

UNIVERSIDADE DO ESTADO DE SANTA CATARINA – UDESC
CENTRO DE ARTES – CEART
PROGRAMA DE PÓS-GRADUAÇÃO – PPGDESIGN

GUSTAVO L. R. JOTA

**APLICAÇÃO DE INDICADORES DE USABILIDADE NO DESENVOLVIMENTO
ÁGIL DE SOFTWARE PARA A MELHORIA DA SATISFAÇÃO DOS USUÁRIOS.**
KPIsUX para a melhoria do NPS.

FLORIANÓPOLIS
2022

GUSTAVO L. R. JOTA

**APLICAÇÃO DE INDICADORES DE USABILIDADE NO DESENVOLVIMENTO
ÁGIL DE SOFTWARE PARA A MELHORIA DA SATISFAÇÃO DOS USUÁRIOS.**

KPIsUX para a melhoria do NPS

Dissertação apresentada como requisito parcial para obtenção do título de mestre em "Métodos para os Fatores Humanos" pelo PPGDesign - Programa de Pós-Graduação em Design da Universidade do Estado de Santa Catarina – Udesc.
Orientador: Prof. Dr. Elton Moura Nickel.

FLORIANÓPOLIS

2022

JOTA, GUSTAVO LOPES RODRIGUES
APLICAÇÃO DE INDICADORES DE USABILIDADE NO
DESENVOLVIMENTO ÁGIL DE SOFTWARE PARA A
MELHORIA DA SATISFAÇÃO DOS USUÁRIOS. : KPIsUX
para a melhoria do NPS. / GUSTAVO LOPES RODRIGUES
JOTA. -- 2022.
95 p.

Orientador: Elton Nickel
Dissertação (mestrado) -- Universidade do Estado de
Santa Catarina, Centro de Artes, Design e Moda, Programa
de Pós-Graduação em Design, Florianópolis, 2022.

1. Indicadores. 2. Usabilidade. 3. NPS. 4. Ágil. 5.
Ergonomia. I. Nickel, Elton. II. Universidade do Estado de
Santa Catarina, Centro de Artes, Design e Moda, Programa
de Pós-Graduação em Design. III. Título.

GUSTAVO L. R. JOTA

**APLICAÇÃO DE INDICADORES DE USABILIDADE NO DESENVOLVIMENTO
ÁGIL DE SOFTWARE PARA A MELHORIA DA SATISFAÇÃO DOS USUÁRIOS:**

KPIsUX para a melhoria do NPS

Dissertação apresentada como requisito parcial para obtenção do título de mestre em "Métodos para os Fatores Humanos" pelo PPGDesign - Programa de Pós-Graduação em Design da Universidade do Estado de Santa Catarina – Udesc.
Orientador: Prof. Dr. Elton Moura Nickel.

BANCA EXAMINADORA

Prof. Dr. Elton Moura Nickel
PPGDesign - UDESC

Membros:

Prof. Dr. Eugênio Andrés Díaz Merino
POSDESIGN – UFSC
PPGEP - UFSC

Profa. Dra. Gabriela Botelho Mager
PPGDesign - UDESC

Florianópolis, 30 de Agosto de 2022.

Aos colegas de profissão que tanto
batalharam e batalham para a
consolidação da profissão em nosso Brasil.

AGRADECIMENTOS

Primeiramente, gostaria de agradecer à minha amada esposa Karine e ao meu filho Pedro que apoiaram este momento importante da minha vida. Ao meu fiel amigo canino Lego que hoje descansa em paz, mas que se manteve firme acompanhando cada progresso deste estudo até a defesa. Ainda, às equipes da Dígitro (Octavio Carradore, Isabella Veloso e Fábio Fernandes) e da Way 2 (Alex Mariano e Jefferson Kobs) que auxiliaram na validação e calibração do experimento de campo, mas em especial à Hotjar e seu CEO Erik Näslund que compreenderam o trabalho proposto e priorizaram seu backlog de produto, permitindo a calibração e validação do experimento e do processo proposto. Finalmente, mas não por último, ao meu colega de magistério e primeiro orientador, Dr. Prof. Alexandre Reis, que me guiou no processo de exploração inicial deste mestrado e, ainda, ao Dr. Prof. Elton Nickel que, nesta segunda e final fase, agrega valor ao ciclo de eterno aprendizado quando orienta a mim, seu outrora professor.

*“Em Deus nós confiamos; todos os outros
devem trazer dados” (DEMIG, [1993])*

RESUMO

A pesquisa propõe um processo que viabilize o projeto de ergonomia e fatores humanos no desenvolvimento de software utilizando metodologia de desenvolvimento Ágil. Os projetos de software sob esta abordagem possuem como principais características a organização adaptativa, entregas evolutivas, curtos prazos e resolvem problemas complexos que, muitas vezes, não são inteiramente conhecidos antes do efetivo início do projeto. Por seu turno, os métodos tradicionais de projetos de produto industriais e gráficos apresentam planejamento anterior ao seu início, inclusive na definição de atividades e fases de aprovação (stage-gate), escopo do problema definido e são geralmente orientados pelo PMI (Project Management Institute). Este estudo propõe a correlação de causa e efeito entre cinco métricas de performance de usabilidade (Sucesso Binário, Algum Sucesso, Erro, Tempo Médio e Eficiência no Tempo) calculadas com dados de uso de uma revista digital (toda a população de usuários) usando abordagem MPA (multi-page application), em relação à satisfação do usuário calculada por amostragem (pesquisa espontânea NPS ou net promoter score), sob uma abordagem aderente ao Ágil (Scrum) pelo período de um ano, cobrindo quatro ciclos de desenvolvimento (Setup e três Releases), em um experimento de campo com a revista digital em produção, seguindo as diretrizes da Lei Geral de Proteção de Dados (LGPD). A conclusão é a de que o constructo KPIsUX pode ser empregado por meio de um processo contínuo de medição, avaliação, decisão e desenvolvimento (MADD) em projetos de software, conforme proposição inicial.

Palavras-chave: Indicadores. Usabilidade. NPS. Ágil. Ergonomia.

ABSTRACT

The research presents a process to enable human factors and ergonomic project in software development using agile methodology. Software projects using agile approach have key characteristics of adaptive organization, evolutionary deliveries, short deadlines and it solve complex problems that are often not fully known before the launch of the endeavor. Unlike industrial and graphic design projects that require planning prior to their start, including the definition of activities, phase approvals, and an outlined problem to be resolved, under a 'stage gate' approach and addressed via PMI (Project Management Institute). This study proposes a correlation of cause and effect between usability performance metrics (Binary Success, Partial Success, Errors, Average Time and Task Success and Time) calculated through usage data of a digital magazine (the entire user population) using MPA structure (multi-page application), in relation to user satisfaction calculated by sampling (NPS spontaneous survey or net promoter score), under an Agile (SCRUM) approach for a period of one year, covering successive 4 development cycles (Setup and 3 Releases) in a field experiment with a digital magazine in production, following the guidelines of Brazilian's LGPD (Law General Data Protection). Concluding that the construct, KPIsUX, can be used through continuous a process of measurement, evaluation, decision and development (MADD) in software according to the initial proposition.

Keywords: Indicators. Usability. NPS. Agile. Ergonomics.

LISTA DE ILUSTRAÇÕES

Figura 1 - Modelo prescritivo para produto industrial.....	15
Figura 2 - " <i>Framework</i> " Scrum para o desenvolvimento de software	21
Figura 3 - Diagrama de aplicação de testes de usabilidade.....	42
Figura 4 - Mapa de decisão para análise de dados contínuos	46
Figura 5 - Categorização de respostas NPS em uma escala 0 a 10	50
Figura 6 - NPS médio entre diversas indústrias (benchmark).	51
Figura 7 - Diagrama de funcionamento do Google Analytics.	58
Figura 8 - Gráfico de fluxo de usuários em um website.	60
Figura 9 - Imagem de uma pesquisa NPS.	61
Figura 10 - Open Refine transformando dados por agrupamento.	62
Figura 11 - Google Data Studio e gráficos para representação de dados.	63
Figura 12 - Representação das macro fases de Setup e Releases no Scrum.	72
Figura 13 - Representação global do processo (KPIsUX).....	73
Figura 14 - Painel de Acompanhamento dos KPIsUX e NPS.	82
Figura 15 - Análise entre as métricas do KPIsUX e o NPS.....	84

LISTA DE GRÁFICOS

Gráfico 1 - Percentual de acertos por tarefa.	53
Gráfico 2 - Representação em barras dos três níveis de sucesso.	54
Gráfico 3 - Representação de Erros em forma de barras.	55
Gráfico 4 - Barras representando a média do tempo nas atividades.	56
Gráfico 5 - Quadro representando o número ideal de passos (três) para a realização de uma tarefa.	57
Gráfico 6 - Gráfico de barras horizontais de esforço no tempo.	57

LISTA DE QUADROS

Quadro 1 - Perspectiva de valores do Ágil	17
Quadro 2 - Comparação macro de modalidades de computação	25
Quadro 3 - Comparação de características entre modelos prescritivos e ágeis.....	27
Quadro 4 - Métodos comportamentais e cognitivos e seus resultados.	38
Quadro 5 - Vantagens e desvantagens de cada um dos métodos.	39
Quadro 6 - Análise de nível de performance.....	44
Quadro 7 – Compilação de questionários padronizados	48
Quadro 8 – Métricas quantitativas de performance mais usadas na usabilidade	52
Quadro 9 – Relatórios disponíveis na plataforma Google Analytics.	59
Quadro 10 – Desdobramento das macros fases até as ferramentas.....	71

LISTA DE TABELAS

Tabela 1 - Coleta de dados e análise no período de 12 meses.	67
Tabela 2 - Parâmetros resultantes do Setup para cálculo dos KPIsUX e definição de metas	75
Tabela 3 – Resultados dos KPIsUX e NPS.....	81
Tabela 4 – Resultados da correlação entre métricas e o NPS.	84

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
GLPD	Lei geral de Proteção dos Dados
ETL	Extração, transformação e carga (de dados)
NPS	Indicador de satisfação (Net Promoter Score)
SPA	Aplicação ou software construído em uma única página (<i>Single-Page Application</i>)
MPA	<i>Aplicação ou software construído em diversas páginas (Multi-Page Application)</i>
SITE	Aplicação publicada na internet
WEB	Abreviação de internet
PDP	Processo de desenvolvimento de produto
KPIS	Indicadores chaves de sucesso
KPISUX	Constructo de indicadores chaves de sucesso em usabilidade
HTTP	Protocolo de transferência de hipertexto (páginas) na internet
ADS	Publicidade online na internet
BI	Aplicação visualização e análise de dados (análise de negócios ou <i>Business Intelligence</i>)

SUMÁRIO

1	CONSIDERAÇÕES INICIAIS	15
2	INTRODUÇÃO	19
3	REVISÃO BIBLIOGRÁFICA.....	32
4	MATERIAIS E MÉTODOS.....	67
5	RESULTADOS E DISCUSSÕES	77
6	CONCLUSÃO E CONSIDERAÇÕES FINAIS	86
	REFERÊNCIAS.....	89
	GLOSSÁRIO.....	91
	ANEXO A – ESTRUTURA DE DADOS DO NPS	92
	ANEXO B– FORMULÁRIO DE PESQUISA ANÔNIMA COM USUÁRIOS.....	93
	ANEXO C– PARECER CONSUBSTANCIADO DO CEP (RESUMIDO)	94

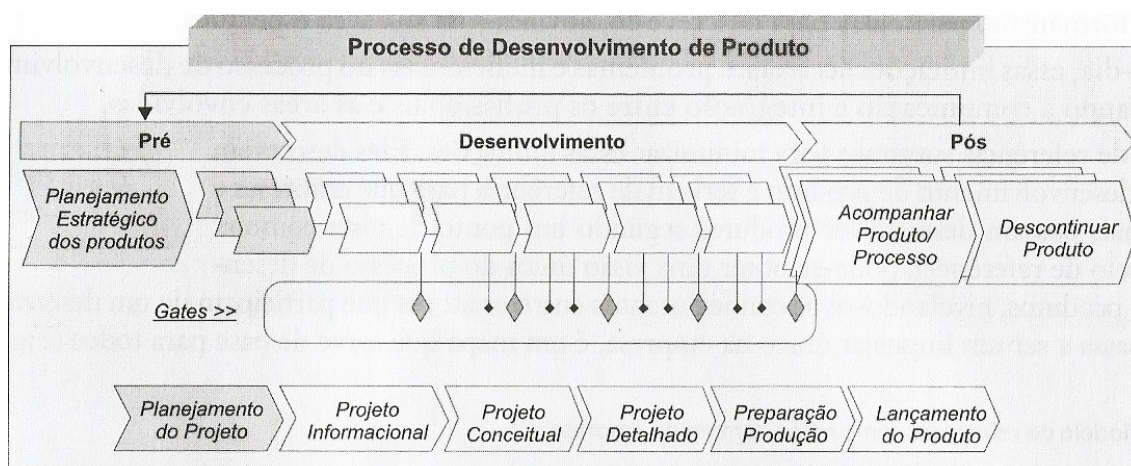
1 CONSIDERAÇÕES INICIAIS

A presente dissertação de mestrado, parte do Programa de Pós-Graduação em Design (PPGDesign) da Universidade do Estado de Santa Catarina (Udesc), possui ênfase em Organização e Fatores Humanos, e propõe-se a aplicar métricas quantitativas de usabilidade do produto e satisfação do usuário no processo de desenvolvimento de software.

O estudo aborda um processo de análise contínua e automatizada de métricas de *eficiência* e *eficácia* de um software (web) em conjunto com pesquisa de satisfação do usuário, com o objetivo de subsidiar, medir e aferir o desempenho contínuo do software ao longo de ciclos de desenvolvimento. Para tanto, é utilizada uma metodologia que possui características distintas das metodologias empregadas em projetos de design, ergonomia e fatores humanos.

Segundo Jota & Maynardes (2018) os modelos de desenvolvimento prescritivos são a base da metodologia tradicional de desenvolvimento de produtos industriais (e design em geral), em contrapartida, os modelos adaptativos são uma abordagem crescente no desenvolvimento de software.

Figura 1 - Modelo prescritivo para produto industrial



Fonte: ROZENFELD et al., 2006, p. 44.

Os modelos prescritivos têm como principal característica o detalhamento das atividades previamente à execução do projeto, sendo orientados a partir de

um modelo de referência. Estes propõem uma estrutura na qual as atividades são divididas em fases (evolutivas e porventura paralelamente), separadas por avaliações com sequência linear, conforme se observa na Figura 2 (Rozenfeld *et al.*, 2006), para problemas previamente conhecidos e dimensionáveis.

Já, os modelos de desenvolvimento adaptativos (ou Ágeis), segundo Sutherland & Schwaber (2017) são abordagens ou estruturas (do inglês *framework*) dentro dos quais as pessoas (equipe) tratam e resolvem problemas complexos e adaptativos conforme seu surgimento. É característica fundamental a autonomia para alterar diversos parâmetros, como a composição da equipe, ferramentas (técnicas ou outros processos), escopo do produto (ou serviço), dentre outros, ao longo dos ciclos de desenvolvimento incremental.

A abordagem Ágil opera dentro de 12 valores, ou princípios: [1] priorização da satisfação do cliente; [2] receptividade a mudanças de valor a qualquer momento; [3] entrega de software funcional dentro de semanas a meses; [4] trabalho conjunto entre pessoal de negócio e desenvolvimento; [5] projeto ao redor de time motivado; [6] troca de informações efetiva-eficiente face a face ; [7] medida de progresso é software funcionando; [8] processos ágeis são sustentáveis e capazes de se manter ao longo do tempo; [9] atenção à excelência e a arquitetura do projeto aprimoram a agilidade; [10] maximização do resultado de entrega e não do esforço; [11] times auto organizados geram as melhores arquiteturas, requisitos e projetos; [12] avaliação constante da eficiência para ajustes (adaptabilidade).

Quadro 1 - Perspectiva de valores do Ágil

Maior priorização de valores	Menor priorização de valores
Indivíduos e interações	Processos e ferramentas
Software em funcionamento	Documentação abrangente
Colaboração com o cliente	Negociação de contratos
Resposta a mudanças	Seguir um plano

Fonte: Adaptado pelo autor. Beck *et al.* (2001) apresentam a priorização (trade-off) de valores do Ágil em contrapartida a valores de abordagens tradicionais (stage-gate ou waterfall).

Este estudo, portanto, propõe e instrumentaliza, por meio de métricas de usabilidade (estruturadas no formato de indicadores de resultado), o uso de um processo contínuo de medição, avaliação, decisão e desenvolvimento para subsidiar as atividades de profissionais de design, desenvolvimento de software, marketing e outros envolvidos, nas suas atividades de desenvolvimento de produtos digitais, no contexto, ou sob uma abordagem Ágil. Sendo a dissertação estruturada em 7 capítulos, incluindo 4 anexos, referências e glossário de termos, visto que duas áreas de conhecimento são exploradas no documento (desenvolvimento de software com a ergonomia e fatores humanos), conforme:

O primeiro capítulo aborda as considerações e apresentações iniciais. O segundo introduz o tema, sua problematização, contexto, escopo do problema, hipótese, variáveis, objetivo e justificativa e, ainda, a delimitação do estudo devido à natureza do experimento de campo utilizado.

O terceiro capítulo detalha os modelos de desenvolvimento prescritivo e Ágil, especificamente o Scrum que é utilizado posteriormente no experimento de campo, modalidade de computação, proteção de dados, métricas quantitativas de usabilidade cognitiva, práticas de pesquisa de usabilidade com usuário, indicadores de uso e acesso (software) e satisfação do usuário.

Por seu turno, o quarto capítulo aborda o experimento de campo, apresenta o processo proposto e as diversas técnicas, ferramentas e instrumentos utilizados.

No quinto capítulo são apresentados os resultados parciais do experimento de campo e sua conclusão.

O sexto capítulo descreve o teste da hipótese por meio de correlação de *Spearman* entre os indicadores de usabilidade (variável independente) e a satisfação dos usuários (variável dependente). Observe-se que tais variáveis interagem evolutivamente a partir de ações do desenvolvimento (variável interveniente) ao longo de sucessivos ciclos de desenvolvimento (variável antecedente).

No sétimo e último capítulo, o presente trabalho é concluído com as considerações finais e propostas para trabalhos futuros.

2 INTRODUÇÃO

A partir da década de 80 (LORENZ, 1986), a indústria Japonesa adotou o artifício de acelerar os ciclos de lançamento de produtos, e, posteriormente, observa-se que diversas indústrias, incluindo as de software, adotaram essa mesma estratégia. Como consequência, imprimiu-se maior competição e pressão sobre as equipes de desenvolvimento das empresas. Atualmente, uma vasta gama de produtos físicos e software são lançados anualmente e o que se percebe é que aos mesmos são incorporados, cada vez mais, funcionalidades de outros produtos e até mesmo serviços. Assim, e por exemplo, aos telefones celulares ou *tablets* foram incorporados gravadores de voz, câmeras fotográficas, tocadores de música, álbum de fotografia, relógios, despertadores, agendas, lanternas e, mais recentemente, equipamentos como dispositivos de navegação GPS (global positional system), dentre outros. Sobre estes equipamentos, incluindo computadores pessoais (PCs), são utilizados websites e aplicações (softwares) que incorporaram ou substituíram a vasta gama de serviços e empresas, como agências de turismo, bancas de jornais, supermercados, bancos, dentre diversos outros.

O pesquisador Cusumano (1997), ao conduzir estudos na Microsoft, observou semelhanças nas abordagens de desenvolvimento de produtos nas empresas de base tecnológica dos EUA na década de 90. Competidores como Yahoo e Netscape, motivados pela necessidade de lançar produtos complexos a nível global e com grande velocidade, adotaram práticas semelhantes, como o desenvolvimento iterativo, projeto simultâneo, vasta gama de protótipos, entregas incrementais, intensa troca de informações, divisão da arquitetura do software em times, dentre outras. Anteriormente, na década de 80, softwares como MS-DOS, Word e Excel possuíam em torno de 10 desenvolvedores e a complexidade de dezenas de milhares de linhas de código. Os produtos sequer eram traduzidos para outras línguas, e a língua inglesa (Norte Americana) era mantida como padrão. Já o Windows 95, introduzido em 1995, possuía cerca de 450 pessoas em sua equipe de desenvolvimento e 11 milhões de linha de código, o que consumia três anos para seu lançamento. Este modelo de

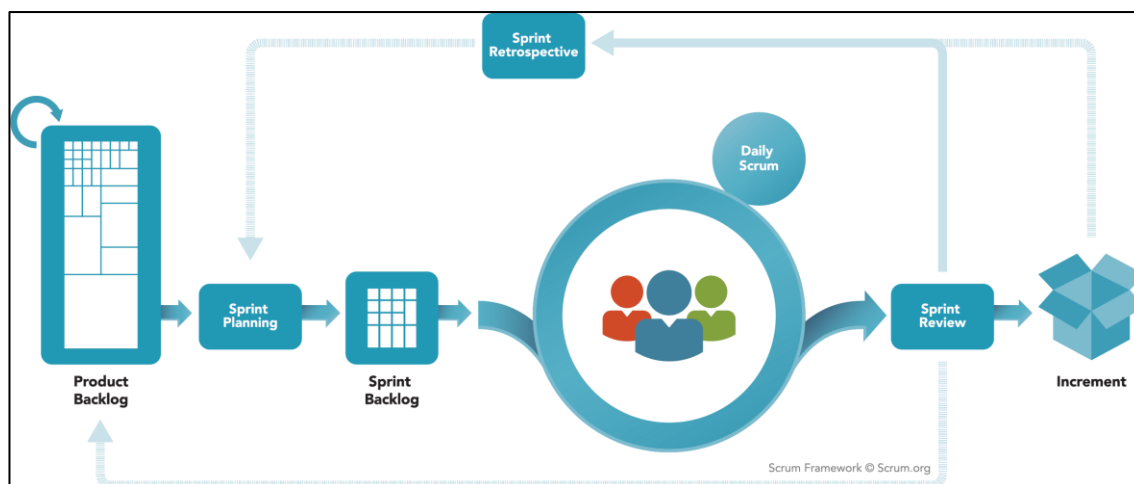
desenvolvimento adaptativo, posteriormente intitulado como Ágil (ou Agile), unificou práticas semelhantes (SCHABER *et al.* 2001) e, desde então, convive com o modelo "Cascata" (ou *Waterfall*), que é baseado em um modelo prescritivo de desenvolvimento, ainda largamente utilizados na indústria de manufatura, bens de consumo, bens de capital, dentre outras. A partir desta teoria, são derivados os modelos de referência propostos por Rozenfeld *et al.* (2006) e Baxter (1998), dentro dos quais está incluindo o projeto de design, ergonomia e fatores humanos.

