.

### 7.3.5. Transcrição das Entrevistas do Estudo de Observação da MIT 2

Por fim, foi realizada a transcrição das entrevistas realizadas com os desenvolvedores. As entrevistas possibilitaram a esta pesquisa ter informações para entender como desenvolvedores utilizaram a técnica MIT 2. A análise dos dados destas entrevistas é apresentada na Subsubseção 7.3.7 Não foram fornecidos os nomes dos desenvolvedores que participaram das entrevistas a fim de preservar a sua identidade.

### 7.3.6. Análise dos Dados Quantitativos do Estudo de Observação da MIT 2

Após a atividade de discriminação, contou-se o número de discrepâncias, falso-positivos e defeitos por desenvolvedor. A Tabela 24 apresenta os resultados de cada desenvolvedor, mostrando o número de discrepâncias relatadas, o número de falso-positivos entre estas discrepâncias, o número atual de defeitos, o tempo, a eficiência e a eficácia de cada desenvolvedor.

Tabela 24. Resultados da inspeção - Estudo de Observação da MIT 2

Participantes	Discrepâncias	Falso-Positivos	Defeitos	Tempo (Hora)	Eficiência (Defeitos/Hora)	Eficácia
P01	9	0	9	0,38	23,48	25,00%
P02	6	0	6	0,58	10,29	16,67%
P03	4	1	3	0,28	10,59	8,33%
P04	8	1	7	0,30	23,33	19,44%
P05	7	0	7	0,45	15,56	19,44%
P06	9	3	6	0,40	15,00	16,67%
P07	9	1	8	0,38	20,87	22,22%
P08	11	3	8	0,43	18,46	22,22%
P09	13	8	5	0,47	10,71	13,89%
P10	13	1	12	0,40	30,00	33,33%
P11	10	4	6	0,62	9,73	16,67%
P12	12	4	8	0,57	14,12	22,22%
P13	12	5	7	0,48	14,48	19,44%
P14	16	6	10	1,02	9,84	27,78%
P15	19	3	16	0,50	32,00	44,44%
Média	10,53	2,67	7,87	0,48	-	21,85%

No total, a inspeção resultou em um conjunto de 36 defeitos de usabilidade, incluindo os 7 defeitos inseridos. Os desenvolvedores que usaram a MIT 2 conseguiram encontrar entre 3 e 16 defeitos entre 0,28 e 1,02 horas. Calculou-se a eficácia neste estudo de observação em 21,85%. Comparando esta medida com a eficácia do grupo de estudantes de graduação que utilizaram a técnica MIT 2 no 1º estudo de viabilidade da MIT 2 (15,87%), pode-se notar que este indicador foi mais alto no estudo de observação. No entanto, não se pode afirmar que a melhoria da eficácia se deve unicamente à melhoria realizada na MIT 2. Outros fatores como: (1) o aumento no conhecimento em avaliações de usabilidade, (2) o fato dos participantes do estudo de observação serem desenvolvedores e os participantes do 1º estudo de viabilidade serem alunos, deve ser considerada.

### **7.3.7. Análise dos Dados Qualitativos do Estudo de Observação da MIT 2**

Após a análise quantitativa, foi feita uma análise específica dos dados qualitativos que foram obtidos através dos comentários dos desenvolvedores na entrevista, tendo por base os procedimentos de *Grounded Theory* (GT) (CORBIN e STRAUSS, 2008).

Os dados qualitativos extraídos das entrevistas foram analisados utilizando um subconjunto das fases do processo de codificação do método GT, apresentado no estudo anterior. Vale lembrar que o objetivo da análise qualitativa neste estudo foi entender como é realizado o processo de aplicação da MIT 2 por desenvolvedores. Decidiu-se não eleger uma categoria central, porque a regra de GT é a circularidade entre as etapas de coleta e análise até que a saturação teórica seja atingida (CORBIN e STRAUSS, 2008). Portanto não foi realizada a codificação seletiva (3ª fase do método GT). As etapas de codificação aberta e axial foram suficientes para entender o motivo de alguns problemas.

A seguir serão discutidos os resultados da pesquisa qualitativa sobre a MIT 2.

#### **a. Ponto de vista sobre o Processo de Aplicação da MIT 2**

Através das entrevistas pode-se notar que os desenvolvedores empregaram a MIT 2 de três diferentes formas, que foram: (i) primeiro o desenvolvedor leu a técnica depois procurou o problema no *mockup* (ver citação abaixo de P08); ou (ii) primeiramente o desenvolvedor procurou o problema no *mockup* e depois leu a técnica (ver citação abaixo de P03); ou (iii) inicialmente o desenvolvedor observou o *mockup*, depois leu a técnica e após isso mudou este modo de aplicação e quando lia um item já ia procurando os problemas (ver citação abaixo de P10).

“*Eu li item a item e tentei achar os problemas em cada mockup*” (Participante 8).

*“Primeiro eu olhei por cima os mockups, eu vi alguns problemas que eu conhecia. Mas depois disso, eu comecei a ler a técnica desde o começo”* (Participante 3).

*“Como são vários itens que tem na técnica, eu olhei o mockup e achei algumas coisas, mas se eu fosse ter que procurar essas coisas (...) uma por uma eu acho que ia demorar mais tempo. Então eu preferi guardar essas coisas que eu achei de errado e (...) eu fui fazendo na ordem de ler cada item e procurar ou relacionar com algum que (...) tinha identificado (...)”* (Participante 10).

Foram coletadas também algumas opiniões sobre a aplicação da técnica. Alguns participantes afirmaram que na hora de aplicar a MIT 2, depois de conhecê-la, é melhor pular alguns itens (ver citação abaixo de P11). Além disso, outros participantes disseram que primeiro observar o *mockup* e depois ler a MIT 2 não é a melhor forma de começar a inspeção (ver citação abaixo de P05). No entanto, um dos participantes acredita que primeiro observar o *mockup* e depois ler a MIT 2, permitiu que ele achasse a aplicação da técnica fácil, pois já sabia onde alguns dos problemas estavam (ver citação abaixo de P06).

*“Fui pulando, teve até uns [itens] que eu não achei muito problema”* (Participante 11).

*“Olhar primeiro [o mockup] eu acho que não é a melhor forma de começar a inspeção”* (Participante 5).

*“Durante a execução dessa avaliação eu percebi que eu comecei a aprender o que significava essa técnica e eu consegui olhar pra um elemento e já saber que problema estava associado (...). Então essa estratégia de olhar o mockup [e] percorrer uma lista pode ser um ponto de partida pra você conseguir memorizar as heurísticas, mas no futuro o que você percebe é que você acaba abandonando isso e você desenvolve uma habilidade”* (Participante 6).

Além disso, foram coletadas algumas opiniões positivas sobre a aplicação da técnica, como: aplicar a MIT 2 não demora (ver citação abaixo de P03); acredita ter encontrado uma maior quantidade de problema utilizando a MIT 2 do que encontraria sem ela (ver citação abaixo de P06) e MIT 2 serve como guia (ver citação abaixo de P12).

*“MIT 2 não toma muito tempo”* (Participante 3).

*“Provavelmente eu encontrei mais problemas utilizando a técnica porque você presta atenção nos detalhes mais pontuais, nos detalhes menores e não tanto no todo da aplicação (...). Teria coisas que eu nem teria olhado por exemplo, se tinha identificação em cima das telas ou não, onde estava, o que estava fazendo, (...) seria uma coisa que eu não teria olhado de jeito nenhum”* (Participante 6).

*“É sempre bom ter um guia para ajudar”* (Participante 12).

Adicionalmente, neste estudo buscou-se obter informações de como os participantes aplicariam a MIT 2 em uma outra oportunidade. Alguns participantes leriam um item de verificação e já procurariam o problema (ver citação abaixo de P03). As justificativas dadas pelos participantes para este modo de aplicação são: (i) senão a inspeção pode ficar mais demorada, (ii) senão depois você esquece, (iii) depois você não sabe onde está o item e (iv) porque isso elimina erros primários. Um dos participantes afirmou que observaria primeiro as heurísticas e procuraria o problema e só depois que encontrasse o problema, procuraria o item de verificação relacionado a ele (ver citação abaixo de P14). No entanto, outros participantes disseram que primeiro analisariam os *mockups* e depois relacionariam os problemas encontrados com a técnica (ver citação abaixo de P11).

*“Eu acho que ler [a MIT 2] e a partir dela ir procurando os problemas dos mockups seria a melhor forma de fazer [a inspeção]”* (Participante 3).

*“Se eu fosse fazer uma nova avaliação, eu não ia perder aquele tempo que eu comecei lendo a [MIT 2]. Se eu fizesse a avaliação pela segunda vez, como eu já tenho o conhecimento prévio, eu olhava (...) em qual dessas heurísticas (...) [o problema] se enquadra. Mas (...) para colocar o item de verificação daí só na segunda vez pegando o número do item”* (Participante 14).

*“A forma mais adequada é você ir olhando, analisando o mockup e relacionando com a técnica”* (Participante 11).

#### **b. Ponto de vista sobre a Facilidade de Uso da MIT 2**

Além da identificação de como os participantes aplicaram e como aplicariam a MIT 2 e suas opiniões sobre o processo de inspeção, foi realizada a coleta de informações relacionadas às dificuldades e facilidades de uso da técnica. A seguir serão apresentadas algumas dificuldades e facilidades ao utilizar a técnica.

Algumas das dificuldades encontradas ao usar a MIT 2 foram: dificuldade de identificar a qual dos itens de verificação o problema encontrado estava relacionado (ver citação abaixo de P06); dificuldade de entender algumas heurísticas pois é necessário primeiro interpretá-las para depois procurar o problema (ver citação abaixo de P09); dificuldade de saber o que se deve fazer quando não se consegue encontrar um problema através de um item de verificação (ver citação abaixo de P02); dificuldade de usar a MIT 2, pois possui muito texto (ver citação abaixo de P13); e dificuldade de lembrar dos itens (ver citação abaixo de P14).

*“O (...) desafio é você conseguir classificar esses erros, isso torna um pouco difícil a execução da técnica, mas ela não invalida a sua aplicação”* (Participante 6).

“*Algumas heurísticas são difíceis de entender e não sei se é por falta de uso ou por não usar isso no dia a dia, (...) tem meio que interpretar o quê que é a heurística para depois procurar, porque não fica tão intuitivo assim*” (Participante 9).

“*Como não tinha identificado nenhum problema através do item de verificação em mockup nenhum, eu não sabia se eu tinha que dizer que não tinha nenhum [problema] ou achar que realmente estava certo porque não tinha*” (Participante 2).

“*Difícil de usar seria porque tem bastante texto*” (Participante 13).

“*Lembrar dos itens foi mais difícil*” (Participante 14).

Além dessas dificuldades, alguns participantes citaram dificuldades específicas relacionadas aos itens de verificação e heurísticas da técnica, como: (i) dificuldade de aplicar o item de verificação 2AE1 pois não se sabia qual a entrada de dados do campo (ver citação abaixo de P13); (ii) a heurística 2AI não está clara (ver citação abaixo de P01); (iii) dificuldade de aplicar o item 2AG2 em *mockups* (ver citação abaixo de P13); e (iv) dificuldade de diferenciar as heurísticas 2AE e 2AH (ver citação abaixo de P06).

“*Não consegui identificar um formato de dado correto (...). Então esse foi um item que eu parei para pensar, para verificar antes*” (Participante 13).

“*Não entendi: "proporcionar as funções que são úteis para o usuário", isso não fica bem claro*” (Participante 1).

“*[Item] 2AG2: verifique se há apoio a tarefas específicas frequentemente repetidas. Eu não consegui identificar isso nos mockups (...) eu parei para pensar, para olhar de novo, para ver se isso existia*” (Participante 13).

“*Senti um pouco de dificuldade com (...) [as heurísticas] “Reconhecimento de diagnóstico e recuperação de erros” e “Prevenção de erros”. Em algumas leituras que eu fiz me pareciam ser a mesma coisa*” (Participante 6).

Alguns participantes também citaram algumas das facilidades encontradas ao se utilizar a MIT 2, como: as heurísticas da MIT 2 estão claras (ver citação abaixo de P14), ter exemplos tornou a aplicação fácil (ver citação abaixo de P03), facilidade de entender a MIT 2 (ver citação abaixo de P05), e MIT 2 é fácil pois possui itens separados por heurística (ver citação abaixo de P03).

“*E a facilidade foi [que as] heurísticas estão bem claras, então fica fácil de você ver o problema no mockup (...)*” (Participante 14).

“*Alguns exemplos que tinham também deixaram fácil (...) de utilizar a técnica*” (Participante 3).

“*Acho que não tive dificuldade nenhuma de entender*” (Participante 5).

*“Acho que tornou fácil a questão de ser bem separada (...), analisar uma heurística de cada vez”* (Participante 3).

### **c. Opinião sobre a Intenção de Uso da MIT 2 em projetos**

Algumas das opiniões sobre o uso da MIT 2 em projetos de desenvolvimento de *software* foram: utilizaria a MIT 2 no início do projeto (ver citação abaixo de P04); e para aplicar a MIT 2 em um projeto será necessário explicar sua vantagem, pois embora haja custo com treinamento e tempo inicial das pessoas, posteriormente terão ganho com as melhorias (ver citação abaixo de P06).

*“Gostaria de tentar utilizar [a MIT 2] no início do projeto para não deixar [a avaliação] para o final”* (Participante 4).

*“[Sobre o uso da técnica em projeto] (...) tem que passar primeiro por uma explicação, um convencimento que ela vai gerar um custo adicional no seu processo de desenvolvimento. Esse custo pode ser de tempo ou de recursos humanos e isso impacta no desenvolvimento do software. Então esse custo tem que ser muito bem explicado para os superiores dizendo que embora você tenha um custo inicial você tem um benefício logo em seguida, o que é difícil é você argumentar para as pessoas que isso é importante e que as organizações possam de fato reconhecer essa importância e assumir que prazos podem ser cedidos ou orçamentos comprometidos por conta dessas melhorias”* (Participante 6).

### **d. Sugestões de Melhoria para a MIT 2**

Uma das sugestões de melhoria para a MIT 2 que merece ser destacada foi: acrescentar um item na técnica que verifique se as cores do sistema e layout do sistema representado estão adequados (ver citação abaixo de P12).

*“Tem alguns pontos [que] (...) tá preto e branco no mockup. (...) não identifiquei a imagem. (...) a questão de layout, a questão das cores também influencia, de certo ponto falta algumas coisas”* (Participante 12).

Outras sugestões de melhoria foram: automatizar a técnica (ver citação abaixo de P10); e permitir que novos itens sejam inseridos na técnica de acordo com a necessidade das pessoas (ver citação abaixo de P6).

*“Acho que [a técnica] seria muito válida se fosse automatizada”* (Participante 10).

*“De alguma forma as pessoas que usam essas técnicas prescritivas tem que ter uma certa liberdade de incluir novos elementos que a técnica não previu”* (Participante 6).

### 7.3.8. Discussão sobre os Resultados Qualitativos do Estudo de Observação

Com relação ao ponto de investigação Processo de Aplicação da MIT 2 (*bullet a* da Subsubseção 7.3.7, pode-se notar que houve participantes que preferiram primeiro observar o *mockup* e depois ler a técnica. Quando eles liam um item de verificação da MIT 2 e lembravam de um possível problema de usabilidade já observado no *mockup*, eles relacionavam o problema com o item de verificação. Para eles, havia a vantagem de, ao ler a técnica, já saber onde alguns dos problemas estavam. Estes desenvolvedores também afirmaram que este pode ser um ponto de partida para o uso da técnica, pois este modo de aplicação também ajuda a memorizar os itens da MIT 2.

Em contrapartida, outros participantes disseram que primeiro observar o *mockup* e depois ler a MIT 2 não é a melhor forma de começar a inspeção. Estes participantes leram um item de verificação e já procuraram o problema, pois para eles a inspeção se torna mais rápida, você não esquece de anotar problemas identificados e não esquece de anotar onde está o item de verificação relacionado ao problema.

Durante a análise do estudo, os pesquisadores notaram que a técnica auxilia na identificação de problemas de usabilidade nos dois modos de aplicação. Além disso, após usar técnica pela primeira vez, os desenvolvedores já possuem o conhecimento prévio dela, e podem pular a leitura de algumas heurísticas. Desta maneira, não há necessidade de estipular uma ordem pré-definida de aplicação da técnica MIT 2. Os desenvolvedores tendem a ajustar a aplicação da técnica para a sua própria maneira de uso. Isso permite que os desenvolvedores se sintam mais à vontade para usar a técnica conforme lhes convém.

Em relação ao ponto de investigação Facilidade de Uso da técnica (*bullet b* da Subsubseção 7.3.7, pode-se observar que houve algumas dificuldades e uma delas precisa ser destacada: dificuldade de identificar a qual dos itens de verificação o problema encontrado estava relacionado. Muitas vezes um problema de usabilidade fere mais que um item de verificação. Portanto, estes problemas podem estar relacionados a mais de um item. Para deixar isso mais claro para o desenvolvedor e para que o mesmo não sinta dificuldade neste sentido durante a inspeção, sugere-se acrescentar uma nota na técnica informando que um problema pode estar relacionado a mais de um item de verificação. Ao usar a técnica, o desenvolvedor só precisa identificar o problema de usabilidade. Não é necessário que o desenvolvedor referencie qual item de verificação feriu o problema identificado.

Outra dificuldade encontrada foi: saber o que fazer quando não se consegue encontrar um problema através de um item de verificação. Desta forma, se um item de verificação não auxiliou a identificar nenhum problema pode ser que seja por três razões: (i)

não há problema no *mockup* relacionado àquele item, (ii) o item de verificação não está claro o suficiente, ou (iii) não há item de verificação para aquele problema. No primeiro caso, será necessário deixar mais claro nas instruções da técnica que caso o desenvolvedor não consiga identificar um problema através de um item de verificação, o mesmo deverá prosseguir com a avaliação. No segundo caso, melhorias no item serão necessárias.

E por fim, pode-se notar que o terceiro ponto de investigação (*bullet c* da Subsubseção 7.3.7) apresentou algumas opiniões dos participantes relacionado ao uso da MIT 2 em projetos de desenvolvimento de *software*. Um dos participantes afirmou que para aplicar a MIT 2 em um projeto seria necessário explicar sua vantagem, pois embora haja custo com treinamento e tempo inicial das pessoas, posteriormente a empresa de *software* teria ganhos com as melhorias. Para alguns desenvolvedores, a avaliação de usabilidade nas fases iniciais do processo de desenvolvimento pode possibilitar vantagens, como: menor índice de retrabalho e menores custos. Isto acontece porque identificam-se problemas de usabilidade mais cedo e consertam-se antes do desenvolvimento (codificação) da aplicação. É mais barato do que consertar algo que já está desenvolvido. Através deste estudo com profissionais da indústria, as empresas de *software* possuem indícios das vantagens e opiniões destes profissionais acerca da avaliação de usabilidade nas fases iniciais do processo de desenvolvimento de *software*.

No geral, nota-se que os resultados deste estudo indicam indícios que a MIT 2 fornece alguns benefícios para desenvolvedores e que possivelmente poderá apoiar empresas de *software* a considerar a usabilidade em seus processos de desenvolvimento. Portanto, faz-se necessário avaliar como a MIT 2 se comporta em um ambiente industrial.

### **7.3.9. Ameaças à Validade do Estudo de Observação da MIT 2**

#### **a. Validade Interna**

Neste experimento, consideraram-se duas principais ameaças que representam um risco para uma interpretação inapropriada dos resultados: (1) efeitos do treinamento e (2) influência do moderador. Pode ter havido um efeito do treinamento se o treinamento na técnica MIT 2 tivesse qualidade diferente para cada desenvolvedor. Controlou-se os efeitos do treinamento, preparando um treinamento único para todos os desenvolvedores. Para diminuir a ameaça com relação à influência do moderador nos resultados do estudo, na reunião de discriminação dois especialistas fizeram a análise das discrepâncias encontradas, julgando se eram defeitos de usabilidade ou não, sem a interferência do moderador.

### **b. Validade Externa**

Três ameaças foram consideradas com relação à limitação da generalização dos resultados: (1) a validade do artefato avaliado como um artefato representativo; (2) o pesquisador inseriu alguns defeitos nos *mockups*; e (3) participantes com necessidade de treinamento. Estas ameaças já foram abordadas em estudos anteriores.

### **c. Validade de Conclusão**

Neste estudo, as principais ameaças que podem afetar a capacidade de se tirar conclusões corretas sobre este estudo são tamanho reduzido e homogeneidade da amostra. Estas ameaças já foram discutidas nos outros estudos relatados nesta tese.

## **7.4. Estudo de Caso na Indústria da MIT 2**

A fim de verificar como a MIT 2 se comporta em um ambiente industrial, realizou-se um estudo de caso na indústria, sendo esta a quarta fase da metodologia apresentada em Shull *et al.* (2001). O objetivo foi verificar se a MIT 2 apoia a antecipação da usabilidade nas fases iniciais do processo de desenvolvimento de *software* em um ambiente industrial. O estudo deve responder a seguinte questão: “A MIT 2 antecipa a usabilidade em um projeto de desenvolvimento de *software* real?”.

### **7.4.1. Planejamento do Estudo de Caso na Indústria da MIT 2**

O estudo foi realizado com um *mockup* que faz parte da especificação de um sistema que gerencia os serviços de assistência técnica e extensão rural e florestal no estado do Amazonas. Foi dado a este sistema o nome fictício Sistema X, por questões de confidencialidade. O *mockup* que foi avaliado nesse estudo faz parte do cadastro de beneficiário, onde é possível cadastrar dados pessoais do beneficiário. Este *mockup* é bastante extenso, uma vez que é necessário cadastrar as informações básicas do beneficiário, além de outras informações como dados do cônjuge e dos filhos. O *mockup* possuía problemas de usabilidade reais que influenciariam no uso do sistema projetado.

A equipe de projeto Sistema X é composta por um gerente de projeto, um analista, dois desenvolvedores e um designer. Antes do estudo ser executado todos da equipe assinaram o TCLE e consentiram participar de uma reunião do projeto com foco em discussões sobre usabilidade de um *mockup* do Sistema X. Além disso, três usuários foram convidados a testar este mesmo módulo do sistema logo após o seu desenvolvimento.

#### 7.4.2. Execução do Estudo de Caso na Indústria da MIT 2

A execução deste estudo de caso foi dividida em duas etapas. Na 1ª etapa ocorreu a realização da reunião de discussão sobre a usabilidade de um *mockup* do Sistema X referente ao cadastro de dados pessoais do beneficiário. Na 2ª etapa ocorreram os testes de usabilidade realizados com três usuários no sistema já desenvolvido.

Durante a 1ª etapa, a pesquisadora que propôs a técnica participou da reunião do projeto com foco na usabilidade do sistema e agiu como moderadora nas discussões. No decorrer da reunião, a pesquisadora lia um item de verificação da MIT 2, e logo em seguida, todos da equipe analisavam o *mockup*, conferiam se existia algum problema de usabilidade relacionado ao item da técnica lido e discutiam possíveis melhorias no *mockup*. Uma das pessoas da equipe anotava os problemas de usabilidade identificados e as sugestões de melhoria. Esta reunião durou em torno de 30 minutos. Após esta reunião, as melhorias necessárias foram realizadas no *mockup*, apresentadas ao cliente e aprovadas por ele para serem desenvolvidas no sistema. Após isso, um dos desenvolvedores da equipe implementou a tela de cadastro de dados pessoais para ser testada na próxima etapa.

Na 2ª etapa, três usuários que fazem parte da instituição que solicitou o desenvolvimento do Sistema X testaram a parte de cadastro de dados pessoais do beneficiário já implementado. Foi dada a estes usuários uma planilha impressa com dados pessoais fictícios de um beneficiário para que eles cadastrassem no sistema. Foi dito a estes usuários que eles podiam expressar suas dificuldades e dúvidas em voz alta durante a execução desta atividade. Durante o teste, a moderadora fez as anotações destas dificuldades e dúvidas em um relatório de usabilidade. O teste de usabilidade foi realizado individualmente com cada usuário e levou cerca de 20 minutos com cada um.

#### 7.4.3. Análise dos Resultados do Estudo de Caso na Indústria da MIT 2

Após a execução deste estudo os dados coletados foram analisados. Na 1ª etapa a equipe do projeto identificou quatro problemas de usabilidade nos *mockups*, descritos a seguir:

- Problema 1: formulário de cadastro muito grande.
  - Sugestão de melhoria: dividir as informações em abas.
- Problema 2: o botão “Cancelar” do lado direito é perigoso, pois o usuário pode clicar sem querer e desperdiçar tudo que ele já tenha digitado. Porém, é necessário fornecer uma saída para o usuário.

- Sugestão de melhoria: eliminação do botão “Cancelar” e inclusão do botão “Voltar” em um lugar menos propenso a erros.
- Problema 3: não fazia sentido após ter preenchido todo o formulário, o usuário ter a opção de “Salvar Rascunho”, já que ele podia voltar neste formulário para fazer alterações através da opção “Editar”.
  - Sugestão de melhoria: exclusão do botão “Salvar Rascunho” e modificação do botão “Salvar Definitivamente” para “Salvar”.
- Problema 4: não fazia sentido ter a opção de cadastrar organização às quais o beneficiário pertencia nesta etapa do cadastro, pois este cadastro é apenas para dados pessoais e isso acabava sobrecarregando o usuário nesta etapa.
  - Sugestão de melhoria: alteração do campo “Organizações às quais pertence” para outra aba, fora do cadastro de dados pessoais.

Na 2ª etapa os usuários identificaram três problemas durante o uso do sistema implementado, descritos a seguir:

- Problema 1: O campo telefone é um campo obrigatório. Porém, os três usuários esqueceram de clicar no botão “+” ao lado deste campo para realizar a ação de inserção do número do telefone. Eles pensavam que digitar o número do telefone bastava. Desta forma, como os usuários não clicaram no botão “+”, eles não conseguiram prosseguir salvando o cadastro.
- Problema 2: Usuário 2 ficou na dúvida se tinha que clicar no ícone da lupa na hora de adicionar um bairro. Mas logo em seguida ele percebeu que não precisava, pois, ao começar a digitar o nome do bairro as opções cadastradas no sistema apareciam na tela. Portanto, o ícone da lupa confundiu o usuário.
- Problema 3: Usuário 3 esqueceu de preencher o campo “Regime de casamento” e este campo não estava sinalizado como obrigatório no Sistema X, no entanto, o sistema não deixava salvar porque este é um campo obrigatório.

Como pode ser observado, a reunião de usabilidade utilizando a MIT 2 permitiu identificar alguns problemas de usabilidade que foram resolvidos antes do sistema ser desenvolvido. Alguns outros problemas foram identificados somente durante o teste de usabilidade com usuários. No entanto, como pode ser notado, alguns destes problemas são problemas que só podem ser identificados durante a interação com a aplicação implementada, pois necessitam de ações do sistema que normalmente não são representadas no *mockup*. Desta forma, destaca-se a importância que MIT 2 teve ao antecipar alguns problemas de usabilidade no início do projeto do Sistema X.

## 7.5. Síntese do Capítulo

Este capítulo relatou os resultados da avaliação empírica da tecnologia MIT 2. No estudo de viabilidade, a MIT 2 foi avaliada em comparação com a Avaliação Heurística. Em seguida, um estudo de observação foi realizado com o propósito de compreender o uso desta técnica. E por fim, um estudo de caso na indústria para analisar como a MIT 2 se comporta em um ambiente industrial foi realizado. Estes experimentos serviram como um meio valioso para a obtenção de feedback dos participantes com o objetivo final de aperfeiçoar a MIT 2. Devido à alta aceitação da técnica MIT 2 pelos participantes do estudo e profissionais de *software*, ela foi considerada adequada para ser utilizada em projetos de *software*.

## CAPÍTULO 8 - ESTUDOS EXPERIMENTAIS DA MIT 3 E UDRT-AD

*Este capítulo apresenta os estudos experimentais que foram realizados para avaliar empiricamente e aperfeiçoar empiricamente as tecnologias MIT 3 e UDRT-AD.*

### 8.1. Introdução

Para avaliar e evoluir tanto a MIT 3 quanto a UDRT-AD foram realizados estudos experimentais. Como pode ser visto na Figura 50, foi realizado um estudo de viabilidade da MIT 3 em comparação à técnica Diretrizes de Usabilidade (DU). A DU possui o mesmo propósito da MIT 3, avaliar a usabilidade através dos Diagramas de Atividades. Após melhorias serem realizadas na MIT 3, um segundo estudo de viabilidade foi realizado, porém desta vez o propósito foi comparar a técnica UDRT-AD com a MIT 3 juntamente com uma abordagem tradicional que auxilie na construção do diagrama de atividades. Estes dois estudos experimentais serão descritos em mais detalhes nas subsecções a seguir, assim como as melhorias realizadas na MIT 3.

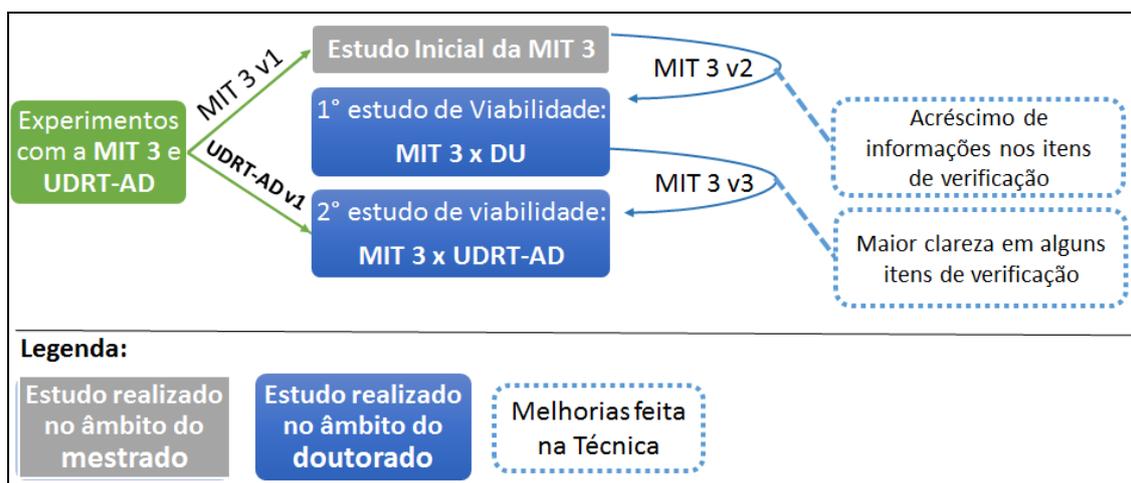


Figura 50. Estudos experimentais realizados com a MIT 3 e UDRT-AD

### 8.2. 1º Estudo de Viabilidade da MIT 3

O objetivo da MIT 3 é apoiar a detecção de defeitos de usabilidade através de Diagrama de Atividades. A técnica MIT 3 foi avaliada através de um estudo de viabilidade

em comparação à Avaliação Heurística (VALENTIM *et al.*, 2013). As Subsubseções a seguir apresentam o detalhamento deste estudo de viabilidade, incluindo a execução das atividades do processo de inspeção, os resultados alcançados e as ameaças à validade.

### 8.2.1. Planejamento do 1º Estudo de Viabilidade da MIT 3

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 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. Os níveis de conhecimento relacionados à IHC, Avaliação e Design são mostrados na Tabela 25 (segunda, terceira e quarta coluna).

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. O Diagrama de Atividades avaliado neste estudo faz parte da especificação de um sistema real, que realiza cadastro, exclusão e edição de dados de doações, e está disponível em Valentim *et al.* (2017c).

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 utilizaria para fazer a inspeção.

### 8.2.2. Execução do 1º Estudo de Viabilidade da MIT 3

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 detecção 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.

### 8.2.3. Coleção e Discriminação do 1º Estudo de Viabilidade da MIT 3

As atividades de coleção e discriminação foram realizadas de forma similar aos outros estudos relatados nesta proposta. Na atividade de discriminação, a lista de discrepâncias foi passada para uma equipe de dois outros pesquisadores da área de IHC, para que esta equipe avaliasse as discrepâncias identificadas pelos participantes do estudo.

### 8.2.4. Análise Quantitativa do 1º Estudo de Viabilidade da MIT 3

Os resultados para cada inspetor que avaliou o diagrama de atividades após a classificação durante a reunião de discriminação são mostrados na Tabela 25.

Tabela 25. Resultado da inspeção no 1º estudo de viabilidade da MIT 3

Part.	Conh. em IHC	Conh. em Avaliações	Conh. em Design	Nº de Disc.	Total de Falso-positivo	Total de Defeitos	Tempo (Hora)	Defeitos /Hora
P01	Baixo	Baixo	Médio	22	1	21	0,83	25,20
P02	Baixo	Baixo	Médio	17	2	15	0,60	25,00
P03	Baixo	Baixo	Baixo	23	1	22	1,32	16,71
P04	Baixo	Baixo	Baixo	22	0	22	0,85	25,88
P05	Médio	Baixo	Baixo	21	4	17	1,00	17,00
P06	Baixo	Baixo	Médio	20	0	20	1,10	18,18
P07	Baixo	Baixo	Baixo	20	0	20	1,25	16,00
P08	Médio	Baixo	Baixo	16	5	11	0,58	18,86
P09	Alto	Alto	Médio	17	0	17	1,27	13,42
P10	Médio	Baixo	Médio	25	2	23	1,12	20,60
P11	Baixo	Baixo	Baixo	12	0	12	0,87	13,85
P12	Médio	Médio	Baixo	16	1	15	1,05	14,29
P13	Médio	Médio	Médio	16	4	12	0,88	13,58
P14	Médio	Médio	Baixo	13	5	8	1,10	7,27
P15	Baixo	Baixo	Baixo	13	1	12	0,93	12,86
P16	Baixo	Baixo	Médio	19	5	14	0,85	16,47
P17	Baixo	Baixo	Baixo	25	3	22	0,75	29,33
P18	Baixo	Baixo	Baixo	22	6	16	0,40	40,00
P19	Baixo	Baixo	Baixo	22	4	18	1,35	13,33
P20	Médio	Baixo	Médio	14	2	12	1,12	10,75
P21	Baixo	Baixo	Baixo	13	3	10	0,63	15,79
P22	Médio	Baixo	Baixo	16	4	12	0,58	20,57
P23	Médio	Baixo	Baixo	13	0	13	0,38	33,91
P24	Baixo	Baixo	Médio	18	4	14	0,83	16,80
P25	Baixo	Médio	Baixo	15	2	13	0,67	19,50

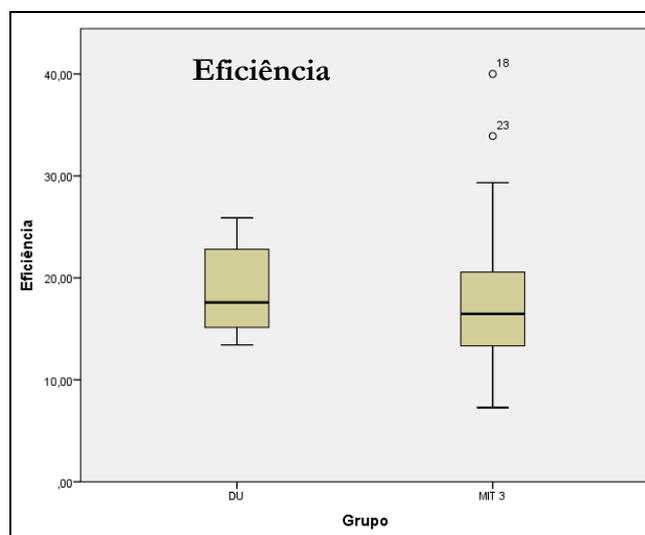
**Legenda:** Part. – Participante; Conh. em IHC – Conhecimento em IHC; Conh. em Avaliações – Conhecimento em Avaliações; Conh. em Design – Conhecimento em Design; Nº de Disc. – Número de Discrepâncias.

Os participantes P01 a P12 aplicaram as Diretrizes de Usabilidade. E os participantes P13 a P25 aplicaram a MIT 3. Ao analisar esta tabela, pode-se verificar 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 detecção, encontrando entre 11 e 23 defeitos. A Tabela 26 apresenta as médias dos indicadores de eficácia e eficiência.

**Tabela 26. Eficácia e Eficiência no 1º Estudo de Viabilidade da MIT 3**

Tecnologia	Diretrizes de Usabilidade	MIT 3
<b>Total de Defeitos</b>	215	176
<b>Média de Defeitos</b>	17,92	13,54
<b>Eficácia</b>	46%	35%
<b>Tempo Médio (horas)</b>	0,99	0,81
<b>Eficiência (defeitos/hora)</b>	18,17	16,79

Para comparar a eficiência das duas amostras, utilizou-se análise de boxplots e o teste não-paramétrico Mann-Whitney (MANN e WHITNEY, 1947), dado o tamanho limitado das amostras. Para realização da análise estatística foi utilizado o *software* SPSS com  $\alpha=0,05$ . A Figura 51 mostra os boxplots comparando a distribuição de eficiência por técnica.



**Figura 51. Boxplots de eficiência do 1º Estudo de Viabilidade da MIT 3**

Ao analisar a Figura 51, pode-se notar que a mediana do grupo das Diretrizes de Usabilidade está um pouco mais alta que a mediana do grupo da 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.

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 52 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.

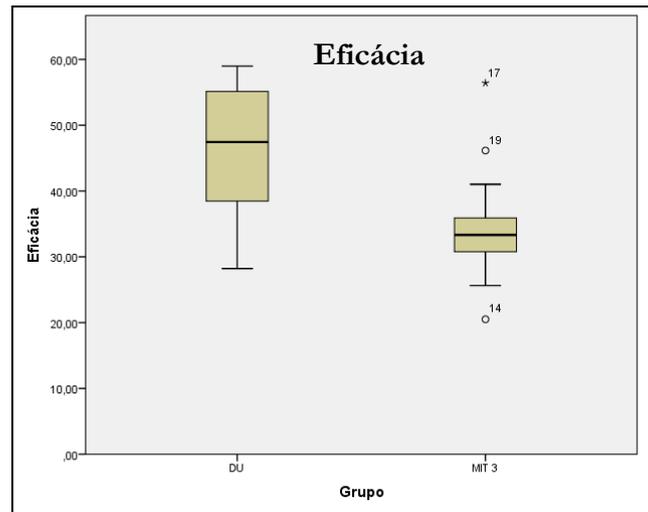


Figura 52. Boxplots de eficácia do 1º Estudo de Viabilidade da MIT 3

### 8.2.5. Análise da Percepção dos Participantes do 1º Estudo de Viabilidade - MIT 3

Neste estudo também se utilizou o modelo de aceitação de tecnologia (*Technology Acceptance Model* – TAM). Os dois indicadores utilizados foram: (1) Percepção sobre Facilidade de Uso e (2) Percepção sobre Utilidade. Eles serão abordados a seguir:

#### a. Percepção sobre a Facilidade de Uso das Técnicas

A Figura 53 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 sentença “foi fácil aprender a utilizar a técnica”. Dois dos doze discordaram das afirmativas “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 sentença “considero a técnica fácil de lembrar”. Já com relação à MIT 3, pode-se destacar que em cada uma das sentenças houve apenas um inspetor – dos treze – que discordou.

