

Universidade Federal de Pernambuco
Departamento de Design
Programa de Pós-Graduação em Design



Dino Lincoln Figueirôa

A AVALIAÇÃO DE ARTEFATOS EM DESIGN E OS PROBLEMAS DECORRENTES DA ALEATORIEDADE

Recife
2012



Dino Lincoln Figueirôa

A AVALIAÇÃO DE ARTEFATOS EM DESIGN E OS PROBLEMAS DECORRENTES DA ALEATORIEDADE

Tese apresentada ao Programa
de Pós-Graduação em Design da
Universidade Federal de
Pernambuco como requisito
parcial para obtenção do título de
Doutor

Orientador: Prof. Dr. Fábio Campos

Recife

2012

Catálogo na fonte
Andréa Marinho, CRB4-1667

- F475a Figueirôa, Dino Lincoln.
A avaliação de artefatos em design e os problemas decorrentes da aleatoriedade/ Dino Lincoln Figueirôa. – Recife: O Autor, 2012.
130 f.: il.: 30 cm.
- Orientador: Fábio Campos.
Tese (Doutorado) – Universidade Federal de Pernambuco, CAC. Design, 2012.
Inclui bibliografia, anexo e apêndice.
1. Design. 2. Projeto de Produto. 3. Espaço (Arte). I. Campos, Fábio (Orientador). II. Título.
- 745.2 CDD (22.ed.) UFPE (CAC 2012-112)



UNIVERSIDADE FEDERAL DE PERNAMBUCO

PROGRAMA DE PÓS-GRADUAÇÃO EM DESIGN

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE TESE DE DOUTORADO ACADÊMICO DE

DINO LINCOLN FIGUEIRÔA SANTOS

“Avaliação de Artefatos em Design e os problemas decorrentes da aleatoriedade”.

ÁREA DE CONCENTRAÇÃO: DESIGN E ERGONOMIA

A comissão examinadora, composta pelos professores abaixo, sob a presidência do primeiro, considera o(a) candidato(a) **DINO LINCOLN FIGUEIRÔA SANTOS** **APROVADO.**

Recife, 05 de outubro de 2012.

Prof. Dr. Leonardo Augusto Gomez Castillo (UFPE)

Prof. Dr. Walter Franklin Marques Correia (UFPE)

Prof. Dr. Wilson Kindlein Júnior (UFRGS)

Prof. Dr. Ricardo Triska (UFSC)

Prof. Dr. Lucian Bogdan Bejan (UFRPE)

Dedicado ao Designer que
projetou e criou o universo
e tudo o que nele há

Agradecimentos

Agradeço ao Senhor, meu **Deus**, pela oportunidade de cursar um doutorado, me sustentando em pé diante da sucessão de dificuldades para concluir este projeto e me dando a honra de ser o primeiro doutor em design das universidades federais.

A minha esposa, **Taciana Figueirôa**, que foi a minha *coacher* nesse processo mas, mais que isso, a minha real motivação para buscar crescer na vida.

Aos meus pais (**Eugênio Lincoln** e **Mirtes Figueirôa**) e demais familiares pela presença, incentivo e todo o apoio.

Ao meu orientador, **Prof. Dr. Fábio Campos**, que por tanto me orientar não só academicamente, mas na vida, acabou se tornando meu padrinho de casamento.

Aos professores da banca examinadora que esforçaram-se para analisar meu trabalho, em alguns casos refazendo os experimentos em seus laboratórios e, demonstrando assim, tamanho respeito com a pesquisa. Obrigado por suas relevantes contribuições, espero ter atendido a todas as correções.

Ao **Prof. André Melo**, **Tenente Bruno de Castro** e ao gênio **Williams Artimã** meus sinceros agradecimentos por realizarem comigo todos os cálculos de espaço amostral, desvios-padrão, graus de confiança, falácia da conjunção, testes de usabilidade e etc.

Ao querido **Pastor Doutor Ney Silva Ladeia**, bem como aos meus demais irmãos da Igreja Batista da Capunga (**Lucas Benvindo**, **Gabriel**, **Grimaldo**, **Edna**, **Roberto**, **Tony**, **Bete** e demais) pelo apoio, presença e principalmente pelas orações sem as quais eu não conseguiria êxito.

Ao **Prof. Dr. Janguê Diniz** por me dispensar do trabalho e, por vezes, mudar a própria agenda para que eu atendesse às exigências do programa de doutorado. Demonstrando com estas atitudes ser um empreendedor verdadeiramente comprometido com a educação.

Aos meus demais **amigos**, **alunos** e **colegas de turma** que contribuíram para conclusão deste árduo trabalho, o meu humilde agradecimento.

RESUMO

A metodologia de design tem auxiliado projetistas a conceber sistematicamente produtos inteligentes, agradáveis e úteis. A presente tese destaca que, neste processo de concepção de artefatos, a fase específica de avaliação configura-se como evento substancialmente impactado pela aleatoriedade. Uma vez que o designer depara-se com múltiplas probabilidades de resultado diante da avaliação de um produto. Assim, esta pesquisa identificou cinco problemas que afetam diretamente a avaliação de artefatos em design, propondo respectivas soluções para os mesmos. Estes problemas foram classificados como: regressão à média; falácia da conjunção; incerteza; espaço amostral e; escalas indutivas. Pelo uso de revisão literária, identificaram-se soluções para a regressão à média; incerteza e para o uso de escalas. Através de estudos de caso, foram realizadas 225 avaliações de artefatos, onde desenvolveram-se soluções para falácia da conjunção e representação do espaço amostral no âmbito do design. Com uso destas propostas de solução estima-se reduzir o impacto do acaso na avaliação de artefatos, aumentando as chances de sucesso do projeto.

Palavras-chave: *avaliação de artefatos, falácia da conjunção, regressão à média, incerteza, espaço amostral, escalas psicométricas.*

ABSTRACT

The design methodology is supporting designers to systematically design intelligent, pleasant and helpful products. This thesis highlights that into this process of designing artifacts, the particular phase of evaluation is substantially impacted by aleatority, once the designer is faced with multiple probabilities during a product evaluation. In the meantime, this study identified five issues that directly affect the evaluation of design artifacts, proposing solutions to them. These problems were classified as regression to the mean; conjunction fallacy; uncertainty; sample space and the; inductive scales. Through literature review, we identified solutions to regression to the mean, uncertainty and the use of scales. Through case studies there were 225 artifacts evaluations, which developed solutions for the conjunction fallacy and representation of the sample space within the design. The use of these proposed solutions reduces the impact of aleatority in the artifacts evaluation, increasing the chances of project success.

Keywords: *artifact evaluation, conjunction fallacy, regression towards the mean, uncertainty, sample space, psychometric scales.*

SUMÁRIO

LISTA DE FIGURAS	11
LISTA DE QUADROS.....	12
1 INTRODUÇÃO	14
1.1 Contextualização da Avaliação de Artefatos na Metodologia de Design	16
1.3 Hipótese e Objetivos	18
1.2.1 Hipótese.....	18
1.2.2 Objetivo Geral.....	18
1.2.3 Objetivos Específicos	18
1.4 Metodologia de Pesquisa	19
1.4.1 Realização de Pesquisa Bibliográfica – Fundamentação Teórica e Estado da Arte.....	19
1.4.2 Realização de Pesquisa Empírica – Estudos de Caso	19
1.4.3 Apresentação das Soluções e Discussões	19
1.5 Justificativas	19
1.5.1 Pervasividade da Avaliação de Artefatos em Múltiplos Paradigmas de Design	19
1.5.2 Aspectos Mercadológicos.....	20
1.6 Normatização da Tese	20
1.7 Organização da Tese	21
2 AVALIAÇÃO DE ARTEFATOS EM DESIGN	22
2.1 Avaliação de Artefatos Quanto à Criatividade	23
2.1.1 Escala Semântica de Produto Criativo	24
2.1.2 Modelo de Análise de Criatividade de Produto	25
2.1.3 Análise de Ideias Criativas por De-Bono.....	26
2.1.4 Avaliação de Artefatos Quanto à Criatividade e a Aleatoriedade.....	27
2.2 Avaliação de Artefatos Quanto à Estética	27
2.2.1 Relação entre Estética, Usabilidade e Emoções.....	28
2.2.2 Avaliação Estética por Aristóteles	29
2.2.3 Avaliação Estética de Websites.....	31
2.2.4 Avaliação de Artefatos Quanto à Estética e a Aleatoriedade	33
2.3 Avaliação de Artefatos Quanto à Experiência com o Usuário	33
2.3.1 Avaliação de Artefatos Quanto à Experiência com o Usuário e a Aleatoriedade.....	36
2.4 Avaliação de Artefatos Quanto à Usabilidade	37
2.4.1 Percurso Cognitivo (“Cognitive Walktought”)	38
2.4.2 Co-Descoberta (“Co-Discovery”)	38

2.4.3 Observações de Campo ("Field Observations")	39
2.4.4 Avaliação Cooperativa ("Cooperative Evaluation").....	39
2.4.5 Diário de Incidentes ("Incident Diaries")	40
2.4.6 As Heurísticas de Jakob Nielsen	41
2.4.7 Outras Avaliações de Usabilidade por Jordan	43
2.4.8 Avaliação de Artefatos Quanto à Usabilidade e a Aleatoriedade	45
2.5 Avaliação de Artefatos Quanto ao Mercado	45
2.5.1 Produtos Complementares e Substitutos	47
2.5.2 Curva de Demanda e Curva de Oferta	48
2.5.3 Benefícios Básicos	51
2.5.3 Ciclo de Vida do Produto	52
2.5.4 Avaliação de Artefatos Quanto ao Mercado e a Aleatoriedade.....	55
3 PROBLEMAS DECORRENTES DA ALEATORIEDADE	56
3.1 Introdução à Aleatoriedade na Avaliação de Artefatos	58
3.2 O Fenômeno da Regressão à Média	61
3.3 A Falácia da Conjunção	69
3.4 Incerteza nas Decisões de Projeto (<i>Lateo</i>)	73
3.5 Representação do Espaço Amostral	77
3.5.1 Introdução ao Problema da Quantidade de Avaliadores em Usabilidade	77
3.5.2 Determinação do Tamanho de uma Amostra.....	78
3.5.4 Controvérsias na Quantidade de Avaliações em Usabilidade.....	82
3.6 Escalas Indutivas	82
3.6.1 Escala Likert	83
3.6.2 Escala de Diferencial Semântico	84
3.6.3 Escala de Stapel.....	85
3.6.4 Escala Pictórica	86
3.6.5 Escala de Guttman	87
3.6.6 Escala de Alpert.....	87
3.6.7 Problema: Divergências Parciais nos Resultados das Escalas	88
4 ESTUDOS DE CASOS.....	90
4.1 Falácia da Conjunção	91
4.1.1 Procedimentos Metodológicos	91
4.1.2 Expondo a Falácia da Conjunção no Design	93
4.1.3 Resultados da Correção da Falácia da Conjunção no Design	97
4.1.4 Considerações Referentes ao Estudo de Caso Sobre a Falácia da Conjunção	101
4.2 Representação do Espaço Amostral	103
4.2.1 Procedimentos Metodológicos	104

4.2.2 Resultados do Estudo de Caso Sobre o Espaço Amostral	107
4.2.3 Considerações Referentes ao Estudo de Caso Sobre o Espaço Amostral.....	111
5 CONCLUSÕES	113
5.1 Conclusões Específicas Sobre os Resultados dos Estudos de Casos	113
5.2 Resumo das Propostas de Solução	114
5.2.1 Regressão à Média.....	115
5.2.2 Falácia da Conjunção	116
5.2.3 Incerteza	116
5.2.4 Espaço Amostral.....	116
5.2.5 Escalas Indutivas.....	117
5.3 Desdobramentos	117
REFERÊNCIAS.....	119
ANEXO A – Pesquisa da Falácia da Conjunção – iCopa	128
APÊNDICE A – Ficha de Avaliação Heurística de Usabilidade	129

LISTA DE FIGURAS

Figura 1 – Eventos aleatórios e determinísticos.....	15
Figura 2 – Redução do impacto da aleatoriedade no design.....	16
Figura 3 – Avaliação de artefatos na metodologia de design.....	18
Figura 4 – Diferentes paradigmas de projeto avaliam protótipos	20
Figura 5 – Telas do game “Corrida Contra o Tempo” (BELFORT, 2011)	35
Figura 6 – Tela de coleta de dados para UX (BELFORT, 2011).....	36
Figura 7 – Concorde – Produto inovador (Foto: Ciorlaus Vicentiu – airliners.net).....	45
Figura 8 – Benefício básico: durabilidade da pilha (Fonte: duracell.com.br).....	52
Figura 9 – Battlefield 2: Special Forces – Manutenção (FIGUEIRÔA, 2009).....	55
Figura 10 – Jatos de treinamento da IAF – Israeli Air Force (Fonte: www.iaf.org.il) ..	62
Figura 11 – Ipad da Apple (Fonte: apple.com.br).....	65
Figura 12 – Message Pad da Apple – “Newton” (Foto: GettyImages).....	66
Figura 13 – Função massa que determina o Lateo (CAMPOS, 2005)	76
Figura 14 – Número de participantes representando o espaço amostral.....	78
Figura 15 – Notações para fórmula do número de amostras	79
Figura 16 – Fórmula do desvio padrão para cinco escores.....	80
Figura 17 – Notações usadas para fórmula de desvio padrão.....	80
Figura 18 – Média aritmética da amostra “A”	80
Figura 19 – Desvio-padrão da amostra “A”	81
Figura 20 – Obtenção do grau de confiança da amostra “A”	81
Figura 21 – Grau de confiança da amostra “B”	82
Figura 22 – Tela do software de pesquisa utilizado.....	94
Figura 23 – Interface do website avaliado	108
Figura 24 – Cálculo estatístico para obtenção do número de avaliadores	110
Figura 25 – Cálculo do nível de confiança da amostra para 33 avaliadores	110

LISTA DE QUADROS

Quadro 1 – Avaliação de Artefatos em Design.....	22
Quadro 2 – Escala psicométrica para avaliação da criatividade de produtos	25
Quadro 3 – Avaliação de produtos via CPAM. Adaptado de CORREIA(2010).....	26
Quadro 4 – Escala estética de Aristóteles.....	29
Quadro 5 – Avaliação sob Escala Aristoteleliana. Adaptado de ROCHA (2007).....	31
Quadro 6 – Atributos da Escala de Diferencial Semântico por PIRAUÁ (2007)	32
Quadro 7 – Entrevista sobre Estética de PIRAUÁ(2007).....	32
Quadro 8 – Análise de Fluxo de Potência para UX. Adaptado de BELFORT(2011) .	34
Quadro 9 – Quadro de avaliação de UX. Adaptado de BELFORT(2011).....	36
Quadro 10 – Quantidade de avaliadores (NIELSEN & LANDAUER, 1993)	42
Quadro 11 – Custo-benefício por avaliadores (NIELSEN & LANDAUER, 1993).....	43
Quadro 12 – Avaliações de usabilidade (JORDAN, 1998)	44
Quadro 13 – Custos de desenvolvimento (BAXTER, 2011).....	47
Quadro 14 – Curva de demanda	48
Quadro 15 – Curva de oferta	49
Quadro 16 – Componentes do preço do produto (BAXTER, 2011)	51
Quadro 17 – Ciclo de vida do produto.....	53
Quadro 18 – Manutenção: estratégia que estende o ciclo de vida do produto	54
Quadro 19 – Problemas decorrentes da aleatoriedade na avaliação de artefatos.....	57
Quadro 20 – Exemplos de má interpretação da aleatoriedade.....	60
Quadro 21 – Mera ilustração do desempenho dos alunos em Regressão à Média...	63
Quadro 22 – Mera ilustração da intuição dos instrutores de voo israelenses	64
Quadro 23 – Aleatoriedade da bolsa (Adaptado de KANITZ, 2008).....	65
Quadro 24 – Produtos Apple (G1, 2011; IG SÃO PAULO, 2011; IDGNOW, 2011) ...	67
Quadro 25 – Ações da Apple mostram crescimento (YAHOO! FINANCES, 2012)....	67
Quadro 26 – Descrição de Bill. Adaptado de TVERSKY & KAHNEMAN(1983)	69
Quadro 27 – Descrição de Linda. Adaptado de TVERSKY & KAHNEMAN(1983).....	70
Quadro 28 – Classificação em ordem decrescente.	71
Quadro 29 – Crenças percentuais das avaliações dos especialistas	74
Quadro 30 – Combinação de opiniões pela regra de Dempster.....	75
Quadro 31 – Normalização pela regra de Dempster	75
Quadro 32 – Incerteza das alternativas mensurada pelo Lateo	76
Quadro 33 – Valores correspondentes de grau de confiança.....	79
Quadro 34 – Amostra "A" da simulação.....	80
Quadro 35 – Amostra "B" da simulação.....	81
Quadro 36 – Escala Likert.....	83
Quadro 37 – Escala de Diferencial Semântico. Adaptado de BRANSALISE (2005) .	84
Quadro 38 – Coleta por Diferencial Semântico. Adaptado de MATTAR (1996)	85
Quadro 39 – Análise por Diferencial Semântico. Adaptado de MATTAR (1996).....	85
Quadro 40 – Escala de Stapel. Adaptado de AGUIAR (2011).....	86
Quadro 41 – Escala pictórica. Adaptado de BELFORT (2011).....	86
Quadro 42 – Exemplo de Escala de Guttman. Adaptado de AGUIAR(2011).....	87
Quadro 43 – Escala Alpert. Adaptado de VELUDO-DE-OLIVEIRA(2001).....	88
Quadro 44 – Escopo para estudos de casos	90
Quadro 45 – DINOTUNES: Descrição e alternativas	93
Quadro 46 – Ordem de probabilidade atribuídas pelos entrevistados do Grupo 1....	95

Quadro 47 – Resultados "Grupo 1"	95
Quadro 48 – ICOPA: Descrição e alternativas	96
Quadro 49 – Ordem de probabilidade atribuídas pelos entrevistados do Grupo 2.....	96
Quadro 50 – Resultados "Grupo 2"	97
Quadro 51 – DINOTUNES com todas as alternativas conjuntivas.....	98
Quadro 52 – Ordem de probabilidade atribuídas pelos entrevistados do Grupo 3.....	99
Quadro 53 – Resultados "Grupo 3"	100
Quadro 54 – Erros de usabilidade reportados	109
Quadro 55 – Variação de porcentagem de erros de usabilidade encontrados	110
Quadro 56 – Projeção das soluções apresentadas na tese	115

1 INTRODUÇÃO

Durante o processo de desenvolvimento de produto, na fase de avaliação de artefatos, o designer tem o objetivo de averiguar a probabilidade de êxito deste produto no mercado. Cabe ao designer justificar à indústria qual dos projetos está mais suscetível ao sucesso. Nestas circunstâncias, a pressão sobre estes projetistas é substancial, uma vez que sua decisão influenciará a linha de montagem, o programa de marketing e subsequentes acontecimentos relativos ao produto em questão (BAXTER, 2011).

Neste processo de avaliação, o projetista dispõe de uma série de técnicas para seu auxílio. Explicita-se nesta tese que dentre estas técnicas, responsáveis por validar todo o processo de criação do design, encontram-se falácias pervasivas decorrentes da aleatoriedade que podem conduzir o designer a erros de projeto.

Assim, é possível observar neste contexto duas amplitudes: de um lado uma sucessão de técnicas de avaliação de artefatos para auxiliar o designer e do outro uma gama de problemas (falácias e fenômenos) que estão permeados pervasivamente nestas técnicas. Por razões de escopo, a presente pesquisa explicita o estado da arte das técnicas de avaliação de artefatos em design e limita-se a apresentar, diante de uma amplitude de problemas projetuais possíveis, aqueles que estão relacionados especificamente à aleatoriedade.

Por melhores que sejam a equipe de projeto, as metodologias empregadas, o equipamento assistivo utilizado, não é possível determinar o exato resultado do projeto, apenas reduzir a quantidade de resultados possíveis. Por vezes, o designer lida inclusive com múltiplas probabilidades de problemas que podem estar afetando o artefato. Ou seja, estamos diante de um evento aleatório (FIGUEIROA et al., 2012; 2008). Diferente de um problema determinístico, onde só existe um único resultado possível. Por exemplo, evento comum em certos problemas de física mecânica onde, uma vez apresentadas as variáveis, só existe um resultado provável no cálculo.

O escopo da presente tese sustenta-se na hipótese de que a avaliação de artefatos configura-se como um evento com substancial impacto do acaso. Portanto, considera-se que existem 2 paradigmas de eventos (acontecimentos): aleatórios e determinísticos, conforme ilustrado na Figura 1.

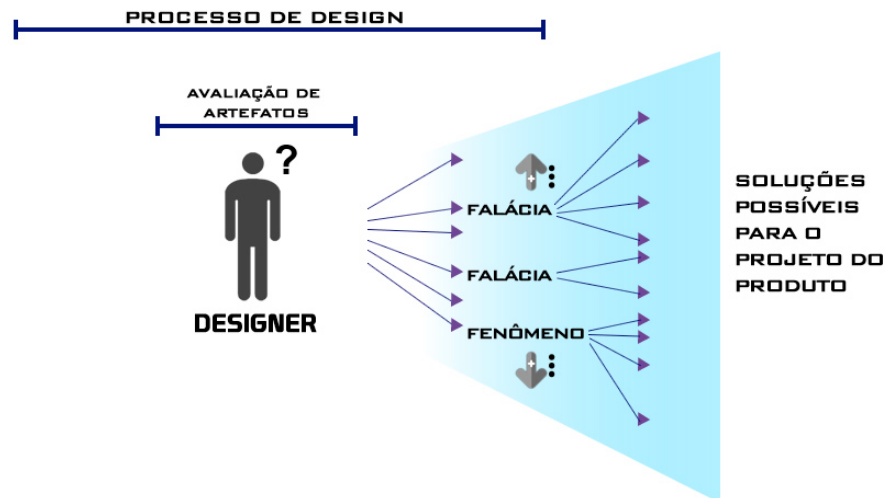


Figura 1 – Eventos aleatórios e determinísticos

A avaliação de artefatos está permeada de problemas decorrentes da aleatoriedade por configurar-se como um evento com múltiplas probabilidades de resultado. Ao contrário de um único resultado pré-determinado, como ocorreria num evento determinístico (Figura 1). Deste modo, teoriza-se que uma maior compreensão dos fenômenos da aleatoriedade que impactam a avaliação de artefatos (em design) deve aumentar relevantemente a precisão desta fase de processual e, por consequência, as chances de sucesso do projeto.

Nestes termos, o problema central abordado por esta tese consiste na redução do impacto do acaso no processo de design, através da maior compreensão sobre os fenômenos relacionados à aleatoriedade durante a fase de avaliação de artefatos. Conforme ilustrado na Figura 2, o designer durante a avaliação de artefatos pode ser negativamente impactado por desconhecer as falácias processuais (decorrentes do desconhecimento dos fenômenos da aleatoriedade) que podem estar ampliando as soluções prováveis, ou mesmo os problemas de projeto prováveis. Ainda descrevendo esta figura, convém esclarecer que a fase anterior a de avaliação de artefatos (a de geração de alternativas) tem por objetivo a expansão de idéias e possibilidades de soluções. Todavia, a fase de avaliação de artefatos tem cunho restritivo, no sentido de filtrar as idéias e soluções mais suscetíveis ao sucesso (destacando aquelas com melhor avaliação e problemas menos críticos) (BAXTER, 2011; LOBACH, 2001).

SEM CONHECIMENTO SOBRE A ALEATORIEDADE NO PROCESSO



COM CONHECIMENTO SOBRE A ALEATORIEDADE NO PROCESSO

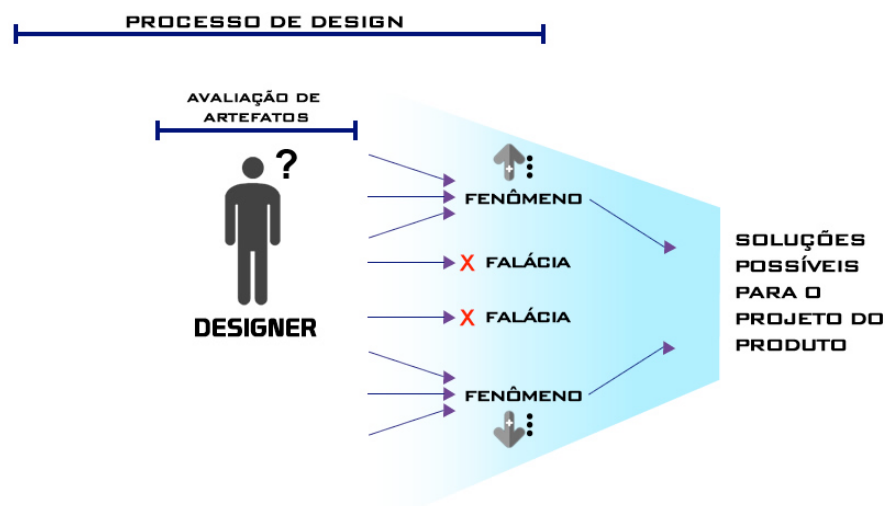


Figura 2 – Redução do impacto da aleatoriedade no design

A seção a seguir apresenta uma introdução à metodologia de design e como a fase de avaliação de artefatos contextualiza-se na mesma, destacando sua relevância.

1.1 Contextualização da Avaliação de Artefatos na Metodologia de Design

Atualmente o design conta com metodologias regidas essencialmente por filosofias cartesianas e aristotelianas. De René Descartes aproveita-se a abordagem dos problemas de projeto similares a problemas de geometria, onde problemas maiores são subdivididos em fases de menor complexidade. Este procedimento é conhecido como *reducionismo cartesiano* e apresenta-se congruente à filosofia platônica da dialética, que deriva de

diairesis e quer dizer “divisão de idéias”. Ou seja, divisão de uma ideia, ou problema, de complexidade relativamente alta em problemas menores, mais simples de abordar. De Aristóteles, o design herda a lógica, a qual encadeiam-se os subproblemas no projeto de artefatos-soluções. Além disso, ainda é possível encontrar nas modernas metodologias de projeto pensamentos norteados em Hegel e Kant (BURDEK, 2006).

Toda esta impregnação filosófica, bem como a científica, foi introduzida ao design inicialmente pela escola Bauhaus nos anos 1930, mas principalmente pela faculdade HfG (*Hochschule für Gestaltung*, ou “Escola Superior da Forma”) nos anos 1970. Estes centros educacionais – ambos na Alemanha – contribuíram sobremaneira para a cientifização do design, sendo aclamadas por muitos autores como as precursoras das atuais metodologias de projeto em uso neste seguimento.

De todos os campos, o da metodologia do design, sem a HfG de Ulm, não seria imaginável. O pensamento sistemático sobre a problematização, os métodos de análise e síntese, a justificativa e a escolha das alternativas de projeto – tudo isto junto, hoje em dia, se tornou o repertório de profissão de design
(BURDEK, 2006)

Numa abordagem sistemática dos problemas de projeto, atualmente diversas metodologias auxiliam o designer na criação de artefatos – físicos ou imateriais – que representem soluções para as atividades cotidianas da sociedade. É manifesto, portanto, que o aprimoramento da atividade do design está diretamente relacionado ao amadurecimento das metodologias de projeto.

Em termos gerais, considerando as metodologias de Lobach (2001) e Burdek (2006), dentre outras, observa-se que suas etapas e técnicas de projeto agrupam-se basicamente nas 4 fases abaixo destacadas na Figura 3 abaixo.



Figura 3 – Avaliação de artefatos na metodologia de design

Após uma maturação considerável das fases de Exploração do Problema e de Geração e Seleção de Alternativas, uma série de imperfeições demandam por pesquisas na fase de Avaliação de Artefatos, também conhecida como “Avaliação de Alternativas” (LOBACH, 2001). É nesta fase que a presente tese propõe reduzir o impacto dos fenômenos correlatos ou essencialmente regidos pela aleatoriedade. Ampliando as chances de sucesso do projetista. Para tal, delineou-se os objetivos de pesquisa expostos a seguir.

1.3 Hipótese e Objetivos

1.2.1 Hipótese

- A existência pervasiva de problemas decorrentes da aleatoriedade impactando negativamente na avaliação de artefatos, durante o processo de design.

1.2.2 Objetivo Geral

- Investigar os fenômenos da aleatoriedade que podem representar problemas no processo de design, especificamente na fase de avaliação de artefatos, contrapresentando propostas de solução.

1.2.3 Objetivos Específicos

Especificamente, a presente pesquisa concentra-se no cumprimento dos seguintes objetivos.

- Descrever do estado da arte referente a avaliação de artefatos no processo de design;
- Identificar, na avaliação de artefatos do processo de design, problemas específicos decorrentes dos fenômenos da aleatoriedade;

- Apresentar propostas de solução contextualizadas em design para os fenômenos decorrentes da aleatoriedade que podem representar problemas nesta atividade projetual.

1.4 Metodologia de Pesquisa

Para o cumprimento dos objetivos da pesquisa e verificação da hipótese abordou-se o desenvolvimento da tese com as etapas abaixo propostas.

1.4.1 Realização de Pesquisa Bibliográfica – Fundamentação Teórica e Estado da Arte

Levantamento de autores (teses, dissertações, artigos, revistas acadêmicas e não-acadêmicas) correlatos à temática de “avaliação de artefatos em design”; “fase de avaliação de alternativas” ou; “aleatoriedade”. Permitindo assim uma visualização do que é o estado da arte da temática abordada, bem como sua fundamentação teórica. Ou seja, estudou-se o estado da arte da avaliação de artefatos em design e dos fenômenos da aleatoriedade que possuam pontos de inflexão.

Como fundamentação teórica, levantou-se ainda um breve histórico do desenvolvimento acadêmico do tema da aleatoriedade e como o mesmo pode ser abordado no contexto da avaliação de artefatos.

1.4.2 Realização de Pesquisa Empírica – Estudos de Caso

Para os problemas decorrentes da aleatoriedade identificados na avaliação de artefatos, conforme hipótese levantada, desenvolveram-se propostas de solução com estudos de casos específicos. Conforme disponibilidade de capital humano, estrutura e tempo coerentes com o presente projeto de doutoramento.

1.4.3 Apresentação das Soluções e Discussões

Apresentação sintetizada e discutida das soluções obtidas para os problemas identificados e detalhados tanto por revisão bibliográfica quanto por estudos de caso.

1.5 Justificativas

Abaixo pontuam-se algumas justificativas para o presente projeto de doutoramento em design.

1.5.1 Pervasividade da Avaliação de Artefatos em Múltiplos Paradigmas de Design

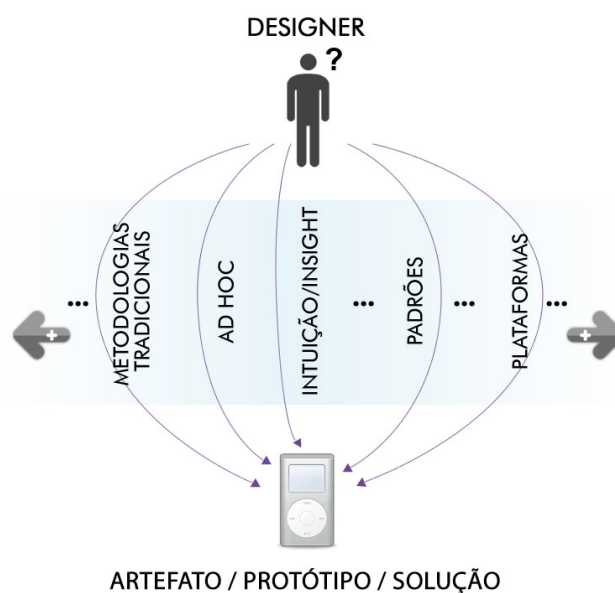


Figura 4 – Diferentes paradigmas de projeto avaliam protótipos

Qualquer que seja a metodologia de design empregada: metodologias tradicionais de segunda geração (LOBACH, 2001); *ad hoc* (CORREIA, 2010); intuição ou *insights*; padrões (FIGUEIRÔA, 2009); plataformas (FIGUEIRÔA, 2009); dentre outras, o designer terá de avaliar seus protótipos. Os problemas de design mapeados e desenvolvidos nesta tese, lidando com avaliação de artefatos, estão pervasivos portanto em qualquer processo de design realizado (Conforme ilustrado na Figura 4).

1.5.2 Aspectos Mercadológicos

Uma das exigências por parte das indústrias, para com os designers, é pela apresentação de dados que permitam verificar a possibilidade de sucesso do projeto (BAXTER, 2011). A execução das avaliações e suas respectivas correções, permitem ampliar significativamente as chances de sucesso do produto no mercado, reduzindo a incerteza envolvida no projeto e portanto os riscos de investimento no mesmo.

Uma avaliação de artefatos imprecisa, portanto, pode acarretar prejuízos financeiros relevantes à indústria, pois estariam lançando ao mercado produtos não plenamente verificados, deixando o resultado final das vendas demasiadamente sujeitos ao acaso.

1.6 Normatização da Tese

O presente documento está pautado nas normas abaixo explicitadas, regidas pela ABNT (Associação Brasileira de Normas Técnicas).