Relativo aos diversos "frameworks" Ágeis, o mais conhecido é o Scrum, popularizado por Schaber & Shuterland (1995), que argumentam que este permite maior velocidade e flexibilidade no desenvolvimento de produtos mais complexos e disruptivos em que os objetivos ainda não são inteiramente conhecidos, ou mesmo estejam em fase de compreensão.

Observado sob a ótica do desenvolvimento tradicional de produtos, é um ambiente de desenvolvimento. Segundo os autores, esta estrutura tem como principais características ser leve, de fácil entendimento, mas, contudo, difícil de dominar. Neste, inclusive é possível a utilização de vários processos ou técnicas, visto que não é uma estrutura rígida ou definitiva. Entretanto, por se tratar de um ambiente de desenvolvimento, necessita de outras ferramentas, técnicas e modelos para auxiliar na sua gestão e condução. Sua organização consiste em uma equipe (time de Scrum) normalmente abaixo de 10 pessoas, associadas a diferentes papéis (funções ou atribuições): eventos dentro do tempo; artefatos para registro de informações e gestão, bem como suas entregas e regras de funcionamento. O objetivo desta estrutura adaptativa é a entrega (desenvolvimento e evolução) e sustentação (manutenção e correção) de produtos complexos através de ciclos incrementais (*sprints*). É, inclusive, utilizada a coordenação de diversos times de Scrum atacando pontos diferentes do mesmo produto (times para diferentes partes da arquitetura do software). Entretanto, por se tratar de um ambiente de desenvolvimento, necessita do uso de diversas outras ferramentas, técnicas e modelos para auxiliar na identificação de elaboração e problemas de projeto para seus artefatos de gestão do produto,

o "backlog do produto" (que é uma visada ampla do que deverá ser feito) e o "backlog" da sprint que trata do que será desenvolvido e entregue (funcionando) no próximo ciclo de desenvolvimento, conforme Figura 3.

Figura 4 - "*Framework*" Scrum para o desenvolvimento de software



Fonte: Home of Scrum (2022). Disponível em
<<https://www.scrum.org/resources/what-is-scrum>>

Em contraposição, os modelos adaptativos apresentavam principalmente desenvolvimento iterativo, entregas incrementais, intensa troca de informações, divisão da arquitetura do produto em times, contato com os usuários, dentre outras, o que permitia o desenvolvimento de complexos produtos em curto espaço de tempo, segundo levantamentos realizados por Cusumano (1997).

Os autores Rubin & Chisnell (2008) ponderam que, conforme a tecnologia penetra no mercado consumidor, mais desafiador se torna o desenvolvimento de produtos de base tecnológica. No início, os desenvolvedores tinham como público-alvo os usuários com grande bagagem técnica (inclusive em computadores), o que facilitava o desenvolvimento de produtos. Contudo, ao longo da expansão da tecnologia no mercado, cada vez mais objetiva-se atingir o público com menor fluência técnica e, conseqüentemente, sem conhecimento suficiente para configurar, instalar, consertar e aprender a utilizar complexos produtos tecnológicos em casa ou no trabalho. A solução proposta pelo autor

para auxiliar no desenvolvimento de produtos complexos é a utilização do UCD (*user centric design*, ou design centrado no usuário), mas com técnicas incorporadas para possibilitar a compreensão e mensuração do uso de produtos e dos fatores ergonômicos associados a estes. Para tanto, é necessário, inclusive, a definição de objetivos e metas de uso, eficiência, eficácia, satisfação e acessibilidade para auxiliar a sanar as dificuldades de desenvolvimento e lançamento de produtos cada vez mais complexos.

O processo de desenvolvimento de produtos, acelerado a partir da obsolescência programada e que embarca cada vez mais funcionalidades e, por consequência, maiores complexidades a serem desenvolvidas em curtos espaços de tempo, traz consigo a preocupação com a satisfação dos clientes. A respeito, é importante destacar que Reichheld (2003) propôs uma métrica para a pesquisa de satisfação, ou NPS (net promoter score), que rapidamente tornou-se popular na tradução da voz do consumidor. Tal métrica de pesquisa expressa, em última instância, sua lealdade, especialmente por sua simplicidade de implementação, como também na sua resposta e mesmo compreensão dos resultados. O NPS é estruturado a partir da probabilidade do cliente (usuário) recomendar, em uma escala de 0 a 10, seu produto, serviço ou marca e, a partir dos resultados, é calculado um único indicador que poderá variar de -100 a 100.

Sauro & Lewis (2016) revelam que pesquisadores e especialistas concordam que a métrica NPS, quando utilizada sozinha, possui limitado valor. Segundo os autores, seria necessário ainda entender os motivos que levaram os respondentes a atribuir determinada nota para, assim, compreender os motivos que afetam positivamente a lealdade e, por fim, se as melhorias percebidas na usabilidade afetariam a lealdade do consumidor, levando a melhores resultados. Ou em outras palavras, uma causa e efeito entre usabilidade e percepção de qualidade.

Diversas indústrias, não apenas a de software, migram para modelos Ágeis, adotando práticas que permitam desenvolver, entregar e manter produtos cada vez mais complexos em escala global. O Scrum, desde 1990, quando

desenvolvido por Sutherland & Schwaber (2017), foi proposto como uma estrutura (do inglês *framework*) dentro da qual pessoas (equipe) podem tratar e resolver problemas complexos e adaptativos, em formato produtivo e criativo, para entregar produtos com valor agregado. O Scrum é uma dentre as diversas técnicas, ferramentas, processos e metodologias intituladas Ágeis. Conforme as indústrias as adotam, é necessário que tanto o designer, como outros profissionais envolvidos no estudo e aplicação da ergonomia de fatores humanos, adaptem suas abordagens de projeto e intervenção para estes novos paradigmas Ágeis de desenvolvimento de produtos e serviços.

2.1 PROBLEMATIZAÇÃO

Os modelos de desenvolvimento prescritivos tradicionalmente utilizados em projetos de produto, devido a sua característica de detalhamento prévio das atividades (anteriormente à aprovação e execução do projeto), preveem, dentre outras disciplinas, o uso de metodologias e ferramentas para o projeto de fatores humanos e design. É possível ressaltar o modelo de referência (ou PDP - Processo de Desenvolvimento Integrado de Produto) apresentado por Rozenfeld *et al* (2006), principalmente voltado para empresas de manufatura e bens de consumo duráveis e de capital, apresenta direcionadores para guiar a adoção (e potencialmente a adaptação) do modelo para qualquer estrutura empresarial multidisciplinar.

Os referidos autores, inclusive, especificam que a ergonomia e estética sejam definidos dentro da fase Projeto Conceitual, na qual são realizadas atividades de busca, criação, representação e seleção de alternativas para os requisitos e metas de projeto anteriormente definidos. Dentre os objetivos citados, estariam a adequação do produto às características físicas e cognitivas do(s) usuário(s), assim como a simplificação do uso e prevenção de erros humanos, dentre outros. Em suma, estes modelos prescritivos, inclusive, contemplam em sua estrutura a aplicação pontual do projeto de design, ergonomia e fatores humanos, evidenciando o tratamento das problemáticas relativas à interface homem-máquina, relativas à disciplina do design.

O Modelo Canvas proposto por Osterwalder (2014) é um exemplo de ferramenta de gestão e acompanhamento utilizado no Scrum e em outros modelos de desenvolvimento Ágil. O Canvas pode ser utilizado na etapa anterior ao desenvolvimento do produto e durante o processo de desenvolvimento para avaliar a aderência (*fit*) do produto e modelo de negócio para determinado mercado. É um modelo qualitativo que permite, em formato incremental: ciclos de análise, proposição de hipóteses, construção de protótipos, desenvolvimento, novamente análise e assim por diante. Ele pode, inclusive, ser utilizado para subsidiar a definição e/ou mudança do escopo de produto ou o modelo de negócio.

O Teste A/B é outra técnica normalmente usada em conjunto com as metodologias Ágeis. Ela é muito utilizada por empresas que provêm software e serviços na internet, como por exemplo, a Microsoft, Google, Apple. Por seu turno, o serviço de "*videostreaming*" Netflix é um expoente, conduzindo diferentes versões de software (funcionalidades, interfaces gráficas e catálogo de títulos de vídeos) nas mais diversas plataformas de apresentação, seja um website na internet, televisão inteligente, telefone inteligente ou aparelho de transmissão acoplado a um televisor). Os experimentos são controlados e conduzidos "em produção", ou seja, eles acontecem na internet, com usuários reais, que sequer são informados de sua participação no experimento. Ao final, a empresa faz uma avaliação a partir de análise quantitativa de qual(is) versão(ões) será(ão) aprovada(s) para adoção em larga escala global, ou mesmo regional, inclusive para cada plataforma de apresentação, segundo aderência a uma meta previamente determinada.

Atualmente, existem basicamente duas vertentes, ou modalidades de computação, sobre as quais as aplicações (software) são desenvolvidas para operar: a computação em nuvem (*cloud computing*) e a computação local (*on-premises*). Existem ainda derivações de arquitetura e até mesmo a mescla destas duas modalidades, o que gera uma gama de hibridizações. A primeira é

definida pela Amazon (ou AWS), que é um fornecedor mundial de plataforma e serviços de computação em nuvem, como o uso de computação (processamento), armazenamento em banco de dados, aplicações (software) e outros recursos de tecnologia da informação sob demanda, por meio de serviços contratados de um fornecedor na internet, que mantém tais recursos para o uso de terceiros (clientes). Já a segunda, possui os recursos supracitados instalados no local (ou sítio), por exemplo da empresa (usuária) ou computador residencial ou veículo (centrais multimídia), sendo que os recursos são mantidos e operados pela própria empresa (ou fornecedor de tecnologia), conforme cita a HP (fornecedora mundial de plataforma de computação em nuvem, software e equipamentos, ou hardware, para *on-premises* e nuvem).

Podemos citar como exemplo de computação em nuvem o Google Docs (editor de texto online) que roda (é processado localmente) sobre um navegador Internet, mas está hospedado na Nuvem do Google. Em contrapartida, o Microsoft Word (editor de texto) é instalado no computador PC ou Mac, rodando sobre o sistema operacional Microsoft Windows ou mesmo o Apple Mac OS. Todavia, um ambiente híbrido seria o produto Microsoft Office 365 que possui tanto o Microsoft Word (online via navegador internet), como o Microsoft Word (instalado no computador), em que ambos podem gravar dados online (Microsoft One Drive), sendo que a versão local (instalada) pode ainda gravar localmente em ou em nuvem (maior complexidade).

Quadro 2 - Comparação macro de modalidades de computação

Necessidades do negócio	Local	Nuvem
Data center com locatário único (para conformidade)	Sim	-
Criptografia de dados altamente segura	Sim	Sim
Hardware personalizável, sistemas desenvolvidos para propósito específico	Sim	-
Capacidade de dimensionamento (tanto para ampliação como redução)	-	Sim
Infraestrutura exige altos investimentos regulares	Sim	-

Preço baseado na utilização, com pagamento conforme o uso	-	Sim
Visibilidade e controle completo dos dados	Sim	-
Backup e recuperação de dados automatizados e integrados	-	Sim
Risco praticamente zero de paralisação	-	Sim

Fonte: Tabela adaptada pelo autor da HPE. Disponível em <<https://www.hpe.com/br/pt/what-is/on-premises-vs-cloud.html>>.

O NPS (*Net Promoter Score*), também conhecido como a 'pergunta definitiva', foi proposto por Reichheld (2013) e é largamente utilizado para medir, a partir de pesquisa com o usuário, a lealdade do consumidor (ou usuário) àquele produto ou serviço (percepção da qualidade). Trata-se, assim, de um indicador único que classifica a percepção em uma escala de detratores (negativo), neutro (indiferente) e promotores (positivo). Esta ferramenta de pesquisa é tradicionalmente utilizada como indicador chave nos modelos Ágeis, uma vez que pode ser facilmente automatizada no software, especialmente naqueles que operam sobre a modalidade de computação em nuvem. Ainda, conforme afirmação em meados da década de 2010, do CEO da Philips Electronics, Gerard Kleisterlee, o NPS era um indicador universal facilmente compreendido pela gestão em toda a empresa e não apenas pelos especialistas.

A partir da interpretação de que diferentes técnicas, ferramentas, processos e mesmo outros modelos possam operar em conjunto dentro das práticas Ágeis, surgiram derivações e novas proposições "híbridas" como o '*Lean Launchpad*' (Blank. 2011). Este, por exemplo, propõe a unificação de diversas abordagens Ágeis, usando diversas ferramentas aderentes ao Ágil, objetivando projetos ainda mais interativos e com entregas de produtos cada vez mais rápidas. Neste modelo, o Canvas e o Scrum, em conjunto com outras abordagens, são utilizados ao longo dos ciclos de desenvolvimento para alcançar a estabilização de um modelo de negócios, e não apenas desenvolver um produto e/ou serviço.

Quadro 3 - Comparação de características entre modelos prescritivos e ágeis.

Característica	Modelos prescritivos	Modelos adaptativos
Definição do problema do produto	Conjunto mínimo necessário para a aprovação do projeto (evolutivo)	Problema de modelo de negócio e/ou produto mínimo (exploratório)
Projeto organizado em macro fases, fases e/ou etapas	Sim (depende do PDP)	Não (ciclos de tempo)
Entrega do produto (objetivo principal)	Ao final do processo de desenvolvimento	Durante o processo de desenvolvimento em formato evolutivo e incremental
Atividades, ferramentas e processos previamente definidos	Estrutura analítica do trabalho (WBS /EDT)	Não
Requisitos ergonômicos definidos	Sim (evolutivo nas fases do projeto)	Não (exploratório)
Tempo de projeto	Medido em anos ou meses	Ciclos medidos em meses, quinzenas ou semanas (sprints e/ou releases)
Modelo de gestão	PMO 6a Edição (10 áreas) PRINCE2 Agile Foundation /2018	Escopo do produto Gestão de recursos Gestão de tempo

Fonte: Compilado pelo autor.

No quadro acima é possível observar algumas diferenças básicas entre os dois modelos, especialmente no que tange à formalização de uma estrutura de projeto (ou empreendimento) devidamente organizado para atingir um fim (conhecido) nos modelos prescritivos versus um modelo que, ao longo do tempo, adapta-se para atingir um fim (pouco conhecido e detalhado).

Seria desejável, portanto, a proposição de uma abordagem que viabilizasse a inserção do projeto de fatores humanos no modelo Ágil a partir desta abertura a novas técnicas, processos e modelos? Como manter íntegros os princípios de agilidade, especialmente as restrições de tempo nos curtos ciclos de desenvolvimento incrementais (*sprints*), que podem durar até um mês,

como previsto no Scrum? Como conciliar o projeto ergonômico no modelo Ágil sem a perda de confiabilidade tradicionalmente obtida nos modelos prescritivos?

Com objetivo de atender a premissa de ambos, desenvolvimento Ágil e o projeto prescritivo em que, naturalmente, incorporam-se os princípios dos fatores humanos, poderia o projeto ergonômico ser viabilizado a partir de uma abordagem automatizada, e por isso, aderente ao Ágil? Poderiam ser utilizados fluxos contínuos de dados de uso do software /usuários como subsídio ao projeto ergonômico ao longo dos ciclos incrementais de desenvolvimento Ágil? Seria possível, a partir destes dados contínuos, calcular métricas de eficiência e eficácia de usabilidade para compor indicadores de performance em usabilidade (intitulando, por conseguinte o constructo KPIsUX) que permitam o acompanhamento da evolução e o estabelecimento de objetivos e metas tanto ergonômicas como de qualidade percebida do produto (software) pelos usuários finais?

2.2 PROBLEMA

Como implementar soluções ergonômicas ao longo dos ciclos de desenvolvimento de software, sem interferir nos princípios da metodologia Ágil.

2.3 HIPÓTESE

O uso de indicadores de performance em usabilidade (KPIsUX) para guiar o desenvolvimento Ágil de software permite a melhora do NPS.

2.4 VARIÁVEIS

Variável independente:

- Métricas de performance (sucesso binário, nível de sucesso, tempo médio na tarefa, perda, sucesso no tempo, erros em alternativas, curva de aprendizado).

Variável antecedente:

- Ciclos de desenvolvimento.

Variável interveniente:

- Mudanças deflagradas no software devido ao uso das métricas de performance.

Variável dependente:

- Indicador de satisfação do usuário NPS (*net promoter score*).

Variáveis de controle:

- Website com arquitetura MPA (*Multi-page application*);
- Uso de modelo Ágil de desenvolvimento (Scrum);
- Modalidade nuvem (*Cloud*).

2.5 OBJETIVOS

2.5.1 Objetivo Geral

Caracterizar a relação de causalidade entre indicadores de performance em usabilidade (constructo KPIsUX) e a qualidade /satisfação percebida pelos usuários finais no website (software) por meio do NPS (*Net promoter score*), introduzindo a utilização de considerações aos fatores humanos no projeto da indústria de software, em compatibilidade com a metodologia Ágil.

2.5.2 Objetivos Específicos

- i. Analisar a aplicação de métricas de usabilidade em formato automatizado para a composição do constructo indicador de performance em usabilidade (ou KPIsUX);
- ii. Definir estrutura mínima de dados para análise contínua dos KPIsUX;
- iii. Definir procedimento de coleta, tratamento e carga de dados;
- iv. Propor um processo aderente ao Ágil (Scrum);

2.6 JUSTIFICATIVA

Empresas de penetração global de software estão realizando a migração de metodologias prescritivas (*waterfall* - cascata) para modelos Ágeis. Esta migração ganhou velocidade a partir da década de 1990, conforme observado nos EUA por Cusumano (1997). Tais empresas apontaram como objetivo o desenvolvimento acelerado de produtos, sendo que eles apresentam crescente complexidade e estão inseridos em contextos altamente competitivos e dinâmicos.

Já a indústria tradicional de produtos físicos, serviços e outros utilizam, vastamente, metodologias integradas prescritivas conforme apresentado por ROZENFELD *et al.* (2006). E é sob este modelo prescritivo que é estruturada, atualmente, a atenção aos princípios de ergonomia e fatores humanos.

É relevante, portanto, que as disciplinas de ergonomia e fatores humanos e seus modelos de referência, metodologias, ferramentas, técnicas e práticas sejam revisitadas e, porventura, adaptadas para serem utilizadas em integração com os modelos Ágeis. Desta forma, instrumentalizando os profissionais de design, ergonomia e fatores humanos a operarem, em contextos de desenvolvimento de produtos (software neste estudo) que possuam características de volatilidade, velocidade e alto grau de incerteza, sem contudo perder a confiança sobre dados de performance de uso, necessários para a condução de desenvolvimento de produtos guiados por dados e fatos.

2.7 DELIMITAÇÕES

O estudo proposto avaliou, em experimento de campo, um website de notícias (revista digital) no formato MPA (*multi-page application*). Ou seja, um website em que as diversas funcionalidades estão distribuídas em diferentes páginas (URLs - *Uniform Resource Locator* ou Localizador Uniforme de Recursos) em uma mescla de diferentes linguagens de programação, neste estudo HTML, CSS e JavaScript.

Não foram utilizadas técnicas de promoção e/ou publicidade digital (propaganda ou Ads) em redes sociais ou ferramentas de pesquisas, a fim de que os resultados do experimento de campo não sofressem interferência.

No estudo foi utilizado como framework Ágil o Scrum. Não tendo sido utilizadas outras ferramentas ou técnicas de identificação e/ou priorização de funcionalidades do backlog do produto.

3 REVISÃO BIBLIOGRÁFICA

Neste capítulo são introduzidos os conhecimentos referentes à aplicação da ergonomia e fatores humanos no desenvolvimento de produtos. Para tanto, a revisão é iniciada com aplicações genéricas, em seguida são discutidas as abordagens padronizadas para, então, culminar na aplicação padronizada em software, com inclusão das ferramentas posteriormente utilizadas no experimento de campo.