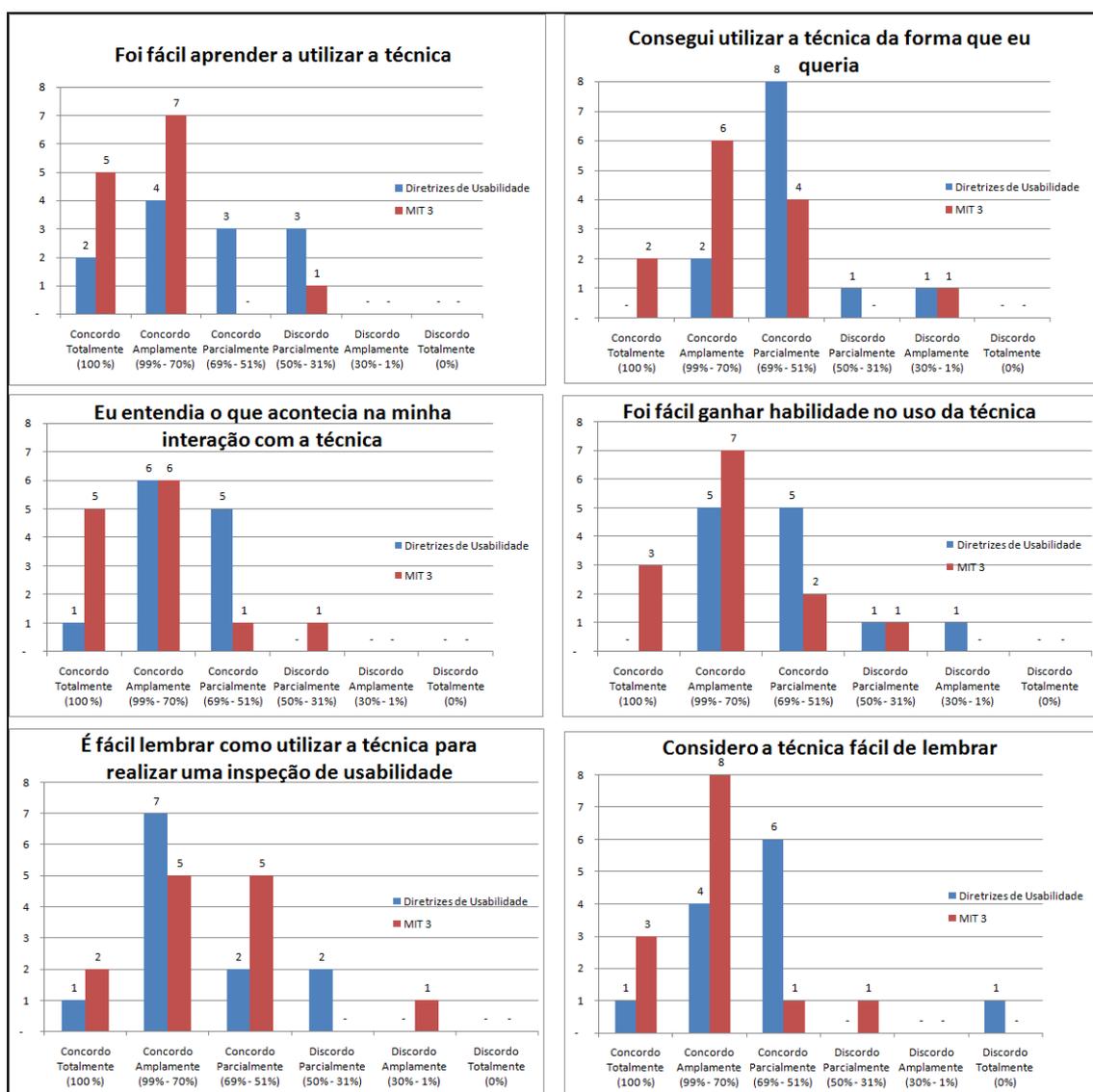


Figura 53. Percepção sobre Facilidade de Uso - 1º Estudo de Viabilidade MIT 3

### b. Percepção sobre a Utilidade das Técnicas

As respostas relacionadas à Percepção sobre Utilidade das técnicas são exibidas na Figura 54. As afirmativas 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 sentença “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 sentenças “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.

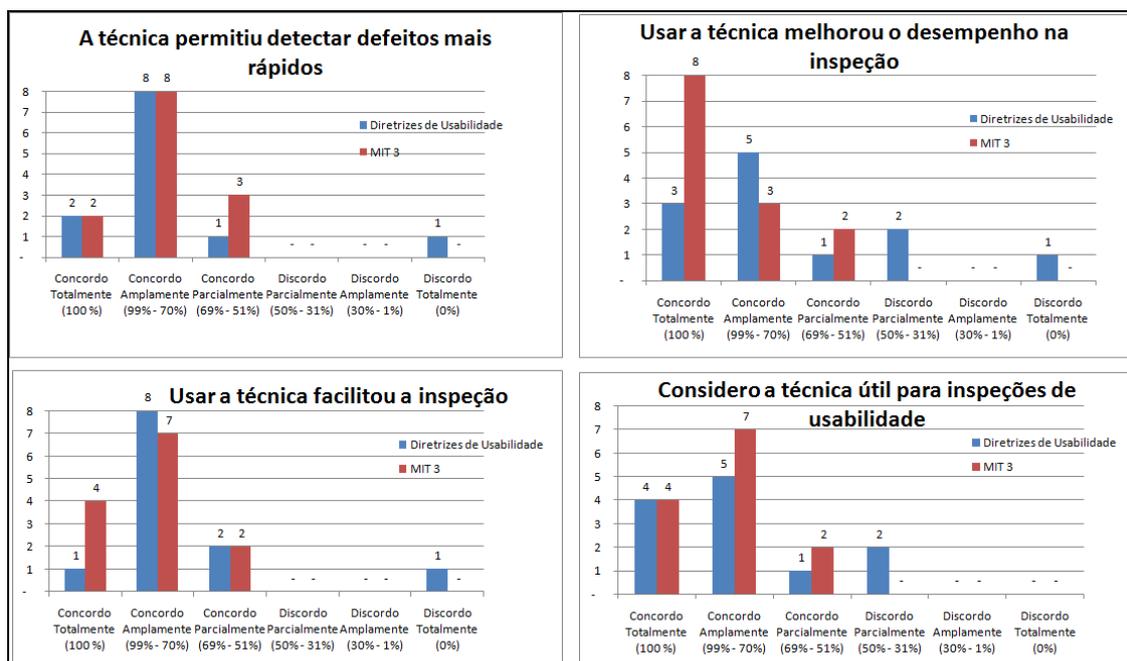


Figura 54. Percepção sobre a Utilidade - 1º Estudo de Viabilidade da MIT 3

## 8.2.6. Ameaças à Validade do 1º Estudo de Viabilidade da MIT 3

### a. Validade Interna

Neste experimento, consideraram-se quatro principais ameaças que representam um risco para uma interpretação inapropriada dos resultados: (1) efeitos do treinamento, (2) classificação da experiência, (3) medição de tempo e (4) influência do moderador. Estas ameaças já foram explicadas em estudos anteriores (ver Subseção 6.4).

### b. Validade Externa

Quatro ameaças à validade externa foram consideradas: (1) participantes foram estudantes de graduação; (2) o estudo foi conduzido em um ambiente acadêmico; (3) a validade do artefato avaliado como um artefato representativo; (4) o pesquisador inseriu alguns defeitos nos *mockups*; e (5) participantes com necessidade de treinamento. Da mesma forma estas ameaças já foram abordadas nos estudos anteriores (ver Subseção 6.4).

### c. Validade de Conclusão

Neste estudo, o principal problema é o tamanho e 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. Devido a estes fatos, há limitação nos resultados, sendo estes considerados indícios e não conclusivos.

#### **d. 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 Fernandez *et al.*, (2012a).

### **8.3. 2º Estudo de Viabilidade da MIT 3 e UDRT-AD**

Um estudo foi realizado a fim de verificar a viabilidade da técnica UDRT-AD e MIT 3. Neste estudo, a técnica UDRT-AD (SILVA *et al.*, 2014b) foi comparada com uma Abordagem Tradicional (AT). Em uma abordagem tradicional, primeiramente os designers modelam os diagramas e, em seguida, realizam uma inspeção de usabilidade nos diagramas modelados utilizando uma técnica de inspeção de usabilidade. Neste estudo, a técnica utilizada para inspecionar os diagramas foi a MIT 3, pois esta foi a técnica base para a definição das diretrizes de usabilidade da UDRT-AD. Nesta tese, a abordagem tradicional combinada com a técnica MIT 3 será chamada de AT&MIT3. As atividades e os papéis que compõem o processo de realização deste estudo serão descritas a seguir.

#### **8.3.1. Planejamento do Estudo de Viabilidade da UDRT-AD e MIT 3**

Nesta etapa, realizou-se a definição do escopo do estudo (preparação das técnicas), elaboração do cenário e questionários pós-modelagem, a seleção e treinamento dos participantes. Todas as atividades desta fase foram realizadas pelo moderador do estudo e revisadas por dois pesquisadores.

Os participantes deste estudo foram alunos voluntários da disciplina de Análise e Projeto de Sistemas (APS), do 5º período do curso de graduação em Sistema de Informação da Universidade Federal do Amazonas. Ao todo, 16 participantes concordaram em participar do estudo. Todos assinaram um Termo de Consentimento Livre e Esclarecido (TCLE) e preencheram um formulário de caracterização.

O formulário de caracterização tinha como objetivo identificar a experiência dos participantes com relação à: experiência em Design de IHC e experiência em desenvolvimento de *software* como Engenheiro de *Software*. As experiências dos participantes foram analisadas e classificadas como Nenhuma (N), Baixa (B), Média (M) ou Alta (A) experiência em Design de IHC ou em desenvolvimento de *software*.

Para classificar a experiência dos participantes no que diz respeito à experiência em Design de IHC, foi considerado: (a) Nenhuma: se o participante não possuía nenhum conhecimento sobre Design em IHC; (b) Baixa: se o participante possuía somente noções de Design em IHC adquirido através de leituras ou palestras; (c) Média: se o participante possuía experiência em projetos ou avaliações de IHC em sala de aula, e (d) Alta: se o participante possuía experiência em projetos ou avaliações de IHC na indústria.

Para classificar a experiência dos participantes com relação à experiência em desenvolvimento de *software* na indústria atuando como Engenheiro de *Software*, foi considerado: (a) Nenhuma: se o participante não possuía experiência em desenvolvimento; (b) Baixa: se o participante possuía no máximo 1 ano de experiência em desenvolvimento; (c) Média: se o participante possuía entre 1 e 4 anos em experiência em desenvolvimento, e (d) Alta: se o participante possuía mais de 4 anos de experiência em desenvolvimento.

Para diminuir o viés de ter participantes mais experientes que outros trabalhando em favor de uma das abordagens, buscou-se dividir os participantes em dois grupos que tivessem a mesma quantidade de participantes e com experiências equivalentes. Com isso, os participantes foram designados para cada grupo de forma aleatória e balanceada, de acordo com o nível de experiência de cada participante. Cada grupo foi composto por oito participantes. A Tabela 27 apresenta as duas equipes e a experiência de cada participante.

**Tabela 27. Caracterização dos participantes do Estudo de Viabilidade da UDRT-AD e MIT 3**

Grupo 1 – UDRT-AD									
Classificação	Participante	01	02	03	04	05	06	07	08
	IHC	B	B	M	N	N	N	N	N
	DS	B	B	N	B	B	B	N	B
Grupo 2 – AT&MIT3									
Classificação	Participante	09	10	11	12	13	14	15	16
	IHC	N	N	B	B	M	B	N	N
	DS	B	N	B	B	A	N	N	N
<b>Legenda:</b> IHC: Experiência em Design de IHC; DS: Experiência em Desenvolvimento como Engenheiro de <i>Software</i> ; N: Nenhuma; B: Baixa; M: Média; A: Alta.									

Um treinamento sobre usabilidade foi realizado para todos os participantes do estudo. Além disso, para o grupo que usou a técnica UDRT-AD, foi realizada uma breve apresentação sobre como utilizar a técnica para projetar os diagramas de atividades. E uma breve apresentação sobre como inspecionar usabilidade através de diagramas de atividades também foi realizada para o grupo que utilizou a técnica MIT 3.

### 8.3.2. Execução do Estudo de Viabilidade da UDRT-AD e MIT 3

Nesta etapa, os participantes foram divididos em dois grupos, que ficaram em ambientes separados. Durante a execução, os artefatos do estudo foram entregues individualmente e não era permitida a comunicação entre os participantes. Os participantes do grupo da UDRT-AD receberam os seguintes artefatos: a técnica UDRT-AD, um cenário para construção dos diagramas e um questionário pós-modelagem. Os participantes do grupo da AT & MIT3, que primeiro devem modelar o Diagrama de Atividades (DA) e, em seguida, avaliar utilizando a MIT3, receberam os seguintes artefatos: instrução para construção do DA, um cenário para construção dos diagramas, a técnica de inspeção (MIT 3), uma planilha para a anotação das discrepâncias encontradas nos diagramas modelados e um questionário pós-modelagem. Os participantes de ambos os grupos utilizaram os mesmos cenários para modelar os diagramas de atividades.

Conforme explicado anteriormente, antes da realização do estudo foi realizado um treinamento sobre usabilidade. Este treinamento visava explicar conceitos de usabilidade, exemplos de aplicações com problemas de usabilidade, e como aplicar um método de inspeção. Neste treinamento, foi abordado o método de Avaliação Heurística (NIELSEN, 1994b) por ser o método base para as duas técnicas.

Durante a realização do estudo, cada grupo recebeu um cenário e, a partir deste cenário, os participantes realizaram a modelagem do diagrama de atividade. Os participantes do Grupo 1 (UDRT-AD) receberam a técnica UDRT-AD para auxiliar a modelar um diagrama de atividades visando à melhoria da usabilidade da aplicação através deste diagrama. Os participantes do Grupo 2 (AT&MIT3) receberam um documento que continha instruções, baseadas na literatura (BOOCH *et al.*, 2005), para que estes pudessem consultar, caso surgissem dúvidas nos passos de construção do DA. O cenário utilizado para modelar o DA tratava-se de uma reserva de um quarto de hotel. Neste cenário, o usuário da aplicação realizava desde o login na aplicação até a confirmação da reserva.

Após realizar a modelagem dos diagramas, os participantes do Grupo 1 (UDRT-AD) respondiam um Questionário Pós-Modelagem que está disponível em Valentim *et al.* (2017c). Este questionário tinha como objetivo verificar a percepção dos participantes ao modelar os diagramas utilizando a UDRT-AD. Já os participantes do Grupo 2 (AT&MIT3), antes de responder este questionário, realizavam uma inspeção de usabilidade no diagrama modelado, utilizando a técnica de inspeção MIT 3. Na inspeção, após detectar um problema de usabilidade utilizando a MIT 3, o participante anotava o defeito na planilha de discrepância.

Vale ressaltar ainda que, durante todo o processo de execução, os participantes realizaram suas atividades individualmente e não receberam nenhum auxílio dos pesquisadores envolvidos.

### 8.3.3. Análise dos Resultados do Estudo de Viabilidade da UDRT-AD e MIT 3

Por fim, nesta última etapa, foi realizada a atividade de análise dos resultados obtidos no estudo. Nesta etapa, os diagramas modelados pelos participantes foram analisados pelos pesquisadores. Para a realização deste processo, retirou-se o nome dos participantes e foi inserido um código para representá-los, a fim de não causar viés durante a análise dos resultados.

A verificação das respostas dos participantes ocorreu em três etapas. Na primeira etapa um pesquisador realizou a verificação da modelagem dos diagramas projetados pelos participantes a partir dos cenários. Na segunda etapa, outro pesquisador revisou a verificação feita, podendo também identificar e apontar novos problemas nos diagramas. Na última etapa, estava presente um terceiro especialista, com alto grau de experiência em modelos de projeto e os três pesquisadores debatiam sobre os problemas encontrados nos diagramas dos participantes.

Para a análise dos dados quantitativos obtidos, foram utilizados os seguintes indicadores: eficácia, tempo de modelagem, corretude e prevenção de erros de usabilidade.

O indicador eficácia foi calculado da seguinte forma:

$$Eficácia (Participante x) = \frac{(nAM - nAnC)}{(nAM - nAnC) + nAO}, \text{ onde:}$$

- Participante X é a referência ao participante do estudo;
- nAM é o número de atividades modeladas pelo Participante X no diagrama de atividades;
- nAnC é o número de atividades que não estão no contexto do cenário utilizado pelo Participante X para realizar a modelagem do diagrama de atividades;
- nAO é o número de atividades omissas no diagrama modelado pelo participante X, isto é, atividades que estão presentes no cenário, mas não estão presentes no diagrama modelado pelo Participante X.

O indicador tempo de modelagem é calculado em horas e representa o tempo total gasto por cada participante para modelar o diagrama de atividades.

O indicador corretude verifica o quão corretos os diagramas modelados estão. Os defeitos encontrados pelos pesquisadores nos diagramas modelados pelos participantes foram classificados de acordo com as categorias de defeitos de Travassos *et al.* (1999). Estas categorias têm sido aplicadas em técnicas para inspeção de modelos de projeto, como OORTs (TRAVASSOS *et al.*, 1999) e ActCheck (DE MELLO *et al.*, 2010). A Tabela 18 apresenta as cinco categorias, com a redefinição dos conceitos para cada tipo de defeito.

**Tabela 28. Categorias de Defeitos (adaptada de Travassos *et al.* (1999))**

<b>Categoria</b>	<b>Definição do defeito</b>
<b>Omissão</b>	Atividades ou elementos necessários foram omitidos no diagrama de atividades.
<b>Fato Incorreto</b>	Algumas atividades ou elementos no diagrama de atividades contradizem informações presentes no cenário utilizado.
<b>Inconsistência</b>	As atividades ou elementos em uma parte do diagrama de atividades estão inconsistentes com outras partes do diagrama de atividades.
<b>Ambiguidade</b>	As atividades ou elementos dos diagramas de atividades são ambíguos, isto é, é possível interpretar as atividades ou elementos de diferentes maneiras, podendo não levar a uma interpretação correta.
<b>Informação Estranha</b>	As atividades ou elementos são modelados, mas não são necessárias ou mesmo utilizadas.

Por fim, o indicador prevenção de possíveis problemas de usabilidade auxiliou a verificar se a técnica UDRT-AD ajuda a prevenir possíveis problemas de usabilidade através dos diagramas modelados. Para isto, realizou-se uma inspeção utilizando a MIT 3 nos diagramas de atividades modelados pelos participantes do grupo que utilizou a UDRT-AD como auxílio. Escolheu-se a MIT 3 como técnica para realizar a inspeção, pois esta é a técnica base para a UDRT-AD, e portanto, é razoável verificar se a UDRT-AD auxiliou a projetar a usabilidade através dos diagramas de atividades.

A seguir serão apresentados os resultados quantitativos do estudo empírico. A análise estatística foi realizada utilizando o *software* SPSS v.21 com  $\alpha = 0,05$ . A escolha da significância estatística e o teste não paramétrico de Mann-Whitney (MANN e WHITNEY, 1947) foi motivada pelo pequeno tamanho da amostra utilizada neste estudo (DYBA *et al.*, 2006), conforme sugerido por Wohlin *et al.* (2000). Utilizou-se os gráficos de boxplot para facilitar a visualização dos dados. Os resultados são apresentados a seguir.

#### **a. Eficácia**

A Tabela 29 apresenta os resultados encontrados em relação ao indicador de eficácia. Nesta tabela é mostrada a quantidade de atividades modeladas (3ª coluna), atividades que não estavam presentes na descrição textual (4ª coluna) e atividades omissas (5ª coluna). E, com base nestas informações, realizou-se o cálculo da eficácia para cada

participante. A Figura 55 apresenta os gráficos de boxplots comparando a eficácia do grupo que utilizou a UDRT-AD e a AT&MIT3.

Tabela 29. Dados quantitativos - Estudo de Viabilidade da UDRT-AD e MIT 3

	#P	nAM	nAnC	nAO	Eficácia	Tempo (h)
Grupo 1 UDRT-AD	01	16	01	01	0,94	0,90
	02	16	-	-	1,00	1,15
	03	16	-	01	0,94	1,03
	04	15	-	02	0,88	0,95
	05	29	04	-	1,00	1,00
	06	15	-	-	1,00	0,93
	07	16	-	-	1,00	1,20
	08	24	-	-	1,00	1,15
<b>Média</b>					<b>0,97</b>	<b>1,04</b>
Grupo 2 AT&MIT3	09	22	02	-	1,00	0,83
	10	17	-	-	1,00	0,60
	11	10	-	03	0,77	0,60
	12	11	-	01	0,92	0,82
	13	9	01	06	0,57	0,35
	14	21	-	-	1,00	0,82
	15	14	-	-	1,00	0,83
	16	13	-	02	0,87	0,63
<b>Média</b>					<b>0,89</b>	<b>0,69</b>
<b>Legenda: #P</b> – Participantes; <b>nAM</b> – Número de Atividades Modeladas; <b>nAnC</b> – Número de Atividades que não estão no contexto; <b>nAO</b> – Número de atividades que estão omissas no diagrama.						

Através da Figura 55, pode-se notar que a mediana do grupo da UDRT-AD está mais alta que a mediana do grupo AT&MIT3. Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p= 0,382$ ). Os resultados encontrados sugerem que as abordagens utilizadas (UDRT-AD e AT&MIT3) auxiliaram quase que de forma similar no processo de modelagem de um diagrama de atividades.

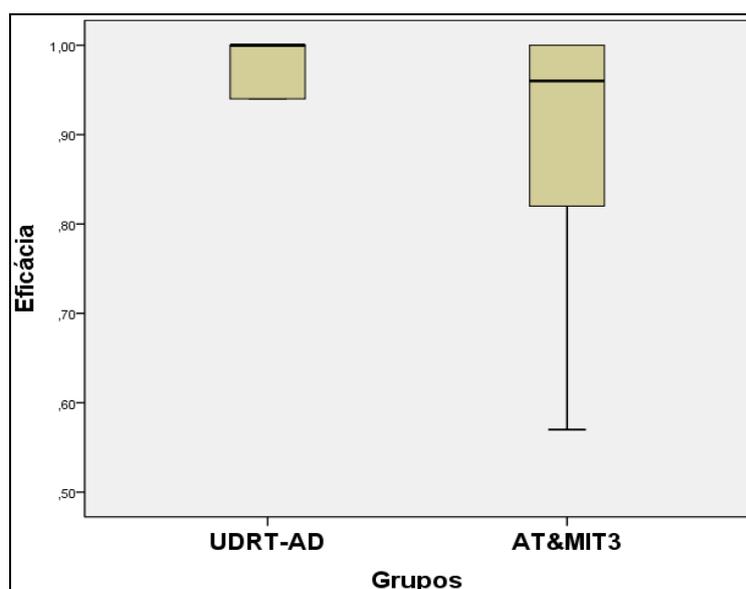


Figura 55. Boxplots da eficácia - Estudo de Viabilidade da UDRT-AD e MIT 3

### b. Tempo de Modelagem

A mesma análise foi aplicada para determinar se havia diferença significativa comparando o indicador tempo de modelagem entre os dois grupos. O gráfico de boxsplot comparando o tempo de modelagem entre os dois grupos é mostrado na Figura 56. Ao analisar a Figura 56, é possível perceber que o grupo da UDRT-AD obteve um tempo de modelagem superior ao grupo da AT&MIT3. Além disso, a mediana do grupo da UDRT-AD está mais alta que a mediana do grupo da AT&MIT3. Isto foi confirmado pelo teste de Mann-Whitney ( $p < 0,001$ ). Portanto, conclui-se que o tempo para modelar um diagrama de atividades visando à melhoria da usabilidade de aplicações interativas é significativamente maior utilizando a UDRT-AD do que utilizando a AT&MIT3.

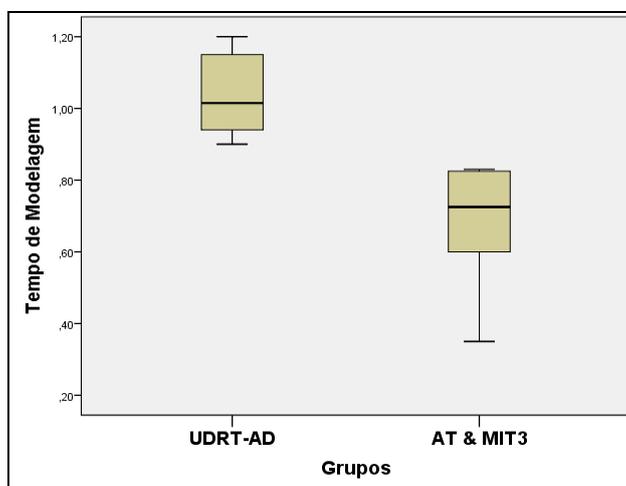


Figura 56. Boxplots do tempo de modelagem do Estudo de Viabilidade da UDRT-AD e MIT 3

### c. Corretude do Diagrama de Atividades

A Figura 57 apresenta o gráfico de boxplot comparando o indicador corretude entre os dois grupos (UDRT-AD e AT&MIT3).

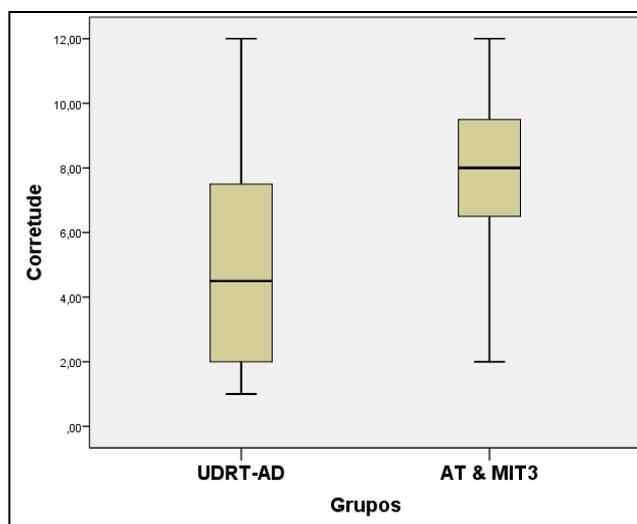


Figura 57. Boxplots com o comparativo da corretude

Nesta figura é possível perceber que a mediana do grupo da AT&MIT3 está mais alta que a mediana do grupo da UDRT-AD. Em outras palavras, foram encontrados menos defeitos nos diagramas dos participantes que utilizaram a UDRT-AD do que os que utilizaram a MIT3&AT. No entanto, quando se compara as duas amostras usando o teste de Mann-Whitney, não se encontra diferença estatística significativa entre os dois grupos ( $p = 0,161$ ). Portanto, conclui-se que não há uma diferença estatisticamente significativa entre o número de defeitos encontrados nos diagramas modelados utilizando a UDRT-AD e o número de defeitos encontrados nos diagramas modelados utilizando a AT&MIT3.

A Tabela 30 apresenta o número e os tipos de defeitos identificados nos diagramas modelados pelos participantes de cada grupo.

**Tabela 30. Número de defeitos – Estudo de viabilidade da UDRT-AD e MIT 3**

Tipos de Defeitos por participante e por grupos											
#P	Omissão		Fato Incorreto		Inconsistência		Ambiguidade		Informação Estranha		TDP
	At.	El.	At.	El.	At.	El.	At.	El.	At.	El.	
<b>GRUPO 1 – UDRT-AD</b>											
P01	01	-	-	01	-	-	-	-	01	-	03
P02	-	-	-	06	-	-	-	-	-	-	06
P03	01	-	-	-	-	-	-	-	-	-	01
P04	02	01	-	-	-	-	-	-	-	-	03
P05	-	-	-	-	-	-	-	-	04	02	06
P06	-	01	-	08	-	-	-	-	-	-	09
P07	-	-	-	12	-	-	-	-	-	-	12
P08	-	01	-	-	-	-	-	-	-	-	01
TD	04	03	0	27	0	0	0	0	05	02	<b>SDG= 41</b>
TTD	<b>07</b>		<b>27</b>		<b>0</b>		<b>0</b>		<b>07</b>		
<b>GRUPO 2 - AT&amp;MIT3</b>											
P09	-	-	-	06	-	-	-	-	02	-	08
P10	-	-	-	06	-	-	-	-	-	-	06
P11	03	-	-	03	-	-	-	-	-	-	10
P12	01	-	-	07	-	-	-	-	-	-	08
P13	06	01	-	01	-	-	-	-	01	-	09
P14	-	03	02	07	-	-	-	-	-	-	12
P15	-	-	-	07	-	-	-	-	-	-	07
P16	02	-	-	-	-	-	-	-	-	-	02
TD	12	04	02	41	0	0	0	0	03	0	<b>SDG=62</b>
TTD	<b>16</b>		<b>43</b>		<b>0</b>		<b>0</b>		<b>03</b>		

**Legenda:** At.: Atividades; El.: Elementos; TDP: Total de Defeitos por Participante; TD: Total de Defeitos por Atividades e Elementos; TTD: Total por Tipo de Defeitos; SDG: Soma dos Defeitos por Grupo.

Estes tipos de defeitos foram classificados segundo a categoria de tipos de defeitos sugerida por Travassos et al. (1999). Os tipos de defeitos foram divididos em problemas relacionados a Atividades e Elementos. Não foram encontrados defeitos de Inconsistência e de Ambiguidade em nenhum dos dois grupos. Em relação aos defeitos de Omissão, no

grupo 1 – UDRT-AD – foram encontrados menos defeitos (07 defeitos sendo 04 de atividades e 03 de elementos) que o grupo 2 – AT&MIT3 – (16 defeitos sendo 12 de atividades e 04 de elementos). O fato do grupo 1 (UDRT-AD) não possuir tantos defeitos de omissão, pode estar relacionado ao fato da técnica UDRT-AD utilizar procedimentos que guiaram os participantes na identificação das atividades e elementos que poderão ser utilizados nos diagramas, a partir do cenário utilizado.

No grupo 1 (UDRT-AD), em relação aos defeitos do tipo Fato Incorreto, foram encontrados menos defeitos (27 – todos os defeitos estavam relacionados a elementos) que no grupo 2 – AT&MIT3 – (43 defeitos – 02 de atividades e 41 de elementos). Ambos os grupos tiveram índices muito alto de problemas encontrados. É possível observar na Tabela 30 que houve poucos defeitos relacionados às atividades e houve muitos defeitos relacionados a elementos. Este tipo de defeito foi detectado porque os participantes não adicionavam colchetes nas condições de guarda (expressões que decidem qual será a próxima ação a ser executada). No grupo 2 (AT&MIT3), por utilizarem apenas um guia contendo instruções sobre a construção do diagrama de atividades, em quase todos os diagramas foram detectados defeitos deste tipo. Já no grupo 1 (UDRT-AD), metade dos diagramas modelados não apresentaram este tipo de defeito. No entanto, o número de defeitos encontrados deste tipo para o grupo 1 (UDRT-AD) foi alto, apesar da UDRT-AD conter uma heurística que guia os participantes a adicionarem colchetes nas condições de guarda. Isto mostra que esta heurística não foi tão eficaz na tentativa de evitar que este tipo de defeito ocorra. Portanto, a técnica deve ser melhorada neste sentido.

Em relação ao defeito Informação Estranha, no grupo 2 foram encontrados menos defeitos (03 defeitos – todos relacionados as atividades) que no grupo 1 (07 defeitos – 05 de atividades e 02 de elementos). Isto pode estar relacionado ao fato da técnica UDRT-AD possuir a diretriz D9 que sugere o seguinte: “Caso o usuário esteja em uma situação de erro deve haver atividades que ajude o usuário a corrigir o erro”. Na tentativa de sanar uma situação de erro, os participantes tentavam criar novas atividades e que exerciam as mesmas ações de outras atividades (duplicadas). Uma das sugestões de melhoria é alterar esta diretriz, de forma a deixá-la mais clara e não causar confusão nos participantes.

#### **d. Prevenção de Possíveis Problemas de Usabilidade**

Este indicador verificou se os DA's modelados pelos participantes do grupo da UDRT-AD poderiam apresentar possíveis problemas de usabilidade na aplicação final. Para isto, duas pesquisadoras com alto grau de experiência na indústria como analistas (8 e 10 anos de experiência), foram selecionadas para atuar com inspetoras dos DA's criados com

o auxílio da técnica UDRT-AD. Para realizar as inspeções, as analistas utilizaram a técnica MIT 3, pois esta foi a técnica base para criar a UDRT-AD. O processo de inspeção foi realizado em dois períodos de tempo: em um primeiro período cada inspetora avaliou dois diagramas e, em outro momento, mais dois diagramas. A Tabela 31 apresenta os resultados das inspeções realizadas nos diagramas modelados com o auxílio da técnica UDRT-AD.

**Tabela 31. Quantidade de problemas de usabilidade do Grupo 1**

Participantes da do Grupo 1	01	02	03	04	05	06	07	08	Total de Defeitos
Quantidade de Possíveis problemas de usabilidade	05	05	07	03	05	09	01	02	37

É possível perceber na Tabela 31 que, mesmo os participantes utilizando as Diretrizes de Usabilidade presentes na UDRT-AD, os diagramas modelados apresentaram um grande número de problemas de usabilidade que podem afetar a aplicação final. Estes problemas, se não forem tratados, podem causar uma má qualidade de uso para os usuários finais. Com isso, pode-se perceber que as diretrizes da UDRT-AD ainda não estão ajudando, da forma esperada, os designers no processo de projeto da usabilidade através dos diagramas elaborados. Portanto, a técnica necessita ser melhorada, no sentido de auxiliar mais os designers a projetarem as aplicações interativas com o mínimo de problemas de usabilidade possíveis, através dos DA's modelados.

#### **8.3.4. Discussão dos Resultados do Estudo de Viabilidade da UDRT-AD e MIT 3**

Em relação à eficácia, os resultados mostraram que apesar da UDRT-AD obter resultados melhores que os resultados da AT&MIT3, o teste estatístico mostrou que não há diferença estatística entre as duas abordagens. O gráfico do boxplot mostrou que os participantes que utilizaram a AT&MIT3 estavam mais distribuídos entre o primeiro e o terceiro quartil do gráfico. Já os participantes do grupo da UDRT-AD estão bem distribuídos entre o primeiro e o segundo quartil do gráfico. Isto pode indicar que utilizando a técnica UDRT-AD os designers tendem a ser mais eficazes, ou seja, tendem a não inserir tantas atividades/elementos que não estão presentes nas descrições textuais. Isto porque a UDRT-AD possivelmente auxilia a identificar mais as atividades/elementos presentes na descrição textual.

Em relação ao tempo de modelagem, os resultados mostraram que se demorou mais tempo utilizando a UDRT-AD do que a AT&MIT3. Conforme dito anteriormente, isto pode estar relacionado ao fato da técnica possuir procedimentos de técnica de leitura. Tais procedimentos fazem com que o designer de *software* siga o passo a passo para

construir o diagrama de atividades. No entanto, a técnica necessita ser melhorada para que haja uma diminuição neste tempo de modelagem, uma vez que isto pode afetar a adoção da UDRT-AD pela indústria, por demandar muito tempo no processo de desenvolvimento.

Com relação à corretude dos diagramas, foram identificados menos defeitos no grupo que utilizou a UDRT-AD para realizar a modelagem. A maioria dos defeitos encontrados, em ambos os grupos, era do tipo Fato Incorreto (27 defeitos no grupo 1 e 41 no grupo 2), ou seja, quando uma atividade/elemento era utilizada de maneira incorreta no diagrama modelado. Outro defeito identificado era do tipo Omissão (07 defeitos no grupo 1 e 16 no grupo 2), ou seja, quando uma atividade/elemento presente na descrição textual não é utilizada no diagrama. Apesar da UDRT-AD possuir as heurísticas e os exemplos de construção que auxiliam no processo de construção do diagrama, identificou-se muitos defeitos no grupo que a utilizou. Isto pode estar relacionado ao fato de algumas destas heurísticas terem causado confusão aos participantes durante o processo de modelagem.

Por fim, com relação à prevenção de problemas de usabilidade, onde somente foram avaliados os diagramas projetados pelo grupo que utilizou a UDRT-AD, verificou-se que os diagramas possuíam pontos com possíveis problemas de usabilidade que podem afetar diretamente a qualidade da aplicação. No diagrama modelado pelo participante 07, foi identificado apenas um possível problema que poderia afetar a usabilidade da aplicação interativa que está sendo modelada através deste diagrama. No diagrama modelado pelo participante 08 foram identificados 2 possíveis problemas, seguido pelo participante 04 com 3 possíveis problemas de usabilidade. Para diminuir ainda mais a quantidade de problemas nos diagramas de atividades serão criados exemplos de aplicação das diretrizes de usabilidade da técnica, visando mostrar como aplicar as diretrizes de usabilidade durante a modelagem do diagrama de atividades.

### **8.3.5. Ameaças à Validade do Estudo de Viabilidade da UDRT-AD e MIT 3**

A seguir serão apresentadas as ameaças relacionadas a este estudo que podem ser classificadas em quatro categorias: validade interna, validade externa, validade de conclusão e validade de constructo.

#### **a. Validade Interna**

Neste estudo foram consideradas quatro principais ameaças que representam um risco de interpretação imprópria dos resultados: (1) efeitos de treinamento; (2) classificação de experiência; (3) medição do tempo; (4) uso dos cenários. As três primeiras ameaças já foram abordadas nos outros estudos relatados. Em relação ao uso dos cenários (4), estes

poderiam afetar o estudo caso os participantes não entendessem o cenário. Esta ameaça foi minimizada utilizando cenários escritos em linguagem natural, onde as tarefas deste cenário estavam explícitas e de fácil entendimento. O fato dos cenários não terem sido validados antes é uma limitação deste estudo.

#### **b. Validade Externa**

Três ameaças foram consideradas: (1) os participantes do estudo foram estudantes de graduação ao invés de analistas ou designers de *software*; (2) o estudo foi realizado em ambiente acadêmico; e, (3) validade do artefato utilizado. Estas ameaças também já foram abordadas nos outros estudos relatados.

#### **c. Validade de Conclusão**

Neste estudo, o maior problema é o tamanho e a homogeneidade da amostra, por serem 16 estudantes e todos de uma mesma instituição. A quantidade de participantes não é ideal do ponto de vista estatístico. O tamanho reduzido de amostras é um problema conhecido em estudos de IHC e ES. Porém, este é um estudo inicial para verificar a viabilidade da técnica. Futuramente, pretende-se realizar novos estudos com mais participantes e que sejam profissionais da indústria. Devido a estes fatos, há limitação nos resultados deste estudo, sendo estes considerados indícios e não resultados conclusivos.