- Figuras e Quadros: NBR14724:2001.
- Referências: NBR6023:2000.
- Sumário: NBR6027:1989-NB85:1987.
- Numeração progressiva das seções de um documento: NBR6024:1989-NB69:1987.
- Resumos (errata, julho 1991): NBR6028:1990-NB88:1987.
- Citações em documento: NBR10520:2001;
- Referências: NBR6023:2000;
- Títulos de lombada: NBR12225:1992;

1.7 Organização da Tese

- **Capítulo I – Introdução (Capítulo atual)** – Apresentação da avaliação de artefatos contextualizada na metodologia de design. Apresentação da pesquisa abordando os problemas decorrentes da aleatoriedade e sua pervasividade do processo de design, especificamente na avaliação de artefatos.
- **Capítulo II – Avaliação de Artefatos em Design** – Estado da arte das áreas de concentração da avaliação de artefatos em design, dividindo-as em 5 grupos: criatividade, estética, experiência com o usuário (UX), usabilidade e mercado.
- **Capítulo III – Problemas Decorrentes da Aleatoriedade** – Evidenciação dos problemas mapeados na avaliação de artefatos em design que são decorrentes da aleatoriedade: regressão à média; falácia da conjunção; incerteza; representação do espaço amostral e; escalas indutivas, bem como as presentes propostas de solução encontradas por revisão literária.
- **Capítulo IV – Estudos de Caso** – Apresentação dos resultados dos experimentos para desenvolvimento de soluções para 2 dos 5 problemas decorrentes da aleatoriedade na avaliação de artefatos: falácia da conjunção e representação do espaço amostral.
- **Capítulo V – Conclusões** – Pareamento de hipóteses levantadas e conclusões obtidas no processo de condução da pesquisa. Resumo das soluções obtidas por revisão literária e estudos de casos.

2 AVALIAÇÃO DE ARTEFATOS EM DESIGN

Este capítulo apresenta especificamente o estado da arte referente às ferramentas processuais que concernem a fase de avaliação de artefatos no processo de design. Destacando em cada um dos tópicos apresentados os aspectos que são considerados importantes na avaliação de um produto. Posteriormente, ao confrontar este capítulo com o próximo (Capítulo 3 – Problemas Decorrentes da Aleatoriedade), é possível observar o quanto os aspectos referentes à aleatoriedade são relevantes para esta atividade e, ao mesmo tempo, o quanto os mesmos tem sido negligenciados pelas presentes metodologias de design.

As técnicas e conceitos envolvendo a avaliação de artefatos foram levantados por revisão bibliográfica e classificados nesta tese em 5 grupos, também chamados de “5 dimensões na avaliação de artefatos”, conforme explicitado no Quadro 1 abaixo (FIGUEIROA et al. 2012; BAXTER, 2011; BELFORT, 2011; BESEMER, 1998; NORMAN, 2008; NIELSEN & LANDAUER, 1993). Para tal procedimento, verificaram-se um total de 39 publicações, distribuídas entre 9 dissertações de mestrado, 1 (uma) monografia, 15 livros de caráter acadêmico e 14 artigos de periódicos e congressos científicos. O referido Quadro não encontra-se com gráfico distribuído em proporção de quantidade de publicações, mas por conveniência estética.



Quadro 1 – Avaliação de Artefatos em Design

Adiante, delinea-se detalhamentos sobre cada uma das áreas de avaliação de artefatos destacados no Quadro 1. Ao final da apresentação de cada uma destas áreas, apresenta-se uma breve consideração sobre sua relação com a tentativa de redução da aleatoriedade no processo de projeto, cujas implicações serão melhor detalhadas no capítulo seguinte (Capítulo 3 – Problemas Decorrentes da Aleatoriedade).

2.1 Avaliação de Artefatos Quanto à Criatividade

A criatividade é o coração do design.
(BAXTER, 2011, pg. 85)

Esta seção da tese apresenta as técnicas CPSS/CPAM de Besemer (1998) e a análise por De-Bono (1967) para avaliação de artefatos em design. Estes procedimentos coletam dados que representam opiniões sobre o quanto um produto é criativo, utilizando-se de escalas específicas de avaliação.

Destacando a importância deste seguimento para o design, alguns autores destacam esta atividade como essencialmente criativa, ou seja, onde produtos são concebidos por intermédio de um processo de criação. O critério de “criatividade” tem sido apontado como um fator crítico de sucesso neste ramo (CORREIA, 2010; BAXTER, 2011). Existem múltiplas definições para criatividade, mas em design este termo usualmente concerne soluções de que representem ideias fora do padrão, o que De-Bono (1967) chama de “Pensamento lateral”.

Este conceito está diretamente relacionado com a bissociação, que de forma sintetizada defende que o conflito de ideias fora do comum (por vezes ridículas ou absurdas) são importantes para a criatividade. Isso se dá pelo funcionamento automático de nossas redes neurais em fazer associações lógicas. Quando enxergamos um objeto, por exemplo, nosso cérebro processa o nome deste objeto (palavras), quando vimos este objeto pela última vez (redes de memória) e quais emoções esse objeto desperta (relacionando o objeto com registros de eventos importantes para nossa memória). Considerando que em vez do objeto temos um problema em mente, uma idéia inovadora precisa excender estes limites da rede lógica pré-estabelecida. Diversas técnicas atualmente ajudam a estimular a criatividade fazendo com que a colisão de idéias distintas ativen diferentes redes de memória, facilitando ter ideias de soluções inovadoras (BAXTER, 2011). A fase de geração de alternativas de produtos, no processo de design, está repleta de técnicas para auxiliar de forma racionalizada o processo criativo, como o brainstorming, método 635, biônica, dentre

outras (BOMFIM, 1984). Mas, uma vez prototipado o produto, como avaliar se o mesmo representa um produto criativo, ou não?

Neste contexto, algumas pesquisas propõem-se auxiliar o designer na verificação mensurada do quão criativo é um artefato, suportando o processo decisório de avaliação de sua probabilidade de sucesso como produto, conforme detalhado abaixo.

2.1.1 Escala Semântica de Produto Criativo

As pesquisas de Besemer & O' Quinn (1986) levaram a criação de um modelo mensurado de avaliação de artefatos sob escala psicométrica denominada CPSS (*Creative Product Semantic Scale* – Escala Semântica de Produto Criativo). Esta metodologia propõe avaliar o artefato quanto a sua criatividade através de pontuações atribuídas por especialistas aos critérios abaixo detalhados.

- **Surpresa** – O quanto o produto apresenta de informações inesperadas para o usuário;
- **Original** – O quanto o produto é incomum relativo a um universo de produtos feitos por pessoas com formação e/ou experiência similares;
- **Lógico** – O quanto o produto é coerente, apresentando regras de uso compreensíveis;
- **Útil** – O quanto o produto apresenta claras aplicações práticas;
- **Valioso** – O quanto o produto preenche uma demanda financeira, física, social ou psicológica;
- **Compreensível** – O quanto o produto comunica-se de forma auto-revelável, sendo cognitivamente amigável ao usuário;
- **Orgânico** – O quanto um produto apresenta-se harmônico entre suas partes;
- **Bem trabalhado** – O quanto um produto aparenta ter sido desenvolvido/retrabalhado para apresentar o nível mais alto de maturidade;
- **Elegante** – O quanto de suavidade e refinamento é expresso pelo produto.

Por intermédio do uso da *Escala de Diferencial Semântico*¹, as autoras acima adaptaram estes critérios criando a escala psicométrica exposta no Quadro 2.

¹ A Escala de Diferencial Semântico apresenta adjetivos contraditórios sobre algo que está sendo avaliado, para que o participante escolha de modo gradual o quanto um produto se aproxima destes adjetivos. Mais detalhes sobre esta escala no Capítulo 3, seção 3.

CREATIVE PRODUCT SEMANTIC SCALE (CPSS)
(BESEMER & O'QUINN, 1986)

SURPREENDENTE	-1---2---3---4---5---6---7-	COMUM
ORIGINAL	-1---2---3---4---5---6---7-	NÃO-ORIGINAL
LÓGICO	-1---2---3---4---5---6---7-	INCOERENTE
ÚTIL	-1---2---3---4---5---6---7-	INÚTIL
VALIOSO	-1---2---3---4---5---6---7-	SEM VALOR
COMPREENSÍVEL	-1---2---3---4---5---6---7-	INCOMPREENSÍVEL
ORGÂNICO	-1---2---3---4---5---6---7-	DESARMONIOSO
BEM TRABALHADO	-1---2---3---4---5---6---7-	POUCO TRABALHADO
ELEGANTE	-1---2---3---4---5---6---7-	DESELEGANTE

Quadro 2 – Escala psicométrica para avaliação da criatividade de produtos

Ao observar o protótipo e suas especificações, um conjunto de especialistas marca planilhas (Quadro 2). Fornecendo ao designer uma avaliação mensurada da criatividade o artefato proposto.

2.1.2 Modelo de Análise de Criatividade de Produto

Um desdobramento do CPSS resultou no CPAM (*Creative Product Avaliation Model – Modelo de Análise de Criatividade de Produto*). Agrupando os critérios do CPSS por categorias específicas, o CPAM permite visualizar a avaliação de criatividade do produto de acordo com os tópicos explicitados abaixo (BESEMER, 1998).

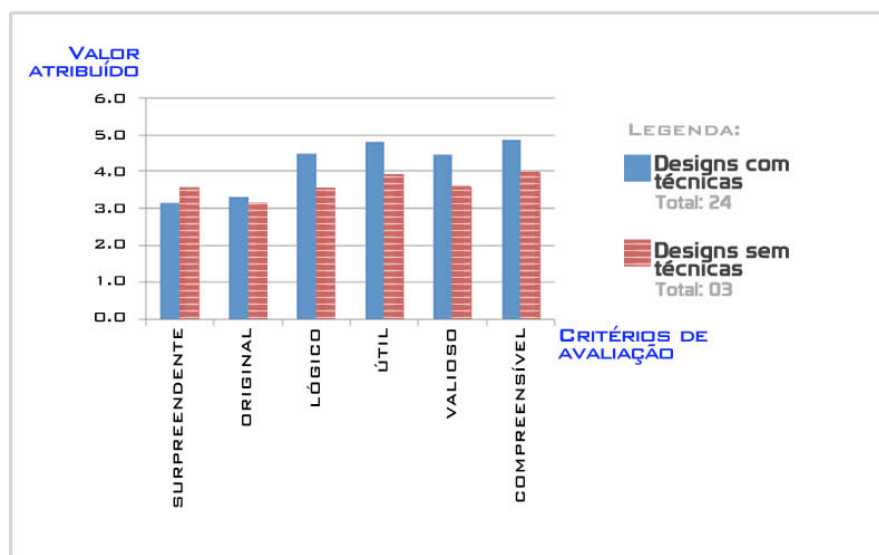
- **Novidade** (Surpresa; Original) – Grau de novidade do produto; Novos conceitos incluídos no produto;
- **Resolução** (Lógico; Útil; Valioso; Compreensível) – Grau em que o produto se enquadra ou atende as necessidades da situação-problema;
- **Estilo** (Orgânico; Bem trabalhado; Elegante) – Grau de quanto o produto apresenta combinação entre suas partes e refinamento.

Abordando problemas metodológicos no processo de design, Correia (2010) investigou as contribuições das técnicas de geração de alternativas na qualidade do produto final. Gerando, numa pesquisa sistematizada, um grupo de projetos de produto com o uso de técnicas e outro grupo sem o uso das mesmas. Considerando a criatividade com o critério a

ser avaliado nestes grupos de produtos, os mesmos foram submetidos à análise de especialistas em design que atribuíram aos projetos valores pela escala CPAM.

A pontuação atribuída para cada grupo de projetos fora a média aritmética dos critérios do CPAM disponibilizados aos especialistas, presente nas categorias novidade (surpresa e original) e resolução (lógico, útil, valioso e compreensível) conforme a Escala de Diferencial Semântico do Quadro 2, onde os valores variam de 1 a 7.

A pesquisa de Correia (2010) resultou num considerável apanhado de dados e projeções. Adaptado e disponibilizado no Quadro 3 explicita-se a análise de criatividade de projetos criados por designers com o uso de técnicas confrontada com o os projetos dos designers sem técnicas.



Quadro 3 – Avaliação de produtos via CPAM. Adaptado de CORREIA(2010)

Neste caso, através da técnica de avaliação de artefatos quanto à criatividade utilizada por Correia (2010) fora possível evidenciar que as técnicas de geração de alternativas (utilizadas na fase anterior à de avaliação) auxiliam designers a conceber produtos mais criativos e também em maior quantidade de alternativas.

2.1.3 Análise de Ideias Criativas por De-Bono

O autor Edward De-Bono (1967) desenvolveu diversas técnicas de geração de alternativas de projeto e, como convém nesta etapa de projeto, centenas de alternativas precisam ser posteriormente filtradas a poucas ideias que passarão por prototipação e serão encaminhadas à fase de avaliação de artefatos. Ou seja, esta técnica avalia idéias ainda não densamente prototipadas, apenas o conceito delas. Portanto, apesar de avaliar quanto

à criatividade, a escala de De-Bono não é utilizada na fase de avaliação de artefatos, e sim na fase anterior que gera e pré-seleciona ideias conforme as conjecturas:

1. Ideia diretamente utilizável;
2. Boa ideia, mas não para nós;
3. Boa ideia, mas não agora (reserva);
4. Precisa ser melhorada;
5. Forte, mas não utilizável;
6. Interessante, mas não utilizável;
7. Pouco valor;
8. Inaproveitável.

As ideias são agrupadas nestas 8 categorias, levando em consideração os benefícios; a viabilidade; os recursos disponíveis (tempo, dinheiro, etc.) para executá-la e o quanto ela se ajusta à empresa ou organização que a desenvolverá.

Apesar de não ser normalmente usada na fase de avaliação de artefatos, os autores Correia (2010) e Melo (2008) destacam a técnica de avaliação de ideias de De-Bono como uma ferramenta processual de avaliação de artefatos. Portanto, cabível de ser reportada nesta seção da presente tese que apresenta o estado da arte das publicações sobre técnicas no seguimento do design usadas para este fim.

2.1.4 Avaliação de Artefatos Quanto à Criatividade e a Aleatoriedade

As supracitadas pesquisas são as principais vias de análise de produtos quanto à criatividade, permeadas no design. Considerando a criatividade com critério importante para avaliação de artefatos, considera-se que os autores acima buscam reduzir o acaso através da redução da incerteza. Uma vez que buscam opiniões de especialistas, orientadas por escalas e parâmetros.

2.2 Avaliação de Artefatos Quanto à Estética

Quando você lava e aplica polimento no seu carro, ele não parece melhor de dirigir?
(NORMAN, 2008)

Nesta seção, apresentam-se as pesquisas que apontam os valores estéticos como críticos na avaliação de um artefato, relacionando os mesmos com fatores emocionais e de

usabilidade. Introduzindo inicialmente a avaliação estética e sua relação com outros critérios de avaliação. Em seguida, o uso da escala de aristóteles e avaliações estéticas de websites.

Para Lobach (2001) estes valores estéticos são dinâmicos: “O produto industrial hoje considerado belo pelo usuário (esteticamente valioso) talvez não seja amanhã, por terem sido estabelecidos novos valores estéticos...”. Esta razão amplia a necessidade da permanência da avaliação quanto à estética durante o projeto de produtos, uma vez que haverá a constante necessidade de balizá-los para o mutável gosto (e consequente senso estético) da sociedade. Neste sentido, apresentam-se abaixo algumas técnicas que objetivam mensurar este senso estético.

2.2.1 Relação entre Estética, Usabilidade e Emoções

A publicação de Norman (2008) representa um dos pilares da discussão sobre impacto dos aspectos estéticos no sucesso de produtos. Considerando que as emoções auxiliam as tomadas de decisão, os aspectos estéticos estariam diretamente ligados às emoções provocadas no usuário ao interagir com um produto e tomar uma opinião a respeito do mesmo. A própria etimologia da palavra estética, de acordo com Aranha e Martins (1990), provêm do grego *aisthesis* que, por sua vez, tem o significado de “faculdade de sentir” ou “compreensão pelos sentidos”.

Os cientistas japoneses Korusu & Kashimura (1995) pesquisaram a influência da estética sobre a funcionalidade de produtos durante a observação dos seus experimentos com 2 projetos de máquinas de autoatendimento bancário. Ambas as propostas apresentavam interfaces de máquinas com as mesmas funções, inclusive número de botões e forma de operar. Foram entrevistadas um total de 256 pessoas objetivando atribuir valores quanto à usabilidade e à estética das máquinas. As que detinham o design esteticamente mais atraente (organizado por princípios da *Gestalt*), demonstravam-se mais fáceis de usar. Em outras palavras, a beleza (critério estético) de um produto facilitaria seu uso (aumentaria o critério de usabilidade).

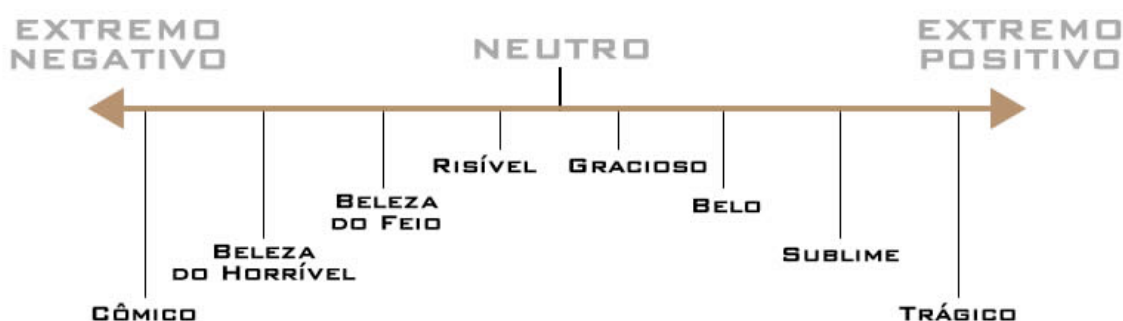
Atraído pelo resultado desta pesquisa, o cientista israelense Noam Tractinsky (1997), considerou que a cultura japonesa, orientada à beleza, teria influenciado os entrevistados. Decidindo re-aplicar a pesquisa em Israel, alegando que uma cultura mais rudimentar no que tange os aspectos estéticos poderia gerar resultados que se contrapusessem à pesquisa dos japoneses. Mesmo assim os resultados em Israel foram ainda mais contundentes, reafirmando que a beleza de um produto facilita seu uso e, portanto,

melhoraria seu funcionamento ante o usuário. Buscando maturar os dados, as pesquisas foram reaplicadas pelo autor gerando o artigo “*What is beautiful is usable*” (TRACTINSKY et al. 2000), onde novamente interfaces de máquinas bancárias consideradas mais belas aumentaram os índices de usabilidade.

De fato, emoções positivas melhoram os índices de aprendizado, curiosidade e criatividade. Pesquisas apontam que pessoas que receberam um pequeno presente, um pacote de doces ou assistiram um pequeno trecho de filmes de comédia obtiveram melhores resultados dos que as que foram privadas destes “estimulantes” antes de realizarem um brainstorming (ASHBY et al. 1999). Nestes termos, de acordo com os autores supracitados, diante de interfaces esteticamente mais agradáveis os usuários teriam uma predisposição à criatividade e maior paciência no aprendizado do manuseio do artefato.

2.2.2 Avaliação Estética por Aristóteles

Lidando com aspectos filosóficos, a abordagem aristoteliana da estética concebeu uma escala de avaliação cujos critérios mensuram a estética de um objeto (seja ele material ou puramente conceitual).



Quadro 4 – Escala estética de Aristóteles

Dividindo a escala do Quadro 4 em duas categorias, Aristóteles classificou como harmônicas as classificações (ROCHA, 2007):

- **Gracioso:** Aquele que as características positivas apresentam-se em pequena intensidade, a beleza em pequenas dimensões, onde apesar da harmonia e proporções agradáveis lhe falta imponentia para ser considerado belo.
- **Belo:** Acrescenta o que falta ao gracioso, as dimensões e a imponentia mas sem exageros, sem sair do padrão ou do contexto.

- **Sublime:** O avanço em relação ao belo; com suas proporções avantajadas causa impacto no observador levando-o a pensar numa esfera mais elevada, em conceitos de nobreza, elevação e exacerbação das características positivas.
- **Trágico:** Maior expressão do harmonioso; Por extrapolar os conceitos da virtude e daquilo que é correto ao ser humano o trágico desperta a sensação de terror e piedade.

Considerando ainda como desarmônicas as seguintes classificações:

- **Risível:** Em paralelo ao gracioso quanto as proporções reduzidas, aponta para os fatores negativos por causar desconforto ou desagrado; Cujos deméritos não são tão graves.
- **Beleza do Feio:** Elementos propositalmente em desarmonia objetivando causar impacto no público; apresenta as conotações negativas presentes no risível, porém potencializadas conotando porte majestoso ou imponência.
- **Beleza do Horrível:** É caracterizada pelas dimensões em grande escala que levam o espectador a confrontar o “lado negro” da vida; retrata os elementos mais terríveis conhecidos pelo ser humano.
- **Cômico:** Visão desarmoniosa do trágico, geralmente exposta de forma caricata com a intensão de distorcer a realidade e não de gerar o riso.

Numa abordagem em design, Rocha (2007) realizou avaliação estética de artefatos utilizando-se da supracitada escala. Através da coleta de dados qualitativos em *Focus Group*, onde produtos eram apresentados a usuários e suas opiniões sobre os mesmos eram interpretadas ante a escala aristotélica, gerou-se os resultados abaixo (Quadro 5).

PRODUTO 1		HORRÍVEL	FEIO	RISÍVEL	GRACIOSO	BELO	SUBLIME
							X
		AVALIAÇÃO ESTÉTICA					
		PRODUTO: MERCEDEZ-BENZ CLASSE C SEDÃ					
PRODUTO 2		HORRÍVEL	FEIO	RISÍVEL	GRACIOSO	BELO	SUBLIME
			X				
		AVALIAÇÃO ESTÉTICA					
		PRODUTO: LAND ROVER DEFENDER					
PRODUTO 3		HORRÍVEL	FEIO	RISÍVEL	GRACIOSO	BELO	SUBLIME
				X			
		AVALIAÇÃO ESTÉTICA					
		PRODUTO: FORD KA					

Quadro 5 – Avaliação sob Escala Aristoleliana. Adaptado de ROCHA (2007)

De acordo com as considerações do referido autor, todas as colocações obtidas no *Focus Group* referente ao Produto 1 (“sensacional”, “fantástico”, “parece um avião”) o enquadravam na classificação de Sublime. Referente ao Produto 2, o “Feio” representa confiabilidade para um carro desta categoria, assim classificado por receber comentários como “mau encarado” e “cara de perverso”. Relativo ao Produto 3, devido aos comentários como “baratinha” e “esquisitinho”, sua classificação tendenciou para Risível. O que está em consonância com a proposta de um carro de baixo consumo de combustível, realização de pequenas tarefas e cumprir pequenos trechos. Estas atribuições diminutivas combinam com a classificação atribuída ao produto na escala estética de Aristóteles.

Com base no desenvolvimento desta escala, Rocha (2007) expôs o que denominou “*Peripatécnica*”. Uma técnica baseada na escala de Aristóteles para avaliação estética de games.

2.2.3 Avaliação Estética de Websites

Pirauá (2007) apresenta uma proposta de avaliação estética específica para websites, denominada “*TAE.web-u*”. Esta técnica se divide em etapas de coleta de dados com os usuários e outras de análise dos mesmos. As 2 primeiras etapas são onde o usuário faz sua avaliação do produto (que pode estar já lançando, mas preferencialmente esteja no estágio

de protótipo tornando mais oportunas eventuais correções). A primeira etapa é realizada uma espécie de coleta por Escala de Diferencial Semântico² de 5 pontos, cujo objetivo é investigar o impacto estético-emocional sobre o usuário (PORTO, 2009) (Quadro 6).

Informações sobre impacto emocional da estética visual sobre o usuário	Interpretação do usuário sobre a estética visual das formas do website
1. Atração – Repulsão 2. Surpresa agradável – Surpresa desagradável 3. Inspiração – Desestímulo 4. Diversão – Tédio 5. Admiração – Desprezo 6. Satisfação – Insatisfação 7. Fascinação – Frustração	1. Masculino – Feminino 2. Inerte – Ativo 3. Quente – Frio 4. Infantil – Maduro 5. Modesto – Sofisticado 6. Moderno – Tradicional 7. Refinado – Rústico 8. Excepcional – Medíocre 9. Sério – Extravagante 10. Original – Comum

Quadro 6 – Atributos da Escala de Diferencial Semântico por PIRAUÁ (2007)

Numa segunda etapa, coleta-se via entrevista outras informações que o autor julga pertinentes ao julgar uma avaliação estética (Quadro 7).

Perguntas	Objetivos
1. Numa escala de 0 a 10, que nota você daria à beleza deste site?	Visa obter a percepção estética global do usuário sobre o website.
2. Numa escala de 0 a 10, que nota você daria ao estilo (formato, tamanho e disposição) dos elementos do site?	Visa obter a percepção do usuário sobre aspectos específicos do website: • estilo (que concentra formato, tamanho e disposição dos elementos visuais); • cores; • fontes, que se refere às tipografias utilizadas.
3. Numa escala de 0 a 10, que nota você daria às cores usadas neste site?	
4. Numa escala de 0 a 10, que nota você daria às fontes que foram usadas?	
5. Você consegue identificar alguma área do site que lhe desagradou?	Busca localizar áreas que sejam fontes de problemas e que não tenham sido indicadas nas respostas anteriores.
6. Quais seriam as suas sugestões para melhorar a aparência deste site?	Busca coletar sugestões dos usuários sobre como melhorar a estética do website.

Quadro 7 – Entrevista sobre Estética de PIRAUÁ(2007)

² A Escala de Diferencial Semântico apresenta adjetivos contraditórios sobre algo que está sendo avaliado. Mais detalhes sobre esta escala no Capítulo 3, seção 3

Segundo Pirauá (2007), com base nestas informações é possível avaliar esteticamente os softwares compreendidos nos websites.

2.2.4 Avaliação de Artefatos Quanto à Estética e a Aleatoriedade

Assim como observado na avaliação de artefatos quanto à criatividade, ao avaliar critérios estéticos os pesquisadores buscam reduzir o acaso no processo de projeto coletando opiniões. Neste caso, coletam-se opiniões de usuários objetivando reduzir a incerteza quanto ao produto estar atendendo suas expectativas. Todavia, observa-se uma preocupação maior com a precisão na educação destas informações, pelo uso também de métodos qualitativos no processo de extração de informações.

Apesar da preocupação em reduzir a incerteza, nenhuma técnica específica lidando com redução de acaso no processo de projeto é observada.

2.3 Avaliação de Artefatos Quanto à Experiência com o Usuário

O design para experiência do usuário, também citado por UX (*User Experience*), objetiva observar os *feedbacks* relativos a satisfação do cliente final na utilização de um produto, através do uso de escalas Likert e Pictóricas. Permitindo assim correções e aprimoramentos quanto às experiências que o usuário sentirá antes de lançar o projeto no mercado. Constitui, portanto, uma ferramenta relativamente poderosa em termos estratégicos, na tentativa de prever o resultado final de aceitação de um produto. A principal especialidade na utilização de UX hoje é o seguimento do design de artefatos digitais, principalmente em aplicativos para celular e games (ROTO et al. 2008).

A satisfação do usuário em interagir com o produto é comumente denominada neste seguimento como “engajamento”, que representa o nível de envolvimento que o produto oferece, atraindo e mantendo o usuário interagindo, agradavelmente.

Convém elucidar que o termo “experiência” empregado neste tipo de avaliação denota “sentimentos e emoções” e não experimento no sentido metodológico/científico. Apesar do design para UX possuir publicações correlatas às pesquisas sobre usabilidade e design, bem como emoções e design, este seguimento de avaliação de artefatos é norteado por três particulares princípios básicos que o distingue (BELFORT, 2011; LAW et al. 2009). Estes levantam os axiomas de que toda experiência com o usuário é:

- **Situada:** Não é possível experimentar algo fora do contexto natural do evento;
- **Efêmera:** Uma vez que a experiência ocorre, não é possível repeti-la pois toda ocorrência é diferente;
- **Subjetiva:** Sendo acessível somente através próprio usuário.

Concentrando teorias de avaliação de comportamento, bem como mensuração de resultados por escala psicométrica, Belfort (2011) desenvolveu uma técnica para avaliação de UX, em especial no ramo de game design. Baseada em TRUE (*Tracking Real-time User Experience*), esta técnica coleta dados relativos ao engajamento do usuário durante a interação do mesmo com um game casual online.

Avaliando o game em 2 frentes, primeiramente um especialista analisa trechos do mesmo diante de um quadro chamado “fluxo de potência”, onde a “potência” determina o nível de engajamento em potencial que aquele trecho do artefato oferece (Ver Quadro 8).



Quadro 8 – Análise de Fluxo de Potência para UX. Adaptado de BELFORT(2011)

Neste quadro o especialista sinaliza um determinado trecho do game quanto ao desafio que ele oferece naquele momento e a habilidade necessária para jogar no mesmo, diante de uma escala Likert³ de 5 pontos (-2,-1,0,1,2) para cada uma das dimensões (desafio e habilidade). Em cada quadrante do Quadro 8 sinaliza-se como o usuário se sente ao posicionar-se nele (Frustrado, medíocre, motivado ou desmotivado). Estudo criado com

³ A Escala Likert é uma ferramenta de avaliação psicométrica pontuada, também usada para avaliação de artefatos. Maiores informações sobre no Capítulo 3, seção 3.

base nos trabalhos de Csíkszentmihályi (1997), onde a área de fluxo do Quadro 8 representa a área de engajamento que o usuário com as habilidades necessárias sendo desafiado estará experimentando felicidade e motivação intrínseca. Assim, trechos do game que recebam de um especialista avaliação menor que 0 (zero) quer em desafio, quer em habilidade, são os trechos que tendenciarão ao desengajamento, ou seja, ao desinteresse ou frustração do usuário. Requerendo atenção e redesign.

Após realizada essa marcação de diversos trechos do game por um especialista, num segundo momento o mesmo é disponibilizado online para que agora os usuários possam opinar jogando, mas somente nos trechos em que o especialista marcou com o fluxo de potência deficiente e onde os designers concentrarão seus esforços para melhorar o engajamento dos usuários, utilizando-se de todas as técnicas de design centrado no usuário disponíveis.

A pesquisa realizada por Belfort (2011) avaliou o game “*Corrida Contra o Tempo*” (Figura 5), onde o jogador ajuda uma personagem a correr de patins por diversos obstáculos utilizando o teclado. Esta pesquisa trouxe, dentre diversos resultados, uma projeção explicitando o quanto a opinião dada pelo especialista relativa ao engajamento em potencial de trechos do game correspondiam ao *feedback* de engajamento dos usuários.



Figura 5 – Telas do game “Corrida Contra o Tempo” (BELFORT, 2011)

Para coletar os dados do usuário ao final de cada um dos trechos selecionados, solicitava-se rapidamente a informação “Como você está se sentindo?”, conforme a tela da Figura 6 abaixo. Esta coleta também seguia a escala psicométrica de Likert de 5 pontos, a mesma escala utilizada para que os especialistas marcassem o fluxo de potência destes trechos.

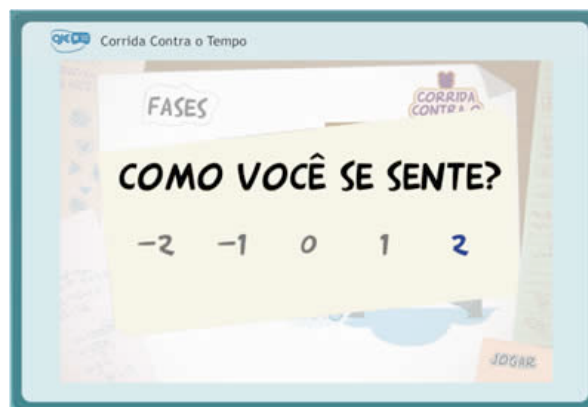
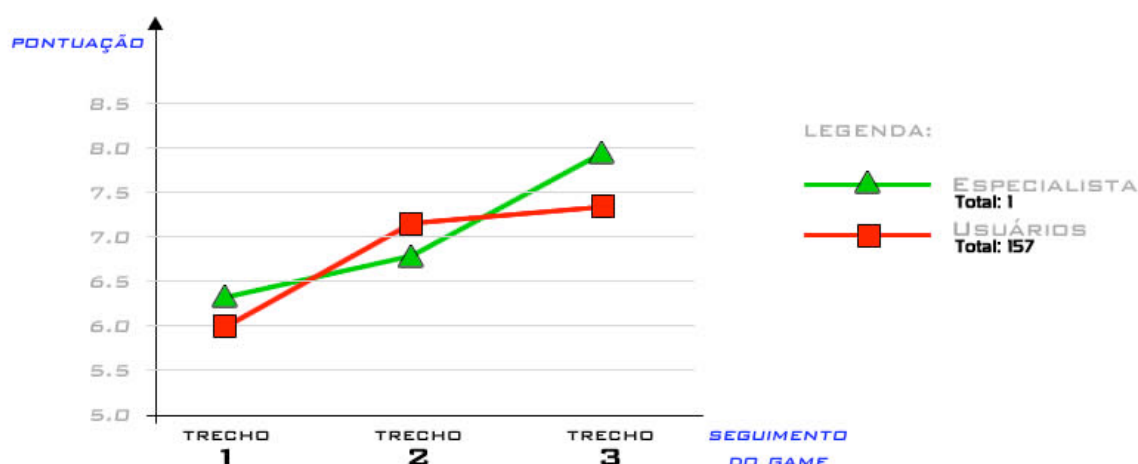


Figura 6 – Tela de coleta de dados para UX (BELFORT, 2011)

Neste caso, 3 trechos foram avaliados e objetivando facilitar a visualização dos resultados, na projeção o autor converteu os resultados de 5 pontos da escala Likert (-2,-1,0,1,2) em 2, 4, 6, 8, 10. Gerando assim o Quadro 9 abaixo, onde fora possível considerar que os especialistas marcaram com relativa proximidade o quanto o game obteria de aceitabilidade pelos usuários.