3.1 ERGONOMIA E FATORES HUMANOS

Ilida (1992) define ergonomia como o estudo da adaptação do trabalho ao homem, incluindo maquinário (equipamento), a todo ambiente que circunda o local em que o trabalho é realizado, com os objetivos práticos de oferecer segurança, satisfação e bem-estar aos trabalhadores no seu relacionamento com os sistemas produtivos. Os estudos em ergonomia segundo aquele autor, concentram-se nos aspectos do comportamento humano no trabalho e outros fatores relativos ao projeto de sistemas de trabalho:

- O homem: incluindo suas características físicas, fisiológicas, psicológicas e sociais, influência do sexo, idade, treinamento e motivação.
- A máquina: sob a ótica de equipamentos para auxílio na realização do trabalho, incluindo equipamentos, ferramentas, mobiliário e instalações.
- O ambiente físico: ambiente em que está contido o homem durante a realização do trabalho, considerando, inclusive, variáveis de temperatura, ruídos, vibrações, luz, cores, gases e outros.
- A informação: referindo-se à comunicação entre os elementos de um sistema, transmissão de informação, o processamento e a tomada de decisões.
- A organização: conjugação dos elementos previamente citados no sistema produtivo, aspectos de horário da jornada de trabalho, turnos e a formação de equipes.
- As consequências do trabalho: questões de controle como inspeção de tarefas, estudos dos erros e acidentes, além de estudos sobre gastos energéticos, fadiga e estresse.

Segundo Lida (1992) a origem da ergonomia foi relatada após o período da II Guerra Mundial (1939 a 1945), fruto da interdisciplinaridade de profissionais que atuaram no período da guerra.

Pahl & Beitz *et al.* (2003) consideram a ergonomia e fatores humanos como parte das características, habilidades e necessidades do indivíduo para cuidar das relações entre o indivíduo e o produto técnico, em busca de equacionar uma configuração entre:

- Adaptação do produto técnico ao indivíduo, mas também pela escolha das pessoas, bem como pela preparação e treinamento.
- Adaptação do indivíduo ao objetivo ou à atividade em/ou com sistemas técnicos.

Os fundamentos da ergonomia (Pahl & Beitz *et al.* 2003) partem de considerações a partir do indivíduo que, em um sistema técnico, pode atuar de diversos modos ou ser afetado pelas ações. Por isso, o projeto de máquinas deve considerar os aspectos:

- Biomecânicos relativos ao manuseio e utilização de produtos técnicos e as posturas e movimentos corporais, com ênfase no cálculo de cargas e forças admissíveis para o indivíduo. É ainda levado em conta aspectos de frequência e duração das atividades, idade, sexo, prática e experiência do indivíduo, dentre outros.
- Fisiológicos relativos aos movimentos corporais requeridos para o manuseio e utilização do produto técnico. A ênfase é relativa ao organismo muscular suportar os esforços considerando a dinâmica, incluindo carregamento, solicitação e fadiga do organismo frente às condições ambientais para a realização do trabalho.
- Psicológicos relativos à percepção, raciocínio, decisão e atuação do indivíduo na utilização do produto técnico, incluindo as perturbações do ambiente no sistema.

Os referidos autores prescrevem a aplicação da ergonomia como uma diretriz especializada no projeto de máquinas na fase de anteprojeto. As diretrizes especializadas são apresentadas como "design para x" (ou DfX), sendo a ergonomia uma dentre as diversas diretrizes técnicas inerentes ao projeto de engenharia de máquinas, que são combinadas naquela fase do projeto, citando especialmente normas técnicas VDI (*Verein Deutscher Ingenieure* ou Associação Alemã de Engenharia), DIN (*Deutsches Institut für Normung* ou Instituto Alemão para Normatização), entre outras fontes.

É possível observar, nas proposições de Pahl & Beitz *et al.* (2003), Rozenfeld *et al.* (2006) e Baxter (1998), que o projeto de ergonomia é alocado ou prescrito para determinada fase do projeto de desenvolvimento do produto, em alinhamento com a proposição 'stage gate' de Cooper (2001). Os referidos autores focam a aplicação da ergonomia no projeto de interfaces homem-máquina e nas interfaces digitais que porventura existam nos produtos industriais. É de particular interesse do estudo a última, identificada como aspectos psicológicos (Pahl & Beitz *et al.* 2003; Baxter, 1998) ou fatores da informação (Iida, 1992).

Já os editores Stanton *et al.* (2005), a partir da compilação do trabalho de pesquisa de diversos autores, apresentam uma atuação mais ampla da ergonomia e dos fatores humanos. Eles expõem, inclusive, que o domínio de atuação da disciplina possui ênfase por vezes implícita na análise da performance humana, segurança e satisfação, e que a disciplina possui tradição no desenvolvimento, aplicação e validação de métodos. Os referidos editores citam ainda que os diversos métodos oferecem aos ergonomistas uma abordagem estruturada para analisar e avaliar problemas de projeto (relativos à ergonomia), delineando, adicionalmente, ao modelo científico-prático que os profissionais necessitam balancear em sua atuação:

Ergonomista como um cientista:

- Ampliação do trabalho de outros;
- Teste de teorias da performance homem-máquina;

- Desenvolvimento de hipóteses;
- Questionamento de todos os fatores envolvidos;
- Uso de rigorosa coleta de dados e técnicas de análise;
- Garantia da repetitividade de resultados;
- Disseminação das descobertas obtidas a partir de estudos.

Ergonomista como um prático:

- Endereçamento de problemas reais;
- Busca do melhor compromisso sobre circunstâncias difíceis;
- Oferta da melhor solução custo-efetiva;
- Desenvolvimento de demonstrações e de protótipos de soluções;
- Análise e validação dos efeitos das mudanças;
- Desenvolvimento de comparações (benchmarks) das melhores práticas;
- Comunicação das descobertas para os interessados.

Os citados editores reúnem ainda um conjunto de seis métodos padronizados que visam auxiliar os profissionais de ergonomia em sua atuação e no equacionamento do modelo científico-prático, sendo eles:

1. Métodos Físicos: análise e avaliação de fatores musculoesqueléticos, incluindo desconforto (medição de), postura, risco do local de trabalho, medição de esforço e fadiga, desordem lombar e predição de riscos das extremidades superiores;
2. Métodos Psicofisiológicos: análise e avaliação psicofisiológica humana, incluindo batimento cardíaco (e sua variação e potencial a partir de eventos relacionados), resposta galvânica da pele, pressão sanguínea, frequência respiratória, movimentos da pálpebra e atividade muscular;
3. Métodos Comportamentais e Cognitivos: análise e avaliação de pessoas, eventos, artefatos e atividades (tarefas), incluindo observação e entrevista, métodos de análise cognitiva de atividades, predição de erros humanos, predição de carga de trabalho e consciência situacional;

4. Métodos de Equipes: análise e avaliação de equipes incluindo treinamento, avaliação, desenvolvimento, comunicação, cognição, tomada de decisão e atividades (em grupo);
5. Métodos Ambientais: análise e avaliação de fatores de ambiente incluindo os térmicos, qualidade do ar (locais fechados), iluminação (locais fechados), ruídos e acústica (medição de), exposição à vibração e habitabilidade;
6. Métodos Macroergonômicos: análise e avaliação de sistemas de trabalho incluindo métodos comportamentais e de organização do trabalho, sistema de manufatura, antropotecnologia, validação e intervenção de posto de trabalho, análise da estrutura e processo dos sistemas de trabalho.

Dos seis métodos apresentados anteriormente pelos editores Stanton *et al.* (2005), os métodos comportamentais e cognitivos, apontam ferramentas e técnicas que permitem operar dentro do contexto Ágil para a avaliação e subsídio ao desenvolvimento de software guiada por dados contínuos (em volume) obtidos diretamente do sistema homem-máquina.

3.1.1 Métodos Comportamentais e Cognitivos

O autor Stanton (2005) apresenta que os métodos comportamentais e cognitivos possuem sua origem nas disciplinas psicológicas, bem como trata da coleta de informações acerca de percepções, processos cognitivos, respostas de indivíduos e outras fontes. Tais fontes podem ser apresentadas em diferentes formas e títulos, dentre eles: erros humanos, tarefas, tempo de tarefa, metas (e submetas), decisões, cargas de trabalho e preferências dos usuários. Os métodos apresentados pelo autor podem ser classificados em quatro categorias:

- Métodos gerais de análise: compostos por observações, entrevistas, protocolos verbais de análise, malha de repertório (repertory grid) e grupo de foco (focus group);

- Métodos de análise cognitiva da tarefa: incluem análise hierárquica da tarefa (HTA), metodologia da alocação da função, método de decisão crítica (CDM) e análise do trabalho cognitiva aplicado (ACWA);
- Método de análise de erros: abordagem de redução e predição sistemática de erros humanos (SHERPA), análise da tarefa para identificação de erros (TAFEI);
- Método para análise da carga de trabalho e situação: índice de carga do trabalho e técnica da classificação da consciência situacional (oriunda da NASA), modelo de compartilhamento de tempo de múltiplos recursos compartilhados (MRTSM), análise do caminho crítico (CPA) e técnica de avaliação global de consciência situacional (SAGAT).

O autor apresenta um total de 15 métodos, que podem ser utilizados em conjunto ou isoladamente, sendo alguns pré-requisitos de outros. Ele ressalta, ainda, a necessidade de condução de pilotos anteriormente ao projeto ergonômico a fim de validar o modelo, formato e padrões de dados necessários para cobrir as saídas (resultados) dos métodos: medição do tempo, erros, análise de tarefas, metas (e objetivos), compreensão e análise de decisões, carga do trabalho e a usabilidade. O quadro 4 a seguir ordena as saídas, métodos e seus objetivos principais.

Quadro 4 - Métodos comportamentais e cognitivos e seus resultados.

		Output						
	Methods	Time	Errors	Tasks	Goals	Decisions	Workload	Usability
General Analysis Methods	Observation							
	Interview							
	Verbal Protocol							
	Rep. Grids							
	Focus Groups							
Cognitive Task Analysis	HTA							
	Function Alloc.							
	CDM							
	ACWA							
Error Analysis	SHERPA							
	TAFEI							
Workload and Situation Analysis	Workload							
	MRTSM							
	CPA							
	SAGAT							

Fonte: Stanton (2005).

O quadro acima classifica em uma escala de cores as saídas (resultados) de dados, conforme sua objetividade principal (preto), secundária (cinza) ou genéricas sem padrão de resultados específico (branco).

Ainda, os autores Stanton *et al.* (2005) cobrem os usos principais e secundárias dos 15 métodos, apontando suas vantagens e desvantagens. Por isso, o uso associado de diversos métodos e o piloto são necessários para a validação do instrumento e método utilizados no projeto ergonômico.

Quadro 5 - Vantagens e desvantagens de cada um dos métodos.

	Métodos	Vantagens	Desvantagens
Métodos gerais de análise	Observação	-Informação objetiva e comparável -Individualização da análise e do sistema	-Uso intenso de recursos -Observador afeta o experimento -Não revela informação cognitiva
	Entrevista	-É familiar para participantes -Permite conferência online -Consistência e detalhe	-Necessita de piloto -Análise consome tempo -Ambiente altera o resultado
	Protocolo Verbal	-Qualidade e volume de conteúdo -Análise mais natural (comportamento) -Indicado para sequência de atividades -Especialistas aumentam confiabilidade	-Consome tempo -Pode afetar o resultado (narrar) -Informação sem foco (estudo) -Pouca resolução (alta demanda)
	Rep. Grid	-Procedimento compreensivo -Não necessita de estatística	-Análise consome tempo -Não provê fatores (análise)
	Focus Group	-Permite quantidade de pessoas -Sem interferências individuais	-Análise consome tempo
Métodos de análise cognitiva da tarefa	HTA	-Uso genérico e adaptável -Respeita a granularidade de detalhe	-Requer especialista e envolvidos -Proporção tempo /complexidade
	Alocação de Função	-Estrutura a decisão (função) -Rastreabilidade de decisões -Garante alocação do usuário/função	-Requer tempo em larga escala -Requer especialista -Requer pessoal de operações
	CDM	-Inteligência sobre incidentes reais -Aprofundamento na problemática -Eficiência na análise (incidentes) -Possibilita cenários (<i>what if</i>)	-Confiabilidade questionável -Usa muitos recursos (data set) -Requer treinamento
	ACWA	-Detalha a rede abstração (FAN) -Prepara para o não planejado -Permite rastreabilidade	-Requer especialista -Pré-requisitos (FAN)
Método de análise de erros	SHERPA	-Procedimento estruturado -Previsão de erros -Verificação e confiabilidade dos dados -Economiza tempo -Predição e redução de erros	-Consome tempo -Pré-requisitos (HTA) -Sem modelo cognitivo (falhas) -Precisão -Generalista
	TAFEI	-Procedimento estruturado -Sustentação teórica -Flexível	-Lento (requer HTA e SSDs) -Requer especialista -Limitado (metas /atividades)
Método para análise da carga de trabalho e situação	MWL	-Efetivo (tarefa 1) -Quantifica cognição (tarefa 2) -Monitoramento contínuo (fisiológico) -Fácil de usar e analisar	-Identifica entre as tarefas -Requer detalhamento (tarefa 2) -Difícil e intrusivo (fisiologia) -Confiabilidade (após tarefa)
	MRTSM	-Identifica fenômenos que afetam a performance -Sustentação e fácil de calcular -Robusto para adaptar -Flexível para administrar	-Requer experiência (em código) -Requer especialista -Medida relativa da interferência -Possui reduzida validação -Não atende multitarefa
	CPA	-Procedimento estruturado -Avalia atividades em paralelo -Dados validam observação	-Consome tempo -Difícil de parametrizar -Restrito (performance de tempo) -Sem medição individualizada -Simplista (tarefas cognitivas)
	SAGAT	-Mede a consciência situacional (SA) -Mede a SA sem interferências -Rapidez na coleta de dados -Diagnóstico preciso na SA e tempo real	-Requer extensa preparação -Requer infraestrutura -Requer habilidade (simulação) -Pode interferir na performance -Não indicado para operação

Fonte: Stanton (2005).

O quadro acima classifica em uma escala de cores as saídas (resultados) de dados, conforme sua objetividade principal (preto), secundária (cinza) ou genéricas sem padrão de resultados específico (branco).

Os autores Rubin & Chisnell (2008) apresentam a usabilidade aplicada à interface digital como uma qualidade inerente a este tipo de produto. E ela pode ser percebida quando há uma ausência de frustração no uso ou quando o usuário pode realizar as atividades do modo esperado, sem hesitação ou dúvidas sobre o uso. Em resumo, um produto é usável quando possui atributos de:

- Utilidade: o produto permite que o usuário alcance seus objetivos;
- Eficiência: a velocidade com que o usuário alcança seus objetivos no uso;
- Efetividade: o quanto o produto se comporta conforme a expectativa dos usuários em termos percentuais de alcance de seus objetivos (sucesso), ou não (erros);
- Aprendizado: é parte da efetividade, sendo relacionada à habilidade de operar o sistema dentro de determinados parâmetros definidos de competência.
- Satisfação: relativa à percepção do usuário, seus sentimentos e opiniões sobre o produto;
- Acessibilidade: sinônimo de usabilidade de forma genérica, sendo a facilidade de acesso aos produtos para atingir determinada meta.

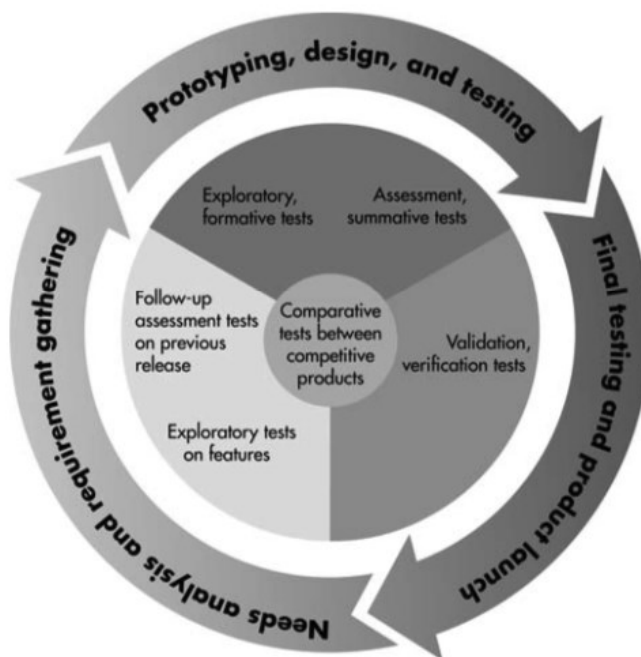
Os autores pontuam que o oposto da usabilidade, ou seja, a dificuldade de uso de produtos, especialmente os tecnológicos, é motivada, principalmente, por cinco motivos: [1] o desenvolvimento é focado no sistema ou máquina; [2] o público-alvo muda ou se adapta; [3] dificuldade de desenvolvimento de produtos fáceis de usar; [4] equipes especializadas não trabalham sempre integradas e, [5] projeto e implementação (execução) nem sempre obtém êxito. Os citados autores propõem, assim, o uso recursivo de uma série de técnicas, métodos e práticas durante o ciclo de desenvolvimento do produto, sendo estes:

- Pesquisa Etnográfica: técnica oriunda da antropologia que usa a observação de usuários nas condições e locais em que normalmente utilizariam o produto;
- Projeto Participativo: aproximação do(s) usuário(s) com o time de desenvolvimento para avaliação de proposições;
- Pesquisa Focus Group: avaliação em grupo de usuários do produto ainda em fases preliminares (ou mesmo protótipos), coletando dados qualitativos de percepção dos usuários;
- Formulários de Pesquisa: aplicação de questionários para uma amostragem de usuários, levando em consideração sua estruturação e validação anteriormente à sua aplicação;
- Navegação "Walk-Throughs": apresentação da estrutura das principais funcionalidades a usuários finais ainda nos estágios iniciais para identificação de dificuldades;
- Sorteio de Cartas: técnica em que os usuários utilizam cartas para auxiliar na avaliação e identificação de nomenclatura e funções do produto (percepções);
- Protótipo em Papel: apresentação e entrevista sobre protótipos impressos, sejam estes wireframes ou rascunhos desenhados à mão, apresentando a navegação do sistema proposto;
- Avaliação de especialistas ou eurística: avaliação do produto ou sistema realizados por um especialista em usabilidade ou fatores humanos;
- Estudos de acompanhamento: estudo após o lançamento do produto com usuários finais utilizando o produto. Para tanto, são aplicados formulários de pesquisa, entrevista e observações para subsidiar a evolução do produto;
- Teste de usabilidade: usa técnicas de coleta de dados empíricos nos usuários que utilizam o produto final, segmentadas em uma abordagem formal na condução de testes experimentais para confirmar ou refutar uma hipótese, e outra, para obter dados ao longo do ciclo de uso para expor deficiências de usabilidade, a fim de gradualmente aprimorar o produto.

Rubin & Chisnell (2015) defendem que a abordagem do "*design centrado no usuário*" (UCD - *user centric design*), no projeto de desenvolvimento do produto, seria a forma de equacionar a dificuldade em desenvolver produtos tecnológicos complexos. A abordagem, de acordo com os autores, é um processo de projeto que busca suportar, ou sustentar, o modo como os usuários realizam determinado conjunto de atividades. Pontuam que o UCD é caracterizado pelo ativo envolvimento na compreensão do usuário e requisitos da tarefa, apropriada alocação de funcionalidades entre usuários e tecnologia, projeto de soluções em formato iterativo e projeto com equipe multidisciplinar. Ressaltam que os testes de usabilidade são técnicas, dentre diversas, para auxiliar na aplicação correta do "*design centrado no usuário*", enfatizando os seguintes princípios:

- Foco iniciando nos usuários e suas tarefas;
- Avaliação e medição do uso do produto;
- Projeto iterativo.

Figura 5 - Diagrama de aplicação de testes de usabilidade



Fonte: Rubin & Chisnell (2015).

Segundo os autores supracitados, a aplicação de diferentes testes de usabilidade deveria ser realizados ao longo do ciclo de desenvolvimento do produto, buscando responder as questões relativas ao seu uso: funcionar; factibilidade do propósito; e o que deveria realizar. O diagrama acima direciona alguns métodos e ferramentas de acordo com as questões existentes.

Os testes de usabilidade deveriam ser realizados com usuários finais que representem a população (público-alvo). Todavia, existiriam limitações e a condução apenas do teste de usabilidade não garantiria que um produto fosse categorizado como usável, especialmente quando observadas as seguintes questões (Rubin & Chisnell, 2008):

- Testes são artificiais: testes em ambiente controlado ou em campo representam as situações reais, e o próprio ato de conduzir o estudo pode alterar o resultado;
- Resultados não provam que o produto funciona: a quantidade de usuários nos testes pode representar estatisticamente a população, mas o teste em si pode mascarar a realidade;
- Participantes dificilmente representam o público-alvo: os participantes são tão representativos quanto a habilidade da equipe em compreender sua composição final.
- Teste de usabilidade pode não ser a melhor técnica: existem outras técnicas especialmente utilizadas para aplicação nos estágios iniciais do desenvolvimento do produto, podendo a avaliação heurística ser mais efetiva em alguns destes casos.