#### **d. Validade de Constructo**

Neste tipo de ameaça, considerou-se a definição dos indicadores: (i) eficácia; (ii) tempo de modelagem, (iii) corretude; e, (iv) prevenção de erros de usabilidade. Estes indicadores foram definidos pelos pesquisadores envolvidos nesta pesquisa e representam aspectos importantes ao se projetar um diagrama de atividade visando à melhoria da usabilidade da aplicação final.

### **8.4. Síntese do Capítulo**

Este capítulo relatou os resultados estudos cujo objetivo foi avaliar as técnicas MIT 3 e UDRT-AD. Os estudos destinaram-se a investigar sobre o projeto/criação e avaliação de modelos utilizados nas fases iniciais do processo de desenvolvimento, visando à usabilidade da aplicação final. A partir dos resultados dos estudos, pode-se notar que a MIT 3 obteve alta aceitação dos participantes e que a UDRT-AD obteve maior eficácia. Desta maneira, elas possivelmente são adequadas para a indústria de *software*.

## CAPÍTULO 9 - ESTUDOS EXPERIMENTAIS DA UDG

*Este capítulo apresenta os estudos experimentais que foram realizados para avaliar empiricamente e aperfeiçoar a tecnologia UDG.*

### 9.1. Introdução

Para avaliar e evoluir a UDG foram realizados estudos experimentais. Como pode ser visto na Figura 58, primeiramente, foi realizado um estudo piloto para avaliar se a UDG estava adequada para o uso nos estudos de viabilidade. Após melhorias serem realizadas na UDG, foram realizados dois estudos de viabilidade da UDG, em comparação a uma abordagem tradicional. Em uma abordagem tradicional, primeiramente os designers constroem os *mockups* e, em seguida, realizam uma inspeção de usabilidade nos *mockups* construídos utilizando uma técnica de inspeção de usabilidade, a Avaliação Heurística. Estes estudos experimentais serão descritos em mais detalhes nas subseções a seguir, assim como as melhorias realizadas na UDG.

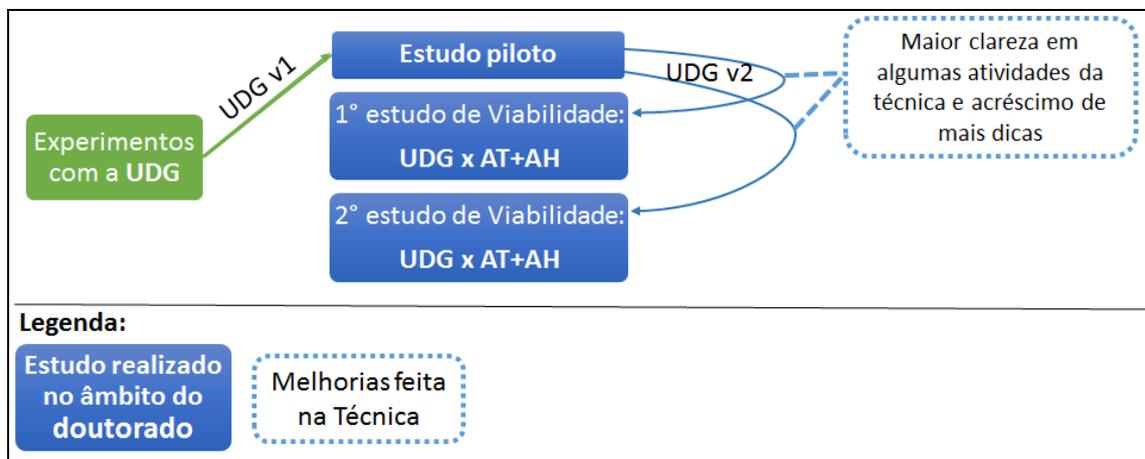


Figura 58. Estudos experimentais realizados com a UDG

### 9.2. Estudo piloto com a UDG

Um estudo piloto foi realizado a fim de verificar se a UDG e os demais artefatos do estudo eram viáveis para serem aplicados nos estudos de viabilidade. Este estudo piloto foi executado com a descrição textual de uma aplicação que apoia o planejamento de viagens, o gerenciamento de diferentes planos, tais como passagens aéreas, reservas de hotéis e reservas de aluguel de carro. A aplicação poderia ter uma representação *Web* ou *mobile*.

Alguns outros artefatos foram definidos para apoiar o estudo: Termo de Consentimento Livre e Esclarecido (TCLE), especificação da tecnologia UDG e um questionário pós-uso, disponível em Valentim *et al.* (2017c).

O estudo piloto foi conduzido com dois mestrandos da área de IHC e ES da UFAM. Os dois participantes receberam orientações de 10 minutos sobre como o estudo piloto ocorreria. As etapas do estudo piloto foram: (1) ler a descrição textual do sistema e extrair duas ou três possíveis tarefas que podem ser feitas no sistema, (2) usar a UDG como apoio para construir os *mockups* referente às tarefas extraídas da descrição textual da aplicação; os participantes durante esta etapa poderiam expressar qualquer dificuldade ou dúvida em relação à técnica e o moderador fazia a anotação dessas informações, e (3) responder um questionário pós-uso da UDG, no qual o participante dava a sua percepção sobre a técnica.

Durante a construção dos *mockups*, cada participante fez esta etapa individualmente utilizando a ferramenta de construção de *mockups* Balsamiq<sup>1</sup>. Após o estudo, os dois participantes entregaram os *mockups* construídos e o questionário pós-uso preenchido. E o moderador forneceu uma lista de problemas e dificuldades identificadas durante o estudo piloto. Alguns dos problemas e dificuldades com a técnica UDG enfrentados pelos participantes são descritos a seguir:

- Falta de significado ou clareza em alguns termos da UDG, como “migalha-de-pão” e “Densidade de Informação” (participante 1).
- Falta de orientação em como deixar claro onde o usuário está na aplicação na versão *mobile* (participante 1).
- Falta de clareza na atividade que pede para “agrupar elementos”, pois o participante interpretou como deixar um elemento bem próximo um do outro, e não no sentido de agrupar por tema ou tipo de conteúdo (participante 1).
- Questionou o que é considerado mensagem na atividade “Verificar se tela/página necessita mostrar mensagem” (participante 2).
- Questionou se alerta também é considerado mensagem na atividade "Verificar qual tipo de mensagem será oferecida pela aplicação" (participante 2).
- Falta de mais dicas referentes ao botão cancelar na heurística “Controle e liberdade do usuário” (participante 2).

Após a análise dos problemas e dificuldades enfrentados com a UDG e observados pelo moderador, foi feita uma análise específica dos dados qualitativos (comentários dos

<sup>1</sup><https://balsamiq.com/>

participantes) contidos nos questionários pós-uso. Algumas das evidências de dificuldades de uso da UDG foram: (i) dificuldade de entender alguns pontos da UDG; (ii) UDG é muito extensa; (iii) UDG tem vários itens e acaba não sendo tão fácil lembrar; e (iv) utilizar a UDG em aplicações maiores talvez demore um pouco.

Alguns comentários sobre a facilidade de uso da UDG foram: (i) após compreender a UDG, foi fácil usá-la para construir *mockups*, (ii) construir *mockup* com a UDG torna mais fácil focar na usabilidade. Sobre a Utilidade da técnica foi dito que: (i) a UDG tem itens sobre usabilidade que não conhecia e, (ii) sem a UDG, durante a construção de *mockups* muitas coisas passam despercebidas. E sobre a Intenção de Uso Futuro da UDG foi dito que: (i) se utilizaria a UDG em projetos reais para tirar alguma dúvida em relação a alguma questão de usabilidade, (ii) utilizaria a UDG em projetos no design da tela e funções e, (iii) em projetos reais pularia algumas partes da UDG que se tem mais domínio.

Melhorias em relação aos problemas e dificuldades acima relatados foram realizadas na UDG e uma nova versão da técnica foi criada. Além disso, a pesquisadora identificou no estudo piloto que era necessário o acréscimo de uma tela inicial contendo instruções para utilizar a UDG, com legenda, dizendo quando uma atividade tem dica ou não e quais são os elementos da técnica. Durante este estudo piloto, percebeu-se que a descrição textual estava extensa e que para o estudo de viabilidade seria mais interessante já deixar definidas as tarefas da aplicação que os participantes considerariam para construir os *mockups*. Portanto, melhorias foram feitas na descrição textual da aplicação. Além disso, como foi necessário explicar todo o procedimento do estudo no início para cada participante e isso se tornaria inviável no estudo de viabilidade, foi pensado em definir um documento contendo instruções de como proceder no estudo.

### 9.3. 1º Estudo de viabilidade da UDG

Com o objetivo de verificar se a nova versão da técnica UDG é viável e se esta técnica possui um desempenho melhor em comparação à uma abordagem tradicional de construção de *mockups* em conjunto com uma avaliação de usabilidade nestes *mockups*, realizou-se um estudo de viabilidade. Na abordagem tradicional, primeiramente os participantes deveriam projetar os *mockups* seguindo as seguintes orientações básicas de construção de *mockups*: (i) a partir das tarefas do sistema crie os *mockups* da aplicação; (ii) inclua os elementos (campos, botões, links, textos informativos, dentre outros) desta aplicação de acordo com a descrição do sistema; (iii) crie a navegação entre as telas projetadas; e (iv) crie as mensagens da aplicação. Em seguida, os participantes deveriam realizar uma inspeção de usabilidade nestes *mockups* utilizando a Avaliação Heurística (AH)

de Nielsen (1994b). Nesta tese, a abordagem tradicional combinada com a Avaliação Heurística será chamada de AT+AH. A abordagem AT+AH foi pensada e definida para ser comparada com a UDG, pois não encontramos na literatura uma abordagem que apoie a construção de *mockups* e apoie a avaliação da usabilidade simultaneamente.

As Subsubseções a seguir apresentam o detalhamento deste estudo de viabilidade.

### 9.3.1. Planejamento do 1º Estudo de Viabilidade da UDG

Este estudo experimental objetivou avaliar empiricamente a Corretude, Completude, Eficácia, Tempo de construção dos *mockups*, Estética dos *mockups*, Facilidade de Uso Percebida, Utilidade Percebida e Intenção de Uso Futuro da UDG quando comparada à abordagem tradicional em conjunto com a Avaliação Heurística (AT+AH).

O indicador Corretude verifica o quão correto os *mockups* projetados estão. Este indicador foi calculado da seguinte forma: quantidade de elementos com defeitos encontrados nos *mockups* elaborados (sem considerar defeitos de omissão) em relação ao total de elementos necessários (que deveria ter) nos *mockups*. Os defeitos encontrados pelos pesquisadores nos *mockups* projetados pelos participantes foram classificados de acordo com as categorias de defeitos de Travassos *et al.* (1999), especificamente os defeitos classificados como fato incorreto, inconsistência, omissão e ambiguidade. A Tabela 32 apresenta as quatro categorias, com a redefinição dos conceitos para cada tipo de defeito identificados em *mockups*.

Tabela 32. Categorias de Defeitos em *mockups* (adaptadas de Travassos *et al.* (1999))

<b>Categoria</b>	<b>Definição do defeito</b>
<b>Omissão</b>	Elementos necessários na aplicação que tem sido omitida do <i>mockup</i> .
<b>Fato Incorreto</b>	Elementos dos <i>mockups</i> que contradizem informações presentes na descrição da aplicação.
<b>Inconsistência</b>	Elementos dentro do <i>mockup</i> que estão inconsistentes com outras partes do(s) <i>mockup(s)</i> .
<b>Ambiguidade</b>	Elementos dos <i>mockups</i> que são ambíguos, ou seja, o usuário pode interpretar os elementos de maneiras diferentes e não poder conduzir a uma interpretação correta.

O indicador Completude verifica o quão completo os *mockups* projetados estão. Este indicador foi calculado da seguinte forma: quantidade de elementos omissos nos *mockups* em relação ao total de elementos necessários (que deveria ter) nos *mockups*.

O indicador Eficácia verifica o quão eficaz os participantes foram ao construir os *mockups*. Este indicador foi calculado como a quantidade de elementos com defeitos encontrados nos *mockups* elaborados (sem considerar defeitos de omissão) em relação ao total de elementos projetados nos *mockups*.

O indicador Tempo de Construção é calculado em minutos e representa o tempo total que cada participante levou para construir os *mockups*.

O indicador Estética avalia aspectos da estética visual dos *mockups*. Este indicador foi baseado em instrumento que contém critérios de avaliação da estética visual de *Websites*, chamado *Visual Aesthetics of Websites inventory* (VisAWI), proposto por Moshagen e Thielsch (2013). A Tabela 33 apresenta os critérios de avaliação (itens) do VisAWI adaptados para *mockups*. Há itens com característica positiva e negativa. Os itens com características negativas possuem o símbolo “(-)” ao lado.

**Tabela 33. Critérios de avaliação da estética visual dos *mockups* (adaptados de Moshagen e Thielsch (2013))**

<b>Fator</b>	<b>Critérios de avaliação (itens)</b>
<b>Simplicidade</b>	1. O layout parece muito denso (-)
	2. O layout é fácil de entender
	3. Tudo muito conectado
	4. A tela parece irregular (-)
	5. O layout parece bem estruturado
<b>Diversidade</b>	6. O layout é inventivo
	7. O design parece sem inspiração (-)
	8. O design é desinteressante (-)
<b>Habilidade</b>	9. O layout parece profissionalmente projetado
	10. O layout não é moderno (-)
	11. A tela foi projetada com cuidado

O indicador Facilidade de Uso Percebida define o grau em que uma pessoa acredita que usar uma tecnologia específica seria livre de esforço. O indicador Utilidade Percebida define o grau em que uma pessoa acredita que a tecnologia poderia melhorar seu desempenho. E por fim, o indicador Intenção de Uso Futuro define o grau em que uma pessoa acredita que utilizaria a tecnologia em projetos futuros. Estes três últimos indicadores (Facilidade de Uso Percebido, Utilidade Percebida e Intenção de Uso Futuro) são calculados pelo grau de aceitação fornecido pelos participantes sobre as afirmativas do Modelo de Aceitação de Tecnologia (*Technology Acceptance Model* - TAM 3), proposto por Venkatesh and Bala (2008) em uma escala de sete pontos. As possíveis respostas foram: concordo totalmente, concordo amplamente, concordo parcialmente, nem concordo nem discordo, discordo parcialmente, discordo amplamente e discordo totalmente.

#### **a. Hipóteses**

O experimento foi planejado e conduzido a fim de testar as seguintes hipóteses (nula e alternativa, respectivamente):

- H01: Não há diferença entre as técnicas UDG e AT+AH com relação ao indicador de corretude.
- HA1: Há diferença entre as técnicas UDG e AT+AH com relação ao indicador de corretude.
- H02: Não há diferença entre as técnicas UDG e AT+AH com relação ao indicador de completude.
- HA2: Há diferença entre as técnicas UDG e AT+AH com relação ao indicador de completude.
- H03: Não há diferença entre as técnicas UDG e AT+AH com relação ao indicador de eficácia.
- HA3: Há diferença entre as técnicas UDG e AT+AH com relação ao indicador de eficácia.
- H04: Não há diferença entre as técnicas UDG e AT+AH com relação ao indicador de tempo de construção.
- HA4: Há diferença entre as técnicas UDG e AT+AH com relação ao indicador de tempo de construção.
- H05: Não há diferença entre as técnicas UDG e AT+AH com relação ao indicador de estética dos *mockups*.
- HA5: Há diferença entre as técnicas UDG e AT+AH com relação ao indicador de estética dos *mockups*.
- H06: Não há diferença entre as técnicas UDG e AT+AH com relação ao indicador facilidade de uso percebida.
- HA6: Há diferença entre as técnicas UDG e AT+AH com relação ao indicador de facilidade de uso percebida.
- H07: Não há diferença entre as técnicas UDG e AT+AH com relação ao indicador utilidade percebida.
- HA7: Há diferença entre as técnicas UDG e AT+AH com relação ao indicador de utilidade percebida.
- H08: Não há diferença entre as técnicas UDG e AT+AH com relação ao indicador intenção de uso futuro.
- HA8: Há diferença entre as técnicas UDG e AT+AH com relação ao indicador de intenção de uso futuro.

### b. Contexto

A mesma descrição textual utilizada no estudo piloto foi usada neste estudo. O experimento foi conduzido com estudantes de graduação de Ciência da Computação da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio).

### c. Seleção de Variáveis

As variáveis independentes foram as técnicas de apoio a construção de *mockups* com boa usabilidade (UDG e AT+AH) e as variáveis dependentes foram indicadores de corretude, completude, eficácia, tempo de construção dos *mockups*, estética dos *mockups*, facilidade de uso percebida, utilidade percebida e intenção de uso futuro.

### d. Seleção de Participantes

Doze estudantes consentiram em participar do estudo. Eles assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) e preencheram um formulário de caracterização que mediu sua experiência em usabilidade e em desenvolvimento de *Software*. O formulário de caracterização foi aplicado para categorizar os participantes como tendo: nenhuma, baixa, média ou alta experiência em usabilidade, seguindo a mesma categorização dos estudos anteriores (ver Subsubseção 6.2.2 - f). De forma análoga, a experiência dos participantes em desenvolvimento de *software* foi classificada seguindo o mesmo padrão. Na Tabela 34, a terceira coluna e a quarta coluna apresentam a categorização de cada participante com relação às suas experiências em usabilidade e desenvolvimento de *software*, respectivamente. Em relação à participação de estudantes de graduação, não foi dada nota em relação ao desempenho neste estudo.

Tabela 34. Categorização dos participantes - 1º estudo de viabilidade da UDG

	#P	Experiência em Usabilidade	Experiência em Desenvolvimento de <i>Software</i>
Grupo 1 AT+AH	01	Nenhuma	Nenhuma
	02	Nenhuma	Nenhuma
	03	Nenhuma	Baixa
	04	Nenhuma	Nenhuma
	05	Nenhuma	Média
	06	Nenhuma	Nenhuma
Grupo 2 UDG	07	Baixa	Média
	08	Baixa	Nenhuma
	09	Nenhuma	Média
	10	Nenhuma	Nenhuma
	11	Nenhuma	Nenhuma
	12	Nenhuma	Nenhuma

### e. Projeto Experimental

Participantes foram divididos em dois grupos: grupo da UDG e grupo da AT+AH. Os participantes foram designados para cada grupo de maneira aleatória e balanceada, de acordo com o nível de experiência de cada participante..

### f. Instrumentação

Alguns artefatos também foram definidos para apoiar o experimento: formulário de caracterização e consentimento, especificação das técnicas UDG e AT+AH impressas em papel, instruções para o experimento, folhas em branco para a construção dos *mockups*, uma planilha de anotação das discrepâncias identificadas (para o grupo da AT+AH) e um questionário pós-uso, disponíveis em Valentim *et al.* (2017c). Além disso, foi utilizada a mesma descrição textual de uma aplicação de apoio ao viajante para ambos os grupos. Todos os artefatos foram validados antes do estudo ser executado por um pesquisador especialista em IHC e ES, para que não houvesse nenhum problema que inviabilizasse a execução do estudo.

### g. Preparação

Os participantes estavam cursando uma disciplina de Interação Humano-Computador. Além disso, para cada grupo, foi feita uma apresentação de 15 minutos sobre a técnica que o grupo aplicaria. Exemplos similares foram apresentados a fim de mostrar como usar ambas as técnicas (UDG e AT+AH).

### 9.3.2. Execução do 1º Estudo de Viabilidade da UDG

No início do experimento, um pesquisador agiu como moderador, sendo responsável por passar as informações sobre o estudo aos estudantes. Após isso, dividiu-se os participantes em dois grupos para usar cada uma das técnicas. Cada participante recebeu os artefatos do experimento. Durante o estudo, cada participante leu a descrição textual da aplicação, leu as instruções das técnicas e leram as técnicas para construir os *mockups*. O grupo da AT+AH também preencheu uma planilha com as discrepâncias (possíveis defeitos) identificados em seus próprios *mockups* após aplicar a Avaliação Heurística. Todos os participantes entregaram no final do experimento os *mockups* construídos, uma planilha contendo as discrepâncias identificadas (grupo da AT+AH), a anotação do tempo total do experimento. Eles também preencheram um questionário pós-uso. Cada inspetor realizou a

atividade de construção dos *mockups* individualmente. Durante esta atividade, os estudantes não receberam ajuda dos pesquisadores envolvidos no experimento.

### 9.3.3. Resultados Quantitativos do 1º Estudo de Viabilidade da UDG

Os *mockups* construídos pelos participantes foram analisados pelos pesquisadores. Para a realização deste processo, retirou-se o nome dos participantes e foi inserido um código para representá-los, a fim de não causar viés durante a análise dos resultados.

A verificação das respostas dos participantes ocorreu em três etapas. Na primeira etapa um pesquisador realizou a verificação dos *mockups* projetados pelos participantes a partir da descrição textual da aplicação em relação à quantidade de defeitos. Na segunda etapa, um especialista, com alto grau de experiência em modelos de projeto revisou a verificação feita, podendo também identificar e apontar novos defeitos nos *mockups*. Na última etapa, estava presente um terceiro especialista, também com alto grau de experiência em modelos de projeto, que revisou a verificação dos *mockups* feita pelo pesquisador em relação à quantidade de defeitos.

A Tabela 35 apresenta os dados quantitativos obtidos neste estudo de viabilidade que servirão de base para o cálculo de alguns dos indicadores definidos anteriormente. Para cada participante é apresentada a quantidade defeitos por tipo de defeito (omissão, fato incorreto, inconsistência e ambiguidade). Além disso, é apresentado o total de elementos com defeitos (sem considerar os defeitos de omissão), total de elementos projetados, total de elementos válidos e total de elementos necessários nos *mockups* de cada participante.

**Tabela 35. Dados quantitativos do 1º Estudo de Viabilidade da UDG**

#P	Tipos de defeitos por participante e por grupo				Total de elementos com defeitos	Total de elementos projetados	Total de elementos válidos	Total de elementos necessários
	Omissão	Fato Incorreto	Inconsistência	Ambiguidade				
<b>Grupo 1 – Abordagem tradicional + Avaliação Heurística (AT+AH)</b>								
01	1	0	0	0	0	15	15	16
02	0	1	1	0	2	21	19	19
03	2	0	1	0	1	12	11	13
04	8	0	0	0	0	4	4	12
05	3	1	0	0	1	37	36	39
06	0	0	0	0	0	15	15	15
<b>Grupo 2 – Usability Design Guidelines (UDG)</b>								
07	7	0	0	1	1	11	10	17
08	3	0	0	1	1	14	13	16
09	1	1	0	0	1	16	15	16
10	1	0	0	0	0	14	14	15
11	2	0	0	0	0	12	12	14
12	7	0	0	0	0	19	19	26

A Tabela 36 apresenta os resultados quantitativos dos indicadores de Corretude, Completude, Eficácia e Tempo de Construção do 1º estudo de viabilidade da UDG.

Tabela 36. Resultados dos indicadores do 1º estudo de viabilidade da UDG

	#P	Corretude (%)	Compleitude (%)	Eficácia (%)	Tempo de construção (minutos)
Grupo 1 AT+AH	01	100,00	93,75	100,00	98
	02	89,47	100,00	90,48	65
	03	92,31	84,62	91,67	108
	04	100,00	33,33	100,00	26
	05	97,44	92,31	97,30	98
	06	100,00	100,00	100,00	83
	<b>Média</b>	<b>96,54</b>	<b>84,00</b>	<b>96,57</b>	<b>79,67</b>
Grupo 2 UDG	07	94,12	58,82	90,91	102
	08	93,75	81,25	92,86	78
	09	93,75	93,75	93,75	85
	10	100,00	93,33	100,00	96
	11	100,00	85,71	100,00	60
	12	100,00	73,08	100,00	108
	<b>Média</b>	<b>96,94</b>	<b>80,99</b>	<b>96,25</b>	<b>88,17</b>

A análise estatística foi realizada utilizando o *software* SPSS v.23 com  $\alpha = 0,05$ . A escolha da significância estatística e o teste não paramétrico de Mann-Whitney (MANN e WHITNEY, 1947) foi motivada pelo pequeno tamanho da amostra utilizada neste estudo (DYBA *et al.*, 2006), conforme sugerido por Wohlin *et al.* (2000). Utilizou-se os gráficos de boxplot para facilitar a visualização dos dados. Os resultados dos indicadores são apresentados a seguir.

#### a. Corretude

Na Tabela 35 é mostrado o total de elementos com defeitos (sem considerar defeitos de omissão) (6ª coluna) e o total de elementos necessários (9ª coluna). O total de elementos necessários é o total de elementos que deveria haver nos *mockups*, que é calculado como o total de elementos válidos na aplicação mais o total de defeitos de omissão. Com base nestas informações, realizou-se o cálculo da corretude que é  $100\% - (\text{total de elementos com defeitos} / \text{total de elementos necessários})$ . A Tabela 36 (3ª coluna) apresenta os resultados encontrados em relação ao indicador de corretude. O grupo da AT+AH obteve uma corretude dos *mockups* de 96,54% e o grupo da UDG obteve corretude de 96,94%. A Figura 59 apresenta os gráficos de boxplots comparando a corretude dos *mockups* do grupo que utilizaram a AT+AH e a UDG.

Através da Figura 59, pode-se notar que a mediana do grupo da AT+AH está mais alta que a do grupo da UDG. Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p = 0,818$ ). Os resultados encontrados sugerem que as abordagens utilizadas (AT+AH e UDG) possibilitam uma corretude dos *mockups* quase que de forma similar (apoiar a hipótese H01).

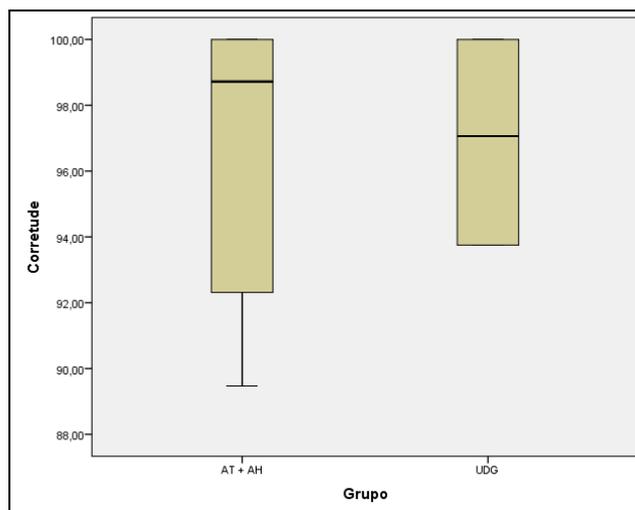


Figura 59. Boxplots da corretude - 1º estudo de viabilidade UDG

### b. Completude

Na Tabela 35 é mostrado o total de defeitos de omissão (2ª coluna) e o total de elementos necessários (9ª coluna). Com base nestas informações, realizou-se o cálculo da completude que é  $100\% - (\text{total de defeitos de omissão} / \text{total de elementos necessários})$ . A Tabela 36 (4ª coluna) apresenta os resultados encontrados em relação ao indicador de completude. A Figura 60 apresenta os gráficos de boxplots comparando a completude dos *mockups* do grupo que utilizaram a AT+AH e a UDG.

Através da Figura 60, pode-se notar que a mediana do grupo da AT+AH está mais alta que a mediana do grupo da UDG. Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p = 0,310$ ). Os resultados encontrados sugerem que as abordagens utilizadas (AT+AH e UDG) auxiliaram quase que de forma similar na completude dos *mockups* construídos (apoia a hipótese H02).

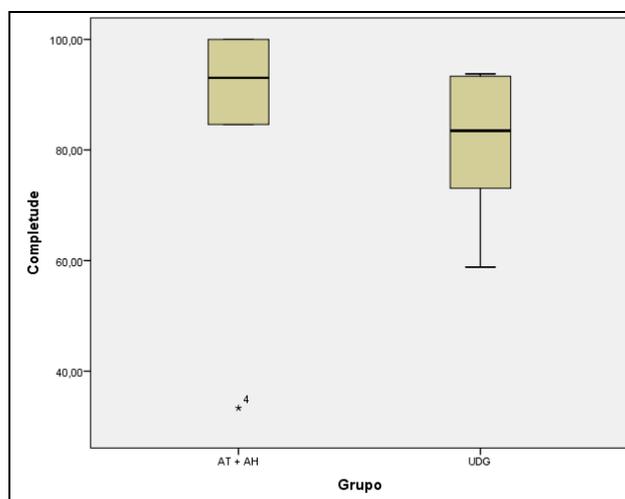


Figura 60. Boxplots da completude - 1º estudo de viabilidade UDG

### c. Eficácia

Na Tabela 35 é mostrado o total de elementos com defeitos (6ª coluna) e o total de elementos projetados (7ª coluna). Com base nestas informações, realizou-se o cálculo da eficácia que é  $100\% - (\text{total de elementos com defeitos} / \text{total de elementos projetados})$ . A Tabela 36 (5ª coluna) apresenta os resultados encontrados em relação ao indicador de eficácia. A Figura 61 apresenta os gráficos de boxplots comparando a eficácia dos *mockups* do grupo que utilizaram a AT+AH e a UDG.

Através da Figura 61, pode-se notar que a mediana do grupo da AT+AH está mais alta que a mediana do grupo da UDG.

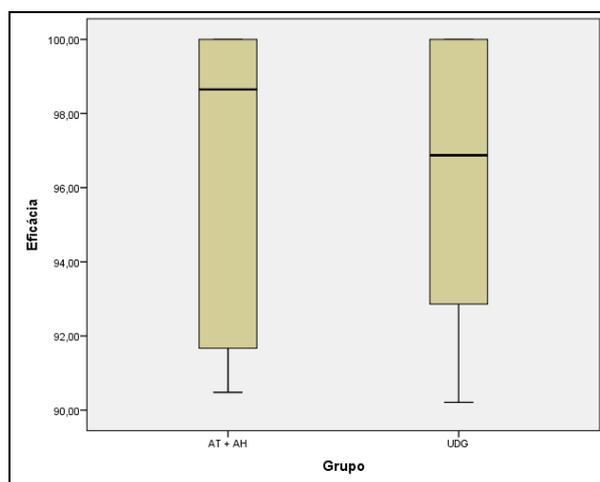


Figura 61. Boxplots da eficácia - 1º estudo de viabilidade UDG

Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p= 0,937$ ). Os resultados encontrados sugerem que as abordagens utilizadas (AT+AH e UDG) auxiliaram quase que de forma similar na construção de *mockups* (apoia a hipótese H03).

### d. Tempo de construção

A Tabela 36 (6ª coluna) apresenta os resultados encontrados em relação ao indicador de tempo de construção que representa o tempo em minutos de construção dos *mockups* por cada participante. A Figura 62 apresenta os gráficos de boxplots comparando o tempo de construção dos *mockups* do grupo que utilizaram a AT+AH e a UDG.

Através da Figura 62, pode-se notar que a mediana do grupo da AT+AH está mais alta que a do grupo da UDG. Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p= 818$ ). Os resultados encontrados sugerem que ambos os grupos (AT+AH e UDG) levaram quase que o mesmo tempo para construir os *mockups* (apoia a hipótese H04).

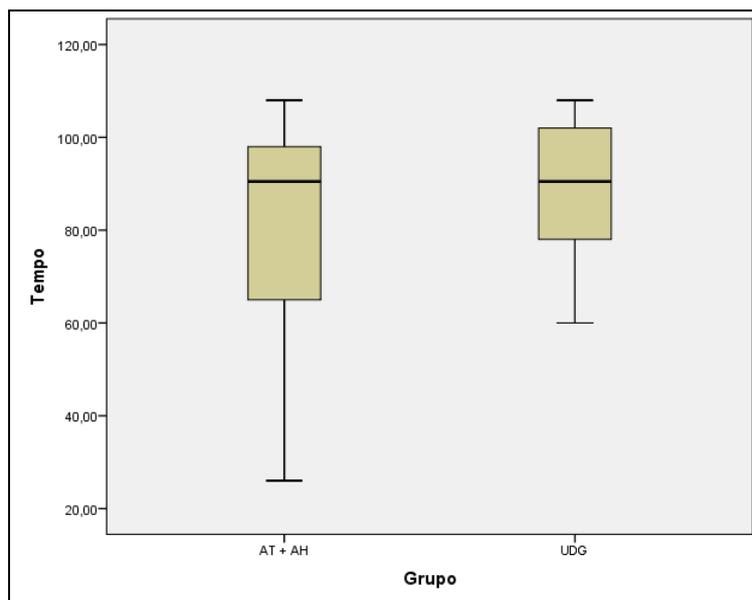


Figura 62. Boxplots do tempo de construção - 1º estudo de viabilidade UDG

### e. Estética

Este indicador foi calculado tendo como base a nota dada por 2 avaliadores em relação à estética dos *mockups* construídos neste estudo. Aos avaliadores foram mostrados os 12 *mockups* projetados por cada participante deste estudo e um formulário com a versão adaptada do VisAWI (ver Tabela 33). Foi solicitado aos participantes que indicassem seu nível de concordância para cada um dos itens. Os participantes classificavam como 0 se o item com característica positiva fosse verdadeiro e 1 se o item com característica positiva fosse falso. Além disso, os participantes classificavam como 1 se o item com característica negativa fosse verdadeiro e 0 se o item com característica negativa fosse falso. Utilizamos esta classificação pois desejávamos obter a quantidade de notas negativas em relação à estética visual dos *mockups*. Na Tabela 37 é mostrado o total de notas negativas em cada item do VisAWI em relação à estética dos *mockups* de participante por avaliador.

Através da Tabela 37, pode-se notar que os *mockups* projetados pelo grupo da AT+AH que obteve maior quantidade de notas negativas pelos dois avaliadores foram os *mockups* do participante P03. E no grupo da UDG, os *mockups* projetados pelo participante P11 foram os que obtiveram maior quantidade de notas negativas pelos dois avaliadores. Além disso, pode-se observar que, no geral, os *mockups* do grupo da UDG foram os que obtiveram menos notas negativas (apoiar a hipótese HA5). Especificamente, os *mockups* projetados pelo participante P08 não tiveram nenhuma nota negativa, mostrando que a estética do mesmo estava adequada para estes dois avaliadores.

Tabela 37. Resultados do indicador Estética - 1º estudo de viabilidade da UDG

#	Grupo	Mockups dos Participantes	Simplicidade					Diversidade			Habilidade			Total
			1	2	3	4	5	6	7	8	9	10	11	
Avaliador 1	Grupo 1 AT+AH	Mockups do P01	1	0	0	0	0	1	0	0	1	1	0	4
		Mockups do P02	0	1	1	1	1	0	0	1	1	1	1	8
		Mockups do P03	0	1	1	1	1	1	1	1	1	1	1	10
		Mockups do P04	0	1	1	1	1	0	0	1	1	1	1	8
		Mockups do P05	1	1	1	0	0	0	1	0	0	0	0	4
		Mockups do P06	1	0	0	0	0	1	0	0	1	1	0	4
	Grupo 2 UDG	Mockups do P07	1	1	1	0	1	1	1	1	1	1	1	10
		Mockups do P08	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P09	0	0	0	0	0	1	0	0	0	0	0	1
		Mockups do P10	0	0	0	0	0	1	0	0	0	1	1	3
		Mockups do P11	1	1	1	1	1	0	1	1	1	1	1	10
		Mockups do P12	1	1	1	1	1	0	0	0	0	0	0	5
Avaliador 2	Grupo 1 AT+AH	Mockups do P01	0	0	0	0	0	1	0	0	1	0	0	2
		Mockups do P02	0	0	0	1	1	0	1	0	1	0	1	5
		Mockups do P03	0	0	0	1	1	1	0	1	1	1	1	7
		Mockups do P04	0	1	1	1	1	1	0	1	1	1	1	9
		Mockups do P05	1	1	0	1	1	0	1	1	1	1	1	9
		Mockups do P06	1	1	1	1	1	0	1	1	1	1	1	10
	Grupo 2 UDG	Mockups do P07	0	0	0	1	1	0	0	0	1	0	1	4
		Mockups do P08	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P09	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P10	1	0	0	1	1	0	0	1	1	1	1	7
		Mockups do P11	1	1	1	1	1	1	1	1	1	1	1	11
		Mockups do P12	1	0	0	1	1	0	0	0	1	1	0	5

#### f. Facilidade de Uso

A Figura 63 apresenta as percepções dos participantes com relação à facilidade de uso da AT+AH e UDG. Pode-se observar que quatro dos seis participantes da UDG (P07, P10, P11 e P12) discordaram da afirmativa “*Interagir com a tecnologia não exige muito do meu esforço mental*”, mostrando que a UDG exige uma concentração maior por parte do usuário para ser utilizada. Nesta mesma afirmativa o participante P03 que utilizou a AT+AH discordou totalmente (apoiar a hipótese HA6).

Além disso, o participante P11 discordou totalmente da afirmativa “*Considero fácil utilizar a tecnologia para fazer o que eu quero que ela faça, apoiar a construção de mockups com boa usabilidade*”, destacando a dificuldade que ele teve ao construir *mockups* com boa usabilidade utilizando a UDG. Este mesmo participante, obteve uma corretude de 100% nos *mockups* e eficácia de 100%. No entanto, este participante foi o que obteve maior quantidade de notas negativas em relação à estética dos seus *mockups*. Isto pode indicar que a UDG precisa ser melhorada no sentido de apoiar mais a construção de *mockups* com boa usabilidade e *mockups* com uma estética visual mais agradável.