Quadro 9 – Quadro de avaliação de UX. Adaptado de BELFORT(2011)

Esse é o resultado do *feedback* obtido por 157 usuários, sendo atribuídas as pontuações 6,58 para o Trecho 1, bem como 6,70 para o Trecho 2 e 7,33 para o Trecho 3 (Quadro 9). Dessa maneira, Belfort (2011) avaliou a experiência com o usuário num produto digital.

2.3.1 Avaliação de Artefatos Quanto à Experiência com o Usuário e a Aleatoriedade

Repetindo as preocupações em reduzir a aleatoriedade do resultado final do produto no mercado observadas nas avaliações de criatividade e estética, os avaliadores de artefatos quanto a experiência com o usuário tentam reduzir o acaso por coleta de opiniões. Todavia, não existe nenhum critério quanto a quantidade de usuários participantes desta coleta de opiniões para se ter um resultado consistente relativa a experiência com o usuário reportada

na pesquisa ser consistente. Ou seja, talvez uma quantidade maior, ou menor, de usuários participantes possa alterar o resultado o que significa que a pesquisa realizada por Belfort (2011) continua sujeita ao acaso.

2.4 Avaliação de Artefatos Quanto à Usabilidade

Nesta seção da tese apresentam-se algumas técnicas de avaliação de usabilidade, destacando a importância dada nestes contextos às falhas de interface apontadas por usuários treinados ou visualizadas por um observador que registra das dificuldades no uso do protótipo por um usuário não-treinado. Na maioria dos casos não existe um senso ou justificativa científica para a quantidade de avaliadores participando do processo, visando uma maior precisão do mesmo.

As avaliações de usabilidade são especialmente bem difundidas no o design de artefatos digitais, dado o número de publicações recorrentes neste setor. Este campo avalia produtos pelo uso de heurísticas (regras) de usabilidade (NIELSEN & TAHIR, 2002). Muitas vezes com controversas quanto a quem deveria ser o avaliador ou quantos avaliadores deveriam participar deste processo, conforme melhor detalhado adiante.

Conceituando o tema, o termo “usabilidade” refere-se a “facilidade do uso”, ou seja, o quão fácil um produto ou sistema é de usar (JORDAN, 1998). O design de uma interface com alto índice de usabilidade está associado ao design de uma “interface amigável ao usuário”, ou seja, ao conceito de “*user-friendly*”. Considerando que um produto – sendo ele material como um telefone móvel ou imaterial como uma tela de software – interage com o usuário pela interface, a avaliação de usabilidade objetiva extrair dados desta interação que possibilitem analisar o quão facilmente a mesma ocorre.

A demanda pela avaliação da usabilidade em produtos digitais emergiu após evidenciar-se que os websites “difíceis de usar” geravam uma frustração e repulsa tal nos usuários, que dificilmente eles optariam por navegar novamente nestas páginas da web (KRUG, 2001). Considerando que a aversão a sites com má usabilidade ocasionava a redução de presença na internet das marcas e dos produtos neles representados, rapidamente os profissionais e pesquisadores de web design responderam apontando mecanismos metodológicos para solucionar este problema (NIELSEN & TAHIR, 2002).

Considerando o advento da internet como é conhecida hoje – e dos websites – a partir dos anos 1990, observa-se desde então o surgimento de uma série de pesquisas objetivando aperfeiçoar a usabilidade. Estas publicações reportam o que ficou conhecido neste seguimento como “testes de usabilidade”. Nos parágrafos a seguir, encontram-se sintetizadas algumas destas técnicas de avaliação.

2.4.1 Percurso Cognitivo (“Cognitive Walktought”)

O percurso cognitivo é utilizado para verificar se um usuário consegue interagir intuitivamente com uma interface com a qual ele se depara pela primeira vez. Assim como em todas as técnicas empíricas de análise de usabilidade, o participante da pesquisa representa o público-alvo (usuário) em potencial do sistema.

Para isto, investiga-se a aprendizagem exploratória do participante ao navegar pela interface de um sistema informatizado, sem qualquer treinamento ou familiarização prévia (MANO & CAMPOS, 2004). Utilizando um objeto (o sistema), e tarefas (conjunto de ações de navegação no cumprimento de um objetivo), o pesquisador tenta responder, durante o teste de usabilidade, as seguintes questões.

- A ação que o usuário pretende realizar está presente na interface?
- A ação correta é evidente o suficiente para o usuário?
- A ação correta está associada ao que o usuário pretende fazer?
- O usuário irá interpretar corretamente o *feedback* do sistema diante da ação executada?

A tarefa executada poderia ser, por exemplo, acessar um sistema informatizado de uma universidade e editar um sumário de aulas, como no estudo de caso de Mano & Campos (2004). Para isto, o usuário deveria intuitivamente executar uma série de ações, interagindo com este sistema, como clicar em links, botões, “*check boxes*”, preencher formulários, e etc. Para cada participante executando a tarefa, gera-se um quadro onde um observador treinado responde binariamente (sim ou não) às perguntas supracitadas. No final, um quadro com o somatório das respostas positivas e negativas revela ao designer aspectos do sistema que precisam de maior atenção no aperfeiçoamento.

2.4.2 Co-Descoberta (“Co-Discovery”)

Com uma abordagem totalmente qualitativa, a técnica da co-descoberta posiciona 2 participantes que normalmente se conhecem, na exploração da interface de um sistema

(JORDAN, 1998). O observador treinado (que registra os eventos e erros de usabilidade que julgar relevantes) pode interferir a qualquer momento, solicitando novas tarefas aos participantes ou mesmo ajudando-os em algum problema encontrado durante a navegação pelo sistema. O fato dos participantes se conhecerem, ou mesmo serem amigos, atua a favor do teste uma vez que diminui a inibição e o constrangimento de apontar as ações que não estão conseguindo executar. Haja vista que as intimidações provocadas por um laboratório e um observador podem resultar em alguns problemas de usabilidade não reportados pelos participantes, principalmente os mais elementares. O observador também pode ausentar-se do ambiente e acompanhar o processo remotamente, pois todo ele é filmado. Desta maneira estima-se registrar os entraves de uso de um produto antes mesmo dele estar no mercado, otimizando as chances de êxito.

2.4.3 Observações de Campo (“Field Observations”)

Reduzindo ainda mais os efeitos de intimidação que um laboratório e uma sessão de testes de usabilidade possam causar nos participantes, aplica-se a técnica da observação em campo (JORDAN, 1998). Nela, a observação ocorre com o sistema sendo utilizado diretamente pelo usuário final, em seu ambiente real. Como uma máquina de autoatendimento bancário, por exemplo. Por não saberem que estão sendo observados, os usuários irão realizar as suas tarefas livremente. Todavia, não sendo designada uma tarefa específica, o observador terá uma margem de interpretação dos dados bastante ampla, no sentido de entender o que o usuário quer fazer e se está obtendo êxito em usar o sistema ou não. Outro impecílio na aplicação desta técnica é a ética. Diversas abordagens tentam apresentar como extrair estes dados do usuário sem que o mesmo saiba que está sendo observado e, ao mesmo tempo, não infringir padrões morais ou éticos. Uma das opções é informar ao usuário, quando ele finalizar suas tarefas no sistema, que os dados da interação foram registrados e solicitar autorização de envio deles para uma pesquisa, visando colaborar com a mesma (BRANDÃO, 2006).

2.4.4 Avaliação Cooperativa (“Cooperative Evaluation”)

Recomendada para protótipos inacabados, a avaliação cooperativa posiciona pesquisador e participantes juntos numa discussão sobre a interface analisada (MIRANDA & MORAES, 2003). Durante a execução deste procedimento o pesquisador faz perguntas direcionadas sobre o entendimento da interface e os participantes também tem abertura para fazerem suas perguntas de volta. Esta abordagem permite uma coleta qualitativa de dados, por vezes, inesperados. Além do denso *feedback* sobre a interface, apontando ao designer experiente o que precisará ser reformulado antes da finalização do protótipo, o baixo custo deste procedimento facilita sua aplicação.

Este teste de usabilidade demanda de uma cautelosa escolha dos participantes que realmente representem o espaço amostral do usuário final. Caso contrário, o participante pode injetar na pesquisa dados qualitativos que consomem tempo de interpretação e não levam a problemas relevantes para a reformulação da interface. As etapas para execução desta pesquisa podem ser divididas no encadeamento abaixo (TEIXEIRA & MORAES, 2004).

- Fase 1: Recrutamento de participantes (1 ou mais);
- Fase 2: Especificar tarefas relevantes no uso da interface;
- Fase 3: Liberar o uso da interface pelos participantes, preferencialmente de forma assíncrona. Os mesmos devem comunicar verbalmente ao pesquisador as dificuldades (problemas) de uso da interface a medida que forem encontradas;
- Fase 4: O pesquisador interpreta seus próprios registros, descreve os problemas acompanhados de solução proposta.

Ainda de acordo com Teixeira & Moraes (2004), o observador deve ser um designer com experiência em design de interfaces, para que possa propor soluções de projeto adequadas ante os problemas identificados pelo teste de usabilidade.

2.4.5 Diário de Incidentes (“Incident Diaries”)

Utilizado com protótipos amadurecidos, o diário de incidentes utiliza formulários resumidos que dividem o processo em duas fases: uma qualitativa e outra quantitativa. Inicialmente o participante preenche discursivamente um formulário onde responde o problema encontrado com breve descrição (incluindo o que o incomoda neste problema) e qual sua sugestão de solução, gerando assim os dados qualitativos os pesquisador (JORDAN 1998).

Num segundo momento, perguntas pertinentes a interface são direcionadas pelo pesquisador num formulário com respostas disponíveis dentro de uma escala Likert de 5 pontos.

Os problemas reportados pelo usuário apenas uma vez são considerados dificuldades de familiarização cognitiva, já os problemas apontados repetidamente representam uma dificuldade de aprendizado que demandam maior atenção no redesign da interface (BRANDÃO, 2006)

2.4.6 As Heurísticas de Jakob Nielsen

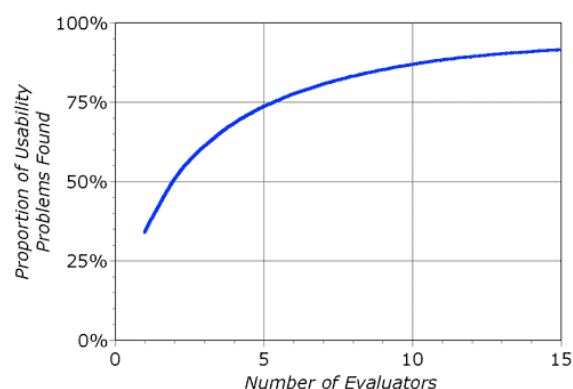
Dentre todas as avaliações de usabilidade possivelmente a mais popular, dada a quantidade de recorrências e citações, seja a baseada nos princípios heurísticos publicados por Nielsen (1993). Este autor postulou um total de 10 heurísticas (regras) que, quando deixam de ser observadas numa interface indicam pontualmente problemas de usabilidade encontrados no sistema, conforme detalhado abaixo.

1. **Feedback** – O sistema deve informar continuamente ao usuário o que ele está fazendo e o que está sendo processado pelo sistema.
2. **Relação análoga entre o sistema e o mundo real** – Ser coerente com o modelo mental do usuário através da utilização de palavras e conceitos familiares ao mesmo. A utilização de palavras e termos técnicos de sistema que não fazem sentido ao usuário provocará seu não entendimento da interface.
3. **Controle do usuário e liberdade** – Possibilitar ao usuário desfazer ações, retornando ao ponto anterior. Criar uma saída fácil de retorno a interface principal para quando o usuário sentir-se perdido.
4. **Consistência e padrões** – Uma mesma ação deve ser apresentada na mesma localização e da mesma maneira. Não devem existir palavras ou situações diferentes que signifiquem a mesma coisa.
5. **Prevenção de erro** – Evitar situações de erro, modificando a interface para que não ocorram.
6. **Reduzir sobrecarga de memória** – É preferível que usuário reconheça imediatamente o que um comando ou ação faz, ao invés de ter que lembrar o que ele faz.
7. **Flexibilidade e eficiência de uso** – Atalhos que propiciem a usuários experientes executarem operações mais rapidamente, sem necessidade de navegar amplamente na árvore informacional.
8. **Estética e design minimalista** – Apresentação da informação que o usuário necessita para entender a ação, nem mais, nem menos.
9. **Diagnóstico de erros** – Mensagens de erro devem ser simples, não intimidadoras e apontar uma solução possível.
10. **Ajuda e documentação** – A princípio a interface deve ser intuitiva e fácil de usar, mas em caso de dúvidas o usuário deve contar com uma documentação de ajuda detalhada acessível sem impecílios.

Ao contrário dos outros testes de usabilidade anteriormente topificados nesta seção, neste caso os participantes são especialistas (designers) treinados, e não potenciais usuários do sistema em análise. Estes participantes que levantam, topificam e detalham os problemas encontrados numa escala por nível de gravidade.

- **Gravidade 0 (Sem importância)** – Não considerado um problema de usabilidade;
- **Gravidade 1 (Superficial)** – Problema estético apenas, não necessita de reparo a não ser que o projeto disponha de tempo extra;
- **Gravidade 2 (Simples)** – Pequeno problema de usabilidade, solucionar com baixa prioridade;
- **Gravidade 3 (Grave)** – Grande problema de usabilidade, importante solucionar. Atribuir alta prioridade;
- **Gravidade 4 (Catastrófico)** – Catástrofe em usabilidade, solucionar imperativamente antes do lançamento do produto.

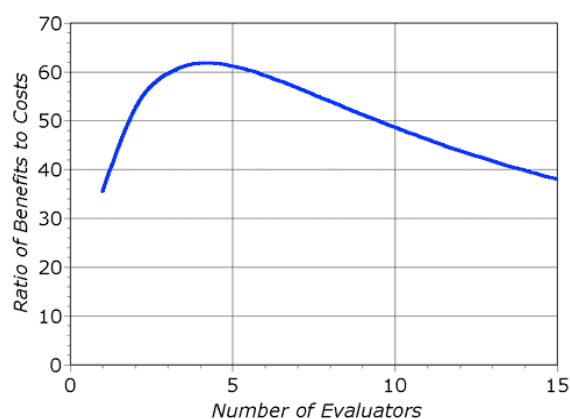
Segundo Nielsen & Landauer (1993) um total de 5 especialistas realizando o teste é o suficiente para identificar cerca 70% dos problemas de probabilidade, chegando a cerca de 85% com 8 avaliadores (Ver Quadro 10). Este posicionamento é precedido por Virzi (1992) com sua pesquisa que reportando até 80% dos problemas encontrados com 5 avaliadores.



Quadro 10 – Quantidade de avaliadores (NIELSEN & LANDAUER, 1993)

Segundo estes autores, à medida que aumenta-se o número de avaliadores da interface, alguns problemas de usabilidade começam a ser identificados por mais de um dos especialistas, tornando-se recorrentes. Chegando ao ponto em que um grande número de avaliadores encontra o mesmo problema. Com base neste argumento, Nielsen & Landauer

(1993) apontaram que após 5 avaliadores, inicia-se um declínio no custo benefício do teste de usabilidade, conforme disposto no Quadro 11. Em contrapartida, Perfetti & Landesman (2001) reportam que com 5 avaliadores encontraram apenas 35% dos problemas de usabilidade de um website específico.



Quadro 11 – Custo-benefício por avaliadores (NIELSEN & LANDAUER, 1993)

Esta controvérsia sobre a quantidade de avaliadores é aprofundada nesta tese, no capítulo 3, Problemas na Avaliação de Artefatos.

Para uma melhor compreensão da aplicação deste teste de usabilidade, no APÊNDICE A encontra-se a critério de exemplo uma das avaliações de um website de acordo com as heurísticas de Nielsen, realizada por um dos especialistas participantes na pesquisa de Melo (2010).

2.4.7 Outras Avaliações de Usabilidade por Jordan

Além dos procedimentos delineados nesta seção que objetivam avaliar a facilidade de uso de websites e demais sistemas informatizados, Jordan (1998) fornece um levantamento detalhado de diversos outros métodos de pesquisa originadas em outros domínios – como marketing e estatística – adaptados aos testes de usabilidade. No Quadro 12, encontram-se sintetizados estes procedimentos, onde normalmente os participantes representam o público-alvo em potencial da interface, ou seja, os usuários finais do produto e o pesquisador é o condutor das coletas de dados. Em alguns casos, os participantes são especialistas com comprovada expertise em design e usabilidade.

Avaliação	Resumo
Entrevistas (Interviews)	Entrevistas sem-estrutura, semi-estruturadas e estruturadas direcionadas pelo pesquisador aos participantes.

Grupo de foco (Focus group)	O pesquisador conduz uma conversa com os participantes, seguindo um roteiro de tópicos e assegurando que todos tenham oportunidade de externar suas opiniões com equilíbrio.
Experimentos controlados (Controlled experiments)	Objetivando testar uma ou mais hipóteses específicas, o pesquisador cria tarefas e as observa sendo executadas por participantes num ambiente artificial, ciente de que eles não responderão exatamente como num ambiente real de uso da interface.
Lista de verificação de checklists (Feature checklists)	Participantes preenchem formulário binário sinalizando quais característica da interface são utilizadas ou não (ao invés de verificar se são fáceis de usar ou não).
Método de avaliação (Valuation method)	Levantamento de o quanto os participantes pagariam por mais uma interfaces que contenha características pré-estabelecidas. Quando mais indivíduos “pagando” por uma determinada característica, maior será o foco do design em otimizar a usabilidade da mesma.
Protocolos “Pensar alto” (Think aloud protocols)	O participante fala o que está fazendo e pensando ao utilizar uma interface por tarefa designada ou exploração livre, a escolha do pesquisador.
Questionários (Questionnaires)	Questionário de perguntas abertas ou fechadas, enviados por email ou presencialmente. Particularmente útil quando o designer não sabe por quais problemas começar.
Registro de conversações (Private camera conversations)	Numa cabine, um participante fala para uma câmera sua opinião sobre uma interface. Uma variação deste procedimento é ter 2 pessoas na mesma cabine debatendo sobre a interface para a câmera.
Registro de uso (Logging use)	Aplicativos (programas de computador) instalados na interface digital para levantamento estatístico de cliques e demais ações dos usuários.
Análise de tarefa (Task analyses)	Levantamento da quantidade de passos necessários para completar uma tarefa. Quanto mais passos, mais complexa se torna a mesma, carecendo de maior atenção quanto a usabilidade.
Avaliação de peritos (Expert appraisals)	Similar a avaliação heurística de Nielsen & Landauer (1993) onde especialistas com relevante expertise em usabilidade avaliam a interface. Todavia, neste caso podem avaliar em conjunto, influenciando as idéias uns dos outros, discutindo e amadurecendo o debate sobre os problemas de uma interface em específico.
Listas de verificação de propriedades (Property checklists)	Lista de critérios de usabilidade que esperam-se ser avaliados na interface entregues a um avaliador com expertise no seguimento. Alguns critérios de programação visual que impactem diretamente na usabilidade também podem constar na lista composta pelo pesquisador, para verificação do especialista.

Quadro 12 – Avaliações de usabilidade (JORDAN, 1998)

2.4.8 Avaliação de Artefatos Quanto à Usabilidade e a Aleatoriedade

Ao contrário do apresentado nas pesquisas sobre avaliação de artefatos quanto a criatividade, estética e experiência com o usuário, durante a avaliação quanto à usabilidade tentou-se reduzir o acaso do processo de projeto ao justificar uma quantidade de usuários que reportariam uma opinião consistente com o público-alvo final (NIELSEN & LANDAUDER, 1993). Todavia, convém validar se esta proposta procede cientificamente, conforme proposto no capítulo seguinte da presente tese.

2.5 Avaliação de Artefatos Quanto ao Mercado

O design é nos dias de hoje tratado como assunto sério nos mais altos escalões das empresas. As decisões estratégicas destas empresas são em muitos casos também decisões de design. (BURDEK, 2006, pg. 363)

Esta seção da tese apresenta uma introdução à importância da avaliação mercadológica em design e os aspectos considerados relevantes neste contexto, posicionando o designer como uma peça fundamental no planejamento estratégico das empresas. Neste âmbito, são avaliados: produtos complementares e substitutos; curvas de demanda e de oferta; benefícios básicos; ciclo de vida do produto, conforme detalhado adiante (FIGUEIROA et al. 2012). Todos estes aspectos visam eduzir o máximo de informações relevantes sobre o produto e o público alvo, reduzindo a incerteza e o acaso no processo de design.



Figura 7 – Concorde – Produto inovador (Foto: Ciorlaus Vicentiu – airliners.net)

Não são raros os casos em que um produto é bem avaliado numa ampla variedade de aspectos, mas sai de circulação por uma má avaliação mercadológica. Por exemplo, o projeto da aeronave Concorde nos anos 1970, como destacado por Lobach (2001). O produto demonstrou-se um relevante avanço científico-tecnológico, sendo a primeira aeronave comercial capaz de desempenhar voo supersônico (Ver Figura 7). Durante a avaliação mercadológica era possível observar diferenciais atrativos em relação aos seus concorrentes mais próximos, como o Boeing 747 (também conhecido por “Jumbo”), cumprindo o trecho Paris-Nova Iorque em apenas 3 horas e trinta minutos, enquanto seu concorrente realizava o mesmo percurso com o dobro do tempo. Além dos voos mais rápidos reduzirem o desgaste físico dos passageiros, o Concorde voava a uma altitude de 60.000 pés, acima das camadas de ar que pudessem causar turbulências. Diversos outros aspectos técnicos e operacionais com vantagens aos clientes podem ser listados. Contudo, o fato é que apesar da grande inovação desse produto cujo projeto custou na casa dos bilhões de dólares, seria necessário uma centena de anos para que ele se tornasse um investimento lucrativo. O Concorde teve seu ciclo de vida limitado a 3 décadas resultando, assim, num fracasso comercial (BAXTER, 2011).

Quanto mais um projeto demandar por recursos, maiores serão as exigências de justificativas (considerando todos os grupos de avaliação de artefatos apresentados nesta tese), inclusive financeiras (Ver Quadro 13). Para Baxter (2011), algumas empresas avaliam seus protótipos frente a detalhadas projeções financeiras e por posicionamento num plano de marketing. Outras, no entanto, adotam métodos menos formalizados, como uma simples discussão em reunião de diretoria. Qualquer que seja o rigor adotado pela empresa, a avaliação das justificativas de sucesso do produto proposto, tratando-se de aspectos mercadológicos, exige o esclarecimento dos questionamentos abaixo pontuados.

- **Custos variáveis do produto** – Insumos por unidade de produto em relação à quantidade vendida. Nestes custos incluem-se: matéria-prima, mão de obra, energia, distribuição e vendas.
- **Custos fixos do produto** – Não relacionados diretamente a quantidade vendida, como compra de maquinário, ferramentas e demais equipamentos, além do salário dos diretores e aluguel do prédio.
- **Meta de preço** – Margem entre o preço do produto e seus custos de produção.

- **Ciclo de vida do produto** – Projeção de qual será o lucro total previsto durante toda a vida do produto no mercado e quanto tempo será necessário para recuperar os custos fixos.

Empresa	Stanley	Hewlett-Packard	Chrysler	Boeing
Produto	Chave de fenda motorizada	Impressora jato de tinta 500	Automóvel Concorde	Aeronave Boeing 777
Quantidade de peças	3	35	10 000	130 000
Prazo de desenvolvimento	1 ano	1,5 anos	3,5 anos	4,5 anos
Equipe de desenvolvimento	3 pessoas	100 pessoas	850 pessoas	6 800 pessoas
Custo do desenvolvimento	\$ 150 mil	\$ 50 milhões	\$ 1 bilhão	\$ 3 bilhões
Preço de venda	\$ 30	\$ 365	\$ 19 000	\$ 130 milhões
Produção anual	100 000	1,5 milhões	250 000	50
Ciclo de vida do produto	4 anos	3 anos	6 anos	30 anos
Custo desenvolvimento/vendas totais	1,2%	3%	3,5%	1,5%

Quadro 13 – Custos de desenvolvimento (BAXTER, 2011)

Estes parâmetros de avaliação são especialmente necessários no lançamento de produtos inovadores, uma vez que reduzem as incertezas envolvidas no projeto e, por vezes, impede que um produto promissor “a frente do seu tempo” entre no mercado no momento errado.

2.5.1 Produtos Complementares e Substitutos

Durante a avaliação da possibilidade de sobrevivência de um produto ao seu lançamento no mercado, convém avaliar o volume de vendas dos produtos correlatos. Neste âmbito, cabe ao designer fazer um levantamento do desempenho comercial dos produtos complementares e substitutos.

- **Produtos complementares** – Bens cujos o aumento das vendas de um implica no aumento das vendas do outro (VARIAN, 2003). O acréscimo das vendas de smartphones, por exemplo, aumentam diretamente as vendas de aplicativos e games para telefones celulares. Impressoras e cartuchos de tinta, armas e munições, carros e acessórios automotivos, dentre outros, são exemplos de bens complementares.

- **Produtos substitutos** – Bens cujos o aumento das vendas de um implica no declínio das vendas do outro (VARIAN, 2003). São, normalmente, produtos concorrentes no fornecimento de uma mesma função ou mesmo benefício ao

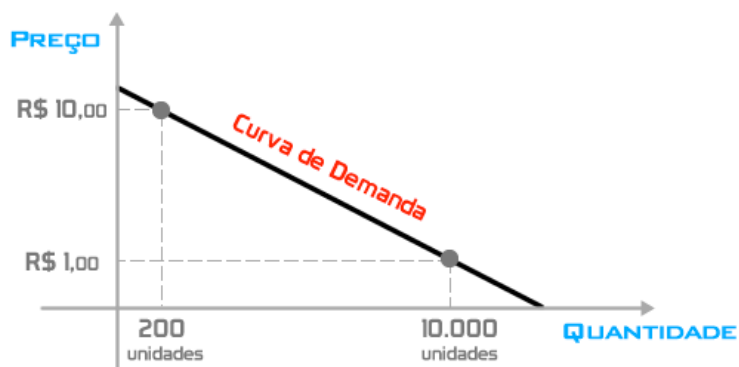
consumidor. Por exemplo, com o declínio dos preços das passagens aéreas uma faixa de usuários deixa de realizar determinados tipos de viagens de ônibus para realizar de avião.

Uma verificação deste levantamento de dados pode indicar ainda tendências de mercado para o correto posicionamento do produto em avaliação.

2.5.2 Curva de Demanda e Curva de Oferta

Durante o projeto de um produto, o designer levanta diversas informações sobre o público-alvo do produto proposto. Estas informações são estrategicamente utilizadas na verificação mercadológica quanto a real demanda dos usuários por este produto, bem como se é vantajoso para a indústria produzir e ofertar o mesmo (BAXTER, 2011). A partir das exemplificações sobre “demanda” e “oferta” (e suas respectivas “curva de demanda” e “curva de oferta”) abaixo delineadas, é possível exemplificar como um produto pode ser mercadologicamente avaliado neste contexto.

Por demanda entende-se a quantidade de produtos que compradores desejam e podem adquirir a diversos níveis de preço. A projeção quantidade por preço chama-se curva de demanda. De acordo com o Quadro 14 abaixo, simulando o preço de um produto fictício como um novo modelo de lápis, ao preço de R\$ 1,00 real, 10 mil pessoas demonstrariam interesse em comprá-lo. Enquanto que este mesmo lápis, se oferecido ao preço de R\$ 10,00 reais, uma quantidade muito menor de pessoas teriam interesse em adquiri-lo. Obviamente, quanto maior o preço, menor a quantidade de pessoas se deslocando ao longo da curva de demanda.

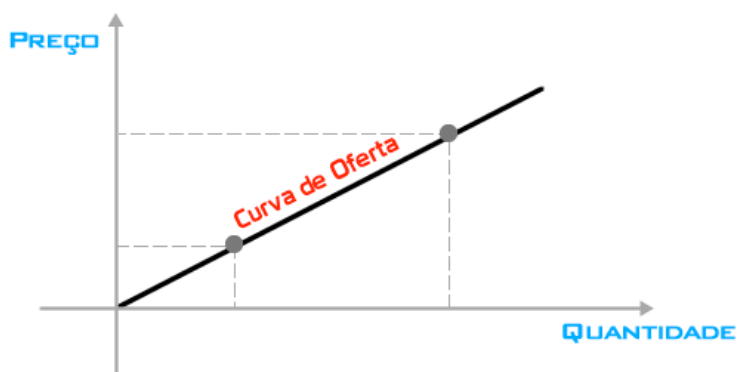


Quadro 14 – Curva de demanda

Enquanto que o preço desloca a quantidade de interessados na compra ao longo da curva de demanda, os demais fatores abaixo detalhados deslocam toda a curva (SILVA, 2001).

- **Renda** – quando a renda do público-alvo declina, a quantidade de demanda acompanha em direta proporção;
- **Preços de produtos relacionados** – O desempenho comercial de produtos complementares ou substitutos influencia a demanda do produto em questão;
- **Gostos** – A moda influencia os gostos temporários do público-alvo, afetando a demanda;
- **Expectativas** – O que se espera do futuro pode determinar o consumo no momento. Por exemplo, ao saber que o preço da gasolina irá aumentar no dia seguinte, aumenta-se a frequência nos postos de gasolina hoje.

Enquanto a curva de demanda é uma tentativa de projetar o interesse (comercial) dos consumidores por um produto, a curva de oferta mostra o potencial interesse das empresas em produzir o mesmo. Por oferta, entende-se a quantidade de um produto ou serviço que os vendedores estão dispostos e são capazes de vender, num determinado período de tempo. Tomando novamente por base a projeção fictícia da curva de oferta de um novo tipo de lápis, quanto maior o preço, maior o lucro, portanto a empresa teria interesse em oferecer em maior quantidade esse produto (Quadro 15).



Quadro 15 – Curva de oferta

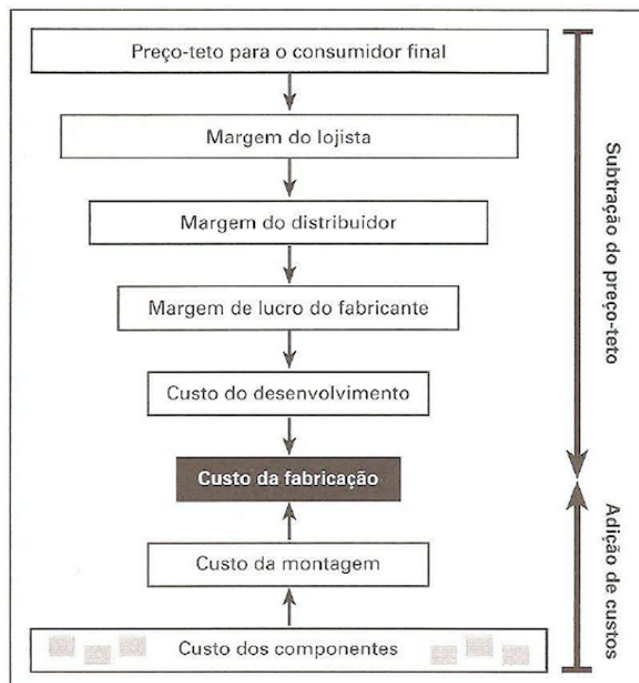
Fixando o preço como um dos fatores determinantes, os demais abaixo pontuados deslocam toda a curva de oferta (SILVA, 2001).

- **Preço dos insumos** – Quanto maior o custo de produção, menor a margem de lucro, reduzindo a quantidade oferecida;

- **Tecnologia** – Avanços tecnológicos diminuem os custos de produção, aumentando a margem de lucro e interesse da empresa em oferecer maior quantidade do produto;
- **Expectativas** – Expectativas futuras quanto ao preço podem alterar a quantidade oferecida.

A projeção combinada da curva de demanda com a curva de oferta permite encontrar o que os economistas chamam de ponto de equilíbrio. Este representa a inflexão dos interesses dos consumidores com a quantidade ideal de produção da empresa. Uma produção acima da quantidade determinada pelo ponto de equilíbrio ocasionará o semelhante à injeção de moeda numa economia. Produtos demasiadamente disponíveis forçam o preço dos mesmos a cair, reduzindo o lucro. Caso uma empresa observe que o ponto de equilíbrio não será o suficiente para retornar os custos de desenvolvimento, deve reprojeter seu produto ou reavaliar seu lançamento neste momento.

