

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE INFORMÁTICA



## **Usabiliphy: Criação e Validação de um Dataset para Estudar a Relação entre a Usabilidade Percebida e Sinais Fisiológicos**

André Nilo Franco Batista Monteiro

**Mestrado em Engenharia Informática**

Dissertação orientada por:  
Prof. Doutor Manuel João Caneira Monteiro da Fonseca



## Agradecimentos

Em primeiro lugar, gostaria de expressar agradecimento ao Professor Doutor Manuel João Fonseca, o meu orientador, pela sua excelente capacidade de organização e apoio contínuo. O seu interesse pelo trabalho e a ajuda mostraram-se essenciais para superar os desafios que surgiram ao longo do processo. Durante a escrita da dissertação, a sua disponibilidade, a abordagem crítica e a introdução de novas ideias incentivaram-me a aperfeiçoar este trabalho até ao mais ínfimo detalhe.

Gostaria também de expressar o meu agradecimento aos amigos e colegas, destacando Eduardo Madeira, Diogo Martins e Khristina Korinets, pela sua disponibilidade e interesse em apoiar e acompanhar o desenvolvimento deste projeto. Além disso, quero agradecer a todos aqueles que generosamente dedicaram o seu tempo e paciência ao processo de testes do Usabiliphy. A contribuição deles foi fundamental.

Por último, mas não menos importante, desejo expressar a minha gratidão à minha família, em particular aos meus pais, irmã, avó, avô e tios pela motivação constante e apoio incondicional ao longo deste trabalho.



*À minha tia Helena.*



## Resumo

As técnicas usadas para registar a usabilidade, experiência de utilização, emoções, etc., percecionadas pelos utilizadores quando usam aplicações, são maioritariamente baseadas em questionários. Apesar destes permitirem obter resultados sólidos, podem tornar-se monótonos, cansativos e até exigentes, podendo desencadear sensações negativas nos utilizadores e consequentemente, afetar os resultados dos questionários tornando-os menos confiáveis.

Uma forma de ultrapassar estes problemas seria através da inferência da usabilidade, experiência de utilização e das emoções percecionadas pelos utilizadores a partir dos seus sinais fisiológicos, enquanto estes usam as aplicações. Para se conseguir fazer esta inferência é necessário ter informação fidedigna que permita relacionar os sinais fisiológicos com as medidas de usabilidade e experiência de utilização. Este trabalho é um primeiro passo neste sentido e com ele pretendemos produzir um *dataset* composto pelos sinais fisiológicos dos utilizadores e pelos resultados de questionários standard, enquanto estes interagem com aplicações interativas. Para isso, realizamos duas experiências com utilizadores onde estes executaram um conjunto de tarefas em aplicações interativas, enquanto recolhíamos os seus sinais fisiológicos e as suas respostas a questionários standard.

No desfecho do estudo, procedemos à validação dos dados por meio de uma análise estatística, a qual revelou que não havia diferenças estatísticas entre as duas experiências. A partir dessa validação, constituímos um *dataset* que inclui sinais fisiológicos (eletroencefalografia, fotopletimografia, acelerômetro, giroscópio) e as pontuações dos questionários standard (*System Usability Scale*, *User Experience Questionnaire Short*, *Single Ease Question*, *NASA task Load Index*, *Usability Metric for User Experience Lite*, *Self-Assessment Manikin*).

Este *dataset* permitirá, no futuro, investigar a relação entre os dois tipos de dados e posteriormente criar modelos para inferir a usabilidade percecionada a partir dos sinais fisiológicos, sem termos de recorrer ao preenchimento de questionários, melhorando as sessões de testes com utilizadores e reduzindo o tempo destas.

**Keywords:** Usabilidade, Sinais Fisiológicos, Dataset, Experiência de Utilização, Emoções



## Abstract

The techniques used to assess usability, user experience, emotions, etc., perceived by users when using applications are predominantly based on questionnaires. Although these questionnaires yield solid results, they can become monotonous, tiresome, and even demanding, potentially triggering negative feelings in users and consequently affecting the reliability of the questionnaires.

One way to overcome these issues would be to infer usability, user experience, and perceived emotions from users' physiological signals while they use the applications. To achieve this inference, it is necessary to have reliable information that allows for a correlation between physiological signals and usability and user experience measures. This work represents a first step in this direction, and our goal is to create a dataset composed of users' physiological signals and the results of standard questionnaires while they interact with interactive applications. To accomplish this, we conducted two user experiments in which participants performed a set of tasks in interactive applications while we collected their physiological signals and their responses to standard questionnaires.

At the end of the study, we validated the data through statistical analysis, which revealed that there was no statistical differences between the two experiments. Based on this validation, we constructed a dataset that includes physiological signals (electroencephalography, photoplethysmography, accelerometer, gyroscope) and scores from standard questionnaires (System Usability Scale, User Experience Questionnaire Short, Single Ease Question, NASA Task Load Index, Usability Metric for User Experience Lite, Self-Assessment Manikin).

This dataset will enable future research to investigate the relationship between these two types of data and subsequently create models to infer perceived usability from physiological signals, eliminating the need for questionnaire completion, improving user testing sessions, and reducing their duration.

**Keywords:** Usability, Physiological Signals, Dataset, User Experience, Emotions



# Conteúdo

<b>Lista de Figuras</b>	<b>xiv</b>
<b>Lista de Tabelas</b>	<b>xv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	1
1.2 Objectivos . . . . .	1
1.3 Aborgadem Usada . . . . .	2
1.4 Estrutura do documento . . . . .	2
<b>2 Fundamentos Teóricos e Trabalho Relacionado</b>	<b>5</b>
2.1 Questionários Standard . . . . .	5
2.1.1 User Experience Questionnaire (UEQ) . . . . .	5
2.1.2 System Usability Scale (SUS) . . . . .	6
2.1.3 Technology Acceptance Model (TAM) . . . . .	7
2.1.4 Questionnaire for User Interaction Satisfaction (QUIS) . . . . .	8
2.1.5 Single Ease Question (SEQ) . . . . .	9
2.1.6 After-Scenario Questionnaire (ASQ) . . . . .	9
2.1.7 The Usefulness, Satisfaction, and Ease of Use Questionnaire (USE) . . . . .	10
2.1.8 NASA Task Load Index (NASA-TLX) . . . . .	10
2.1.9 Subjective Mental Effort Questionnaire (SMEQ) . . . . .	11
2.1.10 Post Study System Usability Questionnaire (PSSUQ) . . . . .	12
2.1.11 Usability Metric for User Experience(UMUX) . . . . .	12
2.1.12 Usability Metric for User Experience LITE (UMUX-LITE) . . . . .	13
2.1.13 Self-Assessment Manikin (SAM) . . . . .	13
2.1.14 Resumo dos Vários Questionários . . . . .	14
2.2 Sinais Fisiológicos . . . . .	17
2.2.1 Eletroencefalografia . . . . .	17
2.2.2 Fotopletismografia . . . . .	18
2.2.3 Acelerómetro . . . . .	18
2.2.4 Giroscópio . . . . .	18
2.3 Emoções . . . . .	18

2.3.1	Modelos de Emoções . . . . .	19
2.4	<i>Datasets com Sinais fisiológicos</i> . . . . .	20
2.4.1	<i>DEAP Dataset</i> . . . . .	20
2.4.2	<i>AMIGOS Dataset</i> . . . . .	22
2.4.3	<i>DREAMER Dataset</i> . . . . .	23
2.4.4	<i>GAMEEMO Dataset</i> . . . . .	24
2.4.5	<i>Overview dos Datasets</i> . . . . .	25
2.5	Sinais Fisiológicos e Questionários Standard . . . . .	25
2.6	Correlação entre Emoções e Métricas de Usabilidade . . . . .	27
2.7	Relação Entre os Fundamentos Teóricos e o Trabalho Relacionado . . . . .	28
2.8	Síntese . . . . .	28
<b>3</b>	<b>Criação do <i>Dataset</i></b> . . . . .	<b>31</b>
3.1	Metodologia . . . . .	31
3.1.1	Aplicações Interativas . . . . .	31
3.1.2	Questionários Standard Usados . . . . .	32
3.1.3	Participantes . . . . .	33
3.1.4	Equipamento . . . . .	34
3.2	Experiência 1 . . . . .	34
3.2.1	Procedimento Experimental . . . . .	34
3.2.2	Dados Recolhidos . . . . .	37
3.3	Experiência 2 . . . . .	38
3.3.1	Aplicações Tecnológicas . . . . .	38
3.3.2	Procedimento Experimental . . . . .	41
3.3.3	Dados Recolhidos . . . . .	42
3.4	<i>Dataset</i> . . . . .	43
3.5	Síntese . . . . .	46
<b>4</b>	<b>Validação do <i>Dataset</i></b> . . . . .	<b>47</b>
4.1	Descrição da validação . . . . .	47
4.2	Resultados . . . . .	47
4.3	Análise dos Resultados e Discussão . . . . .	48
4.4	Síntese . . . . .	53
<b>5</b>	<b>Conclusões e Trabalho Futuro</b> . . . . .	<b>55</b>
5.1	Resumo da dissertação . . . . .	55
5.2	Contribuições e Limitações . . . . .	56
5.3	Trabalho Futuro . . . . .	56
<b>Abreviaturas</b>		<b>58</b>

<b>Bibliografia</b>	<b>63</b>
<b>Índice</b>	<b>64</b>
<b>A Comissão de Ética</b>	<b>65</b>
A.0.1 Formulário de Pedido . . . . .	65
A.1 Folheto Informativo Experiência 1 . . . . .	75
A.2 Folheto Informativo Experiência 2 . . . . .	77
A.3 Procedimento do Trabalho . . . . .	79
<b>B Questionários Experiência 2</b>	<b>83</b>
B.1 Questionário Demográfico . . . . .	83
B.2 Questionário para Cada Tarefa . . . . .	85
<b>C Tarefas das Experiências</b>	<b>99</b>



# Lista de Figuras

2.1 Estrutura de escala do UEQ baseada no estudo [32]. . . . .	6
2.2 Itens do questionário SUS traduzidos para Português pelo estudo [24]. . . . .	7
2.3 O Self-Assessment Manikin (SAM). . . . .	14
2.4 Dispositivo Muse 2 e os seus sensores. . . . .	17
2.5 Emoções principais estabelecidas no <i>Plutchik's emotion wheel</i> . . . . .	19
2.6 Emoções principais estabelecidas no <i>Circumplex Model of Affect</i> do estudo [35]. .	20
2.7 Estrutura para a criação dos <i>Datasets</i> . . . . .	25
3.1 Questionario UEQ-S retirado do estudo [34]. . . . .	32
3.2 Quadrantes SUS e UEQ-S. . . . .	33
3.3 Grupo de sites da primeira experiência. . . . .	35
3.4 <i>Scatter Plot</i> Comparativo: Médias do UEQ-S com as Médias do SUS da Experiência 1. . . . .	36
3.5 <i>Scatter Plot</i> Comparativo: Médias da qualidade pragmática com as Médias do SUS da Experiência 1. . . . .	36
3.6 <i>Scatter Plot</i> Comparativo: Médias da qualidade hedónica com as Médias do SUS da Experiência 1. . . . .	37
3.7 <i>Scatter Plot</i> com a escolha dos 12 sites. . . . .	38
3.8 Sinais do Muse 2 e as suas configurações. . . . .	39
3.9 Aplicação BlueMuse e as definições usadas. . . . .	39
3.10 Grupo de sites da segunda experiência. Cada site possui um número únicos. . .	41
3.11 <i>Scatter Plot</i> Comparativo: Médias do UEQ-S com as Médias do SUS da Experiência 2. . . . .	43
3.12 <i>Scatter Plot</i> Comparativo: Médias da qualidade pragmática com as Médias do SUS da Experiência 2. . . . .	44
3.13 <i>Scatter Plot</i> Comparativo: Médias da qualidade hedónica com as Médias do SUS da Experiência 2. . . . .	44
3.14 Procedimento para a produção do <i>Dataset</i> . . . . .	45
3.15 Abordagem hierárquica para organizar o <i>Dataset</i> . . . . .	45
4.1 Resultados dos testes de <i>Mann-Whitney U</i> e médias do SUS e UEQ-S em cada site e em cada experiência. . . . .	48

4.2 Resultados dos testes de <i>Mann-Whitney U</i> do SEQ, NASA-TLX, UMUX-LITE e SAM em cada site e em cada experiência. . . . .	49
4.3 Resultados das médias do SEQ, NASA-TLX, UMUX-LITE e SAM em cada site e em cada experiência. . . . .	50
4.4 <i>Scatter Plots</i> Comparativos: Valores dos vários questionários da primeira e segunda experiência. . . . .	50
4.5 Referencial que representa os dados do SAM, SUS e UEQ-S. Os dados do SUS são representados por figuras geométricas e podem variar de tamanho conforme a categoria (exemplo: "F-", "F" e "F+"). Os dados do SAM são representados pelos valores no referencial, enquanto os do UEQ-S são identificados por cores. . . . .	51
4.6 <i>Boxplot</i> com os valores do SUS da primeira e segunda experiência. . . . .	52
4.7 <i>Boxplots</i> com os valores do UEQ-S e SEQ da primeira e segunda experiência. . . . .	52
4.8 <i>Boxplots</i> com os valores do UMUX-LITE e NASA-TLX da primeira e segunda experiência. . . . .	53

# **Lista de Tabelas**

2.1 Resumo dos questionários standard com os pontos de medição, escalas, vantagens e desvantagens. . . . .	15
2.1 Resumo dos questionários standard com os pontos de medição, escalas, vantagens e desvantagens. . . . .	16



# Capítulo 1

## Introdução

Neste capítulo apresentamos a motivação para o trabalho desenvolvido, os objetivos que pretendemos atingir e os principais resultados obtidos.