Os referidos autores apresentam um guia para o Ergonomista preparar, executar e analisar os resultados do teste de usabilidade: roteiro orientativo, questionário de checagem, instrumentos de coleta, contrato de acordo e consentimento, questionários pré-testes, cenários de execução de atividades, questionários após-teste e guia de *debriefing*. Ao final do teste, ele apresenta a aplicação de análise, propondo, assim, três etapas: compilação de dados, organização de dados e análise de dados, o que resulta em formas de utilização

do conhecimento para sanar as dificuldades em usabilidade identificadas. A análise é organizada em uma planilha em que são calculados: o tempo médio (média simples) para a conclusão de cada tarefa (testada); mediana para identificação do ponto médio entre as observações; faixa (alta e baixa) de tempo para conclusão da tarefa; desvio padrão do tempo médio para conclusão da tarefa, conforme quadro abaixo, em que as linhas cinzas ressaltam as tarefas que estão fora de critérios (metas) estabelecidos para o teste de usabilidade. Os destaques incluem participantes ⁽²⁾ que não realizaram as tarefas em conformidade com o plano ou ficaram abaixo ⁽¹⁾ do tempo mínimo requerido para o teste de usabilidade.

Quadro 6 - Análise de nível de performance.

Module	Tasks	Percentage of participants performing correctly (within benchmark)	Mean time (minutes)	Standard deviation (minutes)
A	Set temperature pressure	83	3.21	4.38
	Set gas flows	33	12.08	10.15 ¹
	Ignite the QRC	83	2.75	2.68
	Bake out column	100	0.46	0.17
	Set oven temperature program	66	6.54	2.56
	Run QRC checkout	100	0.83	0.34
	Program pressure and temperature	66	5.17	3.46
	Rerun checkout	100	0.29	0.09
B	Load the sample tray	100	0.88	0.28
	Create a subdirectory	33	8.00	3.92
	Create a sequence	66	9.42	7.70
	Start the sequence	66	1.00	0.92
	Stop the sequence	100	1.30	0.89
	Check the report	100 ²	1.38	0.58
	Reintegrate and print report	66	11.67	5.15
	Modify the method	100 ³	2.67	1.86
	Save the method	83	2.92	5.41
	Modify the sequence	66	3.46	3.96
	Restart the sequence	83	1.08	0.89

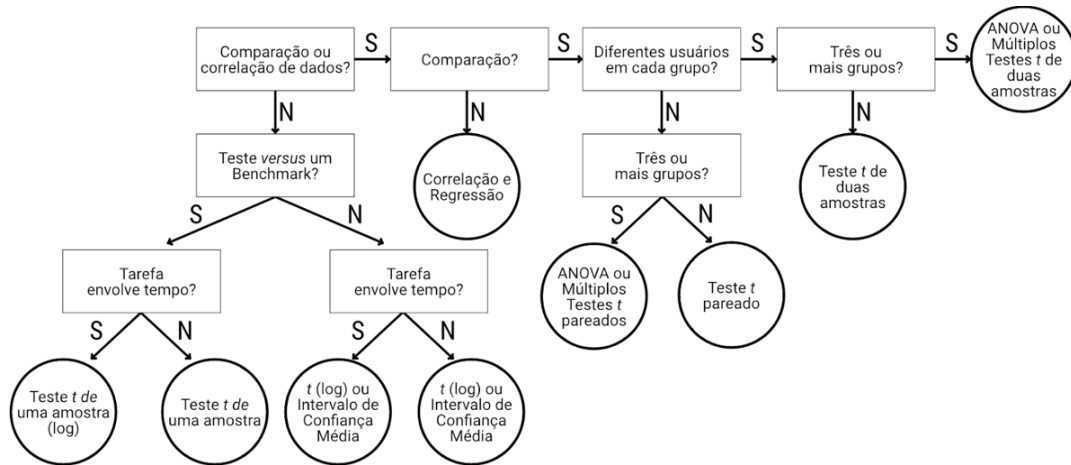
Fonte: Rubin & Chisnell (2008).

Sauro & Lewis (2016), por sua vez, permitem aprofundar ainda mais nos métodos comportamentais e cognitivos aplicados ao software, apresentando uma abordagem pragmática e multidisciplinar acerca da medição da experiência do usuário (UX) por meio de questionários padronizados para teste fim. Ressaltam que, mesmo sendo o design essencialmente qualitativo, diversas atividades analisadas são quantitativas. Assim, é possível medir o impacto de um design bom ou ruim e até mesmo os benefícios associados aos seus esforços. Em sua visão, a análise da experiência do usuário é pertinente aos profissionais de usabilidade, designers, gerentes de produto, profissionais de marketing e desenvolvedores de software. Por fim, os estudos de UX são classificados em três tipos:

- Estudo experimental: os participantes são selecionados aleatoriamente e alocados para diferentes tratamentos (análises), possibilitando maior controle interno e consequente validação do estudo;
- Estudo quasi-experimental: estudo sem alocação aleatória de participantes quando comparando resultados de diferentes grupos usuários, produtos ou versões;
- Estudo de correlação: é examinada a relação entre duas variáveis, embora não seja possível provar a relação de causa e efeito entre elas.

Os três tipos de estudo requerem diferentes estatísticas aplicadas em conformidade com as métricas utilizadas para coleta de dados, resultados esperados, dentre outros parâmetros necessários para a condução do estudo.

Figura 6 - Mapa de decisão para análise de dados contínuos



Fonte: Sauro & Lewis (2016). Tradução do autor.

Destaque para um dos mapas referentes à coleta de dados contínuos, exemplificando (conforme figura acima) a aplicação de diferentes testes estatísticos, de acordo com a análise a ser desenvolvida no estudo da experiência do usuário. Um dos principais objetivos na condução de estudos de usuários, ou experiência dos usuários, é o estabelecimento de alguma relação causal entre o design e o comportamento do usuário (Sauro, 2013). Os três tipos de estudos de UX propostos pelo autor podem utilizar diferentes abordagens na aquisição de informação:

- Coleta de dados: dados de usabilidade podem ser coletados em sessões moderadas em laboratórios ou ainda em campo por meios não moderados ou moderados diretamente com os usuários finais ou amostragem (websites, por exemplo, com a gravação remota de dados de uso das sessões) citando (Albert et al, 2010; Lewis, 2012; Sauro, 2010; Tullis & Albert, 2013);
- Conclusão de atividades: também chamado de taxa de sucesso, é uma das métricas fundamentais de usabilidade (apud Nielsen, 2001) computada como dados binários, a exemplo de sucesso na tarefa (codificado como 1) ou insucesso na tarefa (codificado como 0), mas que pode também ser acrescida de sucesso parcial;

- Problemas de usabilidade: identificação de problemas na interface do usuário (user interface ou UI) que podem ser organizadas em listas, possuir nome, descrições e grau de severidade de acordo com o problema encontrado (apud Rubin, 1004; Dumas & Redish, 1999);
- Tempo na tarefa: medição de quanto tempo o usuário gasta em uma tarefa, podendo ser o tempo de conclusão da tarefa; tempo até a falha ou tempo total na tarefa;
- Erros: é o agrupamento de ações não intencionais, deslizes, erros ou omissões enquanto realizada uma tarefa. O resultado pode ser computado como valor binário, sendo a existência de erro (=1) e a inexistência de erro (=0);
- Satisfação: questionários que meçam a percepção da facilidade de uso de um sistema, imediatamente depois da tarefa (pós-tarefa), ao final de uma sessão de usabilidade (pós-teste) ou à parte de um teste de usabilidade;
- Combinação de escores: a aquisição de diversas métricas em um estudo é mais segura, podendo compor um painel executivo (dashboard) que apresente a correlação entre as diversas métricas;
- Teste A/B: muito utilizado para comparar designs alternativos em páginas da internet (websites), as diferenças podem variar de simples imagens a layouts completos, dentre outros;
- Cliques, visualização de páginas e taxas de conversão: prática habitual em websites e aplicações web de coleta automática de dados, podendo ser a única fonte disponível para o estudo (UX);
- Dados de questionários: uma das formas mais rápida de coletar dados de usuários, combinando respostas abertas, binários e escalas Likert;
- Avaliação: perguntas que usam escalas em respostas fechadas (de acordo ou desacordo) usando escala Likert;
- Net Promoter Score (NPS): avalia a percepção da lealdade do usuário, em uma escala crescente de 0 a 10 (apud Reichheld. 2006), categorizando os participantes em grupos de promotores, passivos e detratores.

- Comentários e dados abertos: analisar e priorizar os comentários de usuários, avaliando os motivos dos usuários serem promotores ou detratores, intuição a partir de estudos de campo, reclamações de produto para o suporte, o motivo de tarefas serem difíceis de concluir;
- Levantamento de requisitos: outra tarefa principal da pesquisa com o usuário é a identificação de funcionalidades de um produto, por meio de análise de comportamento mediante quantificação similar à análise de UI.

Os autores Sauro & Lewis (2016) apontam os questionários como meios práticos para se obter informações de usuários. Todavia, e quando padronizados, eles permitem a avaliação dos participantes com maior confiabilidade, visto que são submetidos a análises psicométricas para sua validação mediante comprovação de confiança, validação e sensibilidade, apresentando as seguintes vantagens: objetividade, replicabilidade, quantificação, economia, comunicação e generalização científica (*apud* Nunnally, 1978). Os autores os classificam em pós-teste ou pós-tarefa.

Quadro 7 – Compilação de questionários padronizados

	Teste	Descrição
Pós-teste	QUIS	Questionário para apurar a satisfação qualitativa do usuário, em interfaces de computador.
	SUMI	Inventário de medição de usabilidade de software com o objetivo de medir a eficiência, efeito, assistência, controle e aprendizado.
	PSSUQ	Questionário de usabilidade de sistema pós-estudo, que resulta em medidas de qualidade do sistema, qualidade da informação e qualidade da interface.
	SUS	Escala de usabilidade do sistema, sendo massivamente adotada deste a década de 1980.
	UMUX	Métrica de usabilidade para experiência do usuário, derivada do SUS alinhada à definição de usabilidade, englobando eficácia, eficiência e satisfação.
	UMUX-Lite	Versão leve baseada no UMUX, com redução de quatro para duas perguntas.
P	ASQ	Questionário após cenário, derivado do PSSUQ cobrindo conclusão da tarefa, satisfação e conclusão e satisfação com informações de suporte.

	SEQ	Questão única simples com derivação reduzida da ASQ .
	SMEQ	Questão do esforço mental subjetivo, pergunta única mediante escala linear segmentada.
	ER	Avaliação de expectativas, é uma variação da SEQ para avaliar a facilidade ou dificuldade em realizar uma tarefa.
	UME	Estimativa da magnitude da usabilidade, refere-se a uma escala linear para dimensionar o estímulo e sua percepção.
	WAMMI	Inventário de medição e análise de website utilizado para medir a atratividade, controle, eficiência, ajuda e aprendizado.
	SUPR-Q	Questionário de avaliação do percentil padronizado de experiência do usuário, voltado para a medição de percepções de usabilidade, credibilidade, aparência e lealdade para websites.

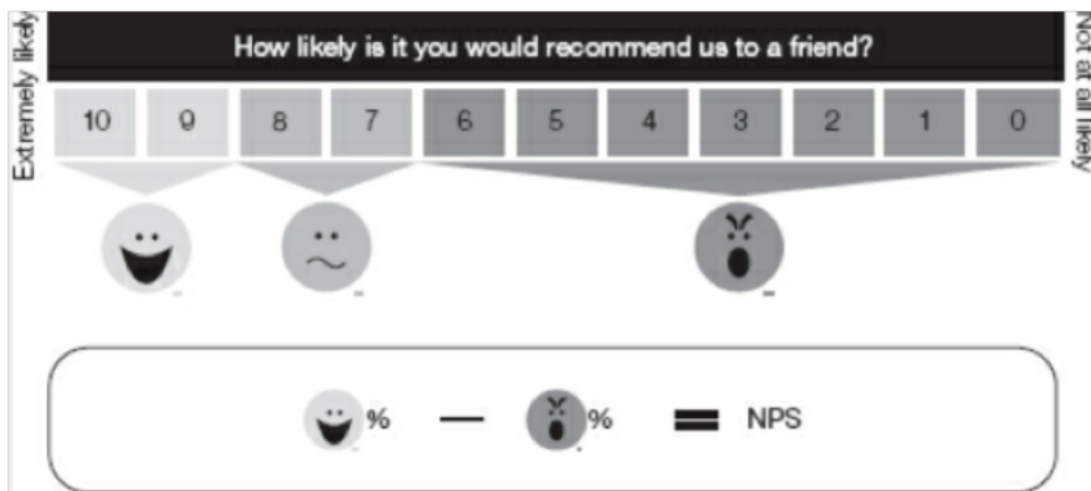
Fonte: compilação do autor de questionários padronizados, agrupados conforme sua aplicação após um teste de usabilidade (pós-teste) ou imediatamente após uma tarefa (pós-tarefa) e suas aplicações principais (Sauro & Lewis, 2016).

Os referidos autores ponderam que deste o início do século XXI ocorre um crescimento no número de questionários voltados para a avaliação de websites com o foco de avaliar a percepção de qualidade, passando pela satisfação até a usabilidade. Mesmo não sendo conhecidos dos pesquisadores, são muito utilizados na avaliação de websites: WEBQUAL, ISQ, GAIS, CSUQ, USE, HQ, EMO, ACSI, CxPi, TAM e o NPS.

O Net Promoter Score, ou NPS, surgiu como uma proposta de redução, ou simplificação frente às complexas abordagens de pesquisa com usuários finais, que exigem equipes especializadas de marketing, psicologia, estatística e programadores, que por muitas vezes tinham como objetivo comum a medição da percepção de lealdade (ou benefício) acerca de determinado produto, serviço ou marca ao qual o usuário, ou consumidor, estaria exposto (Reichheld, 2013). A estrutura proposta pelo autor contempla uma única pergunta, que utiliza na coleta de referência categorizada de 0 a 10 os seguintes parâmetros: valores na faixa de 0 a 6 correspondem a **detratores**, 7 e 8 **passivos** e 9 e 10 **promotores**. Esta estrutura, segundo o proponente, foi validada a partir de diversos estudos comparativos de pesquisas tradicionais previamente realizadas em diferentes segmentos do mercado. A pergunta proposta pelo autor deveria ser uma

aproximação de: *qual a probabilidade de você recomendar esta empresa, ou este produto ou serviço, para um amigo ou colega de trabalho?*

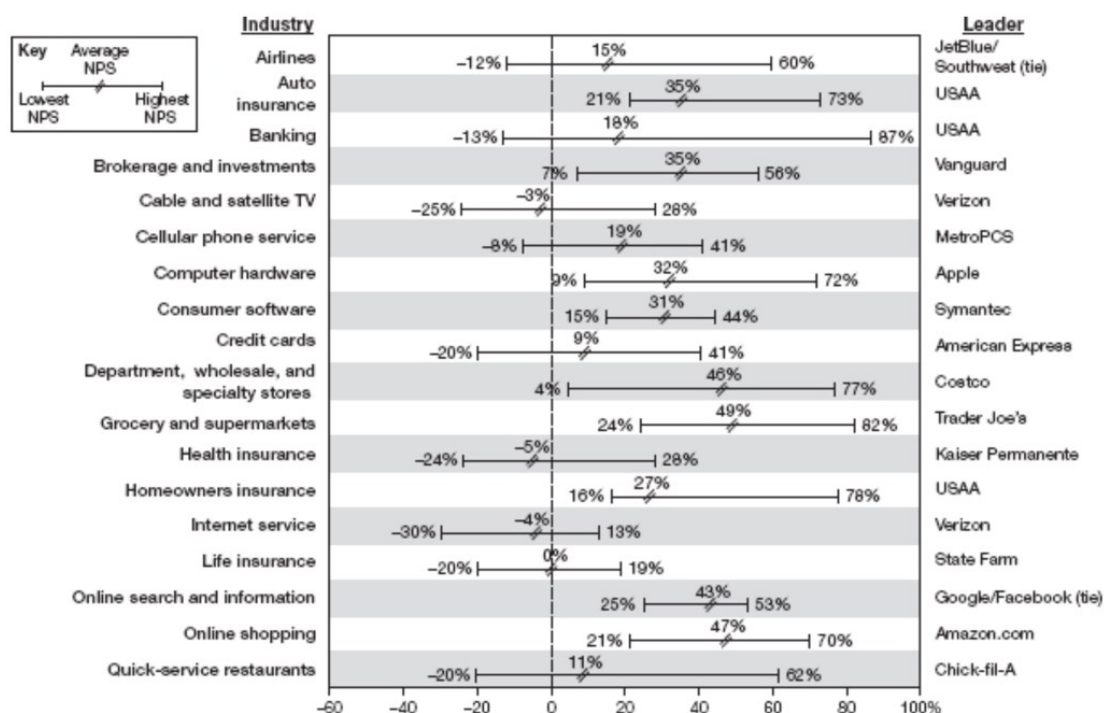
Figura 7 - Categorização de respostas NPS em uma escala 0 a 10



Fonte: Reichheld (2013).

Segundo o autor, diversos fatores como a facilidade de compreensão e resposta assertiva do usuário, assim como facilidade de implementação auxiliaram a tornar a métrica difundida em diversos segmentos, evoluindo de um "Net Promoter score" para um "Net Promoter system of management", ou seja, de uma contabilização de resultado para um formato de gestão. O autor defende, em um estudo de 10 anos em diversos segmentos (realizado em diferentes empresas, inclusive corroborado pela Philips), que empresas com maior quantidade promotores ao invés de detratores, possuem melhores resultados financeiros e crescimento sustentável.

Figura 8 - NPS médio entre diversas indústrias (benchmark).



Fonte: Reichheld (2013).

O NPS, embora largamente utilizado por diversas indústrias, é limitado se avaliado como um valor único, sendo necessário compreender adicionalmente o motivo das respostas dos participantes (Sauro & Lewis, 2016). A fim de avaliar a relação "*lealdade-vendas*" como um resultado da melhoria da usabilidade, foi conduzido um teste usando o SUS e o NPS em mais de 80 produtos de diversas indústrias, em que foi obtida uma correlação positiva de 0.623 (*apud* Lewis, 2012; Sauro 2010), corroborando a tese de que a usabilidade pode impactar o NPS e, por consequência, aumentar o volume de vendas.

Os autores Tullis & Albert (2013) tratam acerca da experiência do usuário (UX) por meio de análises qualitativas e quantitativas. Eles definem que UX (embora definições possam divergir), necessitam de três características para ser um estudo do comportamento do usuário: (i) presença de um usuário envolvido; (ii) um usuário interagindo com um produto; (iii) experiência de interesse do usuário que possa ser observada ou medida. Introduzem, ainda, o conceito de métrica como um meio de medir ou validar um particular fenômeno, bem como

que as métricas de performance são importantes para avaliar a eficácia e eficiência de diferentes produtos.

Quadro 8 – Métricas quantitativas de performance mais usadas na usabilidade

	Medidas	Uso	Opções
Eficácia	Sucesso na tarefa	Mede o sucesso do usuário ao realizar determinada tarefa	Sucesso binário Nível de sucesso
	Erros	Quantifica erros durante a execução da tarefa	-
	Perdição	Mede a quantidade de passos para a realização de uma tarefa	-
Eficiência	Eficiência	Quantidade de esforço necessária para conduzir a atividade	-
	Tempo na tarefa	Quantifica o tempo necessário para completar uma tarefa	-
	Aprendizagem	Medição de como a performance melhora (ou falha) no tempo	-

Fonte: Adaptação de Tullis & Albert (2013).

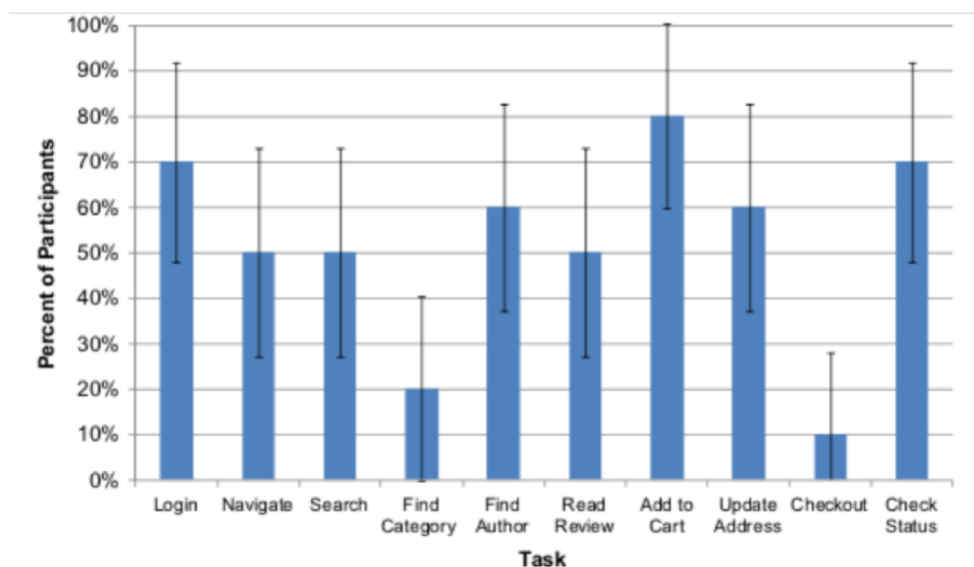
Tullis & Albert (2013) argumentam que métricas são fundamentais na compreensão da magnitude de problemas relacionados à usabilidade, e que apenas saber da existência de um problema em usabilidade não é o suficiente. O autor afirma, ainda, que gerentes sêniores e outros envolvidos em projetos (de produto) normalmente atentam a métricas de performance, especialmente se são apresentadas eficientemente.