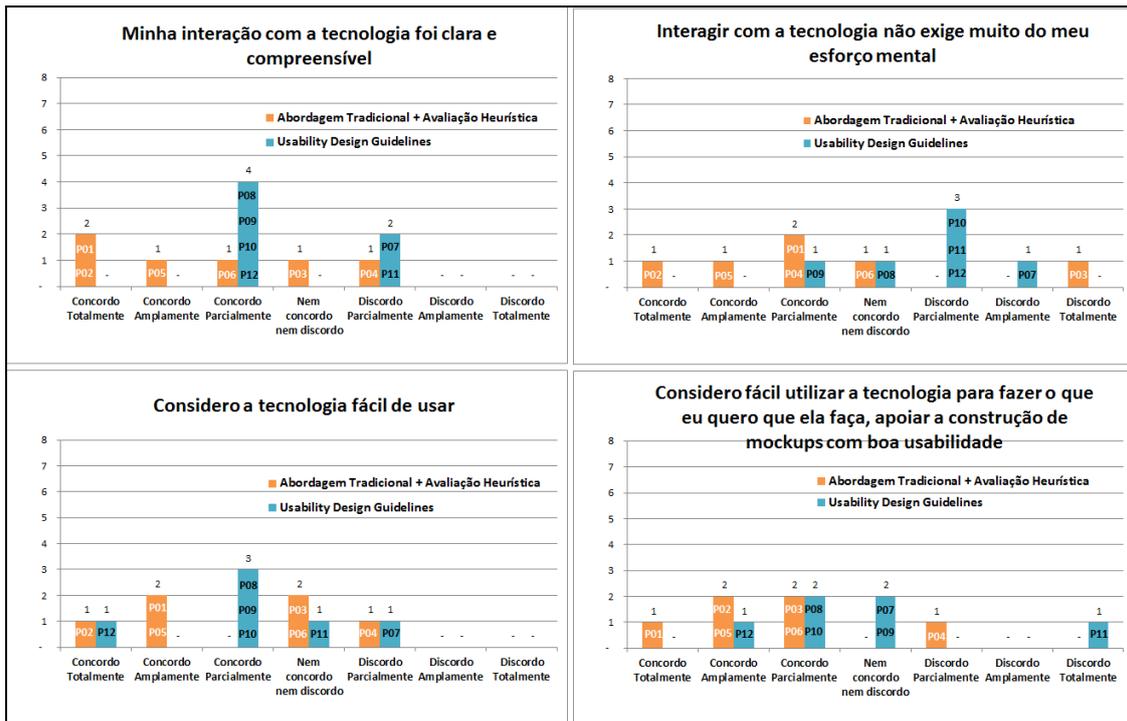


Figura 63. Percepção sobre Facilidade de Uso - 1º estudo de viabilidade UDG

g. Utilidade Percebida

A Figura 64 apresenta as percepções dos participantes com relação à utilidade percebida da AT+AH e UDG.

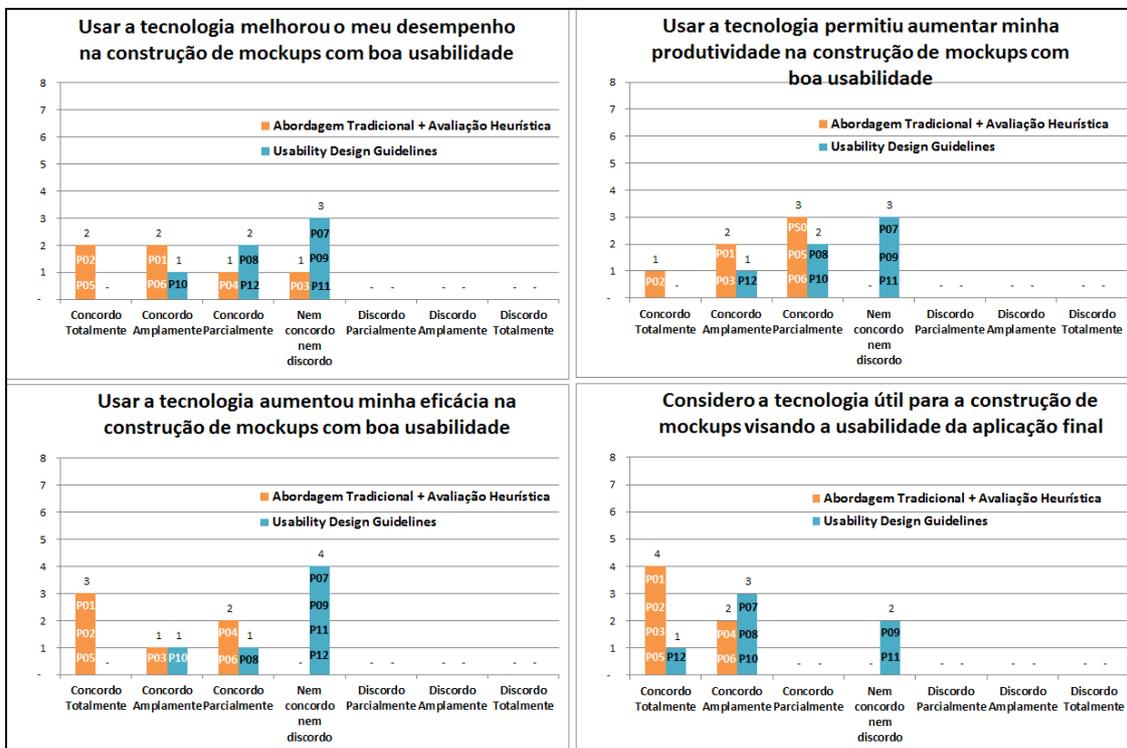


Figura 64. Percepção sobre Utilidade - 1º estudo de viabilidade UDG

Verificou-se que nenhum dos participantes, tanto da AT+AH quanto da UDG, discordou das afirmativas deste indicador. Isto pode indicar que todos os participantes deste estudo consideram que ambas as abordagens são úteis para construir mockups com boa usabilidade (apoia a hipótese H07).

#### h. Intenção de Uso Futuro

A Figura 65 apresenta as percepções dos participantes com relação à intenção de uso futuro da AT+AH e UDG. Pode-se observar que nenhum dos participantes, tanto da AT+AH quanto da UDG, discordou das afirmativas deste indicador. Isto pode indicar que os participantes pretendem usar futuramente estas abordagens para a construção de *mockups* com boa usabilidade (apoia a hipótese H08).

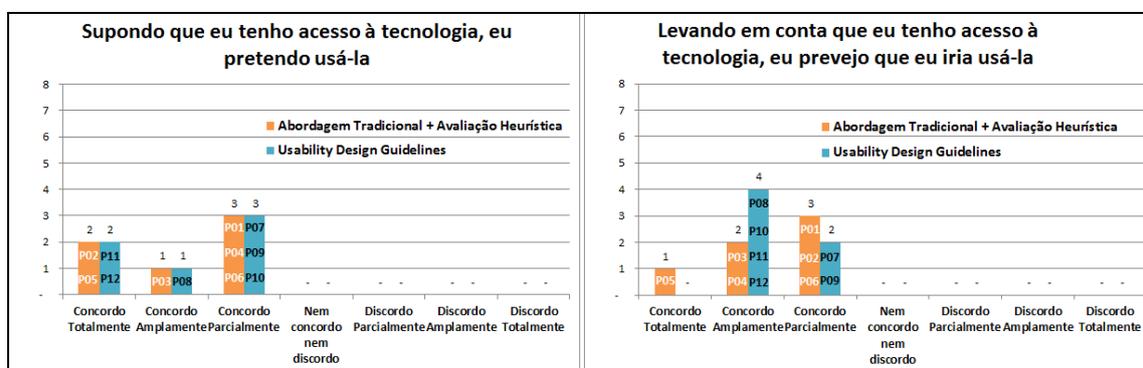


Figura 65. Percepção sobre Intenção de Uso - 1º estudo de viabilidade UDG

#### 9.3.4. Resultados Qualitativos do 1º Estudo de Viabilidade da UDG

Realizou-se uma análise dos dados qualitativos (comentários adicionais dos participantes) contidos nos questionários pós-uso. Como o intuito é evoluir a UDG, analisou-se somente os comentários feitos pelos participantes que utilizaram esta tecnologia.

O participante P11 afirmou: “*tive dificuldades em entender a UDG*”. Porém este participante não informou quais foram suas dificuldades, o que impede de identificarmos onde a UDG precisa ser melhorada. Além disso, este participante comentou “*acredito que pelo computador será muito mais agradável ou em trabalhos de grupo*”. Percebeu-se através deste comentário que a técnica não foi tão agradável de ser utilizada impressa em papel. Portanto, no próximo estudo pretende-se utilizar a técnica no computador.

O participante P12 disse que “*A UDG tem padrões fáceis de seguir, porém por isso torna difícil aplicar algo fora desse padrão. Poderia ser mais aberta a estilos diferentes*?”. No entanto, as informações sobre estilos são fornecidas apenas como dicas. O projetista pode seguir um estilo que melhor lhe agrada.

O participante P07 sugeriu “*Ter uma explicação mais enxuta e objetiva que ajude quem está iniciando agora*”. A UDG é extensa, pois possui muitas dicas de usabilidade. No entanto, o projetista não precisa aplicar todas elas, apenas as que são convenientes ao contexto da aplicação. Portanto, diminuir a técnica implicaria na exclusão destas dicas, o que não é viável, uma vez que estas dicas apoiam a construção dos *mockups* com boa usabilidade.

Como a quantidade de participantes no 1º estudo de viabilidade foi pequena e não se obteve dados qualitativos satisfatórios para a melhoria da UDG, um novo estudo de viabilidade foi realizado. Este estudo será descrito a seguir.

## **9.4. 2º Estudo de viabilidade da UDG**

No 2º estudo de viabilidade a UDG foi comparada à abordagem tradicional de construção de *mockups* em conjunto com as heurísticas de usabilidade proposta por Nielsen (1994b), como no estudo anterior. As Subsubseções a seguir apresentam o detalhamento deste estudo de viabilidade.

### **9.4.1. Planejamento do 2º Estudo de Viabilidade da UDG**

Este estudo experimental objetivou validar empiricamente a Corretude, Completude, Eficácia, Tempo de construção dos *mockups*, Estética dos *mockups*, Facilidade de Uso Percebida, Utilidade Percebida e Intenção de Uso Futuro da UDG quando comparada à abordagem tradicional em conjunto com a Avaliação Heurística (AT+AH).

#### **a. Hipóteses**

As hipóteses deste estudo são as mesmas definidas no 1º estudo de viabilidade da UDG (Subseção 9.3.1 - a).

#### **b. Contexto**

O experimento foi executado com uma descrição textual de um sistema que gerencia a compra de passagens aéreas para servidores de uma universidade. Foi utilizada esta descrição, pois ela é mais simples que a descrição do sistema utilizada no 1º estudo de viabilidade. Foi feita esta mudança de descrição textual entre os estudos para que os participantes tivessem mais tempo para projetar os *mockups*. O experimento foi conduzido com estudantes de graduação da disciplina de Análise e Projeto de Sistemas do curso de Ciência da Computação da Universidade Federal do Amazonas (UFAM).

### c. Seleção de Variáveis

As variáveis independentes foram as técnicas de apoio a construção de *mockups* com boa usabilidade (UDG e AT+AH) e as variáveis dependentes foram indicadores de corretude, completude, eficácia, tempo de construção dos *mockups*, estética dos *mockups*, facilidade de uso percebida, utilidade percebida e intenção de uso futuro.

### d. Seleção de Participantes

Trinta e cinco estudantes consentiram em participar do estudo. Eles assinaram o TCLE e preencheram um formulário de caracterização que mediu sua experiência em usabilidade e em desenvolvimento de *Software*. O formulário de caracterização foi aplicado para categorizar os participantes como tendo: nenhuma, baixa, média ou alta experiência em usabilidade, seguindo a mesma categorização dos estudos anteriores (ver Subsubseção 6.2.2 - f). De forma análoga, a experiência dos participantes em desenvolvimento de *software* foi classificada seguindo o mesmo padrão. Na Tabela 38, a segunda, terceira, quinta e sexta coluna apresentam a categorização de cada participante em ambos os grupos com relação às suas experiências em usabilidade e desenvolvimento de *software*, respectivamente. Em relação à participação de estudantes de graduação, eles não receberam nota em relação ao seu desempenho neste estudo.

Tabela 38. Categorização dos participantes - 2º estudo de viabilidade da UDG

Grupo 1 - AT+AH			Grupo 2 - UDG		
#P	Experiência em Usabilidade	Experiência em Desenvolvimento de <i>Software</i>	#P	Experiência em Usabilidade	Experiência em Desenvolvimento de <i>Software</i>
01	Alta	Média	18	Baixa	Baixa
02	Média	Alta	19	Nenhuma	Média
03	Baixa	Média	20	Nenhuma	Baixa
04	Baixa	Baixa	21	Baixa	Nenhuma
05	Baixa	Nenhuma	22	Baixa	Nenhuma
06	Baixa	Nenhuma	23	Baixa	Nenhuma
07	Baixa	Nenhuma	24	Baixa	Nenhuma
08	Baixa	Nenhuma	25	Baixa	Nenhuma
09	Baixa	Nenhuma	26	Baixa	Nenhuma
10	Baixa	Nenhuma	27	Baixa	Nenhuma
11	Baixa	Nenhuma	28	Baixa	Nenhuma
12	Baixa	Nenhuma	29	Baixa	Nenhuma
13	Baixa	Nenhuma	30	Baixa	Nenhuma
14	Baixa	Nenhuma	31	Nenhuma	Nenhuma
15	Baixa	Nenhuma	32	Nenhuma	Nenhuma
16	Nenhuma	Nenhuma	33	Nenhuma	Nenhuma
17	Nenhuma	Baixa	34	Baixa	Nenhuma
			35	Baixa	Nenhuma

### e. Projeto Experimental

Participantes foram divididos em dois grupos: grupo da UDG e grupo da AT+AH. Os participantes foram designados para cada grupo de maneira aleatória e balanceada, de acordo com o nível de experiência de cada participante. O grupo da AT+AH foi composto por 17 participantes e o grupo da UDG por 18 participantes.

### f. Instrumentação

Os artefatos deste estudo foram os mesmos do 1º estudo de viabilidade. No entanto, a UDG foi utilizada no computador através de um pdf clicável. Além disso, foi utilizado a mesma descrição textual de uma aplicação de apoio ao viajante para ambos os grupos, porém a descrição foi simplificada.

### g. Preparação

Os participantes estavam cursando uma disciplina de Análise e Projeto de Sistema e a maioria deles já tinham participado da disciplina de Interação Humano-Computador em um semestre anterior.

#### 9.4.2. Execução do 2º Estudo de Viabilidade da UDG

No início do experimento, um pesquisador agiu como moderador, sendo responsável por passar as informações sobre o estudo aos estudantes. Além disso, três outros pesquisadores estiveram apoiando a execução do experimento, realizando atividades como entrega de artefatos, conferência da assinatura dos nomes e anotação do tempo. Dividiu-se os participantes em dois grupos para usar cada uma das técnicas. Cada participante recebeu os artefatos do experimento. Durante o estudo, cada participante leu a descrição textual da aplicação, leu as instruções das técnicas e leram as técnicas para construir os *mockups*. O grupo da AT+AH também preencheu uma planilha com as discrepâncias (possíveis defeitos) identificados em seus próprios *mockups* após aplicar a Avaliação Heurística. Todos os participantes entregaram no final do experimento os *mockups* construídos, uma planilha contendo as discrepâncias identificadas (grupo da AT+AH), a anotação do tempo total do experimento. Eles também preencheram um questionário pós-uso. Cada inspetor realizou a atividade de construção dos *mockups* individualmente. Durante esta atividade, os estudantes não receberam ajuda dos pesquisadores envolvidos no experimento.

### 9.4.3. Resultados Quantitativos do 2º Estudo de Viabilidade da UDG

Os *mockups* construídos pelos participantes foram analisados pelos pesquisadores. Para a realização deste processo, retirou-se o nome dos participantes e foi inserido um código para representá-los, a fim de não causar viés durante a análise dos resultados. Para a análise dos dados quantitativos obtidos, foram utilizados os seguintes indicadores: correteza, completude, eficácia, tempo de construção, estética, facilidade de uso, utilidade percebida e intenção de uso futuro. A definição de cada um destes indicadores foi apresentada na Subseção 9.3.1.

A Tabela 39 apresenta os dados quantitativos obtidos neste estudo de viabilidade que servirão de base para o cálculo dos indicadores definidos anteriormente.

**Tabela 39. Dados quantitativos do 2º Estudo de Viabilidade da UDG**

#P	Tipos de defeitos por participante e por grupo				Total de elementos com defeitos	Total de elementos projetados	Total de elementos válidos	Total de elementos necessários
	Omissão	Fato Incorreto	Inconsistência	Ambiguidade				
<b>Grupo 1 – Abordagem tradicional + Avaliação Heurística (AT+AH)</b>								
01	3	0	0	1	4	21	20	23
02	1	2	0	0	3	22	20	21
03	0	1	0	0	1	25	24	24
04	0	1	0	0	1	28	27	27
05	1	1	0	0	2	25	24	25
06	0	1	1	0	2	32	30	30
07	1	1	0	0	2	20	19	20
08	1	0	0	0	1	29	29	30
09	0	0	0	0	0	21	21	21
10	1	0	0	0	1	23	23	24
11	0	0	0	0	0	19	19	19
12	1	0	0	0	1	15	15	16
13	3	0	0	0	3	21	21	24
14	1	0	0	0	1	16	16	17
15	0	0	0	0	0	19	19	19
16	3	0	0	0	3	19	19	22
17	1	0	0	0	1	29	29	30
<b>Grupo 2 – Usability Design Guidelines (UDG)</b>								
18	3	0	0	0	3	13	13	16
19	1	1	0	0	2	16	15	16
20	1	2	0	0	3	30	28	29
21	1	0	0	0	1	19	19	20
22	2	0	0	0	2	25	25	27
23	0	0	0	0	0	22	22	22
24	1	0	0	0	1	17	17	18
25	2	0	0	1	3	11	10	12
26	0	0	0	0	0	24	24	24
27	4	0	0	0	4	34	34	38
28	2	0	0	0	2	12	12	14
29	0	0	0	0	0	24	24	24
30	1	0	0	0	1	29	29	30
31	1	0	0	0	1	28	28	29
32	1	0	0	0	1	35	35	36
33	1	0	0	0	1	25	25	26
34	1	0	0	0	1	18	18	19
35	0	0	0	0	0	47	47	47

Para cada participante é apresentada a quantidade defeitos por tipo de defeito (omissão, fato incorreto, inconsistência e ambiguidade). Além disso, é apresentado o total de elementos com defeitos (sem considerar os defeitos de omissão), total de elementos projetados, total de elementos válidos e total de elementos necessários nos mockups de cada participante.

A Tabela 40 apresenta os resultados quantitativos dos indicadores deste estudo.

**Tabela 40. Resultados dos indicadores do 2º estudo de viabilidade da UDG**

	#P	Corretude (%)	Completo (%)	Eficácia (%)	Tempo de construção (min)
Grupo 1 - AT+AH	01	95,65	86,96	95,24	26
	02	90,48	95,24	90,91	41
	03	95,83	100,00	96,00	34
	04	96,30	100,00	96,43	54
	05	96,00	96,00	96,00	30
	06	93,33	100,00	93,75	32
	07	95,00	95,00	95,00	34
	08	100,00	96,67	100,00	55
	09	100,00	100,00	100,00	39
	10	100,00	95,83	100,00	22
	11	100,00	100,00	100,00	28
	12	100,00	93,75	100,00	22
	13	100,00	87,50	100,00	37
	14	100,00	94,12	100,00	39
	15	100,00	100,00	100,00	38
	16	100,00	86,36	100,00	25
	17	100,00	96,67	100,00	51
	<b>Média</b>	<b>97,80</b>	<b>95,53</b>	<b>97,84</b>	<b>35,71</b>
Grupo 2 - UDG	18	100,00	81,25	100,00	24
	19	93,75	93,75	93,75	34
	20	93,10	96,55	93,33	29
	21	100,00	95,00	100,00	28
	22	100,00	92,59	100,00	24
	23	100,00	100,00	100,00	19
	24	100,00	94,44	100,00	32
	25	91,67	83,33	90,91	19
	26	100,00	100,00	100,00	29
	27	100,00	89,47	100,00	68
	28	100,00	85,71	100,00	36
	29	100,00	100,00	100,00	34
	30	100,00	96,67	100,00	29
	31	100,00	96,55	100,00	38
	32	100,00	97,22	100,00	60
	33	100,00	96,15	100,00	47
	34	100,00	94,74	100,00	50
	35	100,00	100,00	100,00	62
	<b>Média</b>	<b>98,81</b>	<b>94,08</b>	<b>98,78</b>	<b>36,78</b>

A análise estatística foi realizada utilizando o *software* SPSS v.23 com  $\alpha = 0,05$ . Utilizou-se os gráficos de boxplot para facilitar a visualização dos dados. Os resultados dos indicadores são apresentados a seguir.

### a. Corretude

A Tabela 40 (3ª coluna) apresenta os resultados encontrados em relação ao indicador de corretude. O grupo da AT+AH obteve uma corretude dos mockups de 97,80% e o grupo da UDG obteve corretude de 98,81%. A Figura 66 apresenta os gráficos de boxplots comparando a corretude dos mockups do grupo que usaram a AT+AH e a UDG.

Através da Figura 66, pode-se notar que a mediana do grupo da AT+AH está no mesmo nível da mediana do grupo da UDG. No entanto, houve uma dispersão mínima do grupo da UDG em relação à corretude. Isto pode indicar que a UDG guiou mais a construção de *mockups* corretos. Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p=0,318$ ). Os resultados encontrados sugerem que as abordagens utilizadas (AT+AH e UDG) possibilitam uma corretude dos *mockups* quase que de forma similar (apoia a hipótese H01).

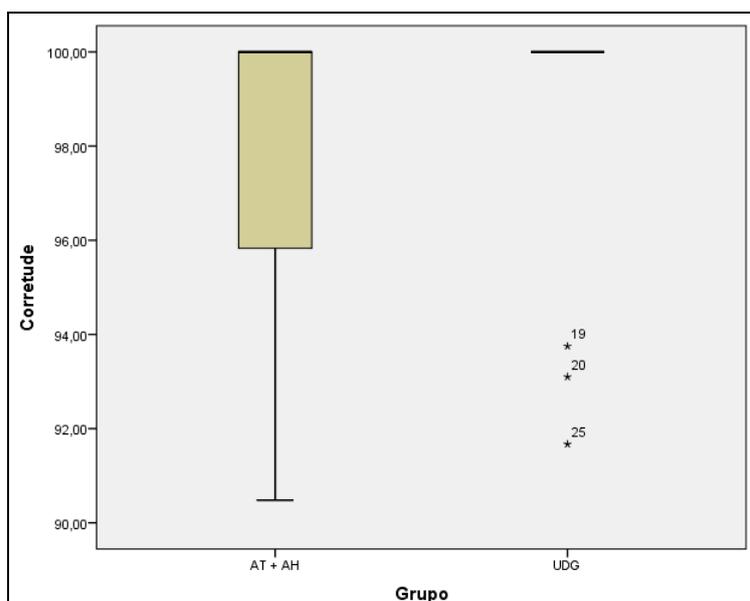


Figura 66. Boxplots da corretude - 2º estudo de viabilidade UDG

### b. Completude

A Tabela 40 (4ª coluna) apresenta os resultados encontrados em relação ao indicador de completude. O grupo da AT+AH obteve uma completude dos *mockups* de 95,53% e o grupo da UDG obteve completude de 94,08%. A Figura 67 apresenta os gráficos de boxplots comparando a completude dos *mockups* do grupo que utilizaram a AT+AH e a UDG.

Através da Figura 67, pode-se notar que a mediana do grupo da AT+AH está quase no mesmo nível que a mediana do grupo da UDG.

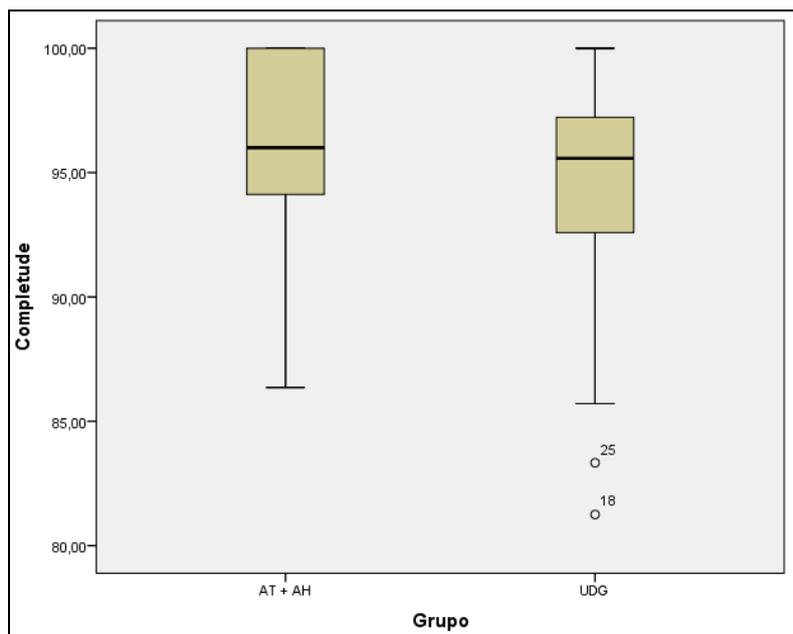


Figura 67. Boxplots da completude - 2º estudo de viabilidade UDG

Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p= 0,443$ ). Os resultados encontrados sugerem que as abordagens utilizadas (AT+AH e UDG) possibilitam uma completude dos mockups quase que de forma similar (apoia a hipótese H02).

### c. Eficácia

A Tabela 40 (5ª coluna) apresenta os resultados encontrados em relação ao indicador de eficácia. O grupo da AT+AH obteve uma eficácia dos mockups de 97,84% e o grupo da UDG obteve eficácia de 98,78%. A Figura 68 apresenta os gráficos de boxplots comparando a eficácia dos mockups do grupo que utilizaram a AT+AH e a UDG.

Através da Figura 68, pode-se notar que a mediana do grupo da AT+AH está no mesmo nível da mediana do grupo da UDG. No entanto, houve uma dispersão mínima do grupo da UDG em relação à eficácia. Isto pode indicar que o grupo da UDG obteve uma maior eficácia na construção de *mockups*. Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p= 0,318$ ). Os resultados encontrados sugerem que as abordagens utilizadas (AT+AH e UDG) possibilitam uma eficácia na construção dos *mockups* quase que de forma similar (apoia a hipótese H03).

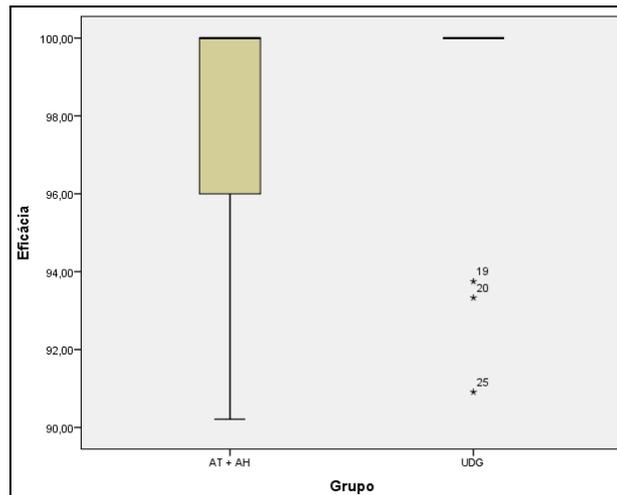


Figura 68. Boxplots da eficácia - 2º estudo de viabilidade UDG

#### d. Tempo de construção

A Tabela 40 (6ª coluna) apresenta os resultados encontrados em relação ao indicador de tempo de construção que representa o tempo em minutos de construção dos mockups por cada participante. O grupo da AT+AH levou em média 35,71 minutos para construir mockups com boa usabilidade e o grupo da UDG levou em média 36,78 minutos. A Figura 69 apresenta os gráficos de boxplots comparando o tempo de construção dos mockups do grupo que utilizaram a AT+AH e a UDG.

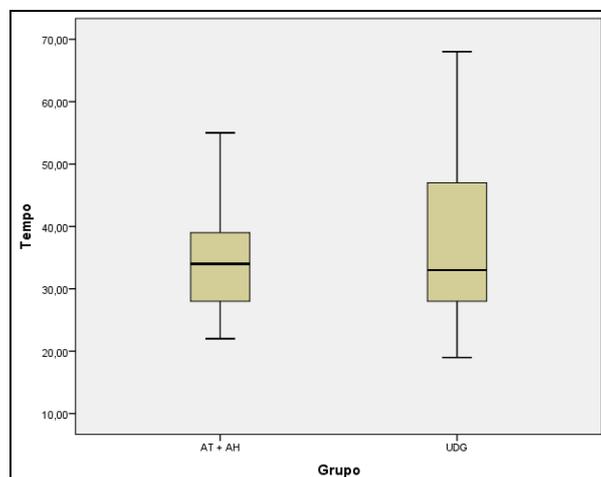


Figura 69. Boxplots do tempo de construção - 2º estudo de viabilidade UDG

Através da Figura 69, pode-se notar que a mediana do grupo da AT+AH está no mesmo nível da mediana do grupo da UDG. Ao comparar as duas amostras usando o teste estatístico de Mann-Whitney, não foi encontrada diferença estatística significativa entre os grupos ( $p = 0,782$ ). Os resultados encontrados sugerem que ambos os grupos (AT+AH e UDG) levaram quase que de forma similar o mesmo tempo para construir os *mockups* (apoia a hipótese H04).

### e. Estética

Este indicador foi calculado tendo como base a nota dada por 2 avaliadores em relação à estética dos *mockups* construídos neste estudo. Aos avaliadores foram mostrados os 35 *mockups* projetados por cada participante deste estudo e um formulário com a versão adaptada do VisAWI (ver Tabela 33). Na Tabela 41 (última coluna) é mostrado o total de notas negativas dadas pelo avaliador 1 em relação à estética dos *mockups* de participante. Na Tabela 42 é mostrado o total de notas negativas dadas pelo avaliador 2.

Tabela 41. Resultados do indicador Estética do 2º estudo de viabilidade da UDG do Avaliador 1

#	Grupo	Mockups dos Participantes	Simplicidade					Diversidade			Habilidade			Total
			1	2	3	4	5	6	7	8	9	10	11	
Avaliador 1	Grupo 1 AT+AH	Mockups do P01	0	0	0	0	0	1	1	1	1	1	1	6
		Mockups do P02	0	0	0	0	0	0	1	0	0	0	0	1
		Mockups do P03	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P04	1	0	0	0	0	0	0	0	0	0	0	1
		Mockups do P05	1	1	1	0	1	1	1	0	1	1	0	8
		Mockups do P06	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P07	1	0	0	0	0	1	1	0	1	1	1	6
		Mockups do P08	0	1	1	0	1	1	1	1	1	1	0	8
		Mockups do P09	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P10	0	0	0	0	0	0	0	0	0	0	1	1
		Mockups do P11	1	1	1	0	1	1	1	1	1	1	1	10
		Mockups do P12	0	0	0	0	0	0	0	0	0	0	1	1
		Mockups do P13	0	0	0	0	0	1	0	0	0	0	0	1
		Mockups do P14	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P15	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P16	0	0	0	0	0	0	0	0	0	0	1	1
		Mockups do P17	0	0	0	0	0	0	0	0	0	0	0	0
	Mockups do P18	0	0	0	0	0	1	1	0	0	0	0	2	
	Mockups do P19	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P20	0	1	1	0	0	0	0	0	0	0	1	3	
	Mockups do P21	0	0	0	0	0	1	1	0	0	1	0	3	
	Mockups do P22	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P23	0	0	0	0	0	0	0	0	0	0	1	1	
	Mockups do P24	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P25	0	1	1	1	1	1	1	1	1	1	1	10	
	Mockups do P26	0	0	0	0	0	0	0	0	0	1	1	2	
	Mockups do P27	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P28	0	0	0	0	0	1	1	1	1	1	0	5	
	Mockups do P29	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P30	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P31	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P32	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P33	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P34	0	0	0	0	0	0	0	0	0	1	0	1	
	Mockups do P35	0	0	0	0	0	0	0	0	0	0	0	0	
	Grupo 2 UDG													

Através da Tabela 41 e Tabela 42, pode-se notar que os *mockups* projetados pelo grupo da AT+AH que obteve maior quantidade de notas negativas pelos dois avaliadores

foram os *mockups* do participante P11. No entanto, analisado os dados dos indicadores de corretude e completude, percebeu-se que este participante obteve 100% em cada um destes indicadores. Isto pode indicar que embora a AT+AH não apoie tanto na construção de *mockups* com boa estética, esta abordagem permite a construção de *mockups* corretos e completos. E no grupo da UDG os *mockups* projetados pelo participante P25 foram os que obtiveram maior quantidade de notas negativas pelos dois avaliadores. Este mesmo participante levou apenas 19 minutos (menor tempo de todos) para construir seus *mockups*. Isto mostra que este participante talvez não tenha se dedicado tanto à estética dos *mockups*.

Tabela 42. Resultados do indicador Estética do 2º estudo de viabilidade da UDG do Avaliador 2

#	Grupo	Mockups dos Participantes	Simplicidade					Diversidade			Habilidade			Total
			1	2	3	4	5	6	7	8	9	10	11	
Avaliador 2	Grupo 1 AT+AH	Mockups do P01	0	0	0	0	0	1	1	0	0	1	1	4
		Mockups do P02	1	0	1	1	0	1	1	0	1	1	0	7
		Mockups do P03	0	0	0	0	0	1	0	0	0	0	0	1
		Mockups do P04	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P05	0	0	0	0	1	0	0	0	1	1	0	3
		Mockups do P06	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P07	0	1	1	1	1	1	1	0	1	1	1	9
		Mockups do P08	0	0	0	1	1	1	0	0	1	1	0	5
		Mockups do P09	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P10	0	0	0	1	1	1	1	0	1	1	1	7
		Mockups do P11	0	1	1	1	1	1	1	1	1	1	1	10
		Mockups do P12	0	0	1	1	1	0	0	0	1	0	1	5
		Mockups do P13	0	0	0	0	0	0	0	0	0	0	0	0
		Mockups do P14	0	0	0	0	0	0	0	0	0	1	0	1
		Mockups do P15	0	0	0	0	0	0	0	0	0	0	0	0
	Mockups do P16	0	1	1	1	1	1	0	0	1	0	1	7	
	Mockups do P17	0	0	0	1	0	1	1	0	0	1	0	4	
	Mockups do P18	0	1	1	0	1	1	1	1	1	0	0	7	
	Mockups do P19	0	0	0	1	1	1	1	0	0	1	0	5	
	Mockups do P20	0	1	1	0	0	1	1	0	1	1	1	7	
	Mockups do P21	0	0	0	0	0	1	1	0	0	1	0	3	
	Mockups do P22	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P23	0	0	1	1	1	1	1	1	1	1	1	9	
	Mockups do P24	0	0	1	1	1	1	0	0	0	1	0	5	
	Mockups do P25	0	1	1	1	1	1	1	1	1	1	1	10	
	Mockups do P26	0	0	1	1	1	1	1	1	1	1	1	9	
	Mockups do P27	0	0	0	0	0	0	0	0	0	0	0	0	
	Mockups do P28	0	0	0	1	0	1	1	1	0	1	0	5	
	Mockups do P29	0	0	0	0	1	1	0	0	0	1	0	3	
	Mockups do P30	0	0	0	0	0	0	0	0	0	1	0	1	
	Mockups do P31	0	0	0	0	0	0	0	0	0	1	0	1	
	Mockups do P32	0	0	0	0	0	1	0	0	0	1	0	2	
	Mockups do P33	0	0	1	1	1	1	0	0	0	1	0	5	
	Mockups do P34	0	0	0	0	0	0	0	0	0	1	0	1	
	Mockups do P35	0	1	0	1	1	0	0	0	0	0	0	3	
	Grupo 2 UDG													

Além disso, pode-se observar que, para o avaliador 2, os *mockups* do grupo da UDG foram os que obtiveram mais notas negativas. No entanto, para o avaliador 1 os *mockups* do grupo da UDG foram os que obtiveram menos notas negativas (apóia a hipótese H05).

Especificamente, os *mockups* projetados pelos participantes P22 e P27 não tiveram nenhuma nota negativa, mostrando que a estética do mesmo estava adequada para estes dois avaliadores.

#### f. Facilidade de Uso

A Figura 70 apresenta as percepções dos participantes com relação à facilidade de uso da AT+AH e UDG. Pode-se observar na Figura 70 que quatro dos dezoito participantes da UDG (P20, P24, P30 e P34) discordaram da afirmativa “*Interagir com a tecnologia não exige muito do meu esforço mental*”, mostrando que para estes participantes a UDG exige um concentração maior. Nesta mesma afirmativa quatro participantes do grupo da AT+AH discordaram, sendo que um deles (P15) discordou totalmente. Como a mesma quantidade de participantes da UDG e da AT+AH discordaram nesta afirmativa, apoia-se a hipótese H06. Além disso, dois participantes da UDG (P24 e P30) discordaram amplamente da afirmativa “*Considero a tecnologia fácil de usar*”, destacando a dificuldade que eles tiveram ao utilizar a UDG para construir *mockups* com boa usabilidade.

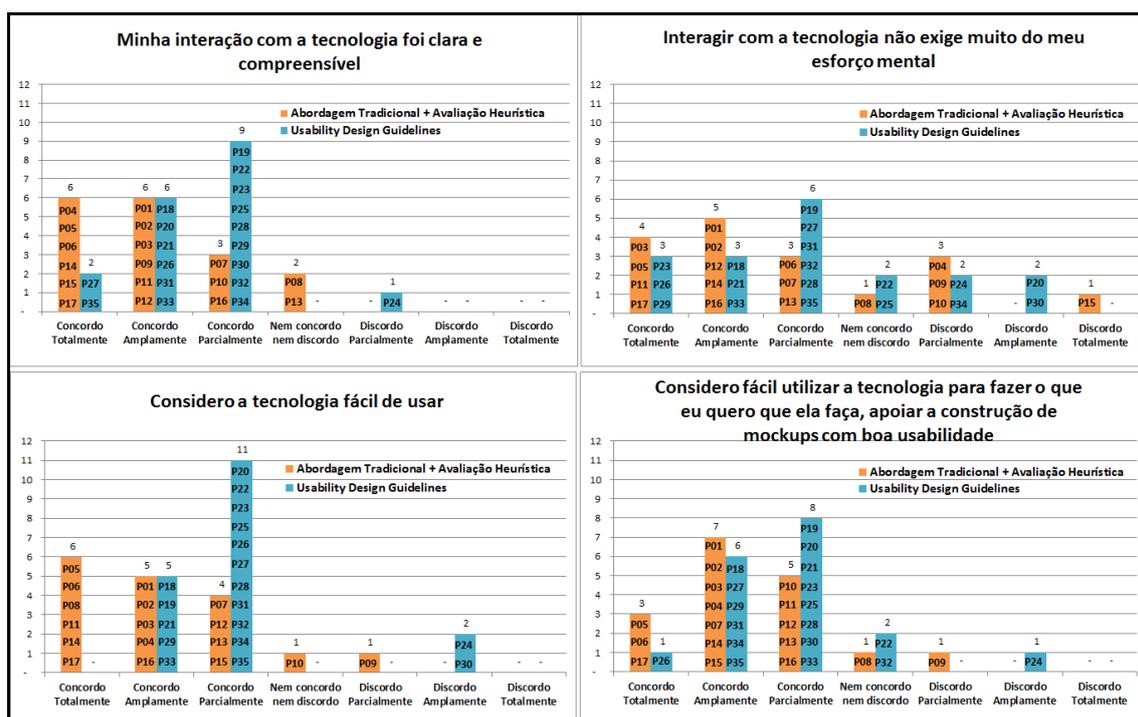


Figura 70. Percepção sobre Facilidade de Uso - 2º estudo de viabilidade UDG

#### g. Utilidade Percebida

A Figura 71 apresenta as percepções dos participantes com relação à utilidade percebida da AT+AH e UDG. Verificou-se que a maioria dos participantes que discordaram das afirmativas deste indicador foram participantes do grupo da UDG. Isto

indica que alguns participantes deste estudo consideram que a tecnologia UDG não é tão útil em seu propósito (apoia a hipótese HA7). Isto pode ter acontecido porque a técnica é muito extensa.