### 1.1 Motivação

Sempre que é necessário recolher informação sobre a usabilidade de uma aplicação, tipicamente realizamos testes com utilizadores. Estes necessitam de uma abordagem manual para serem realizados, recorrendo, por exemplo, a questionários standard para recolher a usabilidade percecionada pelos utilizadores. Estas abordagens tornam-se demoradas, podendo levar os utilizadores a ficarem saturados e a realizarem os testes com menos rigor, comprometendo os dados e levando a resultados errados. Além destes problemas existem outros como o enviesamento das respostas, que pode ser provocado por várias causas como por exemplo o cansaço, ou a dificuldade de alguns utilizadores em entenderem certas nuances dos questionários.

Para ultrapassar estes problemas, podemos adaptar uma abordagem que permite a inferência automática da usabilidade a partir das respostas fisiológicas dos utilizadores. No entanto, para conseguirmos fazer esta inferência, são necessários dados *groundtruth* confiáveis, para conseguirmos criar e validar os modelos de inferência.

### 1.2 Objectivos

O principal objetivo deste trabalho é produzir um *dataset* composto pelos sinais fisiológicos e pelos resultados dos questionários standard, preenchidos pelos utilizadores enquanto interagem com uma ou mais aplicações. Para atingirmos este objetivo principal, delineamos uma série de sub objetivos interligados que nos guiaram ao longo deste projeto, sendo esses:

- Escolha dos sites / aplicações para serem usadas durante as experiências;
- Classificação dos sites em termos de usabilidade e experiência de utilização;

- Recolha de ambos os tipos de dados (questionários e sinais fisiológicos);
- Validação dos dados recolhidos;
- Criação e disponibilização do *dataset*.

### 1.3 Abordagem Usada

Iniciámos o processo com uma seleção empírica de diversos sites, visando a escolha de um conteúdo amplo em termos de usabilidade e experiência de utilização que seria submetido a testes. Através desta abordagem, conseguimos abranger uma vasta gama de possibilidades, proporcionando uma base sólida para os passos subsequentes.

A segunda etapa consistiu em criar uma experiência envolvendo utilizadores, que realizaram uma série de tarefas nesses sites. Adicionalmente, os participantes responderam a questionários standard, que são tipicamente utilizados para avaliar a usabilidade, a experiência de utilização, entre outros aspetos.

Na terceira etapa, realizamos uma nova experiência com utilizadores, utilizando um novo subconjunto dos sites anteriores, que foi indentificado a partir dos resultados da experiência anterior. Nesta etapa, incorporámos a captação de sinais fisiológicos através do dispositivo Muse2. Este dispositivo permitiu-nos monitorizar em tempo real o movimento corporal, através do acelerómetro, a frequência cardíaca, via PPG, a respiração, por meio da combinação de PPG e acelerómetro, assim como a atividade cerebral, através dos sinais EEG. Estas informações foram complementadas com os questionários standard. A compilação dos dados recolhidos foi utilizada para a criação do nosso *dataset*.

Para garantir a qualidade e integridade do nosso *dataset*, submetemos o mesmo a um processo de validação. Nesse processo, realizamos uma comparação entre os dados dos questionários recolhidos na primeira e na segunda experiência, para verificarmos se existia coerência nos scores dos vários questionários de cada um dos sites.

### 1.4 Estrutura do documento

Este documento é constituído por mais 4 capítulos.

No Capítulo 2 são apresentados os fundamentos teóricos necessários para a realização do trabalho, como a explicação dos vários questionários usados, os sinais fisiológicos com que trabalhámos e as emoções. Apresentamos ainda os vários trabalhos relacionados que nos ajudaram a entender o melhor caminho a seguir. Através destes conseguimos entender quais os melhores questionários standard a usar e as melhores abordagens para as diversas partes da investigação.

No Capítulo 3 descrevemos o processo usado para a criação do *dataset*, descrevendo a metodologia usada, as duas experiências realizadas e os passos para a produção do *dataset*.

No Capítulo 4 apresentamos a validação do *dataset*, fazendo primeiro a descrição dos passos realizados, seguida dos resultados obtidos e da sua análise.

Finalmente, no Capítulo 5, apresentamos uma breve conclusão e o trabalho futuro.



## Capítulo 2

# Fundamentos Teóricos e Trabalho Relacionado

Para alcançar os objetivos previamente mencionados, é necessário que compreendamos o funcionamento e a utilização de cada questionário, bem como a interpretação de cada sinal fisiológico e as emoções. Além disso, é essencial adquirir conhecimento sobre os vários *datasets* já existentes e compreender a estrutura dos diferentes trabalhos relacionados que incluem dados fisiológicos e questionários standard, analisando a organização e a metodologia das experiências que conduziram.

## 2.1 Questionários Standard

Nesta secção descrevemos os vários questionários standard mais usados para recolher a usabilidade, experiência de utilização, utilidade e emoções percecionadas pelos utilizadores.

### 2.1.1 User Experience Questionnaire (UEQ)

O *User Experience Questionnaire* (UEQ) [33]<sup>1</sup> é um questionário utilizado para medir e avaliar a experiência de utilização em relação a um determinado produto, sistema ou serviço.

O UEQ consiste em seis dimensões-chave da experiência de utilização: atratividade, eficiência, estimulação, novidade, confiabilidade e perspicuidade. Estas dimensões resultam num total de 26 itens de avaliação (ver Figura 2.1). Cada dimensão é geralmente medida através de quatro itens relacionados, e os utilizadores são solicitados a avaliar cada item numa escala de sete pontos, variando de 1- "Discordo Fortemente" a 7- "Concordo Fortemente".

Estas avaliações fazem com que o UEQ seja capaz de fornecer informações sobre diferentes aspectos da experiência de utilização. A dimensão da atratividade, por exemplo, destina-se à percepção subjetiva da agradabilidade e beleza do produto. A dimensão da eficiência mede o quanto bem o produto permite que os utilizadores alcancem os seus objetivos de forma rápida e eficaz. A dimensão da perspicuidade avalia a clareza e facilidade de compreensão do produto, enquanto a dimensão da confiabilidade refere-se à confiança e consistência das interações.

---

<sup>1</sup><https://www.ueq-online.org/>

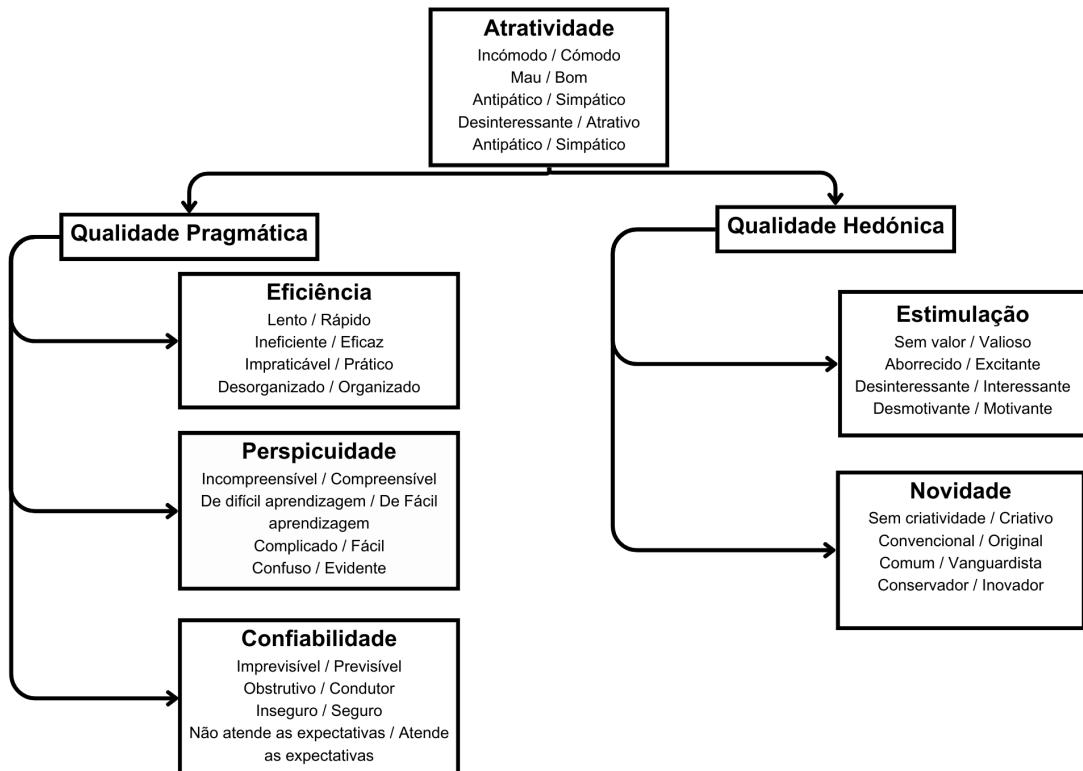


Figura 2.1: Estrutura de escala do UEQ baseada no estudo [32].

Uma das vantagens do UEQ é poder fornecer uma avaliação global da experiência de utilização, permitindo uma compreensão mais profunda e holística. Além disso, o UEQ permite que os investigadores e projetistas consigam identificar pontos fortes e áreas de melhoria ao longo do tempo. Isto acontece devido ao UEQ poder ser implementado em diferentes estágios do ciclo de desenvolvimento de um produto, desde a fase de prototipagem até à fase de lançamento.

O UEQ pode ser aplicado em conjunto com outras técnicas de pesquisa, como entrevistas, observação de utilizadores e testes de usabilidade, para obter uma visão mais completa da experiência de utilização.

Existe uma versão mais pequena do UEQ (UEQ-S) que contém apenas 8 itens de avaliação. Esta versão usa a mesma escala e cobre praticamente todo o espetro da qualidade do produto mesmo possuindo menos itens de avaliação.

### 2.1.2 System Usability Scale (SUS)

O *System Usability Scale* (SUS) [18]<sup>2</sup> é um questionário que é utilizado para avaliar a usabilidade de sistemas, produtos ou interfaces.

Este questionário curto e padronizado consiste em 10 afirmações relacionadas com a usabilidade (ver Figura 2.2). Os utilizadores devem responder numa escala de Likert de cinco pontos, variando de 1-”Discordo Totalmente” a 5-”Concordo Totalmente”. Os itens do SUS são cuidadosamente formulados para abordar diferentes aspectos da usabilidade, como a efetividade, a

<sup>2</sup><https://stuart-cunningham.github.io/sus/>

eficiência e a satisfação geral do utilizador. A pontuação final do SUS é calculada somando-se as

Questionário SUS
Acho que gostaria de utilizar este produto com frequência.
Considero o produto mais complexo do que necessário.
Achei o produto fácil de utilizar.
Acho que necessitaria de ajuda de um técnico para conseguir utilizar este produto.
Considero que as várias funcionalidades deste produto estavam bem integradas.
Achei que este produto tinha muitas inconsistências.
Suponho que a maioria das pessoas aprenderia a utilizar rapidamente este produto.
Considero o produto muito complicado de utilizar.
Senti-me muito confiante a utilizar este produto.
Tive que aprender muito antes de conseguir lidar com este produto.

Figura 2.2: Itens do questionário SUS traduzidos para Português pelo estudo [24].

respostas de cada item, com ajustes específicos para alguns itens invertidos. O resultado varia de 0 a 100, sendo que uma pontuação mais alta indica uma melhor usabilidade percebida.

O SUS é uma ferramenta versátil e flexível que pode ser aplicada em várias fases de desenvolvimento de produtos ou sistemas. Ele pode ser usado para avaliar sistemas já existentes, protótipos ou até mesmo para comparar diferentes designs e iterar em melhorias. Além disso, o SUS pode ser utilizado em diferentes contextos, como aplicações móveis, websites, software empresarial e dispositivos eletrónicos.

Uma das vantagens do SUS é conter uma escala fácil de administrar e poder ser utilizado tanto com uma amostra pequena tanto com uma grande. O questionário é curto e rápido de ser respondido pelos utilizadores, tornando conveniente tanto para os mesmos como para os investigadores. Além disso, o SUS fornece uma medida global de satisfação com o sistema, permitindo a comparação entre diferentes sistemas ou designs, facilitando a análise e interpretação dos resultados.

No entanto, é importante saber que o SUS avalia a usabilidade percebida pelo utilizador e não fornece informações detalhadas sobre problemas específicos de usabilidade. Como resultado, pode ser necessário combinar o SUS com outras técnicas de pesquisa, como testes de usabilidade, para identificar e resolver problemas específicos do sistema.

### 2.1.3 Technology Acceptance Model (TAM)

Como o próprio nome diz, o *Technology acceptance model* (TAM) [23] [36] é um modelo teórico utilizado para compreender e prever a aceitação e adoção de novas tecnologias pelos utilizadores. O TAM usa uma escala Likert de 1 a 7, sendo 1-”Discordo Totalmente” e 7-”Concordo Totalmente”.

Este modelo baseia-se em duas variáveis principais: Utilidade Percebida (PU) e Facilidade de

Uso Percebida (PEOU). A Utilidade Percebida diz respeito ao grau em que uma pessoa acredita que a utilização de um determinado sistema melhoraria o seu desempenho profissional. Por outro lado, a Facilidade de Uso Percebida está relacionada com o grau em que uma pessoa acredita que a utilização de um determinado sistema seria isenta de esforço.

O TAM, afirma que a intenção do utilizador em aceitar uma tecnológica advém destas duas variáveis. Além disso, o TAM refere que experiências anteriores do utilizador, influências sociais e características demográficas vão influenciar na relação entre estas duas variáveis.

Uma das principais contribuições deste modelo é oferecer *insights* sobre os fatores que afetam a aceitação e a adesão de uma tecnologia.

Apesar do que oferece, o TAM é um modelo teórico simplificado e a adesão a uma nova tecnologia pode ser influenciada por uma diversidade de fatores, como fatores pessoais, contextuais, culturais e organizacionais. Por este motivo, quando usado em contextos práticos, o TAM pode ser complementado com outros métodos para uma melhor compreensão da adoção da tecnologia.