Assim com o designer trabalha em comunicação com a equipe de engenharia na checagem de materiais e viabilidade técnica de construção de um produto, se faz necessário consultar o engenheiro de produção e, por vezes, um consultor técnico financeiro para projetar com precisão os dados de demanda. Avaliando assim a viabilidade mercadológica do produto proposto. Para facilitar o prognóstico do preço do produto consideram-se os fatores destacados no Quadro 16.



Quadro 16 – Componentes do preço do produto (BAXTER, 2011)

Convém destacar novamente que conforme aumenta o valor do projeto, maior a necessidade de dados que reduzam as incertezas quando ao seu resultado no mercado. Portanto, dependendo destes valores, uma análise de demanda pode ser crucial para justificar a avaliação financeira de um produto.

2.5.3 Benefícios Básicos

Os produtos devem oferecer aos consumidores uma nítida vantagem sobre os produtos concorrentes, também conceituados nesta seção como produtos substitutos. Cada característica do produto que representa um diferencial nesta competição chama-se benefício básico (BAXTER, 2011). No caso do mercado de pilhas elétricas, normalmente o único benefício básico oferecido é o tempo de duração (Figura 8). Sendo esta outra função do benefício básico, dar argumentos para que os publicitários concentrem nele os esforços de propaganda do produto.

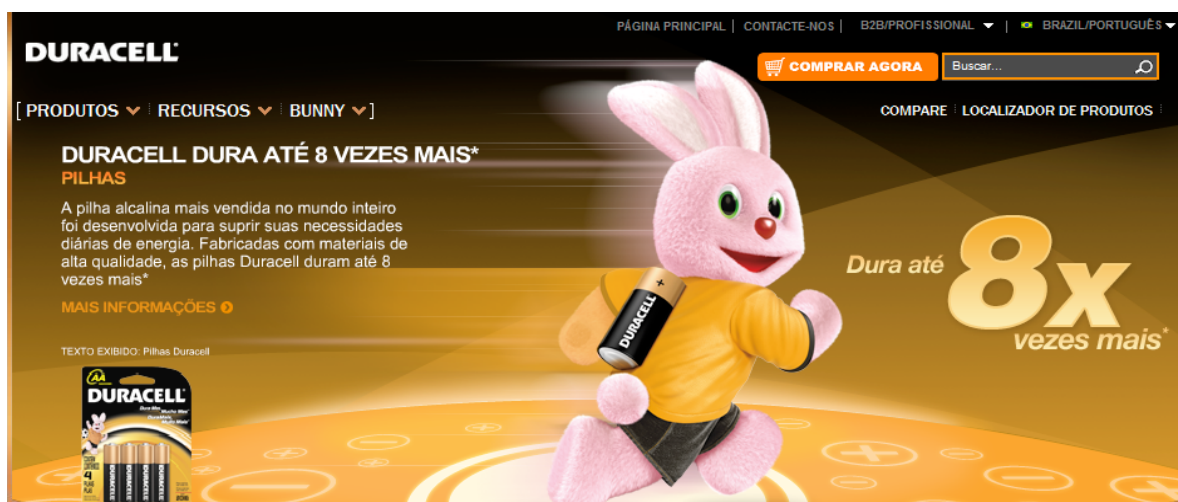


Figura 8 – Benefício básico: durabilidade da pilha (Fonte: duracell.com.br)

Às vezes os benefícios básicos são especificados desde o briefing do produto. Mas a *fase de geração de alternativas*⁴ pode gerar inovações no produto ou mesmo propor algo totalmente novo, com novos benefícios básicos estrategicamente mais vantajosos. Sendo necessária uma re Checagem na fase de avaliação tanto para verificação se os benefícios básicos exigidos no briefing foram atingidos quanto para visualizar se os novos benefícios atingem o público-alvo e destacam o produto dos concorrentes.

2.5.3 Ciclo de Vida do Produto

Existem duas análises de ciclo de vida do produto ortogonalmente dispostas, mas com exatamente a mesma nomenclatura e provenientes na engenharia de produção. Uma refere-se a análise cadeia de valor do produto, ou seja, todos os estágios do nascimento a morte do mesmo desde a coleta dos materiais de fabricação, passando pela distribuição e venda até o descarte. Conceito de análise vinculado à sustentabilidade (FERREIRA, 2004).

O outro conceito, o que estamos adotando na presente tese, é a análise do ciclo de vida comercial do produto (KOTLER & KELLER, 2006). A correta observação do ciclo de vida proposto, e as estratégias de mercado para cada fase do mesmo, objetivam permitir a visualização de todo o retorno financeiro que o produto fornecerá ao longo de sua existência no mercado. Desse modo, a diretoria da empresa considerará em quanto tempo terá o retorno de seus investimentos e se a margem de lucro em potencial é satisfatória. Avaliando, assim, comercialmente o produto que só possui densidade detalhes suficientes para este tipo de análise a partir da fase de avaliação de artefatos, no processo de design.

⁴ Segunda etapa normalmente reconhecida no processo de design, onde são geradas e selecionadas várias alternativas de solução para o problema.

Os produtos apresentam um ciclo de vida que segue o paradigma abaixo pontuado (REIS, 2007).

- Produtos têm vida limitada;
- As vendas dos produtos passam por estágios diferentes, com oportunidades e problemas diferentes;
- Os lucros sobem e descem nos diferentes estágios;
- Os produtos necessitam de diferentes estratégias de acordo com cada estágio;

Neste sentido, os produtos apresentam o ciclo de vida conforme apresentado no Quadro 17 e detalhado em seguida.



Quadro 17 – Ciclo de vida do produto

- **Lançamento** – Fase de implementação do produto no mercado, os consumidores estão passando a tomar conhecimento do produto;
- **Crescimento** – Período em que substancial quantidade de consumidores adere ao produto e que se espera conquistar o público-alvo previsto. Esta fase é crucial para o estabelecimento do produto no mercado ou fracasso dos investimentos;
- **Maturidade** – Baixa no crescimento das vendas, uma vez que o público-alvo em potencial foi conquistado. Tende a estabilização das vendas e/ou declínio em função dos esforços da concorrência;
- **Declínio** – Produto atinge sua obsolescência e é substituído pelo concorrente mais inovador.

Visando prolongar o ciclo de vida do produto (e aumentar o retorno dos investimentos), é comum recorrer às estratégias de manutenção na fase do declínio, permitindo estender o período de vendas enquanto que a tendência seria o produto sair do mercado (Quadro 18).



Quadro 18 – Manutenção: estratégia que estende o ciclo de vida do produto

Esta estratégia é amplamente aplicada no mercado de games. O game design atualmente adota o paradigma de projeto baseado em plataformas, que permite uma rápida reconfiguração do produto, modificando-o para estender os recursos do game (como a quantidade de estágios). Ou ainda reconfigurar o game gerando um produto totalmente novo a partir da mesma plataforma (FIGUEIRÔA, 2009; FIGUEIRÔA et al. 2007).

Encontra-se esta estratégia, por exemplo, com o game *Battlefield 2* da *EA Games*. Prestes ao entrar no declínio, a produtora lançou o pacote de expansão *Battlefield 2: Special Forces* que estendia o game original com mais armamentos e combates noturnos (Figura 9), impulsionando novamente as vendas. Diversos outros pacotes de expansão foram lançados sempre que as vendas tendenciavam ao declínio (*Armored Fury* e *Euro Force*), quase dobrando o tempo do produto no mercado. Quando a plataforma *Battlefield 2* em si começou a ser sufocada tecnologicamente pelos concorrentes ao longo dos anos, a *EA Games* então passou a pensar no desenvolvimento de uma nova versão completa do game. Mas só depois de ter estendido bastante seu faturamento pela estratégia de manutenção do ciclo de vida do produto.



Figura 9 – Battlefield 2: Special Forces – Manutenção (FIGUEIRÔA, 2009)

Assim, um dos parâmetros verificados durante a análise do produto é não só sua possibilidade de retorno financeiro ao longo da vida útil no mercado, mas também quais estratégias propostas para manutenção do ciclo de vida deste tipo de artefato.

2.5.4 Avaliação de Artefatos Quanto ao Mercado e a Aleatoriedade

Uma vez que a incerteza no processo de design aumenta os riscos de projeto, a preocupação de redução dela demonstra-se presente na avaliação de artefatos quanto ao mercado (FIGUEIROA et al., 2012; BAXTER, 2011; BURDEK, 2006). Todavia, nenhuma técnica específica é apresentada para redução do acaso neste processo.

Após esta classificação referente ao estado da arte do que consiste na avaliação de artefatos em design, no capítulo seguinte apresentam-se os problemas decorrentes da aleatoriedade presentes prevasivamente em todos os grupos de avaliação aqui destacados:

- seja pelo uso de escalas sem a consciência de que as mesmas podem induzir os resultados da avaliação;
- seja pelo desconhecimento da quantidade de usuários que devem fazer a avaliação para ter maior ciência da opinião do público-alvo sobre o produto;
- seja pela mensuração da incerteza existente no processo de avaliação ou;
- seja em outros fenômenos aleatórios ou falácias que conduzem os resultados das avaliações ao acaso.

3 PROBLEMAS DECORRENTES DA ALEATORIEDADE

Conforme outrora explicitado nesta tese, é possível observar uma grande amplitude de problemas de projeto na fase de avaliação de artefatos. Por limitações de escopo, a presente pesquisa concentrou-se em mapear os problemas decorrentes da aleatoriedade, sob a justificativa da avaliação de artefatos configurar-se como um evento aleatório (ou seja, evento com várias probabilidades de resultado).

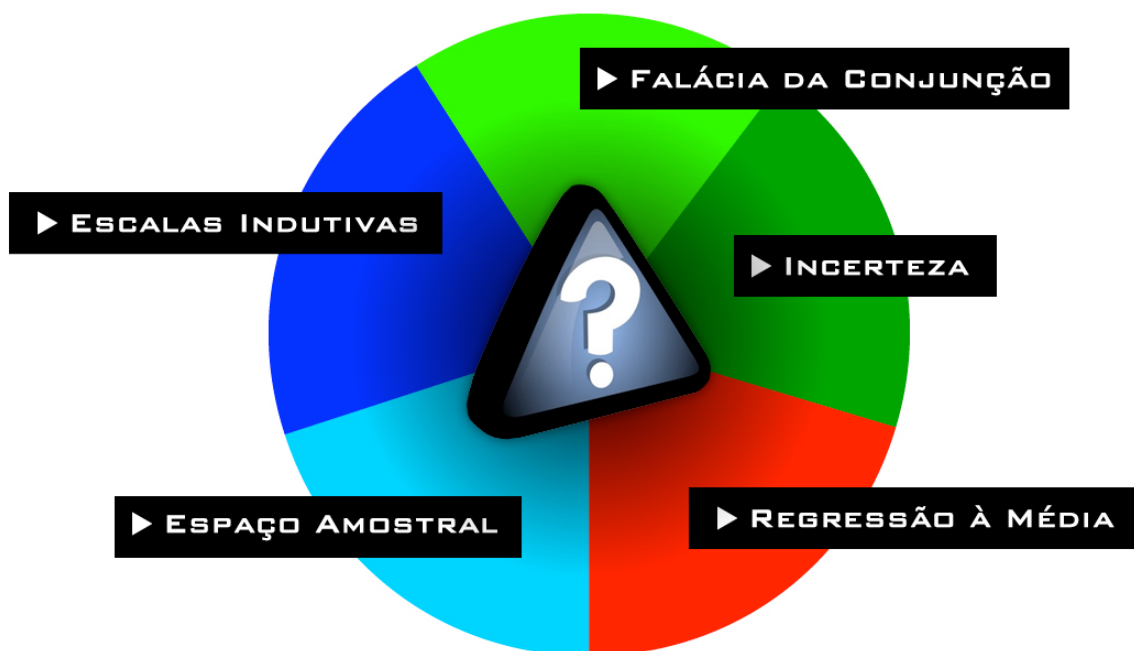
Pervasivamente, nos campos de avaliação de artefatos apresentados no capítulo anterior (Criatividade, estética, experiência com o usuário, usabilidade e mercado), observam-se recorrentemente as características abaixo pontuadas.

- **Precisão da educação:** O recorrente uso de métodos de educação de informações sobre o produto e a opinião do usuário final, objetivando reduzir o acaso na aceitação do produto por parte destes usuários (FIGUEIREDO, 2012; AGUIAR, 2011; BELFORT, 2011; CORREIA, 2010; MELO, 2008; ROCHA 2007; BESEMER, 1998);
- **Redução da incerteza:** O tentativa redução da incerteza sobre o resultado final do projeto pelo aumento de informações sobre o produto e mercado (FIGUEIRÔA et al. 2008, 2012; BAXTER, 2011; CAMPOS, 2005);
- **Quantidade de avaliadores:** A preocupação sobre quantos avaliadores devem participar do processo de projeto com intuito verificar as correções necessárias para potencializar a aceitação do produto (FAULKNER, 2003; PERFETTI & LANDESMAN, 2001; NIELSEN & LANDAUER, 1993; VIRZI, 1992);

Todos estes artifícios possuem um objetivo comum: reduzir a aleatoriedade e aumentar a certeza sobre informações críticas para o sucesso do projeto. Suprimindo assim o impacto do acaso no processo de design.

Neste contexto, o presente capítulo apresenta os problemas identificados na avaliação de artefatos agrupando-os em 5 seções, conforme ilustrado no Quadro 19. Ainda sobre este quadro, as 5 seções do gráfico apresentado estão igualitárias por conveniência estética e não por propoção de conteúdo dos temas. Foram utilizadas publicações formais (científicas) para embasar a existência destes problemas e a solução deles, acrescidas de publicações informais (sem validação científica) com caráter ilustrativo, sem prover qualquer embasamento aos temas abordados. Totalizando 43 publicações distribuídas entre 34

publicações acadêmicas (teses, dissertações, monografias, periódicos, livros de caráter acadêmico e artigos científicos) e 9 relevantes publicações informais (artigos e revistas).



Quadro 19 – Problemas decorrentes da aleatoriedade na avaliação de artefatos

A tese materializa que os 5 temas destacados no Quadro 9 interferem de forma impactante ao desempenho designer na tarefa de avaliar seus artefatos.

- **Fenômeno da regressão à média** (KAHNEMAN & TVERSKY, 1974): onde os eventos regidos pela aleatoriedade tendem a um resultado;
- **Falácia da conjunção** (TVERSKY & KAHNEMAN, 1983): onde problemas cognitivos interferem nas escolhas de resultados prováveis durante uma avaliação;
- **Incerteza nas decisões de projeto** (FIGUEIRÔA et al. 2008): onde a incerteza proveniente de crenças divergentes ou conflitantes pode ampliar a aleatoriedade dos resultados;
- **Má representação do espaço amostral** (FAULKNER, 2003): onde o resultado do design está demasiadamente sujeito ao acaso por não saber a opinião dos usuários finais sobre o produto;
- **Escalas indutivas** (AGUIAR, 2011): onde amplia-se o impacto do acaso no processo de design pela possibilidade da escala de avaliação induzir as opiniões.

A seguir o delineamento de como o designer pode compreender melhor a avaliação de artefatos através de uma maior compreensão dos eventos regidos pela aleatoriedade

através da introdução à aleatoriedade, apresentação de cada um dos problemas pontualmente destacados acima e contextualização dos mesmos no âmbito do design.

Nestes termos, apresenta-se um estado da arte sobre estes problemas, e as atuais tentativas de correção dos mesmos no âmbito do design e em temas correlatos.

3.1 Introdução à Aleatoriedade na Avaliação de Artefatos

Alea jacta est (A sorte está lançada)

JÚLIO CÉSAR

... meu conhecimento venceu o azar ...

(GEROLAMO CARDANO em ORE *apud* MLODINOW, 2009)

Esta seção apresenta o desenvolvimento histórico da aleatoriedade, seu impacto na avaliação de produtos e seu potencial de influenciar o design. No capítulo anterior apresentaram-se diversos critérios considerados relevantes pelos designers durante o projeto de produtos. Criatividade, inovação, beleza, usabilidade, funcionalidade, preço, satisfação: quais destes fatores são de fato determinantes na avaliação de um produto? Ou, como justificar na avaliação de um artefato que o mesmo tende a ser bem sucedido a partir de determinadas correções?

Nesta busca dos designers por justificativas e embasamentos diversos para seus protótipos, corriqueiramente desconsidera-se que o processo de antever o sucesso de um produto envolve uma amplitude de variáveis difíceis de serem controladas e, por vezes, desconhecidas. Sendo estas variáveis responsáveis por tornar este processo aleatório. Ou seja, um evento onde existem várias probabilidades de resultado, ao contrário de um resultado pré-determinado como aconteceria num evento determinístico.

A probabilidade como é conhecida hoje provém da publicação “*O livro dos jogos de azar*” de Gerolamo Cardano, século XVI. O que representa uma pesquisa a frente do seu tempo, onde sequer o sinal matemático de “igual” existia. O autor descreve em sua autobiografia que num dos seus muitos momentos de infortúnio ao longo da vida: “ ...[fui] forçado a apostar novamente nos dados para poder sustentar minha esposa; e com isso, meu conhecimento venceu o azar, e conseguimos comprar comida e viver...” (ORE *apud* MLODINOW, 2009; WYKES, 1969).

A expressão “...meu conhecimento venceu o azar...” denota o uso da ciência na redução do impacto do acaso nos resultados de nossas escolhas, ampliando nossas chances de sucesso em eventos permeados de aleatoriedade.

Esse recurso de redução dos efeitos da aleatoriedade, nas escolhas de projeto, é o foco da presente tese. Devido à abordagem muitas vezes determinística dos eventos aleatórios, os designers estão utilizando-se de justificativas que podem não representar nenhuma significância para o resultado final de sucesso, ou fracasso, do produto na avaliação do mesmo.

O comportamento flutuante do consumidor, o timing de mercado, a infidelidade do protótipo, o surgimento de produtos complementares ou substitutos, dentre tantos outros fatores a serem considerados, levam o projetista a uma tentativa natural de racionalização do processo. O talento, a habilidade e as técnicas são fatores que tem seu papel fundamental nas decisões de projeto: a de ampliadores das chances de êxito. Contudo, conforme explicitado em *cases* a seguir, uma melhor compreensão da aleatoriedade é um fator igualmente importante.

A incompreensão ou desconhecimento das variáveis envolvendo o sucesso de um produto que fez a Hewlett-Packard (HP) e a ATARI avaliarem negativamente projeto de produto proposto por Steve Jobs nos anos 1970. Este produto viria a se chamar cerca de 2 anos mais tarde “*Apple II*”, tornando-se o computador mais bem-sucedido da história, passando 16 anos no mercado (ISAACSON, 2010; MLODINOW, 2009).

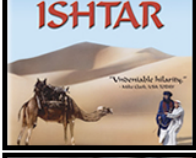
De igual modo, 9 editores avaliaram e rejeitaram o primeiro livro da série “*Harry Potter*”. O que os deixou de fora de uma tiragem 400 milhões de cópias vendidas em 68 idiomas, ao longo de 200 territórios, além dos *royalties* referentes aos R\$ 6 bilhões de dólares em bilheterias de cinema. Este produto tornara a autora britânica, J.K. Rowling, a única em seu país com a fortuna calculada em mais de \$ 1 bilhão de dólares, além do ator que interpreta o protagonista da série na versão cinematográfica, Daniel Radcliffe, como o mais rico do seu país com idade menor que 30 anos (REVISTA VEJA, 2011).

Como os avaliadores não perceberam, em sucessivas rejeições, o sucesso proeminente desses produtos? Por outro lado, apostando como fatores determinantes de sucesso para um filme ter um elenco de grande reputação ante o público, investimentos pesados e uma excelente produção, Hollywood apresentou o filme “*Ishtar*”. Esta produção, com atores

relativamente caros como Dustin Hoffman e Warren Beatty, recebeu \$ 55 milhões de dólares em investimentos, retornando a contra-intuitiva bilheteria de \$ 14 milhões de dólares. Outro fiasco desta indústria, que demonstra o poder da má compreensão da aleatoriedade na avaliação de produtos, é materializado pelo filme “O Último Grande Héroi”, com o protagonista Arnold Schwarzenegger. Este produto recebeu \$ 85 milhões de dólares em investimentos e obteve a bilheteria de \$ 50 milhões de dólares.

Não obstante, mediante um orçamento inicial de \$ 13 milhões de dólares, a Universal Studios rejeitou um projeto que fora aceito pela 20th Century Fox. Com o cachê inicial de apenas \$ 200 mil dólares, somado aos posteriores royalties, a indústria permitiu ao jovem diretor George Lucas produzir “Star Wars: A New Hope”, que rendeu \$ 460 milhões de dólares (Ver Quadro 20).

CASES DE ALEATORIEDADE NA AVALIAÇÃO DE PRODUTOS

	HARRY POTTER (LIVRO) PRODUTO	Avaliado negativamente por 9 editoras AVALIAÇÃO -	400 milhões de cópias vendidas RESULTADO +
	APPLE II PRODUTO	Avaliado negativamente por 2 empresas (HP e ATARI) AVALIAÇÃO -	6 milhões de unidades vendidas RESULTADO +
	STAR WARS: A NEW HOPE PRODUTO	Avaliado negativamente pela Universal Studios AVALIAÇÃO -	Investimento: US\$ 13 milhões Bilheteria: US\$ 460 milhões RESULTADO +
	ISHTAR PRODUTO	Avaliado positivamente por investimentos e grande elenco AVALIAÇÃO +	Investimento: US\$ 55 milhões Bilheteria: US\$ 14 milhões RESULTADO -
	O ÚLTIMO GRANDE HÉROI PRODUTO	Avaliado positivamente por investimentos e grande elenco AVALIAÇÃO +	Investimento: US\$ 80 milhões Bilheteria: US\$ 50 milhões RESULTADO -

Quadro 20 – Exemplos de má interpretação da aleatoriedade

Estes números e resultados contra-intuitivos tem chamado a atenção de muitos pesquisadores, como perguntas como: “Como avaliar o potencial de sucesso de um produto?” (MLODINOW, 2009). Oportunamente, esta racionalização ante múltiplas variáveis está presente numa relevante amplitude das decisões humanas abordada em diversos segmentos científicos (KAHNEMAN & TVERSKY, 1974). Isso significa a densa contribuição

de vários domínios científicos que estudam as decisões ante a aleatoriedade. Permitindo assim uma melhor compreensão deste corpo de conhecimento que, apesar de as primeiras abordagens nos tornarem para Século XVI, com as obras de Gerolamo Cardano, ainda hoje carecerem de muitos desdobramentos.

Adiante, explicitam-se fenômenos relacionados à aleatoriedade e sua contextualização no design, conforme pontualmente destacados no início do presente capítulo.

3.2 O Fenômeno da Regressão à Média

Esta seção da tese apresenta uma introdução do fenômeno da regressão à média através das pesquisas e relatos de Kahneman & Tversky (1974). Apresentando em seguida um case em design da empresa Apple, permitindo ilustrar respectivamente o comportamento deste fenômeno em eventos aleatórios e seu impacto no âmbito do design.

Uma vez que o objetivo central da avaliação de artefatos é prever a probabilidade de sucesso de um produto no mercado, as pesquisas de Kahneman & Tversky (1974) demonstram-se substancialmente relevantes ao tema. Por estas pesquisas, sobre a influência da aleatoriedade no resultado de eventos probabilísticos, o Prof. Kahneman fora contemplado com o prêmio Nobel de Economia.

Ilustrando as descobertas do supracitado autor, é conveniente relatar o caso ocorrido durante os anos 1960, quando o mesmo era docente da Universidade Hebraica, ministrando aulas de psicologia a um grupo de instrutores de aviação da IAF (Israeli Air Force). Objetivando aperfeiçoar a didática de voo dos israelenses promovia-se o pensamento de Skinner (1953) sobre a eficiência de elogiar bons comportamentos ao invés de punir os maus, no sentido de que esta metodologia seria a melhor maneira possível de contribuir para um mais eficiente aprendizado por parte dos alunos da aviação militar.

Contradizendo o conceito de Skinner (1953) – apesar do mesmo ser amplamente referenciado pela eficiência desta metodologia – os instrutores israelenses conseguiram externar em sua aparentemente genuína experiência que punir o mau desempenho dos alunos, ao invés de elogiar quando os mesmos executavam corretamente as manobras, era a forma que produzia melhores resultados no aprendizado. De acordo com eles, após reclamar de um erro de pilotagem do aluno, o mesmo tendia a melhorar na manobra seguinte, conforme o relato de Kahneman abaixo.

Quando eu terminei meu entusiasmado discurso, um dos mais experientes instrutores na sala levantou sua mão e fez seu próprio discurso curto, que começou por dizer que reforço positivo pode ser bom para pássaros, mas negou que seria o ideal para cadetes aviadores. Ele disse, “Em muitas ocasiões eu elogiei cadetes por uma boa execução de uma manobra acrobática, e geralmente quando eles tentavam novamente, eles faziam pior. Por outro lado, eu frequentemente grito com cadetes quando eles executam mal uma manobra, e geralmente eles fazem melhor na próxima vez. Então por favor não nos diga que reforço positivo funciona e punição não, porque oposto é o que acontece”.

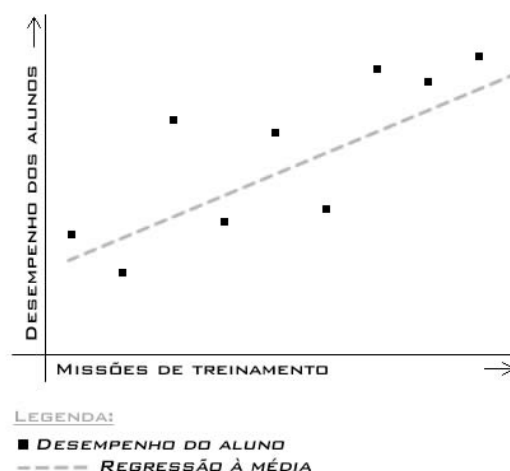
(KAHNEMAN, 2002, tradução livre)

Haja vista ali se formarem os que são considerados por muitos os melhores pilotos militares do mundo, tornava a contra-argumentação dos métodos dos instrutores uma tarefa difícil (Figura 10).



Figura 10 – Jatos de treinamento da IAF – Israeli Air Force (Fonte: www.iaf.org.il)

Após algumas abordagens, Kahneman (2002) identificou o que ele chamou de “... a mais gratificante *heureka* da sua carreira”. Com pesquisa publicada na *Psychological Review*, o referido autor descreve o que estava ocorrendo com os cadetes israelenses: o fenômeno da regressão à média (KAHNEMAN & TVERSKY, 1973). Considerando que múltiplos fatores contribuem no aprendizado, muitos destes representam variáveis difíceis de controlar ou mesmo desconhecidas. Sob a mecânica de que os eventos afetados pela aleatoriedade tendem a efetuar oscilações de resultados em torno de uma média, ilustra-se no Quadro 21 a oscilação do desempenho dos alunos durante as missões de treinamento.



Quadro 21 – Mera ilustração do desempenho dos alunos em Regressão à Média

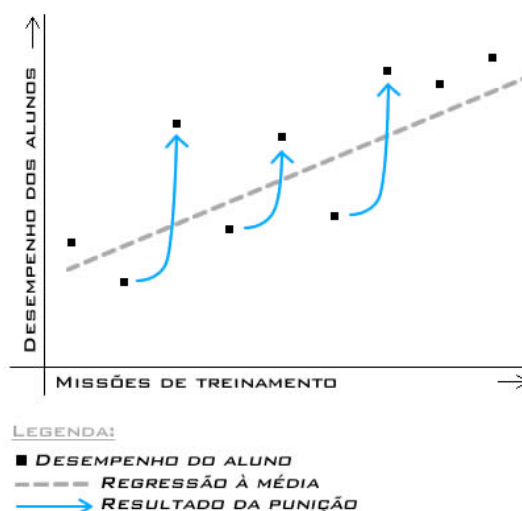
Convém destacar que apesar do Quadro 21 ser ilustrativo, de fato o fenômeno da regressão à média apresenta-se de forma linear, conforme o proposto por Galton (1886).

Exemplificando, os fatores que também poderiam estar contribuindo para cada manobra (com diferentes ponderações de contribuição) poderíamos listar: a manutenção da aeronave em uso; as condições meteorológicas relativas ao vento ou turbulência; o estado psicológico, emocional ou físico do aluno; a diferença de performance da aeronave de acordo com peso da mesma, decorrente da quantidade de combustível; as colocações dos instrutores de voos anteriores; dentre diversos outros.

Regressão é inevitável em manobras de voo porque o desempenho não é perfeitamente confiável e progresso entre sucessivas manobras é lento. Assim, os pilotos que se saíram excepcionalmente bem numa tentativa são susceptíveis a deteriorar a seguinte, independentemente da reação dos instrutores para o sucesso inicial.

(KAHNEMAN & TVERSKY, 1973, pg. 251, tradução livre)

Aos resultados de desempenho acima da média eram atribuídos pelos instrutores pelas punições quando o aluno executava uma manobra sem grande aproveitamento, conforme ilustra-se no Quadro 22 abaixo.



Quadro 22 – Mera ilustração da intuição dos instrutores de voo israelenses

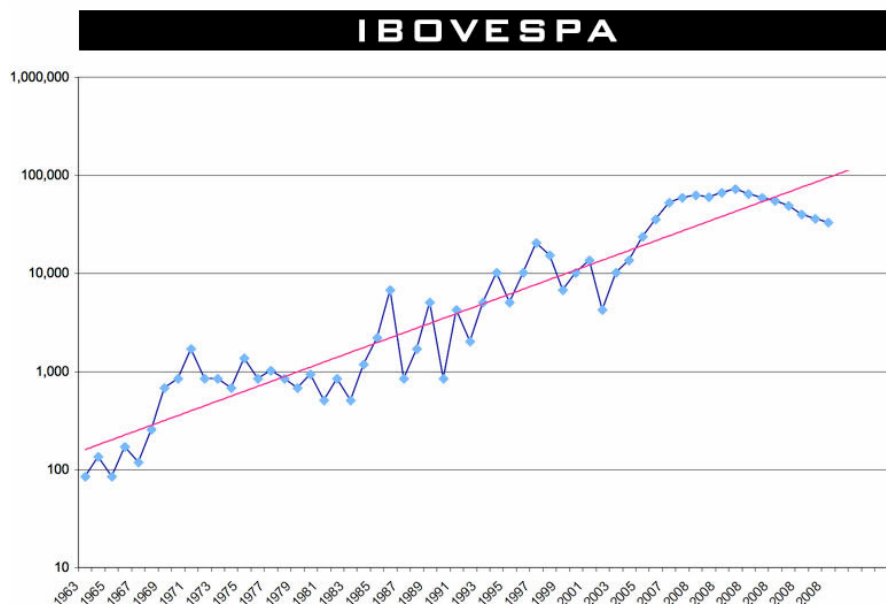
Portanto, de acordo com estes autores, o fenômeno da regressão à média era o principal atuante na oscilação de desempenho dos alunos e não as punições dos instrutores, uma vez que as mesmas consistiam apenas uma diante de uma multiplicidade de variáveis contribuintes no resultado do evento. É possível deduzir a consistência deste fenômeno ao observar sua recorrência em eventos envolvendo aleatoriedade, permeados nos resultados apontados por diversas outras publicações, muitas delas da área médica (GALTON, 1886).

As primeiras pesquisas sobre este tema se iniciaram na medicina, quando Galton(1886) observou uma regressão à média na altura dos pacientes em relação ao tamanho da população, e não somente dos pais. O levantamento de dados estatísticos levou este pesquisador a encontrar um coeficiente (uma constante) nestes padrões e inclusive descrever a média encontrada matematicamente.

De relevante, passamos a ter, nestes estudos, a possibilidade de entender que os eventos permeados de aleatoriedade não são regidos por forças totalmente randômicas, mas que os resultados tendem a uma média. Esta tendência é um “norte”, um direcionamento para onde os resultados buscarão concentrar-se dinamicamente.

Enriquecendo as ilustrações e a compreensão do fenômeno, verificam-se os investimentos na bolsa de valores, característicos pela aleatoriedade. Ao comprar uma ação não é possível determinar o quanto ela irá valorizar (ou desvalorizar) num dado espaço de tempo. Todavia, os investidores utilizam todas as estratégias disponíveis para antecipar-se às flutuações da bolsa e, com isso, multiplicar seus rendimentos. Um dos conceitos considerados nestas estratégias é o de que o mercado de ações tende a apresentar

corriqueiramente casos de regressão à média, conforme demonstrado por pesquisas empíricas recentes (BUSH & CANNING, 2006) (Quadro 23).



Quadro 23 – Aleatoriedade da bolsa (Adaptado de KANITZ, 2008)

Nas avaliações de produtos em design, devido às múltiplas probabilidades consideradas como sucesso ou fracasso, os eventos carecem de um olhar pela perspectiva da aleatoriedade. Por exemplo, nos dias atuais a empresa Apple, fundada por Steve Jobs, é amplamente conhecida pelo sucesso dos seus produtos inovadores, como o *iPad* e o *iPhone* (Figura 11).