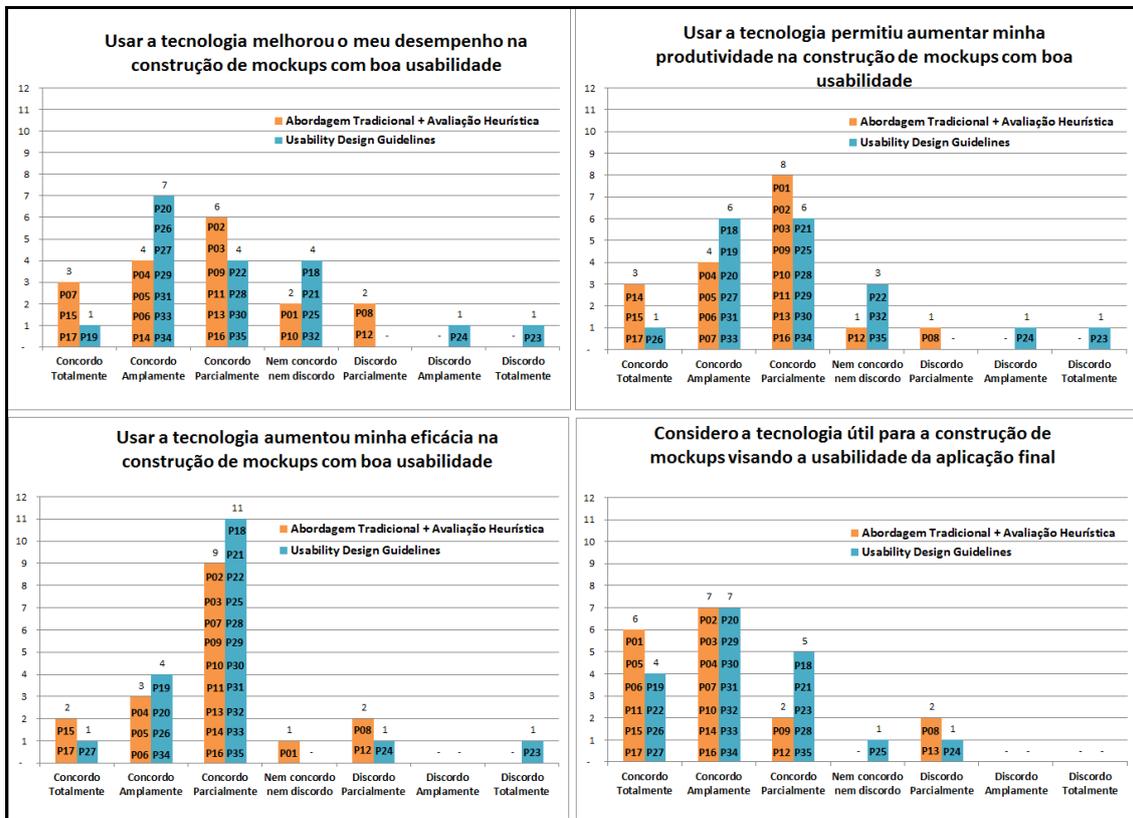


Figura 71. Percepção sobre Utilidade - 2º estudo de viabilidade UDG

### h. Intenção de Uso Futuro

A Figura 72 apresenta as percepções dos participantes com relação à intenção de uso futuro da AT+AH e UDG.

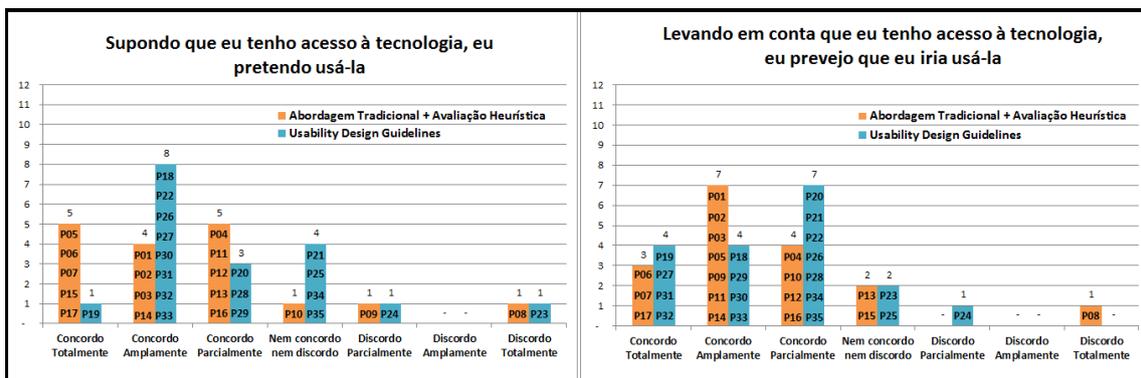


Figura 72. Percepção sobre Intenção de Uso - 2º estudo de viabilidade UDG

Pode-se observar que a mesma quantidade de participantes, tanto da AT+AH quanto da UDG, discordou das afirmativas deste indicador. Além disso, a maioria dos

participantes de ambos os grupos concordaram com as duas afirmativas. Isto pode indicar que os participantes possuem a intenção de usar estas tecnologias em trabalhos futuros.

#### 9.4.4. Resultados Qualitativos do 2º Estudo de Viabilidade da UDG

Foi realizada uma análise específica dos dados qualitativos (comentários adicionais dos participantes) contido nos questionários pós-uso. Esta análise foi feita usando os procedimentos do método *Grounded Theory* (GT) (CORBIN e STRAUSS, 2008).

Os dados qualitativos que foram extraídos dos questionários pós-uso foram analisados usando um subconjunto de fases do processo de codificação sugerido por Corbin e Strauss (2008), conforme já descrito em estudos anteriores. O propósito da análise neste estudo foi entender a percepção dos participantes através de sua experiência usando a UDG. A seguir será apresentado os resultados e análise qualitativa deste estudo.

##### a. Pontos de Vista sobre as Dificuldades de Uso da UDG

Algumas das dificuldades que foram coletadas neste experimento quando se usou a UDG foram: a UDG tem textos longos que dificultam o entendimento (ver citação de P27 abaixo); a UDG exige esforço mental (ver citação de P34 abaixo); aplicar a UDG em um curto período de tempo prejudica no desempenho (ver citação de P31 abaixo); cansou de ler as *guidelines* e foi ler as dicas (ver citação de P30 abaixo).

*“Em alguns momentos as descrições eram longas demais, o que dificulta o entendimento”.*  
(Participante P27).

*“A técnica exige esforço mental, pois a leitura e entendimento dos diagramas reque atenção e concentração”.* (Participante P34).

*“Apesar de ser muito boa, a aplicação da técnica em um curto período de tempo prejudica no desempenho e no resultado obtido”.* (Participante P31).

*“Honestamente, comecei a ler, cansei e usei as dicas”* (Participante P30).

Pode-se observar que houve algumas dificuldades. Uma delas foi que a UDG exigiu esforço mental para ser aplicada. A UDG é uma técnica extensa e por isso exige muito esforço por parte do projetista para ler e desempenhar o que a técnica orienta a fazer. Como dito no estudo de viabilidade anterior, se reduzirmos a técnica, eliminando dicas, ela pode não cumprir com seu papel de apoiar a construção de *mockups* com boa usabilidade. No entanto, baseado no comentário deste participante, refletiu-se e teve-se a ideia de criar uma versão desta técnica baseada em *checklist*. Acredita-se que utilizando perguntas que orientem a construção torne a técnica menos cansativa.

Além disso, outra dificuldade de uso foi que ao aplicar a UDG em um curto período de tempo prejudicou o desempenho do participante. Como a técnica UDG é longa, a mesma pode ser utilizada em partes nos projetos acadêmicos ou industriais. Por exemplo, ao utilizar a técnica em uma segunda vez e já conhecendo a mesma, pode-se pular algumas *guidelines* e usar apenas as *guidelines* que se tenha tido maior dificuldade de memorizar ou compreender. Ou seja, ao já conhecer a técnica, o projetista não precisa ler necessariamente todas as *guidelines* e dicas.

#### **b. Ponto de Vista sobre a Utilidade da UDG**

Algumas dos pontos de vista dos participantes sobre a utilidade da UDG neste experimento foram: a UDG ajudou a entender como deve ser construído o *mockup* a fim de melhorar a experiência do usuário (ver citação de P34 abaixo); a UDG ajudou na revisão do que foi projetado como esboço no início do estudo (ver citação de P27 abaixo); a UDG facilita o desenvolvimento de aplicações interativa (ver citação de P14 abaixo); a UDG ajuda a lembrar de detalhes que se acaba esquecendo (ver citação de P35 abaixo).

*“(...)[a técnica] me ajudou a entender o que devia estar contido no mockup, a fim de ajudar a melhorar a experiência do usuário”.* (Participante P34).

*“(...)[a técnica] auxiliou quanto a revisão daquilo que eu projetei inicialmente em pontos de esquecimento”.* (Participante P27).

*“O emprego da técnica auxiliou bastante na percepção do usuário e facilita o desenvolvimento de boas aplicações interativas.”.* (Participante P14).

*“Auxilia na compreensão de alguns detalhes que possam passar despercebido”* (Participante P35).

Um dos pontos de vista sobre a utilidade da UDG que pode ser destacado é que a UDG ajuda a lembrar de detalhes que se acaba esquecendo durante a construção de *mockups* com boa usabilidade. Evitar que pontos importantes sejam esquecidos e apoiar profissionais não especialistas em usabilidade na construção de *mockups* com boa usabilidade são alguns dos objetivos da UDG.

#### **c. Sugestões de Melhoria para a UDG**

Uma das sugestões de melhoria para a UDG que merece ser destacada foi: melhorar a tela inicial para orientar mais o início de aplicação da técnica (ver citação abaixo de P27). Pretende-se fazer melhorias na tela inicial de forma que esta seja mais explicativa e intuitiva possível. Uma das ideias é acrescentar uma lista resumida das *guidelines* de

usabilidade logo no início, para que o projetista entenda a técnica contém este tipo de *guidelines* como apoio à construção de *mockups*.

“(…) melhorar a descrição inicial do processo pois me senti um pouco desorientada em como iniciar”. (Participante P27).

Outras sugestões de melhorias foram: ter mais exemplos visuais nas dicas (ver citação abaixo de P20) e ter menos textos (ver citação abaixo de P27). Em trabalhos futuros, pretende-se acrescentar mais exemplos visuais de *mockups*, tanto exemplos bons quanto ruins, em relação à usabilidade. A ideia é que estes exemplos apoiem ainda mais a criação dos *mockups*.

“Acho que deveria conter mais exemplos visuais, ilustrando cada uma das dicas que ele fornece”. (Participante P20).

“Diminuir as descrições”. (Participante P27).

## 9.5. Ameaças à Validade do 1º e 2º Estudo de Viabilidade da UDG

Nestes estudos houve algumas ameaças que podem afetar a validade de seus resultados. Nesta Subseção será discutido sobre estas ameaças, categorizando-as usando a mesma abordagem de Wohlin *et al.*, (2000): interna, externa, de conclusão e de constructo.

### 9.5.1. Validade Interna

Nestes experimentos, considerou-se quatro principais ameaças que representam um risco à interpretação dos resultados: (1) classificação da experiência, (2) medição do tempo, (3) influência do moderador e; (4) uso das descrições textuais. As três primeiras ameaças já foram abordadas nos outros estudos relatados e teve-se o mesmo cuidado para que estas ameaças não influenciassem os resultados do estudo. E por fim, em relação ao uso das descrições textuais, estas poderiam afetar o estudo caso os participantes não as entendessem. Esta ameaça foi minimizada utilizando descrições textuais escritas em linguagem natural, onde as tarefas do usuário estavam explícitas e de fácil entendimento.

### 9.5.2. Validade Externa

Cinco ameaças foram consideradas: (1) participantes foram estudantes de graduação ao invés de analistas ou designers de *software*; (2) o estudo foi conduzido em um ambiente acadêmico; e (3) a validade do artefato utilizado. Estas ameaças também já foram abordadas nos outros estudos relatados e teve-se o mesmo cuidado para que estas ameaças não influenciassem os resultados do estudo.

### 9.5.3. Validade de Conclusão

No 1º estudo de viabilidade da UDG, o principal problema é o tamanho da amostra, por serem 12 estudantes. O pequeno número de participantes não é o ideal do ponto de vista estatístico. Tamanho da amostra é um problema conhecido em estudos IHC e ES. No 2º estudo de viabilidade, o tamanho da amostra não foi considerado um problema, uma vez que houve uma quantidade alta de participantes (35 alunos de graduação). Outra ameaça é a homogeneidade da amostra, pois os participantes em cada estudo são todos estudantes de uma mesma instituição. Devido a estes fatos, há limitação dos resultados, sendo estes considerados indícios e não conclusivos.

### 9.5.4. Validade de Constructo

A validade de constructo pode ter sido influenciada pelos indicadores que foram aplicados na análise quantitativa: corretude, completude, eficácia, tempo de construção, estética, facilidade de uso, utilidade percebida e intenção de uso. Estes indicadores foram definidos pelos pesquisadores envolvidos nesta pesquisa e representam aspectos importantes ao se projetar *mockups* visando à usabilidade da aplicação final.

## 9.6. Síntese do Capítulo

Este capítulo relatou os resultados de dois experimentos cujo objetivo foi avaliar a corretude, completude, eficácia, tempo de construção dos *mockups*, estética, facilidade de uso, utilidade percebida e intenção de uso ao utilizar a tecnologia UDG em comparação com uma abordagem tradicional em conjunto a Avaliação Heurística de Nielsen. Estes experimentos serviram como um meio valioso para a obtenção de feedback dos participantes com o objetivo de melhorar a UDG.

No geral, pode-se perceber que em ambos os experimentos da UDG e AT+AH os resultados quantitativos foram similares. Isto pode ter acontecido porque as tarefas do usuário que foram escolhidas para serem projetadas nos *mockup* eram simples (cadastro e consulta de viagem). Portanto, é interessante executar novos estudos com tarefas mais complexas para observar como ambas as técnicas se comportam. Porém, isto não inviabiliza os estudos executados, uma vez que o intuito foi analisar a viabilidade de uso de ambas as técnicas e analisar como as mesmas se comportam no apoio à construção de *mockups* com boa usabilidade.

## CAPÍTULO 10 - CONCLUSÕES E PERSPECTIVAS FUTURAS

*Este capítulo apresenta conclusões desta pesquisa, resumindo sua motivação e proposta, apresentando as suas contribuições. As perspectivas futuras fornecem a direção para que seja dada continuidade a esta pesquisa.*

### 10.1. Conclusões

O foco desta pesquisa foi apoiar a avaliação de usabilidade de modelos utilizados nas fases iniciais no processo de desenvolvimento de *software*. Com este propósito, no mestrado foi definido o conjunto de técnicas chamada MIT (*Model Inspection Technique for Usability Evaluation*), que são técnicas de inspeção de usabilidade específicas para avaliação de modelos de análise e projeto de *software* e no doutorado estas técnicas foram evoluídas através dos estudos experimentais. Além disso, esta pesquisa de doutorado tem como foco apoiar o projeto de modelos utilizados nas fases iniciais do processo de desenvolvimento de *software*, já visando à usabilidade da aplicação final. Com este propósito, foram definidas duas técnicas chamadas UDRT-AD (*Usability Design Reading Technique for Activity Diagrams*) e UDG (*Usability Design Guidelines*). A técnica UDRT-AD tem o objetivo de auxiliar na construção do Diagrama de Atividades, focando em questões de usabilidade a fim de obter aplicações com maior qualidade. A UDG tem como objetivo apoiar a construção de *mockups* visando à usabilidade da aplicação final. Estas técnicas foram desenvolvidas de forma que pudessem ser empregadas pelos próprios envolvidos no processo de desenvolvimento de *software* (como analistas, projetistas e avaliadores) na avaliação e projeto de seus modelos.

As tecnologias propostas foram avaliadas experimentalmente através de uma família de experimentos. Os resultados dos estudos realizados proporcionaram indícios de que as tecnologias apresentam bom nível de eficácia, eficiência, completude, corretude, facilidade de uso, utilidade e intenção de uso futuro. A seguir são apresentados os requisitos que guiaram a proposta das tecnologias e os resultados relacionados, medidos através do estudo experimental executado:

- Ser fácil de aprender e de utilizar: o tempo de apresentação/treinamento das tecnologias propostas variou entre 15 e 20 minutos. Logo após o treinamento todos os participantes foram capazes de desempenhar a atividade de detecção de problemas de usabilidade (com as MITs) e a atividade de construção de modelos com boa usabilidade (com a UDRT-AD e UDG). Além disso, as técnicas obtiveram grande aceitação pelos participantes dos estudos. Portanto, as tecnologias propostas podem ser aplicadas após um curto treinamento, que pode variar entre 15 e 20 minutos, e são consideradas fáceis de aprender e usar.

- Apresentar bom nível de eficácia: a eficácia média do grupo de participantes que utilizou a MIT 1 no 1º estudo de viabilidade foi de 33,06% enquanto que a eficácia média utilizando AH foi de 19,44%. No 2º estudo de viabilidade a eficácia média do grupo de participantes que utilizou a MIT 1 foi de 11,47% enquanto que a eficácia média utilizando UCE foi de 7,12%. Em relação ao 1º estudo de viabilidade com a MIT 2, a eficácia média do grupo de participantes que utilizou a MIT 2 foi de 15,87% enquanto que a eficácia média utilizando AH foi de 15,24%. No estudo de observação a MIT 2 obteve uma eficácia média de 21,85%. No 1º estudo de viabilidade da MIT 3, a eficácia média do grupo de inspetores que utilizou a MIT 3 foi de 35% enquanto que a eficácia média utilizando as Diretrizes de Usabilidade foi de 46%. No 2º estudo de viabilidade da AT&MIT 3 e UDRT-AD, a eficácia média do grupo de inspetores que utilizou a AT&MIT 3 foi de 89% enquanto que a eficácia média utilizando UDRT-AD foi de 97%. Por fim, a UDG obteve uma eficácia de 96,25% no 1º estudo de viabilidade e uma eficácia de 98,78% no 2º estudo de viabilidade em comparação à AT+AH que obteve uma eficácia de 96,57% no 1º estudo de viabilidade e uma eficácia de 97,84% no 2º estudo de viabilidade.

- Apresentar bom nível de eficiência: a eficiência média do grupo de participantes que utilizou a MIT 1 no 1º estudo de viabilidade foi de 13,25 defeitos/hora enquanto eficiência média utilizando AH foi de 12,57 defeitos/hora. No 2º estudo de viabilidade a eficiência média do grupo de participantes que utilizou a MIT 1 foi de 18,20 defeitos/hora enquanto eficiência média utilizando o UCE foi de 14,10 defeitos/hora. Em relação ao 1º estudo de viabilidade com a MIT 2, a eficiência média do grupo de participantes que utilizou a MIT 2 foi de 14,02 defeitos/hora enquanto que a eficiência média utilizando AH foi de 15,08 defeitos/hora. No estudo de observação, os desenvolvedores que usaram a MIT 2 conseguiram encontrar entre 3 e 16 defeitos entre 0,28 e 1,02 horas. No 1º estudo de viabilidade da MIT 3, a eficiência média do grupo de inspetores que utilizou a MIT 3 foi de 16,79 defeitos/hora enquanto que a eficiência média

utilizando as Diretrizes de Usabilidade foi de 18,17 defeitos/hora. No 2º estudo de viabilidade da AT&MIT 3 e UDRT-AD não foi calculado a eficiência e sim o tempo de modelagem, portanto, o tempo de modelagem médio do grupo de participantes que utilizou a AT&MIT 3 foi de 0,69 horas enquanto que o grupo da UDRT-AD levou cerca de 1,04 horas. Por fim, o tempo médio de construção dos *mockups* pelo grupo que utilizou a UDG foi de 88,17 minutos no 1º estudo de viabilidade e de 79,67 minutos pelo grupo da AT+AH. E no 2º estudo de viabilidade o grupo que usou a UDG levou cerca de 36,78 minutos em comparação ao grupo da AT+AH que levou cerca de 35,71 minutos.

- Oferecer uma boa relação custo-benefício na sua aplicação: os custos destas tecnologias são baixos, uma vez que necessitam de poucas pessoas as aplicando. Desta forma, uma ou duas pessoas conseguem encontrar a maior parte dos problemas de usabilidade ou conseguem projetar com boa usabilidade. Além disso, as tecnologias propostas podem ser aplicadas em um curto período de tempo, como visto nos estudos experimentais. Por exemplo, a média de tempo para aplicar a MIT 1 foi de 17,07 minutos no 2º de viabilidade. A média de tempo para usar a MIT 2 foi de 23,78 minutos no estudo de viabilidade. E a média de tempo para aplicar a MIT 3 foi de 0,81 horas (ou 48,6 minutos) no 1º estudo de viabilidade. Já o grupo que aplicou a UDRT-AD levou cerca de 1,04 hora (ou 62,4 minutos). Por fim, o tempo médio de uso da UDG no 2º estudo de viabilidade foi de 36,78 minutos. Além disso, os materiais necessários para aplicar estas tecnologias são de baixo custo, como papel, lápis, leitor de pdf (para o uso da UDG), ferramentas opcionais como processador de texto e planilha eletrônica (para anotação de problemas de usabilidade) e Balsamiq versão grátis (para a construção de *mockups*). Portanto, as tecnologias propostas oferecem uma boa relação custo-benefício, se comparadas à abordagens convencionais onde a identificação de problemas de usabilidade é realizada quando o *software* já está pronto, gerando assim retrabalho e desperdício de tempo e dinheiro das empresas de *software*.

## 10.2. Contribuições

As principais contribuições desta tese de doutorado são descritas nas Subsubseções abaixo:

### 10.2.1 Desenvolvimento de tecnologias que apoiam o projeto e avaliação de modelos das fases iniciais, visando à usabilidade da aplicação

- Definição de tecnologias específicas para apoiar o projeto e avaliação de modelos utilizados nas fases iniciais do processo de desenvolvimento, que

podem ser empregadas pelos próprios envolvidos nos projetos de desenvolvimento de software.

- Elaboração de recursos de apoio à realização de inspeções de usabilidade de modelos em organizações de desenvolvimento de software: instruções para a inspeção de usabilidade com as MITs e planilhas para relato de discrepâncias; e recursos de apoio à construção de modelos visando à usabilidade: instruções para os procedimentos dos experimentos.

#### **10.2.2 Realização de estudos secundários voltados para questões características de projeto e avaliação de modelos visando à usabilidade**

- Realização de um mapeamento sistemático para caracterizar tecnologias que apoiam a integração entre as áreas de Engenharia de Software e Interação Humano-Computador.
- Realização de um mapeamento sistemático para caracterizar tecnologias que auxiliam no projeto ou avaliação de modelos visando à usabilidade da aplicação.
- A realização destes Mapeamentos Sistemáticos possibilitou:
  - Definição de um protocolo para realização de mapeamento sistemático que pode ser utilizado/estendido em novas pesquisas sobre este tema;
  - Estabelecimento de uma base de conhecimento sobre tecnologias que melhoram a usabilidade no processo de desenvolvimento de *software*.
  - Estabelecimento de uma base de conhecimento sobre tecnologias que auxiliam no projeto ou avaliação de modelos visando à usabilidade.

#### **10.2.3 Definição de pacote para planejamento, execução e análise de estudos experimentais**

- Avaliação e evolução das tecnologias propostas e dos recursos de apoio ao uso das tecnologias;
- Disseminação do conhecimento sobre realização dos estudos experimentais, incluindo análise qualitativa tendo por base os procedimentos do método Grounded Theory (CORBIN e STRAUSS, 1998) e avaliação da aceitação de tecnologia.

### **10.3. Perspectivas Futuras**

A realização desta pesquisa resultou no desenvolvimento de tecnologias para apoio ao projeto e avaliação de modelos de análise e projeto. Este resultado abre novas

perspectivas de pesquisa, que podem ser exploradas em trabalhos futuros. Alguns desses trabalhos futuros são detalhados a seguir.

### **10.3.1. Extensão dos Mapeamentos Sistemáticos da Literatura**

Pretende-se realizar uma análise dos estudos secundários identificados no 1º e 2º Mapeamento Sistemático da Literatura descritos nesta tese. Isto permitirá que novas tecnologias que apoiam a usabilidade nas fases iniciais do processo de desenvolvimento sejam identificadas e analisadas.

### **10.3.2. Propostas de novas tecnologias que apoiam o projeto de outros modelos**

Novas tecnologias poderão ser propostas a fim de apoiar o projeto de outros modelos de análise e outros modelos de projeto, visando à usabilidade da aplicação final. Identificou-se, através dos mapeamentos sistemáticos, uma carência de tecnologias com este propósito, por exemplo: há poucas tecnologias para apoiar a construção de modelos na fase de análise de um processo de desenvolvimento de software. Este tipo de tecnologia permite uma forma proativa de considerar a usabilidade no processo de desenvolvimento de *software*. Além disso, poderão ser construídas tecnologias computacionais, como ferramentas, para apoiar o uso das tecnologias descritas nesta tese.

### **10.3.3. Adaptações nas tecnologias propostas**

Nos estudos anteriores realizados com a tecnologia UDG, percebeu-se que se esta tecnologia tivesse sido construída para ser um *checklist*, talvez ela fosse mais intuitiva e fácil de usar, além de se tornar mais concisa. Adaptações podem ser feitas neste sentido na técnica e uma nova versão da UDG pode ser gerada. Consequentemente, novos estudos devem ser realizados afim de avaliar a nova versão da técnica em comparação com a versão original para verificar qual tem melhor eficácia, eficiência e satisfação de uso. Além disso, seria interessante o desenvolvimento de uma ferramenta que apoie a aplicação das tecnologias propostas nesta pesquisa. A ideia é que esta ferramenta seja gratuita de forma que não encareça o uso das tecnologias tanto em ambiente acadêmico quanto industrial.

### **10.3.4. Triangulação dos dados dos estudos experimentais realizados**

Triangulação é um procedimento que combina diferentes métodos de coleta de dados, distintas amostras, diferentes perspectivas teóricas e diferentes momentos no tempo, para consolidar as conclusões a respeito do fenômeno que está sendo investigado. Este tipo de procedimento permitiria uma maior análise dos dados coletados nos

experimentos executados, além da identificação de outras lacunas de pesquisa. Portanto, futuramente, pretende-se realizar a triangulação dos dados dos experimentos realizados.

#### **10.3.5. Realização de novos estudos experimentais em outros contextos**

Outros estudos experimentais poderão ser realizados a fim de avaliar as tecnologias propostas em outros contextos de aplicação. Alguns contextos não foram explorados nesta pesquisa de doutorado, por exemplo: aplicação das tecnologias em projetos de desenvolvimento ágil e aplicação das tecnologias no contexto de aplicações móveis e desktops. Este tipo de estudo permitiria caracterizar o comportamento das tecnologias propostas nestes e em outros contextos de uso.

#### **10.3.6. Realização de estudo de caso em ciclo de vida real e na indústria**

SHULL *et al.* (2001) apontam que estudos de caso em Ciclo de Vida Real e na Indústria devem ser realizados para identificar se existem problemas de integração na aplicação da tecnologia proposta em um ciclo de vida real e em um ambiente industrial. Com a realização deste estudo, pode-se observar indícios sobre a efetividade das tecnologias propostas. O propósito destes estudos é: (1) caracterizar a aplicação da tecnologia no contexto de um ciclo de vida de desenvolvimento e (2) identificar se a aplicação da tecnologia possui alguma interação negativa com o ambiente industrial.

Nesta pesquisa executou-se um único estudo de caso com a MIT 2 em uma indústria de *software*. Futuramente, pretende-se executar mais estudos de caso na indústria com a MIT 2, assim como, com as demais tecnologias propostas nesta tese (MIT 1, MIT 3, UDRT-AD e UDG).

## REFERÊNCIAS

- ABRAHÃO, S., INSFRAN, E., 2006. "Early Usability Evaluation in Model Driven Architecture Environments". In International Conference on Quality Software, Beijing, pp. 287 - 294.
- ANDERSON, J., 1983. "The architecture of cognition". Harvard University Press.
- ARDITO C., BUONO P., CAIVANO D., COSTABILE M. F., LANZILOTTI R., BRUUN A., STAGE J., 2011. "Usability evaluation: a survey of software development organizations". In International Conference on Software Engineering and Knowledge Engineering (SEKE 2011), Miami, pp. 282-287.
- BACIM, F., POLYS, N., CHEN, J., SETAREH, M., LI, J., MA, L., 2010. "Cognitive scaffolding in Web3D learning systems: a case study for form and structure". In International Conference on Web 3D Technology, pp 93 - 100.
- BASIL, V., ROMBACH, H., 1988. "The TAME project: Towards improvement-oriented software environments". In IEEE Transactions on Software Engineering, v. 14 (6), pp. 758 - 773.
- BASIL, V., GREEN, S., LAITENBERGER, O., LANUBILE, F., SHULL, F., SØRUMGÅRD, S., ZELKOWITZ, M., 1996. "The Empirical Investigation of Perspective-Based Reading". Empirical Software Engineering, v.1(2), pp. 133 - 164.
- BEN AMMAR, L., MAHFOUDHI, A., KACEM, Y. H., 2013. "Empirical evaluation of an early understandability measurement method". In International Conference on Control, Decision and Information Technologies, pp. 454 - 457.
- BLANDFORD, A., KEITH, S., BUTTERWORTH, R., FIELDS, B., FURNISS, D., 2007. "Disrupting digital library development with scenario informed design". In Interacting with Computers, v. 19 (1), pp. 70 - 82.
- BOEHM, B., 2006. "A view of 20th and 21st century software engineering". In Proc. of the International Conference on Software Engineering, pp. 12-29.
- BOLCHINI, D., GARZOTTO, F., 2007. "Quality of Web usability evaluation methods: an empirical study on MiLE+". In International Workshop on Web Usability and Accessibility, pp. 481 - 492.
- BOOCH, G., RUMBAUGH, J., JACOBSON, I., 2005. "Modeling Language user Guide". In The Addison-Wesley Objectc Technology Series, 2a edição, 484 páginas.
- BUDGEN, D., TURNER, M., BRERETON, P., KITCHENHAM, B., 2008. "Using mapping studies in software engineering". In Psychology of Programming Interest Group, Lancaster University, pp. 195 - 204.
- CALEFATO, F., LANUBILE, F., MINERVINI, P., 2010. "Can Real-Time Machine Translation Overcome Language Barriers in Distributed Requirements Engineering?". In IEEE International Conference on Global Software Engineering, pp. 257 - 264.
- CARD, S., MORAN, T., NEWELL, A., 1983. "The Psychology of Human-Computer Interaction". Hillsdale, New Jersey: Erlbaum.
- CARVAJAL, L., 2012. "Usability-oriented Software Development Process". Tese (Doutorado) - Universidad Politécnica de Madrid, 272 páginas.
- CARVER, J., JACCHERI, L., MORASCA, S., SHULL, F., 2003. "Issues in Using Students in Empirical Studies in Software Engineering Education". In International Symposium on Software Metrics, Sydney, pp. 239 - 249.
- CHEN, Y., KUTAR, M., HAMILTON, A., 2007. "Modelling and matching: a methodology for eplanning system development to address the requirements of

- multiple user groups". In International Conference on Online Communities and Social Computing, pp. 41 - 49.
- CONTE, T., MASSOLAR, J., MENDES, E., TRAVASSOS, G., 2007a. "Web Usability Inspection Technique Based on Design Perspectives". In Brazilian Symposium on Software Engineering, v. 1, pp. 394 - 410.
- CONTE, T., MASSOLAR, J., MENDES, E., TRAVASSOS, G., 2007b. "Usability Evaluation based on Web Design Perspectives". In International Symposium on Empirical Software Engineering and Measurement, pp. 146 – 155.
- CONTE, T., MASSOLAR, J., MENDES, E., TRAVASSOS, G., 2009a. "Web usability inspection technique based on design perspectives". In IET Software, v. 3 (2), pp. 106 – 123.
- CONTE, T., CABRAL, R., TRAVASSOS, G., 2009b. "Applying Grounded Theory in Qualitative Analysis of an Observational Study in Software Engineering: An Experience Report". In INFOCOMP Journal of Computer Science, v.2(1), pp.58-69.
- CORBIN, J., STRAUSS, A., 2008. "Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory". Sage, 3ª edição, 400 páginas.
- CYBIS, W., BETIOL, A., FAUST, R., 2007. "Ergonomia e Usabilidade: Conhecimentos, Métodos e Aplicações". São Paulo: Novatec, 3ª edição, 496 páginas.
- DAVIS, F., 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology". In MIS Quarterly, v. 13 (3), pp. 319-340.
- DA CRUZ, A., 2014. "Refining Use Cases Through Temporal Relations". In International Conference on Software Paradigm Trends, v. 1, pp. 95 - 102.
- DA SILVA, T., SILVEIRA, M., 2010. "Validação de um Método para Identificação de Problemas de Usabilidade a partir de Diagramas UML". In Brazilian Symposium on Human Factors in Computing Systems, pp. 179 - 188.
- DE ANGELI, A., MATERA, M., COSTABILE, M. F., GARZOTTO, F., PAOLINI, P., 2003. "On the Advantages of a Systematic Inspection for Evaluating Hypermedia Usability". In International Journal of Human-Computer Interaction, v. 15(3), pp. 315 - 335.
- DE MELLO, R., PEREIRA, W., TRAVASSOS, G., 2010. "Activity Diagram Inspection on Requirements Specification". In Simpósio Brasileiro de Engenharia de Software, pp. 168-177.
- DE SA, M., CARRIÇO, L., 2008. "Lessons from early stages design of mobile applications". In International Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 127 - 136.
- DE SA, M., CHURCHILL, E., 2012. "Mobile augmented reality: Exploring design and prototyping techniques". In International Conference on Human-computer interaction with mobile devices and services, pp. 221 - 230.
- DONAHUE, G. M., 2001. "Usability and the Bottom Line". In Journal of IEEE Software, v. 18 (1), pp. 31 - 37.
- DYBA, T., KAMPENES, V., SJOBERG, D., 2006. "A systematic review of statistical power in software engineering experiments". In Information and Software Technology, v. 48 (8), pp. 745 - 755.
- FABO, P., DURIKOVIC, R., 2012. "Automated Usability Measurement of Arbitrary Desktop Application with Eyetracking". In International Conference on Information Visualisation, Montpellier, França, pp. 625 – 629.
- FAILY, S., FLÉCHAIS, I., 2010. "Barry is not the weakest link: eliciting secure system requirements with personas". In British Computer Society Conference on Human-Computer Interaction, pp. 124 - 132.

- FERNANDES, P., CONTE, T., BONIFÁCIO, B., 2012. "WE-QT: A Web Usability Inspection Technique to Support Novice Inspectors". In Brazilian Symposium on Software Engineering, pp. 11 - 20.
- FERNANDEZ, A., 2009. "WUEP: Un Proceso de Evaluación de Usabilidad Web Integrado en el Desarrollo de Software Dirigido por Modelos". Dissertação (Mestrado) - Universitat Politècnica de València, 173 páginas.
- FERNANDEZ, A., INSFRAN, E., ABRAHÃO, S., 2011a. "Usability evaluation methods for the web: A systematic mapping study". In Information and Software Technology, v.53 (8), pp. 789 - 817.
- FERNANDEZ, A., ABRAHÃO, S., INSFRAN, E., 2011b. "A web usability evaluation process for model-driven web development". In International Conference Advanced Information Systems Engineering, pp. 108-122.
- FERNANDEZ, A., ABRAHÃO, S., INSFRAN, E., MATERA, M., 2012a. "Further Analysis on the Validation of a Usability Inspection Method for Model-Driven Web Development". In International Symposium on Empirical Software Engineering and Measurement, pp. 153 - 156.
- FERNANDEZ, A., 2012b. "A Usability Inspection Method for Model-driven Web Development Processes". Tese (Doutorado) - Universitat Politècnica de València, 371 páginas.
- FERNANDEZ, A., ABRAHÃO, S., INSFRAN, E., 2013. "Empirical validation of a usability inspection method for model-driven Web development". In The Journal of Systems and Software, v. 86, pp. 161 - 186.
- FERRÉ, X., JURISTO, N., WINDL, H., CONSTANTINE, L., 2001. "Usability Basics for Software Developers". In: Journal of IEEE Software, v. 18 (1), pp. 22 - 29.
- FERRÉ, X., JURISTO, N., MORENO, A., 2004. "Improving Software Engineering Practice with HCI Aspects". In International Conference on Software Engineering Research and Applications, Los Angeles, v. 3026, pp. 349 - 363.
- FERRÉ, X., JURISTO, N., MORENO, A., 2005. "Framework for Integrating Usability Practices into the Software Process". In International Conference Product Focused Software Process Improvement, v. 3574, pp. 202 - 215.
- FLEISS, J., 1981. "Statistical Methods for Rates and Proportions". John Wiley & Sons, Nova York, 2ª edição, 800 páginas.
- FØLSTAD, A.; KNUTSEN, J., 2010. "Online user feedback in early phases of the design process: Lessons learnt from four design cases". In Advances in Human-Computer Interaction, artigo ID 956918, 9 páginas.
- FURTADO, E., FURTADO, V., VASCONCELOS, E., 2007. "A conceptual framework for the design and evaluation of affective usability in educational geosimulation systems". In International Conference on Human-Computer Interaction, pp.497-510.
- GARLAND, R., 1991. "The Mid-Point on a Rating Scale: Is it Desirable?" In Marketing Bulletin 2, 1991, pp. 66 - 70.
- GOLDEN, E., 2009. "Helping software architects design for usability". In Symposium on Engineering interactive computing systems, pp. 317 - 320.
- GORDILLO, A.; BARRA, E.; AGUIRRE, S.; QUEMADA, J., 2014. "The usefulness of usability and user experience evaluation methods on an eLearning platform development from a developer's perspective: A case study". In Frontiers in Education Conference, pp. 1 - 8.
- HASSENZAHN, M., DIEFENBACH, S., GÖRITZ, A., 2010. "Needs, affect, and interactive products – Facets of user experience". In Interacting with computers, v. 22 (5), pp. 353 - 362.