#### 2.1.4 Questionnaire for User Interaction Satisfaction (QUIS)

O *Questionnaire for User Interaction Satisfaction* (QUIS) <sup>[15]</sup><sup>3</sup> é uma ferramenta de medição projetada para avaliar a satisfação subjetiva de um utilizador com a interface humano-computador (HCI).

O QUIS inclui um questionário demográfico, uma medida de satisfação geral com o sistema e uma medida de fatores específicos da interface, como visibilidade do ecrã, terminologia e informações do sistema, fatores de aprendizagem e capacidades do sistema.

O QUIS exige que os utilizadores respondam a perguntas através de uma escala Likert de 1 a 10. A escala contém palavras opostas em cada extremidade, permitindo que os utilizadores expressem o seu nível de concordância ou discordância com cada afirmação. Depois dos utilizadores responderem a todas as perguntas, estas então são analisadas e calculadas para quantificar qual foi a satisfação geral do utilizador com a interação com o sistema que foi avaliado.

O QUIS além de ser simples e rápido de utilizar, consegue fornecer informações detalhadas e específicas sobre os pontos fortes e fracos de um sistema tanto quando a amostra é grande tanto quando é pequena. Além disso quando é utilizada uma amostra correta ela fornece dados poderosos e também informações úteis sobre as opiniões e atitudes dos utilizadores em relação ao sistema.

No entanto, o QUIS é apenas uma das várias ferramentas existentes que avalia a satisfação do utilizador. Com o avanço da tecnologia e por limitar-se ao domínio da interação homem-computador o uso de outras técnicas juntamente do QUIS é importante.

<sup>3</sup><https://garyperlman.com/quest/quest.cgi?form=QUIS>

### 2.1.5 Single Ease Question (SEQ)

O *Single Ease Question* (SEQ) [30]<sup>4</sup> é um questionário simples e eficaz para avaliar a facilidade de utilização após a conclusão de uma tarefa ou cenário de interação com um sistema, produto ou serviço.

O SEQ contém apenas uma pergunta simples ("Em geral, quanto difícil ou fácil foi a tarefa de concluir?"), direta e fácil de entender, que requisita aos utilizadores para avaliar a facilidade com que conseguiram realizar uma determinada tarefa.

Uma escala Likert de 1 a 7, sendo 1-”Muito difícil” e 7-”Muito fácil”, é usada pelo utilizador para fornecer a sua resposta. Esta abordagem simplificada permite uma avaliação rápida e fácil sobre a facilidade de utilização sem sobrecarregar o utilizador devido à simplicidade do questionário.

O SEQ é frequentemente utilizado em pesquisas de experiência de utilização e testes de usabilidade porque fornece um indicador geral da facilidade de utilização percebida pelos utilizadores.

A simplicidade do SEQ é um dos seus principais benefícios. Por ser uma pergunta única, os utilizadores podem responder de forma rápida e sem complicações. Isto torna o SEQ especialmente útil em estudos que buscam obter feedback rápido e direto sobre a facilidade de utilização de um produto ou serviço.

No entanto, é importante lembrar que o SEQ possui algumas limitações que advêm de conter apenas uma questão. Este questionário não fornece informações específicas sobre os elementos que contribuem para a facilidade ou dificuldade percebida. Combinar o SEQ com outros métodos ou perguntas acaba então por ser aconselhável se o objetivo for obter informação detalhada. Esta combinação acaba por não trazer grandes problemas pois a sua integração é fácil.

### 2.1.6 After-Scenario Questionnaire (ASQ)

O *After-Scenario Questionnaire* (ASQ) [21] é um questionário utilizado para avaliar o grau de dificuldade com que um utilizador percepciona uma tarefa num teste de usabilidade, sendo administrado imediatamente após a conclusão da mesma.

O ASQ foi projetado para conseguir feedback completo dos utilizadores sobre uma experiência específica. Ele é aplicado imediatamente, enquanto as informações ainda estão frescas na mente do utilizador.

Este questionário é composto por três itens que avaliam a facilidade de uma tarefa, o tempo que a tarefa demorou a ser concluída e o nível de apoio recebido ao longo do processo. Estas perguntas são respondidas numa escala de likert com 7 pontos de classificação, sendo 1-”Discordo Totalmente” e 7-”Concordo Totalmente”.

As respostas obtidas no ASQ podem ser usadas para melhorar o design, a usabilidade e a funcionalidade do sistema, visando uma experiência de utilização mais agradável.

Uma das vantagens do ASQ é ser uma ferramenta de avaliação relativamente rápida e fácil de administrar devido à sua simplicidade. Por ser aplicado logo após a conclusão da tarefa, facilita

<sup>4</sup><https://measuringu.com/seq10/>

a obtenção de feedback mais preciso e detalhado, o que é útil para identificar problemas urgentes que requerem atenção imediata. Além disso, é eficaz na comparação de diferentes abordagens ou designs em tarefas específicas, ajudando na tomada de decisões que levem a um design mais eficaz.

No entanto, o ASQ possui certas limitações, a sua aplicação logo após a conclusão, apesar de trazer uma maior precisão pode existir influência do momento emocional do utilizador, podendo afetar os resultados recolhidos das respostas. Além disso, o ASQ não captura informação alguma sobre a experiência ao longo do tempo, já que este questionário apenas se concentra no momento pós-interação. Combinar o ASQ com outros métodos poderá ajudar a conseguir ter uma melhor validação da informação recolhida.

### 2.1.7 The Usefulness, Satisfaction, and Ease of Use Questionnaire (USE)

O *The Usefulness, Satisfaction, and Ease of Use Questionnaire (USE)* [14]<sup>5</sup> é um questionário utilizado para avaliar a usabilidade de sistemas, produtos ou serviços.

Existem trinta perguntas neste questionário, onde cada uma delas está relacionada a uma dimensão. A primeira dimensão, “Usefulness”, avalia a impressão do utilizador sobre o valor e a utilidade do sistema, levando em consideração os objetivos e necessidades do mesmo. A dimensão de “Satisfaction” mede a satisfação do utilizador com o sistema, levando em consideração as expectativas existentes e experiência de uso. A “Ease of Use”, analisa o quanto simples e natural é usar o sistema. Por fim, a última dimensão, “Ease of Learning”, avalia a facilidade com que os utilizadores conseguem aprender a utilizar o sistema ou produto.

Os utilizadores são solicitados a responder às perguntas do questionário classificando cada item numa escala de Likert de 1 a 7, sendo 1-”Discordo Totalmente” e 7-”Concordo Totalmente”. Estas respostas são analisadas e os resultados calculados para fornecer uma medida quantitativa da percepção de utilidade, satisfação, facilidade de utilização e facilidade de aprendizagem do sistema avaliado. Estas informações são essenciais para o aprimoramento do design e da usabilidade, visando a criação de sistemas mais eficazes e satisfatórios.

O USE tem como vantagem ter uma abordagem abrangente e estruturada, e devido a isso permite uma avaliação sistemática da usabilidade. Além disso, o mesmo é de fácil administração e pode ser aplicado em vários contextos, desde sistemas de software até produtos físicos ou serviços.

Combinar o USE com outras técnicas de avaliação fornece uma visão mais completa e poderá identificar problemas que não são capturados neste questionário.

### 2.1.8 NASA Task Load Index (NASA-TLX)

O *NASA Task Load Index (NASA-TLX)* [16]<sup>6</sup> é um questionário para medir e efetuar uma avaliação subjetiva da carga mental de trabalho pelos utilizadores durante a execução de uma tarefa. O NASA-TLX tem como objetivo fornecer uma medida quantitativa da carga de trabalho

<sup>5</sup><https://garyperlman.com/quest/quest.cgi?form=USE>

<sup>6</sup><https://www.keithv.com/software/nasatlx/nasatlx.html>

mental e física enfrentada pelos utilizadores. Ele apoia-se no princípio de que a carga de trabalho pode ser dividida em seis dimensões principais:

- **Exigência mental** - Relacionada com a quantidade de atividade mental e preceptiva que a tarefa requer;
- **Exigência Física** - Relacionada com a quantidade de atividade física que a tarefa requer;
- **Exigência Temporal** - Relacionada com o nível de pressão temporal sentido;
- **Desempenho** - Relacionado com o ponto que o utilizador se sente satisfeito com o seu nível de desempenho;
- **Esforço** - Indica o grau de esforço mental e físico que o utilizador tem de realizar para obter um nível de desempenho;
- **Frustação** - Relacionada com o ponto em que o utilizador se sente inseguro ou irritado durante a tarefa.

Através do NASA-TLX, os utilizadores são solicitados a avaliar cada uma destas dimensões atribuindo pontuações numa escala de Likert de 1 a 20, sendo 1-”Muito Baixa” e 20-”Muito alta”. Todas as dimensões seguem este padrão, exceto a dimensão de Desempenho, que possui uma escala que varia de 1-”Perfeito” a 20-”Fracasso”. Estas pontuações podem assumir outra forma, sendo esta através de uma distribuição de 100 pontos. As pontuações são então ponderadas e combinadas para fornecer uma pontuação global da carga de trabalho percebida.

Uma das vantagens do NASA-TLX é a sua capacidade de fornecer uma medida extensa e objetiva da carga de trabalho percebida pelos utilizadores. Ele permite que se consiga identificar como uma tarefa ou determinado sistema afeta o utilizador tanto em termos de carga mental, como física e temporal e quais são os fatores que podem levar a um aumento da carga ou insatisfação.

No entanto, é importante ressaltar que o NASA-TLX é uma ferramenta subjetiva, baseada nas percepções dos utilizadores, e não em medidas objetivas. Além disso, como todas as ferramentas de avaliação já referidas anteriormente, o NASA-TLX possui certas limitações, como a dependência da capacidade dos utilizadores em avaliar e pontuar corretamente cada dimensão.

### 2.1.9 Subjective Mental Effort Questionnaire (SMEQ)

O *Subjective Mental Effort Questionnaire* (SMEQ) [30] <sup>7</sup> é um questionário de um item que avalia o esforço mental percebido pelos utilizadores durante a realização de uma tarefa ou atividade.

O SMEQ é projetado para medir o esforço mental e a carga cognitiva necessários para concluir uma tarefa específica. Ele baseia-se na percepção subjetiva dos utilizadores, permitindo que expressem o nível de esforço mental que experimentaram durante a realização da tarefa. Os participantes são solicitados a avaliar atribuindo pontuações numa escala de 9 divisões que vai de 1-”Sem Qualquer Esforço” a 9-“Esforço Extremo”, sendo estas inseridas numa linha que vai de 0 a 150.

<sup>7</sup>[https://plos.figshare.com/articles/figure/Subjective\\_mental\\_effort\\_questionnaire/4881470](https://plos.figshare.com/articles/figure/Subjective_mental_effort_questionnaire/4881470)

Ao analisar a resposta do SMEQ, é possível obter uma medida subjetiva do esforço mental experimentado pelos utilizadores durante a realização da tarefa. Estas informações podem ser úteis para identificar tarefas ou atividades que exigem um esforço mental excessivo. O SMEQ é particularmente útil quando se pretende medir e regular a carga cognitiva necessária para uma tarefa. Tal como o SEQ, o SMEQ é rápido e fácil de administrar. O número praticamente contínuo de opções de resposta na escala é outra vantagem, pois as respostas dos utilizadores são mais precisas devido a existir mais opções de escolha.

É importante lembrar que o SMEQ varia de acordo com os fatores individuais porque é baseado na percepção do utilizador. Como resultado, para obter uma avaliação mais abrangente e precisa, o SMEQ deve ser combinado com outras métricas objetivas de desempenho e feedback qualitativo. Além disso, a escala do questionário pode ser confusa tanto para os participantes como para os investigadores, uma vez que não é uma escala uniforme e fechada como a que estamos normalmente habituados a ver.

### 2.1.10 Post Study System Usability Questionnaire (PSSUQ)

O *Post Study System Usability Questionnaire* (PSSUQ) [29]<sup>8</sup> é um questionário utilizado para avaliar a satisfação do utilizador com a usabilidade do sistema, produto ou serviço.

Ele contém 16 itens sobre aspectos da usabilidade, tais como a utilidade do sistema, qualidade da informação e a qualidade da interface.

Os utilizadores respondem às várias questões do PSSUQ atribuindo pontuações numa escala de likert de 1 a 7, que varia de 1-”Concordo Totalmente” a 7-”Discordo Totalmente”. As respostas são posteriormente analisadas e o resultado geral é calculado através da média das respostas.

O objetivo do PSSUQ é obter feedback dos utilizadores sobre a sua experiência de utilização do sistema avaliado.

A maior vantagem do PSSUQ é a sua capacidade de fornecer uma visão geral da usabilidade percebida pelos utilizadores. Além disso, por ser bastante abrangente consegue fazer uma avaliação completa da experiência de utilização.

Podemos encontrar diferentes versões do PSSUQ disponíveis online e a maneira de diferenciá-las é observar o número de perguntas. Existem 18 perguntas na primeira versão, 19 questões na segunda versão e 16 questões na terceira versão.

### 2.1.11 Usability Metric for User Experience(UMUX)

A *Usability Metric for User Experience* (UMUX) [22][12] é uma métrica utilizada para a avaliação subjetiva da usabilidade percebida de um sistema, produto ou serviço.

O UMUX é projetado para capturar a percepção global do utilizador em relação à usabilidade e satisfação com o sistema. Ele baseia-se em perguntas simples e diretas, nas quais os utilizadores atribuem pontuações numa escala de Likert de 7 pontos para indicar o nível de concordância com as afirmações apresentadas.

<sup>8</sup><https://cdn.uiuxtrend.com/wp-content/uploads/PSSUQ-Questionnaire-PDF-Template.pdf>

Destina-se a ser semelhante ao SUS, mas as perguntas são mais curtas e direcionadas. Além disso o UMUX é direcionado para a definição de usabilidade da ISO 924-11 (eficácia, eficiência e satisfação). As respostas são depois combinadas e calculadas para obter uma pontuação geral da usabilidade geral de um sistema.