Figura 11 – Ipad da Apple (Fonte: apple.com.br)

Mas ao verificar o histórico da Apple é possível visualizar o quanto o fenômeno da regressão a média se fez presente. Fazendo com que de forma totalmente imparcial às justificativas de avaliação dos produtos feitas por Jobs, os resultados dos mesmos no mercado seguissem sucessivas oscilações de sucessos e fracassos. Quem contempla hoje o sucesso do *Ipad*, que só no ano de lançamento – 2010 – rendeu \$ 9,5 bilhões em lucros para a Apple, muitas vezes não lembra-se do seu antecessor, o *Newton* (Figura 12). Um produto repleto de inovações tecnológicas, em proporções próximas as do *iPad* para seus dias, mas que fracassou no mercado.



Figura 12 – Message Pad da Apple – “Newton” (Foto: GettyImages)

Outro fracasso marcante para a empresa foi o *Apple Hockey Mouse*, que fora avaliado por Steve Jobs – o mesmo que criou o *iPad* – como “o melhor mouse já criado”. Entretanto, sua rápida remoção das prateleiras e as avaliações da crítica diziam o contrário, com expressões como “design terrível” e “pesadelo da ergonomia”. Seu desenho redondo provocava confusão ao usuário, só sendo possível entender a direção do dispositivo ao movimentar o mesmo e ter o *feedback* do ponteiro do mouse movendo-se na tela. (IDG NOW, 2011)

E assim, mesmo com designers de reconhecimento no mercado, que ofereciam produtos inovadores e com grande apelo ao consumo, a Apple se viu regida – também – pela aleatoriedade conforme exposto no Quadro 24. A empresa flutua entre grandes sucessos e fracassos, mas mesmo assim chegou a ser considerada a empresa mais admirada da América do Norte em várias categorias: inovação, gerenciamento e qualidade dos produtos/serviços (CNNMONEY, 2008). Ou seja, o desempenho da empresa considerando

como ponto de partida a sua fundação até onde chegou a receber essas premiações – entre os sucessos e fracassos – denota uma média crescente (Quadros 24 e 25).

Produto (Nome)	Lançamento (Ano)	Tempo no mercado (Anos)	Resultado (Considerado)
APPLE I	1976	1	FRACASSO
APPLE II	1977	16	SUCESSO
APPLE III	1980	4	FRACASSO
Lisa	1983	3	FRACASSO
PowerBook	1991	16	SUCESSO
Newton	1993	6	FRACASSO
Macintosh TV	1993	1	FRACASSO
Quicktake	1994	3	FRACASSO
Pippin	1996	2	FRACASSO
TAM	1997	1	FRACASSO
Apple USB Mouse	1998	2	FRACASSO
iMac	1998	14(*)	SUCESSO
Mac OS X	1998	14(*)	SUCESSO
Power Mac G4 Cube	2000	1	FRACASSO
iPod	2001	11(*)	SUCESSO
iPod Hi-Fi	2006	1	FRACASSO
iPhone	2007	5(*)	SUCESSO
iPad	2010	2 (*)	SUCESSO
(*) produtos em uso nos presentes dias			

Quadro 24 – Produtos Apple (G1, 2011; IG SÃO PAULO, 2011; IDGNOW, 2011)



Quadro 25 – Ações da Apple mostram crescimento (YAHOO! FINANCES, 2012)

Segundo o autor Issacson (2011), estas oscilações de resultados da Apple fez com que a diretoria tenta-se resolver o problema, racionalizando o processo e procurando culpados. Neste âmbito de busca por soluções, a diretoria acabou demitindo o designer e fundador da empresa, Steven Jobs, em 1985. De acordo com a crítica de Mlodinow (2009), esta metida ante o evento visivelmente impactado pela aleatoriedade foi equivocada, assim como os gritos dos instrutores israelenses no outro caso apresentado nesta seção. Ou seja, o resultado das vendas de novos produtos estava oscilando aleatoriamente sobre regressão a uma média, e no momento da demissão de Jobs esta oscilação apenas estava em declínio, assumindo posteriormente a natural tendência de ascender a média.

Segundo o físico Mlodinow (2009) é esta mesma incompreensão que faz um técnico de beisebol ser demitido após alguns fracassos do time (mesmo anteriormente tendo acumulado algumas vitórias), ou ainda o executivo de Hollywood ser dispensado após o lançamento de alguns filmes com baixa bilheteria (mesmo este executivo possuindo um histórico de bons filmes). Após estas demissões, com os novos contratados e resultados crescentes (resultantes da regressão à média), as empresas são induzidas a acreditar na falácia de que suas ações diretas reverteram os resultados, quanto que foram importantes apenas como mais um dos fatores influenciantes. Até que os novos contratados, que elevaram os resultados a níveis considerados de sucesso, começam a declinar seus desempenhos e regressar a média, passando para um resultado considerado de fracasso.

O que esta corrente de pensamento levantada por físicos, médicos, economistas e psicólogos defende é que ao analisar a oscilação de resultados em torno de uma média – seja de vendas computadores, bilheterias de filmes ou placares esportivos – deve-se levar em consideração que os resultados tiveram, talvez, tanta influência dos fenômenos da aleatoriedade quanto suas ações diretas. Sendo todos estes componentes contribuintes, mas não determinantes dos resultados. Associar demasiada ponderação as ações diretas de avaliação, descreditando os efeitos do acaso, levam as avaliações de qualquer área a surpresas, por vezes, catastróficas (MLODINOW, 2009; GALTON 1886).

Concluindo, considerando os eventos acima descritos e as pesquisas de Galton (1886), a avaliação de um produto deve relevar o histórico de desempenho do mesmo. Uma vez que os eventos regidos pela aleatoriedade tendem fortemente a uma média de resultados, o designer poderá ter mais informações que antevêm a tendência referente ao produto específico em projeto. Desse modo, a regressão à média, por fornecer um direcionamento

mesmo que impreciso do resultado provável, representa uma solução informacional para um que seria considerado, antes, um evento totalmente randômico.

3.3 A Falácia da Conjunção

... em situações que envolvam o acaso, nossos processos cerebrais costumam ser gravemente deficientes.

(MLODINOW, 2009, pg. 7)

Partindo do princípio de que o processo de avaliação de um protótipo envolve decisões projetuais, pesquisas recentes apontam para problemas cognitivos que interferem diretamente nessas escolhas. Dentre estes problemas, destaca-se a falácia da conjunção. Fenômeno relacionado aos processos cerebrais que induz o avaliador a tomar parte de escolhas de projeto que são diretamente opostas às alternativas matematicamente mais prováveis. Fazendo, assim, com que avaliadores de projeto tenham a natural propensão a errar.

As pesquisas de Tversky & Kahneman (1983) elucidam este fenômeno. Um grupo de 88 estudantes da UBC (University of British Columbia) foram convidados a participar de um experimento onde sua tarefa era classificar numa escala de 1 a 8 as alternativas apresentadas diante dos 2 briefings. Sendo que a classificação “1” deveria se atribuída à alternativa mais provável e “8” à mais improvável. Nos Quadros 26 e 27 abaixo encontram-se estes briefings, que possuem os personagens fictícios Bill e Linda, bem como as alternativas daquilo que seria mais provável referente aos mesmos.

Briefing 1

Bill tem 34 anos de idade. Ele é inteligente, mas sem imaginação. Compulsivo e geralmente apático. Era bom em matemática na escola, mas fraco nas matérias sociais e humanas.

AFIRMAÇÃO

Bill é um físico que joga poker por hobby

Bill é um arquiteto

Bill é um contador

Bill toca jazz por hobby

Bill é um repórter

Bill é um contador que toca jazz por hobby

Bill escala montanhas por hobby

Quadro 26 – Descrição de Bill. Adaptado de TVERSKY & KAHNEMAN(1983)

Briefing 2

Linda tem 31 anos de idade, é solteira, franca e muito inteligente. Coursou filosofia na universidade. Enquanto era estudante, ela se preocupava profundamente com questões de discriminação e justiça social, e participou de protestos contra armas nucleares.

AFIRMAÇÃO

Linda é professora da escola elementar.

Linda trabalha em uma livraria e tem aulas de ioga.

Linda participa ativamente do movimento feminista.

Linda é assistente social psiquiátrica

Linda é membro da Liga de Mulheres Eleitoreiras (League of Women Voters).

Linda é caixa de banco.

Linda é vendedora de seguros.

Linda é caixa de banco e participa ativamente do movimento feminista.

Quadro 27 – Descrição de Linda. Adaptado de TVERSKY & KAHNEMAN(1983)

Nas descrições dos quadros acima, Bill foi intencionalmente descrito para representar um contador e não alguém que toca jazz por hobby. Assim como Linda fora descrita para representar uma militante do movimento feminista e não como uma caixa de banco.

Os pesquisadores afirmam que suas expectativas foram atendidas, pois os participantes responderam apontando como intuitivamente mais prováveis respostas que em alguns casos eram matematicamente as mais improváveis. O Quadro 28 abaixo mostra especificamente a ordem de algumas afirmações relevantes para a pesquisa, conforme classificadas pelos participantes. Ou seja, todas as alternativas foram classificadas pelos participantes, mas o Quadro 28 mostra apenas algumas das opções para facilitar a visualização da falácia de probabilidade em estudo. Esta foi a ordem apontada por 87% dos participantes para Bill e 85% dos participantes para Linda.

Bill
1. Bill é um contador (A)
2. Bill é um contador que toca jazz por hobby (A&B)
3. Bill toca jazz por hobby (B)
Linda
1. Linda participa do movimento feminista (C)
2. Linda é caixa de banco e participa do movimento feminista (C&D)
3. Linda é caixa de banco (D)

Quadro 28 – Classificação em ordem decrescente.

As respostas da maioria dos participantes afirmando intuitivamente que é mais provável Bill ser um contador que toca jazz por hobby (A&B) do que Bill simplesmente tocar jazz por hobby(B), bem como Linda ser caixa de banco que e militante feminista (C&D) ter maior chance do que ela ser somente caixa de banco (D), violam a primeira lei da probabilidade. De acordo com Freund (2006): “a probabilidade de ocorrência simultânea de dois eventos independentes é simplesmente o produto de suas probabilidades respectivas”. Em outras palavras, “a probabilidade de que dois eventos ocorram nunca pode ser maior do que a probabilidade de que cada evento ocorra individualmente” (MLODINOW, 2009). Esta lei é conhecida como *regra da conjunção*.

Na pesquisa de Tversky & Kahneman (1983) eles ainda aprofundaram o estudo, dessa vez tentando induzir os participantes para que eles não violassem a “regra da conjunção”. A descrição de Linda foi novamente apresentada, mas agora a um grupo de 142 universitários. Desta vez fora-lhes perguntado – nessa ordem – qual das alternativas era mais provável:

- Linda é caixa de banco. (D)
- Linda é caixa de banco e militante do movimento feminista. (C&D)

Apesar de inverter a ordem das alternativas na apresentação aos participantes, a pesquisa reporta que: “a manipulação não teve efeito”. No final, 85% dos entrevistados diziam “C&D” era mais provável que “D”.

Os dois cientistas estenderam-se nesta investigação, apresentando um briefing de diagnóstico médico a um grupo de 66 residentes. Mais uma vez a resposta matematicamente menos provável fora atribuída como a intuitivamente mais provável.

Diante de um diagnóstico de embolia pulmonar, 91% dos médicos afirmaram a combinação de um sintoma raro com um sintoma comum, em vez de somente um sintoma comum ou um raro. Os sintomas de dispneia (falta de ar) e hemiparesia (paralisia parcial) combinados foram atribuídos como mais prováveis do que simplesmente hemiparesia.

E assim se repetiu o que estes pesquisadores denominaram como “falácia da conjunção” em sucessivas investigações quanto a pesquisas políticas, decisões de risco e vereditos jurídicos. Em tempo, os advogados experimentam esta mesma falácia nos tribunais. Os resultados obtidos por Fox & Birke (2002) apontam que os juízes consideram intuitivamente como mais prováveis as contingências descritas com maior combinação de detalhes. Avaliações de mercado realizadas em avaliações de imóveis indicaram a mesma falácia (MARTINS, 2011).

Se os detalhes que recebemos se adequarem à imagem mental que temos de alguma coisa, então, quanto maior o número de detalhes numa situação, mais real ela parecerá, e, portanto, consideraremos que será mais provável. [...] O que é mais provável: que um réu, depois de encontrar um corpo, deixe a cena do crime, ou que um réu, depois de encontrar um corpo, deixe a cena do crime porque teme ser acusado pelo macabro assassinato? [...] É mais provável que uma empresa aumente suas vendas no ano que vem ou que aumente as vendas no ano que vem porque a economia geral passará por um bom ano? Em todos os casos, embora a segunda opção seja menos provável que a primeira, pode parecer mais provável.
(MLODINOW, 2009, p. 33)

Este fenômeno, que leva nosso sistema cognitivo a considerar com maiores chances de sucesso alternativas que matematicamente tendem ao fracasso, impacta sobremaneira no design de produto, em especial, na avaliação de protótipos.

Esses vieses na avaliação de eventos compostos são particularmente importantes para o planejamento, visto que a realização bem sucedida de um negócio, tal como **o desenvolvimento de um novo produto, possui caráter conjuntivo: para ter sucesso, cada série de eventos deve, necessariamente, ocorrer**; mesmo quando a probabilidade de cada evento ocorrer for alta, a probabilidade geral de haver sucesso pode ser baixa, em função do grande número de eventos; observou-se, assim, que a tendência geral a **superestimar a probabilidade de eventos conjuntivos leva a um**

otimismo infundado na avaliação de probabilidade de um plano dar certo ou de um projeto complementar-se no prazo.
(FERREIRA, 2008, pg 165, marcação própria)

Nos estudos de caso, a presente tese investiga os impactos da falácia da conjunção no design, solucionando este problema ao gerar conhecimentos que antecipem os designers às armadilhas cognitivas durante a avaliação de seus artefatos.

3.4 Incerteza nas Decisões de Projeto (*Lateo*)

Considerando que incerteza e aleatoriedade são temáticas diretamente ligadas, esta seção da tese propõe contribuir na redução da aleatoriedade no processo de design ao apresentar os avanços no modelamento da incerteza na fase de avaliação de artefatos. Para tal, introduz-se as pesquisas que se desdobraram a partir da *Teoria das Evidências* de Dempster-Shafer, o recurso matemático que modela a incerteza (o *Lateo*), bem como sua aplicação no design (FIGUEIROA et al., 2008).

Na fase de avaliação, o designer busca verificar a probabilidade de sucesso de um único produto isoladamente, ou de confrontar a probabilidade de sucesso de alguns protótipos propostos, comparativamente. Escolhendo, entre algumas poucas opções, qual projeto terá maiores condições de obter sucesso e, assim, prosseguir com este para a fase de apresentação.

Para melhor mensurar as probabilidades de resultado nas decisões de projeto, a *Teoria das Evidências* vem sendo amplamente aplicada em diversas áreas, incluindo robótica, inteligência artificial, gerência de risco, processamento de imagem, dentre outros (FERSON et al., 2003). Na avaliação de artefatos, esta apresenta-se como uma alternativa viável.

Por razões óbvias, o emprego dos números no processo de design transmite ao projetista a expressão de valores que doutra maneira seriam intangíveis. Desde as raízes do design sendo abordado por metodologia científica na HfG de Ulm que buscam-se mecanismos matemáticos para embasar projetos. Satisfazendo, desta forma, a exigência da indústria por adequação aos seus processos sistêmicos e rigidamente justificados (BURDEK, 2006). Sendo a *Teoria das Evidências* de Shafer (1976), uma alternativa adotada como viável para justificativas projetuais baseadas em probabilidade, como as que podemos encontrar durante a avaliação de artefatos (FERSON et al., 2003).

Todavia, Campos (2005) identificou falhas crônicas na *Teoria das Evidências* que restringem seu uso, incluindo no design. Resultados contra-intuitivos (ou seja, decisões que parecem ilógicas) surgem quando as opiniões (avaliações) combinadas envolvem alto grau de conflito ou quando as evidências possuem concentração de crenças em hipóteses distintas, conforme melhor detalhado adiante. (CAMPOS & CAVALCANTE, 2003).

Através da pesquisa de Campos (2005) foi criado um mecanismo matemático chamado *Lateo*, que é proposto por este autor para corrigir os problemas contra-intuitivos da teoria de Dempster-Shafer através de um modelamento numérico da incerteza nas decisões projetuais.

A publicação de Figueirôa et al. (2008) demonstrou como a *Teoria das Evidências* pode provocar resultados contra-intuitivos no processo de design, em especial na avaliação de artefatos, e como *Lateo* pode auxiliar o designer a justificar suas escolhas de projeto através do modelamento numérico da incerteza.

Para melhor compreender o problema envolvendo esta teoria no processo de design, é oportuno expor um exemplo ilustrativo. Consideremos que uma equipe de designers está projetando um novo tipo de interface digital, e deseja avaliar quais das 3 propostas de tela é mais adequada, submetendo as mesmas à avaliação de 2 especialistas. Denominaremos ilustrativamente estas telas de LCD (crystal líquido), HOLO (holografia) e 3D (tridimensional) nos quadros abaixo. O Quadro 29 representa as massas de crença atribuídas pelos especialistas para cada uma das opções de tela, onde 1,0 representa 100%.

Primeiro especialista	Segundo especialista
LCD = 0,985	LCD = 0
HOLO = 0,015	HOLO = 0,01
3D = 0	3D = 0,99

Quadro 29 – Crenças percentuais das avaliações dos especialistas

O Quadro 30 apresenta a combinação das massas de crenças pela regra de Dempster, onde basicamente multiplicam-se os valores correspondentes a cada opção, gerando uma nova matriz com os valores combinados. Estes valores resultantes estão dispostos ortogonalmente.

Crenças	LCD = 0,985	HOLO = 0,015	3D = 0
LCD = 0	0	-	-
HOLO = 0,01	-	0,00015	-
3D = 0,99	-	-	0
	$\Sigma = 0$	$\Sigma = 0,00015$	$\Sigma = 0$

Quadro 30 – Combinação de opiniões pela regra de Dempster

Após a normalização da matriz do Quadro 30 para uma escala centígrada, obtêm-se os valores de crença para cada uma das opções de tela, exposto no Quadro 31 abaixo.

Opção	Crença
LCD	0%
HOLO	100%
3D	0%

Quadro 31 – Normalização pela regra de Dempster

Conforme evidente no Quadro 31, a alternativa que recebeu menor crença de sucesso pelos avaliadores (HOLO) foi exatamente a que a *Teoria das Evidências* externou como melhor avaliada pelos especialistas, com considerável disparidade. Afinal, como poderia a tela holográfica (HOLO) ser a melhor avaliada se os especialistas atribuíram pouca ou nenhuma massa de crença a ela? Bem como as outras opções que receberam considerável massa de crença, serem retornadas com 0% como avaliação final. Esse resultado que procede contra a intuição acontece na regra de Dempster sempre que posicionam-se massas de crenças extremamente conflitantes, ou divergentes.

Esta falha pode levar robôs guiados pela *Teoria das Evidências* a ações equivocadas, bem como designers que tentem embasar suas escolhas de projeto a optar por alternativas que conflitam diretamente com o que seria uma opção plausível.

Isto acontece porque, pela regra de Dempster, as crenças atribuídas às alternativas disjuntas são distribuídas para crenças conjuntas, escondendo a incerteza subjetiva contida nas opções. Aplicando o *Lateo*, que é o resultado do somatório ortogonal das massas de crença dividido por $(1 + \log(1/k))$, conforme proposto por Campos (2005), obtemos o resultado exposto no Quadro 30.

Opção	Crença
LCD	0%
HOLO	3%
3D	0%
Lateo	97%

Quadro 32 – Incerteza das alternativas mensurada pelo Lateo

Agora, o conflito das avaliações dos especialistas – que intuitivamente elevaria a incerteza do designer – é devidamente expressa com um relevante valor percentual atribuído ao *Lateo*. Sinalizando, neste caso, que o designer deveria buscar integrar à equação as massas de crenças de mais especialistas a fim de adquirir uma avaliação final com menor nível de incerteza subjetiva envolvida.

O valor do *Lateo* obtido no Quadro 32 foi resultado a aplicação da função massa disposto na Figura 13 abaixo.

$$LATEO \quad \Lambda = m_1 \psi m_2(\Theta) = (\chi \cdot m_1 \oplus m_2(\Theta)) + 1 - \sum_{\substack{A \subset \Theta \\ A \neq \emptyset}} m_1 \psi m_2(A)$$

Figura 13 – Função massa que determina o Lateo (CAMPOS, 2005)

A formação do designer não inclui, normalmente, familiaridade com notações matemáticas desde nível. Impedindo assim que os mesmos tenham acesso a esta ferramenta metodológica que reduz o impacto negativo da aleatoriedade pelo monitoramento da incerteza envolvida. Em outras palavras, não obstante o problema da *Teoria das Evidências*, os designers também não tem acesso à solução pela complexidade matemática envolvida no Lateo, que normalmente envolve conhecimentos e notações não-familiares ao mesmo. Neste sentido, desenvolveu-se a ferramenta e-Lateo. Um software que abstrai o designer da complexidade matemática para cálculo da incerteza envolvida no processo (MONTEIRO, 2009).

Neste sentido, com base nos autores destacados nesta revisão bibliográfica, apresenta-se como modelar a incerteza nas decisões de projeto em design e, com isso, reduzir o impacto da aleatoriedade nesta atividade.

3.5 Representação do Espaço Amostral

Esta seção da tese apresenta discussões sobre quantos avaliadores devem participar de uma pesquisa, objetivando melhor conhecer a opinião dos usuários sobre um produto (FIGUEIRÔA et al., 2012). Para tal, abordam-se as discussões já em curso sobre o método de Nielsen (2006) para avaliação de usabilidade de websites, que engloba a quantidade de usuários necessários para este processo. Neste ínterim, apresenta-se ainda uma proposta de como este método pode ser validado cientificamente, utilizando-se do cálculo de espaço amostral de Triola (1999), de acordo com o descrito na pesquisa de Figueirôa et al. (2012).

3.5.1 Introdução ao Problema da Quantidade de Avaliadores em Usabilidade

Uma das maiores controvérsias, atualmente, sobre a temática abordada nesta tese encontra-se nas discussões sobre quem ou quantas pessoas deveriam avaliar o artefato, visando conhecer sua probabilidade de sucesso, sob diversos aspectos.

Delimitando a discussão ao design de artefatos digitais, nesse caso websites, este assunto encontra-se em franco desenvolvimento principalmente nas avaliações de usabilidade deste tipo de produto. O problema é claro: dado que um produto como um website é utilizado, às vezes, por milhões de usuários demograficamente muito diversificados, qual seria uma quantidade aceitável de usuários avaliando este produto de modo que representassem a opinião destes milhões? Em outras palavras, um grupo de poucos usuários cuja sua avaliação do protótipo seja coerente com a avaliação dos milhões de usuários finais do produto após lançado no mercado. Afinal, se houver discrepância relevante entre a opinião destes 2 grupos, avaliação de usabilidade do website teria sido em vão e o resultado do produto estaria demasiadamente sujeito ao acaso.

Este conceito é abordado na estatística. O grupo que concerne todos os usuários finais do produto chama-se “espaço amostral” e o que define o pequeno grupo de usuários testando o protótipo denomina-se “amostra”.

A amostra é uma representação nivelada de todo o espaço amostral, ela é uma coleta simplificada que reflete características representando o todo. Por exemplo, para se mensurar um dado estatístico referente à média salarial de uma dada população – vamos supor de designers em território brasileiro – seria inviável coletar os números referentes aos salários de todos os designers deste grupo, espalhados por todas as partes do país sem

deixar faltar nenhum sequer. Para, só então, calcular a média aritmética destes salários e, com isso, saber exatamente a média salarial.

Assim como este procedimento, incontáveis outros processos estatísticos seriam inviáveis coletando dados de todo o espaço amostral. Para isto aplica-se o procedimento de coleta de uma pequena amostra calculada que represente, mediante níveis de confiança variáveis, o todo. É viável obter amostras relativamente pequenas com níveis de confiança considerados aceitáveis, dado o custo-benefício do método: coletar relativamente poucos dados – ao invés de todos os dados – e assim obter resultados próximos do que seria coletando-se todo o espaço amostral.

Entretanto, a principal referência em testes de usabilidade, Nielsen (2006; 1993), afirma que os 5 primeiros usuários são suficientes para encontrar 75% dos problemas de usabilidade de um website. Considerando que estes são produtos às vezes acessados por um espaço amostral de milhões de usuários, esta amostra de apenas 5 usuários seria confiável para representa-lo? Esta proposta parece bastante atrativa para o mercado, sendo amplamente aceita pelo mesmo, uma vez que baixa consideravelmente os custos (humanos e financeiros) para avaliação de seus softwares. Todavia, esta abordagem procede cientificamente? A seguir, uma simulação que introduz o cálculo desta amostra em termos estatísticos, bem como a confrontação do modelo de análise de Nielsen (2006; 1993) com outros autores.

3.5.2 Determinação do Tamanho de uma Amostra

Consideramos a equação e notações para cálculos referentes ao tamanho da amostra dispostos nas Figuras 14 e 15 (FIGUEIROA, 2012; TRIOLA, 1999; LEVIN, 1987). Nos exemplos adotados aqui, a amostra representa as coletas de avaliações feitas pelos usuários. Em termos gerais, quanto maior a amostra, menor é o erro amostral. Ou seja, quanto maior a amostra, mais provável que seus valores se aproximem do espaço amostral inteiro.

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E} \right)^2$$

Figura 14 – Número de participantes representando o espaço amostral

Onde:

$$\therefore \begin{cases} n = \text{tamanho mínimo da amostra.} \\ Z_{\alpha/2} = \text{valor crítico referente ao grau de confiança desejado.} \\ \sigma = \text{desvio – padrão da amostra.} \\ E = \text{erro máximo de estimativa.} \end{cases}$$

Figura 15 – Notações para fórmula do número de amostras

O valor crítico que representa o grau de confiança da amostra ($Z_{\alpha/2}$) é uma constante interpretada conforme o Quadro 33 abaixo. Por exemplo, se $Z_{\alpha/2} = 2,0$ significa que a amostra tem 95,44% de confiança (probabilidade de fidelidade) com o espaço amostral em questão.

$Z_{\alpha/2}$	%	$Z_{\alpha/2}$	%	$Z_{\alpha/2}$	%	$Z_{\alpha/2}$	%
0,0	0,00	1,0	68,26	2,0	95,44	3,0	99,74
0,1	7,96	1,1	72,86	2,1	96,42	3,1	99,80
0,2	15,86	1,2	76,98	2,2	97,22	3,2	99,86
0,3	23,58	1,3	80,64	2,3	97,86	3,3	99,90
0,4	31,08	1,4	83,84	2,4	98,36	3,4	99,94
0,5	38,30	1,5	86,64	2,5	98,76	3,5	99,96
0,6	45,16	1,6	89,04	2,6	99,06	3,6	99,96
0,7	51,60	1,7	91,08	2,7	99,30	3,7	99,98
0,8	57,62	1,8	92,82	2,8	99,48	3,8	99,98
0,9	63,18	1,9	94,26	2,9	99,62	3,9	100,0

Quadro 33 – Valores correspondentes de grau de confiança

Consideradas estas formalidades matemáticas, adotemos uma simulação hipotética sob caráter elucidativo sobre como se comportaria estatisticamente o modelo proposto por Nielsen: os 5 primeiros usuários encontram aproximadamente 75% dos erros de usabilidade que afetariam a interação com o website. Para tal, considera-se a simulação ilustrativa abaixo (Quadro 34).

Participante	Porcentagem de erros de usabilidade encontrados (Avaliação)
A ₁	30 %
A ₂	55 %
A ₃	20 %
A ₄	80 %
A ₅	70%

Quadro 34 – Amostra "A" da simulação

Para cálculo do desvio padrão, conforme consta na equação explicitada na Figura 14, utilizou-se a fórmula e notações dispostas nas Figuras 16 e 17.

$$\sigma = \sqrt{\frac{(A_1 - M_a)^2 + (A_2 - M_a)^2 + (A_3 - M_a)^2 + (A_4 - M_a)^2 + (A_5 - M_a)^2}{M}}$$

Figura 16 – Fórmula do desvio padrão para cinco escores

Onde:

$$\therefore \begin{cases} \sigma = \text{desvio} - \text{padrão.} \\ M_a = \text{média aritmética dos resultados da pesquisa.} \\ A_1, A_2, A_3, A_4, A_5 = \text{Escores associados a cada membro da pesquisa.} \\ M = \text{quantidade de usuários participantes da pesquisa.} \end{cases}$$

Figura 17 – Notações usadas para fórmula de desvio padrão

Para obter o desvio-padrão é, portanto, necessário extrair a média aritmética da amostra que consta no Quadro 34, conforme exposto na Figura 18 abaixo.

$$M_a = \frac{30 + 55 + 20 + 80 + 70}{5} \rightarrow M_a = \frac{255}{5} \rightarrow M_a = 51$$

Figura 18 – Média aritmética da amostra “A”

Procedendo com a verificação, obtemos o seguinte desvio-padrão (σ_A) (Figura 19).

$$\sigma_A = \sqrt{\frac{2620}{5}} \rightarrow \sigma_A = \sqrt{524} \rightarrow \sigma_A = 22,8910$$

Figura 19 – Desvio-padrão da amostra "A"

Em posse destes números, é possível obter o grau de confiança da amostra "A" do Quadro 34, considerando uma amostra de 5 usuários ($n_A = 5$) realizando o teste, conforme a linha de raciocínio explicitada na Figura 20.

$$n_A = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E} \right)^2 \rightarrow 5 = \left(\frac{Z_{\alpha/2} \cdot 22,891}{5} \right)^2 \rightarrow Z_{\alpha/2} = \sqrt{\frac{125}{523,998}} \rightarrow Z_{\alpha/2} = 0,4884$$

Figura 20 – Obtenção do grau de confiança da amostra "A"

Convertendo $Z_{\alpha/2} = 0,4884$ de acordo com a escala do Quadro 33, obtemos um grau de confiança em aproximadamente 40%. Número consideravelmente baixo, mesmo para uma simulação.

Ainda para efeito ilustrativo, consideremos o mesmo método de Nielsen, mas supondo que os usuários avaliadores sejam designers extremamente bem qualificados em detecção de problemas de usabilidade e agora obtivemos a avaliação de usabilidade do website composta no Quadro 35.

Participante	Porcentagem de erros de usabilidade encontrados (Avaliação)
B ₁	100 %
B ₂	100 %
B ₃	80 %
B ₄	60 %
B ₅	60%

Quadro 35 – Amostra "B" da simulação

Assim, repetindo exatamente o procedimento supracitado, obtivemos novamente desvio-padrão ($\sigma_B = 17,8885$) e o grau de confiança conforme explicitado na Figura 21.

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E} \right)^2 \rightarrow 5 = \left(\frac{Z_{\alpha/2} \cdot 17,8885}{5} \right)^2 \rightarrow Z_{\alpha/2} = \sqrt{\frac{5 \cdot 25}{319,9984}} \rightarrow Z_{\alpha/2} = 0,625$$

Figura 21 – Grau de confiança da amostra "B"

Conferindo o Quadro 33, com $Z_{\alpha/2} = 0,625$ o grau de confiança da amostra "B" refletiu aproximadamente 45%. Novamente, um valor com o qual estatisticamente seria impossível externar a afirmação que 5 usuários refletiriam com segurança o espaço amostral.

3.5.4 Controvérsias na Quantidade de Avaliações em Usabilidade

Provavelmente devido ao relevantemente baixo grau de confiança obtido quando considera-se apenas 5 usuários testando um produto complexo como um website (extensível a quantidade de usuários avaliando outros tipos de produtos), que exista tanta disparidade de informações quando lidamos com o assunto. Baixando a credibilidade da avaliação de artefatos neste contexto.

Faulkner (2003) afirma que avaliando a usabilidade de websites pelo modelo de Nielsen obteve resultados que variavam (para cada grupo de 5 usuários participando da pesquisa) ora 55% dos problemas, ora 99% dos problemas detectados. Enquanto Virzi (1992) reforça Nielsen & Landauer (1993) destacando que detecta 80% dos erros com os primeiros 4 a 5 usuários, Perfetti & Landesman (2001) reporta que ao aplicar o teste de Nielsen encontrou apenas 35% dos problemas com os 5 primeiros usuários.

O baixo grau de confiança deve ser levando em consideração como um dos fatores, senão o principal, desta oscilação de resultados. Muito embora o controverso modelo de Nielsen de avaliação de websites pareça comercialmente viável pela relativamente pouca quantidade de recursos necessários, carece de maiores aprofundamentos quanto a sua validade científica. Sua abordagem específica para usabilidade é desdobrável para todas as outras metodologias mensuradas de avaliação de artefatos. Demandando assim aprofundamentos quanto a sua real viabilidade, conforme a presente tese propõe, mas sob a ótica de métodos científicos que melhor embasem estes posicionamentos.