- HAYNES, S., 2009. "It's what it's in: Evaluating the usability of large-scale integrated systems". In *Extended Abstracts on Human Factors in Computing Systems*, pp. 3051-3060.
- HORNBÆK, K., HØEGH, R. T., PEDERSEN, M., STAGE, J., 2007. "Use Case Evaluation (UCE): A Method for Early Usability Evaluation in Software Development". In *International Conference on Human-Computer Interaction*, pp. 578 - 591.
- HORNBÆK, K., 2010. "Dogmas in the assessment of usability evaluation methods". In *Behaviour & Information Technology*, v. 29 (1), pp. 97 - 111.
- ISSA, T., ISAIAS, P., 2015. "Sustainable Design: HCI, Usability and Environmental Concerns". Springer, Londres, 181 páginas.
- ISO/IEC, 1998. ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals, Geneva.
- ISO/IEC, 2002. ISO/IEC Standard 12207:1995: International Standard: Information Technology. Software Life Cycle Processes. Amendment 1. Amd.1:2002. ISO, Geneva, Switzerland.
- ISO/IEC, 2011. ISO/IEC 25010: Systems and software engineering - SquaRE - Software product Quality Requirements and Evaluation: System and Software Quality Models).
- IVORY, M., HEARST, M., 2001. "The state of the art in automating usability evaluation of user interfaces". In *ACM Computing Surveys* 33 (4), pp. 470 - 516.
- JOHN, B., BASS, L., GOLDEN, E., STOLL, P., 2009. "A responsibility-based pattern language for usability-supporting architectural patterns". In *Symposium on Engineering interactive computing systems*, pp. 3 - 12.
- JOHNS, R., 2005. "One Size Doesn't Fit All: Selecting Response Scales For Attitude Items". In *Journal of Elections, Public Opinion, and Parties*, v. 15 (2), pp. 237 - 264.
- JURISTO, N., MORENO, A., 2001. "Basics of Software Engineering Experimentation". Kluwer Academic Publishers, 420 páginas.
- JURISTO, N., MORENO, A., SÁNCHEZ, M., BARANAUSKAS, M., 2007. "A Glass Box Design: Making the Impact of Usability on Software Development Visible". In *Proc. of the International Conference on Human-Computer Interaction*, Rio de Janeiro, v. 4663, p. 541 - 554.
- KARPPINEN, K.; LIINASUO, M., 2008. "Usability promotion in a technical project with sparse resources - A case study". In *Workshop on the Interplay between Usability Evaluation and Software Development*.
- KIERAS, D. E., 2009. "Model-based evaluation", In *Sears, A. and Jacko, J. (Eds.), The Human-Computer Interaction: Development Process*, pp. 294 - 310.
- KITCHENHAM, B., CHARTERS, S., 2007. "Guidelines for performing systematic literature reviews in software engineering". In *EBSE Technical Report EBSE-2007-01*, Software Engineering Group Department of Computer Science Keele University.
- KITCHENHAM, B., 2008. "The role of replications in empirical Software Engineering – a word of warning". In *Empirical Software Engineering*. v. 13 (2), pp. 219 - 221.
- LANUBILE, F., MALLARDO, T., CALEFATO, F., 2003. "Tool support for Geographically Dispersed Inspection Teams". In *Software Process Improvement and Practice*, v. 8, pp. 217 - 231.
- LAUGHERY, K. R., 1989. "Micro SAINT - A tool for modeling human performance in systems". In *G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. Van Breda (Eds.), Applications of human performance models to system design*. Nova York: Plenum Press, pp. 219 - 230.
- LAUESEN, S., 1997. "Usability Engineering in Industrial Practice". In *International Conference on Human-Computer Interaction*, Sydney, v. 2, pp. 15 - 22.

- LEWIS, C., 1982. Using the "Thinking Aloud Method". In *Cognitive Interface Design*. IBM T.J. Watson Research Center, 12 páginas.
- LUNA, E., PANACH, J., GRIGERA, J., ROSSI, G., PASTOR, O., 2010. "Incorporating usability requirements in a test/model-driven web engineering approach". In *Journal of Web Engineering*, v. 9 (2), pp. 132 - 156.
- MAFRA, S., 2006. "Definição de uma Técnica de Leitura Baseada em Perspectiva (OO-PBR) apoiada por Estudos Experimentais". Dissertação de Mestrado, COPPE/UFRJ, Rio de Janeiro.
- MAGUIRE, M., 2001. "Methods to support human-centred design. In *International Journal of Human- Computer Studies*", v. 55(4), pp. 587 - 634.
- MANN, H., WHITNEY, D., 1947. "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other". In *Annals of Mathematical Statistics*, v. 18, pp. 50 - 60.
- MATERA, M., COSTABILE, M., GARZOTTO, F., PAOLINI, P., 2002. "SUE Inspection: An Effective Method for Systematic Usability Evaluation of Hypermedia". In *IEEE Transactions on Systems, Man and Cybernetics, Part A*, v. 32 (1), pp. 93 - 103.
- MATERA, M., RIZZO, F., CARUGHI, G. T., 2006. "Web Usability: Principles and Evaluation Methods". In: Mendes, E., Mosley, N. (eds), *Web Engineering*, Chapter 5, New York, Springer Verlag.
- MEMMEL, T., REITERER, H., 2008. "Inspector: method and tool for visual UI specification". In *International Conference on Computer Aided Design of User Interfaces*, pp. 161-174.
- MENDES, M., FURTADO, E., 2014. "The DIRCE approach for Requirements Elicitation and Design of Interaction in innovative systems". In *Latin American Computing Conference*, pp. 1-10.
- MOLINA, F., TOVAL, A., 2009. "Integrating usability requirements that can be evaluated in design time into Model Driven Engineering of Web Information Systems". In *Journal Advances in Engineering Software, Oxford*, v. 40 (12), pp. 1306 - 1317.
- MORAGA, M. A., CALERO, C., PIATTINI, M., DIAZ, O., 2007. "Improving a portlet usability model. *Software Quality Control*". In *Journal Software Quality Control*, Hingham, v. 15 (2), pp. 155 - 177.
- MOSHAGEN, M., THIELSCH, M. T., 2013. "A short version of the visual aesthetics of websites inventory". *Behaviour & Information Technology*, v. 32(12), pp. 1305-1311.
- MÜLLER, M. J., HASLWANTER, J.H., DAYTON, T., 1997. "Participatory Practices in the Software Lifecycle". In: Helander, Martin G.; Landauer, Thomas K.; Prabhu, Prasad V. (eds), *Handbook of Human-Computer Interaction*, 2a edição, Elsevier, pp. 255 - 297.
- NAMAHN, 2001. "Using eye tracking for usability testing: a research note by Namahn", 12 páginas.
- NIELSEN, J., 1989. "What do users really want?". In *International Journal of Human-Computer Interaction*, v. 1(2), pp. 137 - 147.
- NIELSEN, J., 1993. "Usability engineering". Boston, MA, USA: AP Professional.
- NIELSEN, J., 1994a. "Guerrilla HCI: using discount usability engineering to penetrate the intimidation barrier". Academic Press, Inc., Orlando, pp. 245 - 272.
- NIELSEN, J., 1994b. "Heuristic evaluation". In: Jakob Nielsen, Mack, R. L. (eds), *Usability inspection methods, Heuristic Evaluation*, New York, pp. 25 - 62.
- NORMAN, D., DRAPER S., 1986. "User centered system design: New Perspectives in Human-Computer Interaction". In Hillsdale N. J., Lawrence Erlbaum , 540 páginas.
- OLSEN, M., HEDMAN, J., VATRAPU, R., 2011. "E-wallet properties". In *International Conference on Mobile Business*, pp. 158 - 165.

- OLSINA, L., ROSSI, G., 2002. "Measuring Web Application Quality with WebQEM". In *Journal of IEEE Multimedia*, v. 9 (4), pp. 20 - 29.
- OR, C., TAO, D., 2012. "Usability study of a computer-based self-management system for older adults with chronic disease". In Ed. Gunther Eysenbach. In *Journal of Medical Internet Research*, v. 1(2), artigo 13.
- ORMENO, Y., PANACH, J., CONDORI-FERNANDEZ, N., PASTOR, O., 2013. "Towards a proposal to capture usability requirements through guidelines". In *International Conference on Research Challenges in Information Science*, Paris, pp. 1-12.
- PAELKE, V., NEBE, K., 2008. "Integrating agile methods for mixed reality design space exploration". In *Conference on Designing interactive systems*, pp. 240-249.
- PALANQUE, P., BASNYAT, S., NAVARRE, D., 2007. "Improving interactive systems usability using formal description techniques: application to healthcare". In *Human-computer Interaction and Usability Engineering of the Austrian Computer Society Conference on HCI and Usability for Medicine and Health Care*, pp. 21 - 40.
- PANACH, J., VALVERDE, F., PASTOR, O., 2007. "Improvement of a Web Engineering Method through Usability Patterns". In *International Conference on Web Information Systems Engineering*, v. 4832, pp. 441 - 446.
- PANACH, J.; CONDORI-FERNÁNDEZ, N.; VALVERDE, F.; AQUINO, N.; PASTOR, O., 2008. "Towards an early usability evaluation for web applications". In *International Conference on Software Process and Product Measurement*, pp. 32-45.
- PANACH, J., JURISTO, N., PASTOR, O., 2013. "Including functional usability features in a Model-Driven Development method". In *Computer Science and Information Systems*, v. 10 (3), pp. 999 - 1024.
- PANACH, J., JURISTO, N., VALVERDE, F., PASTOR, Ó., 2015. "A framework to identify primitives that represent usability within Model-Driven Development methods". In *Journal Information and Software Technology*, v. 58, pp. 338 - 354.
- PANKRATIUS, V., 2011. "Automated usability evaluation of parallel programming constructs". In *International Conference on Software Engineering*, pp. 936 - 939.
- PEISCHL, B., FERK, M., HOLZINGER, A., 2013. "Integrating user-centred design in an early stage of mobile medical application prototyping a case study on data acquisition in health organisations". In *International Conference on e-Business*, pp. 1 - 11.
- PETERSEN, K., VAKKALANKA, S., KUZNIARZ, L., 2015. "Guidelines for conducting systematic mapping studies in software engineering: An update". In *Information and Software Technology*, v. 64, pp. 1 - 18.
- PINELLE, D., GUTWIN, C., 2002. "Groupware walkthrough: adding context to groupware usability evaluation. In *Conference on Human Factors in Computing Systems*, pp. 455-462.
- PRATES, R., BARBOSA, S., 2003. "Avaliação de Interfaces de Usuário – Conceitos e Métodos". In *Jornada de Atualização em Informática do Congresso da Sociedade Brasileira de Computação*, Campinas, Capítulo 6, v. 2, pp. 245 - 293.
- PRESSMAN, R.S., 2006. "Engenharia de Software", McGraw-Hill, 6ª edição, 780 páginas.
- POLSON, P., LEWIS, C., RIEMAN, J., WHARTON, C., 1992. "Cognitive walkthroughs: a method for theory-based evaluation of user interfaces". In *International Journal of Man-Machine Studies*, v. 36, n. 5, pp. 741 - 773.
- RAFLA, T., ROBILLARD, P., DESMARAIS, M., 2007. "A method to elicit architecturally sensitive usability requirements: its integration into a software development process". In *Software Quality Journal*, v. 15 (2), pp. 117 - 133.
- RAMLÍ, R., JAAFAR, A., 2010. "Design and development of e-RUE as a web-based evaluation tool". In *International Symposium in Information Technology*, pp. 1268 - 1273.

- ROCHA, H., BARANAUSKAS, M., 2003. "Design e Avaliação de Interfaces Humano-Computador". NIED, UNICAMP, Campinas, 244 páginas.
- RIIHIAHO, S., 2002. "The pluralistic usability walk-through method". In *Ergonomics in Design: The Quarterly of Human Factors Applications*, v. 10(3), pp. 23-27.
- RIVERO, L., CONTE, T., 2012. "Using an Empirical Study to Evaluate the Feasibility of a New Usability Inspection Technique for Paper Based Prototypes of Web Applications". In *Brazilian Symposium on Software Engineering*, pp. 81 - 90.
- RIVERO, L., CONTE, T., 2013. "Improving usability inspection technologies for web mockups through empirical studies". In *International Conference on Software Engineering and Knowledge Engineering*, pp. 172 - 177.
- RYU, Y., LEE, Y., SMITH-JACKSON, T., NUSSBAUM, M., TOMIOKA, K., 2006. "User centered design and testing methods for the development of the menu system for mobile phones". In *Human Factors and Ergonomics Society Annual Meeting*, v. 50 (18), pp. 2183 - 2186.
- SALVADOR, C., NAKASONE, A., POW-SANG, J., 2014. "A systematic review of usability techniques in agile methodologies". In *Euro American Conference on Telematics and Information Systems*, artigo 17.
- SANTOS, F., CONTE, T., 2011. "Evoluindo um Assistente de Apoio à Inspeção de Usabilidade através de Estudos Experimentais". In *Ibero-American Conference on Software Engineering*, Rio de Janeiro, v. 1, pp. 197 - 210.
- SANTOS, G., ROCHA, A., CONTE, T., BARCELLOS, M., PRIKLADNICKI, R., 2012. "Strategic Alignment between Academy and Industry: A Virtuous Cycle to Promote Innovation in Technology". In *Brazilian Symposium on Software Engineering*, Natal, v. 1, pp. 196 - 200.
- SAUER, C., JEFFERY, D., LAND, L., YETTON, P., 2000. "The Effectiveness of Software Development Technical Review: A Behaviorally Motivated Program of Research". In *IEEE Transactions on Software Engineering*, v. 26 (1), pp. 1 - 14.
- SEFFAH, A., METZKER, E., 2004. "The obstacles and myths of usability and software engineering". *Commun ACM*, v. 47 (12), pp. 71 - 76.
- SEFFAH, A., DONYAEE, M.; KLINE, R.; PADDA, H. K., 2006. "Usability measurement and metrics: A consolidated model". In *Software Quality Journal*, v. 14(2), pp. 159 - 178.
- SEFFAH, A., DJOUAB, R., ANTUNES, H., 2011. "Comparing and Reconciling Usability-Centered and Use Case-Driven Requirements Engineering Processes". In *Australasian Conference on User interface*, Los Alamitos, pp. 132 - 139.
- SHACKEL, B., 1991. "Whence and Where - A Short History of Human-Computer Interaction". In H-J. Bullinger (Ed) *Human Aspects in Computing: Design and Use of Interactive, Systems, and Work with Terminals*; Amsterdam, Elsevier, pp. 3 - 18.
- SHNEIDERMAN, B., PLAISANT, C., 2010. "Designing the User Interface: Strategies for Effective Human-Computer Interaction". Pearson, 6a edição, 606 páginas.
- SHULL, F., CARVER, J., TRAVASSOS, G. H., 2001. "An empirical methodology for introducing software processes". *ACM SIGSOFT Software Engineering Notes*, v. 26 (5), pp. 288 - 296.
- SHULL, F., CARVER, J., VEGAS, S., JURISTO, N., 2008. "The role of replications in Empirical Software Engineering". In *Empirical Software Engineering*, v. 13 (2), pp. 211 - 218.
- SIEK, K., KHAN, D. U., ROSS, S. E., HAVERHALS, L. M., MEYERS, J., CALI, S. 2011. "Designing a Personal Health Application for Older Adults to Manage Medications: A Comprehensive Case Study". In *Journal of Medical Systems*, v. 35 (5), pp. 1099 - 1121.

- SWIERENGA, S. J., ABUJARAD, F., DENNIS, T., POST, L., 2011. "Real-world user-centered design: the michigan workforce background check system". In International Conference on Human-Computer Interaction, pp. 325 - 334.
- SILVA, W., VALENTIM, N., CONTE, T., 2014a. "Mapeamento Sistemático para a unificação das Tecnologias de ES e IHC através da Inclusão da Usabilidade no processo de desenvolvimento de software". In USES Technical Report USES TR-USES-2014-0001, Manaus. Disponível em: <http://uses.icomp.ufam.edu.br/wp-content/uploads/2017/02/TR-USES-2014-0001.pdf>
- SILVA, W., VALENTIM, N., CONTE, T., 2014b. "Projetando Diagramas de Atividade visando a Usabilidade de Aplicações Interativas". In Simpósio Brasileiro Sobre Fatores Humanos em Sistemas Computacionais, Foz do Iguaçu, v. 1. pp. 349 - 352.
- SILVA, W.; VALENTIM, N.; CONTE, T., 2015. "Integrating the Usability into the Software Development Process: A Systematic Mapping Study". In International Conference on Enterprise Information Systems, Barcelona, pp. 105 - 113.
- SIVAJI, A., ABDULLAH, M. R., DOWNE, A. G., AHAMD, W., 2013. "Hybrid Usability Methodology: Integrating Heuristic Evaluation with Laboratory Testing across the Software Development Lifecycle". In International Conference on Information Technology: New Generations, pp. 375 - 383.
- SOMMERVILLE, I., 2007. "Engenharia de Software". São Paulo: Pearson – Addison Wesley, 8ª Edição, 552 páginas.
- SPAN, M., SMITS, C., GROEN-VAN DE VEN, L., JUKEMA, J., VERNOOIJ-DASSEN, M., HETTINGA, M., JANSSEN, R., EEFSTING, J., 2014. "An interactive web tool to facilitate shared decision making in dementia: Design issues perceived by caregivers and patients". In International Journal on Advances in Life Sciences, v. 6 (3-4), pp. 107 - 121.
- STOLL, P., BASS, L., GOLDEN, E., JOHN, B., 2009. "Supporting usability in product line architectures". In International Software Product Line Conference, pp. 241-248.
- Guide to the Software Engineering Body of Knowledge (SWEBOK), 2004. Institute of Electrical and Electronics Engineers.
- TOGNAZZINI, B., 1992. "TOG on Interface". Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, ISBN: 020160842, 331 páginas.
- TOGNAZZINI, B., 2014. "First Principles of Interaction Design (Revised & Expanded)". Disponível em: <http://asktog.com/atc/principles-of-interaction-design/>. Acessado em: 20 de Junho de 2016.
- TOHIDI, M., BUXTON, W., BAECKER, R., SELLEN, A., 2006a. "Getting the right design and the design right". In Conference on Human Factors in Computing Systems, pp. 1243 - 1252.
- TOHIDI, M., BUXTON, W., BAECKER, R., SELLEN, A., 2006b. "User sketches: A quick, inexpensive, and effective way to elicit more reflective user feedback". In Nordic Conference on Human-computer Interaction: changing roles, pp. 105 - 114.
- TRAVASSOS, G., SHULL, F., FREDERICKS, M., BASILI, V., 1999. "Detecting defects in object-oriented designs: using reading techniques to increase software quality". In ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications, v. 34 (10), pp. 47 - 56.
- TRAVASSOS, G., SHULL, F., CARVER, J., BASILI, V., 2002. "Reading Techniques for OO Design Inspections". University of Maryland, 56 páginas.
- Usability Guidelines (USABILITYGUIDELINES), 2016. U. S. Department of Health & Human Services. Disponível em: <https://webstandards.hhs.gov/guidelines/>. Acessado em: 20 de Junho de 2016.
- VALENTIM, N., DA SILVA, T., SILVEIRA, M., CONTE, T., 2013. "Estudo Comparativo entre Técnicas de Inspeção de Usabilidade sobre Diagramas de

- Atividades". In Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais, Manaus, pp. 92 - 101.
- VALENTIM, N., CONTE, T., 2014. "Improving a Usability Inspection Technique Based on Quantitative and Qualitative Analysis". In Simpósio Brasileiro de Engenharia de Software (SBES 2014), Maceió, pp. 171 - 180.
- VALENTIM, N., MALDONADO, J., CONTE, T., 2015a. "Evaluating an Inspection Technique for Use Case Specifications: Quantitative and Qualitative Analysis". In International Conference on Enterprise Information Systems, v. 3, pp.13 - 24.
- VALENTIM, N., CONTE, T., ESTACIO, B., PRIKLADNICKI, R., 2015b. "How do software engineers apply an early usability inspection technique? A qualitative study". In International Conference on Software Engineering and Knowledge Engineering, Pittsburgh, USA: KSI Research Inc. and Knowledge Systems Institute Graduate School, v. 1. pp. 686 - 691.
- VALENTIM, N., RABELO, J., ORAN, A., CONTE, T., 2015c. "A Controlled Experiment with Usability Inspection Techniques Applied to Use Case Specifications: Comparing the MIT 1 and the UCE Techniques". In ACM/IEEE International Conference on Model Driven Engineering Languages and Systems, Canadá, v.1, pp. 206 - 215.
- VALENTIM, N., SILVA, W., CONTE, T., 2017a. "Mapeamento Sistemático para a geração de um Framework que Projetam e/ou Avaliam a Usabilidade nos Estágios Iniciais do Processo de Desenvolvimento de Software". USES Technical Report TR-USES-2017-0003, Manaus. Disponível em: <http://uses.icomp.ufam.edu.br/wp-content/uploads/2016/02/TR-USES-2017-0003.pdf>
- VALENTIM, N., SILVA, W., CONTE, T., 2017c. "Material de suporte do Framework para a Antecipação da Usabilidade". In USES Technical Report USES TR-USES-2017-0005, Manaus. Disponível em: <http://uses.icomp.ufam.edu.br/wp-content/uploads/2017/02/TR-USES-2017-0005.pdf>
- VAZ, V., TRAVASSOS, G., CONTE, T., 2012. "Empirical assessment of WDP tool: A tool to support web usability inspections". In Conferencia Latinoamericana en Informática, pp. 1 - 9.
- VELMOUROGAN, S., DHAVACHELVAN, P., BASKARAN, R., RAVIKUMAR, B., 2014. "Software development Life cycle model to build software applications with usability". In International Conference on Advances in Computing, Communications and Informatics, pp. 271 - 276.
- VENKATESH, V., BALA, H., 2008. "Technology Acceptance Model 3 and a Research Agenda on Interventions". In Decision Sciences, v. 39 (2), pp. 273 - 315.
- VERMEEREN, A., LAW, E., ROTO, V., OBRIST, M., HOONHOUT, J., VÄÄNÄNEN-VAINIO-MATTILA, K., 2010. "User experience evaluation methods: current state and development needs". In Nordic Conference on Human-Computer Interaction: Extending Boundaries, pp. 521 - 530.
- WANIA, C., ATWOOD, M., 2009. "Pattern languages in the wild: exploring pattern languages in the laboratory and in the real world". In International Conference on Design Science Research in Information Systems and Technology, artigo 12.
- WÖHLIN, C., RUNESON, P., HÖST M., OHLSSON, M., REGNELL, B., WESSL, A., 2000. "Experimentation in software engineering: an introduction". Kluwer Academic Publishers, 236 páginas.
- WYK, E., DE VILLIERS, R., 2008. "Usability context analysis for virtual reality training in South African mines". In Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries: Riding the Wave of Technology, pp. 276 - 285.

YEN, P.-Y.; BAKKEN, S., 2012. "Review of health information technology usability study methodologies". In *Journal of the American Medical Informatics Association*, v. 19(3), pp. 413 - 422.

## APÊNDICE A – TÉCNICA MIT 1

MIT 1 – Avaliação de Usabilidade em Casos de Uso	
<b>Objetivo:</b>	Verificar a usabilidade em Casos de Uso
<b>Instruções:</b>	<p>1. Antes de começar a inspeção, observe as seguintes diretrizes:</p> <ol style="list-style-type: none"> <li>Caso o Caso de Uso possua <b>alto nível de detalhamento</b> (como mensagens de erro, textos informativos, advertências, nome de telas, nome de campos, nome de botões, nome de links, nome de opções, entre outros), comece a inspeção utilizando as heurísticas da <b>MIT1 - Alto Detalhamento</b>.</li> <li>Caso o Caso de Uso possua <b>baixo nível de detalhamento</b> (ou seja, não descreve as mensagens de erro, textos informativos, advertências, nome de telas, nome de campos, nome de botões, nome de links, nome de opções, entre outros), comece a inspeção utilizando as heurísticas da <b>MIT1 - Baixo Detalhamento</b>.</li> <li>Caso o Caso de Uso possua <b>nível médio de detalhamento</b> (com apenas <b>alguns</b> dos itens citados acima), utilize as heurísticas da <b>MIT1 - Alto Detalhamento</b>.</li> </ol> <p>2. Para cada Caso de Uso a ser inspecionado observe e anote em qual FP<sup>4</sup>, FA<sup>5</sup>, FE<sup>6</sup> ou RN<sup>7</sup> você identifica um problema de usabilidade seguindo as heurísticas abaixo:</p>
<b>MIT1 - Alto Detalhamento</b>	
<b>1AA. Visibilidade do Status do Sistema</b>	
<b>1AA1</b>	Verifique se há algum texto no FP, FA e FE que informa em que parte do sistema o usuário se encontra;
<b>1AA2</b>	Verifique se há algum texto no FP, FA e FE que informa ao usuário o que foi realizado após uma persistência de dados. Por exemplo: quando há alteração ou exclusão de algo, uma mensagem de texto é apresentada.
<b>1AB. Concordância entre o sistema e o mundo real</b>	
<b>1AB1</b>	Verifique se os nomes de campos, telas, botões, links, mensagens de erros e textos informativos no FP, FA, FE e RN apresentam conceitos familiares aos usuários, ou seja, segue as convenções do mundo real;
<b>1AB2</b>	Verifique se as opções ou campos informados pelo sistema no FP, FA e FE estão apresentados em uma ordem natural e lógica segundo os conceitos do domínio do problema.
<b>1AC. Controle e liberdade ao usuário</b>	
<b>1AC1</b>	Verifique se o usuário, através do FA e FE, pode desfazer ou refazer algo que envolva persistência de dados no sistema. Por exemplo: pode excluir ou alterar dados inseridos.
<b>1AD. Consistência e padrões</b>	
<b>1AD1</b>	Verifique se todos os campos, telas, botões e links possuem nomes que indicam

<sup>4</sup> FP: Fluxo Principal

<sup>5</sup> FA: Fluxo Alternativo

<sup>6</sup> FE: Fluxo Exceção

<sup>7</sup> RN: Regras de Negócio

	funcionalidades diferentes no FP, FA, FE e RN, ou seja, não tenham nomes diferentes, mas com a mesma funcionalidade;
<b>1AD2</b>	Verifique se as mensagens de erro são apresentadas de forma consistente com os padrões apresentados no FP <sup>8</sup> , FA <sup>9</sup> e FE <sup>10</sup> ;
<b>1AD3</b>	Verifique se os nomes de campos, telas, botões e links no FP, FA, FE e RN <sup>11</sup> apresentam um único sentido, sem ambiguidades.
<b>1AE. Prevenção de Erros</b>	
<b>1AE1</b>	Verifique se há alguma advertência do sistema que alerta o usuário através de mensagens ou textos informativos de que o que ele está fazendo pode ser inapropriado naquele momento no FP, FA, FE;
<b>1AE2</b>	Verifique se todas as opções, botões e links disponíveis possuem nomes que definem claramente que resultados ou estados serão atingidos no FP, FA, FE e RN.
<b>1AE3</b>	Verifique se os dados ou opções obrigatórias FP, FA, FE e RN estão claramente definidos;
<b>1AF. Reconhecer ao invés de lembrar</b>	
<b>1AF1</b>	Verifique se os nomes das opções, campos, telas e links são informados de forma que o usuário não tenha que se lembrar quais são eles no FP, FA, FE e RN.
<b>1AG. Flexibilidade e eficiência de uso</b>	
<b>1AG1</b>	Verifique se há nomes de links, nome de telas ou nome de campos abreviados para usuários mais experientes no FP, FA, FE e RN;
<b>1AG2</b>	Verifique se há como passos serem realizados de forma mais rápida e eficiente para usuários mais experientes no FP, FA, FE e RN;
<b>1AG3</b>	Verifique se o sistema permite que o usuário realize o que ele deseja com facilidade através dos passos do FA e FE;
<b>1AG4</b>	Verifique se o sistema permite ao usuário formas de acesso às tarefas principais no FA e FE. Por exemplo: estando no FA ou FE o usuário consegue voltar ao FP quando for viável;
<b>1AG5</b>	Verifique se as formas de acesso oferecidas pelo sistema no FA e FE diminuem o esforço de ações físicas dos usuários.
<b>1AH. Reconhecimento, diagnóstico e recuperação de erros</b>	
<b>1AH1</b>	Verifique se as mensagens de erros utilizam linguagem simples de ser entendida pelos usuários, ou seja, de acordo com o perfil do usuário no FP, FA e FE. Por exemplo: mensagens de erro que não tenha códigos;
<b>A1H2</b>	Verifique se todas as mensagens de erros não culpam ou intimidam o usuário no FP, FA, FE;
<b>A1H3</b>	Verifique se as mensagens de erro ajudam o usuário a corrigir o erro no FP, FA e FE. Por exemplo: mensagens com indicação de recuperação;
<b>A1H4</b>	Verifique se o sistema mostra como acessar as soluções alternativas que são apresentadas em algumas mensagens de erro no FP, FA e FE.

---

<sup>8</sup> FP: Fluxo Principal

<sup>9</sup> FA: Fluxo Alternativo

<sup>10</sup> FE: Fluxo Exceção

<sup>11</sup> RN: Regras de Negócio

<b>1AI. Proporcionar funções que são úteis para o usuário</b>	
<b>1AI1</b>	Verifique se todo campo, opção, tela, botão ou link que seja essencial está presente no FP <sup>12</sup> , FA <sup>13</sup> , FE <sup>14</sup> e RN <sup>15</sup> .
<b>1AI2</b>	Verifique se tudo que está presente no FP, FA, FE e RN é útil.
<b>MIT1 - Baixo Detalhamento</b>	
<b>1BA. Visibilidade do Status do Sistema</b>	
<b>1BA1</b>	Verifique no FP, FA, FE se há alguma informação em que parte do sistema o usuário se encontra;
<b>1BA2</b>	Verifique se no FP, FA e FE há alguma informação para o usuário do que foi realizado após uma persistência de dados.
<b>1BB. Concordância entre o sistema e o mundo real</b>	
<b>1BB1</b>	Verifique se as opções ou campos informados pelo sistema no FP, FA e FE estão apresentados em uma ordem natural e lógica segundo os conceitos do domínio do problema.
<b>1BC. Controle e liberdade ao usuário</b>	
<b>1BC1</b>	Verifique se o usuário, através do FA e FE, pode desfazer ou refazer algo que envolva persistência de dados no sistema. Por exemplo: pode excluir ou alterar dados inseridos.
<b>1BE. Prevenção de Erros</b>	
<b>1BE1</b>	Verifique se os dados ou opções obrigatórias estão claramente definidos;
<b>1BG. Flexibilidade e eficiência de uso</b>	
<b>1BG1</b>	Verifique se há como passos serem realizados de forma mais rápida e eficiente para usuários mais experientes no FP, FA, FE e RN;
<b>1BG2</b>	Verifique se o sistema permite que o usuário realize o que ele deseja com facilidade através dos passos do FA e FE;
<b>1BG3</b>	Verifique se o sistema permite ao usuário formas de acesso às tarefas principais no FA e FE. Por exemplo: estando no FA ou FE o usuário consegue voltar ao FP quando for viável;
<b>1BG4</b>	Verifique se as formas de acesso oferecidas pelo sistema no FA e FE diminuem o esforço de ações físicas dos usuários.
<b>1BI. Proporcionar funções que são úteis para o usuário</b>	
<b>1BI1</b>	Verifique se todo campo, opção, tela, botão ou link que seja essencial para o usuário está presente no FP, FA, FE e RN;
<b>1BI2</b>	Verifique se tudo que está presente no FP, FA, FE e RN é útil.

---

<sup>12</sup> FP: Fluxo Principal

<sup>13</sup> FA: Fluxo Alternativo

<sup>14</sup> FE: Fluxo Exceção

<sup>15</sup> RN: Regras de Negócio

## APÊNDICE B – TÉCNICA MIT 2

<b>MIT 2 – Avaliação de Usabilidade dos <i>Mockups</i></b>	
<b>Objetivo:</b>	Verificar a usabilidade dos <i>Mockups</i>
<b>Instruções:</b>	<p>1. Antes de começar a inspeção, observe as seguintes diretrizes:</p> <ol style="list-style-type: none"> <li>a. Caso o <i>Mockup</i> apresente as mensagens do sistema (mensagens de erro, textos informativos, advertências) comece a inspeção utilizando as heurísticas da <b>MIT2 - Alto Detalhamento</b>.</li> <li>b. Caso o <i>Mockup</i> <b>NÃO</b> apresente as mensagens do sistema, comece a inspeção utilizando as heurísticas da <b>MIT2 - Baixo Detalhamento</b>.</li> </ol> <p>2. Para cada <i>Mockups</i> a ser inspecionado observe e anote qual problema de usabilidade você encontrou seguindo as heurísticas a seguir:</p>
<b>MIT2 - Alto Detalhamento</b>	
<b>2AA. Visibilidade do Status do Sistema</b>	
<b>2AA1</b>	Verifique se há informações textuais ou nome nos <i>Mockups</i> que informa em que parte do sistema o usuário se encontra.
<b>2AA2</b>	Verifique se há algum texto informativo ou mensagem que informa ao usuário o que foi realizado após uma persistência de dados (alteração, exclusão, etc).
<b>2AB. Concordância entre o sistema e o mundo real</b>	
<b>2AB1</b>	Verifique se os símbolos, palavras, frases, nome de botões, links, mensagens de erros e textos informativos estão expressas em uma representação facilmente compreendida pelo usuário, ou seja, possuem uma linguagem familiar;
<b>2AB2</b>	Verifique se as opções do sistema estão apresentadas em uma ordem natural e lógica segundo os conceitos do domínio do problema.
<b>2AC. Controle e liberdade ao usuário</b>	
<b>2AC1</b>	Verifique se o usuário tem as opções de desfazer ou refazer algo que ele tenha escolhido;
<b>2AC2</b>	Verifique se o usuário tem a opção de cancelar o que está realizando, ou se há opções similares que permitam ao usuário utilizar saídas em caso de escolhas erradas ou para sair de um estado ou local não esperado.
<b>2AD. Consistência e padrões</b>	
<b>2AD1</b>	Verifique se a terminologia (nomenclatura), gráficos e símbolos estão consistentes com o sistema representado;
<b>2AD2</b>	Verifique se há padrões adotados em relação a layout, formatação e controles. Por exemplo, botão “Salvar” no canto superior direito em um <i>Mockup</i> e em outro no canto inferior direito;
<b>2AD3</b>	Verifique se todos os botões/links possuem nomes/ícones que indicam funcionalidades diferentes, ou seja, não há botões/links com nomes/ícones diferentes, mas com a mesma funcionalidade;
<b>2AD4</b>	Verifique se as mensagens de erro são apresentadas de forma consistente com os padrões.

<b>2AE. Prevenção de Erros</b>	
<b>2AE1</b>	Verifique se os dados obrigatórios na entrada de dados estão claramente definidos;
<b>2AE2</b>	Verifique se há alguma indicação do formato correto para uma entrada de dados específica;
<b>2AE3</b>	Verifique se há alguma advertência do sistema que alerta o usuário através de mensagens ou textos informativos de que o que ele está fazendo pode ser inapropriado naquele momento;
<b>2AE4</b>	Verifique se todas as opções, botões e links disponíveis possuem nomes que definem claramente que resultados ou estados serão atingidos.
<b>2AF. Reconhecer ao invés de lembrar</b>	
<b>2AF1</b>	Verifique se os nomes das opções, campos, botões e links são informativos e se os ícones/figuras minimizam o esforço físico e cognitivo do usuário de se lembrar;
<b>2AF2</b>	Verifique se é fácil reconhecer/visualizar a opção que deve ser usada para atingir o objetivo desejado;
<b>2AG. Flexibilidade e eficiência de uso</b>	
<b>2AG1</b>	Verifique se a disposição dos elementos aumenta a eficiência de uso, minimizando o esforço de ações físicas;
<b>2AG2</b>	Verifique se há apoio a tarefas específicas frequentemente repetidas;
<b>2AG3</b>	Verifique se há diferentes formas de acesso às tarefas principais, ou seja, possibilidade de utilização de aceleradores ou atalhos;
<b>2AG4</b>	Verifique se há facilidade para entrada de dados, sejam eles simples ou complexos estruturalmente;
<b>2AG5</b>	Verifique se há mecanismos de busca de informação que auxiliem a entrada de dados obrigatória;
<b>2AG6</b>	Verifique se há como as funcionalidades do sistema serem realizadas de forma mais rápida e eficiente para usuários mais experientes.
<b>2AH. Reconhecimento, diagnóstico e recuperação de erros</b>	
<b>2AH1</b>	Verifique se as mensagens de erros estão claramente visíveis para o usuário;
<b>2AH2</b>	Verifique se as mensagens de erros utilizam linguagem simples de ser entendida pelos usuários, ou seja, de acordo com o perfil do usuário.
<b>2AH3</b>	Verifique se todas as mensagens de erros não culpam ou intimidam o usuário;
<b>2AH4</b>	Verifique se as mensagens de erro ajudam o usuário a corrigir o erro.
<b>2AH5</b>	Verifique se o sistema mostra como acessar as soluções alternativas que são apresentadas em algumas nas mensagens de erro.
<b>MIT-2AI. Proporcionar funções que são úteis para o usuário</b>	
<b>2AI1</b>	Verifique se toda funcionalidade do sistema, botão ou link essencial está presente;
<b>2AI2</b>	Verifique se toda funcionalidade do sistema, botão ou link presente é útil.

## APÊNDICE C – TÉCNICA MIT 3

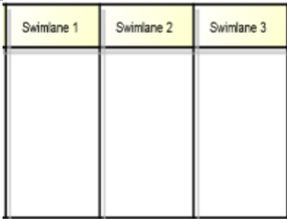
<b>MIT 3 – Avaliação de Usabilidade dos Diagramas de Atividades</b>	
<b>Objetivo:</b>	Verificar a usabilidade dos Diagramas de Atividades
<b>Instruções:</b>	1. Para cada Diagrama de Atividades a ser inspecionado observe e anote qual problema de usabilidade você encontrou seguindo as heurísticas baixas:
<b>3B. Concordância entre o sistema e o mundo real</b>	
<b>3B1</b>	Verifique se as atividades estão apresentadas em uma ordem natural e lógica segundo os conceitos do domínio do problema;
<b>3B2</b>	Verifique se os nomes das atividades utilizam termos (palavras) que seguem as convenções do mundo real, ou seja, que englobam tanto as convenções do domínio do problema quanto às convenções de terminologia de aplicações semelhantes.
<b>MIT-3C. Controle e liberdade ao usuário</b>	
<b>3C1</b>	Verifique se o usuário através de alguma atividade consegue alterar dados que ele inseriu através de outra atividade;
<b>3C2</b>	Verifique se o usuário através de alguma atividade consegue excluir dados que ele inseriu através de outra atividade.
<b>3C3</b>	Verifique se o usuário através de alguma atividade consegue cancelar, voltar ou desfazer algo.
<b>MIT-3D. Consistência e padrões</b>	
<b>3D1</b>	Verifique se todas as atividades indicam funcionalidades diferentes, ou seja, não há atividades com nomes diferentes, mas com a mesma funcionalidade.
<b>MIT-3E. Prevenção de Erros</b>	
<b>3E1</b>	Verifique se há alguma atividade em que o sistema previne algum tipo de erro.
<b>MIT-3F. Reconhecer ao invés de lembrar</b>	
<b>3F1</b>	Verifique se é fácil reconhecer a atividade que deve ser usada para atingir o objetivo desejado.
<b>MIT-3G. Flexibilidade e eficiência de uso</b>	
<b>3G1</b>	Verifique se o usuário possui diferentes formas de acesso às atividades principais;
<b>3G2</b>	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.
<b>MIT-3H. Reconhecimento, diagnóstico e recuperação de erros</b>	
<b>3H1</b>	Verifique através das atividades se o sistema ajuda o usuário a sair de uma situação de erro;
<b>3H2</b>	Verifique através das atividades se o sistema ajuda a corrigir o erro. Por exemplo: atividades que indicam a recuperação do erro.