No entanto, é importante salientar que o UMUX é uma métrica subjectiva, que se baseia nas opiniões dos utilizadores. Portanto, recomenda-se a combinação do UMUX com outros métodos de investigação, como testes de usabilidade, observações e análise qualitativa, para obter uma compreensão mais completa e exacta da experiência de utilização.

Cada item individual do UMUX tem um intervalo de 0 a 6 após a recodificação, dando um total que pode chegar no máximo a 24, pois são 4 perguntas. Caso queiramos um intervalo de 0 a 100 similar ao do SUS, a pontuação UMUX de um participante é a soma das quatro perguntas divididas por 24 e depois multiplicadas por 100. As pontuações de cada participante são depois usadas para encontrar uma pontuação média do UMUX.

### 2.1.12 Usability Metric for User Experience LITE (UMUX-LITE)

O *Usability Metric for User Experience LITE* (UMUX-LITE) [22] é um questionário simplificado do UMUX e como tal serve para o mesmo propósito. O UMUX-LITE é projetado para ser rápido e fácil de aplicar. Ele consiste em apenas duas perguntas que abordam a usabilidade percebida do sistema e a intenção do utilizador em usá-lo novamente. As duas perguntas do UMUX-LITE são:

- Os recursos [deste sistema] atendem aos meus requisitos;
- [Este sistema] é fácil de usar.

Os utilizadores respondem a ambas as perguntas atribuindo uma pontuação numa escala de Likert que varia de 1 a 7, onde 1 representa "Discordo Totalmente" e 7 "Concordo Totalmente". Posteriormente, as respostas são analisadas e calculadas para fornecer uma avaliação quantitativa da experiência de utilização em relação à usabilidade do sistema.

Uma das vantagens do UMUX-LITE é a sua simplicidade e rapidez de aplicação. Como envolve apenas duas perguntas, é possível obter um feedback rápido dos utilizadores sem exigir muito tempo ou esforço. Apesar disso, é importante ressaltar que o UMUX-LITE é uma métrica simplificada e não fornece dados detalhados de diferentes aspectos da usabilidade. Ele concentra-se principalmente na satisfação geral do utilizador, deixando de lado outros aspectos importantes da usabilidade, como a eficiência, a eficácia e a facilidade de aprendizagem. Se quisermos mais informação a recomendação passa por combinar este questionário com outras técnicas de avaliação existentes.

### 2.1.13 Self-Assessment Manikin (SAM)

O *Self-Assessment Manikin* (SAM) [4] <sup>9</sup>, ao contrário dos questionários anteriores, é utilizado para recolher informações sobre o estado emocional dos utilizadores. Este utiliza um conjunto de

<sup>9</sup>[https://www.researchgate.net/figure/The-Self-Assessment-Manikin-SAM-used-in-the-questionnaire\\_fig1342704678](https://www.researchgate.net/figure/The-Self-Assessment-Manikin-SAM-used-in-the-questionnaire_fig1342704678)

figuras / desenhos que representam diferentes estados emocionais. Geralmente, são três elementos emocionais: Valência (positiva ou negativa), Excitação (alta ou baixa) e Dominância (controlo ou falta de controlo) (ver Figura 2.3). Os utilizadores são solicitados a classificar cada categoria fa-

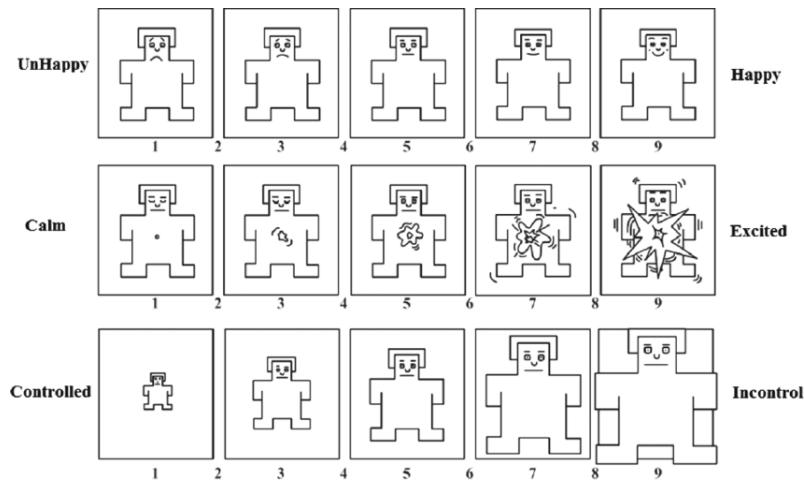


Figura 2.3: O Self-Assessment Manikin (SAM).

zendo um círculo na imagem que melhor representa a maneira como se sentem naquele momento.

Esta abordagem visual e simplificada que o SAM possui, permite aos utilizadores expressar as suas emoções de maneira rápida e intuitiva, evitando que tenha de usar palavras complexas para expressar os seus sentimentos. Além disso por ter um design não-verbal, pode ser lido por qualquer pessoa, independentemente da idade, idioma ou outros fatores educacionais.

O SAM é bastante utilizado em pesquisas de usabilidade e design de interfaces para capturar as respostas emocionais dos utilizadores. Ele fornece informações valiosas sobre como os utilizadores percebem e reagem aos estímulos visuais.

Tal como os questionários anteriores, o SAM é uma ferramenta subjetiva, que se baseia nas percepções e interpretações individuais de cada utilizador e por isso os resultados podem variar consoante diversos fatores como cultura, contexto e experiências pessoais. Por estes motivos a interpretação dos resultados tem de ser feita de maneira cautelosa.

#### 2.1.14 Resumo dos Vários Questionários

Na próxima secção, criámos uma tabela que resume todos os questionários utilizados na nossa pesquisa. Esta tabela foi dividida em quatro pontos-chave: Pontos de Medição, Escala, Vantagens e Desvantagens. Esta abordagem permitirá uma análise das principais características de cada questionário, facilitando a compreensão.

Tabela 2.1: Resumo dos questionários standard com os pontos de medição, escalas, vantagens e desvantagens.

Citação	Questionários	Pontos de Medição	Escala	Vantagens	Desvantagens
[33]	User Experience Questionnaire (UEQ)	Atratividade, eficiência, estimulação, novidade, confiabilidade e perspicuidade	26 itens cada um com escala de likert de 1 a 7.	Dados bastante enriquecedores, engloba todos os pontos da experiência de utilização e é dividido em várias secções.	Demorado, por conter uma grande quantidade de itens pode tornar-se demasiado exaustivo.
[18]	System Usability Scale (SUS)	Efetividade, eficiência e satisfação.	10 perguntas cada uma com escala de likert de 1 a 5.	Escala muito fácil de administrar, pode ser usado em amostras pequenas e diferencia entre utilizável e inutilizável os sistemas.	Nenhuma informação precisa sobre os pontos fracos e a comparação entre dois sistemas é um tanto desfocada.
[23]	Technology Acceptance Model (TAM)	Utilidade Percebida e facilidade de uso percebida.	12 itens cada um com escala de likert de 1 a 7.	Mais fácil de usar e menos dispendioso de aplicar, tem qualidade, é estatisticamente fiável e é um modelo robusto e poderoso para prever a aceitação.	Ignora algumas construções teóricas importantes e não reflete a variedade de ambientes de tarefas e restrições dos utilizadores.
[15]	Questionnaire for User Interaction Satisfaction (QUIS)	Satisfação geral e fatores específicos da interface.	21 itens cada um com escala bipolar de 1 a 10.	Eficaz mesmo quando o tamanho da amostra é pequeno, fornece dados poderosos, é bastante simples, aplicação rápida e fornece feedback útil das opiniões e atitudes do utilizador sobre o sistema.	Está limitado ao domínio HCI.
[30]	Single Ease Question (SEQ)	Dificuldade ou facilidade de fazer uma determinada ação numa aplicação.	Uma única pergunta com uma escala de likert de 1 a 7.	O questionário mais simples, funciona igualmente bem ou melhor que questionários mais complicados, fácil interpretação e dados precisos.	Não dá abertura por só conter 1 pergunta e requer validação.
[21]	After-Scenario Questionnaire (ASQ)	O grau de dificuldade com que um utilizador percebe uma tarefa (Similar ao SEQ)	3 itens, cada um com escala de likert de 1 a 7.	Curto, simples e com uma boa distribuição.	Falta de contexto, não abrange toda a experiência de utilização.
[14]	The Usefulness, Satisfaction, and Ease of Use Questionnaire (USE)	Utilidade, facilidade de uso, facilidade de aprendizagem e satisfação.	30 itens cada um com escala de likert de 1 a 7.	Um questionário com uma abordagem abrangente e estruturada. Os itens contêm descrições inequívocas e relevantes	Poucos trabalhos reportaram a fiabilidade ou validade da sua utilização. Contém uma grande quantidade de itens, o que o pode tornar-se demasiado exaustivo.

Tabela 2.1: Resumo dos questionários standard com os pontos de medição, escalas, vantagens e desvantagens.

Citação	Questionários	Pontos de Medição	Escala	Vantagens	Desvantagens
[16]	NASA Task Load Index (TLX)	Mede a carga de trabalho mental percebida e a eficácia de uma tarefa.	6 itens cada um com uma escala de 1 a 20.	Estimativa rápida e simples da carga de trabalho mental do utilizador, pode ser usado em vários domínios, alivia a carga de trabalho de quem analisa.	Demorado e difícil e os utilizadores podem ficar inclinados a fazer correlação do desempenho da tarefa e as classificações de carga de trabalho.
[30]	Subjective Mental Effort Questionnaire (SMEQ)	Esforço mental que o utilizador sente que está a ser usado numa tarefa.	9 itens que são inseridos numa linha de base de 0 a 150.	Dados mais enriquecedores, útil quando queremos regular a carga cognitiva necessária para uma certa tarefa.	Escala bastante complexa tanto para os participantes como para os próprios investigadores.
[29]	Post Study System Usability Questionnaire (PSSUQ)	Avaliar a satisfação do utilizador com a usabilidade do sistema.	16 itens cada um com uma escala de likert de 1 a 7.	Flexível, sensível e é eficaz em amostras de tamanho reduzido, pode ser dividido em vários scores.	Questionário pode se tornar bastante complexo e difícil de entender.
[22]	Usability Metric for User Experience (UMUX)	Avaliação subjetiva da usabilidade percebida num determinado aplicativo.	4 itens cada um com uma escala de 1 a 7.	Simples de administrar e pode ser utilizado para medir todas as fases do ciclo de vida de um produto.	Combina questões positivas e negativas que cria um teste de dois fatores, quando a intenção é um teste de apenas um.
[22]	Usability Metric for User Experience LITE (UMUX -LITE)	Avaliação subjetiva da usabilidade percebida num determinado aplicativo.	2 itens (só os itens positivos do UMX) cada um com uma escala de 1 a 7.	Fixa os principais problemas do UMX, utilizando apenas os itens positivos, diminuindo a confusão que possa existir.	A abordagem simplista pode levar à falta de captura de nuances detalhadas da experiência de utilização e pode ser menos eficaz na identificação de problemas específicos de usabilidade ou design.
[4]	Self-Assessment Manikin (SAM)	Expressa 3 elementos emocionais (Prazer, excitação e domínio).	3 itens cada um com uma escala de 1 a 9.	Design não verbal o que torna o questionário de fácil acesso.	Por vezes difícil de entender em determinados contextos.

## 2.2 Sinais Fisiológicos

Nesta secção descrevemos resumidamente os vários sinais fisiológicos que necessitamos de entender para fazer o uso correto do dispositivo Muse2. O Muse 2 é um dispositivo de *neurofeedback* que destaca-se pela sua capacidade de monitorar sinais fisiológicos fundamentais, oferecendo uma visão aprofundada do estado mental e emocional dos utilizadores. Este dispositivo possui sensores EEG extremamente precisos que permitem a análise das respostas cerebrais registando as ondas cerebrais em tempo real. A frequência cardíaca e os padrões de respiração, por exemplo, são monitorados pelo Muse 2 além da atividade cerebral. Estes dados são transmitidos para uma aplicação móvel ou para uma plataforma de análise.

O Muse 2 é um dispositivo pequeno e leve que pode ser facilmente colocado na cabeça dos participantes do estudo. Isto o torna adequado para uso em laboratórios, ambientes clínicos ou até mesmo em investigações de campo.

A Figura 2.4 representa o dispositivo Muse 2 e os vários sensores utilizados para recolher esses sinais.

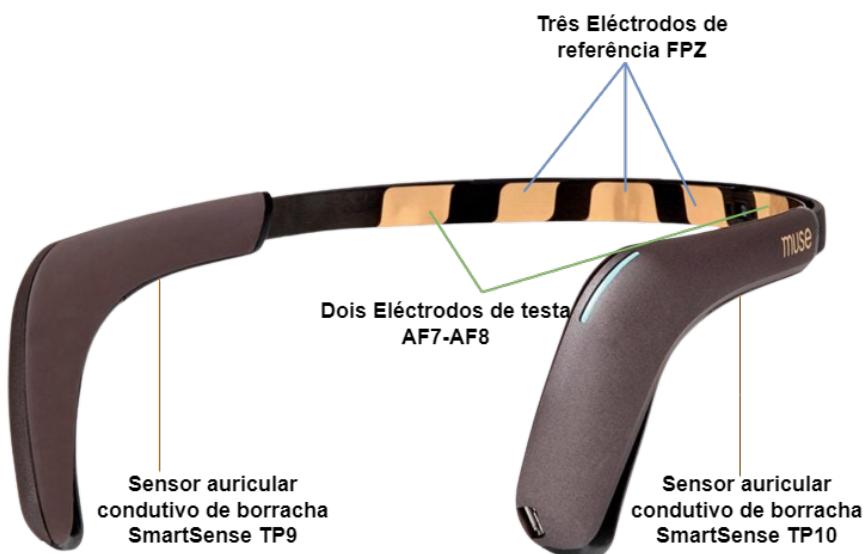


Figura 2.4: Dispositivo Muse 2 e os seus sensores.