3.6 Escalas Indutivas

Para a avaliação de um artefato no processo de design é comum que seja conduzida uma pesquisa (não necessariamente acadêmica) para coleta de dados. Normalmente, o designer que conduz este levantamento de informações expõe o protótipo aos participantes, que podem ser especialistas ou usuários em potencial do produto proposto, dependendo do

método de avaliação empregado. Com base nos *feedbacks* destes participantes o artefato é então avaliado e considera-se se atingiu os objetivos demarcados quanto aos que se esperava no planejamento do produto.

Conforme explicitado ao longo desta tese, esta metodologia é amplamente aplicada na avaliação de artefatos: expor o protótipo aos participantes e colher seus *feedbacks*. Para o exercício desta coleta, observa-se um largo emprego de escalas mensuradas. Estas, por sua vez condensam de forma parametrizada a subjetividade dos participantes da pesquisa em relação ao produto, servindo de filtro informacional para o projetista. Sendo também conhecidas como escalas psicométricas.

Especificamente no design, as escalas mais comumente encontradas encontram-se abaixo destacadas.

3.6.1 Escala Likert

Desenvolvida pelo sociólogo e psicólogo Rensis Likert (1932) e, possivelmente, a mais popular das escalas na avaliação de produtos dado o substancial número de recorrências em pesquisas de design (FIGUEIREDO, 2012; CORREIA, 2010; BELFORT, 2011; AGUIAR, 2010). De acordo com Brandalise (2005) o consumidor avalia produtos e serviços conforme suas experiências e influências sociais e a escala Likert, também conhecida como “escala somada”, tem especial adequação a este contexto.

Esta escala psicométrica fora concebida para mensurar a atitude do entrevistado frente a alternativas objetivas bipolares (LIKERT, 1932). Em outras palavras, diante de um questionamento o entrevistado recebe normalmente 5 opções (escala de 5 pontos) de respostas objetivas que se repelem entre si linearmente em torno de uma âncora neutra central, conforme no Quadro 36 abaixo.

Como você avalia a experiência proporcionada pelo produto?
1. Muito satisfeito
2. Satisfeito
3. Indiferente
4. Insatisfeito
5. Muito insatisfeito

Quadro 36 – Escala Likert

A coleta de dados pode oferecer ao entrevistado uma série de perguntas com suas respectivas alternativas bipolares. Para efeito de análise, para cada uma das respostas é atribuído um valor: +2, +1, 0, -1, -2; respectivamente, sendo +3, +2, +1, 0, -1, -2, -3 caso toda a coleta obedeça uma escala de 7 pontos (valores inteiros). Sendo ainda possível valores de 1 a 5; de 1 a 7; de 5 a 1 ou; de 7 a 1 (valores naturais). A média aritmética dos pontos obtidos pelas respostas de cada um dos questionamentos levantados volve a avaliação.

3.6.2 Escala de Diferencial Semântico

Técnica desenvolvida para apuração da atitude do entrevistado, em relação ao objeto de pesquisa, revelando “força” e “direção” da atitude numa escala de 7 pontos. O contínuo é ancorado nas extremidades por um par de adjetivos bipolares, sendo o centro da escala a alternativa neutra (BRANDALISE, 2005; CORREIA, 2010) (Ver Quadro 37).

Como você avalia o aparelho utilizado?						
Inamistoso	—	—	—	—	—	Amistoso
Moderno	—	—	—	—	—	Antiquado
Eficiente	—	—	—	—	—	Ineficiente
Lento	—	—	—	—	—	Rápido
Agradável	—	—	—	—	—	Desagradável

Quadro 37 – Escala de Diferencial Semântico. Adaptado de BRANDALISE (2005)

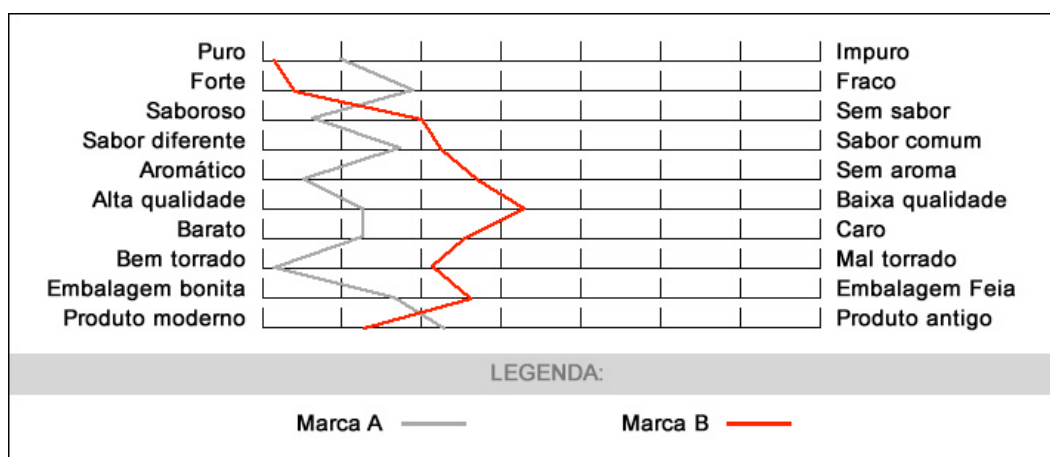
A exposição da escala ao entrevistado – na qual ele irá marcar sua avaliação – pode ser acompanhada, ou não, de números. Apesar de que, segundo Mattar (1996), o emprego da numeração na escala tem gerado controvérsias, no sentido de afetar a opinião dos usuários. Evitando ainda a indução das respostas, Brandalise (2005) recomenda não concentrar aspectos negativos exclusivamente num dos lados da escala.

Brandalise (2005) destaca que cerca de 50 adjetivos bipolares foram levantados para melhor expressar a avaliação dos entrevistados. Uma vez que adjetivos que não sejam consideravelmente bipolares, ou que não sejam assim interpretados pelo modelo mental dos entrevistados, podem gerar resultados infiéis. A autora destaca que o uso da Escala de Diferencial Semântico é especialmente útil para avaliação de produtos e marcas concorrentes, uma vez que constrói gráficos comparativos claros. Permitindo a análise

vantajosa de várias unidades simultaneamente, conforme exemplificado nos Quadros 38 e 39 abaixo.

Com relação à marca de café A, qual sua opinião sobre os seguintes atributos:									
Puro	-	-	-	-	-	-	-	-	Impuro
Forte	-	-	-	-	-	-	-	-	Fraco
Saboroso	-	-	-	-	-	-	-	-	Sem sabor
Sabor diferente	-	-	-	-	-	-	-	-	Sabor comum
Aromático	-	-	-	-	-	-	-	-	Sem aroma
Alta qualidade	-	-	-	-	-	-	-	-	Baixa qualidade
Barato	-	-	-	-	-	-	-	-	Caro
Bem torrado	-	-	-	-	-	-	-	-	Mal torrado
Embalagem bonita	-	-	-	-	-	-	-	-	Embalagem Feia
Produto moderno	-	-	-	-	-	-	-	-	Produto antigo

Quadro 38 – Coleta por Diferencial Semântico. Adaptado de MATTAR (1996)



Quadro 39 – Análise por Diferencial Semântico. Adaptado de MATTAR (1996)

Montadas basicamente sobre Diferencial Semântico, a escala de avaliação de criatividade CPSS, modelou adjetivos que de acordo com Besemer & O'Quin (1986) permitem o modelamento numérico e gráfico da avaliação da avaliação de produtos.

3.6.3 Escala de Stapel

Representa um aprofundamento, uma modificação, da Escala de Diferencial Semântico, estendendo o contínuo a uma escala de 11 pontos, aumentando a precisão e poder de expressão do entrevistado (AGUIAR, 2011). Dentre as diferenças da escala de Stapel para sua precursora, emergem os 11 pontos ao invés de 7, sempre visivelmente enumerados na escala, partindo de +5, passando por 0 e terminando em -5, com valores inteiros (VELUDO-DE-OLIVEIRA, 2001). Cabe destacar ainda que os adjetivos, ou frases descritivas, são testados de modo unipolar conforme melhor representado no Quadro 40.

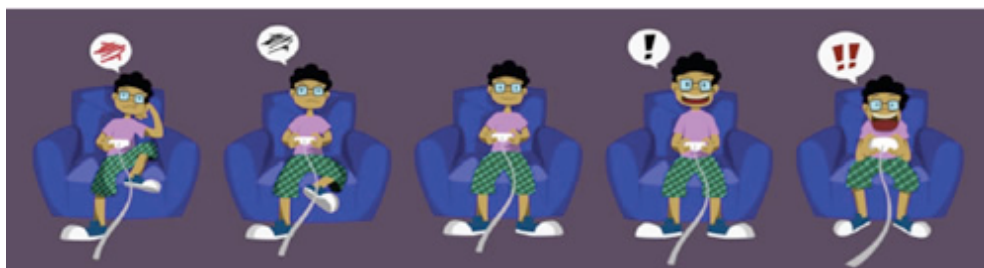
Avalie os aspectos do game XYZ marcando um valor de acordo com sua opinião:			
+5	+5	+5	+5
+4	+4	+4	+4
+3	+3	+3	+3
+2	+2	+2	+2
+1	+1	+1	+1
Divertido	Bonito	Difícil	Dá vontade de jogar
-1	-1	-1	-1
-2	-2	-2	-2
-3	-3	-3	-3
-4	-4	-4	-4
-5	-5	-5	-5

Quadro 40 – Escala de Stapel. Adaptado de AGUIAR (2011)

A projeção comparativa para análise das avaliações coletadas via escala Stapel podem ser expostas de modo semelhante ao realizado na Escala de Diferencial Semântico (Quadro 37).

3.6.4 Escala Pictórica

O uso de escalas pictóricas, apesar de pouca – ou nenhuma – formalização no design, demonstra-se especialmente útil na coleta de experiências e emoções do usuário (TWYMAN, 1985). Nesta linha argumentativa, a pesquisa de Belfort (2011) aplicou escala pictórica para coletar o *feedback* da experiência de usuários durante sua interação com o uso de um game online. Integrando os valores da escala nos moldes de Likert (1932).



Quadro 41 – Escala pictórica. Adaptado de BELFORT (2011)

Em vez de apresentar a escala para coleta conforme disposto Quadro 36, nesta ocasião utilizou-se a escala pictórica disposta no Quadro 41. Sendo a cada um dos

pictogramas atribuídos os valores de -2 a +2, passando por 0 (zero) como âncora neutra e utilizando valores inteiros.

3.6.5 Escala de Guttman

Escala similar à Likert, porém com uma composição par no número de âncoras (de 1 a 6, por exemplo). A forma como é apresentada ao usuário, traduz essencialmente a intensidade da atitude do mesmo frente o que está sendo avaliado. Esta técnica pode ter uma ou uma série de perguntas, cada qual dispondo de uma composição para resposta objetiva conforme explicitado no Quadro 42.

Você achou o game XYZ divertido? <i>Selecione uma das alternativas abaixo</i>					
SIM	SIM	SIM	NÃO	NÃO	NÃO

Quadro 42 – Exemplo de Escala de Guttman. Adaptado de AGUIAR(2011)

Para posterior análise do pesquisador, cada resposta é traduzida por um equivalente numérico natural, gerando as projeções gráficas assim como nas demais escalas supracitadas.

3.6.6 Escala de Alpert

Utilizada particularmente para comparar marcas e produtos, avaliando uma série de atributos dos mesmos julgados relevantes pelo pesquisador, mediante 3 dimensões abaixo topificadas.

1. **Importância:** deste parâmetro para o entrevistado;
2. **Satisfação:** proporcionada pelo produto quanto a este parâmetro;
3. **Diferença:** entre o produto A e o produto B tratando-se deste parâmetro.

Gerando, com base nestas diretrizes, um quadro de avaliação exposto ao entrevistado, conforme apresentado no Quadro 43. Para cada atributo o entrevistado aloca uma pontuação de 1 a 5, sendo que 1 significa “nenhuma importância / nenhuma satisfação / nenhuma diferença” e 5 significa “totalmente importante / total satisfação / totalmente diferente”, dispostas num contínuo.

Avalie, segundo uma escala de 1 a 5, em que 1 significa nenhum e 5 significa totalmente: 1) a importância do atributo apresentado em relação a uma sala de cinema; 2) a satisfação que o cinema A proporciona quanto ao atributo apresentado; 3) a diferença do atributo apresentado do cinema A para o cinema B.			
Atributos	Importância	Satisfação	Diferença A e B
Pontualidade quanto ao início do filme.	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Variedade da programação de filmes.	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Atualização da programação de filmes.	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Conforto da sala de cinema.	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Luminosidade da sala de cinema.	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Som da sala de cinema.	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5

Quadro 43 – Escala Alpert. Adaptado de VELUDO-DE-OLIVEIRA(2001)

Para análise desta avaliação por parte do pesquisador considera-se o ranking como a multiplicação dos valores das dimensões, podendo cada atributo, portanto, variar de 125 pontos (5x5x5) a 1 ponto (1x1x1). Com base nestes números gera-se a projeção gráfica da avaliação dos atributos de um dado produto ou serviço.

3.6.7 Problema: Divergências Parciais nos Resultados das Escalas

Nesta subseção apresentam-se, por revisão literária, pesquisas apontando que o uso das escalas supracitadas e que são amplamente empregadas na avaliação de produtos podem estar induzindo os resultados da educação.

A pesquisa de Aguiar (2011) aponta divergências entre os resultados de um mesmo produto quando avaliado por escalas diferentes, utilizando de uma metodologia de análise de escalas através comparação dos desvios-padrão. Apesar de retornar resultados preliminares (que carecem de repetições da aplicação da metodologia para maior consistência dos dados), esta pesquisa levanta este relevante resultado. Considerando o uso de escalas e projeções como principais meios de educação das avaliações de artefatos, a precisão desta coleta de dados é crucial para que o designer encontre as correções críticas necessárias para o bom desempenho do produto, bem como na verificação do quanto o mesmo atende as exigência do mercado. Neste sentido, uma divergência de resultados entre as escalas pode comprometer a avaliação, levando o projetista a focar seus esforços em irrelevantes problemas no produto enquanto àqueles que demandam real atenção não foram devidamente pontuados em determinada escala. Característica permanente em

eventos aleatórios (várias possibilidades de resultado demandando atenção) e que podem ter sua multiplicidade de resultados ampliada pela imprecisão da escala.

Além da flutuação de resultados de um mesmo produto diante de escalas diferentes, outras controvérsias no uso das escalas demandam a atenção do designer. A escala de diferencial semântico é uma das escalas mais utilizadas neste âmbito de acordo com Correia (2010) e Besemer (1998) (Vide Quadro 37). Todavia, observando o potencial desta escala em induzir os resultados, Brandalise (2005) e Mattar (1996) conceberam as recomendações quanto ao não-uso de números na exposição da escala aos avaliadores. Bem como a não concentração de adjetivos considerados negativos em apenas um dos lados da tabela, evitando a indução dos resultados (BRANDALISE, 2005; MATTAR, 1996). Deste modo, estas são as recomendações levantadas até então no uso das escalas de avaliação.

No capítulo a seguir, apresentam-se as pesquisas desenvolvidas por estudos de caso objetivando gerar soluções para os problemas da falácia da conjunção e do espaço amostral, bem como as justificativas para este escopo metodológico.

4 ESTUDOS DE CASOS

Uma vez mapeados os 5 potenciais problemas decorrentes da aleatoriedade⁵ na avaliação de artefatos, subselecionou-se 2 destes problemas para desenvolver respectivos aprofundamentos: a falácia da conjunção e a representação do espaço amostral (Ilustrado no Quadro 44 abaixo).

A justificativa para tal seleção enreda-se tanto pela relevância dos problemas no domínio do design, dado o número de publicações recorrentes, quanto as condições metodológicas favoráveis pra uma pesquisa de doutorado (envolvendo tempo hábil, capital humano e laboratórios disponíveis).

FENÔMENO/PROBLEMA	PROPOSTA DE SOLUÇÃO
REGRESSÃO À MÉDIA	REVISÃO BIBLIOGRÁFICA
FALÁCIA DA CONJUNÇÃO	EXPERIMENTO
INCERTEZA	REVISÃO BIBLIOGRÁFICA
ESPAÇO AMOSTRAL	EXPERIMENTO
ESCALAS INDUTIVAS	REVISÃO BIBLIOGRÁFICA

Quadro 44 – Escopo para estudos de casos

Se faz conveniente destacar que só é possível a plena compreensão dos procedimentos abaixo aplicados com a leitura prévia do Capítulo 3 da tese (Problemas Decorrentes da Aleatoriedade). Principalmente tratando-se das seções 3.3 e 3.5: A Falácia da Conjunção e; Representação do Espaço Amostral, respectivamente.

No mesmo capítulo 3 (Problemas Decorrentes da Aleatoriedade) apresenta-se apenas uma hipótese de que a falácia da conjunção afetaria o processo de design, enquanto que, referente ao espaço amostral, demonstra-se que existe uma inconsistência na quantidade

⁵ Considera-se nesta tese, conforme amplamente definido ao longo do documento, o “aleatório” como o evento com multiplicidade de resultados prováveis.

de avaliadores. Por esta razão, o estudo de caso referente a falácia da conjunção tornou-se mais extenso que o do espaço amostral, pois fora necessário primeiro executar procedimentos que confirmassem a hipótese para depois propor uma solução metodológica para o problema. Enquanto que no estudo de caso do espaço amostral, fora apresentado apenas a solução metodológica proposta.

Como cada um dos estudos de caso tem abordagens processuais diferentes, os detalhes quanto a metodologia adotada encontra-se descrita dentro de cada um dos mesmos.

4.1 Falácia da Conjunção

Este estudo de caso investiga a hipótese de Ferreira (2008), de que a falácia da conjunção estaria presente na avaliação de produtos. Conduzindo, assim, experimentos controlados para expor este problema no processo de design, bem como prover sugestões de mecanismos metodológicos que evitem esta falácia (FIGUEIRÔA, et al. 2012). Após 185 entrevistas com estudantes de graduação em design, chegou-se à resultados que permitem entender o fenômeno e minimizar o problema o problema no projeto de artefatos digitais. A correção proposta é extensível a outros seguimentos de design e mesmo a todos os outros domínios nos quais a falácia da conjunção for identificada.

4.1.1 Procedimentos Metodológicos

A metodologia adotada nesta pesquisa, para verificar a existência da falácia da conjunção no processo de design, é baseada na criada por Tversky & Kahneman (1983) no estudo de caso que ficou conhecido como “Linda”⁶. Sendo características processuais deste experimento ter uma breve descrição de uma conjectura (de perfil de um personagem ou de acontecimento de uma situação específica) e apresentação de alternativas de resposta, solicitando ao entrevistado que as classifiquem por ordem de probabilidade, de acordo com o detalhamento abaixo.

- **Evento provável:** Uma das alternativas é intencionalmente o evento mais provável em relação a conjectura apresentada (Denominado evento “A” ou “alternativa provável”);
- **Evento improvável:** Uma das alternativas é intencionalmente o evento mais improvável em relação a conjectura apresentada (Denominado evento “B” ou “alternativa improvável”);

⁶ Para visualizar o experimento ver Capítulo 3, Seção 3.3, Quadro 27

- **Evento conjuntivo:** Uma das alternativas é um evento conjuntivo que apresenta intencionalmente a alternativa mais provável e a mais improvável combinadas (Denominado evento “A&B” ou “evento conjuntivo”);
- **Eventos opcionais:** Demais alternativas podem ser posicionadas se os pesquisadores julgarem necessário, para camuflar ao entrevistado o fato de que as 3 alternativas supracitadas estão sendo monitoradas.

Nestes termos, espera-se que os entrevistados violem intuitivamente a primeira lei da probabilidade, escolhendo o evento conjuntivo como mais provável do que um evento isolado. Em outras palavras, espera-se que os entrevistados avaliem como mais provável a alternativa “A&B” do que a alternativa “B”, que contradiz a “regra da conjunção”. Pois os mesmos estariam afirmando (erroneamente) que um evento conjuntivo tem mais probabilidade de ocorrer do que um evento isolado⁷.

Para classificar as alternativas, da mais provável à mais improvável, solicita-se que os entrevistados atribuam valores numéricos naturais para cada uma delas. Sendo 1 atribuído à alternativa julgada mais provável e de modo crescente até a mais improvável. Por exemplo, se a descrição (conjectura) apresentada apresenta 4 alternativas de resposta, o entrevistado deve classificá-las de 1 a 4.

Ao final, a média aritmética das classificações atribuídas a cada uma das alternativas gera uma escala. Em ordem crescente é possível observar a alternativa que foi julgada pelo grupo entrevistado como a mais provável (menor valor na escala) até a mais improvável (maior valor na escala).

Ao realizar o experimento de Tversky & Kahneman (1983) para expor a falácia da conjunção no processo de design e contemplar os dados, fora possível projetar uma tentativa de solução para o problema. Assim sendo, este estudo de caso se divide em duas etapas. Na primeira apresenta-se como os designers podem projetar equivocadamente seus projetos tentando corrigir problemas que não existem, por caírem na falácia aqui em discussão. Na segunda etapa apresenta-se como corrigir este problema cognitivo balanceando os eventos conjuntivos.

⁷ De acordo com Freund (2006), bem como Tversky & Kahneman (1983), a probabilidade de eventos isolados ocorrerem separadamente sempre será maior que a destes eventos ocorrerem de forma combinada. Em outras palavras, a interseção destes conjuntos de eventos nunca será maior do que cada um dos conjuntos.

4.1.2 Expondo a Falácia da Conjunção no Design

Esta etapa da pesquisa resultou num total de 185 entrevistas válidas, realizadas em cursos superiores de design, design gráfico e web design, respectivamente nas instituições UFPE⁸, UNIBRATEC⁹ e FMR¹⁰. Foram projetadas 3 descrições (conjecturas) com suas respectivas alternativas, sendo 2 delas para exposição do problema no design (expostas nesta subseção do documento) e 1 para sugerir como corrigir o problema da falácia da conjunção neste contexto (exposta na próxima subseção do documento) (FIGUEIROA, 2012).

Dando prosseguimento aos procedimentos descritos acima, foram inicialmente entrevistados 50 estudantes do curso de web design da FMR, apresentando-lhes a descrição e as 5 alternativas explicitadas no Quadro 45. Estes entrevistados foram denominados “Grupo 1”. Solicitou-se que atribuissem valor 1 para a alternativa considerada mais provável, seguindo a escala de números naturais até o valor 5, atribuído a alternativa considerada mais improvável.

DINOTUNES é um website de venda de álbuns de música, constituído de tecnologias Flash, CSS e HTML 5. Este site recebeu uma má avaliação quanto a satisfação dos usuários que alegam encontrar problemas para comprar as músicas.
☐ O website se parece com uma loja de computadores
☐ O website apresenta problemas de usabilidade
☐ O website tem pouca divulgação nas redes sociais
☐ O website tem muitos botões
☐ O website tem pouca divulgação nas redes sociais e problemas de usabilidade

Quadro 45 – DINOTUNES: Descrição e alternativas

Para tal procedimento, utilizou-se um software gratuito de coleta dados chamado Survey Monkey¹¹. Os entrevistados responderam à pesquisa nos computadores do laboratório de web design da referida faculdade e seguiram as instruções conforme exposto na Figura 22. Somente era permitido atribuir um dos valores da escala de 1 a 5 para cada uma das alternativas. Ou seja, propositalmente o software fora configurado para não aceitar, por exemplo, duas alternativas ou mais com valor 1.

⁸ Universidade Federal de Pernambuco

⁹ Faculdade de Tecnologia IBRATEC

¹⁰ Faculdade Marista Recife

¹¹ <http://www.surveymonkey.com>

Bem vindo Survey

www.surveymonkey.com/s/9WHNLR5

LEIA COM ATENÇÃO:

Após ler o enunciado, enumere as afirmações de 1 a 5 sendo 1 MAIS PROVÁVEL e o valor 5 MAIS IMPROVÁVEL.

Importante: Atribuir um número para cada afirmação, sem repetir.

***1. DINOTUNES é um website de venda de álbuns de música, constituído de tecnologias Flash, CSS e HTML 5. Este site recebeu uma má avaliação quanto a satisfação dos usuários que alegam encontrar problemas para comprar as músicas.**

	1	2	3	4	5
O website se parece com uma loja de computadores	☐	☐	☐	☐	☐
O website apresenta problemas de usabilidade	☐	☐	☐	☐	☐
O website tem pouca divulgação nas redes sociais	☐	☐	☐	☐	☐
O website tem muitos botões	☐	☐	☐	☐	☐
O website tem pouca divulgação nas redes sociais e problemas de usabilidade	☐	☐	☐	☐	☐

Seg.

Figura 22 – Tela do software de pesquisa utilizado

É de relevante destaque que a descrição foi intencionalmente concebida para representar um problema de usabilidade facilmente identificável por designers, e não um problema de divulgação. Ou seja, os eventos observados eram especificamente:

- O website apresenta problemas de usabilidade (Evento “A”);
- O website tem pouca divulgação nas redes sociais (Evento “B”);
- O website tem pouca divulgação nas redes sociais e problemas de usabilidade (Evento conjuntivo “A&B”).

No Quadro 46 encontram-se as médias de pontuação atribuídas pelos entrevistados para cada alternativa. Quanto menor o valor da média, maior foi a atribuição de probabilidade dada pelos entrevistados, exatamente como postulado pelo procedimento de Tversky & Kahneman (1983).

Alternativa	Média
O website apresenta problemas de usabilidade (A)	1,92
O website tem pouca divulgação nas redes sociais e problemas de usabilidade (A&B)	2,64
O website tem muitos botões	2,98
O website tem pouca divulgação nas redes sociais (B)	3,48
O website se parece com uma loja de computadores	3,98
Total de 50 amostras	

Quadro 46 – Ordem de probabilidade atribuídas pelos entrevistados do Grupo 1

Como é possível observar no Quadro 47, o evento conjuntivo resultou em 2º lugar na escala. Das 50 amostras, em 76% (38 delas) os entrevistados julgaram o evento conjuntivo (A&B) como mais provável que o evento isolado (B) (Quadro 47 abaixo). Violando a regra da conjunção.

DINOTUNES - GRUPO 1																			
(a) O website se parece com uma loja de computadores																			
(b) O website apresenta problemas de usabilidade																			
(c) O website tem pouca divulgação nas redes sociais																			
(d) O website tem muitos botões																			
(e) O website tem pouca divulgação nas redes sociais e problemas de usabilidade																			
Amostra						Amostra						Amostra							
a	b	c	d	e		a	b	c	d	e		a	b	c	d	e			
1	5	1	4	2	3		19	4	1	5	3	2		37	5	1	4	3	2
2	2	5	4	1	3		20	5	2	3	4	1		38	3	5	2	1	4
3	5	1	4	2	3		21	5	1	3	4	2		39	5	2	3	4	1
4	2	1	3	4	5		22	5	3	4	2	1		40	5	1	3	2	4
5	5	1	3	2	4		23	1	4	3	2	5		41	5	1	4	2	3
6	4	5	1	3	2		24	3	1	4	2	5		42	5	3	2	4	1
7	5	1	3	4	2		25	5	1	4	2	3		43	5	3	2	4	1
8	4	1	3	5	2		26	1	3	2	5	4		44	4	1	5	2	3
9	4	2	5	1	3		27	5	1	4	2	3		45	1	5	2	4	3
10	3	1	5	4	2		28	3	1	5	2	4		46	5	4	2	1	3
11	5	1	4	3	2		29	1	2	5	4	3		47	3	1	5	4	2
12	5	1	3	4	2		30	5	1	4	3	2		48	5	2	1	4	3
13	4	3	2	5	1		31	4	1	5	2	3		49	5	1	3	4	2
14	5	1	4	3	2		32	4	1	5	3	2		50	3	1	5	2	4
15	4	5	2	3	1		33	2	1	5	3	4							
16	3	1	5	2	4		34	5	1	4	3	2							
17	4	3	2	5	1		35	5	1	3	4	2							
18	3	1	4	2	5		36	5	4	2	3	1							

Quadro 47 – Resultados "Grupo 1"

Outra coleta de dados foi incursada na UFPE (Universidade Federal de Pernambuco), com 42 estudantes do bacharelado em design, denominados para efeitos de organização desta pesquisa como “Grupo 2”. Seguindo o mesmo protocolo, porém desta vez utilizando formulário em papel (ao invés do software) com escalonamento de 1 a 6 (do mais provável ao mais improvável, nesta ordem) (ANEXO A). Apresentando a descrição e as alternativas presentes no Quadro 49.

Designers estão projetando o iCopa, um aplicativo para smartphones que apresenta simultaneamente várias informações relevantes sobre a copa do mundo. O protótipo do produto tem recebido muitas reclamações dos seus avaliadores, que normalmente se perdem na sua interface.
O aplicativo apresenta uma interface complexa, com muitos botões
O aplicativo parece um jornal digital
O aplicativo tem pouca divulgação
O aplicativo não apresenta uma interface com temática esportiva
O aplicativo tem pouca divulgação e apresenta uma interface complexa, com muitos botões
O aplicativo é composto por uma tecnologia ultrapassada

Quadro 48 – ICOPA: Descrição e alternativas

Novamente a descrição intencionalmente relata um problema de interface e não de divulgação. Foi obtido por parte dos entrevistados o escalonamento resultante no Quadro 50.

Alternativa	Média
O aplicativo apresenta uma interface complexa, com muitos botões (A)	1,86
O aplicativo tem pouca divulgação e apresenta uma interface complexa, com muitos botões (A&B)	2,48
O aplicativo parece um jornal digital	3,64
O aplicativo tem pouca divulgação (B)	3,90
O aplicativo não apresenta uma interface com temática esportiva	4,26
O aplicativo é composto por uma tecnologia ultrapassada	4,88
<i>Total de 42 amostras</i>	

Quadro 49 – Ordem de probabilidade atribuídas pelos entrevistados do Grupo 2

De acordo com o Quadro 50, novamente o evento conjuntivo teve destaque em relação em relação aos demais. Em 85,7% das amostras (36 das 42 coletadas), o evento “A&B” foi atribuído como mais provável que o evento “B” (Quadro 51 abaixo). Violando novamente a regra da conjunção.

ICOPA - GRUPO 2																						
(a) O aplicativo apresenta uma interface complexa, com muitos botões																						
(b) O aplicativo parece um jornal digital																						
(c) O aplicativo tem pouca divulgação																						
(d) O aplicativo não apresenta uma interface com temática esportiva																						
(e) O aplicativo tem pouca divulgação e apresenta uma interface complexa, com muitos botões																						
(f) O aplicativo é composto por uma tecnologia ultrapassada																						
Amostra								Amostra								Amostra						
51	1	2	5	6	3	4		66	2	6	5	4	1	3		81	1	3	5	6	2	4
52	1	4	3	5	2	6		67	1	3	5	2	4	6		82	2	4	3	6	1	5
53	2	6	3	5	1	4		68	2	4	3	5	1	6		83	1	4	3	5	2	6
54	1	3	6	5	2	4		69	2	4	6	1	3	5		84	3	6	4	5	1	2
55	2	5	3	4	1	6		70	2	6	3	4	1	5		85	2	5	3	6	1	4
56	3	1	2	4	5	6		71	2	4	6	5	1	3		86	1	3	4	5	2	6
57	1	2	5	3	4	6		72	4	1	2	3	5	6		87	1	3	6	2	5	4
58	3	4	2	5	1	6		73	3	6	4	2	1	5		88	2	1	4	6	3	5
59	1	4	6	2	5	3		74	1	4	3	5	2	6		89	2	5	3	6	1	4
60	1	2	6	4	5	3		75	3	1	2	5	6	4		90	1	2	5	4	3	6
61	2	6	4	3	1	5		76	1	5	3	6	2	4		91	1	3	6	4	2	5
62	2	4	3	6	1	6		77	4	2	1	3	5	6		92	4	1	2	3	6	5
63	1	5	6	4	3	2		78	2	5	4	3	1	6								
64	1	4	5	2	3	6		79	2	5	3	4	1	6								
65	2	1	4	5	3	6		80	2	4	3	6	1	5								

Quadro 50 – Resultados "Grupo 2"

Concluindo os resultados preliminares de exposição da falácia da conjunção em design com um total de 92 amostras, das quais a relevante maioria (74 delas) violou a regra da conjunção, resultando numa média de 80,4% dos casos.

4.1.3 Resultados da Correção da Falácia da Conjunção no Design

Observando que as alternativas isoladas em si eram classificadas pelos usuários com uma ordem de probabilidade coerente e que as alternativas conjuntivas recebiam um otimismo infundado, remodelou-se o procedimento com a descrição e as alternativas expostas no Quadro 52.

DINOTUNES é um website de venda de álbuns de música, constituído de tecnologias Flash, CSS e HTML 5. Este site recebeu uma má avaliação quanto a satisfação dos usuários que alegam encontrar problemas para comprar as músicas.
O website se parece com uma loja de computadores e apresenta problemas de usabilidade
O website tem pouca divulgação nas redes sociais e se parece com uma loja de computadores
O website apresenta problemas de usabilidade e tem muitos botões
O website tem muitos botões e pouca divulgação nas redes sociais

Quadro 51 – DINOTUNES com todas as alternativas conjuntivas

Posicionando, portanto, todas as alternativas com eventos conjuntivos. Os eventos que por hora apresentam-se combinados, são visualizados de forma isolada como:

- A1) O website apresenta problemas de usabilidade (Evento provável);
- A2) O website tem muitos botões (Evento provável);
- B1) O website se parece uma loja de computadores (Evento improvável);
- B2) O website tem pouca divulgação nas redes sociais (Evento improvável).