A seguir, o detalhamento e demonstração das métricas quantitativas segundo os referidos autores:

a) Sucesso Binário

Forma comum de medição de sucesso na realização de uma atividade: os usuários concluem ou não determinada atividade. Ele é apropriado quando o sucesso de um produto depende do usuário concluir uma atividade ou conjunto de atividades. Um exemplo de produto físico seria a utilização de um desfibrilador para ressuscitar uma pessoa que sofreu um ataque cardíaco, ou mesmo a compra de um livro em um website; ou seja, em ambos os exemplos, sucesso é concluir a atividade, pois o sucesso parcial não é válido.

Gráfico 1 - Percentual de acertos por tarefa.



Fonte: Tulus & Albert (2013).

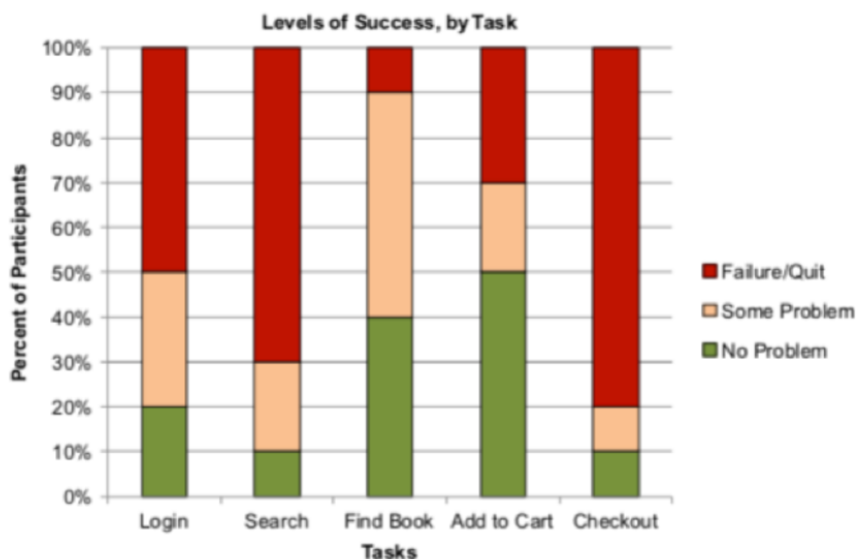
As barras de erro representam, neste exemplo, 90% de intervalo de confiança para cada atividade. O sucesso na realização de uma atividade pode ser representado conforme o gráfico da figura acima. Contudo, é também importante mensurar o insucesso, que pode ser classificado em: Desistência, Moderador intervém, Tempo excedido e Erro.

b) Nível de Sucesso

Utilizado quando existirem níveis parciais entre o insucesso e sucesso, sendo necessário, entretanto, definir os diferentes níveis de sucesso.

Tradicionalmente, são usados três níveis: sucesso completo, mesmo que com assistência; sucesso parcial, mesmo que com assistência; e falha total em que o usuário supõe ter concluído, mas não o fez ou desistiu.

Gráfico 2 - Representação em barras dos três níveis de sucesso.



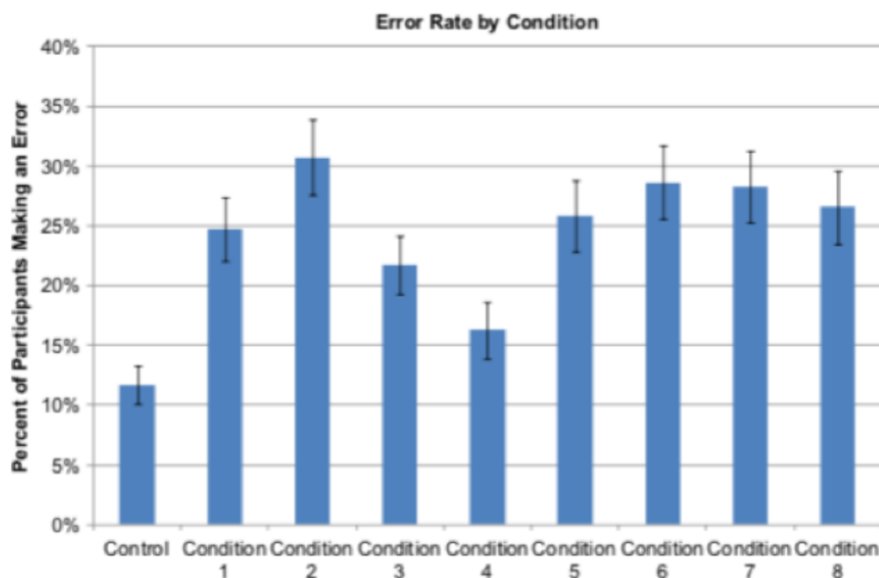
Fonte: Tulis & Albert (2013).

Estes níveis poderiam ser observados em conjunto com os valores resultantes do Tempo na Tarefa, possibilitando mensurar e até mesmo definir o insucesso que consta da métrica Sucesso Binário e, ainda, dos Níveis de Sucesso, para uma análise com tomada de decisão mais criteriosa e assertiva.

c) Erros

Em usabilidade, erros e problemas são questões diferentes, embora estejam relacionados entre si. Problema de usabilidade é a raiz da questão, enquanto erros são os possíveis resultados de um problema. Em essência, erros são frutos de ações indevidas que levam a falhas na execução de uma tarefa (atividade). Eles podem ser tanto identificados como classificados, e sua medição permitirá a compreensão de quando uma ação indevida, ou um conjunto delas, resultará na falha de execução.

Gráfico 3 - Representação de Erros em forma de barras.



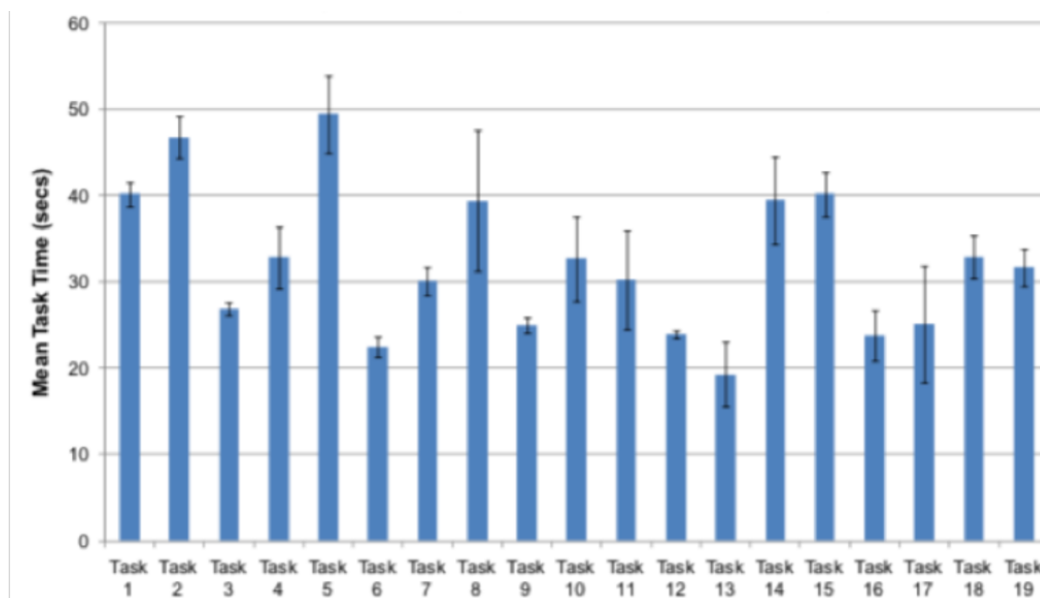
Fonte: Tulus & Albert (2013).

No gráfico acima é representado um exemplo baseado em teclado padrão (de computador), e sua taxa de erros, comparado a taxas de erros de outros modelos (alternativas) de teclados (em versão digital na tela). A atividade ministrada aos usuários foi a entrada de sua senha de acesso, e o único erro delineado foi o seu preenchimento incorreto (erro na digitação). Neste caso apresentado pelo autor, foi utilizado um grupo de controle (Control) e diversas outras condições de uso (*Condition 1 a Condition 8*) do teclado. Ou seja, uma avaliação comparativa na usabilidade do teclado.

d) *Tempo na Tarefa*

Esta métrica é também conhecida como eficiência, em que, na maioria das situações, quanto mais rápida a execução da atividade, melhor. Ela é determinante em produtos utilizados rotineiramente pelos usuários sendo que, no entanto, existem casos de uso em que ocorre o inverso: quanto maior o tempo gasto na atividade, melhor. Como exemplo, cite-se os casos de jogos e websites, em que o objetivo é normalmente 'gastar' ou 'investir' tempo na realização (prazerosa) daquelas atividades.

Gráfico 4 - Barras representando a média do tempo nas atividades.



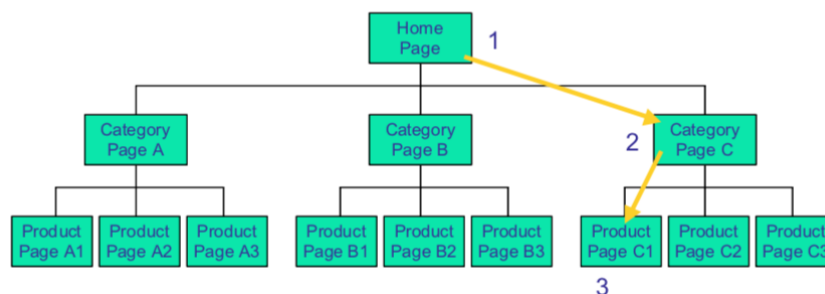
Fonte: Tulus & Albert (2013).

O 'Tempo na Tarefa' mede o tempo gasto entre o início e o fim de uma atividade, e é expresso normalmente em minutos e segundos. Nota-se o uso da média geométrica ao invés da média, que também é bastante utilizada. Todavia, a média pode apresentar resultado alterado, visto que alguns usuários podem utilizar mais tempo do que outros. Desta forma, a média geométrica é mais assertiva ao eliminar distorções de picos e vales de duração.

e) Perdição (ou Lostness)

Métrica de eficiência usada no estudo de comportamento na internet, calculada a partir de três valores: número de páginas *diferentes* na realização da tarefa; número *total* de páginas visitadas enquanto a tarefa é realizada, incluindo as repetidas; número mínimo necessário de páginas para realizar a tarefa (*apud* Smith, 1996).

Gráfico 5 - Quadro representando o número ideal de passos (três) para a realização de uma tarefa.



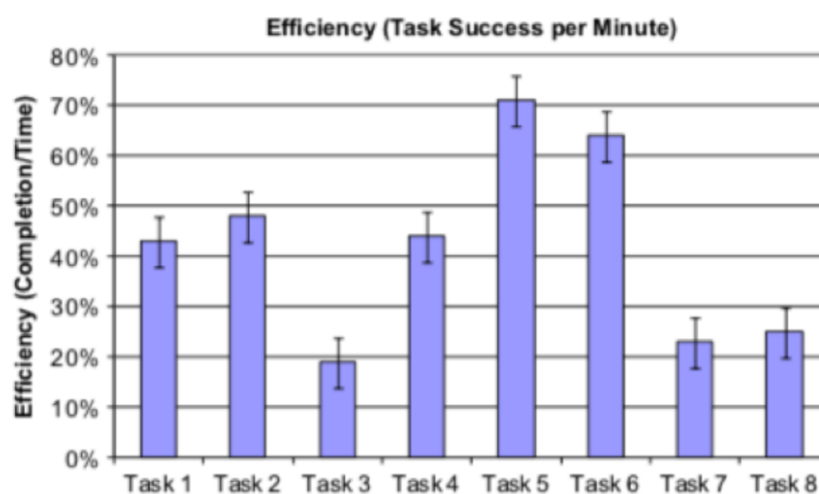
Fonte: Tulis & Albert (2013). Diagrama representando o número ideal de passos (três) necessários para a realização de uma determinada tarefa.

O escore perfeito de perda seria o valor 0 (zero), representando perda mínima. É, contudo, aceitável uma variação de até 0,4, o que representa pouco problema de navegação. Contudo, valores acima de 0,5 significam que o usuário está perdido (*apud* SMITH, 1996).

f) Sucesso na Tarefa

A relação "Sucesso na Tarefa e o Tempo" é uma métrica de eficiência (*apud* ISO/IEC 25062:2006) citada como o "fundamento da medida de eficiência": ela expressa a razão da taxa de conclusão da tarefa pelo tempo.

Gráfico 6 - Gráfico de barras horizontais de esforço no tempo.



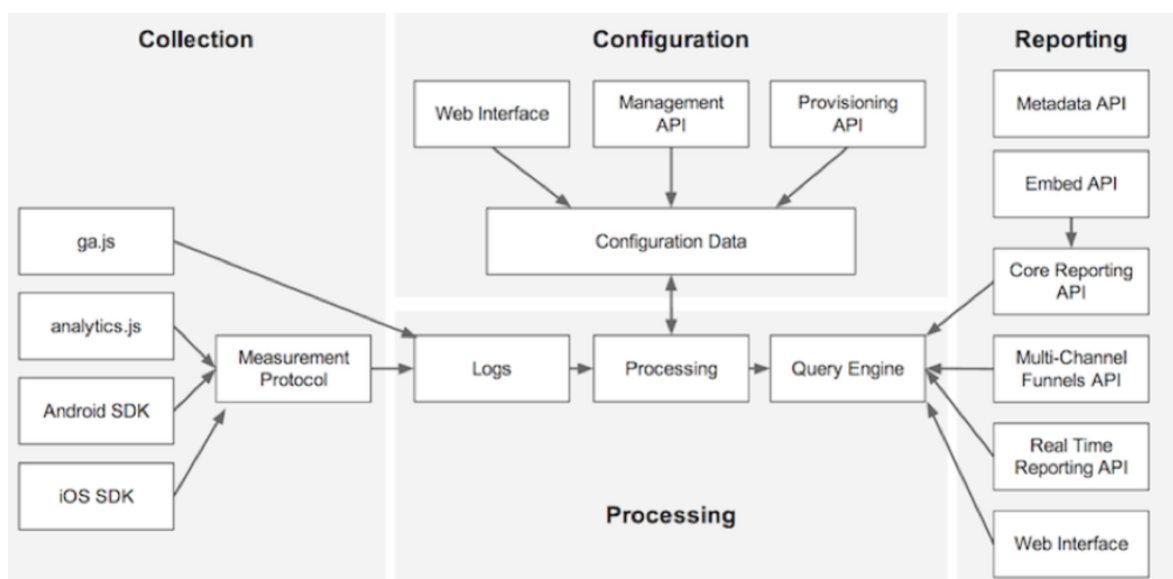
Fonte: Tulis & Albert (2013).

Uma variação da relação supracitada é a "eficiência média", que é computada por meio da contagem de todas tarefas realizadas com sucesso e sua divisão pelo tempo total de todas as tarefas (tanto sucesso como insucesso), permitindo, por exemplo, a compreensão da taxa de atividades pelo tempo.

3.2 FERRAMENTAS E INSTRUMENTOS PARA ANÁLISE DE DADOS

O *Google Analytics* é uma dentre várias reconhecidas ferramentas de análise de acesso a websites e softwares. A ferramenta (gratuita), permite a compreensão do acesso de usuários a seus softwares em tempo real, sejam eles aplicações corporativas, websites, *Apps* para *tablets* ou *smartphone*, dentre outros. O seu funcionamento pressupõe uma configuração inicial e, após a mesma, sucederá em coleta automatizada do fluxo de dados (*Collection*) de múltiplas fontes como, por exemplo: [i] requisições HTTP do usuário em um website, que é o protocolo utilizado para a transferência de hipertexto pela internet (*Hypertext Transfer Protocol*); [ii] informações do sistema (computador) e navegador internet relativos ao usuário; e [iii] internet *cookies*, que são pequenas porções de código enviadas pela instrução do Google para o computador (ou dispositivo que acessa a aplicação) durante o uso da mesma.

Figura 9 - Diagrama de funcionamento do Google Analytics.



Fonte: Google. Diagrama que apresenta a composição em quatro grandes blocos da ferramenta: coleta, configuração, processamento e relatórios. Acessível em:
<<https://developers.google.com/analytics/devguides/platform/>>

A configuração permite programar diversos parâmetros da coleta e processamento para a(s) aplicação(ões). O processamento é executado a partir dos dados armazenados (logs) e coletados por meio das configurações de parâmetros, e o resultado fica, então, disponível para consultas (*query engine*).

A última porção do diagrama da Figura 7 são os relatórios em que são apresentados a análise, cruzamento e visualização de relatórios oriundos daquelas três fontes macro de dados, a exemplo de: data e horário, duração, identificação da página, geolocalização, sexo, idade, fluxo do usuário, dentre diversas outras variáveis qualitativas e quantitativas que podem ser também exportadas para diversos fins.

Quadro 9 – Relatórios disponíveis na plataforma Google Analytics.

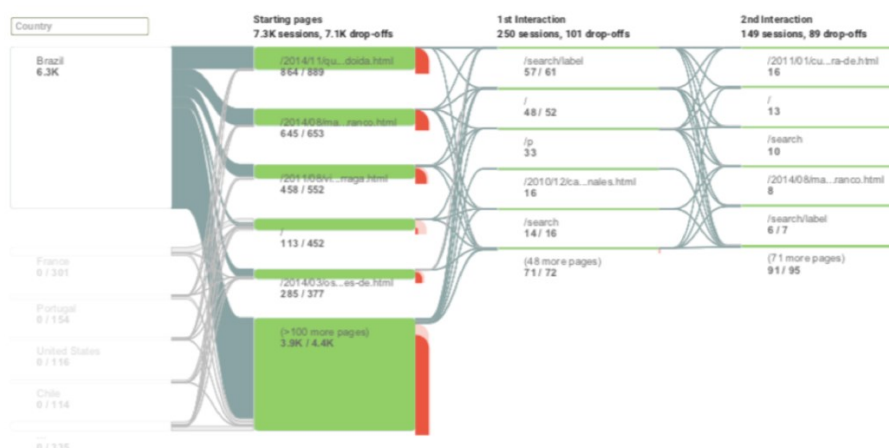
Dados	Objetivo	Natureza de análises e relatórios
Público-alvo	Conhecer o público que acessa a aplicação	Demografia, geografia, grupos de usuários, comportamento de acesso, tecnologia de acesso, benchmark.
Publicidade	Gerenciar a campanha publicitária	Criação de grupos de marketing, informações demográficas de interesse publicitário e segmentação demográfica.
Aquisição	Compreender o ciclo de aquisição, comportamento e conversão de usuários	Canais e seus parâmetros, origem do tráfego (dimensões), dados da campanha publicitária, análises de custo de campanhas, integração com redes sociais e ferramentas de busca
Comportamento	Avaliação da velocidade da aplicação frente a interação com usuários	Latência englobando o tempo de carregamento de cada página, execução e carregamento da página e serviços correlacionados e, ainda, o tempo de resposta do navegador em relação a aplicação.
Conversão	Gerenciar metas de campanhas	Definição de metas, fluxos percorridos pelo usuário na aplicação.

Fonte: Adaptação do autor. Principais relatórios que trazem informações anonimizadas, ou de cunho meramente estatístico, de natureza qualitativa e quantitativa.

Os diversos relatórios do *Google Analytics* permitem tanto a análise diretamente na plataforma, como em outras a partir de exportação ou integração automatizada de dados. Todavia, a capacidade de análise é restrita às métricas e filtros disponíveis na plataforma que possuem como foco a análise de marketing. Um exemplo é o relatório de fluxos percorridos pelos usuários que permite um retrato da jornada dos mesmos. Contudo, não há permissão para análises através dos fluxos, ou seja, não é apresentado como os usuários se comportam ao longo do uso do produto/aplicação.

A aplicação Google Analytics declara sua conformidade com regulações de privacidade internacionais, inclusive a LGPD vigente no Brasil deste 16 de Agosto de 2020, e detalha, adicionalmente, como configurar determinados parâmetros conforme a necessidade da coleta de dados.

Figura 10 - Gráfico de fluxo de usuários em um website.



Fonte: Google Analytics.

A figura acima contabiliza a quantidade de usuários em formato anônimo e o seu fluxo de navegação em um website, representando 100% dos usuários (ou seja a população) que acessaram o website no período entre 3 de junho de 2018 a 3 de junho de 2019. A métrica Taxa de Conversão baseia-se no fluxo de acessos entre uma URL para outra, conforme objetivos traçados de direcionamento de acessos. Por exemplo: Taxa de Conversão (em percentual)

do total de acessos da primeira página de um website de notícias em relação a página de assinatura, ou matéria de destaque, dentre outras.

O *Hotjar*, por sua vez, é uma ferramenta (website) que permite a gravação de dados de uso qualitativos e quantitativos em formato anonimizado em websites, aplicativos web e aplicativos para telefones celulares /tablets. A versão gratuita (2020) do programa possibilita a gravação de tempo de uso da sessão, localização geográfica, quantidade de cliques por página ou eventos, mapa de calor (*heatmap*), automação de pesquisas através de formulários digitais com o usuário final, dentre outras ferramentas e seus respectivos dados qualitativos e quantitativos.