### 2.2.1 Eletroencefalografia

A Eletroencefalografia (EEG) [3] é um método de monitoramento eletrofisiológico que é utilizado para registar a atividade elétrica do cérebro. É uma ferramenta importante pois permite a análise e o registo das ondas cerebrais.

Trata-se de um método normalmente não-invasivo, com elétrodos colocados no couro cabeludo, muito embora haja alguns métodos utilizados em aplicações específicas que são invasivos. Estes elétrodos são conectados a um equipamento especializado para registar as atividades elétricas cerebrais. Estes captam os sinais elétricos gerados pelos neurónios no cérebro e mede as flutuações de tensão resultante da corrente iônica.

Através da EEG, é possível analisar diferentes frequências de ondas cerebrais, como as ondas alpha (ocorrem em estados mentais relaxados e no sono), beta (ocorrem no estado normal de vigília), delta (ocorrem em momentos de sono profundo) e tetha (ocorrem em meditação e relaxamento profundo). Mesmo a EEG sendo uma técnica extremamente útil, é importante reconhecer alguns dos problemas que possui. A atividade elétrica registada pela EEG mostra indiretamente o

que acontece no cérebro, e a interpretação dos resultados requer uma compreensão especializada. Além disso, a sensibilidade da EEG a artefactos como movimentos musculares e interferências elétricas externas pode afetar a qualidade dos dados que são recolhidos.

### 2.2.2 Fotopletismografia

A Fotopletismografia (PPG) [27] é uma técnica de monitorização não invasiva utilizada para medir as mudanças de volume sanguíneo em determinada região do corpo. Esta técnica é usada para monitorar a frequência cardíaca e avaliar a função cardiovascular.

A PPG é baseada na detecção da luz transmitida ou refletida por tecidos vasculares, como a pele. Um feixe de luz que se dirige para estes tecidos é recolhido por um fotodíodo (um semicondutor que converte a luz em corrente elétrica). O sensor deteta alterações na absorção da luz à medida que o volume sanguíneo nas artérias se modifica devido ao fluxo sanguíneo.

Estas variações na absorção da luz são convertidas num sinal gráfico conhecido como fotopletismograma. A PPG mostra as variações do volume sanguíneo em relação ao tempo, permitindo a identificação dos picos correspondentes aos batimentos cardíacos. Com base nessas informações, é possível calcular a frequência cardíaca e analisar outros parâmetros relacionados à função cardiovascular, como a variabilidade da frequência cardíaca.

Mesmo que a PPG seja uma técnica de monitorização que é acessível e não invasiva, é importante ter em consideração as suas desvantagens. Fatores como movimentos do corpo, temperatura e níveis de iluminação podem interferir na qualidade dos dados obtidos pela PPG.

### 2.2.3 Acelerómetro

O acelerómetro (ACC) [11] é um dispositivo que mede a vibração ou aceleração do movimento de uma estrutura. O acelerómetro geralmente consiste num sensor que deteta as forças de aceleração aplicadas a ele.

Quando um objeto, como um smartphone, é movido, o acelerómetro dentro dele deteta as mudanças na velocidade e direção desse movimento. Os acelerómetros de três eixos são os mais populares. Eles são construídos como um sistema de três acelerómetros independentes, cada um dos quais mede a aceleração numa direção diferente (X, Y e Z).

### 2.2.4 Giroscópio

Um giroscópio (GYR) [11] é um dispositivo capaz de orientar-se usando a gravidade da Terra. Os dispositivos conhecidos como sensores giroscópicos podem registar a velocidade angular, ou a mudança no ângulo de rotação por unidade de tempo. Em sua maioria, a velocidade angular é expressa em graus por segundo.

Quando utilizado como sensor, é um dispositivo projetado que serve para medir/detetar as mudanças na orientação angular de um objeto. Eles funcionam com princípios parecidos aos dos tradicionais, mas a sua aplicação volta-se para a medição e o registo de movimento ou mudança de orientação.

## 2.3 Emoções

As emoções indicam um estado afetivo subjetivo que ocorre em resposta a algo que experimentamos, algo que surge espontaneamente. É um estado que muitas vezes é muito breve, durando segundos ou minutos, e o seu acontecimento acaba por ter uma causa ou razão específica.

Os sentimentos, por outro lado, podem ser vistos como uma representação subjetiva das emoções,

mas tendo a diferença fundamental que os sentimentos são experimentados conscientemente, enquanto as emoções podem ser manifestadas de forma tanto consciente como de forma subconsciente.

### 2.3.1 Modelos de Emoções

Para conseguirmos estimar emoções é necessário que os investigadores consigam descrevê-las. Por este motivo ao longos dos anos foram criados modelos que apesar de diferentes tentam abordar da melhor maneira o assunto. Uma das abordagens mais conhecidas são as seis emoções básicas que são propostas por Ekman [9], sendo estas raiva, repugnância, medo, felicidade, tristeza e surpresa. Esta abordagem deveu-se ao estudo da relação entre expressões faciais e emoções. Foram escolhidas estas emoções por serem consideradas como emoções base pois a manifestação externa acaba por não depender de nenhum fator externo como por exemplo cultura ou experiência pessoal.

Outra abordagem que veio depois da referida anteriormente é a *Plutchik's emotion wheel* [28] (ver Figura 2.5). Este modelo contém mais emoções que o anterior, 8 emoções (raiva, medo, tris-

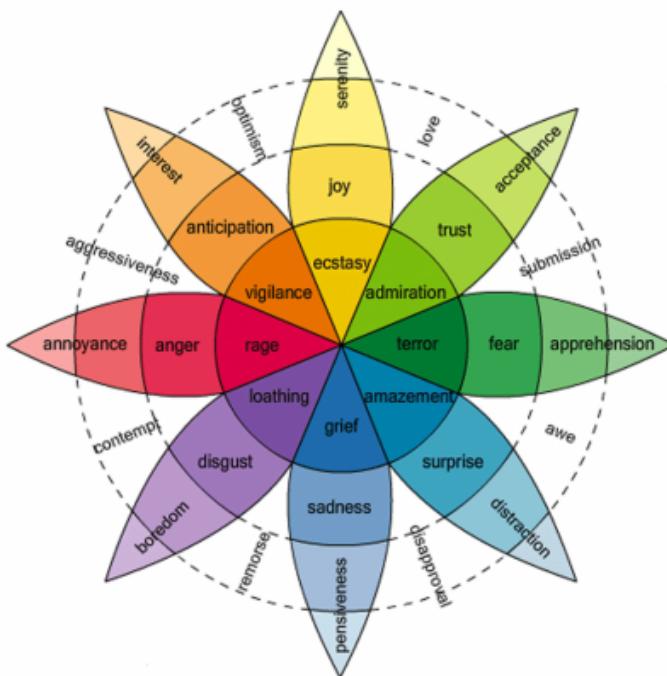


Figura 2.5: Emoções principais estabelecidas no *Plutchik's emotion wheel*.

teza, repugnância, surpresa, curiosidade, aceitação e alegria) e a partir destas emoções pode haver a criação de outras. Esta abordagem acaba por ser transmitida em forma de flor com oito pétalas de diferentes cores (podendo também existir o modelo em cone), e cada pétala acaba por estar situada num sítio específico para mostrar e identificar de forma clara as diferentes emoções e as relações que elas vão estabelecer entre si.

Mais abordagens foram feitas com mais emoções, mas podem trazer alguns problemas, sendo um deles que muitas vezes os dados não são os suficientes para emoções que são menos comuns, o que pode trazer um modelo bastante desequilibrado. Outro dos modelos mais utilizados é a perspetiva dimensional. Quando falamos de uma perspetiva dimensional, as emoções são mapeadas num espaço de 3 dimensões, onde temos valência, excitação e a dominância como dimensões. A

valência refere-se a aversão ou atratividade de estímulos que as pessoas acabam por experimentar e varia de negativo (desagradável) a positivo (agradável) ou de aversivo a atraente.

A excitação representa a ativação do corpo humano, indo de sonolento (calmo) a excitado (agitado). A dominância corresponde à força da emoção. Como a dominância é a dimensão menos compreendida, alguns investigadores usam espaços apenas com as dimensões de excitação e valência. O modelo mais conhecido e mais usado nesse espaço 2D é o *Circumplex Model of Affect* [10] (ver Figura 2.6) sendo que o eixo vertical representa a excitação e o eixo horizontal representa a valência.

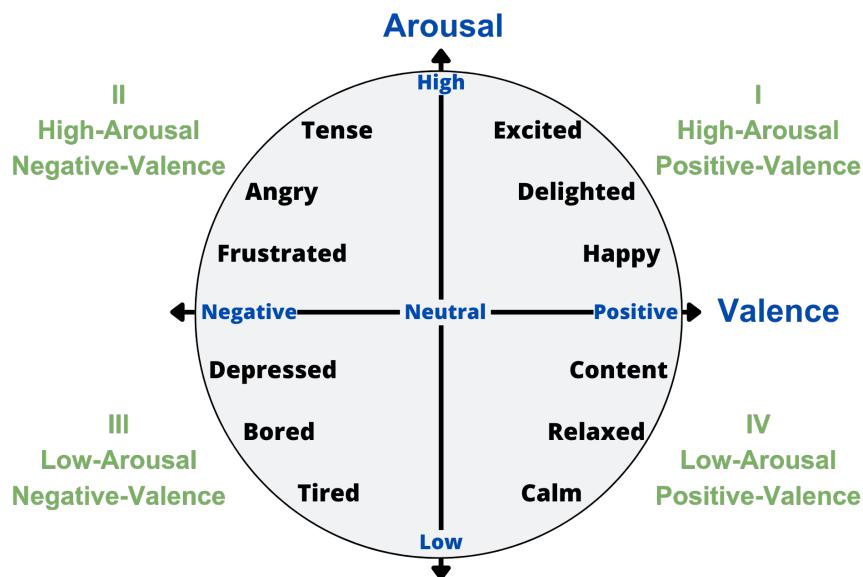


Figura 2.6: Emoções principais estabelecidas no *Circumplex Model of Affect* do estudo [35].

O espaço valência-excitação pode ser subdividido em quatro quadrantes, nomeadamente, baixa excitação / baixa valência (LALV), baixa excitação / alta valência (LAHV), alta excitação / baixa valência (HALV), e alta excitação / alta valência (HAHV).

## 2.4 Datasets com Sinais fisiológicos

Nesta secção descrevemos de maneira concisa o procedimento e os resultados obtidos pelos principais *datasets* que contêm sinais fisiológicos e dados de questionários standard.

### 2.4.1 DEAP Dataset

O *dataset* DEAP [19] teve como objetivo explorar a possibilidade de classificar as dimensões das emoções induzidas pela exibição de vídeos para vários utilizadores. Inicialmente, os autores realizaram uma seleção de estímulos, que consistia em clips de música ou clips de filmes / documentários. Foram escolhidos 120 estímulos iniciais, sendo metade deles selecionados de forma semi-automática e a outra metade manualmente. Posteriormente, foi determinado um minuto de cada clip para cada estímulo. Por fim, conduziu-se uma experiência de avaliação subjetiva, na qual foram selecionados 40 estímulos finais para avaliação.

Os dados foram obtidos de um grupo de 32 participantes. Eles foram convidados a assistir a uma série de clips de estímulo enquanto os seus sinais fisiológicos eram monitorizados. Os participantes tiveram de sentar-se confortavelmente numa cadeira e foram orientados a evitar muitos movimentos durante a experiência. Foi usado o sistema *Biosemi ActiveTwo* para registar os sinais fisiológicos, incluindo eletroencefalografia (EEG), eletrocardiografia (ECG), eletromiografia (EMG), respostas galvânicas da pele (GSR) e frequência cardíaca (HR). Além destes sinais, os participantes tiveram de preencher o questionário SAM de modo a avaliar, a valência, excitação, dominância e o gostar ou não gostar.

Nesta pesquisa, os estímulos foram escolhidos de forma a provocar emoções nos quatro quadrantes do espaço valência-excitabilidade (alta valência-alta excitação (HVHA), alta valência-baixa excitação (HVLA), baixa valência-baixa excitação (LVLA), baixa valência-alta excitação (LVHA)). O início da experiência compreendeu uma fase inicial de dois minutos destinada ao registo de dados de referência, enquanto os participantes foram orientados a permanecerem num estado de descontração. Durante esta fase, uma imagem simples foi apresentada no ecrã para que os participantes concentrassem a sua atenção nela, enquanto recebiam instruções para permanecerem calmos e tranquilos. Em seguida, os 40 vídeos foram apresentados em 40 tentativas.

Inicialmente, os autores tinham o objetivo de compreender como a estimulação afetiva afetava as avaliações subjetivas obtidas dos participantes no estudo. Os resultados demonstraram que os estímulos de baixa excitabilidade induziram avaliações de valência diferentes daqueles de alta excitabilidade. Da mesma forma, os estímulos de baixa valência induziram avaliações de excitabilidade diferentes dos estímulos de alta valência. A análise das avaliações dos participantes revelou que as escalas de "gostar", "valência", "excitação", "dominância" e "familiaridade" estavam intercorrelacionadas de forma complexa. Foi observada uma forte correlação positiva entre "gostar" e "valência" e entre "dominância" e "valência". Além disso, "excitação" e "dominância" apresentaram uma correlação positiva moderada, assim como "excitação" e "gostar". "Familiaridade" também revelou correlações positivas moderadas com "gostar" e "valência".

A ordem que foram apresentadas as músicas teve um pequeno efeito nas avaliações de gostar e dominância e não apresentou relação significativa com as outras. Para fazer a comparação das avaliações subjetivas e dos dados fisiológicos, foram calculados os coeficientes de correlação de Spearman entre as mudanças de potência nas diferentes bandas de frequências e as classificações dos participantes. Além disto, foram calculados os valores de p para os testes de correlação positiva e negativa. Esta análise foi realizada separadamente para cada participante, e os valores de p resultantes foram combinados usando o método de Fisher. Os resultados da análise revelaram várias correlações significativas entre as mudanças de potência nas diferentes bandas de frequência do EEG e as classificações subjetivas dos participantes.