Assim espera-se que o evento conjuntivo que combina alternativas prováveis sejam classificados pelos entrevistados como o mais provável (“A1&A2”). Bem como os eventos que combinam intencionalmente alternativas prováveis com improváveis (“A1&B1” e “A2&B2”, não necessariamente nesta ordem) recebam classificações mediadas. Porfim, espera-se que a alternativa que combina os eventos improváveis (“B1&B2”) fique por último nesta classificação de probabilidade. Lembrando que os eventos “B” são os intencionalmente projetados para serem considerados improváveis, e os eventos “A” de igual modo para parecerem prováveis diante da descrição.

Foram entrevistados, para esta etapa, um 93 estudantes de design da UNIBRATEC (Faculdade de Tecnologia IBRATEC), denominados “Grupo 3”. Como resultado obteve-se a classificação disposta no Quadro 53 abaixo.

Alternativa	Média
O website apresenta problemas de usabilidade e tem muitos botões (A1&A2)	1,61
O website se parece com uma loja de computadores e apresenta problemas de usabilidade (A1&B1)	2,45
O website tem muitos botões e pouca divulgação nas redes sociais (A2&B2)	2,88
O website tem pouca divulgação nas redes sociais e se parece com uma loja de computadores (B1&B2)	3,07
<i>Total de 93 amostras</i>	

Quadro 52 – Ordem de probabilidade atribuídas pelos entrevistados do Grupo 3

Em contraposição dos casos anteriores, e mesmo às tentativas de manipulação de Tversky & Kahneman (1983), agora os entrevistados responderam erroneamente em apenas 17,2% dos casos. Ou seja, apresentando as alternativas com balanceamento de conjunção eles passaram a responder corretamente em 77 das 93 entrevistas (82,8%), onde o evento mais provável (“A1&A2”) foi assim classificado em relação ao evento mais improvável (“B1&B2”).

Em síntese, os resultados obtidos reduziram uma média de 80,4% de escolhas equivocadas (ocasionadas pela falácia da conjunção) para 17,2% com a proposta de correção, lidando com problemas comuns no design de artefatos digitais. A totalidade da coleta de resultados referente ao Grupo 3 encontra-se no Quadro 54 a seguir.

GRUPO 3 – DINOTUNES – ALTERNATIVAS CONJUNTIVAS																
(a) O website se parece com uma loja de computadores e apresenta problemas de usabilidade																
(b) O website tem pouca divulgação nas redes sociais e se parece com uma loja de computadores																
(c) O website apresenta problemas de usabilidade e tem muitos botões																
(d) O website tem muitos botões e pouca divulgação nas redes sociais																
Amostra	a	b	c	d		Amostra	a	b	c	d		Amostra	a	b	c	d
93	1	3	2	4		124	2	4	1	3		155	2	3	4	1
94	2	1	4	3		125	3	2	1	4		156	4	2	3	1
95	1	4	2	3		126	4	3	1	2		157	1	3	2	4
96	3	4	1	2		127	2	4	1	3		158	4	3	1	2
97	3	1	2	4		128	1	4	2	3		159	2	3	1	4
98	2	3	1	4		129	3	4	1	2		160	2	3	1	4
99	1	4	2	3		130	2	3	1	4		161	1	3	4	2
100	4	2	1	3		131	2	4	1	3		162	4	3	1	2
101	2	4	1	3		132	4	3	1	2		163	2	4	1	3
102	4	3	1	2		133	1	3	2	4		164	4	3	2	1
103	2	1	3	4		134	4	2	3	1		165	2	4	1	3
104	4	2	3	1		135	4	3	1	2		166	2	3	1	4
105	2	4	1	3		136	4	1	2	3		167	1	3	2	4
106	3	4	1	2		137	2	3	1	4		168	2	4	1	3
107	1	3	2	4		138	4	3	2	1		169	2	4	1	3
108	3	4	1	2		139	4	3	2	1		170	4	2	3	1
109	2	4	1	3		140	1	4	2	3		171	3	1	2	4
110	1	2	4	3		141	1	4	2	3		172	1	4	2	3
111	4	3	1	2		142	4	3	1	2		173	1	4	2	3
112	1	3	2	4		143	3	2	1	4		174	2	4	1	3
113	2	4	1	3		144	2	4	1	3		175	4	1	2	3
114	3	4	2	1		145	1	3	2	4		176	2	4	1	3
115	2	3	1	4		146	2	4	1	3		177	3	4	1	2
116	2	4	1	3		147	1	3	2	4		178	4	3	1	2
117	4	2	3	1		148	2	3	1	4		179	3	2	1	4
118	3	4	1	2		149	4	3	1	2		180	3	2	1	4
119	3	1	4	2		150	2	3	1	4		181	2	1	3	4
120	4	3	1	2		151	2	3	1	4		182	1	4	3	2
121	4	3	1	2		152	2	4	1	3		183	2	3	1	4
122	1	3	2	4		153	2	3	1	4		184	3	4	1	2
123	2	4	1	3		154	1	3	2	4		185	4	2	1	3

Quadro 53 – Resultados "Grupo 3"

Os 185 resultados foram coletados, contabilizados e analisados ao longo de 4 meses, conforme a disponibilidade de turmas de designers. Durante a coleta em papel, apesar do aviso no formulário (ANEXO A) e das instruções verbais transmitidas aos entrevistados,

alguns entregavam com alternativas em branco ou com todas as alternativas com um mesmo valor. Sendo estes formulários descartados e não entrando para a contabilização dos dados acima apresentados.

4.1.4 Considerações Referentes ao Estudo de Caso Sobre a Falácia da Conjunção

Após as entrevistas com os Grupos 1 e 2 fora explanado aos estudantes que eles caíram na falácia da conjunção com uma apresentação do fenômeno, mas alguns ainda mantinham seu posicionamento de que o evento conjuntivo era o mais provável. Exatamente como ocorrido na pesquisa de Tversky & Kahneman (1983) houve o caso de um estudante externar seu posicionamento de que somente matematicamente o evento conjuntivo (no caso do iCopa – Grupo 2) era mais improvável, mas que na “vida real” isto não procedia. Ele tentou explicar que ao combinar o evento improvável (B) com o mais provável (A), ele aumentou a credibilidade do evento (B). Este posicionamento somente reforça o quanto a presente pesquisa (bem como as demais sobre o fenômeno em estudo) é relevante, pois demonstra que a má compreensão dos eventos probabilísticos é intrínseca nos nossos processos cerebrais, conforme criticado por Mlodinow (2009, p.7). Na prática, ao selecionar um evento que já é improvável de acontecer, neste caso “um problema de divulgação” tornou-se ainda mais improvável ao colocar uma condição para que o mesmo ocorra (qualquer que seja ela, provável ou não) (FREUND, 2006).

Caso os protótipos de produto descritos nas investigações realizadas nos Grupos 1 e 2 fossem situações reais, os designers tomariam decisões de projeto crucialmente equivocadas. Investindo tempo, dinheiro e demais recursos que nestas circunstâncias são raros (e portanto custosos) para resolver um problema que não existia.

Por exemplo, no caso do DINOTUNES (Quadro 46), com base no resultado da avaliação (Quadro 47), seriam investidos recursos também na divulgação do aplicativo em redes sociais, algo relativamente caro em relação a testes de usabilidade. Direcionando o designer para um problema que não tem nenhuma relação com o produto. Afinal, qual a ligação lógica ou operacional entre o problema de usabilidade de uma interface e a divulgação deste produto?

Situação igualmente ocorrida no caso do produto fictício iCopa (Quadro 49). Os usuários já estariam usando o produto (portanto já conheceriam o mesmo), por que então investir mais recursos em divulgação, ao invés de concentrar esforços no redesign da interface? De

acordo com os resultados obtidos no Quadro 50, devido a falácia da conjunção era exatamente esta decisão de projeto que seria equivocadamente tomada pelo designer.

Apesar dos dados não serem plenamente conclusivos, considera-se que sejam relevantes evidências da existência da falácia da conjunção impactando negativamente o processo de design. Bem como a proposta de correção, pois ao posicionar todas as alternativas como conjuntivas, obteve-se resultados substancialmente relevantes, uma vez que houve redução de 80,4% para 17,2% das respostas equivocadas.

O fato de posicionar as alternativas com todas as opções contendo eventos conjuntivos foi um modo de experimentar o quanto o balanceamento dos eventos isolados se faz importante para um correto raciocínio do avaliador. Considerando que o experimento não possui a finalidade de “curar” o processo cognitivo que ocasiona a falácia da conjunção, mas de contornar ao mesmo afirmando aos designers que balanceiem a combinação de características dos seus protótipos ao apresentá-lo a um avaliador.

Cabe aqui o alerta de que o efeito oposto à correção sugerida nesta tese também deve funcionar. Assim como nos Grupos 1 e 2, obteve-se otimismo infundado às alternativas conjuntivas, mesmo que descrevessem problemas que não tinham conexão com o produto descrito, em posse deste conhecimento um designer pode apresentar alternativas de projeto combinando os eventos que ele deseja ter maior aceitação. Gerando assim resultados positivos exatamente para as alternativas que ele desejar, sendo estes resultados induzidos pela falácia da conjunção intencionalmente posicionada.

Esta discussão desdobra-se, obviamente, para todos os outros domínios onde a falácia da conjunção tem seu efeito mapeado. Talvez por este otimismo atribuído opções com maior número (não necessariamente qualidade) de combinações apresentadas, que técnicas de vendas avançadas tenham sido moldadas para que o máximo de sentidos do potencial comprador seja estimulado, objetivando a aquisição do produto frente a concorrentes (ZIGLAR, 2002). Assim, o comprador que viu uma propaganda do produto e experimentou o mesmo, tendência a comprá-lo mais do que um produto que ele apenas viu. Daí os programas de *test-drive* nas concessionárias automotivas, *promoters* com amostras grátis de perfumes, dentre outros.

Trazendo a discussão novamente ao campo do design, estas mesmas técnicas de vendas (ocasionadas pela falácia da conjunção ou apenas como agentes ampliadores da

mesma) podem ser mais uma vez usadas pelo projetista na apresentação de sua proposta de produto à diretoria da empresa. Uma vez que este caracteriza-se um processo similar ao de venda.

Porfim, considera-se com base nos resultados obtidos e nas discussões vinentes, que a falácia da conjunção interfere no projeto de produtos, mas que balanceando as combinações nas alternativas é possível suprimir a falha cognitiva que ocasiona o problema. Convém considerar ainda que a falácia da conjunção pode ser intencionalmente manipulada para gerar otimismo por certas alternativas, desde que posicione-se nas mesmas alguma combinação de eventos.

4.2 Representação do Espaço Amostral

Uma das propostas desta tese é introduzir à metodologia de design um modelo estatístico de verificação da quantidade de avaliadores necessários para testar um produto (ou protótipo de produto). De modo que a opinião dos mesmos ofereça alguma consistência em relação a opinião dos usuários finais.

Devido as discussões já em amadurecimento no campo da avaliação de usabilidade de websites, realizamos na *Seção 3.5.2* uma simulação do modelo de proposto por Nielsen (2006; 1993). O mesmo afirma que 5 usuários treinados são suficientes para encontrar 70% dos erros de usabilidade que afetariam os usuários finais de um website (NIELSEN, 2006; NIELSEN & LANDAUER, 1993). Desconsiderando que o público alvo desse tipo de produto pode chegar na ordem dos milhões de usuários, estariam estes 5 refletindo a opinião dos usuários finais?

A presente tese propõe a aplicação de um método estatístico para validar o modelo de Nielsen (2006) e, ao menos na simulação exposta na *Seção 3.5.2*, ele se demonstrou cientificamente incoerente (FIGUEIRÔA et al., 2012). A oscilação de resultados reportadas por Faulkner (2003), Perfetti & Landesman (2001), dentre outros, que obtinham resultados entre 35 e 80% dos erros encontrados com os primeiros 5 usuários foram associadas ao baixo nível de confiança estatística do modelo em questão.

Sendo assim, a pergunta presente nas avaliações de usabilidade em websites – extensível a todas as outras rotinas de avaliação de artefatos presentes no design – permanece: quantos avaliadores são necessários testando um produto para conhecer a

opinião dos usuários finais? Acredita-se que em posse da informação referente a opinião (ou comportamento) dos usuários finais ante um produto pode-se aumentar as chances de sucesso do designer em projetar algo especificamente para os mesmos, reduzindo os efeitos do acaso no resultado final.

A seguir, o detalhamento dos experimentos conduzidos visando introduzir validação científica quanto número de usuários necessários na avaliação de um produto, durante o processo de design do mesmo.

4.2.1 Procedimentos Metodológicos

Considerando a proposta de Nielsen (2006), que afirma que 5 avaliadores seriam suficientes para encontrar os erros que afetariam a usabilidade por parte dos usuários finais de um website, aplicamos um método estatístico de Triola (1999) na verificação desta conjectura.

Utilizando o mesmo procedimento da simulação realizada *Seção 3.5.2*, considerou-se a “opinião dos avaliadores” como a “amostra”. O objetivo é verificar com que nível de confiança a amostra reflete os “usuários finais do produto”, neste caso denominados “espaço amostral”.

Para escolha do método de validação estatística, considerou-se que a maioria dos designers projetam produtos destinados a serem fabricados em massa (carros, computadores, relógios e demais objetos de uso pessoal), ou para serem usados em massa (games, websites e demais softwares). Assim sendo, o método mais adequado de cálculo do nível de confiança da amostra é a que aborda o conceito do “espaço amostral infinito” (TRIOLA, 1999). Assim selecionado quando o número de pessoas participantes na amostra é inferior a 5% do número que corresponde ao espaço amostral.

Em outras palavras, dificilmente um designer terá como entrevistar 5% dos milhões de usuários finais de um *tablet*, carro ou website. Assim sendo, o cálculo do número de usuários (amostra) pode ser abordado de acordo com o detalhado na Figura 14, presente na *Seção 3.5.2*.

Outros autores classificam o “espaço amostral infinito” como todo aquele compreendido por mais de 100.000 elementos (POCINHO & FIGUEIREDO, 2004). Em ambos os casos,

quer seja por quantidade elementos ou por porcentagem da amostra em relação ao total de elementos, nosso experimento se enquadra na amostragem por “espaço amostral infinito”.

Para realização deste procedimento, considerou-se ainda os tipos de amostragem disponíveis (POCINHO & FIGUEIREDO, 2004). Estas técnicas de coleta de amostras dividem-se em 2 grupos: probabilísticas e não-probabilísticas.

Classificam-se por tipos de amostragem probabilísticas:

- **Amostragem aleatória simples**

Consiste em elaborar uma lista numérica de elementos de onde se tira, com a ajuda de uma tabela de números aleatórios, uma série de números para constituir a amostra.

- **Amostragem aleatória estratificada**

É uma variante da amostra aleatória simples. Esta técnica consiste em dividir a população alvo em subgrupos homogêneos chamados “estratos” e a seguir seleciona-se de forma aleatória uma amostra de cada “estrato”. A amostragem aleatória estratificada é utilizada quando a população inteira é reconhecida por certas características precisas, tais como a idade, o sexo, a incidência de uma condição de saúde, tudo isto para assegurar a melhor representatividade possível.

- **Amostragem em grupos**

Consiste em retirar de forma aleatória os elementos por cachos (grupos) em vez de unidades. É útil quando os elementos da população estão naturalmente por grupos ou quando não é possível obter uma listagem de todos os elementos da população-alvo.

- **Amostragem sistemática**

Consiste quando existe uma lista ordenada de elementos da população. Esta técnica consiste de uma quantidade (denominada “K”) de elementos dessa lista, sendo o primeiro elemento da amostra retirado ao acaso.

De acordo com Pocinho & Figueiredo (2004), classificam-se ainda por amostragens “não-probabilísticas”:

- **Amostragem accidental ou de conveniência**

É formada por sujeitos facilmente acessíveis, que estão presentes num determinado local e momento preciso. Exemplo: pessoas hospitalizadas. Um investigador pode ter acesso a uma unidade hospitalar para constituir uma amostra de pacientes hospitalizados.

- **Amostragem por cotas**

Idêntica à amostragem aleatória estratificada diferenciando-se desta apenas pelo fato dos sujeitos não serem escolhidos aleatoriamente no interior de cada estrato ou de cada grupo.

- **Amostragem de seleção racional**

Tem por base o julgamento do investigador para constituir uma amostra de sujeitos em função do seu carácter típico. Exemplo: Estudo de caso específico com indivíduos com uma determinada disfunção visual.

- **Amostragem por redes**

Toma-se por base redes sociais, amizades e contatos próximos. Quando o investigador encontra sujeitos que satisfazem os critérios escolhidos pede-lhes que indiquem outras pessoas de características similares. Basicamente é uma seleção auxiliada por uma rede de indicações.

Dadas as condições de nossa pesquisa, utilizou-se da disponibilidade de turmas de design (elementos naturalmente agrupados) que assimilariam com maior precisão as heurísticas de Nielsen e considerou-se a ausência de uma listagem de todos os usuários finais do produto. Caracterizando, portanto, o procedimento como “amostragem por grupos”, sob cálculo de “espaço amostral infinito”.

Em síntese, este procedimento metodológico consiste especificamente em:

- Realizar testes preliminares de usabilidade reais utilizando o modelo de Nielsen (2006; 1993) com suas 10 heurísticas¹² e grupos 5 usuários avaliando um website. O formulário aplicado para avaliação é o padrão para testes deste tipo, conforme ilustrado no APÊNDICE A;

¹² As 10 heurísticas de Jakob Nielsen encontram-se detalhadas na Seção 2.4.6 desta tese. Em síntese, são elas: 1- Feedback; 2 - Relação análoga entre o sistema e o mundo real; 3 - Controle do usuário e liberdade; 4 - Consistência e padrões; 5 - Prevenção de erro; 6 - Reduzir sobrecarga de memória; 7 - Flexibilidade e eficiência de uso; 8 - Estética e design minimalista; 9 - Diagnóstico de erros; 10 - Ajuda e documentação.

- Calcular com base no modelo de “espaço amostral infinito” e “amostragem por grupos” a quantidade de usuários necessários para ter uma opinião estatisticamente válida (POCINHO & FIGUEIREDO, 2004; TRIOLA, 1999;);
- Verificar se o número de usuários apontado pelo cálculo supracitado revela um nível de confiança satisfatório em representar a opinião do espaço amostral (usuários finais);
- Com base neste experimento, expor um modelo com o qual designers podem embasar a quantidade de avaliadores de seus protótipos, atendendo o “requisito científico”¹³ mencionado por Burdek (2006) para que designers e indústria possam estabelecer um diálogo comum.

4.2.2 Resultados do Estudo de Caso Sobre o Espaço Amostral

Um total de 40 alunos de design, na disciplina de Interface Homem-Máquina da Faculdade Marista Recife e da UNIBRATEC foram introduzidos às 10 heurísticas de usabilidade de Nielsen. Posteriormente, os mesmos avaliaram a usabilidade do website <http://www.ufpe.br> (Figura 23).

¹³ Burdek (2006) destaca que o advento das metodologias de design na HfG em Ulm, Alemanha, nos anos 1970, se deu pela necessidade de embasar seus projetos diante das exigências da indústria. Para tal, iniciaram-se o uso de métodos científicos na metodologia de design, como meio de provar tanto a geração dos projetos quanto a avaliação do porque eles seriam bem sucedidos quando lançados no mercado.

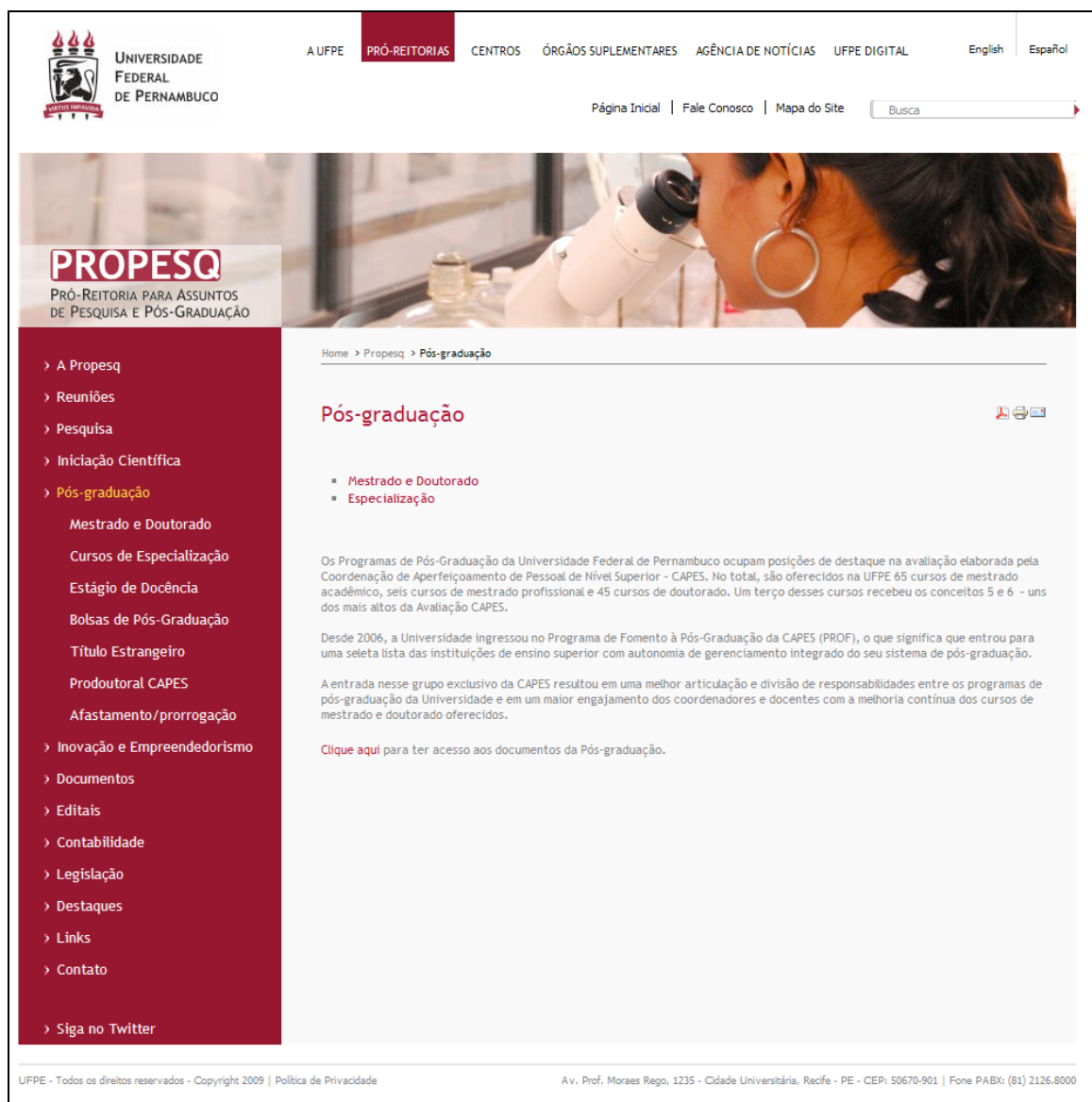


Figura 23 – Interface do website avaliado

Retornando, ao final do processo, a seguinte quantidade de erros presentes no Quadro 55. Estes resultados individuais de avaliação foram sequencialmente posicionados, em grupos de 5.

Grupo	ID Avaliador	Erros (Porcentagem dos erros)	Desvio-padrão (DP), e nível de confiança ($Z_{\alpha/2}$)	Grupo	ID Avaliador	Erros (Porcentagem dos erros)	Desvio-padrão (DP), e nível de confiança ($Z_{\alpha/2}$)
A	1	50 (98%)	DP:	E	21	11 (22%)	DP:
	2	51 (100%)	28,673		22	8 (16%)	24,352
	3	21 (41%)	$Z_{\alpha/2}$:		23	23 (45%)	$Z_{\alpha/2}$:
	4	20 (39%)	0,78		24	43 (84%)	0,92
	5	21 (41%)	(56%)		25	17 (33%)	(64%)
B	6	27 (53%)	DP:	F	26	26 (51%)	DP:
	7	10 (20%)	27,702		27	25 (49%)	16,849
	8	50 (98%)	$Z_{\alpha/2}$:		28	25 (49%)	$Z_{\alpha/2}$:
	9	25 (49%)	0,81		29	6 (12%)	1,33
	10	43 (84%)	(58%)		30	10 (20%)	(81%)
C	11	22 (43%)	DP:	G	31	14 (27%)	DP:
	12	23 (45%)	13,832		32	25 (49%)	14,161
	13	11 (22%)	$Z_{\alpha/2}$:		33	29 (57%)	$Z_{\alpha/2}$:
	14	25 (49%)	1,62		34	9 (18%)	1,58
	15	33 (65%)	(90%)		35	19 (37%)	(88%)
D	16	50 (98%)	DP:	H	36	3 (6%)	DP:
	17	15 (29%)	28,479		37	19 (37%)	15,005
	18	21 (41%)	$Z_{\alpha/2}$:		38	25 (49%)	$Z_{\alpha/2}$:
	19	50 (98%)	0,79		39	21 (41%)	1,49
	20	30 (59%)	(57%)		40	21 (41%)	(86%)

* A margem de erro (E) para cálculo do nível de confiança ($Z_{\alpha/2}$) foi de 10 pontos.

** Para maiores informações sobre cálculo de nível de confiança ($Z_{\alpha/2}$), ver Quadro 33

Quadro 54 – Erros de usabilidade reportados

O grupo “A” da supracitada tabela é composto pelos alunos considerados por parte dos professores como os de maior *expertise* e por isso foi o que encontrou a maior quantidade de erros (51). Sendo, exclusivamente para efeito de cálculo deste experimento, este número assumido como o máximo de erros do website em análise e o grupo “A”, portanto, como grupo de controle (âncora) dos demais.

Convém destacar que mesmo com uma alta taxa de erros encontrados, a amostragem do grupo “A” retornou um baixo grau de confiança (56%). Considera-se apenas valores acima de 92,5% como relevantes para representar a opinião do espaço amostral, uma vez que esta confiança evolui exponencialmente (Observar Quadro 33).

Nestes termos, é possível ainda visualizar ao final desta coleta de resultados a seguinte síntese de erros encontrados por grupos de 5 usuários avaliando o website (Quadro 56).

Grupo	A (controle)	B	C	D	E	F	G	H
Erros encontrados	100%	98%	65%	98%	84%	51%	57%	49%

Quadro 55 – Variação de porcentagem de erros de usabilidade encontrados

Visando introduzir o cálculo de validação estatística desta amostragem, tomemos por exemplo ainda o grupo “A”. Ao seguir o raciocínio materializado nas equações da Figura 24 abaixo, é possível deduzir que a quantidade de pessoas opinando sobre o website em análise seria de 32,8 (arredondando temos 33 pessoas).

$$n_A = \left(\frac{Z_{\alpha/2} \cdot \sigma_A}{E} \right)^2$$

$$n_A = \left(\frac{2 \cdot 28,673}{10} \right)^2$$

$$n_A = 32,8$$

Figura 24 – Cálculo estatístico para obtenção do número de avaliadores

Verificando a consistência desta informação, utilizamos os grupos “B”, “C”, “D”, “E”, “F”, “G” e “H”, apanhando 33 amostras de avaliação (do avaliador número 6 ao 38). O desvio-padrão desta amostragem é 24,542 e os demais cálculos referentes à verificação do nível de confiança da mesma encontram-se na Figura 25 abaixo.

$$n_A = \left(\frac{Z_{\alpha/2} \cdot \sigma_A}{E} \right)^2$$

$$33 = \left(\frac{Z_{\alpha/2} \cdot 24,542}{10} \right)^2$$

$$Z_{\alpha/2} = \frac{10 \sqrt{33}}{24,542}$$

$$Z_{\alpha/2} = 2,34$$

Figura 25 – Cálculo do nível de confiança da amostra para 33 avaliadores

Com base nestes números é possível afirmar que com 33 usuários avaliando o website obtem-se uma amostragem que reflete o que seria encontrado pelo espaço amostral de usuários deste produto com um nível de confiança estatístico de 98% (Interpretando $Z_{\alpha/2} = 2,34$ pelo Quadro 33).

4.2.3 Considerações Referentes ao Estudo de Caso Sobre o Espaço Amostral

As discussões desta seção direcionam-se a 2 temas pertinentes aos resultados obtidos pelo estudo de caso sobre o espaço amostral: as 10 heurísticas de Nielsen e; o cálculo do número de avaliadores.

Relativo às heurísticas não restaram dúvidas da razão de sua aceitabilidade ante ao mercado. Elas proveram, em todas as amostras, a detecção de problemas na interface que afetariam a usabilidade do website e, se este for considerado pelo designer do produto em questão como um fator crítico de sucesso, a proposta de Nielsen trás contribuições relevantes.

Não obstante, haja vista as discussões presentes neste seguimento sobre a quantidade de avaliadores, boa parte delas encabeçadas por Perfetti & Landesman (2001), utilizou-se oportunamente este âmbito de discussão para introduzir a verificação da quantidade de avaliadores. Apesar das heurísticas de Nielsen serem relevantemente eficientes para o melhoramento da interface e, portanto, procederem bastante aceitas pelo mercado, a quantidade afirmada por este autor de 5 usuários serem suficientes para representar o espaço amostral foi rebatida pelo modelo estatístico apresentado.

De acordo com os resultados obtidos, determinar um número fixo de avaliadores para todos os casos não procede cientificamente. Sendo necessário obter dados preliminares de uma amostragem menor (no caso o grupo “A”) para então calcular qual seria uma quantidade plausível de participantes na avaliação.

Seja através da simulação na Seção 3.5.3 ou no estudo de caso com avaliações reais de usabilidade, fora possível apresentar aos designers um modelo estatístico com o qual os mesmos podem justificar para sua indústria – ou seus dirigentes – a quantidade de avaliadores de um protótipo. Sendo possível balizar a busca de respostas de perguntas como: “Quantos jogadores devem experimentar um game em desenvolvimento para saber o quanto este produto está atendendo os critérios estabelecidos para satisfazer o público-alvo?”

Convém destacar que o método apresentado para este cálculo não é conclusivo, representando apenas uma das várias ferramentas metodológicas com embasamento acadêmico disponíveis que podem auxiliar o designer nesta tarefa. Sendo estas introduções e experimentações fortes tendências de desdobramento da presente tese deste que, considera-se, ter apresentado um primeiro passo neste sentido, ampliando a fronteira do conhecimento do design nesta demanda da avaliação de artefatos.

5 CONCLUSÕES

A presente tese levantou a hipótese de que a avaliação de artefatos em design configura-se como um evento aleatório e, portanto, herda os problemas decorrentes da aleatoriedade condensados em fenômenos e falácias. Estima-se ter alcançado o objetivo da pesquisa em demonstrar esta tese, bem como o de apresentar propostas de solução para os 5 problemas pervasivos na avaliação de artefatos oriundos da aleatoriedade através de revisão literária e estudos de casos.

Este capítulo apresenta uma seção de conclusões específicas referentes aos estudos de casos e em seguida uma síntese de conclusões referentes aos 5 problemas classificados pela tese como decorrentes da aleatoriedade e interferindo na avaliação de artefatos em design. Por fim, apresentam-se propostas de desdobramentos da presente pesquisa.

5.1 Conclusões Específicas Sobre os Resultados dos Estudos de Casos

Os estudos de caso conduzidos permitiram verificar a hipótese de que fenômeno da falácia da conjunção impacta o processo de design, na fase de avaliação de artefatos, conforme indicado por Ferreira (2008). Este impacto pode ser negativo no sentido de levar designers a problemas que não tem ligação direta com o protótipo em análise, como ocorrido em nosso experimento. Bem como, com base nos resultados obtidos, é possível considerar que em posse do conhecimento a respeito deste fenômeno o designer pode manipular a avaliação do seu artefato. Uma vez que os avaliadores tendenciarão a optar pelas alternativas que apresentarem propostas conjuntivas, ao invés de propostas isoladas. A correção para este erro é proposta pela presente tese ao balancear todas as alternativas com eventos conjuntivos. Este procedimento reduziu o impacto da falácia da conjunção de 80,4% para 17,2% dos casos aqui desenvolvidos. Esta correção deve ser válida para outros domínios onde este fenômeno foi identificado, uma vez que seguiu-se o procedimento de Tversky & Kahneman (1984), o mesmo aplicado noutros seguimentos.