A aplicação permite criar, editar modelos (*templates*) ou usar pesquisas padronizadas como o NPS. As pesquisas criadas pelo usuário, ou editadas, permitem o download de dados em formatos universais (CSV por exemplo). Já a pesquisa NPS é automaticamente calculada na plataforma, mas também permite o download dos dados. A imagem abaixo (Figura 9) apresenta a customização do formulário NPS no bloco da esquerda e a forma com que a pesquisa NPS será apresentada na interface gráfica do usuário final à direita. Adicionalmente, pode ser configurado o comportamento de apresentação da pesquisa.

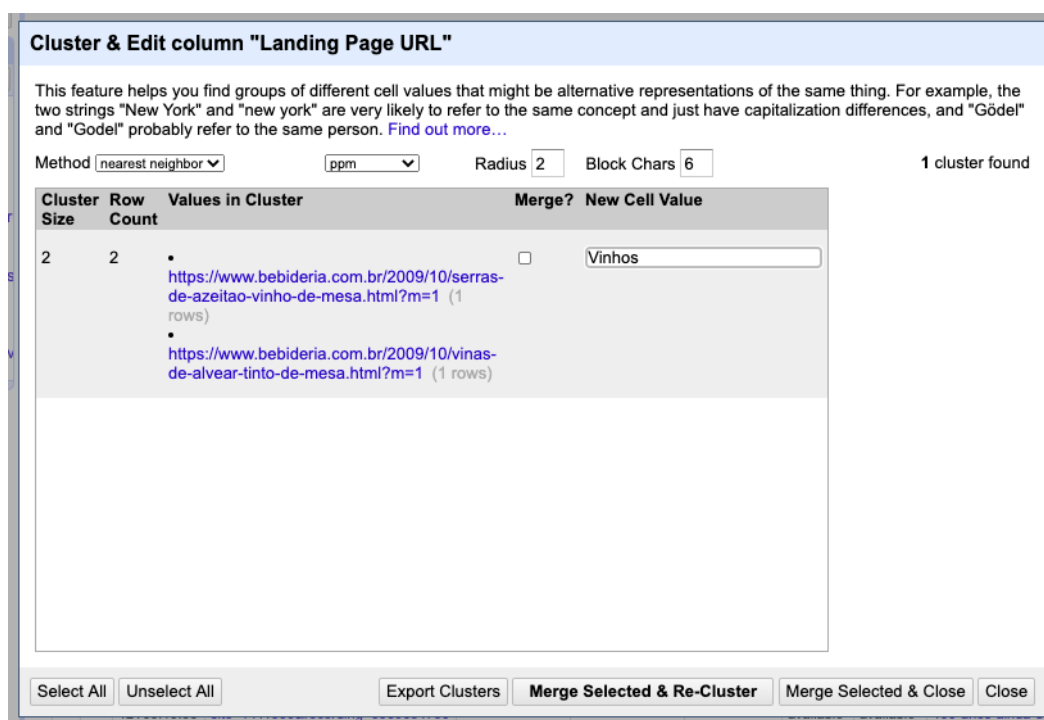
Figura 11 - Imagem de uma pesquisa NPS.

The image shows two side-by-side panels from the Hotjar interface. The left panel is the configuration editor for a Net Promoter Score (NPS) question. It features a header with '+ Add question here', a question type dropdown set to 'Net Promoter Score®', and a 'Required' status with a green checkmark. The main question text is 'How likely are you to recommend Hotjar to a friend?'. Below this, there are labels for 'LOW SCORE LABEL' (Not likely at all) and 'HIGH SCORE LABEL' (Extremely likely). At the bottom, there is a dropdown menu for 'After this question, go to:' with 'Next question' selected, and a '+ Add Question' button. The right panel is a preview of how the survey will appear to the user. It shows the same question text, a horizontal scale from 0 to 10 with the number 10 selected, and the 'Not likely at all' and 'Extremely likely' labels. The Hotjar logo is visible at the bottom left of the preview, and a green 'Next >' button is at the bottom right. Below the preview, there are links for 'Preview mode: Desktop | Phone' and a 'Reset preview' button.

Fonte: Hotjar. Acessível em: <<https://help.hotjar.com/hc/en-us/articles/360021788474--How-to-Setup-your-NPS-Survey>>.

O *Open Refine* é uma ferramenta que automatiza o processo de extração, transformação e carga de dados, ou ETL (do inglês *extract, transform and load*). A ferramenta pode ser incorporada a uma arquitetura de software em Nuvem ou on-premises permitindo, via configuração ou programação, que o processo de ETL ocorra automaticamente, sem a necessidade de supervisão humana. Para tanto, são coletados dados qualitativos e quantitativos, e a transformação é realizada de forma automática. Essa ferramenta possui ainda a capacidade de integrar em diversas fontes de dados através de diversos métodos para, posteriormente, realizar o tratamento, padronização e, mesmo se necessário, a análise supervisionada sobre grandes volumes de dados. Ao final, e se automatizada, ela pode entregar dados para um sistema de armazenamento, análise ou apresentação de dados.

Figura 12 - Open Refine transformando dados por agrupamento.



Fonte: O autor.

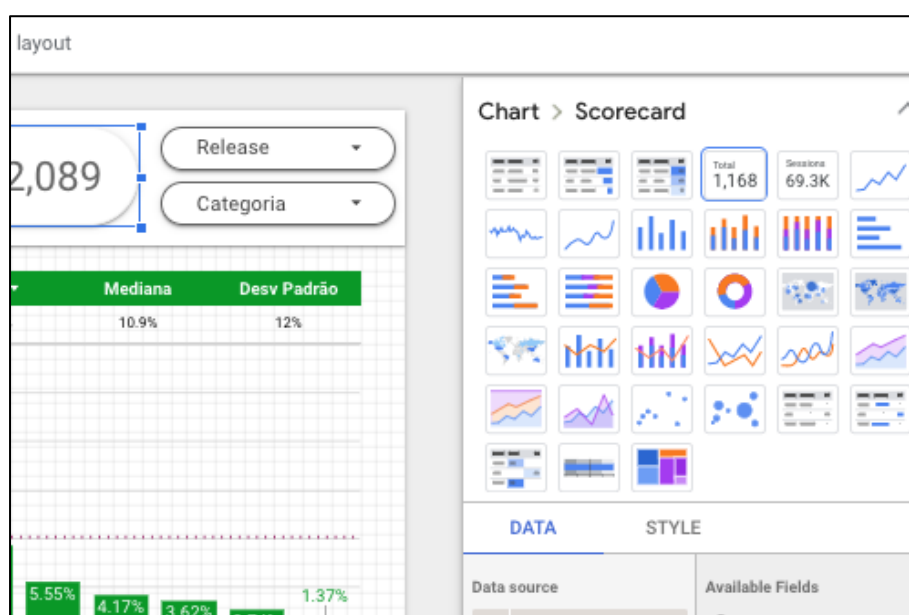
Acima na figura 10, temos em destaque um processo, via algoritmo cognitivo, de agrupamento (ou *clustering*) que possibilita higienizar, transformar,

limpar e padronizar dados grande volume de dados no intervalo de minutos. A ferramenta é gratuita e *open-source* (código aberto) e pode ser baixada da internet e instalada on-premises ou em nuvem. Acesso em: <<https://openrefine.org/>>.

O Google Planilhas (Sheets), por sua vez, é uma planilha digital online que permite a compilação, cálculo e apresentação de informações em formatos gráficos. Possui ainda diversas fórmulas para cálculo matemático, técnicas de tratamento de dados qualitativos, ferramentas de estatística descritiva e possibilita o uso de scripts proprietários (linguagem de programação) de seu desenvolvedor (Google). Os dados são armazenados online (nuvem) e toda plataforma de aplicações de produtividade.

O Google Data Studio é uma plataforma de análise de dados de negócios (*business intelligence* ou BI) com foco em análises de marketing. Possui diversos modelos (templates) de análise e apresentação de dados, inclusive conectores de diversas naturezas, sendo alguns gratuitos e outros mediante pagamento adicional.

Figura 13 - Google Data Studio e gráficos para representação de dados.



Fonte: O autor. Imagem da tela do Google Analytics com foco nos diferentes tipos de gráficos disponíveis na versão (2o semestre de 2020). Os gráficos podem ser customizados e agregar diferentes fontes de dados.

O software é online, nuvem da Google e permite, a depender da fonte de dados, cálculos, apresentações e interação em tempo real. Possui integração nativa com todos os produtos de sua fabricante e tem uma linguagem de programação (*script*) sobre a qual é possível criar fórmulas matemáticas, manipular variáveis qualitativas, dentre outros.

Em suma, diversos formulários, técnicas, ferramentas e métodos apresentados cobrem, cada um com sua particularidade, a coleta de dados que subsidiam a análise de usabilidade de um produto, seja ele físico ou digital, de forma qualitativa ou quantitativa. Contudo, sua operacionalização, conforme demonstrado, pode incorrer na necessidade de um prazo (potencialmente longo) de planejamento do experimento da sua execução, o que inviabiliza seu uso contínuo junto à metodologia Ágil, ao longo dos ciclos de desenvolvimento do produto.

Conforme Iida (1992), a ergonomia é uma disciplina multidisciplinar que, desde seu surgimento no pós-guerra, vem se especializando e é aplicada dentro de processos de desenvolvimento de produtos e máquinas (abordagem em stage-gate), sob a divisão clássica de aspectos biomecânicos, fisiológicos e psicológicos. Posteriormente, a disciplina se torna componente normatizado, conforme citação de Pahl & Beitz *et al.* (2003), dentro da estrutura de projetos de máquinas (e sistemas), ou modelos de referência de desenvolvimento integrado de produtos, conforme defendido por Rozenfeld *et al.* (2006) ou mesmo Baxter (1998). Em contrapartida, Stanton *et al.* (2005) apresenta uma atuação mais ampla (flexível) dos profissionais encarregados da ergonomia em um projeto (pacificando entre o cientista e o prático) que culminam em seis métodos (físicos, psicofisiológicos, comportamentais/cognitivos, equipes, ambientais e macroergonômicos). Em especial foco do estudo, os métodos comportamentais e cognitivos possuem 15 métodos, cada qual com sua particularidade, mas especialmente avaliados no quesito objetividade. Tal quesito leva à ponderação

tanto da capacidade (de cada método) de quantificar resultados, como na possibilidade de utilizar menos tempo para sua implantação e consequente uso (possível aderência aos modelos Ágeis que desenvolvem produtos em ciclos muito curtos - abaixo de 1 mês). Em linha semelhante de amplitude de aplicação e menor rigidez, Rubin & Chisnell (2015) defendem a abordagem do UCD, enfatizando foco (usuário/tarefas), avaliação (com medição) e projeto iterativo, em um formato exploratório, porém objetivo.

Já Sauro & Lewis (2016) defendem que os estudos de experiência do usuário na aplicação (ou UX), embora qualitativos na maioria, permitem análises quantitativas medindo o impacto de um bom design (causa e efeito) e classificam os estudos em: (i) experimental (participantes aleatórios), (ii) quasi-experimental (comparando resultados), (iii) de correlação (relação entre variáveis sem prova de relação de causa e efeito), propondo, adicionalmente, testes estatísticos e questionários padronizados para este fim. Para tanto, são apresentados seis questionários pós-teste e sete pós-tarefa, mas devido à aceleração do mercado, é constante o surgimento de novos questionários, especialmente com foco na avaliação de websites.

Tullis & Albert (2013), na mesma linha de atuação, propõem métricas qualitativas e quantitativas. No estudo por eles realizado o foco está nos dados quantitativos em aderência ao Ágil, com destaque para as métricas eficácia e eficiência (ou performance), cada qual com três ferramentas que permitem mensuração objetiva do comportamento do usuário no uso da aplicação.

As ferramentas apresentadas por Tullis & Albert (2013) para a medição da performance, se integradas a ferramentas tecnológicas (coleta, transformação, cálculo e visualização de dados), poderiam permitir a medição de resultados próximas ao tempo real para fins de comparação com outros produtos (benchmarks), entre versões do mesmo produto (aferição de resultados), tomada de decisão para melhoria contínua do software, dentre outros. Isto porque, parte dos principais desafios para os métodos ergonômicos e de fatores humanos concentram-se no desenvolvimento de: (i) métodos que

se integrem com outros métodos; (ii) relação dos métodos com a teoria da ergonomia; (iii) desenvolvimento de método fáceis de usar; (iv) provisão de evidências de confiabilidade e validação; (v) demonstração de que os métodos levam a intervenções custo-efetivas e (vi) o encorajamento de ética na aplicação dos métodos (Stanton *et al.* 2005).

Os desafios apresentados Stanton *et al.* (2005) e outros autores citados previamente, seja para ferramentas diversas, processos ou métodos, se equacionados sem confrontar os princípios Ágeis (neste caso em particular o Scrum), poderiam ser utilizados, resultando em modelos híbridos conforme já proposto por Blank (2011) no "*Lean Launchpad*".

Portanto, as cinco métricas de performance, quais sejam, Sucesso Binário, Algum Sucesso, Erro (eficácia) e Tempo Médio e Eficiência no Tempo (eficiência), propostas por Tullis & Albert (2013) foram selecionadas uma vez que permitem a coleta de dados em formato contínuo de aplicações online (websites, aplicações SaaS, aplicativos para smartphones ou *tablets*, dentre outros). Ademais, tais métricas permitem, inclusive, o processamento e apresentação automatizada das métricas supracitadas que compõe o constructo KPIsUX. Adicionalmente, para a medição da satisfação do usuário foi selecionado o NPS (net promoter score), por se tratar de um indicador reconhecido pela indústria de software. Contudo, e quando utilizado sozinho, pode apresentar distorções ou mesmo dificuldade para identificar as causas de um bom ou ruim score. Desta forma, a seleção deste conjunto de cinco métricas de usabilidade, em conjunto com o indicador de satisfação, visa facilitar, além da aplicação do método proposto em um experimento de campo, a replicação do experimento apresentado e, por conseguinte, o modelo proposto por outros pesquisadores e profissionais que realizam pesquisa em usabilidade /UX.

4 MATERIAIS E MÉTODOS

O estudo proposto avaliou, em experimento de campo, um website de notícias no formato MPA (*multi-page application*). Não foram utilizadas técnicas de promoção e/ou divulgação digitais (propaganda) a fim de evitar eventual interferência nos resultados. Para tanto, foi utilizado o Scrum (framework) Ágil, não sendo avaliado neste a aplicação em abordagens tradicionais *stage-gate*. O website foi avaliado durante o período de 12 meses (tabela 1) que inicia com o Setup, que é a implantação das ferramentas de medição e pesquisa NPS, prosseguindo com a Release 1, Release 2 e Release 3.

Tabela 1 - Coleta de dados e análise no período de 12 meses.

Fases	Período	KPIsUX (população)	NPS (amostragem)	Erro Amostral (NPS)
Setup	Nov/2019 à Dez/2019	234	16	20%
Release 1	Fev/2020 à Abr/2020	456	22	18%
Release 2	Mai/2020 à Jul/2020	887	15	22%
Release 3	Jul/2020 à Out/2020	522	16	21%

Fonte: O autor. Conjunto de dados da população utilizados para calcular os KPIsU, amostra de dados usados para calcular o NPS, juntamente com o erro amostral associado do NPS naquela população em cada período, ou ciclo temporal.

O conjunto de dados utilizados para análise dos KPIsUX foram todas as páginas (*page title*) de todo o website em conjunto com o seu tempo de uso (*session duration*) associado (compondo desta forma toda a população ou 100% dos acessos de cada período), exportadas do Google Analytics para uma planilha eletrônica conforme procedimento detalhado posteriormente.

Já os dados do NPS foram obtidos por meio de pesquisa aleatória anônima por meio do Hotjar, que apresentava uma pesquisa aos usuários durante o uso (acesso) do website, vide procedimento detalhado posteriormente.

4.1 PLANEJAMENTO DO EXPERIMENTO

Para esta avaliação, foi considerada a Lei Geral de Proteção de Dados Pessoais (LGPD ou LGPDP) nº 13.709/2018, que estabelece padrões para o uso seguro, sigiloso e divulgação de dados pessoais sobre todos os sistemas digitais (software) que operam no Brasil, garantindo os fundamentos relacionados a privacidade apresentados na lei, conforme Capítulo 1, art.2. Desta forma, a coleta de dados foi realizada de forma anonimizada, priorizando-se a seleção das ferramentas e técnicas aderentes a este fim, inclusive com alcance internacional, conforme legislação GDPR (proteção de dados da Europa) e a CCPA (política da Califórnia/EUA).

O experimento foi submetido ao CEP, explanando que o mesmo consiste em uma coleta de dados contínua efetuada por meio de dois conjuntos de dados, que são combinados a fim de se obter os indicadores de usabilidade (constructo KPIsUX) e o NPS, sendo eles: [a] dados oriundos de amostragem aleatória simples na pesquisa NPS; em conjunto com [b] dados contínuos da plataforma automatizada (Google Analytics). Em ambos, a coleta temporal foi realizada de forma anônima, visto que não existem usuários logados no website (não é necessário usuário, nome e senha para acessar o website). Acerca do público-alvo, ou Personas, segundo dados do Google Analytics que coleta dados no site observado neste experimento há mais de 10 anos, os usuários são compostos por adultos, predominantemente do sexo feminino entre 18 e 65 anos, sendo:

- 76% entre 25 e 34 anos;
- 33% entre 35 e 44;
- 24% entre 18 e 24 anos;
- 22% entre 45-55 anos; e

- 10% acima de 65 anos.

Foi pontuado ao comitê de ética que os benefícios diretos, ao website, é a implementação de serviços mais assertivos e alinhados ao público do mesmo. Contudo, o aspecto mais importante é o alinhamento com o objetivo do estudo, que é o desenvolvimento de um processo de medição automatizada (ou mais automatizada possível), que permita ser implantado em larga escala na indústria de software, e deste modo expandindo os benefícios a diversos usuários de nosso mercado, que certamente levará ao aprimoramento dos produtos nacionais (digitais) e sua competitividade no mercado nacional e internacional.

Já os riscos para os participantes do experimento foram minimizados desde a coleta de dados (ambos procedimentos de pesquisa e dados contínuos) até os procedimentos seguintes de tratamento de dados e cálculos automatizados (que também são automatizados, ou seja, não-supervisionados) dos indicadores de performance em usabilidade (KPIsUX) e satisfação (NPS), resultando em informações totalmente desassociadas dos indivíduos (usuários).

Estas particularidades da legislação LGPD, em conjunto com o método de anonimização de dados e a finalidade do estudo - método contínuo de análise estatística de indicadores de performance em usabilidade KPIsUX e o NPS - foi submetida ao CEP (Comitê de Ética), que aprovou o protocolo de pesquisa, conforme Parecer 4.369.865 (Anexo C).

4.2 MODO DE OPERAÇÃO DO PROCESSO PROPOSTO

O processo é composto de duas macro fases: Setup e Releases. A primeira contempla os procedimentos da instalação até a definição de objetivos e metas dos KPIsUX (e porventura o NPS). Já a segunda contempla sucessivos ciclos com procedimentos de medição, avaliação, decisão e desenvolvimento (ou o acrônimo MADD) utilizando os KPIsUX e o NPS como indicadores de resultados para comparação (release a release) integrados ao Scrum.

Segue abaixo a lista descritiva dos procedimentos e atividades dentro das macros fases e, na sequência, o desdobramento esquemático dos procedimentos, técnicas e as ferramentas utilizadas no experimento de campo.

Setup - Definição de Metas:

- Instalação e configuração das ferramentas;
- Parametrização de frequência de publicação da pesquisa;
- Coleta inaugural de dados;
- Extração de dados de acesso;
- Padronização dos dados de acesso;
- Armazenamento de dados;
- Interpretação de resultados parametrização (KPIsUX);
- Programação e parametrização para cálculo;
- Definição de objetivos e metas (KPIsUX e o NPS);
- Decisão de mudanças, melhorias e novas implementações.

Release(s) - Procedimento recorrente de MADD (Medição, Avaliação, Decisão e Desenvolvimento):

- Medição de dados KPIsUX e o NPS;
- Avaliação de resultados dos KPIsUX e o NPS (referentes às decisões do ciclo anterior);
- Decisão de mudanças, melhorias e novas implentações;
- Desenvolvimento de novo ciclo;
- Avaliação e/ou Revisão de Metas e Hipótese de Desenvolvimento (backlogs).

Anteriormente à macro fase Setup, e durante a estabilização de processos e calibração de instrumentos, foram avaliadas alternativas de processos de coleta, extração, transformação e carga de dados (ETLs – *extract, transform and load*), incluindo aplicações de Comunicação Unificada (U.C.) e Gestão de Energia Elétrica. Todavia, e ao longo deste processo, as ferramentas Hotjar, Google Analytics e Google DataStudio sofreram atualizações que permitiram

enxugar os processos, eliminando etapas conforme quadro 10 abaixo, para a computação automatizada (ou o mais automatizada possível) dos KPIsUX e do NPS. Portanto, os mesmos poderiam ser implantados no Scrum em outras configurações, visando maior ou menor automação com os sistemas/produtos em observação, para diversos outros fins.