- Para a excitação, foram encontradas correlações negativas nas bandas de frequência theta, alpha e gamma, indicando que um aumento na excitação estava associada a uma diminuição na potência nessas bandas.
- Para a valência, foram encontradas correlações positivas nas bandas de frequência theta e alpha, indicando que um aumento na valência estava associado a um aumento na potência nessas bandas. Estes efeitos foram localizados principalmente nas regiões occipitais do cérebro.
- Para o gostar, encontraram correlações positivas em diversas bandas de frequência, sendo estas o theta e o alpha, mostrando que uma maior afeição associava um aumento na potência destas bandas.

Os resultados mostram que as mudanças que ocorreram nos dados EEG estavam relacionadas com as respostas emocionais dos participantes aos estímulos musicais. No entanto, as correlações

observadas foram maior parte das vezes de magnitude moderada. Além disso existiu uma alta variabilidade entre os vários participantes. Logo, os resultados demonstraram que uma abordagem onde cada participante tivesse uma classificação específica seria mais apropriada em vez de uma classificação geral. Isto aconteceu devido à variabilidade individual nas ativações cerebrais. Além de todas as técnicas que foram mencionadas, os autores ainda fizeram uso de técnicas de visualização, como mapas de calor e gráficos de dispersão. Os autores tentaram melhorar o desempenho geral combinando os resultados de cada um dos dados (como sinais EEG e sinais periféricos). No entanto, o resultado da fusão não proporcionou uma melhoria significativa em relação ao uso individual de cada um deles.

#### 2.4.2 AMIGOS *Dataset*

O *dataset* AMIGOS [26], produziu o seu *dataset* em função dos efeitos, traços de personalidade e humor através de sinais fisiológicos. Diferente dos outros *datasets*, este provoca efeitos usando vídeos curtos e longos em diferentes configurações, sendo a primeira com participantes individualmente e outra com grupos de participantes.

O *dataset* permite o estudo multimodal das respostas afetivas dos indivíduos em relação à sua personalidade e humor, e também à análise de como essas respostas são afetadas pelas configurações feitas neste estudo. As configurações são compostas pela atividade ser feita em grupo ou individual, juntamente com a duração do vídeo, seja curta ou longa.

Os vídeos curtos foram selecionados com base nas avaliações de valência e excitação de 72 voluntários, que os classificaram em diferentes quadrantes de valência-excitação. Escolheram três vídeos de cada quadrante e adicionaram mais quatro vídeos de outros estudos para abranger todos os quadrantes. Quanto aos vídeos longos, selecionaram oito trechos de filmes bem avaliados no IMDb e, em seguida, os investigadores classificaram, escolhendo quatro deles.

Recolheram dados em duas configurações: na primeira, 40 participantes assistiram a 16 vídeos emocionais curtos sozinhos. Na segunda configuração, o mesmo número de pessoas assistiu a 4 vídeos longos, alguns sozinhos e outros em grupos. Durante essas configurações, gravaram os sinais dos participantes (Electroencefalograma (EEG), Eletrocardiograma (ECG) e Resposta Galvânica da Pele (GSR), também conhecida como atividade eletrodérmica (EDA)), bem como os vídeos que eles assistiram.

Os participantes avaliaram suas emoções em cada vídeo por meio de autorrelato usando a Escala de Afeto Positivo e Negativo (PANAS) e preencheram também dados como a valência, a excitação, a dominância e perguntas sobre familiaridade e preferência, além de uma seleção de emoções básicas. Também obtiveram anotações externas de vários avaliadores sobre os níveis de valência e excitação de cada participante em cada vídeo. Estas cinco dimensões foram classificadas numa escala de 1 (baixo) a 9 (alto), enquanto as emoções básicas, como felicidade, medo e surpresa, foram registadas em valores binários.

Para analisar os dados, os autores inicialmente usaram as anotações externas de valência e excitação dos participantes, mapeando-as numa escala de 0 a 1. A conclusão da análise mostrou uma alta concordância entre os avaliadores, o que indica que as anotações externas eram confiáveis. Os investigadores usaram o coeficiente de correlação de Spearman para comparar essas anotações externas com as autoavaliações dos participantes. Eles encontraram correlações significativas, indicando que as anotações externas podem prever com precisão o estado afetivo dos participantes, principalmente em relação à valência.

Além disso, foram apresentadas distribuições dos dados que mostravam a diferença na valência e excitação percebida pelos participantes em relação aos vídeos curtos. A familiaridade dos participantes com os vídeos também foi calculada usando o coeficiente de correlação de Spearman, e concluiu-se que isso não teve um grande efeito nos estados afetivos relatados pelos participantes.

Os autores também compararam os valores entre os dois grupos de vídeos longos. Os resultados revelaram que o contexto social desempenha um papel importante na forma como as pessoas expressam suas emoções, particularmente em relação à valência e excitação. Essas diferenças são mais pronunciadas em clips com emoções mais intensas e variam dependendo do contexto social, seja individual ou em grupo. Em contraste, a personalidade e o afeto positivo não tiveram um impacto geral nas respostas de valência e excitação dos participantes durante as experiências com vídeos curtos e longos.

Finalmente, quando houve a correlação das emoções com os dados fisiológicos, os resultados sugeriram que os diversos dados, como o EEG, ECG e o GSR, podem ser úteis para prever emoções e traços de personalidade. O EEG foi o sinal que melhor apresentou desempenho. Os dados visuais também apresentaram ser eficazes na previsão de emoções.

#### 2.4.3 DREAMER Dataset

O DREAMER [17] contém sinais EEG e ECG que foram registados durante estímulos audiovisuais através de 18 clips de filme a 25 participantes. Estes clips contêm uma ampla gama de emoções, sendo que dois de cada continham uma das 9 emoções: diversão, excitação, felicidade, calma, raiva, nojo, medo, tristeza e surpresa. Todo o processo foi bastante similar ao feito pelo *dataset* DEAP. Para estabelecer um estado emocional de referência um clip de filme neutro foi mostrado antes de cada clip emocional. Os sinais EEG e ECG foram gravados, sendo que depois para todos os sinais EEG foram atribuídos 3 estados binários sendo estes:

- Baixa/alta valência.
- Baixa/alta excitação.
- Baixa/alta dominância.

Além da gravação desses sinais, os participantes tiveram que avaliar a valência, excitação e dominância para cada filme visualizado, usando o SAM.

Para começar a analisar as avaliações feitas, os autores inicialmente buscaram entender se havia alguma anormalidade ou variação inexplicada entre as respostas dos participantes. Para avaliar a concordância entre os participantes, foi calculado o coeficiente de variação (CV) para as avaliações de todos os clips de filme. Os resultados indicaram baixa variabilidade nas avaliações dos participantes, com CVs próximos de zero para todas as escalas. Isso sugere que as avaliações dos participantes estavam em grande concordância.

As médias das avaliações dos participantes para valência e excitação foram apresentadas em gráficos e tabelas, mostrando uma ampla gama de classificações e um conjunto de dados equilibrado. As classificações dos participantes estavam alinhadas com as emoções pretendidas pelos clips de filmes. Além disso, foram realizadas análises de correlação, utilizando o coeficiente de correlação de Spearman, para explorar as relações entre as escalas de avaliação. Houve uma forte correlação positiva entre excitação e dominância, indicando que os participantes se sentiram empoderados quando excitados e impotentes quando entediados. No entanto, não houve correlação significativa entre valência e excitação, e a correlação entre valência e dominância foi fraca e não significativa.

Os autores também avaliaram os dados em relação às avaliações de outro estudo que utilizou exatamente os mesmos estímulos audiovisuais. Eles confirmaram que houve correlações bastante significativas e baixa variação entre as avaliações feitas.

Em relação à comparação entre os sinais fisiológicos e as emoções, o estudo obteve que existiu uma forte correlação entre os sinais (EEG e ECG) e as emoções. Esta correlação usou a análise de correlação de Spearman. Para tentar obter valores mais precisos, os autores realizaram uma fusão entre os vários sinais. Apesar da tentativa os resultados da fusão desses sinais não mostraram melhorias significativas na precisão em relação ao uso dos sinais separadamente. Além disso,

a correlação entre os sinais foi forte, o que indica que ambos fornecem informações bastante semelhantes sobre as emoções.

#### 2.4.4 GAMEEMO Dataset

O *dataset* GAMEEMO [1] inclui sinais de EEG baseados em jogos de computador. Estes dados vieram trazer uma abordagem diferente dos outros conjuntos referidos anteriormente, pois a ideia principal por trás desta abordagem é que os jogos de computador são eficazes não só de forma física, mas também de forma psicológica. Nos jogos de computador, os participantes não só observaram os estímulos, mas também experimentaram o cenário em primeira pessoa.

O conjunto de dados EEG resultou da exposição a estímulos audiovisuais que corresponderam a quatro jogos diferentes. Cada jogo foi jogado por 5 minutos e cada um estava rotulado de acordo com os seus géneros como aborrecidos, calmos, horrorizados e engraçados. Eles foram apresentados aos participantes na mesma ordem. Os dados foram recolhidos de 28 indivíduos, usando um dispositivo portátil contendo 14 canais EEG.

Para determinar se os sinais representavam o que o utilizador sentia, foi implementado o formulário SAM para cada assunto. Para a análise dos dados do SAM, em primeiro lugar, foi calculado o desvio padrão, a variação, a média e a distribuição das classificações. Depois de feitos os cálculos, a relação entre as classificações foi descrita.

Devido a existir movimentos corporais, expressões faciais e a linguagem corporal quando os utilizadores jogaram os jogos, os dados EEG que os autores recolheram tiveram de ser processados com cuidado para eliminar artefactos de movimento. Os jogadores além dos dados EEG também responderam a perguntas relacionadas às experiências que tiveram durante os jogos. O objetivo da recolha de dados foi investigar como os sinais de EEG podem ser úteis para a compreensão das respostas emocionais das pessoas durante o jogo de computador.

Os resultados mostraram que os jogos escolhidos despertaram os sentimentos desejados e receberam avaliações consistentes dos participantes.

Para analisar os dados, foram realizadas análises estatísticas, incluindo ANOVA (Análise de Variância) e o teste de Friedman, a fim de avaliar os efeitos dos jogos nas respostas emocionais dos participantes. Os resultados indicaram diferenças significativas nas avaliações de valência e arousal entre os diferentes jogos.

Análises de frequência-tempo e extração de características dos sinais de EEG foram realizadas para relacionar as emoções relatadas pelos participantes com os sinais fisiológicos captados pelo EEG. A utilização da Discrete Wavelet Transform (DWT) foi necessária para a análise de frequência-tempo. Essa transformação divide os sinais originais do EEG em quatro sub-sinais. Em seguida, várias características relevantes foram extraídas desses sub-sinais, incluindo características estatísticas, características caóticas e características de análise de frequência-tempo. Estas características foram selecionadas para identificar possíveis indicadores dos estados emocionais dos participantes.

Utilizando as características extraídas dos sinais EEG, os investigadores aplicaram três algoritmos de classificação diferentes. Estes algoritmos foram treinados para aprender a relação entre os padrões nos sinais de EEG e as emoções relatadas pelos participantes nos questionários. As duas tarefas de classificação foram: estimar se as emoções eram positivas ou negativas e classificar a dimensão de excitação e valência das emoções.

Os resultados sugerem que os sinais EEG capturaram informações relevantes para a classificação das respostas emocionais dos participantes. No entanto, os resultados variaram dependendo do canal de EEG e do algoritmo de classificação utilizado. Olhando para o desempenho dos classificadores, concluíram que na classificação binária de emoções (positivas vs. negativas) houve um maior desempenho face a classificação multiclasse (distinguir entre várias emoções específicas). Além disso, os classificadores apresentaram uma taxa de sucesso geralmente superior

na classificação de emoções negativas em comparação com emoções positivas.

#### 2.4.5 Overview dos *Datasets*

Apesar de todos os *datasets* terem as suas particularidades, todos acabam por seguir uma linha cronológica de atividades bastante parecida. A Figura 2.7 representa uma possível cronologia geral dos *datasets* descritos anteriormente.

Como podemos ver todos eles começam com a seleção do conteúdo, que ainda é abrangente.

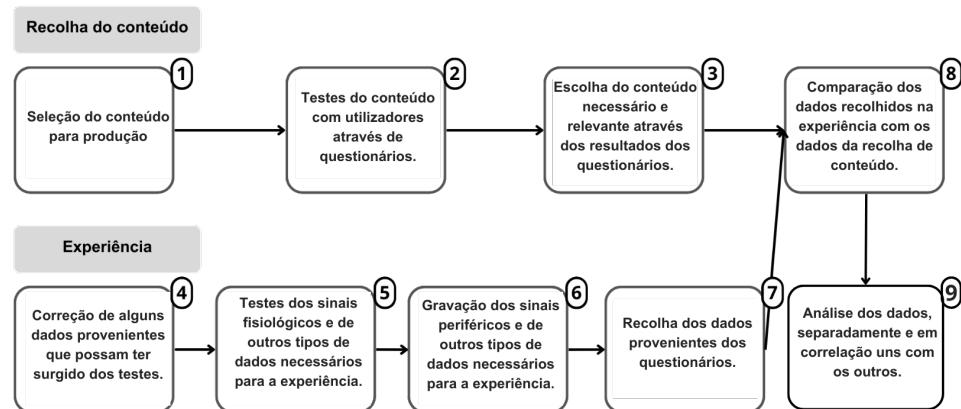


Figura 2.7: Estrutura para a criação dos *Datasets*.

Em seguida fazem testes para diminuir a seleção inicial e para conseguirem chegar mais próximo do conteúdo desejado. De notar que estes testes muitas vezes só acontecem fazendo uso dos questionários standard e outras ferramentas pois é um período de seleção de conteúdo para ter diversidade e não um período da experiência principal.