Ainda nos estudos de caso, foi possível demonstrar como embasar estatisticamente a quantidade de usuários que avaliam um produto, demonstrando como calcular este número com validade estatística. Convém destacar que a quantidade de avaliadores necessários pode variar em cada caso. Com base nos cálculos de espaço amostral apresentados é possível justificar a quantidade de avaliadores diante da indústria, que de acordo com Burdek (2006) exige métodos tangíveis de projeto. Assim demonstrou-se viável a

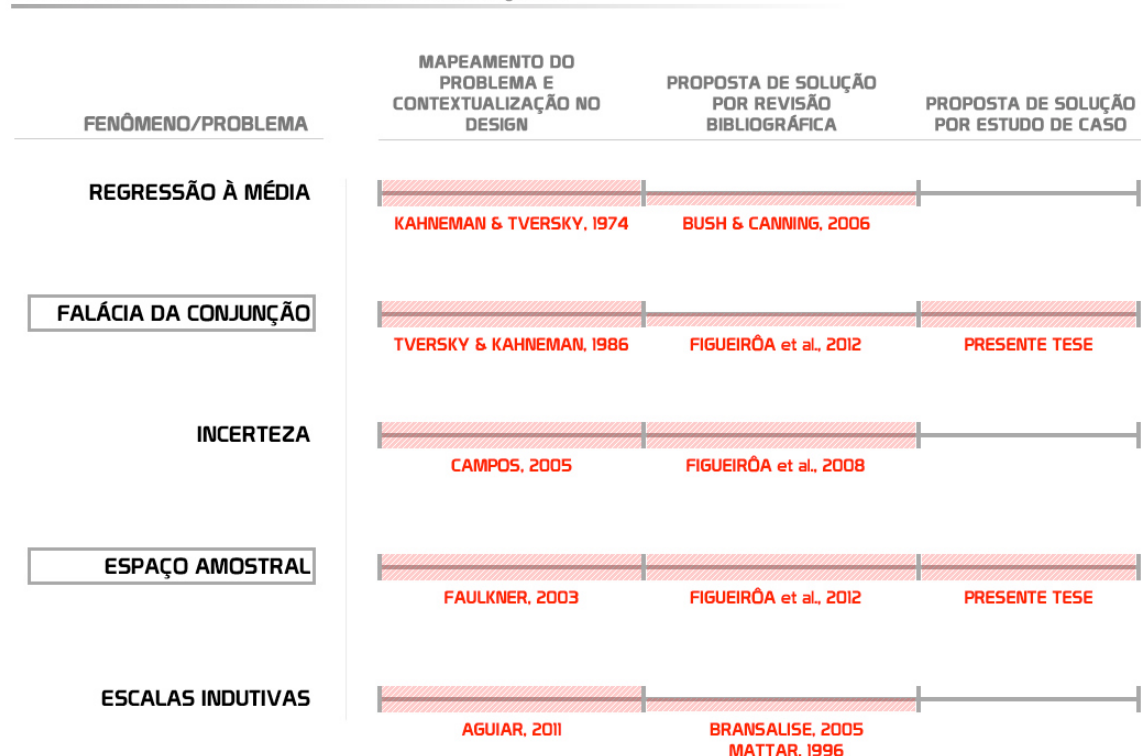
combinação do procedimento de cálculo para representação do espaço amostral por níveis de confiança combinado com um procedimento de avaliação de produto (TRIOLA, 1999). Na tese, utilizaram-se oportunamente as discussões sobre esta quantidade de usuários avaliadores já vigentes no campo dos testes de usabilidade. Ao combinar o cálculo de representação do espaço amostral de Triola (1999) com os testes de usabilidade de Nielsen (2006; 1993) foi possível ilustrar aos designers como justificar a quantidade de avaliadores.

Estes experimentos conduzidos na verificação do espaço amostral de usuários no campo da usabilidade utilizando-se do método de Nielsen (2006) conduziu-nos a algumas conclusões sobre o mesmo. As 10 heurísticas de Jakob Nielsen reafirmaram-se úteis no design, especialmente na fase de avaliação de artefatos, detectando problemas cruciais dos de interface que precisavam ser corrigidos. Estas heurísticas revelaram-se um importante mecanismo de detecção destes problemas e em momento algum dos experimentos a tese encontrou argumentos para se contrapor as mesmas. Todavia, as afirmações de Nielsen de que 5 usuários são suficientes para encontrar 70% dos erros de usabilidade que afetariam o espaço amostral (usuários finais) demonstrou-se cientificamente inconsistente. Ou seja, apesar método proposto pelo referido autor ser mercadologicamente atraente (pois reflete um baixo custo na execução pela pouca quantidade de avaliadores), ele não apresentou nos experimentos realizados consistência científica. Enfim, a presente tese não contrapõe-se as 10 heurísticas de Nielsen (nem se propõe a isso), mas apresenta evidências de que a quantidade de avaliadores usados neste método precisa passar por um crivo estatístico, conforme a proposta de cálculo apresentada.

5.2 Resumo das Propostas de Solução

O Quadro 56 abaixo apresenta o mapeamento dos problemas e as soluções mapeadas ou desenvolvidas pela presente tese, no que tange a redução da aleatoriedade na avaliação de artefatos. Em seguida, detalha-se com maior profundidade as conclusões referentes a cada uma das soluções.

RESUMO DO DESENVOLVIMENTO DE SOLUÇÕES



Quadro 56 – Projeção das soluções apresentadas na tese

Ainda sobre o Quadro 56, as propostas de solução por revisão bibliográfica apresentadas para regressão à média, falácia da conjunção e espaço amostral encontram-se com menor espessura ilustrando que estas propostas apresentam menos densidade de conteúdo em relação as demais ou não propõem uma solução plenamente aplicável. Em outras palavras, considera-se que estas propostas de solução carecem de desdobramentos para validá-las.

Em síntese, o designer irá reduzir a aleatoriedade no processo de design ao seguir as recomendações desenvolvidas por pesquisas acadêmicas detalhadas abaixo.

5.2.1 Regressão à Média

- **Problema:** Randomicidade dos eventos regidos pela aleatoriedade (KAHNEMAN & TVERSKY, 1974).
- **Solução:** Observar o histórico de desempenho de mercado de um produto ajudará o designer a indentificar qual resultado ele está tendenciando ao regredir à média (BUSH & CANNING, 2006).

Apesar de Bush & Canning (2006) propor uma avaliação mercadológica de produtos diante da regressão à média, estudos de caso mais específicos em design são demandados.

5.2.2 Falácia da Conjunção

- **Problema:** Avaliadores de projeto tendem a uma natural propensão ao erro pela falácia da conjunção, quando diante de questões aleatórias (TVERSKY & KAHNEMAN, 1986).
- **Solução:** Balancear as alternativas todas as alternativas de resposta com eventos conjuntivos, ou todas com eventos isolados, inibirá o avaliador de cair na falácia da conjunção (FIGUEIRÔA et al., 2012).

A proposta de Figueirôa et al (2012) de balancear as alternativas apresentadas ao avaliador com eventos conjuntivos foi realizada via estudo de caso pela tese e bem sucedida.

5.2.3 Incerteza

- **Problema:** A incerteza e a aleatoriedade são diretamente proporcionais (CAMPOS, 2005).
- **Solução:** Modelar a incerteza através do Lateo ajudará o designer a evitar os resultados contra-intuitivos da Teoria das Evidências, reduzindo a aleatoriedade (FIGUEIRÔA et al., 2008).

O problema da incerteza identificado e desenvolvido por Campos (2005) através do Lateo foi experimentado especificamente no design por Figueirôa et al. (2012), explicitando como reduzir a aleatoriedade proveniente da incerteza no projeto de produto.

5.2.4 Espaço Amostral

- **Problema:** Como definir a quantidade de avaliadores durante a avaliação de um artefato, visando melhor conhecer a opinião dos usuários finais do produto (FAULKNER, 2003).

- **Solução:** Justificar a quantidade de avaliadores participantes pelo do procedimento de cálculo de grau de confiança da amostra (FIGUEIRÔA et al., 2012).

A proposta de Figueirôa et al. (2012) de embasar a quantidade de avaliadores de um produto de modo que a opinião dos espaço amostral de usuários fosse melhor representada, reduzindo o acaso no processo de design, foi desenvolvida pela presente tese com um estudo de caso.

5.2.5 Escalas Indutivas

- **Problema:** O uso de escalas psicométricas pode estar induzindo os resultados, aumentando o acaso da avaliação (AGUIAR, 2011).
- **Solução:** Não numerar as escalas de diferencial semântico e nem concentrar todos os adjetivos negativos ou positivos em um único lado da escala (BRANDALISE, 2005; MATTAR, 1996).

Apesar de Brandalise (2005) e Mattar (1996) apresentarem propostas para reduzir a indução das escalas de diferencial semântico, muitas das escalas que podem estar induzindo os resultados demonstradas em Aguiar (2011) demandam por soluções.

5.3 Desdobramentos

Nos experimentos sobre a falácia da conjunção, procedeu-se da forma mais fiel possível ao procedimento de Tversky & Kahneman (1983). Inclusive pelo detalhe dos participantes serem estudantes universitários. Compete agora prolongar o presente estudo, reaplicando este procedimento ao entrevistar designers profissionais experientes. Verificando assim se a falácia continua procedendo com os mesmos resultados ao comparar os dados obtidos entre profissionais e estudantes.

A presente tese abre ainda linhas de pesquisas para múltiplos procedimentos na avaliação de artefatos, principalmente no que tange os problemas levantados neste domínio que, por limitações de uma tese de doutorado, não puderam todos receber um estudo de caso específico. Por exemplo, no que tange a regressão à média se faz necessário demonstrar matematicamente a fórmula desenvolvida por Galton (1886) que prova o

fenômeno, ao reaplicá-la em histórias de vendas de produtos, com estudos de caso em design. Permitindo uma visualização parametrizada da ocorrência deste fenômeno.

A ilustração de Aguiar (2011) sobre as diferenças de resultados obtidos sobre o mesmo produto, quando utilizavam-se escalas diferentes carecem de mais investigações sobre o quanto cada escala induz os resultados, permitindo uma comparação mais clara. O que a presente tese fez foi expor este problema apresentando um estado da arte sobre o mesmo o posicionando-o como um efeito ampliador da aleatoriedade inerente a este tipo de atividade (avaliação). Assim sendo, demanda-se por procedimentos que permitam gerar recomendações para evitar esta falácia no processo de design.

O uso do procedimento baseado em espaço amostral infinito apresentado nesta tese pode ter maior consistência se for aplicado em outros seguimentos da avaliação de artefatos que não sejam especificamente usabilidade ou Nielsen (2006; 1993).

Neste sentido, é possível sugerir novas pesquisas. Concluindo que com os dados de estados da arte e resultados de experimentos obtidos nesta tese que a avaliação de artefatos demanda de mais estudos científicos que objetivem reduzir (não eliminar) o impacto da aleatoriedade no processo de design.

Em síntese, como sugestões de novos trabalhos, propõe-se: abordar o cálculo da regressão à média em cases de design; a observação da falácia da conjunção ao entrevistar designers experientes (comparando com os resultados obtidos de estudantes de design desta tese); a verificação do procedimento de cálculo de espaço amostral proposto nesta tese em outras especialidades do design que não seja a usabilidade, como nas escalas CPSS para avaliação de criatividade e; a uma verificação mais profunda da proposta de Aguiar (2011) ao verificar um número maior de escalas de avaliação de artefatos, comparando seus respectivos resultados e apontando as possíveis contradições na avaliação de um mesmo produto por escalas diferentes.

REFERÊNCIAS

- AGUIAR, B. *Uso de Escalas de Autorrelato na Análise de Games*. Dissertação (Mestrado em Design). Recife: Universidade Federal de Pernambuco, Departamento de Design, 2011.
- ALEXANDER, C. *A Pattern Language*. Nova Iorque: Oxford University Press, 1977.
- ARANHA, M. L. A; MARTINS, M. H. P. *Filosofando - Introdução à Filosofia*. São Paulo: Moderna, 1990. p. 443.
- ASHBY, F. G.; ISEN, A. M.; TURKEN, A. U. A Neuropsychological Theory of Positive Affect and Its Influence on Cognition. *Psychological Review*, n. 106, p. 529-550, 1999.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR14724: informação e documentação – trabalhos acadêmicos - apresentação*. Rio de Janeiro, 2001.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR6023: informação e documentação - referências - elaboração*. Rio de Janeiro, 2000.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR12225: títulos de lombada*. Rio de Janeiro, 1992.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR6027: sumário*. Rio de Janeiro, 1989.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR6024: numeração progressiva das seções de um documento*. Rio de Janeiro, 1989.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR6028: resumos*. Rio de Janeiro, 1990.
- BAXTER, M. *Projeto de Produto: Guia Prático para o Desenvolvimento de Novos Produtos*. 3ª Ed. São Paulo: Editora Edgard Blücher Ltda., 2011
- BELFORT, R. *TR²UE: Design para Experiência de Usuário. Aplicando Princípios de XU na Avaliação de Jogos Casuais*. Dissertação (Mestrado em Design). Recife: Universidade Federal de Pernambuco, Departamento de Design, 2011.
- BESEMER, S.P. & O'QUIN, K. Analyzing creative products: refinement and test of a judging instrument. *Journal of Creative Behavior*, 20, p. 115 – 26. 1986.

BESEMER, S. P. Creative product analysis matrix: Testing the model structure and a comparison among products – Three novel chairs. *Creativity Research Journal*, 11, p. 333-346, 1998.

BOMFIM, G. A. Metodologia para desenvolvimento de projeto. Campina Grande: Editora Gráfica Rodier Ltda, 1984.

BRANDALISE, L. T. Modelos de medição de percepção e comportamento: Uma revisão. In: Laboratório de Gestão, Tecnologia e Informação – UFSC, Florianópolis, 2005. Disponível em: <www.lgti.ufsc.br/brandalise.pdf>. Acesso em: 01 fev 2012.

BRANDÃO, E. R. *Publicidade On-line, Ergonomia e Usabilidade: O Efeito de Seis Tipos de Banner no Processo Humano de Visualização do Formato do Anúncio na Tela do Computador e de Lembrança da sua Mensagem*. Dissertação (Mestrado em Design). Rio de Janeiro: PUC-Rio, Departamento de Artes & Design, 2006.

BÜRDEK, B. História teoria e prática do design de produtos. São Paulo: Blücher, 2006.

BUSH, H.F; CANNING, M.D. Regression Towards The Mean Versus Efficient Market Hypothesis: An Empirical Study. *Journal of Business & Economics Research*. Vol. 4, N.12, 2006.

CAMPOS, F. Uma Extensão à Teoria Matemática da Evidência. Tese (Doutorado em Ciência da Computação). Recife: Universidade Federal de Pernambuco, Centro de Informática, 2005.

CAMPOS, F.; CAVALCANTE, S. An extended approach to dempster-shafer theory. *IEEEIRI-2003*, Las Vegas, USA, 2003.

CNNMONEY. *Top 20 Most Admired Companies*. America's Most Admired Companies 2008. 2008. Disponível em: <http://money.cnn.com/galleries/2008/fortune/0802/gallery.mostadmired_top20.fortune/index.html>. Acesso em: 20 fev 2012.

CORREIA, A. *O Papel das Técnicas Explícitas de Geração de Alternativas no Projeto de Produtos*. Dissertação (Mestrado em Design). Recife: Universidade Federal de Pernambuco, Departamento de Design, 2010.

CSÍKSZENTMIHÁLYI, Mihály. Finding Flow. The Psychology of Engagement With Every Day Life. New York: HarperCollins Publishers, 1997.

DAVID, F. N. Gods, Games and Gambling: A History of Probability and Statistical Ideas. Nova York: Dover, 1998. p. 24-6.

DE-BONO, E. El pensamiento lateral: Manual de creatividad. Paidós Plural: México, 1998.

DE-BONO, E. Criatividade levada a sério: como gerar ideias produtivas através do pensamento lateral. São Paulo: Pioneira, 1967.

FAULKNER, L. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments and Computers* 2003, 35 (3), p 379- 383. 2003.

FERREIRA, V. R. M. Psicologia Econômica: Estudo do Comportamento Econômico e da Tomada de Decisão. Rio de Janeiro: Elsevier, 2008.

FERREIRA, J. V. R. Gestão Ambiental - Análise do ciclo de vida dos produtos. Viseu: (ESTV/IPV) Instituto Politécnico de Viseu, 2004.

FERREIRA, A. B. H. Dicionário Aurélio Eletrônico Século XXI. Versão 3.0, Rio de Janeiro: Editora Nova Fronteira; Lexicon Informática Ltda, nov. 1999.

FERSON, S. et al. Constructing probability boxes and Dempster-Shafer structures. Sand Report. Unlimited Release, jan 2003.

FIGUEIREDO, T. Coleção de Bolsas de Mão para Primavera Verão 2012: Utilizando Metodologia de Design para Gerar Esboços Criativos. TCC (Graduação). Universidade Federal de Pernambuco, Departamento de Design, 2012.

FIGUEIRÔA, D. L., CHAVES, W., CAMPOS, F., CORREIA, W. Representação do Espaço Amostral de Usuários na Avaliação de Artefatos Digitais. In: P&D Design 2012 | 10º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design, 2012, São Luis. *Anais do 10º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design*. São Luis: UFMA, 2012.

FIGUEIRÔA, D. L., CAMPOS, F., CORREIA, W. A Falácia da Conjunção no Processo de Design. In: P&D Design 2012 | 10º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design, 2012, São Luis. *Anais do 10º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design*. São Luis: UFMA, 2012.

FIGUEIRÔA, D. L., CAMPOS, F., CORREIA, W. Avaliação Mercadológica de Artefatos no Processo de Design: Uma Abordagem Estratégica. In: P&D Design 2012 | 10º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design, 2012, São Luis. *Anais do 10º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design*. São Luis: UFMA, 2012.

FIGUEIRÔA, D. L. *Design de Artefatos Digitais Baseados em Padrões e Plataformas*. Dissertação (Mestrado). Universidade Federal de Pernambuco, Departamento de Design, 2009.

FIGUEIRÔA, D. L., VILAR NETO, E., NUNES, I., CORREIA, A., ALVES, H., CAMPOS, F., NEVES, A. M. M. Técnicas Matemáticas aplicadas ao Processo de Design. In: P&D Design 2008 | 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design, 2008, São Paulo. *Anais do 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design*. São Paulo: SENAC, 2008.

FIGUEIRÔA, D. L., CAMPOS, F., NEVES, A. M. M. O Paradigma do Projeto baseado em Plataforma aplicado ao Web Design. In: P&D Design 2008 | 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design, 2008, São Paulo. *Anais do 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design*. São Paulo: SENAC, 2008.

FIGUEIRÔA, D. L., CAMPOS, F., CHAVES, W., NEVES, A. M. M. Game Design Baseado em Plataforma. In: P&D Design 2008 | 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design, 2008, São Paulo. *Anais do 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design*. São Paulo: SENAC, 2008.

FIGUEIRÔA, D. L., CAMPOS, F., NEVES, A. M. M. O Paradigma do Projeto Baseado em Plataformas Aplicado ao Game Design. In: SBGAMES 2007 | VI Simpósio Brasileiro de Jogos para Computador e Entretenimento Digital. São Leopoldo, 7 a 9 nov. 2007. *Anais do VI Simpósio Brasileiro de Jogos para Computador e Entretenimento Digital*. São Leopoldo: UNISINOS, 2007.

FOX, C. R.; BIRKE, R., Forecasting trial outcomes: lawyers assign higher probabilities to possibilities that are described in greater details. *Law and Human Behavior*, 26, n.2, p.159-73, abr 2002.

FREUND, J. E. Estatística Aplicada: Economia, Administração e Contabilidade. 11ª Ed. Porto Alegre: Bookman, 2006.

G1. *Relembre os sucessos e fracassos da história da Apple*. Tecnologia e Games, 2011. Disponível em: < <http://g1.globo.com/tecnologia/noticia/2011/08/relembre-os-sucessos-e-fracassos-da-historia-da-apple.html>>. Acesso em: 20 fev 2012.

GALTON, F. Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland*, Vol. 15, p. 246-263, 1886.

HOLLAND, B. K. What Are the Chances? Voodoo Deaths, Office Gossip, and Other Adventures in Probability. Baltimore: Johns Hopkins University Press, 2002. p.24.

IDG NOW. *Nem tudo foi sucesso: 5 fracassos da Apple sob comando de Steve Jobs*. Computação Pessoal, 2011. Disponível em: <http://idgnow.uol.com.br/computacao_pessoal/2011/10/06/nem-tudo-foi-sucesso-5-fracassos-da-apple-sob-comando-de-steve-jobs/>. Acesso em: 20 fev 2012.

IG SÃO PAULO. *Galeria de fotos: Sucessos e fracassos da empresa de Steve Jobs*. Tecnologia, 2011. Disponível em: <<http://tecnologia.ig.com.br/galeria-de-fotos-sucessos-e-fracassos-da-empresa-de-steve-jobs/n1597259556137.html>>. Acesso em: 20 fev 2012.

ISAACSON, W. *Steve Jobs – A Biografia*. São Paulo: Editora Schwarcz Ltda, 2010.

JORDAN, P. *An Introduction to Usability*. Londres: Taylor and Francis, 1998.

KOTLER P.; KELLER K., *Administração de Marketing*, 12 edição, São Paulo: Pearson Prentice Hall, 2006.

KAHNEMAN, D. *Daniel Kahneman - Autobiography*. Nobelprize.org, 2002. Disponível em: <http://www.nobelprize.org/nobel_prizes/economics/laureates/2002/kahneman-autobio.html>. Acesso em: 20 fev 2012.

KAHNEMAN, D.; TVERSKY, A. Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 80 , No. 4, p.273-251, 1973.

KANITZ, S. *Índice Bovespa 1963 até hoje*. 2008. Disponível em: <<http://estatisticas.melhores.com.br/2008/12/indice-bovespa-1963-at%C3%A9-hoje.html>>. Acesso em: 20 fev 2012.

KLINE, M *Mathematical Thought from Ancient to Modern Times*. Nova Iorque: Oxford University Press, 1972.

KRUG, S. *Não me faça pensar*. São Paulo: Editora Market Books, 2001.

- KUROSU, M.; KASHIMUR, K.. Apparent usability vs. inherent usability: Experimental analysis on the determinants of the apparent usability. *Conference on Human Factors and Computing Systems*. New York: ACM Press, 292–9, 1995.
- LAW, E. et al. Understanding, Scoping and Defining User eXperience: A Survey Approach. In: *CHI'09 proceedings*. Boston: CHI'09 proceedings, 2009.
- LEVIN, Jack. Estatística Aplicada a Ciências Humanas. 2ª Ed. São Paulo: Editora Harbra Ltda, 1987.
- LIKERT, Rensis. A technique for the measurement of attitudes. *Archives of psychology*. No. 140, p. 1-55, 1932.
- LÖBACH, B. Desenho industrial: bases para a configuração dos produtos industriais. São Paulo. Edgar Blücher Ltda, 2001.
- MACDONALD, N. What is web design?. Hove: Rotovision, 2003.
- MANO, A.; CAMPOS, J. C. Aplicação de um Cognitive Walkthrough – Estudo de Caso. 1ª *Conferência Nacional em Interação Pessoa-Máquina*. Lisboa, jul 2004.
- MARTINS, J. S. *Decisões de Investimentos: Uma Abordagem Comportamental no âmbito da Indústria Moveleira de Lagoa Vermelha*. Monografia (Bacharelado em Administração). Porto Alegre: Universidade Federal do Rio Grande do Sul, Escola de Administração, 2011.
- MATTAR, F. N. Pesquisa de Marketing. São Paulo: Atlas, 1996.
- MELO, A. *Avaliação Ergonômica de Websites para Uso por Pessoas com Deficiência Visual Através da Interface Humano-Computador*. Dissertação (Mestrado em Design e Ergonomia). Recife: Universidade Federal de Pernambuco, Departamento de Design, 2010.
- MELO, E. V. V. *Aplicação de Técnicas de Exploração do Espaço Criativo ao Design de Jogos Digitais*. Dissertação (Mestrado em Design). Recife: Universidade Federal de Pernambuco, Departamento de Design, 2008.
- MIRANDA, F. & MORAES, A. Avaliação da Interface de um Site de Comércio Eletrônico Através da Técnica Avaliação Cooperativa. In: *Congresso Internacional de Ergonomia e Usabilidade, Design de Interfaces e Interação Humano-Computador*, 2. 2003, Rio de Janeiro: LEUI – Laboratório de Ergonomia e Usabilidade de Interfaces em Sistemas Humano-Tecnologia, 6 e 7, p. 8, jun 2003.

MLODINOW, L. O Andar do Bêbado – Como o Acaso Determina Nossas Vidas. Rio de Janeiro: Jorge Zahar, 2009.

MONTEIRO, R. A. *E-Lateo – Combinação e Representação do Conhecimento*. Dissertação (Mestrado). Universidade Federal de Pernambuco, Centro de Informática, 2009.

NIELSEN, J. Quantitative Studies: How Many Users to Test?. Useit.com, Alertbox, 2006. Disponível em: http://www.useit.com/alertbox/quantitative_testing.html. Acesso em: 17 jun 2012.

NIELSEN, J.; TAHIR, M. Homepage: Usabilidade. Rio de Janeiro: Campus, 2002.

NIELSEN, J.; LANDAUER, T.K. A mathematical model of the finding of usability problems. In: *Proceedings of ACM INTERCHI'93 Conference*. Amsterdam: ACM Press, p. 206-213, 1993.

NIELSEN, J. Usability Engineering. Boston: Academic Press, 1993.

NORMAN, D. A. Design emocional: por que adoramos (ou detestamos) os objetos do dia-a-dia. Rio de Janeiro: Rocco, 2008.

PERFETTI, H.; LANDESMAN, L. *Eight is Not Enough*. User Interface Engineering. 2001. Disponível em: <http://www.uie.com/articles/eight_is_not_enough>. Acesso em: 20 fev 2012.

PIRAUÁ, J. *TAE.web-u: A Construção de uma Técnica para Avaliação Estética de Websites com Usuários*. Dissertação (Mestrado em Design). Recife: Universidade Federal de Pernambuco, Departamento de Design, 2007.

POCINHO, M.; FIGUEIREDO, J.P. Estatística e Bioestatística. Lisboa: Madeira, 2004.

PORTO, M. P. *Investigação da aplicabilidade da TAE.web-u (Técnica de Avaliação Estética de Websites por Usuários) no contexto dos desenvolvedores de websites*. Dissertação (Mestrado em Design). Curitiba: Universidade Federal do Paraná. Setor de Ciências Humanas, Letras e Artes, 2009.

REIS, F. O. A. *O Ciclo de Vida do Produto e as Estratégias de Mercado na Gestão de Marcas – Sandálias Havaianas – Um Estudo de Caso*. Monografia (Bacharelado em Engenharia de Produção). Juiz de Fora: Universidade Federal de Juiz de Fora, 2007.

REVISTA VEJA. *Os Números Mágicos de Harry Potter. A série de sete livros e oito filmes que alcançou níveis históricos de vendas.* Cinema. 2011. Disponível em: <<http://veja.abril.com.br/noticia/celebridades/os-numeros-magicos-de-harry-potter>>. Acesso em : 20 fev 2012.

ROCHA, D. *Peripatécnica: Uma Técnica para Avaliação Estética em Design.* Dissertação (Mestrado em Design). Recife: Universidade Federal de Pernambuco, Departamento de Design, 2007.

ROTO, V. et. al. User Experience Evaluation in Nokia. Florence: *CHI'08 proceedings*, 2008.

SHAFER, G. A Mathematical Theory of Evidence. [S.l.]: Princeton University Press, 1976. ISBN: 0-691-08175-1.

SILVA, C. R. L. Economia e mercados: introdução à economia. São Paulo: Saraiva, 2001

SKINNER, B. F. Science and Human Behavior. Nova Iorque: MacMillian, 1953.

TEIXEIRA, E. A. S.; MORAES, A. Avaliação Cooperativa da Interface de Sites Hiper mídias Focados na “Banda Larga”. In: *Congresso Internacional de Ergonomia e Usabilidade, Design de Interfaces e Interação Humano-Computador*, 3. 2004, Rio de Janeiro: LEUI – Laboratório de Ergonomia e Usabilidade de Interfaces em Sistemas Humano-Tecnologia, p. 6, 27 e 28 mai 2004.

THURSTONE, L. L. Attitudes can be measured. *American Journal of Sociology*, 33, 529-54, 1928.

TRACTINSKY, N.; SHOVAL-KATZ, A.; IKAR, D. What is Beautiful is Usable. *Interacting with Computers* (13), p. 127-145, 2000.

TRACTINSKY, N. Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues, *CHI 97 Conference Proceedings*, Nova York: ACM Press, p. 115-12, 1997.

TRIOLA, M. F. Introdução à Estatística. 7ª Ed. Rio de Janeiro: LTC, 1999.

TVERSKY, A.; KAHNEMAN, D. Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90 , 293-315, 1983.

TWYMAN, M. Using pictorial language: a discussion of the dimensions. In: DUFTY, M. et. al. *Designing usable text.* Orlando: Academic Press, 1985.

VARIAN, H. R. Microeconomia. Rio de Janeiro: Elsevier, 2003.

VELUDO-DE-OLIVEIRA, T. M. . Escalas de mensuração de atitudes: Thurstone, Osgood, Stapel, Likert, Guttman, Alpert. Administração On Line (São Paulo), São Paulo, v. 2, 2001.

VIRZI, R. A. Refining the test phase of usability evaluation: how many subjects is enough? In: *Human Factors*, 457-468, 1992.

WYKES, A. Doctor Cardano: Physician Extraordinary. London: Frederick Muller, 1969.

YAHOO! FINANCES. *AAPL Basic Chart*. Apple Inc (AAPL), 2012. Disponível em: <<http://finance.yahoo.com/q/bc?s=AAPL&t=my>>. Acesso em: 20 fev 2012.

ZIGLAR, Z. Os Segredos da Arte de Vender, 6ª Edição. Rio de Janeiro: Record, 2002.

ANEXO A – Pesquisa da Falácia da Conjunção – iCopa



UFPE – Departamento de Design

Pesquisa Doutorado

Amostra N _____

Leia atentamente cada enunciado e numere de 1 a 6 as alternativas a seguir, sendo:

1 = o mais provável

6 = o mais improvável

Deve ser marcado um número para cada alternativa.

Designers estão projetando o iCopa, um aplicativo para smartphones que apresenta simultaneamente várias informações relevantes sobre a copa do mundo. O protótipo do produto tem recebido muitas reclamações dos seus avaliadores, que normalmente se perdem na sua interface.

- () O aplicativo apresenta uma interface complexa, com muitos botões
 - () O aplicativo parece um jornal digital
 - () O aplicativo tem pouca divulgação
 - () O aplicativo não apresenta uma interface com temática esportiva
 - () O aplicativo tem pouca divulgação e apresenta uma interface complexa, com muitos botões
 - () O aplicativo é composto por uma tecnologia ultrapassada
-

APÊNDICE A – Ficha de Avaliação Heurística de Usabilidade

Extraído como parte da pesquisa de André Melo (2010).

UFPE – Universidade Federal de Pernambuco		Qtd Heurísticas desobedecidas	
URL: www.ufpe.br/ufpenova		Sem importância: 0	
Inspetor: 1		Superficial: 0	
Ambiente: Windows XP		Simples: 17	
Navegador: FireFox		Grave: 20	
Data da realização: 16/12/2009		Catastrófico: 0	
Local do problema de usabilidade	Problema	Heurística desobedecida	Criticidade
Todas as páginas	Não é dividida em seção, não tem zona de salto.	Atalhos, minimizar a sobrecarga de memória do usuário, saídas claras marcadas	Grave
Página de notícias	A seção de notícias não é indicada, você não sabe se realmente esta nela.	Saídas claras marcadas, ajuda e documentação	Grave
Todas as páginas	Não tem recurso para listar um curso, tem um combo, mas não tem opção por cursos.	Atalhos, consistência	Grave
Todas as páginas	Os banners não possuem descrição textual.	Feedback, falar a linguagem do usuário, diálogos simples e naturais	Grave
Todas as páginas	Não possuem tags de atalho.	Atalhos	Grave
Todas as internas	Nas internas só se mantém o primeiro menu o resto da estrutura muda.	Minimizar a sobrecarga de memória do usuário, Consistência	Simples
Página principal	Aparece uma mensagem de javascript que não faz sentido.	Boas mensagens de erro	Grave
Todas as páginas	Não há áudio descrição da marca.	Feedback, falar a linguagem do usuário, diálogos simples e naturais	Simples
Todas as páginas	Um link chamado links links sem descrição do seja.	Feedback, falar a linguagem do usuário, diálogos simples e naturais	Grave
Todas as páginas	Demora muito grande para achar informações sobre cursos da graduação.	Feedback, consistência	Simples
Em algumas páginas	Em algumas páginas não há opção de voltar.	Atalhos, prevenir erros, saídas claramente demarcadas	Simples
Todas as páginas	Os links poderiam ter uma descrição	Feedback,	Simples

	mais detalhada.	consistência	
Todas as páginas	Links diferentes aparecem nas internas.	Minimizar a sobrecarga de memória do usuário	Simple
Todas as páginas	Não há lógica na navegação.	Diálogos simples e naturais, falar a linguagem do usuário	Grave
Todas as páginas	Faltam cabeçalhos para as seções.	Atalhos	Simple
Todas as páginas	O botão de busca não possui descrição textual.	Feedback, falar a linguagem do usuário, diálogos simples e naturais	Simple
Todas as páginas	Testei o fale conosco e não tive resposta.	Feedback, boas mensagens de erro, prevenir erros	Grave