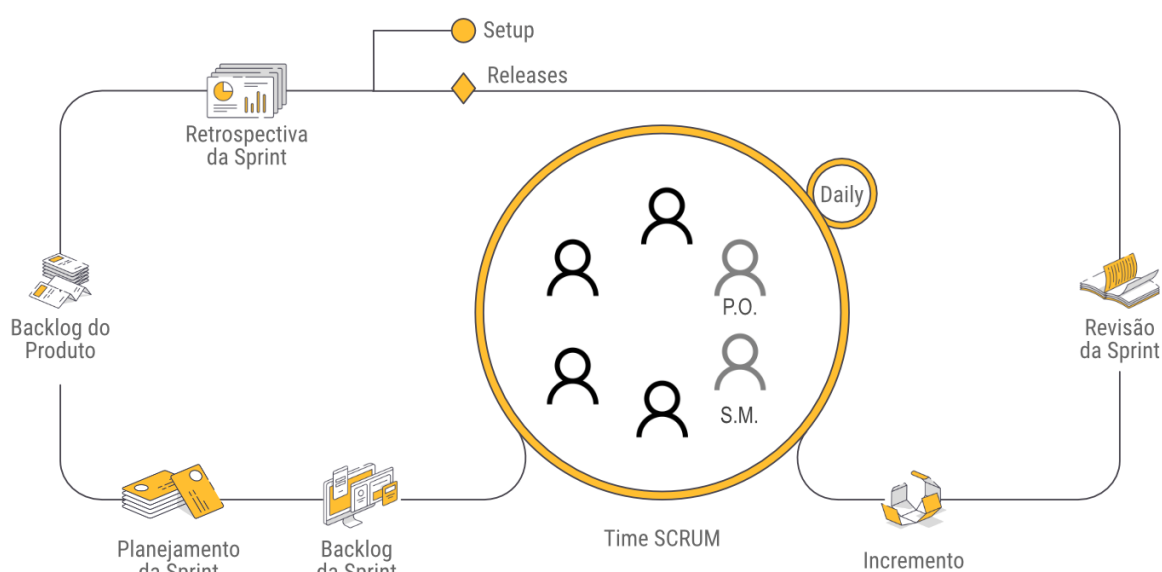
Quadro 10 – Desdobramento das macros fases até as ferramentas.

			Procedimentos	Ferramenta(s)
Macro Fases	Setup		Instalação e configuração	Google Analytics Hotjar
			Definição de frequência de publicação da pesquisa NPS	Hotjar
			Coleta inaugural de dados por determinado período	Google Analytics
			Fusão, armazenamento e padronização de dados	Google Planilhas
				Open Refine
				Hotjar
			Interpretação de dados para parametrização e a programação dos KPIsUX e NPS	Google DataStudio
			Definição de objetivos e metas	N/A
			Decisão a partir de dados (data-driven)	Backlog do Produto
				Backlog da Sprint
			Desenvolvimento em um determinado ciclo	SCRUM
	Releases	R1 a Rn (MADD)	Medição de dados por determinado período	Google Analytics
				Hotjar
			Avaliação dos resultados dos KPIsUX e NPS	Google DataStudio
			Decisão a partir de dados (<i>data-driven</i>)	Backlog do Produto
				Backlog da Sprint
			Desenvolvimento em um determinado ciclo	Scrum

Fonte: O autor.

Nesta proposição, e independentemente da configuração, o Setup e as Releases foram incorporados ao Scrum sem grandes impactos. Isto porque, a primeira fase possui processos adicionais e, em contrapartida, as fases subsequentes seriam repetições cíclicas, até que se verifica uma alteração de paradigma que resulte na reavaliação de objetivos ou reprogramação de metas para os KPIsUX e NPS.

Figura 14 - Representação das macro fases de Setup e Releases no Scrum.

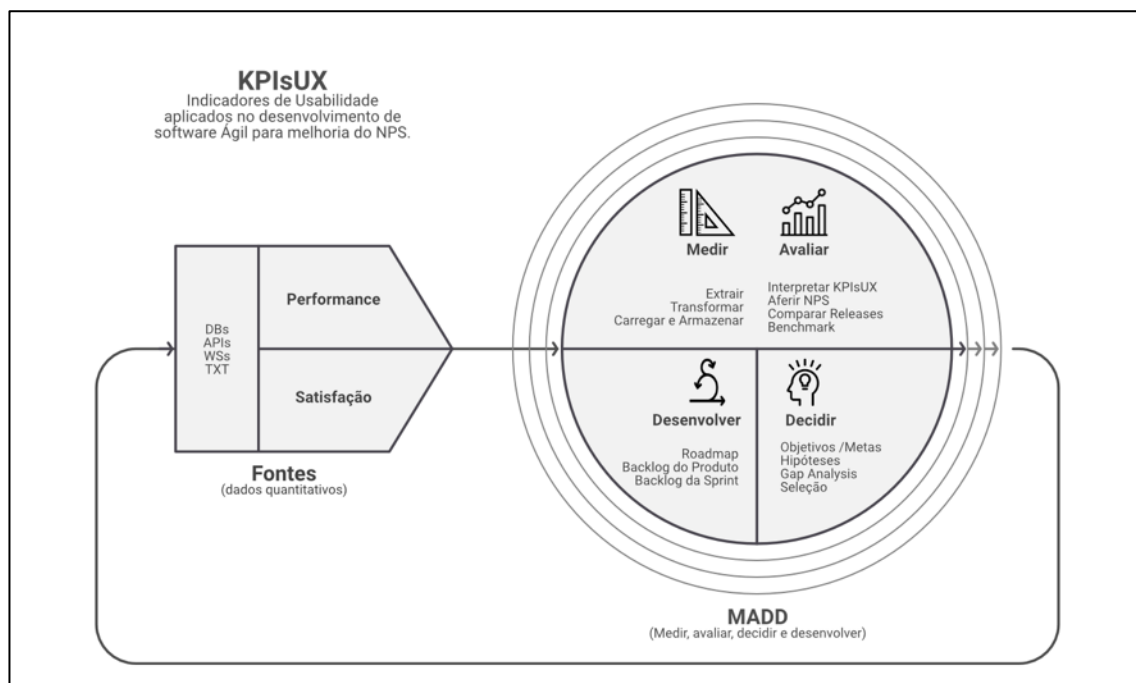


Fonte: O autor. O Setup cobre a instalação de ferramenta e a interpretação dos dados resultantes de um período para subsidiar a tomada de decisão. As Releases cobrem continuamente os períodos dos Ciclos (Releases ou mesmo Sprints) para a tomada de decisão até que seja necessária a redefinição de algum parâmetro de cálculo dos KPIsUX.

Conforme a figura 12 acima, o Scrum é um processo contínuo no tempo, sendo o Setup representado pelo círculo amarelo, um desvio de continuidade que ocorrerá no início da aplicação deste método ou em um tempo futuro em que seja necessária a reprogramação dos KPIsUX. Ao longo do processo de desenvolvimento em que o Scrum é utilizado, será realizado continuamente o rito de avaliação de resultados, representado pelo losango amarelo, para subsidiar a tomada de decisão nos devidos *eventos* do Scrum: Planejamento da

Sprint, Revisão da Sprint ou Retrospectiva da Sprint. Tais eventos, sua vez, afetam os artefatos *Backlog* de Produto ou *Backlog* da Sprint que, após terem seus itens selecionados, desenvolvidos e publicados (produção) resultam no objetivo final do Scrum que é o Incremento (funcional). A abordagem proposta introduz um princípio adicional em que o Incremento deverá possuir ou respeitar (meta) a performance de uso (KPIsUX) em conjunto com satisfação do usuário (NPS) previamente definidos e acordados pelo time Scrum. Isso permitirá que os KPIsUX e NPS sejam medidos e comparados continuamente ao longo das Releases (ou mesmo Sprints), informando a melhoria ou degradação da usabilidade do produto (o todo) ou do Incremento (a parte). Uma forma global de representação da operacionalização do processo proposto (constructo KPIsUX), NPS (ou outro modelo de aquisição de dados de satisfação) integrados ao Scrum (ou outro modelo Ágil) é a visualização do processo por três grandes blocos: [1] entrada de fontes de dados qualitativos; [2] procedimentos cíclicos e [3] reavaliação de objetivos e metas dos KPIsUX e NPS.

Figura 15 - Representação global do processo (KPIsUX).



Fonte: O autor.

Em suma, e segundo a representação global, a Macro fase Setup é composta pela entrada de dados (Fontes) que alimentam um primeiro procedimento MADD (representado pelo primeiro círculo interno em cinza) que, posteriormente, segue uma segunda Macro fase Release 1, que repete o procedimento MADD (representada pelo segundo círculo externo ao cinza). Por sua vez, este último se repete indefinidamente até que uma mudança de paradigma deflagre a necessidade de revisão de objetivos, metas e, finalmente, reinicie o Setup para seguir as Releases.

4.3 APLICAÇÃO NO EXPERIMENTO

No experimento de campo, após o Setup, foi realizada uma coleta de dados pelo período de dois meses e, ao seu término, foram executados os processos de procedimentos de “Fusão, armazenamento e padronização de dados”, “Interpretação de dados para parametrização e a programação dos KPIsUX e NPS” e “Definição de objetivos e metas” (conforme quadro de desdobramento de processos e ferramentas). A definição de objetivos e metas ocorreram de forma interativa pelos membros do time Scrum. Ainda, houve a particularidade do uso do tempo, que é uma medida da eficiência transformada em medida de eficácia devido à natureza do website que é uma revista digital – convencendo que baixo tempo de uso (ou leitura) significava erro. Desta forma, qualquer leitura de conteúdo abaixo de 30 segundos foi considerada um “Erro”, que conforme proposto por Tulis & Albert (2013), softwares como jogos ou revistas digitais se beneficiam de maior tempo de uso (engajamento). Como resultado, foram definidos tanto os parâmetros a serem implementados enquanto código na ferramenta de visualização (Google Data Studio) como os objetivos de melhora dos KPIsUX (Sucesso Binário, Algum Sucesso, Erros, Tempo Médio e Eficiência no Tempo) e a identificação do NPS percebido (satisfação). Abaixo o quadro de parâmetros de cálculo automatizado das métricas de performance.

Tabela 2 - Parâmetros resultantes do Setup para cálculo dos KPIsUX e definição de metas das Releases subsequentes

KPIsUX	Sucesso Binário	Algum Sucesso	Erros	Tempo Médio (s)	Eficiência no Tempo
Parâmetro	Acima de 60 segundos	Entre 29 e 59 segundos	Abaixo de 30 segundos	-	-
Valores resultantes	60 (ocorrências)	23 (ocorrências)	149 (ocorrências)	00:01:32 (HH:MM:SS)	28,8% (percentual)

Fonte: O autor. A primeira linha contém os parâmetros para cálculo após iterações de alinhamento entre os membros do Scrum, que resultam na segunda linha em valores quantizados das métricas.

As seguintes convenções de nomenclatura foram utilizadas no experimento a fim de manter o devido sigilo de informações do website:

- Página inicial: Home
- Funcionalidade existente antes do experimento: FNC1
- Funcionalidade (nova proposta ao longo do desenvolvimento): FNC2
- Funcionalidade (nova proposta ao longo do desenvolvimento): FNC3
- Categorias de conteúdo (existentes antes do experimento): CTG1 a CTG7

Ainda, foi determinado como não-escopo: as atividades e processos de formulação, priorização e definição de hipóteses de desenvolvimento (e funcionalidades). A métrica de Perdição (eficácia) foi removida da análise (conforme explicação posterior) e a Curva de Aprendizado foi obtida como um subproduto, como pode ser observado posteriormente na ilustração referente ao Painel de Acompanhamento dos KPIsUX e NPS.

Os dados para avaliação, download e testes, juntamente com o painel de visualização e análise da proposição estão disponíveis respectivamente em:

- Dados para avaliação: <<https://www.kaggle.com/datasets/gustavojota/ux-kpis> >
- Painel de visualização: <<https://datastudio.google.com/s/kqS76OwoEYs>>

5 RESULTADOS E DISCUSSÕES

Após a etapa intitulada Setup, o time de desenvolvimento do website (composto por designers de produto e gráficos, analistas de sistema e programadores) mediu (por meio de coleta de dados do período) e avaliou os valores de performance e satisfação respectivamente, por meio dos KPIsUX (variável independente) e do NPS (variável dependente), que eram o reflexo do período anterior ao início do experimento, em que outras variáveis e processos operavam aleatoriamente para o subsídio à tomada de decisão no Scrum (variável de controle). Ou seja, o Setup permitiu medir a eficiência dos usuários no website anteriormente ao experimento aqui proposto (uso dos KPIsUX em conjunto com o processo MADD).

Após a interpretação dos resultados de performance do Setup anteriormente ao experimento, métricas de usabilidade (KPIsUX) e satisfação (NPS), foi definido um conjunto de alterações a serem desenvolvidas e implementadas (variável interveniente) com o objetivo de elevar, ou seja, melhorar ambos os valores de KPIsUX e NPS.

Backlog priorizado para desenvolvimento a partir da decisão do Setup:

1. **Reestruturar Home:** Alteração da arquitetura de informação da primeira página, composta por matéria da capa (principal), matérias de destaque (secundárias) e espaço para publicidade digital.
2. **Redesign da marca e aplicações:** Renovação estilizada de cores e geometria da marca, assim como sua aplicação e assinatura ao longo do website.
3. **Ferramenta de busca:** Inclusão de ferramenta de busca na “Home” e páginas subsequentes com parâmetros adicionais que permitiam a busca estruturada por categorias de matérias.
4. **Estrutura de acesso rápido:** Alteração do menu lateral de acesso rápido às matérias de um formato histórico com nome das matérias,

para um formato histórico resumido da quantidade de matérias disponíveis naquele período (mês a mês).

5. **Diretriz de redação:** Uma vez que foram identificadas diversas matérias com tempo reduzido de leitura (engajamento), que resultavam em “Erros” nos KPIsUX, foi definido que as matérias deveriam possuir conteúdo o suficiente para leituras pontuais acima de 2 minutos a fim de resultar na elevação da taxa de “Sucesso” global no site e nos KPIsUX). Permitindo desta forma maior tempo de engajamento do usuário e potencial de receita oriunda de publicidade digital.
6. **Categorização de matérias:** Foram reestruturados todos os índices de matérias e marcadores associados para um formato que reduziu a quantidade de categoria de matérias, visando facilitar a localização de matérias e inclusive a avaliação através dos KPIsUX e MADD, para melhoria contínua do website.
7. **Cards das matérias:** O layout dos (boxes ou quadros) das matérias da primeira página (tanto o destaque como as secundárias) foi alterado maximizando a foto de cada uma delas em contraposição à redução da quantidade de texto introdutório contido nestes “Cards”.
8. **Alteração de layout do site:** Conjunto de alterações na “grid” e consequente disposição de diversos elementos principais e secundários do layout do site (afetando todas as suas partes).

Na Release 1, primeiro ciclo da macro fase Releases, foram utilizados *sprints* semanais, iniciando-se em Fev/2020, para implementação contínua das alterações listadas no *Backlog do Produto*. Ao longo do ciclo foram observadas, a partir das medições, as alterações e implementações que ao seu final apresentaram o incremento de todas as métricas de Eficácia (Sucesso, Algum Sucesso e inclusive os Erros com elevação de 46%). Já na aferição do NPS foi observada uma queda de 36% em relação ao Setup, atingindo 32 pontos (anteriormente eram 50 pontos).

É importante observar que os “Erros” não reduziram. Uma hipótese desenvolvida pelo time de desenvolvimento, foi que existiam (e existem) uma grande quantidade de matérias de conteúdo reduzido, ou seja, que permitem rápida leitura, sendo estas relativas ao dicionário de bebidas, onde as mesmas possuíam tamanho aproximado entre 370 a 500 caracteres. Foi decidido que as mesmas não deveriam ser “apagadas” ou reescritas, visto que o website possuía histórico elevado de SEO (Search engine optimization) que permite aos buscadores internet encontrar conteúdo a partir de buscas (ex. Google ou Bing).

A avaliação principal, diante de elevada queda do NPS, foi que o ciclo de desenvolvimento entregou melhorias aquém da expectativa dos usuários, inclusive com pequena elevação de apenas cinco segundos de tempo médio de uso (Setup era de um minuto e trinta e dois segundos, elevando-se para um minuto e trinta e sete segundos. Sendo este o principal objetivo traçado inicialmente, visando maior tempo de engajamento. No entanto, houve elevação da eficiência no tempo, saltando de 67% para 150% (oriunda mais da elevação das taxas de “Sucesso” do que apropriadamente de maior tempo de uso /duração média). A decisão para a próxima release ou ciclos de desenvolvimento foi, portanto, de buscar a recuperação do NPS (satisfação percebida pelos leitores /usuários), revisitando as entregas da release anterior visando um conjunto de melhorias e diretrizes trivialmente corretivas).

Backlog priorizado para desenvolvimento a partir da decisão da Release 1:

1. **Refinar layout de todo o site:** Foram implementados melhorias no acabamento do layout relativas a “grids” visando melhor espaçamento, ou “respiro” entre elementos em conjunto com o uso de palheta de cores com menor contraste entre os diversos elementos.
2. **Nova alternativa de Home:** A primeira página sofreu alteração no formato com maior destaques de fotos em detrimento de texto (foi utilizado apenas o título nas matérias secundárias), inclusive sua disposição e quantidade (redução da quantidade de matérias secundárias). Já a matéria principal possuía foto ampliada, contendo título e corpo do texto parcial (com link para a matéria na íntegra).

3. **Aprimorar acesso às categorias:** O acesso rápido às matérias antigas foi novamente reestruturado, tornando-se um elemento secundário na lateral do layout.
4. **Implementar funcionalidades:** Foram desenvolvidas duas novas funcionalidades, uma primeira relativa (FNC 2) buscando maior interatividade dos usuários finais entre si (uma espécie de socialização) que deveria resultar em maior tempo de uso (engajamento) e uma segunda (FNC 3) relativa ao recrutamento de usuários e o agendamento para que os usuários utilizassem a primeira funcionalidade (FNC 2).

Na Release 2, também utilizando *sprints* semanais, foram observadas ao longo e término da medição a elevação de 89% de Erros (ressaltando que trata da leitura de conteúdo de reduzido tamanho e não das falhas inerentes ao uso do website), a elevação de permanência (tempo médio saindo de um minuto e trinta e sete para dois minutos e vinte e seis segundos) puxada pela nova funcionalidade (FNC 2 conforme gráfico de barras do relatório) implementada, com a recuperação do NPS que atingiu 67 pontos, elevando-se a patamar superior à percepção dos usuários finais, conforme medição no Setup.

Pondera-se que este NPS tão elevado poderia ser atribuído majoritariamente a respostas posteriores ao uso da FNC 2, haja visto que o tempo médio desta funcionalidade atingiu uma média de quatro minutos e trinta e um segundos no período.

Neste segundo ciclo, a avaliação geral foi que o design do site havia estabilizado (mediante aferição do NPS), ou poderia estar sendo influenciado em demasia pela funcionalidade FNC 2, e que diante destes resultados novas alterações e aprimoramentos poderiam ser implementados. Foi decidido, então, o aprimoramento e desenvolvimento de um serviço piloto a partir de nova funcionalidade (FNC2), haja vista o elevado tempo médio de quatro minutos e trinta e um segundos e, pontualmente, atingiu nos níveis de sucesso cento e oito sucessos, doze algum sucesso, mas que apresentou em contrapartida cinquenta e dois erros. Foi ainda decidida a eliminação de outra funcionalidade (FNC 3) que apresentava apenas trinta e cinco segundos de tempo médio.

Backlog priorizado para desenvolvimento a partir da decisão da Release 2:

1. **Pivotar funcionalidade:** A funcionalidade FNC 2 foi aprimorada visando tanto a redução da taxa de erros como a busca por maior tempo de engajamento do usuário final, através de um serviço acessório de delivery de produtos (vinhos para degustação às cegas).
2. **Desativar funcionalidade:** Observou-se que a funcionalidade FNC 3 praticamente não atraiu usuários, portanto foi desativada, sendo parcialmente incorporada à funcionalidade FNC 2 para cumprir com o propósito de realização de pedidos e agendamento das degustações de vinhos.

Após a implementação das hipóteses de desenvolvimento na Release 3, que é a última da série, foi possível o incremento do NPS com elevação de 12%, atingindo 75 pontos e a elevação dos Erros em 73%. A funcionalidade pivotada (FNC 2) atingiu seis minutos de tempo médio e pontualmente, nos níveis de sucesso, vinte e cinco sucessos, um algum sucesso e seis erros apenas. Foram realizadas outras avaliações e tomadas de decisão para os próximos ciclos que, contudo, não foram acompanhadas devido ao término do período de análise do experimento proposto.

Tabela 3 – Resultados dos KPIsUX e NPS.

Métricas	Eficácia			Eficiência		Satisfação
	Sucesso Binário	Erros	Algum Sucesso	Tempo Médio	Eficiência no Tempo	NPS
Setup	62	149	23	01:32	67%	50
Release 1	147	218	91	01:37	150%	32
Release 2	318	414	145	02:26	217%	67
Release 3	209	222	91	02:13	157%	75

Fonte: O autor. Compilação dos resultados obtidos ao longo de 12 meses do Setup (Nov/2019) até a última Release 3 (Nov/2020).