A maioria destes *datasets* verifica as informações que foram recolhidas após a conclusão da experiência, durante a qual também são registados os sinais fisiológicos. Eles fazem isso comparando-os com os dados que foram recolhidos durante a fase de recolha de conteúdo, com informações de outros conjuntos de dados ou com diversos tipos de dados que foram registados durante a experiência, caso existam. Este processo de validação é crucial porque garante que o passo seguinte, que envolve a verificação da correlação entre os dados, seja confiável.

## 2.5 Sinais Fisiológicos e Questionários Standard

Apesar dos estudos a relacionar os sinais fisiológicos com questionários standard serem ainda bastante recentes, existem alguns que usam a atividade cerebral dos utilizadores para avaliar a experiência de utilização numa determinada aplicação.

Um dos poucos estudos na área, que não é recente [25], revelou que poderia existir uma boa correlação entre os sinais EEG e os questionários de usabilidade. Os autores realizaram uma experiência onde pediram aos utilizadores para produzirem atividades no Excel e depois de concluirem todas as tarefas, preencher o questionário QUIS. Estas tarefas foram feitas em duas versões de Excel diferentes. Uma que os utilizadores já estavam mais habituados e outra que era completamente nova para os mesmos pelo pouco uso. Quando fizeram a comparação dos dados do questionário com os dados do EEG conseguiram verificar que existia uma correlação significativa entre eles.

Um estudo mais recente [8], que tira partido do anterior, fez uma experiência parecida, mas usando dois tipos de smartphones onde cada utilizador realizou três tarefas em cada um deles.

Durante a realização das tarefas recolheram-se os sinais cerebrais de cada participante e no final da experiência os mesmos avaliaram a experiência de utilização através de um questionário. Este era um questionário modificado e não um questionário standard, contendo perguntas para identificar, o fluxo, o prazer, a facilidade de uso, o controle, a funcionalidade e finalmente a imersão. Cada pergunta era respondida usando uma escala de Likert de 1-”Discordo Totalmente” a 7-”Concordo Totalmente”. Os resultados da análise de dados mostraram que os participantes concluíram as três tarefas mais rapidamente usando um smartphone com uma maior pontuação de UX. Além disso houve uma correlação significativa entre os resultados do EEG e os itens referidos anteriormente do questionário de UX. Os autores concluíram que as diferenças produzidas no EEG podem ser tomadas como um indicador de avaliação em relação ao que o utilizador transmitiu, através dos questionários, ao usar um determinado produto sem qualquer interrupção.

Outro estudo recente [7] focou-se na captação da atividade cerebral através do EEG para analisar os dados de um grupo de pessoas, enquanto interagiam com um jogo clássico 2D chamado Snake. Este estudo foi dividido em duas interações onde a variável que mudava era onde os utilizadores jogaram o jogo. Na primeira interação jogaram o jogo no computador com teclado. Na segunda houve o envolvimento do dispositivo móvel. Ambas as interações tiveram o mesmo processo, os utilizadores tinham de fechar os olhos por 90 segundos para haver uma neutralização das emoções, depois dessa neutralização havia uma interação com o sistema interativo durante cerca de 120 segundos. Os participantes responderam ainda aos questionários SAM, NASA-TLX, e SUS.

A análise dos dados EEG indicou que as emoções dos participantes foram afetadas pelo tipo de interação com os sistemas. Especificamente, a segunda opção parece ter causado uma resposta emocional mais intensa em comparação com a primeira opção. Além disso, a avaliação da usabilidade dos sistemas também apontou para a primeira opção como sendo mais fácil de usar do que a opção 2, com pontuações mais altas no questionário SUS. Isso sugere que os questionários standard conseguiram captar informações semelhantes às obtidas pelos sinais fisiológicos, destacando a sua utilidade na avaliação da experiência de utilização.

Um trabalho anterior feito pelos mesmos autores e com um procedimento praticamente igual [6], apenas com a diferença que neste caso os utilizadores interagiram com dois sistemas interativos, revelou que existia uma correlação entre os questionários SUS, NASA-TLX e SAM e as características dos sinais EEG. Esta correlação indica que esta técnica poderá ser uma alternativa que permite a avaliação da interação de sistemas. Além deste trabalho referido anteriormente muitos outros como [5] seguem mais ou menos os mesmos procedimentos mudando apenas a quantidade ou os questionários standard que usam e os dispositivos para extrair os sinais EEG. Os estudos anteriores mostram ainda que sistemas de captura de sinais de baixo custo conseguem retirar dados suficientes para obter resultados bastante similares aos dados que são retirados dos sistemas sofisticados.

A procura da carga mental do utilizador durante diferentes tarefas de interação aparece em mais estudos, como o do Frey [13], em que o foco do estudo era investigar os estados mentais dos jogadores durante a interação com um labirinto digital num ambiente virtual, utilizando dois dispositivos de entrada diferentes: um teclado convencional e uma interface sensível ao toque (*touch screen*). Uma particularidade deste trabalho está no facto dos autores usarem diferentes níveis de carga de trabalho mental quer através da mudança de dispositivos, quer manipulando o ambiente virtual (como por exemplo, através da profundidade, número de direções, velocidade e orientação espacial). Através destas mecânicas criaram 4 níveis de jogo: fácil, médio, difícil e ultra. Os níveis de carga de trabalho mental aumentavam consoante a dificuldade do nível.

Para verificarem que os níveis produziam a carga de trabalho mental pretendida os autores realizaram um teste piloto. Neste não gravaram os sinais fisiológicos (EEG), mas usaram o questionário NASA-TLX para recolher a carga de trabalho. Os resultados mostraram que cada nível de dificul-

dade induzia um nível diferente de carga de trabalho.

O estudo principal foi então realizado, gravando os sinais EEG com os mesmos níveis de dificuldade do teste piloto, mas com a adição dos dois dispositivos de interação. As hipóteses que este estudo principal tinha eram:

- Se a carga de trabalho medida pelo EEG era maior no sistema tático e se aumentava também com a dificuldade, refletindo o que aconteceu no teste piloto.
- Se o foco que os participantes atribuíram aos vários sons diminuía à medida que a dificuldade aumentava.
- Se a condição tática induzia um maior número de erros em relação ao teclado.

Praticamente todas estas hipóteses foram validadas na análise dos dados. Os resultados do NASA-TLX nos testes piloto foram replicados no estudo principal.

Os sinais PPG não foram usados em praticamente nenhum estudo que os relacione com os questionários standard, tanto para tentarem entender a experiência de utilização como a usabilidade.

O estudo [37] usa um questionário simples para entender a diversão, o desafio e o esforço nas operações que os utilizadores fazem, sendo este aplicado depois da realização da tarefa. Além deste questionário foi feito um outro para entender o nível de conhecimento que os utilizadores já tinham em relação ao jogo Plants vs. Zombies.

Sem contar a utilização do questionário fizeram o uso do *Stress-Induced Vascular Response Index* (sVRI) que serviu para avaliar dinamicamente a carga cognitiva dos jogadores durante os jogos que foram realizados. Os questionários tiveram um papel secundário neste trabalho, mas houve comparações que mostraram que quanto maior a carga cognitiva, menor a diversão dos utilizadores. Também fizeram a comparação entre os dados retirados do nível de conhecimento do jogo e a carga cognitiva entre os diferentes utilizadores nas diferentes ações.

## 2.6 Correlação entre Emoções e Métricas de Usabilidade

Apesar de existirem poucos estudos sobre a relação entre emoções e métricas de usabilidade, o estudo [31] procura fazer esta relação, mais especificamente os autores realizaram testes de usabilidade em larga escala, usando um design *within-subject* (um design onde a mesma pessoa testa todas as condições) entre dois sites de usabilidade variável. As emoções foram avaliadas por meio de análise de sentimento baseada em texto, análise de emoção baseada em fala e análise de emoção por vídeo.

As questões de pesquisa deste trabalho foram três. A primeira delas busca investigar se há uma relação entre os resultados das ferramentas de reconhecimento de sentimentos / emoções e as métricas de usabilidade. A segunda questão de pesquisa procura determinar se ocorrem diferenças nos resultados das ferramentas de reconhecimento de sentimentos / emoções ao comparar dois produtos com níveis variáveis de usabilidade. Por último, a terceira questão aborda a capacidade dos modelos de regressão, baseados nos resultados das ferramentas de reconhecimento de sentimentos / emoções, para prever as métricas de usabilidade. Foi estabelecido como métricas de usabilidade, a taxa de conclusão da tarefa, o SUS e finalmente o tempo para completar uma tarefa.

Na inferência estatística da primeira questão de pesquisa (correlações entre métricas de usabilidade e sentimentos ou métricas de emoções), identificaram que existia uma correlação positiva significativa entre os valores do SUS e o resultado da análise de sentimento textual (quanto maior a valência maior é o score do SUS). Para a taxa de conclusão da tarefa não conseguiram encontrar nenhuma correlação significativa. Além da correlação com o resultado da análise de sentimento textual, para os valores do SUS, duas categorias de emoção da análise emocional baseado na

fala mostraram correlação significativa, sendo estas a neutralidade e a felicidade. Finalmente a correlação entre os valores do SUS e a análise de emoção baseada no rosto foi nula, não houve correlação significativa nem na correlação nem na taxa de conclusão da tarefa.

A explicação que tiveram para estes resultados foi que ou as ferramentas de reconhecimento de emoções têm limitações inerentes ligadas à tarefa específica de testes de usabilidade, ou que as ferramentas e métodos utilizados não são sofisticados o suficiente, ou então não se encaixa na tarefa específica.

## 2.7 Relação Entre os Fundamentos Teóricos e o Trabalho Relacionado

Depois da análise das componentes dos fundamentos teóricos e do trabalho relacionado entendemos que os questionários como o SUS, o NASA-TLX e o SAM acabam por transparecer nos trabalhos relacionados. Os estudos e os *datasets* na sua generalidade antes de produzirem qualquer tipo de resultados produzem validações, por este facto, adicionar este passo é bastante importante, pois só com uma boa validação os resultados poderão ser precisos. Esta validação nos trabalhos relacionados ocorreu sempre de três maneiras, fazendo a comparação dos dados retirados de testes anteriores com os dados da experiência principal, ou então compararam com outros trabalhos já existentes na área, ou ainda com diferentes tipos de dados que foram registados durante a experiência.

Pode ser interessante, além destas duas validações, usar, caso haja, questionários que completem / validem os que foram usados. O resto da estrutura dos *datasets* é bastante coesa podendo haver pequenas diferenças em como recolher a informação da experiência ou na recolha do conteúdo para ser usado na experiência.

Estes trabalhos mostram que para produzir dados de testes que possam ser equiparados e que não tragam inconsistências, precisamos de ter sistemas bastante similares entre si ou o mesmo e alterar certas configurações dentro dele que achamos que irão provocar o que pretendemos.

## 2.8 Síntese

Neste capítulo, exploramos diversos tópicos que desempenham um papel fundamental no nosso estudo. Iniciamos por analisar em detalhe os questionários standard, explicando cada um deles em pormenor, abordando a sua estrutura, escalas de avaliação, aplicações específicas e as vantagens e desvantagens associadas a cada um.

Em seguida, procedemos a uma análise dos sinais fisiológicos disponibilizados pelo dispositivo Muse2.

Após essa análise dos sinais fisiológicos, aprofundamos a nossa análise sobre as emoções, bem como os modelos teóricos que têm vindo a ser desenvolvidos ao longo do tempo para uma compreensão mais aprofundada desta dimensão.

Discutimos ainda conjuntos de dados que incluem sinais fisiológicos, como o DEAP, o AMIGOS, o DREAMER e o GAMEEMO, abordando os respetivos procedimentos experimentais, os questionários e os sinais fisiológicos utilizados, bem como as análises aplicadas aos dados.

Para além dos conjuntos de dados, examinamos estudos que procuram utilizar os dados fisiológicos para obter as mesmas respostas que são obtidas dos questionários standard. Nestes estudos, encontramos uma notável semelhança entre os dois tipos de dados.

Por fim, investigamos a correlação entre as emoções e as métricas de usabilidade, observando que, embora existam poucos estudos nesta área, alguns já conseguiram estabelecer uma ligação entre esses dois elementos.

O capítulo conclui, enfatizando as interações e ligações entre estes tópicos, sublinhando a importância de uma abordagem integrada para o nosso estudo.



# Capítulo 3

## Criação do *Dataset*

No decorrer deste capítulo, descrevemos o processo de construção do *dataset*, que é o foco principal do nosso trabalho e apresentamos a metodologia usada para a criação do mesmo.

Descrevemos as duas experiências realizadas, onde a primeira teve como propósito a seleção do conteúdo, sendo posteriormente também relevante para a validação do *dataset*. Já a segunda experiência teve como objetivo extrair o conteúdo necessário para a produção do *dataset*.

É importante destacar que ambas as experiências apresentadas neste capítulo foram aprovadas pela Comissão de Ética de Ciências <sup>1</sup>. Apresentamos os detalhes das etapas da elaboração da experiência subsequente responsável pela colheita de dados destinados ao nosso *dataset*. Por fim, explicamos como todo o *dataset* foi moldado com a informação recolhida anteriormente. Cada uma destas fases desempenha um papel fundamental na asseguração da qualidade e robustez dos resultados alcançados.

### 3.1 Metodologia

Nesta secção, apresentamos a metodologia detalhada que usamos para produzir os dois estudos que foram fundamentais para a criação do nosso *dataset*.

Antes de apresentarmos os estudos é necessário descrever tudo o que os compõe, as fontes de dados para a experiência, os questionários adotados, o processo de seleção dos participantes e qual o equipamento usado para fazer a produção das experiências.

#### 3.1.1 Aplicações Interativas

Como aplicações interativas para o nosso trabalho, escolhemos sites de compras online pela similaridade nas suas operações. Optamos, deliberadamente por sites que apresentam métodos de funcionamento semelhantes ou idênticos, para garantir a consistência nas tarefas a pedir aos participantes e nos resultados obtidos nas tarefas. As tarefas pedidas nas experiências estão descritas no anexo <sup>2</sup>.