<b>MIT-3I. Proporcionar funções que são úteis para o usuário</b>	
<b>3I1</b>	Verifique se todas as atividades são úteis;
<b>3I2</b>	Verifique se todas as atividades úteis estão presentes.

## APÊNDICE D – TÉCNICA UDRT-AD

### PROCEDIMENTOS

- Para aplicação da técnica, primeiramente devem ser realizadas uma leitura completa da descrição textual, de forma a entender o domínio do problema que é apresentado.
- Após a leitura, tente construir o Diagrama de Atividade utilizando os seguintes passos: Identificação dos Atores, Identificação do nó de Início, Identificação das Atividades, Agrupamento das Atividades, Transição das Atividades, Identificação dos Nós de Fim;
- Cada passo possui três itens que auxiliam na construção dos diagramas, que são:
  - a. As **Heurísticas** para identificar os elementos do diagrama de atividade, a partir da descrição textual, e como transformá-lo em um elemento do diagrama;
  - b. As **Instruções** para criar os elementos do diagrama; e,
  - c. As **Diretrizes de Usabilidade** apoiam a incluir a usabilidade em cada parte do processo de construção do diagrama;
- Em seguida, verifique a consistência do diagrama projetado através dos itens de verificação.

IDENTIFICAÇÃO DAS SWINLANES		
Heurística	Elemento	Instruções
<p><b>H1 – Atores envolvidos na Descrição</b></p> <p><i>Os atores podem ser identificados através das responsabilidades/atividades que podem assumir na descrição. São os responsáveis pela execução de alguma atividade.</i></p>	 <p style="text-align: center;"><i>Swimlane</i></p>	<p>✓ <b>Identifique os papéis</b> dos atores e os <b>transforme</b> em <i>swimlanes</i>:</p> <p>O nome das <i>swimlanes</i> deve estar de acordo com o papel que cada ator irá realizar no sistema;</p>



#### Diretrizes de Usabilidade:

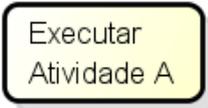
- ✓ Cada *swimlane* deve estar **descrita de forma correta e compreensível**, de acordo com o domínio do problema:
  - Caso seja necessário, modifique o nome da *swimlane* para que **o termo utilizado siga o padrão já utilizado** em aplicações semelhantes e também atenda o domínio do problema;

NÓ DE INÍCIO		
Heurística	Elemento	Instruções
<p><b>H2 – Nó de Início</b></p> <p><i>Os nós de início podem ser identificados através da primeira ação que um ator realiza para executar uma atividade.</i></p>	 <p style="text-align: center;"><i>Nó de Início</i></p>	<p>✓ <b>Insira</b> o nó de início na <i>swimlane</i> responsável pelo início das atividades;</p> <p>✓ O <b>primeiro fluxo</b> sairá do nó de início e irá até a <b>primeira atividade</b> identificada;</p>



#### Diretrizes de Usabilidade:

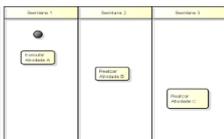
- ✓ Os fluxos do nó de início devem seguir a ordem lógica apresentada no domínio do problema;

IDENTIFICAÇÃO DAS ATIVIDADES		
Heurística	Elemento	Instruções
<p><b>H3 – Ações envolvidas na Descrição</b></p> <p>São as ações que o ator pode/deve realizar. As ações podem ser identificadas através de: (a) verbos que estão na 3ª pessoa do singular (faz, realiza, executa) ou (b) verbos no infinitivo (fazer, realizar, executar);</p>	 <p>Atividade</p>	<p>✓ <b>Transforme</b> cada ação identificada em uma <b>atividade</b>:</p> <ul style="list-style-type: none"> <li>○ <b>Reescreva</b> o verbo das ações identificadas <b>no infinitivo</b>, por exemplo: realiza → realizar;</li> <li>○ <b>Agrupe</b> o verbo ao <b>seu complemento</b>: realizar → “realizar cadastro do aluno”;</li> </ul>



#### Diretrizes de Usabilidade:

- ✓ Cada atividade deve estar **compreensível** e **descrita corretamente**, de acordo com o domínio do problema;
- ✓ Cada atividade deve ser de **fácil reconhecimento** de qual seu objetivo no diagrama;
- ✓ O nome das atividades deve estar **de acordo com a descrição textual**.
  - Caso seja necessário, modifique-a para que o termo utilizado siga o padrão já utilizado em aplicações semelhantes e também atenda o domínio do problema;
- ✓ O nome das atividades devem representar funcionalidades diferentes que vão ser realizadas na aplicação, ou seja, não deve existir atividade que possuem nomes diferentes, mas com a mesma funcionalidade;

AGRUPAMENTO DAS ATIVIDADES		
Heurística	Elemento	Instruções
<p><b>H4 – Alocação das atividades</b></p> <p>Cada atividade deve estar alocada ao seu swimlane responsável que deve realizá-la no diagrama.</p>	 <p>Alocação de Atividades</p>	<p>✓ <b>Insira</b> cada atividade se encontra em seu swimlane responsável.</p> <p>✓ <b>Verifique</b> se cada swimlane possui as suas respectivas atividades pelas quais são responsáveis.</p>



#### Diretrizes de Usabilidade:

- ✓ Caso exista alguma atividade em que seja necessário inserir, alterar ou excluir deve existir outra atividade que permita que o usuário possa voltar ou cancelar a ação realizada;
- ✓ Caso, em alguma atividade, o usuário estiver numa situação de erro, deve ser criada uma atividade que permita que ele consiga sair desta situação;
- ✓ Caso o usuário esteja numa situação de erro deve haver atividades que ajudam o usuário a corrigir o erro;
- ✓ Caso, em alguma atividade, exista uma tomada decisão, deve haver atividades perguntando para o usuário se o mesmo deseja realizar a ação;

TRANSIÇÃO DAS ATIVIDADES		
Heurística	Elemento	Instruções
<p><b>H5 – Fluxo Sequencial</b></p> <p>São as atividades realizadas sem sequência, ou seja, primeiro realiza uma atividade e, em seguida outra.</p>	 <p>Transição Sequencial</p>	<p>✓ <b>Direcione</b> uma seta simples da primeira atividade à segunda atividade;</p>



#### Diretrizes de Usabilidade:

- ✓ Os fluxos das atividades devem estar em uma ordem lógica e de acordo com o domínio do problema;

Heurística	Elemento	Instruções
<p><b>H6 – Condições de Decisão</b></p> <p><i>São as atividades que permitem o ator escolher entre duas ou mais ações nas quais deseja realizar.</i></p>	<p><b>Tomada de Decisão</b></p>	<ul style="list-style-type: none"> <li>✓ <b>Crie</b> uma condição e <b>direcione</b> uma seta simples da decisão as atividades que devem ser escolhidas;</li> <li>✓ <b>Crie</b> Condições de Guarda para cada atividade que resulta da condição de decisão;</li> </ul>



**Diretrizes de Usabilidade:**

- ✓ Os fluxos das atividades devem estar em uma ordem lógica e de acordo com o domínio do problema;
- ✓ Os termos utilizados devem ser de fácil compreensão e entendimento;
- ✓ Os nomes das condições de guarda devem ser de fácil reconhecimento pelos usuários;

Heurística	Elemento	Instruções
<p><b>H7 – Ponto de Merge</b></p> <p><i>Identifique se há atividades nas quais, independente das outras atividades estarem sendo realizadas, em certo momento se agrupam e tem como saída somente uma atividade.</i></p>	<p><b>Merge</b></p>	<ul style="list-style-type: none"> <li>✓ <b>Direcione</b> os fluxos que devem se encontrar para o ponto de merge;</li> <li>✓ <b>Direcione</b> o fluxo de saída do ponto de merge para outra atividade ou outro elemento do diagrama;</li> </ul>



**Diretrizes de Usabilidade:**

- ✓ Os fluxos das atividades devem estar em uma ordem lógica e de acordo com o domínio do problema;

Heurística	Elemento	Instruções
<p><b>H8 – Atividades em Paralelo – Bifurcação e Sincronização</b></p> <p><i>Estas são as ações identificadas na descrição que podem acontecer paralelamente. Ou seja, quando duas ou mais atividades necessitem ser realizadas ao mesmo tempo.</i></p>	<p><b>Bifurcação e Sincronização</b></p>	<ul style="list-style-type: none"> <li>✓ Na <b>Bifurcação insira</b> uma linha na horizontal para que esta englobe todas as atividades que devem ser executadas em paralelo; <ul style="list-style-type: none"> <li>○ <b>Direcione um fluxo</b> sequencial da última atividade para esta linha horizontal desenhada;</li> </ul> </li> <li>✓ Na <b>Sincronização insira</b> também uma linha na horizontal abaixo de todas as atividades que foram executadas em paralelo; <ul style="list-style-type: none"> <li>○ <b>Direcione todos os fluxos</b> das atividades em paralelo para esta linha horizontal;</li> <li>○ <b>Direcione um fluxo</b> sequencial da sincronização para a próxima atividade;</li> </ul> </li> </ul>



**Diretrizes de Usabilidade:**

- ✓ Os fluxos das atividades devem estar em uma ordem lógica e de acordo com o domínio do problema;

NÓ DE FIM		
Heurística	Elemento	Instruções
<p><b>H9 – Nós de Fim.</b></p> <p><i>É identificado somente quando o ator realiza a última ação no fluxo das atividades.</i></p>	 <p><b>Nó de Fim</b></p>	<p>✓ <b>Insira</b> um nó de fim:</p> <ul style="list-style-type: none"> <li>○ Caso alguma atividade necessita ser finalizada antes do término do fluxo;</li> <li>○ Caso o fluxo chegou ao ponto final;</li> </ul>



#### Diretrizes de Usabilidade:

- ✓ Os fluxos do nó de início devem seguir a ordem lógica apresentada no domínio do problema;

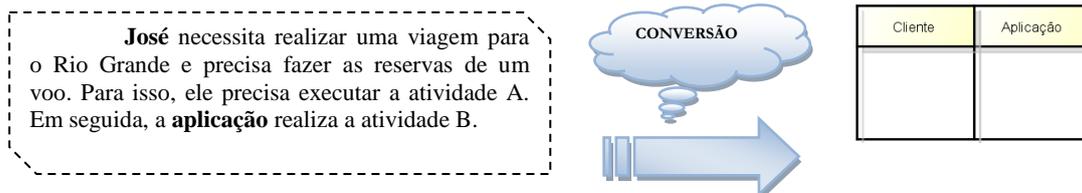
Inspeção do Diagrama Projetado	
Descrição	Itens de Verificação
<p>Esta inspeção deve ser realizada no final da construção, de forma a verificar se o diagrama criado com esta técnica está de acordo com a descrição textual utilizada;</p>	<ol style="list-style-type: none"> <li>1. Verifique se as atividades utilizadas no diagrama estão de acordo com as atividades que estão sendo informadas na descrição textual. <ul style="list-style-type: none"> <li>○ Verifique se os nomes das <i>swimlanes</i>/atividades estão de acordo com a descrição e são de fácil reconhecimento e entendimento;</li> <li>○ Verifique se está faltando alguma atividade que está presente na descrição e não está presente no diagrama;</li> <li>○ Verifique se o diagrama apresenta atividades que estão fora do escopo da descrição;</li> </ul> </li> <li>2. Verifique se os fluxos das atividades do diagrama estão de acordo com as atividades do diagrama: <ul style="list-style-type: none"> <li>○ Verifique se o fluxo das atividades foi modelado de acordo com o apresentado na descrição;</li> <li>○ Verifique se as decisões estão sendo utilizadas no local correto;</li> <li>○ Verifique se as condições de guarda estão inseridas de maneira correta;</li> <li>○ Verifique se as atividades paralelas estão sendo utilizadas de forma concorrente e não sequencialmente;</li> </ul> </li> </ol>

# APÊNDICE E – EXEMPLOS DE APLICAÇÃO DA TÉCNICA UDRT-AD

## EXEMPLO DE APLICAÇÃO DAS HEURÍSTICAS PARA CONSTRUÇÃO DOS DIAGRAMAS DE ATIVIDADES

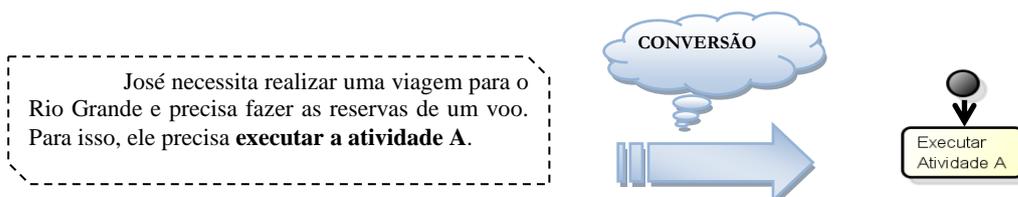
### H1 – Atores envolvidos na Descrição

**Exemplo 1:** Identificação de swimlanes para o diagrama



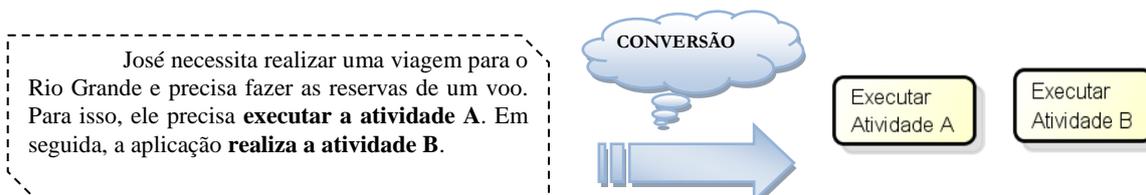
### H2 – Nó de Início

**Exemplo 2:** Criação do nó inicial do diagrama

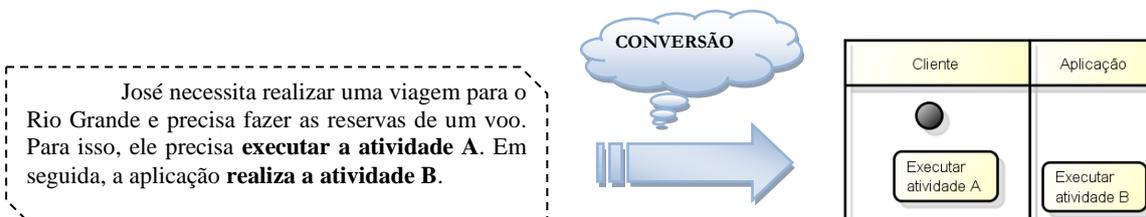


### H3 – Ações envolvidas na Descrição

**Exemplo 3:** Identificação das ações das atividades do diagrama.



### H4 – Alocação das atividades

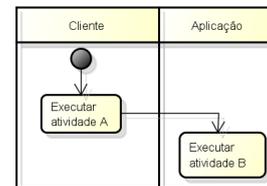


**Exemplo 4:** Alocação de cada atividade em seu swimlane responsável.

### H5 – Fluxo Sequencial

**Exemplo 5:** Identificação do fluxo sequencial das atividades do diagrama

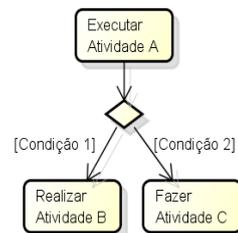
José necessita realizar uma viagem para o Rio Grande e precisa fazer as reservas de um voo. Para isso, ele precisa **executar a atividade A**. Em seguida, a aplicação **realiza a atividade B**.



### H6 – Condições de Decisão

**Exemplo 6:** Exemplo de uma condição do diagrama

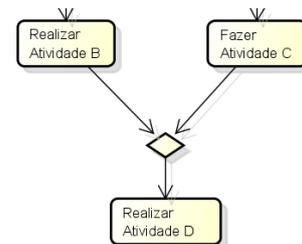
José necessita realizar uma viagem para o Rio Grande e precisa fazer as reservas de um voo. Para isso, ele precisa **executar a atividade A**. Em seguida, José escolhe entre **realizar a atividade B** ou **faz a atividade C**.



### H7 – Ponto de Merge

**Exemplo 7:** Exemplo de um ponto de Merge no diagrama

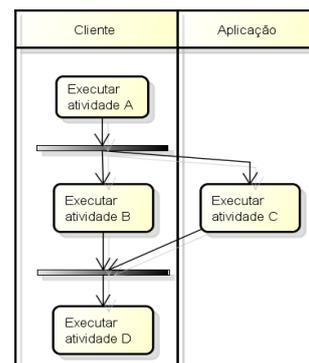
José necessita realizar uma viagem para o Rio Grande e precisa fazer as reservas de um voo. Para isso, ele precisa **executar a atividade A**. Em seguida, José escolhe entre **realizar a atividade B** ou **fazer a atividade C**. No entanto, independente do caminho que José seguir, ele sempre **realiza a Atividade D**.



### H8 – Atividades em Paralelo – Bifurcação e Sincronização

**Exemplo 8:** Exemplo de atividades paralelas.

José necessita realizar uma viagem para o Rio Grande e precisa fazer as reservas de um voo. Para isso, ele precisa **executar a atividade A**. Enquanto José **realiza a atividade B**, a **atividade C é executada pela aplicação**. Após **realizar estas duas atividades**, José deve **executar a Atividade D**.



**H9 – Nós de Fim.**

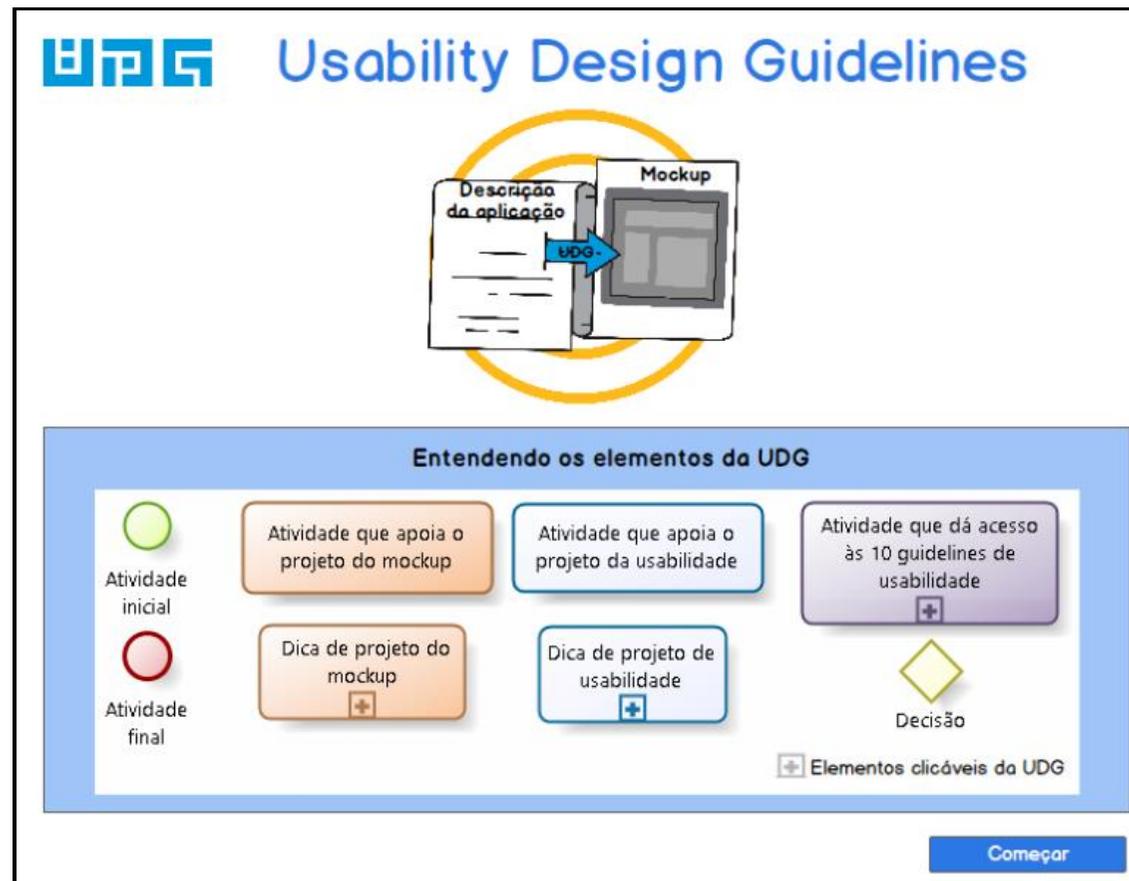
**Exemplo 9:** *Exemplo de criação do nó final.*

José necessita realizar uma viagem para o Rio Grande e precisa fazer as reservas de um voo. Para isso, ele precisa **executar a atividade A**. Enquanto José **realiza a atividade B**, a **atividade C é executada pela aplicação**. Após **realizar estas duas atividades**, José **deve executar a Atividade D** para finalizar a atividade de realizar reserva.

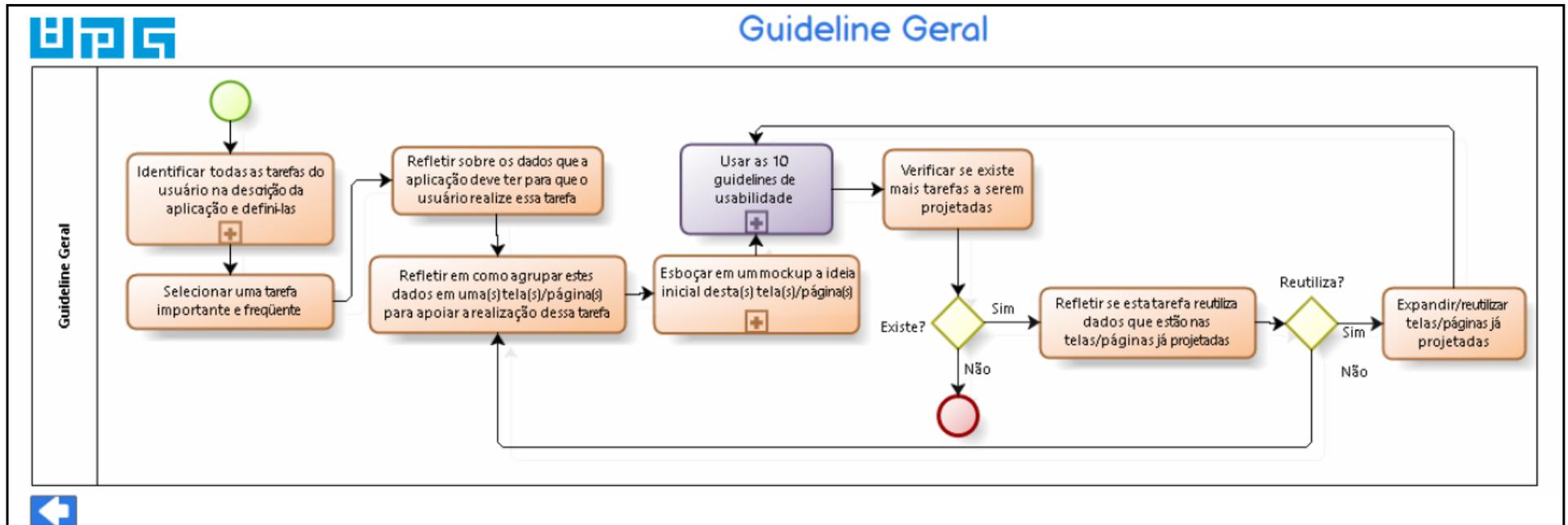


# APÊNDICE F – TÉCNICA UDG

## Tela 1



## Tela 2



## Tela 3

[Voltar para a guideline](#)

### **Dicas - Tarefas da aplicação:**

As **tarefas da aplicação** podem ser obtidas através das seguintes etapas adaptadas de Snyder (2003):

1. Liste os **objetivos e realizações do usuário** que podem ser alcançados usando a aplicação. Dica: O que as pessoas usarão nesta aplicação? O que é importante para eles? Como é que a sua aplicação tomará a vida deles melhor?
2. Liste suas **questões** - problemas, riscos e outras preocupações sobre a aplicação.
3. Priorize suas questões. Dica: Olhe para a sua lista e se pergunte: "Qual destas realizações são mais importantes para o sucesso da aplicação?"
4. Escolha uma **realização do usuário** que abrange uma ou mais questões importantes e transforme-os em uma tarefa usando o seguinte template.
  - a. **Nome da tarefa e número.** Dê a cada tarefa um nome descritivo breve e um número.
  - b. **Objetivos/saídas.** O que os usuários terão realizado quando tiver terminado a tarefa? Como eles vão saber que a tarefa está completa?
  - c. **Entradas.** Liste todas as informações ou os recursos tangíveis e intangíveis que um usuário precisa para completar esta tarefa.
  - d. **Premissas.** As premissas são as condições e pré-requisitos que estão em vigor no início da tarefa.
  - e. **Passos.** Anote os passos que você espera que o usuário passe para completar a tarefa. Isso ajuda a identificar os pedaços do mockup que você precisará criar. Dica: já anote alguns campos que você acha que o mockup necessitará.

Repita a etapa 4 até que você tenha um conjunto de tarefas que cubra suas questões.

Referência: Carolyn Snyder. "Paper Prototyping: The Fast and Easy Way to Define and Refine User Interfaces". Morgan Kaufmann; 1 edition, 2003, 408 páginas.

[Voltar para a guideline](#)

## Tela 4

[Voltar para a guideline](#)
[Próxima dica](#)

### Dicas - Tela/página: (1/2)

O que uma tela/página normalmente tem?

- A **marca** ou o **logotipo** no alto, à esquerda da página.
- A **barra de navegação** mais importante no alto, com links dispostos horizontalmente, ou no lado esquerdo da página, com links dispostos verticalmente.
- A **ferramenta de busca** interna no alto, na área central ou na parte direita da página.
- **Área de conteúdo**, na área central da página, onde pode ser localizada com mais facilidade.
- **Menus** dropdown e **atalhos** para páginas localizadas em áreas mais profundas no alto da área de conteúdo ou ao longo do lado esquerdo




[Voltar para a guideline](#)
[Próxima dica](#)

Referência: Barbara Ballard. "Designing the Mobile User Experience". Wiley; 1 edition, 2007, 260 páginas.

## Tela 5

[Voltar para a guideline](#)

### **Dicas - Elementos da tela/página: (2/2)**

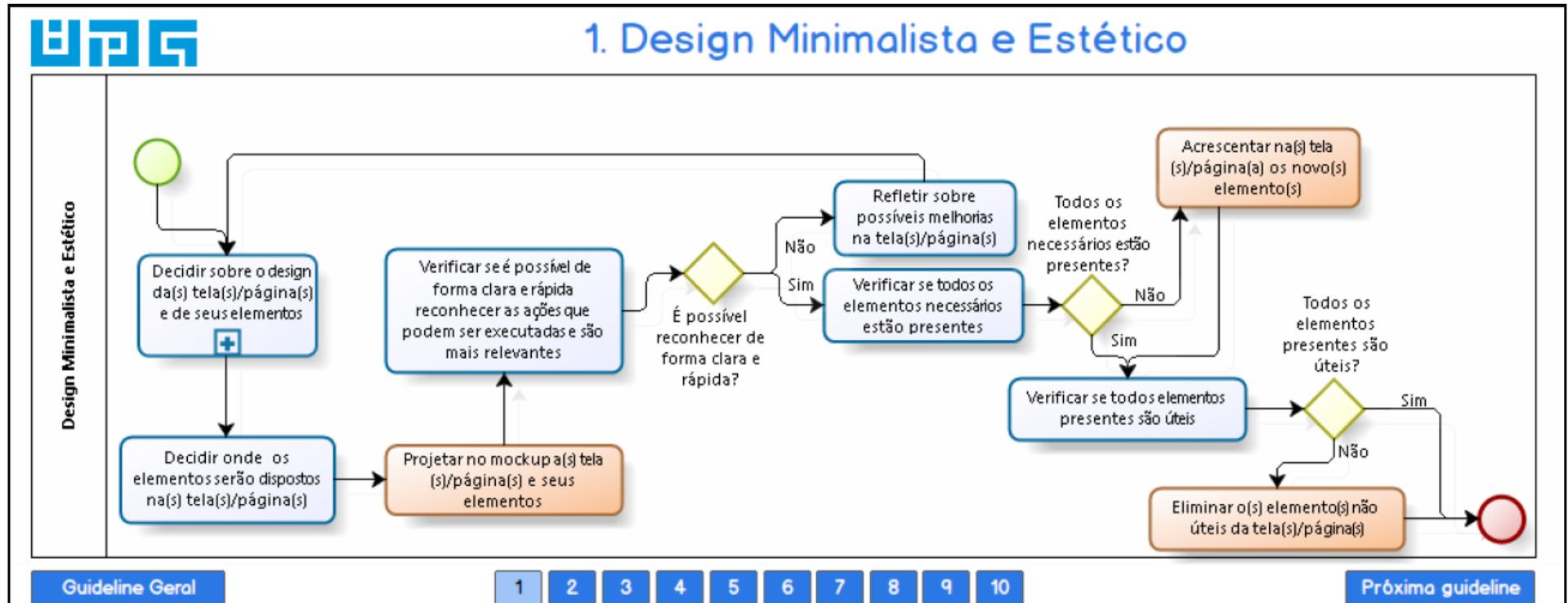
Elementos da tela/página são **botões, campos, links, contoles, textos informativos**, entre outros.

Alguns elementos utilizados na tela/página pode ser do tipo:

Form controls	Name
	Command button
	Combo box
	Check box
	Spin button
	List box
	Option button
	Scroll bar
	Text box
	Label
	Image

[Voltar para a guideline](#)

## Tela 6



## Tela 7

[Voltar para a guideline](#)[Próxima dica](#)

### **Dicas - Design minimalista e estético da tela/página: (1/2)**

- A **tela/página não** devem conter **informação irrelevante** ou raramente necessária.
- Combinação de cores, tamanho, fonte do texto e do fundo **não deve tornar a leitura difícil**.
- O **texto** deve ser posicionado de modo a ser **visível em qualquer situação**.
- A **tela/página inicial** deve transmitir a **qualidade da aplicação**. Ela deve de forma bem visível comunicar o propósito e valor da aplicação, enfatizar o que a aplicação oferece e como ela difere de suas principais concorrentes. Você deve limitar a quantidade de texto na página inicial.
- Não criar telas/páginas com **elementos desordenados e em excesso**, pois isto pode diminuir o desempenho do usuário ao tentar encontrar certas informações na aplicação.
- O layout da tela/página deve ajudar os usuários a **encontrar e utilizar a informação mais importante**. As informações devem ser apresentadas na ordem em que é mais útil para os usuários.
- Quanto ao tamanho da tela/página, **use telas/páginas mais curtas** para as **telas/páginas de destino** e **telas/páginas de navegação**, ou qualquer tela/página que requer uma navegação rápida. Use **telas/páginas mais longas** quando for necessária **leitura ininterrupta**, para coincidir com a estrutura de um documento em papel ou para as tornar conveniente para download e impressão.
- **Comunicar quaisquer alterações previstas no design da tela/aplicação** para os usuários antes de acontecer a alteração.

[Voltar para a guideline](#)[Próxima dica](#)

## Tela 8

[Voltar para a guideline](#)

### **Dicas - Design minimalista e estético dos elementos: (2/2)**

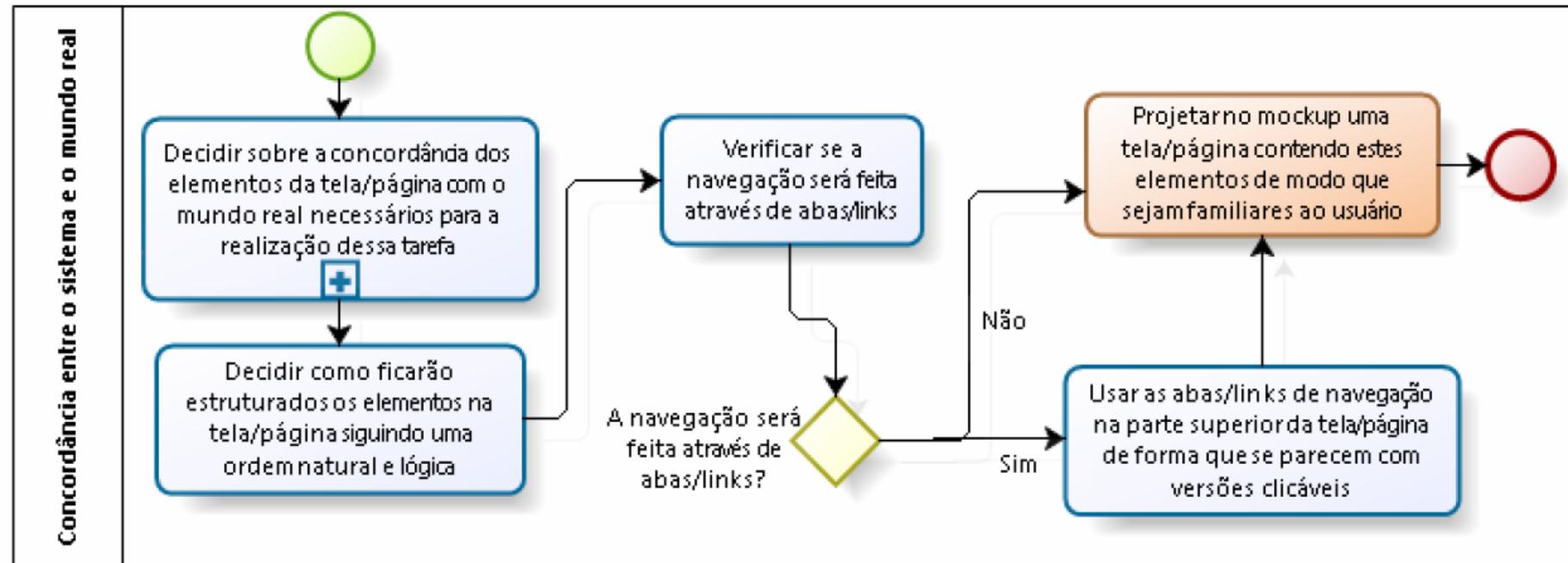
- **Links, botões e controles** devem aparecer somente se suas **ações associadas podem ser realizadas**.
- Use **elementos grandes** para **funções importantes** (botões grandes são mais rápidos).
- Usar no elemento um **tamanho de fonte** que possa ser **lido em displays padrões**.
- Usar no elemento uma **cor de fonte com alto contraste** em relação ao seu fundo.
- Usar no elemento uma **fonte** que seja **fácil de ser lida**.
- Use **pequenos elementos para funções** que você prefere que os **usuários não executem**.
- Qualquer elemento na página inicial, que deve imediatamente **atrair a atenção dos usuários**, deve ser colocado na **parte superior**. Todo o conteúdo, navegação e opções críticas devem estar no topo da tela/página.
- Coloque os **menus de navegação principais** no **painel esquerdo**.
- As **tabelas de dados** devem ter **títulos** de linha e coluna claros, **concisos** e **precisos**.
- Use **frames** quando determinadas **funções devem permanecer visíveis na tela** quando o usuário acessa outras informações da aplicação. Porém, mais de três frames em uma tela/página pode ser confuso para os usuários menos experientes com aplicações.
- Use os recursos para **atrair a atenção do usuário com cautela** e apenas quando eles são altamente relevantes.

[Voltar para a guideline](#)

## Tela 9



## 2. Concordância entre o sistema e o mundo real



Guideline Geral

1

2

3

4

5

6

7

8

9

10

Próxima guideline

## Tela 10

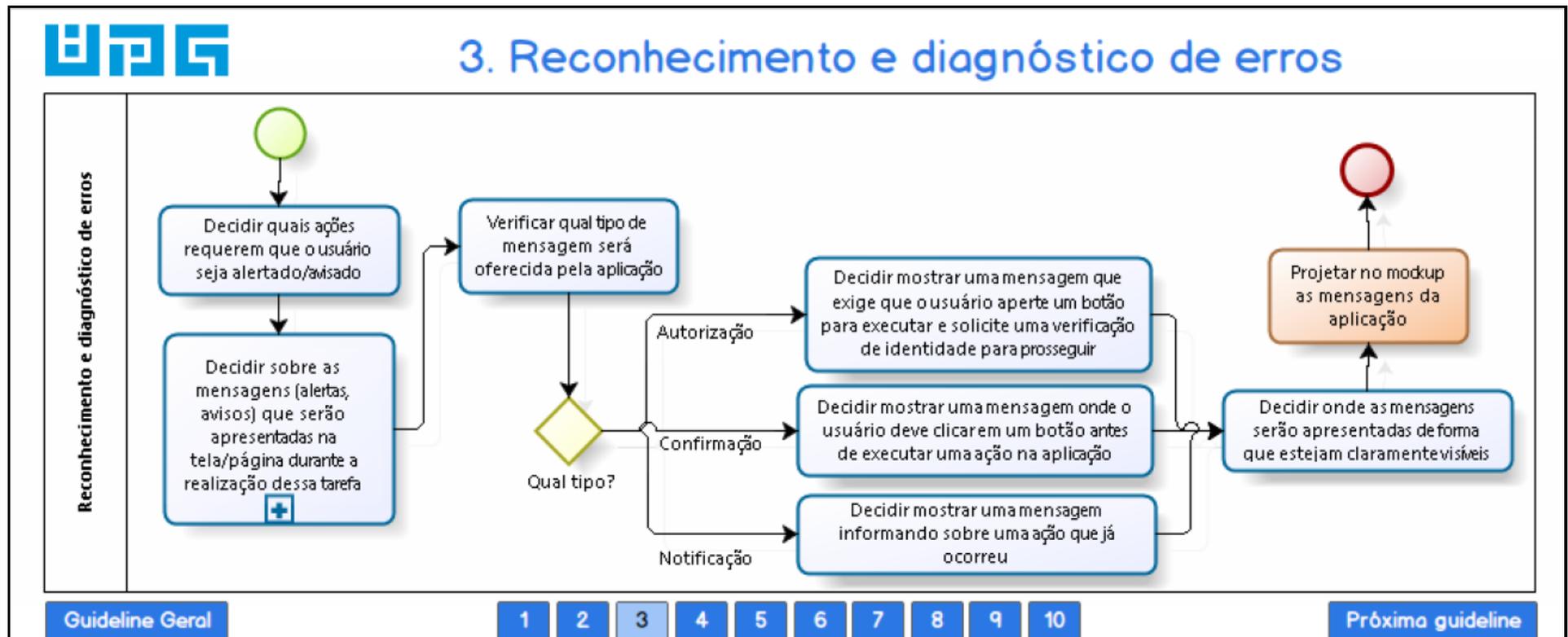
[Voltar para a guideline](#)

### **Dicas - Concordância entre o sistema e o mundo real:**

- Usar nos elementos **metáforas do mundo real** para ajudar a tornar a **interação mais natural** para o usuário. Ou seja, a aplicação deve utilizar a linguagem do usuário, ao invés de termos orientados ao sistema.
- Usar nos elementos um **tamanho, cor e tipo de fonte** que sigam as normas e **convenções mais familiares para os usuários**.
- Usar nestes elementos uma **navegação que seja familiar ao usuário**. Exemplo: links, abas. Porém tome cuidado, os usuários podem ficar confusos sobre o uso de abas quando elas não parecem ser abas. Além disso, aba são usadas quando a navegação entre as telas/páginas pode acontecer em qualquer ordem. Já o passo-a-passo, a navegação deve seguir uma ordem.
- Os **dados e informações** devem ser exibidos em um formato que **não requer conversão pelo usuário**.