Na conclusão do período de acompanhamento de 12 meses foi possível aferir uma média de satisfação de 59 pontos, com flutuações ao longo do período, estas refletindo a “qualidade percebida” ou a voz do consumidor em relação às decisões do projeto do produto (website). Os KPIsUX tiveram o papel de medir (explicitar) de forma quantitativa o comportamento do usuário na utilização do website (tarefa real). Ainda, vale ressaltar que as métricas Eficiência no Tempo e Eficiência Tempo Médio, relacionadas à duração (tempo de leitura) corresponderam às observações de Tullis & Albert (2013) em que as publicações digitais buscam maior tempo de permanência de seus usuários (para conotação de sucesso), ou seja, maior tempo de leitura. No experimento, o mesmo pôde ser comprovado especialmente no desenvolvimento, manutenção e aprimoramento (ou cancelamento) de determinadas funcionalidades e características (layout e marca) do website a partir e ao longo das Releases (CTG7, FNC2 e FNC3).

Figura 16 - Painel de Acompanhamento dos KPIsUX e NPS.



Fonte: O autor.

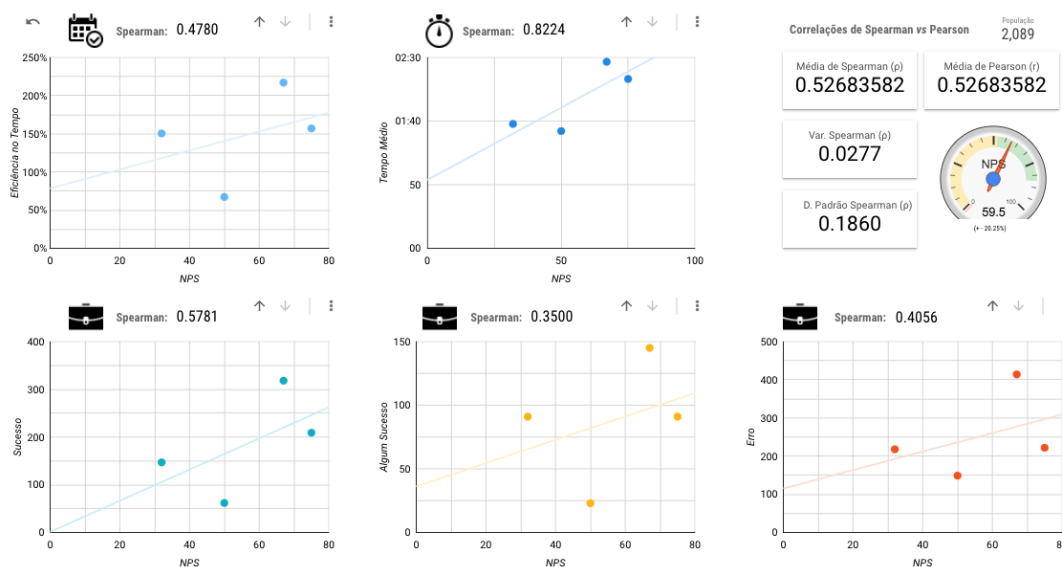
No início do experimento foi construído um Dashboard, ou painel de acompanhamento, conforme figura 16 acima utilizando a ferramenta de BI (ou business intelligence neste caso do Google Data Studio), para permitir a melhor compreensão da relação de causa e efeito entre os indicadores de usabilidade

e a satisfação dos usuários de cada entrega do desenvolvimento (e em cada release). Esta ferramenta, possui recursos de aplicação de múltiplos “filtros” ou seleções de características que permite ao designer e outros profissionais envolvidos no desenvolvimento, através de cliques, navegar através dos dados para melhor compreensão de fenômenos pontuais (ao nível de apenas uma funcionalidade) ou em uma visão mais ampla (nível macro).

O painel apresenta na lateral esquerda superior as macrofases (Setup e Releases) em um gráfico de barras horizontal expondo o valor do resultado NPS no gráfico de barras horizontal, sendo o *backlog* do produto a ser desenvolvido nas *sprints* de cada ciclo em um formato de tabela. Ao centro, estão dispostos em gráficos de barras verticais as métricas de eficiência no tempo e tempo médio e, abaixo delas, está a curva de aprendizado (todos em tons de azul), que são obtidos através dos tempos médios apresentados em uma série temporal. Finalmente, na direita estão os níveis de sucesso, representados por um gráfico de pizza que apresenta a proporção de sucesso/algum sucesso/erros de cada ciclo e mesmo funcionalidade, juntamente com o gráfico de barras horizontal empilhado em percentual que permite visualizar a composição de cada funcionalidade e seus valores intrínsecos.

A partir dos dados consolidados do período de experimento foi realizada uma análise multivariada de fatores (Manly & Alberto, 2019) para compreender o grau de correlação entre as métricas (ou parcelas) que continuam o constructo KPIsUX (cada uma das variáveis independentes) frente à variável moderadora NPS que o estudo visa observar. O Dashboard traz ainda os gráficos de dispersão de cada uma das métricas versus o NPS ao longo dos quatro ciclos de desenvolvimento, com a possibilidade de alteração da ‘medida’ de ciclo por cada funcionalidade/categoria para avaliação.

Figura 17 - Análise entre as métricas do KPIsUX e o NPS.



Fonte: O autor.

Por conseguinte, foi aplicada a correlação de *Spearman* (ρ) individualmente em cada métrica *versus* o NPS e o consolidado do estudo obtido a partir da média sobre todas as correlações obtidas. Não foi realizada a análise de correlação entre as próprias métricas de usabilidade uma vez que todas são cálculos desdobrados da mesma fonte de dados (duração e local) e o objetivo era a observação da flutuação do NPS mediante os KPIsUX.

Tabela 4 – Resultados da correlação entre métricas e o NPS.

Correlação vs Métricas	Eficácia			Eficiência	
	Sucesso	Algum Sucesso	Erros	Sucesso no Tempo	Tempo Médio
Spearman ($\rho=0.5268$)	0.5781	0.3500	0.4056	0.4056	0.8224

Fonte: O autor.

As correlações das métricas de usabilidade foram individualmente obtidas ao longo dos ciclos de desenvolvimento em relação aos NPS dos mesmos ciclos.

Para avaliação consolidada foi usada a média das correlações individuais em que a correlação de Spearman e Pearson apresentaram-se equivalentes.

Em suma, o modelo proposto de uso dos KPIsUX para melhoria da satisfação percebida através do NPS, resultou em correlação forte ($\rho=0.5268$) como resultado da média aritmética (e $\rho=0.5032$ em média geométrica) das variáveis que compõe o constructo KPIsUX *versus* a variável NPS. Assim, a hipótese apresentada foi validada ao considerar-se as delimitações da natureza do objeto em estudo, ferramentas usadas no experimento e quantidade de ciclos de desenvolvimento observados, além de outras variáveis de ambiente relativas à qualidade da equipe de desenvolvimento, tecnologia utilizada e, certamente, outras relativas a mercado.

Finalmente, o processo contínuo de medição, avaliação, decisão e desenvolvimento (MADD), que utiliza aqui dos KPIsUX como facilitadores para tomada de decisão em um projeto sob a metodologia Ágil, mostrou-se robusto e eficiente para utilização em conjunto com o Scrum. Certamente, este processo MADD seja a maior contribuição do presente estudo, por permitir ampliar o leque de aplicações conforme o objeto, seja ele relativo à aspectos de design, como de fatores humanos e ergonomia.

6 CONCLUSÃO E CONSIDERAÇÕES FINAIS

O processo proposto revelou-se capaz de traduzir métricas de usabilidade, que explicam como o trabalho real é executado pelo usuário final, em indicadores de performance de uso do software (KPIsUX), que utilizados dentro de um ciclo com procedimentos de medição, avaliação, decisão e desenvolvimento (MADD) permitiram guiar o PDP em busca da melhoria de satisfação dos usuários (NPS). Para tanto, foi apresentado um modelo de referência para utilização em conjunto com o Scrum que utilizou um conjunto de dados (data set) facilmente obtido a partir de ferramentas padrões e gratuitas de mercado, eliminando desta forma barreiras financeiras e de conhecimento para adoção do processo proposto.

Desta forma, é possível adotar os processos, procedimentos e técnicas de forma continuada, seja em abordagens estatísticas por amostragem ou censo, a partir da coleta de dados de toda a população, ou ainda, um misto de ambos como demonstrado neste experimento de campo. Destaca-se ainda que o processo MADD e os indicadores KPIsUX podem ser usados tanto em conjunto (como neste experimento) como de forma independente, utilizando de outras ferramentas quantitativas ou qualitativas de acordo com a preferência dos envolvidos no desenvolvimento ou mesmo de acordo com o deságio do projeto. Adicionalmente, as ferramentas demonstradas no experimento foram selecionadas por apresentarem tanto a capacidade de lidar com dados possuidores de características de big data (elevado volume, variedade e velocidade), como também por serem padrões da indústria internacional que facilitam a aquisição daquela estrutura de dados proposta, sua disseminação e escalabilidade de acordo com a natureza da aplicação em estudo.

Ainda, as ferramentas são aderentes às legislações nacionais e internacionais de sigilo e privacidade, possuindo licenciamento flexível, gratuito ou mesmo de código aberto. Desta forma, foram eliminadas barreiras de custo ou tecnologias proprietárias para a adoção em escala pelas empresas que busquem o aprimoramento de seus produtos com aferição da satisfação dos

seus usuários e clientes, sobre os esforços dispendidos no projeto do produto / serviços baseados em software sob desenvolvimento ágil.

O processo proposto foi aplicado em uma revista digital, baseando-se nas premissas da natureza da aplicação em que um maior tempo de permanência era o objetivo do negócio (permitindo maior tempo de permanência do usuário no site e consequente exposição à publicidade e propaganda de terceiros). Desta forma, é importante que no futuro seja avaliada a aplicação do s KPIsUX e do MADD em outras naturezas de aplicações, tais como: sistema de gestão corporativos, aplicações de produtividade, comunicação, e-commerce, jogos e outros, especialmente se utilizaram a abordagem de arquitetura SPA (*single-page application*) em que toda a aplicação ‘roda’ em uma mesma página (URL). Ainda, um dos delimitadores foi a execução do experimento de campo em um sistema que opera na modalidade de computação em nuvem pública (*cloud*) e sob arquitetura MPA (*multi-page application*) em que existem diversas páginas sobre as quais as partes do software é escrito.

Portanto, é necessário avaliar a adaptação, com consequente seleção de ferramentas ou técnicas específicas que permitam o uso do processo proposto de KPIsUX em instalações locais (*on-premises*) e mesmo embarcadas em veículos (outros meios de transporte ou equipamentos), visto que atualmente eles possuem sistema de computação (isolados) similar a uma instalação local tradicional (e inclusive à *fog-computing*, que é o *pré-processamento distribuído oriundos de diversos dispositivos, antes de ser levado à nuvem*).

Importante ponderar que, devido à natureza do website do experimento, a duração foi utilizada como parâmetro para estabelecer regras para as métricas de Eficácia: Sucesso, Algum Sucesso e Erros. Contudo, em softwares de outras naturezas, conforme Tulis & Albert (2013), a duração pode não ser sempre o parâmetro mais adequado para convencionar as medidas relativas de eficácia, podendo inclusive ser necessário explicitar o Sucesso Binário ao contrário do aqui foi demonstrado, em que o Sucesso foi parte indissociável do nível de sucesso.

Conforme o volume de dados para cálculo dos KPsUX e NPS crescem, seria uma evolução natural o incremento de novas fontes de dados ao "data set" apresentado para que sejam aprimorados os conhecimentos acerca de preferência de uso, comportamento estatístico dos usuários (*behavior analytics*), Personas Analytics (Personas baseadas em dados estatísticos de múltiplas fontes por exemplo).

É importante considerar, ainda, a aplicação *down-top* (da micro funcionalidade à macro funcionalidade ou função global) do processo MADD e KPIsUX, (ou outras métricas de análise) de forma a fragmentar a aplicação em parcelas menores, ou seja a análise por módulos, funcionalidades, interfaces gráficas (telas), dentre outros, conforme defendido por REICHHELD (2019), em que o NPS é usado como um sistema e não um indicador isolado (haja visto que diversos autores citados no estudo questionam o uso do NPS isoladamente). Tal aplicação fragmentando o software ou produto em parcelas menores, inclusive, viabilizaria o uso da métrica Perdição (*Lostness*) apresentada por Tullis & Albert (2013) visto que a mesma efetua a medição pontual de cada funcionalidade em relação ao todo, e que nesta proposição foi descartada por ser muito trabalhosa de operacionalizar e de ser medida e acompanhada, tornando-se uma potencial confrontadora dos princípios ágeis e da dinâmica 'acelerada' do Scrum.

Finalmente, e em face de todo o observado, é recomendado o aprofundamento de estudo e aplicação das abordagens Ágeis nas disciplinas do design, visto que abarcam, de forma flexível e adaptativa, soluções complexas que estão em constante mutação. Adicionalmente, também seria indicada a adoção de cadeiras básicas de lógica, programação e análise de dados para instrumentalizar os futuros profissionais na condução de análises quantitativas em usabilidade e na tomada de decisão referente ao projeto de produtos baseando-se cada vez mais em dados e fatos.

REFERÊNCIAS

- AMAZON. **What is Cloud Computing**. <<https://aws.amazon.com/what-is-cloud-computing/>>. Acesso em: Maio de 2019.
- BAXTER, Mike. **Projeto de Produto**. 3a edição. São Paulo: Blucher, 2011.
- COOPER, Robert G. **Winning at New Products: Creating Value Through Innovation**. New York: Basic Books, 2001.
- CUSUMANO, Michael A. **How Microsoft makes large teams work like small teams**. Cambridge: MIT Sloan Management Review, 1997.
- Google. Disponível em:
<<https://developers.google.com/analytics/resources/concepts/gaConceptsTrackingOverview>> Acesso em: Abril de 2019.
- Hewlett Packard. **Definição de No Local e Computação em Nuvem**.
<<https://www.hpe.com/br/pt/what-is/on-premises-vs-cloud.html>>. Acesso em: Maio de 2019.
- HOME OF SCRUM. Disponível em: <<https://www.scrum.org/resources/what-is-scrum>>. Acesso em: Outubro de 2022.
- Iida, Itiro. **Ergonomia - Projeto e Produção**. São Paulo: Blucher, 1992.
- JOTA, Gustavo; BUSON, Marcos A. ; *et al.* **A otimização do Processo de Planejamento do Projeto de Produtos Industriais, pela Segmentação de Mercado e Arquitetura do Produto**. In: Congresso Brasileiro de Gestão de Desenvolvimento de Produtos (CBGDP). Belo Horizonte: 2007.
- JOTA, Gustavo; MAYNARDES, Ana Claudia; **"Discussão sobre o uso de metodologias ágeis no desenvolvimento de produtos industriais e digitais"**, p. 2196-2210 . In: Anais do 13º Congresso Pesquisa e Desenvolvimento em Design (2018). São Paulo: Blucher, 2019.
- LEI FEDERAL <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>
- LORENZ, C. **The Design Dimension**. Oxford: Basil Blackwell, 1986.
- Página inicial do Google Analytics**. Disponível em:
<https://support.google.com/analytics/answer/7421425?hl=pt-BR&ref_topic=3544907> Acesso em: Abril de 2019.
- PMI. PMBOK. **A Guide to the Project Management Body of Knowledge**:

Fifth Edition. Newton Square: Project Management Institute, 2013.

REICHHELD, Frederick. **The One Number You Need to Grow.** Disponível em: <<https://hbr.org/2003/12/the-one-number-you-need-to-grow>>. Acesso em: Abril de 2019.

ROZENFELD, Henrique. FORCELLINI, Fernando A. AMARAL, Daniel C.

TOLEDO, José C. SILVA, Sergio L. ALIPRANDINI, Dário H. SCALICE, Régis K.

Gestão de Desenvolvimento de Produtos: Uma referência para a Melhoria do Processo. São Paulo: Saraiva, 2006.

STANTON, Neville. HEDGE, Alan. BROOKHUIS, Alan. SALAS, Eduardo.

HENDRICK, Hal. **Handbook of Human Factors and Ergonomics Methods.**

Boca Raton: CRC PRESS LLC, 2005.

SUTHERLAND, Jeff. SCHWABER, Ken. **Guia do Scrum: Um guia definitivo para o Scrum** - As regras do jogo. Disponível em:

<<https://www.scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-Portuguese-Brazilian.pdf>>. Acesso em: Novembro de 2017.

TULIS, Tom; ALBERT. Bill. **Measuring the User Experience - Collecting, Analyzing and Presenting Usability Metrics.** 2nd Edition. Waltham-MA:

Morgan Kaufmann - Elsevier, 2013.

GLOSSÁRIO

Sprint: ciclo curto de desenvolvimento de software.

Release: ciclo de desenvolvimento de software longo que engloba diversos ciclos menores (sprint).

Cloud: modalidade de computação em nuvem.

Dataset: conjunto estruturado de dados selecionados para determinada aplicação (cálculo e/ou análise dentre outros).

Waterfall: traduzido de 'Cascata', é uma abordagem, ou metodologia de desenvolvimento linear “*stage-gate*”.

Stage Gate: etapa final ou intermediária no desenvolvimento com avaliação de resultados.

Data-Driven: em tradução livre, processo guiado por dados.

ANEXO A – ESTRUTURA DE DADOS DO NPS

O “*dataset*” (estrutura de dados) abaixo é um exemplo da coleta de dados do Hotjar para o cálculo do NPS. A macro fase demonstrada abaixo é o de Setup:

Number	User	Date Submitted	Country	Source URL	Device	Browser	OS	O quanto você indicaria a Bebideria.com.br para um amigo?	Do que você sente falta na Bebideria?
1	314ae6fa	2019-11-08 9:43:41	Brazil	https://www.bebideria.com.br/2009/10/vinas-de-alvear-tinto-de-mesa.html?m=1	phone	Chrome Mobile 78.0.3904	Android 8.1.0	10	Não há nada a reclamar. Excelente estoque.
11	01163ebe	2019-11-25 2:41:53	Brazil	https://www.bebideria.com.br/?m=1	phone	Chrome Mobile 78.0.3904	Android 8.1.0	7	Não sei

De todos os campos coletados pelo software, apenas os campos abaixo foram utilizados no experimento:

- "O quanto você indicaria a Bebideria.com.br para um amigo?": que é a pergunta NPS.
- Date Submitted: data de coleta do dado.

ANEXO B – FORMULÁRIO DE PESQUISA ANÔNIMA COM USUÁRIOS

A pesquisa NPS (Net Promoter Score) aplicada no experimento é estruturada a partir de derivações de: *"How likely is it that you would recommend this company, or this product or service, to a friend or colleague?"* Reichheld (2019). Em tradução (do autor): Qual a probabilidade de você recomendar esta empresa, ou produto, ou serviço, para um amigo ou colega? Tais perguntas foram apresentadas aos usuários aleatoriamente, sem a exigência de qualquer tipo de identificação (armazenadas no Hotjar, podendo ser recuperadas para revisão).

Tabela de perguntas

Ciclo	Pergunta	Resposta
Setup	O quanto você indicaria a Bebideria.com.br para um amigo?	Baixo escore: Não Indicaria!
		Alto escore: Sim indicaria muito!
Release 1	Curtiu o visual e a página de cursos? Indicaria a Bebideria.com.br a um amigo?	Baixo escore: Não Indicaria
		Alto escore: Indicaria Sim!
Release 2	O quanto você curtiu o novo visual e a Degustação Interativa da Bebideria?	Baixo escore: Não curti
		Alto escore: Curti muito!
Release 3	O quanto você curtiu o novo visual e a Degustação Interativa da Bebideria?	Baixo escore: Não curti
		Alto escore: Curti muito!

Imagem da tela da Revista Digital com um dos questionários, inclusive com mensagem de política de privacidade e explicações sobre a pesquisa NPS.



ANEXO C – PARECER CONSUBSTANCIADO DO CEP (RESUMIDO)

Título da Pesquisa: Aplicação de KPIsU no desenvolvimento ágil de software para a melhoria do NPS.

Pesquisador: GUSTAVO LOPES RODRIGUES JOTA

Versão: 2

CAAE: 38257320.1.0000.0118

Instituição Proponente: Centro de Artes

Patrocinador Principal: Financiamento Próprio

Considerações Finais a critério do CEP: A Diretoria APROVA o Protocolo de Pesquisa e informa que, qualquer alteração necessária ao planejamento e desenvolvimento do Protocolo Aprovado ou cronograma final, seja comunicada ao CEP via Plataforma Brasil na forma de EMENDA, para análise sendo que para a execução deverá ser aguardada aprovação final do CEP. A ocorrência de situações adversas durante a execução da pesquisa deverá ser comunicada imediatamente ao CEP via Plataforma Brasil, na forma de NOTIFICAÇÃO. Em não havendo alterações ao Protocolo Aprovado e/ou situações adversas durante a execução, deverá ser encaminhado RELATÓRIO FINAL ao CEP via Plataforma Brasil até 60 dias da data final definida no cronograma, para análise e aprovação. Lembramos ainda, que o participante da pesquisa ou seu representante legal, quando for o caso, bem como o pesquisador responsável, deverão rubricar todas as folhas do Termo de Consentimento Livre e Esclarecido - TCLE - apondo suas assinaturas na última página do referido Termo.

Situação do Parecer: Aprovado

Necessita Apreciação da CONEP: Não