[Voltar para a guideline](#)

## Tela 11



## Tela 12

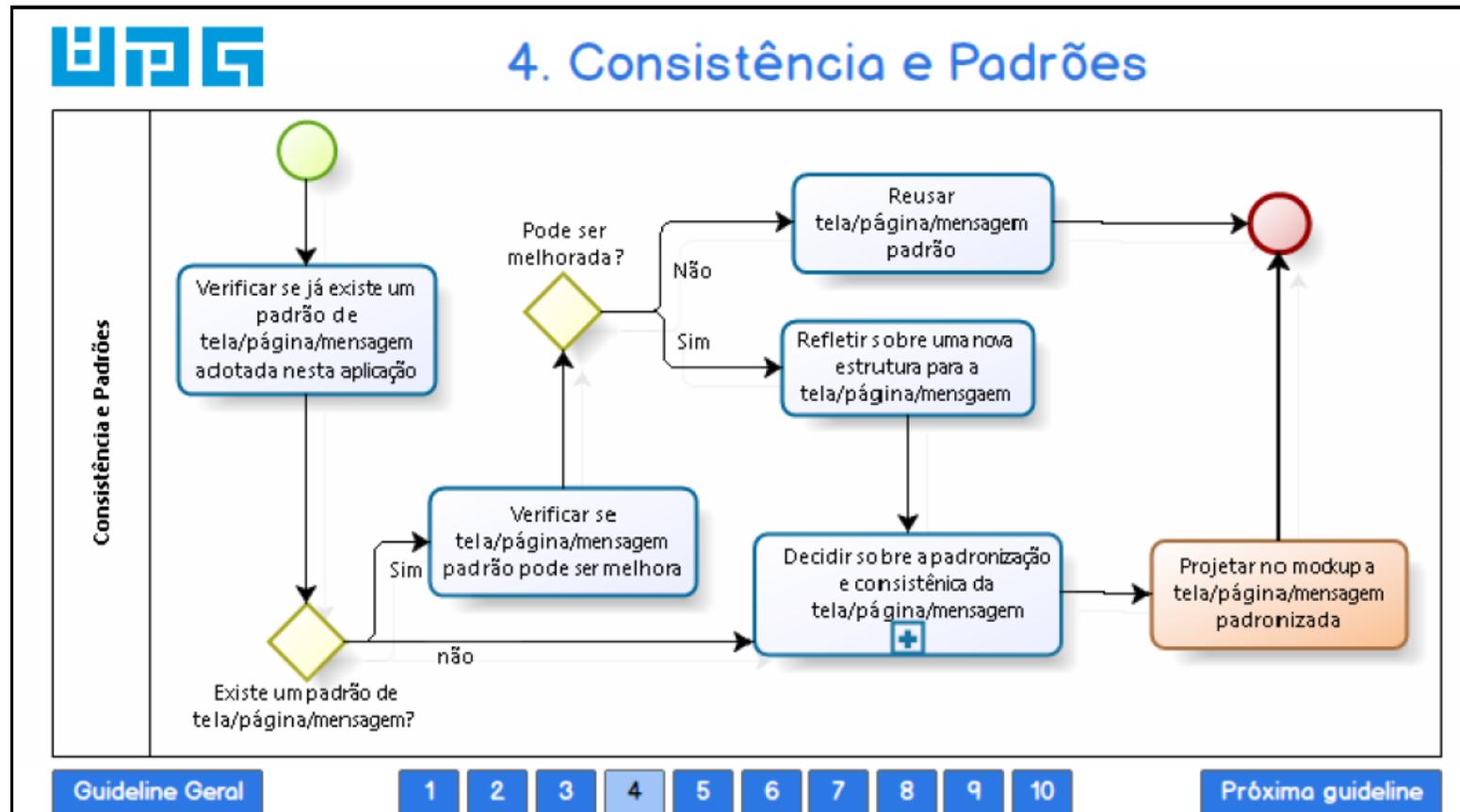
[Voltar para a guideline](#)

### **Dicas - Reconhecimento e diagnóstico de erros:**

- As **mensagens** devem **orientar a interação do usuário** com a aplicação.
- As **mensagens** devem ser **úteis e significativas** para o usuário interagir corretamente com a aplicação e corrigir o erro.
- As **mensagens** devem ser expressas em **linguagem clara (sem códigos)** indicando precisamente o problema ou situação.
- As **mensagens não** devem **culpar ou intimidar o usuário**.
- As **mensagens** devem mostrar como **acessar as soluções alternativas** para sair do estado de erro.

[Voltar para a guideline](#)

## Tela 13



## Tela 14

[Voltar para a guideline](#)[Próxima dica](#)

### **Dicas - Padronização da tela/página: (1/3)**

- O layout da tela/página deve buscar ter a mesma estrutura que a tela/página padrão, diferenciando apenas de dispositivo para dispositivo.



### **Dicas - Consistência da tela/página:**

- Os **títulos** devem **corresponder a ações** que foram realizadas para acessar eles.
- Os **elementos da tela/página** devem ser **posicionados de forma consistente**. Por exemplo: alinhe visualmente os elementos da página, verticalmente ou horizontalmente.

[Voltar para a guideline](#)[Próxima dica](#)

## Tela 15

Voltar para a guideline
Próxima dica

### **Dicas - Padronização dos elementos da tela/página: (2/3)**

- Um **elemento da tela/página** deve ter sempre a mesma **representação/símbolo, terminologia/nome, fonte e cor** quando usado em outra tela/página.
- Um **elemento da tela/página** deve ter sempre o **mesmo comportamento** quando usado em outra tela/página.
- Um **elemento da tela/página** deve estar na **mesma ordem/sequência/posição** quando usado em outra tela/página.
- **Elementos diferentes** devem ter uma representação/símbolo, terminologia/nome que indique **funcionalidades diferentes**.
- **Elementos diferentes** devem ter **comportamentos diferentes**.
- Opção de **Cancelar**
  - É padrão ter um botão Cancelar que fecha qualquer caixa de diálogo e descarta quaisquer alterações que o usuário possa ter feito dentro dessa caixa de diálogo.
  - Família Windows coloca o Cancelar depois do OK/Salvar 
  - Família Apple coloca o Cancelar antes do OK/Salvar.
- Opção de **Voltar**
  - O logotipo da aplicação em todas as telas/páginas pode permitir que o usuário volte para a tela/página inicial. Isso faz com que os usuários se sintam seguros e protegidos.
  - Em aplicações móveis, o usuário pode voltar também através de uma aba chamada Tela Inicial ou através do menu com a opção Início.

Referências:

<https://www.nngroup.com/articles/reset-and-cancel-buttons/>

<http://asktog.com/atc/principles-of-interaction-design/>

<https://www.nngroup.com/articles/ok-cancel-or-cancel-ok/>

Voltar para a guideline
Próxima dica

## Tela 16

[Voltar para a guideline](#)

### **Dicas - Padronização da mensagem: (3/3)**

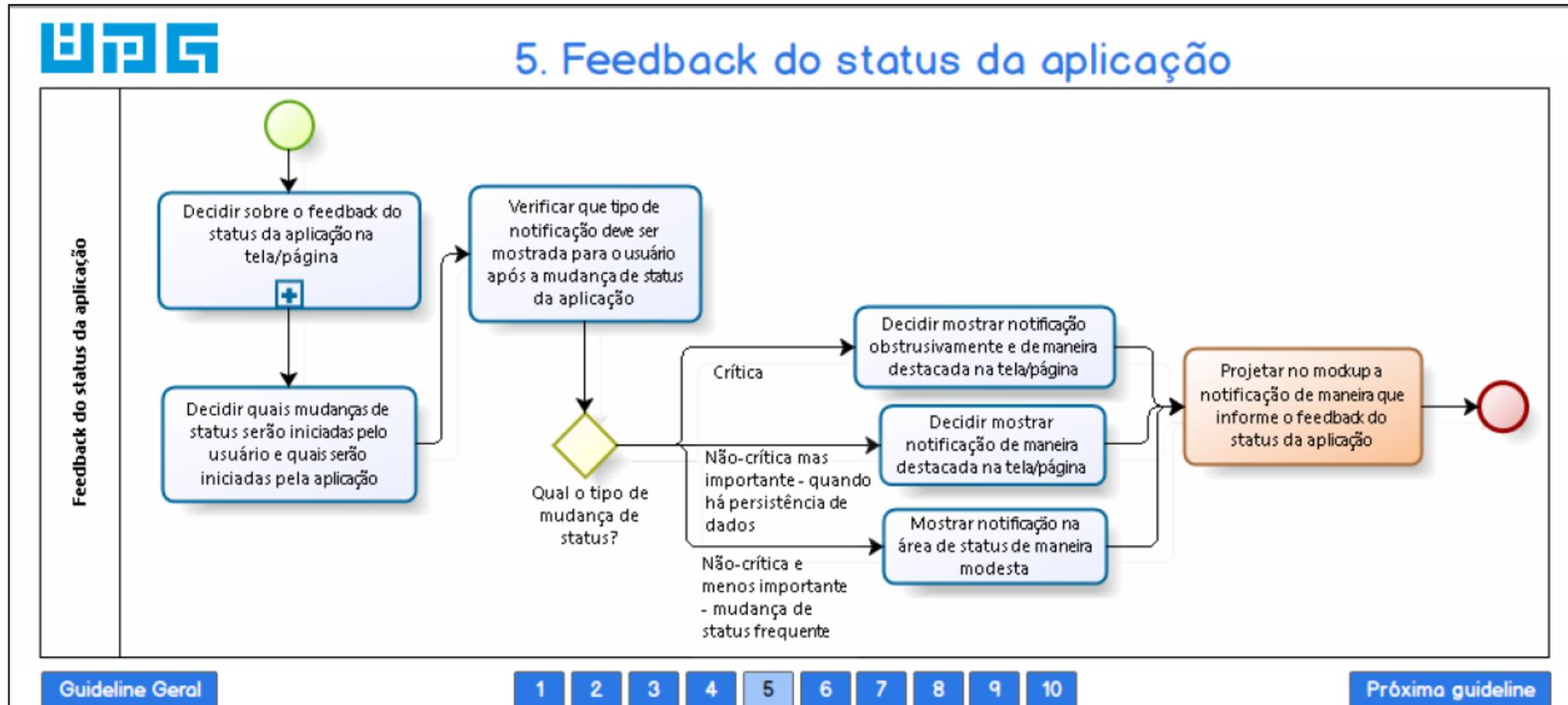
- O **layout da mensagem** deve **ter a mesma estrutura** de acordo com os seu tipo (autorização, confirmação ou notificação).
- Uma **mensagem** deve **ter sempre o mesmo conteúdo** quando aparecer em outra situação semelhante a qual ela foi acionada.

### **Dicas - Consistência da mensagem:**

- O **título da mensagem** deve **corresponder as ações que foram realizadas para se chegar a ela**.
- As **mensagens** de devem ser **apresentadas/posicionadas de forma consistente**. Por exemplo: sobreposta à tela página e centralizada.

[Voltar para a guideline](#)

## Tela 17



## Tela 18

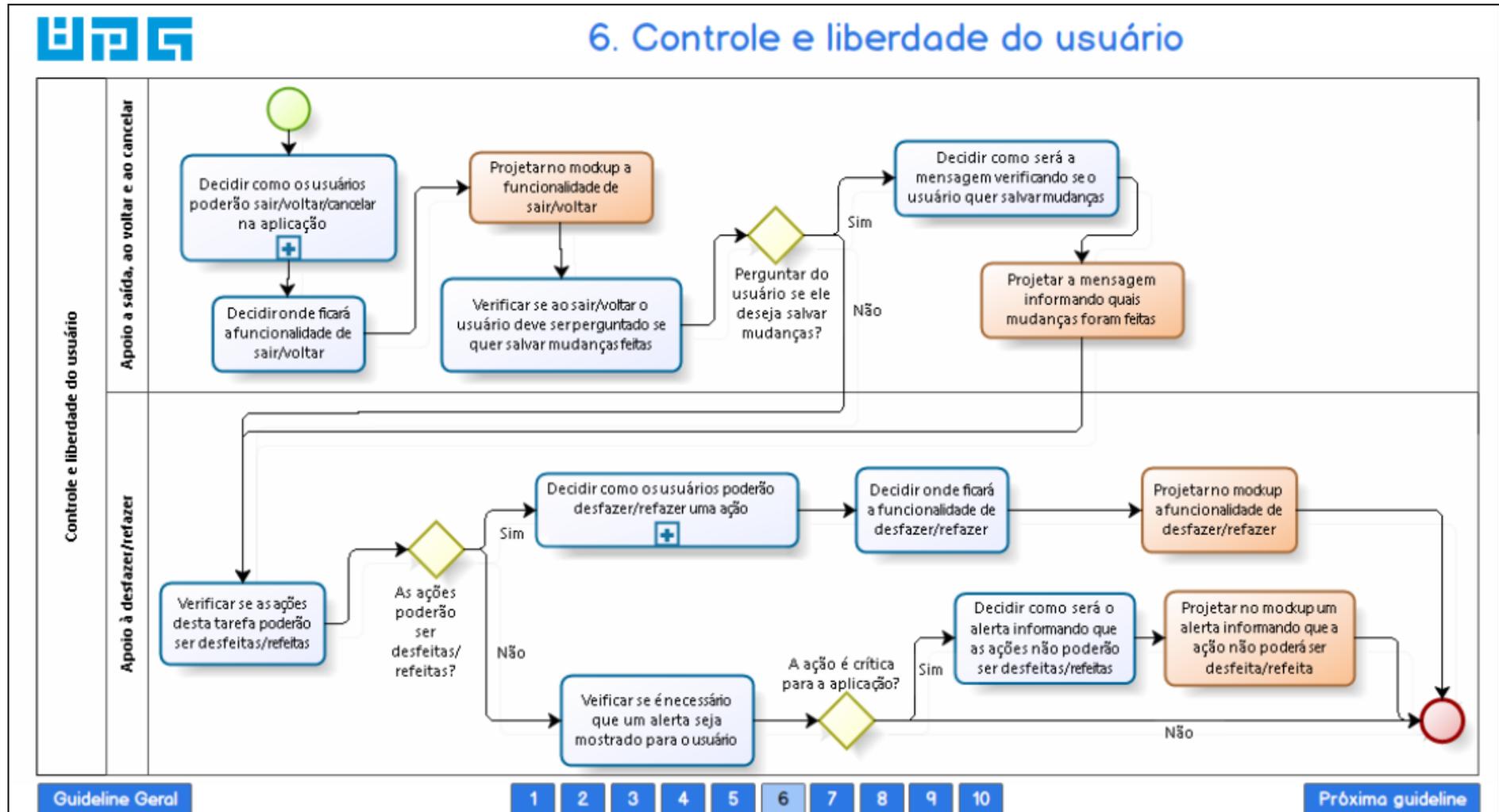
[Voltar para a guideline](#)

### **Dicas - Feedback do status da aplicação:**

- Deve haver um **elemento que informe o contexto** em que o usuário está localizado dentro da aplicação.
- Deve haver um **elemento que informe onde o usuário tem estado** durante a sessão. Dicas: na aplicação web pode-se usar a "migalha de pão" (breadcrumbs) ou links destacados nas subseções. Exemplo de migalha-de-pão: Página Inicial > Produtos > Detalhe do Produto. Na aplicação móvel pode-se destacar a aba que o usuário se encontra.
- Deve haver um **elemento que informe sobre o que está acontecendo** na aplicação
- Usar **elemento** que informe quando for a **primeira vez que o usuário usa a aplicação** (opcional).
- Usar **elemento** que informe **onde o usuário esteve quando parou na última sessão** (opcional).
- Se uma **área da aplicação for reformulada** ou até mesmo **uma nova aplicação for gerada**, é importante que **as pessoas reconhecem imediatamente que algo grande mudou**.

[Voltar para a guideline](#)

## Tela 19



## Tela 20

[Voltar para a guideline](#)

### Dicas - Botão Fechar/Sair:

- Ofereça uma botão de fechar/sair visível na aplicação.
- Botão Fechar/Sair com [X] no canto (superior direito/esquerdo, inferior direito/esquerdo) 
- Apenas um [X] no canto (superior direito/esquerdo, inferior direito/esquerdo) 
- Apenas um botão Fechar/ Sair 
- Clique em qualquer lugar da aplicação para fechar/sair

### Dicas - Botão Voltar:

- Ofereça uma botão de voltar visível na aplicação 
- Criar uma **maneira fácil e óbvia** para o usuário **voltar rapidamente à tela/página inicial** de qualquer ponto da aplicação.

### Dicas - Botão Cancelar:

- É necessário oferecer **liberdade ao usuário para cancelar** determinada ação.
- Exemplo de botão Cancelar 

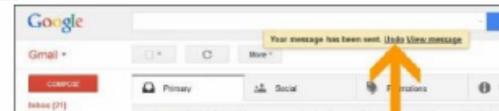
[Voltar para a guideline](#)

## Tela 21

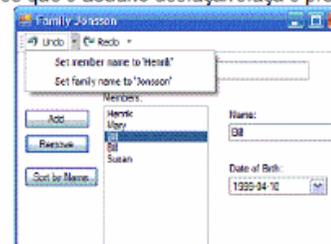
Voltar para o guideline

### Dicas - Botão Desfazer/Refazer:

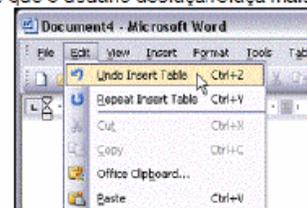
- As **unidades de reversibilidade** podem ser:
  - **uma única ação**, como a de enviar email (Gmail fornece a opção de undo depois de enviar um email). Esta opção é usada quando deseja-se que o usuário desfça/refaça uma única ação.



- **uma tarefa de entrada de dados**, como o preenchimento de um nome. Esta opção é usada quando permite-se que o usuário desfça/refaça o preenchimento de um campo.



- **um grupo completo de ações**, como a inserção de vários itens no Word. Esta opção é usada quando deseja-se que o usuário desfça/refaça mais de uma ação.

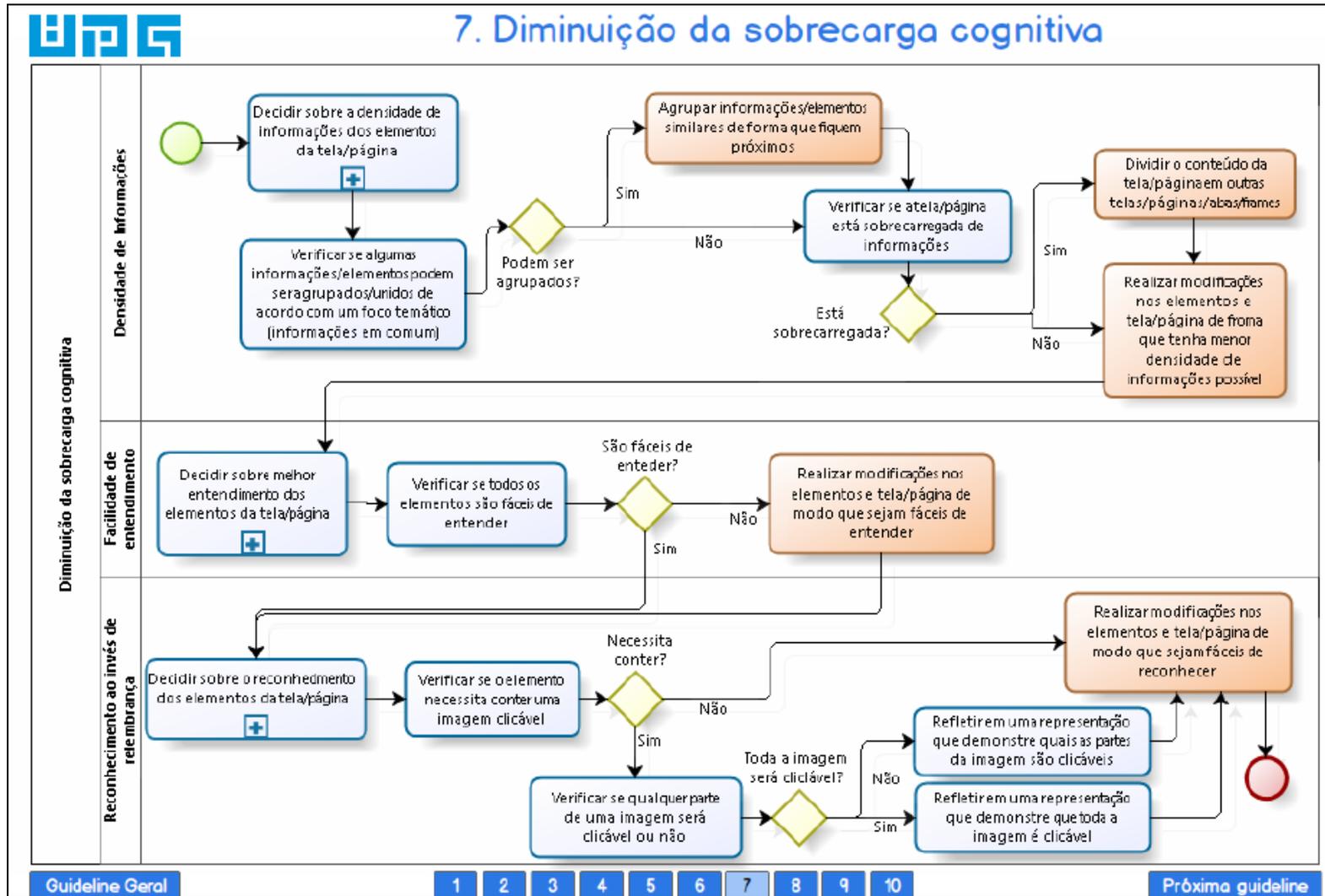


Referência: <http://www.cs.umd.edu/~ben/goldenrules.html>

Voltar para o guideline

## Tela 22

### 7. Diminuição da sobrecarga cognitiva



## Tela 23

[Voltar para a guideline](#)

### Dicas - Densidade de informação:

- **Densidade** é a porcentagem da tela/página preenchida com texto e gráficos. Em outras palavras, a densidade determina a quantidade de algo existente em um espaço delimitado na tela/página.
- Use frames/abas para **agrupar informações de acordo um foco temático**.
- Use **poucos elementos por grau de ângulo visual** dentro de um grupo distinto visualmente na tela/página.
- **Limitar a quantidade de espaço em branco** (áreas sem texto, gráficos, etc.) nas telas/páginas que são utilizados para a digitalização e pesquisa.
- Se a **velocidade de leitura** é mais importante, usar comprimentos de linha mais longos (75-100 caracteres por linha). Se a **aceitação da aplicação** é o mais importante, usar comprimentos de linha mais curtos (cerca de 50 caracteres por linha).

[Voltar para a guideline](#)

## Tela 24

[Voltar para a guideline](#)

### **Dicas - Facilidade de entendimento:**

- Usar nos elementos uma terminologia/nome e/ou representação/símbolo e cores eficazes que seja **fácil de entender para todo tipo de usuário**.
- Usar neste elemento uma representação/símbolo, terminologia/nome que **indique claramente a sua função** ou destino.
- Usar **rótulos significativos, bordas e espaço em branco** permite que os usuários identifiquem e entendam um conjunto de elementos mais rápido.

[Voltar para a guideline](#)

## Tela 25

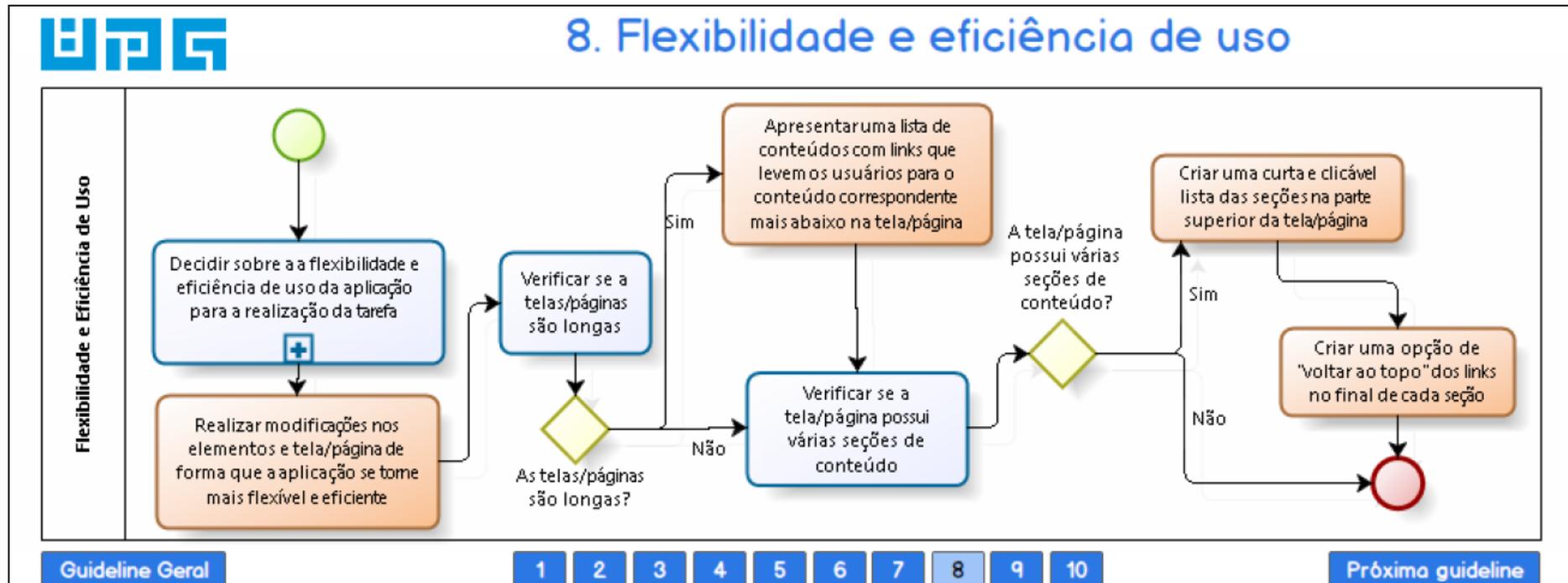
[Voltar para a guideline](#)

### **Dicas - Reconhecimento ao invés de relembração:**

- Usar no elemento o **mesmo tamanho, cor e tipo de fonte** que foi empregado no grupo de elementos a qual este elemento pertence.
- Usar uma terminologia/nome e/ou representação/símbolo no elemento que **reduza a carga cognitiva e o esforço físico de lembrar**.
- **Evitar que os usuários tenham que se lembrar de informações** a partir de uma tela/página e, em seguida, usar a informação em outra tela/página.
- Um **link** deve ser reconhecido como um **elemento clicável**.
- **Controles** e outros elementos necessários para o sucesso do uso de software **devem ser visivelmente acessíveis** em todos os momentos.

[Voltar para a guideline](#)

## Tela 26



## Tela 27

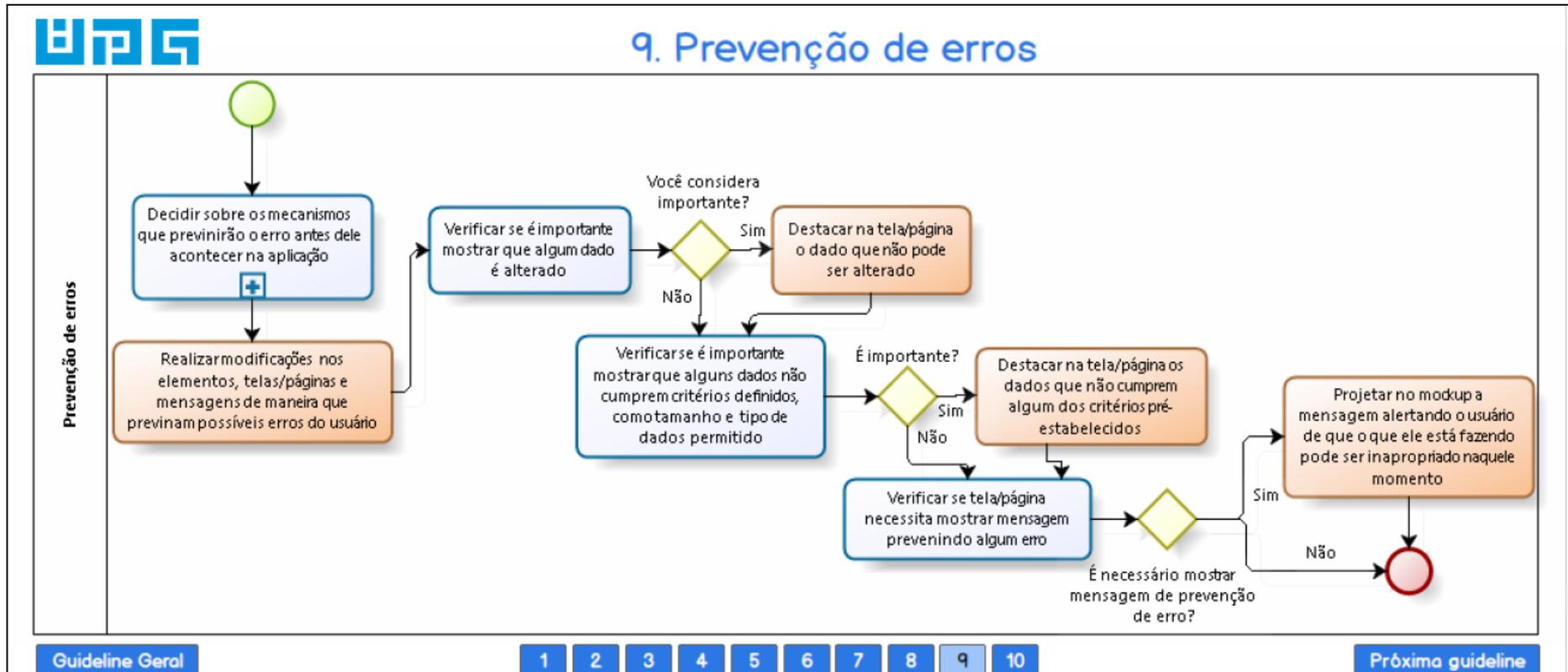
[Voltar para a guideline](#)

### **Dicas - Flexibilidade e eficiência de uso:**

- Reconhecer as necessidades de **diversos usuários**, como diferenças entre **novatos e especialistas** e diferentes faixas etárias, facilitando a transformação de conteúdo.
  - Os **nomes/nomenclatura** de elementos podem ser **abreviados** para usuários mais experientes.
  - Os **usuários experientes** devem poder "**cortar caminho**" em ações frequentemente repetidas.
  - Os **usuários experientes** desejam estar no **comando da aplicação**. Eles não gostam de surpresas ou mudanças no comportamento familiar. Eles ficam incomodados com sequências tediosas de entrada de dados, dificuldade em obter as informações necessárias, e incapacidade de chegar no seu resultado desejado.
  - É normal fazer o **ambiente** do usuário **mais simples** quando eles estão **aprendendo**, escondendo caminhos e recursos mais avançados, revelando-os quando os usuários vir a precisar deles e saber como lidar com eles.
- A **disposição dos elementos** deve **aumentar a eficiência de uso**, minimizando o esforço físico.
  - Usar um **layout de tela/página** apropriada para **eliminar a necessidade dos usuários** terem que **rolar a tela/página horizontalmente**.
  - **Não exigir** que os usuários **rolem as telas/páginas** que são apenas de navegação. As telas/páginas de navegação não devem conter mais do que uma tela/página de informações.
  - **Usar mais a rolagem de telas/páginas** quando os usuários terão que **ler um texto grande**, que necessita de bastante esforço do usuário para compreensão.
  - Deve ser possível **passar de uma tela/página para outra**, selecionando links, **sem ter que rolar a tela/página para baixo ou para cima** para encontrar informações importantes.
  - Usar **menus "sequenciais"** para **tarefas simples** e use **menus "simultâneos"** para **tarefas que exigirão vários usos do botão Voltar**. No "sequencial", após cada seleção um menu se abre.
  - Usar **menus estáticos** para obter a **velocidade mais rápida** possível ao acessar itens de menu.
  - Um **conteúdo importante** deve poder ser acessado a partir de **mais de um elemento de controle ou link**. Isso pode ajudar alguns usuários a encontrar o que precisam.
  - Organize **listas e tarefas** em uma **ordem que melhor facilite o desempenho** do usuário de forma eficiente e bem sucedida.
- Use **mudanças de cor** para **indicar** aos usuários **quando um link foi visitado**. Isso ajuda os usuários a entender quais partes de um site eles visitaram, aumentando a capacidade de encontrar informações.

[Voltar para a guideline](#)

## Tela 28



## Tela 29

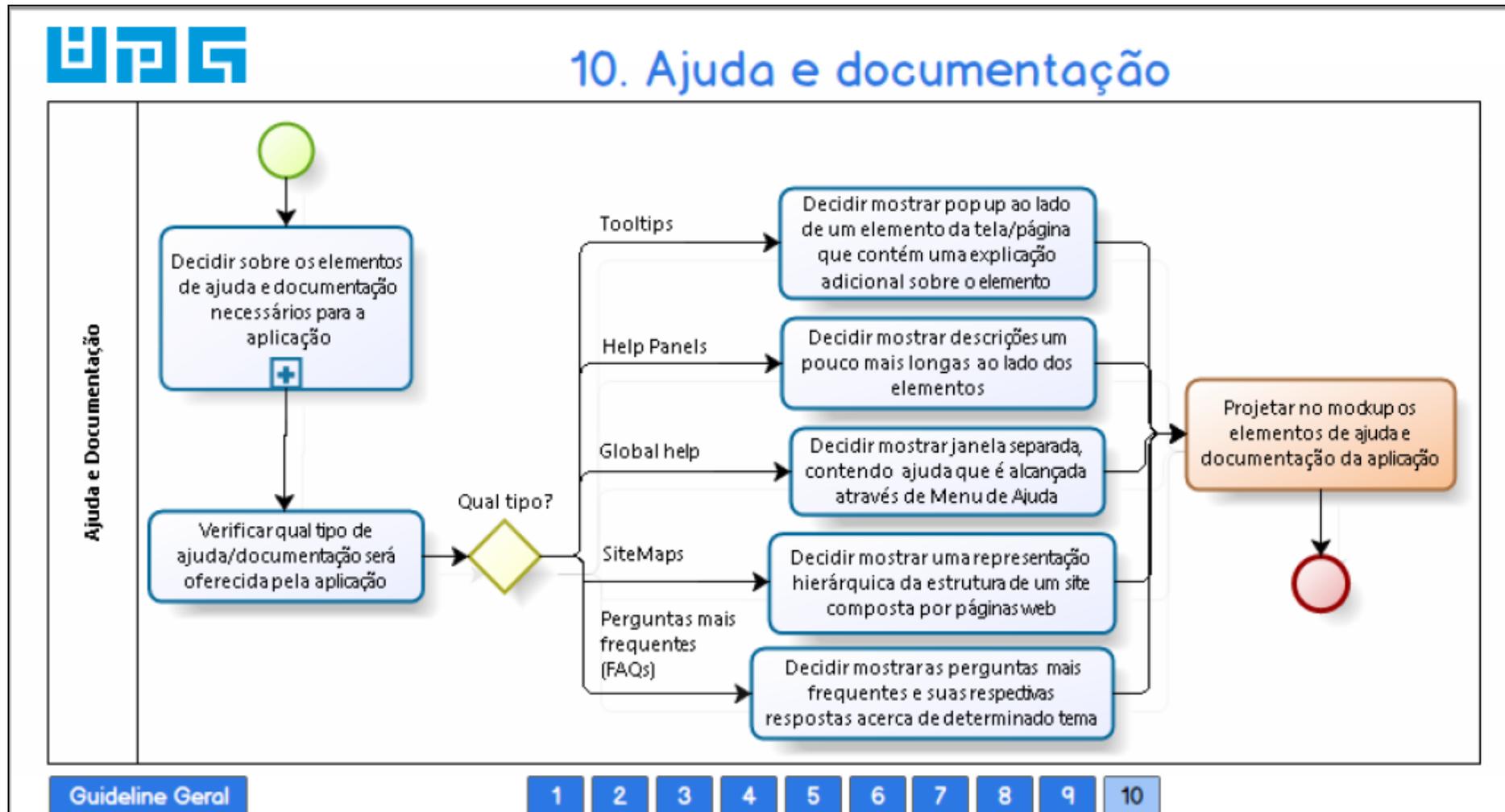
[Voltar para a guideline](#)

### Dicas - Prevenção de erros:

- Usar **elementos** que possuam **nome/nomenclatura** que definem claramente **que resultados, estados, ações ou conteúdos serão atingidos**.
- Usar **título descritivo, único e conciso** para cada tela/página da aplicação.
- Os **dados obrigatórios na entrada de dados** devem estar **claramente definidos**.
- Deve haver alguma **indicação do formato correto** para uma **entrada de dados** específica.  
Dica: colocar máscara de datas e CPF, por exemplo.
- **Distinguir visualmente elementos da tela/página importante**, isto é, destaque itens que exigem a atenção do usuário, especialmente quando esses itens são exibidos com pouca frequência.

[Voltar para a guideline](#)

## Tela 30



## Tela 31

[Voltar para a guideline](#)

### Dicas - Ajuda e Documentação:

- Usar **elementos** que listem **passos concretos** do que deve ser seguido pelo usuário.
- Usar **elementos** que descrevam as **possíveis ações para todo tipo de usuário** (inexperiente ou experiente).
- **Instruções para uso** do sistema devem estar **visíveis e facilmente recuperáveis** quando necessário.

[Voltar para a guideline](#)