Considerando que diferentes sites podem variar significativamente em termos de fluxo de trabalho e procedimentos, a inclusão de plataformas com métodos divergentes poderia levar a resultados discrepantes nos questionários preenchidos pelos participantes. Isso se deve ao fato de que as formas distintas de realizar uma determinada atividade em cada site poderiam influenciar consideravelmente as respostas fornecidas.

---

<sup>1</sup>Os documentos submetidos encontram-se no Apêndice A

<sup>2</sup>Os documentos submetidos encontram-se no Apêndice C

Ao escolher sites de compras com funcionamento similar, busca-se minimizar a variabilidade dos dados recolhidos, proporcionando uma base mais sólida. Esta abordagem permite uma comparação mais direta e precisa das respostas dos participantes, sem que as diferenças nos métodos de funcionamento dos sites introduzam problemas nos resultados. Os sites para este estudo foram selecionados após um processo de escolha que integrou testes iniciais, tais como testes de desempenho e usabilidade, bem como testes piloto.

Num primeiro momento, avaliamos a ampla diversidade de sites usando critérios como usabilidade e experiência de utilização, escolhendo um total de 44. Depois, realizámos testes com utilizadores para confirmar se os resultados demonstrados se refletiam na opinião dos mesmos. Esta abordagem resultou na seleção de 24 sites para a experiência inicial, assegurando não só qualidade inerente, mas também uma experiência confiável para os utilizadores.

### 3.1.2 Questionários Standard Usados

Após uma análise aprofundada dos diversos questionários standard, identificámos aqueles que melhor se adequavam a este projeto. Com o objetivo de encontrar um equilíbrio entre a medição da usabilidade, evitando sobrecarregar o utilizador, e a deteção de extremos de usabilidade (tanto elevada como muito baixa), optámos por recolher dados através do questionário *System Usability Scale* (SUS) e também por utilizar o *Usability Metric for User Experience LITE* (UMUX-LITE) para complementar e conferir maior validade aos dados.

A revisão dos trabalhos relacionados revelou-nos a inevitabilidade da utilização do *NASA Task Load Index* (NASA-TLX). A maioria desses estudos emprega-o devido à sua fácil correlação com os sinais fisiológicos, dada a avaliação subjetiva da carga de trabalho mental.

Dado que reconhecemos a ligação entre a experiência de utilização e a usabilidade, considerámos relevante enriquecer o nosso estudo ao incorporar o *User Experience Questionnaire* (UEQ). No entanto, apesar do desejo de incluir o UEQ, compreendemos que, devido à sua complexidade e extensão, poderia tornar-se excessivamente exaustivo para o utilizador. Como solução para esta questão, optámos por utilizar o UEQ-S, uma versão mais concisa, conforme demonstrado na Figura 3.1.

obstructive	o o o o o o o o	supportive	1
complicated	o o o o o o o o	easy	2
inefficient	o o o o o o o o	efficient	3
clear	o o o o o o o o	confusing	4
boring	o o o o o o o o	exitng	5
not interesting	o o o o o o o o	interesting	6
conventional	o o o o o o o o	inventive	7

Figura 3.1: Questionário UEQ-S retirado do estudo [34].

Tivemos ainda a preocupação de verificar se todos ou quase todos os pontos do UEQ-S são independentes em relação à usabilidade. Ao constatar que não existe esta dependência, o nosso próximo objetivo consistiu em criar quatro quadrantes distintos num espaço bidimensional, com os valores do SUS no eixo dos y e os valores do UEQ-S no eixo dos x, como é ilustrado na Figura 3.2.

Estes quadrantes irão auxiliar-nos a identificar, dentro de cada um deles, os extremos da relação entre a usabilidade e a experiência de utilização.

É pertinente realçar que a identificação de pontos de referência nos quadrantes de elevado SUS

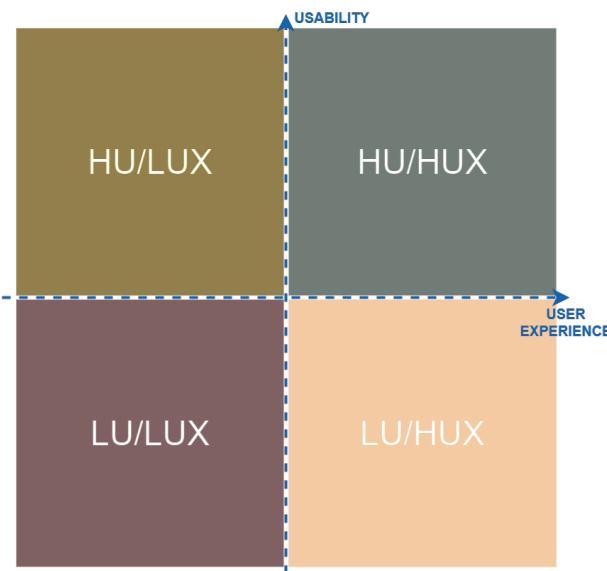


Figura 3.2: Quadrantes SUS e UEQ-S.

e baixo UEQ-S, bem como nos quadrantes de alto UEQ-S e baixo SUS, é um desafio pois muitas vezes o utilizador acaba por inferir uma interdependência subjacente entre o SUS e o UEQ-S, caso exista uma experiência de utilização positiva implica automaticamente uma alta usabilidade e vice-versa.

Por último, selecionamos mais dois questionários, o *Self-Assessment Manikin* (SAM), dada a correlação existente entre emoções e usabilidade, juntamente com a sua ampla utilização em estudos semelhantes e o *Single Ease Question* (SEQ) por ser um questionário de apenas uma pergunta e que nos trás dados importantes como a facilidade de utilização.

### 3.1.3 Participantes

Os participantes foram convocados por meio de divulgação em redes sociais, grupos institucionais e conhecidos. Para selecionar participantes para o estudo, os critérios de inclusão e exclusão foram rigorosamente seguidos. Critérios de Inclusão:

1. **Idade superior a 18 anos:** Os candidatos teriam que ser adultos para participar na experiência.
2. **Uso diário de um computador:** Os participantes deveriam utilizar computadores regularmente como principal meio de acesso à internet.
3. **Tenha utilizado alguns sites de compras ocasionalmente:** Era fundamental que os candidatos tivessem alguma experiência prévia em compras online, mesmo que esporadicamente, para que pudessem fornecer feedback significativo sobre a experiência geral de compra.
4. **Possuir um nível básico de inglês:** Foi necessário os participantes terem um nível básico de inglês para facilitar a interação com alguns dos sites que são apresentados.
5. **Escolaridade mínima secundária ou superior:** Este critério foi estabelecido para garantir um certo nível de educação formal dos participantes.

Na busca por uma amostra representativa para o estudo, abordamos cerca de 200 potenciais participantes. Enfrentamos desafios, como a falta de disponibilidade devido a compromissos pessoais e profissionais, resistência ao uso do dispositivo Muse2 e percepção de que o tempo das experiências era demasiado longo. Desses 200 pessoas abordadas, conseguimos recrutar 78 participantes para as experiências.

### 3.1.4 Equipamento

Os equipamentos empregues nas duas experiências consistem em computadores, cuja origem poderá ser o dispositivo pessoal do utilizador ou um equipamento fornecido pelo investigador para a realização das tarefas nos sites em estudo. Não se requer um computador de elevada potência, uma vez que as tarefas a executar nos sites não são computacionalmente pesadas, consistindo maioritariamente na abertura e execução de tarefas simples. Portanto, a especificação do computador não se mostra crítica.

Para a gravação dos dados recolhidos pelo dispositivo Muse 2, bem como para a exibição dos gráficos que representam as várias formas de onda dos sinais fisiológicos durante o decorrer da experiência, utilizamos um segundo computador com um processador Ryzen 5 5600H e uma placa gráfica RTX 3050.

Finalmente usamos o Muse 2 que assume um papel vital na captação dos diversos sinais fisiológicos enquanto os utilizadores interagem com as atividades propostas durante a experiência.

## 3.2 Experiência 1

Nesta seção, exploramos detalhadamente a primeira experiência realizada, aprofundando o Procedimento Experimental que ocorreu e os Dados Recolhidos. O intuito principal desta experiência foi selecionar criteriosamente 12 sites a serem incluídos na próxima fase da pesquisa.

### 3.2.1 Procedimento Experimental

Nesta experiência, os participantes, num total de 20 (13 participantes do sexo masculino e 7 do sexo feminino), realizaram tarefas em seis sites de compras, selecionados a partir de um de quatro grupos.

A experiência foi conduzida num ambiente confortável, livre de ruídos, com iluminação excelente e sem qualquer tipo de distração. Caso algum utilizador já estivesse familiarizado com um site do grupo atribuído, este site era substituído por um de outro grupo. Se esta situação ocorresse, o site que não foi utilizado pelo primeiro utilizador do grupo seria atribuído a outro participante. Após completar cada uma das tarefas, os utilizadores responderam aos questionários selecionados anteriormente sobre as tarefas que realizaram.

Além dos questionários, que foram mencionados na Secção 3.1.2, os utilizadores preencheram também um questionário demográfico. Este questionário foi colocado no início do processo pois quisemos simplificar a aceitação de todos os pontos relacionados com o Formulário de Consentimento Informado, adicionando uma opção neste formulário onde os utilizadores poderiam aceitar todos os pontos do mesmo. O primeiro passo que foi feito antes da realização de toda a experiência referida anteriormente foi explicar a cada participante a experiência, os seus objetivos, os vários passos que a compuseram, que o que pretendíamos era recolher os dados e não avaliar o seu desempenho e que todos os dados seriam recolhidos de forma anónima. Além disso, não existia tempo limite para cada sessão e que poderiam desistir em qualquer momento. Avisámos também que tinham a opção de fazer uma pequena pausa entre cada tarefa, se assim o desejassem.

Os sites que foram utilizados nesta experiência passaram por uma pré-seleção já referida anteriormente. Esta pré-seleção reuniu 24 sites que foram distribuídos como já referido anteriormente em 4 grupos de 6, como exibido na Figura 3.3.

Grupo 1	Grupo 2	Grupo 3	Grupo 4
Instant Gaming <a href="http://www.instant-gaming.com">www.instant-gaming.com</a>	Eneba <a href="http://www.eneba.com">www.eneba.com</a>	Kinguin <a href="http://www.kinguin.net">www.kinguin.net</a>	Kiabi <a href="http://www.kiabi.pt">www.kiabi.pt</a>
Blinkee <a href="http://www.blinkee.com">www.blinkee.com</a>	Toronto Cupcake <a href="http://www.torontocupcake.com">www.torontocupcake.com</a>	Penny Juice <a href="http://www.pennyjuice.com">www.pennyjuice.com</a>	Airbnb <a href="http://www.airbnb.com">www.airbnb.com</a>
Ralph Lauren <a href="http://www.ralphlauren.pt">www.ralphlauren.pt</a>	Frank and Oak <a href="http://www.frankandoak.com">www.frankandoak.com</a>	Aboutyou <a href="http://www.aboutyou.pt">www.aboutyou.pt</a>	Roverp6parts <a href="http://www.roverp6parts.com">www.roverp6parts.com</a>
Create and Barrel <a href="http://www.crateandbarrel.com">www.crateandbarrel.com</a>	Vitra <a href="http://www.vitra.com">www.vitra.com</a>	WayFair <a href="http://www.wayfair.com">www.wayfair.com</a>	Grainne Morton <a href="http://www.grainnemorton.co.uk">www.grainnemorton.co.uk</a>
Lings Cars <a href="http://www.lingscars.com">www.lingscars.com</a>	Lexa Autolease <a href="http://www.lexautolease.co.uk">www.lexautolease.co.uk</a>	Pure Vehicle Leasing <a href="http://www.purevehicleleasing.co.uk">www.purevehicleleasing.co.uk</a>	Zenni Optical <a href="http://www.zenni optical.com">www.zenni optical.com</a>
Mesh Label <a href="http://www.meshlabel.com">www.meshlabel.com</a>	Connection <a href="http://www.connection.com">www.connection.com</a>	Allbirds <a href="http://www.allbirds.eu">www.allbirds.eu</a>	Dollar Tree <a href="http://www.dollartree.com">www.dollartree.com</a>

Figura 3.3: Grupo de sites da primeira experiência.

Estes grupos foram divididos de forma equitativa em termos de dificuldade em relação ao que pudemos inferir da pré-seleção. Antes de implementar a experiência com os 20 utilizadores, foram conduzidos testes piloto com 2 utilizadores para garantir a adequação do procedimento. O feedback obtido durante os testes piloto foi utilizado para aprimorar várias etapas do processo. Depois desse feedback algumas mudanças foram feitas ao desenho inicial da experiência, como por exemplo, adicionar uma restrição onde os participantes não podiam usar a barra de pesquisa. O objetivo desta restrição era avaliar se os vários filtros disponíveis nos sites satisfaziam os diversos requisitos e se o site apresentava uma organização eficiente.

Uma outra alteração implementada visou evitar a possibilidade de dados influenciados pelo cansaço ou aborrecimento dos utilizadores. Dado que o conjunto de seis sites exigia um tempo médio considerável para conclusão, conforme indicado nos testes piloto, que foram demorados, optámos por dividir a realização das seis tarefas ao longo de dois dias. Cada dia englobou a execução de tarefas em três sites. Para mitigar possíveis interferências nos resultados, estabelecemos um intervalo mínimo de cinco dias entre os dois dias de testes. Este intervalo serviu para reduzir variáveis externas que poderiam comprometer a validade dos resultados, tal como a familiarização excessiva com o funcionamento nestes tipos de sites, o que poderia distorcer as avaliações.

Todo este processo teve como objetivo principal recolher respostas dos questionários relativos aos diversos sites avaliados. Um dos propósitos era reduzir a lista inicial de 24 sites para um total de 12, com base nos resultados obtidos. Além disso, os resultados também serviram para ajudar na validação do *dataset*.

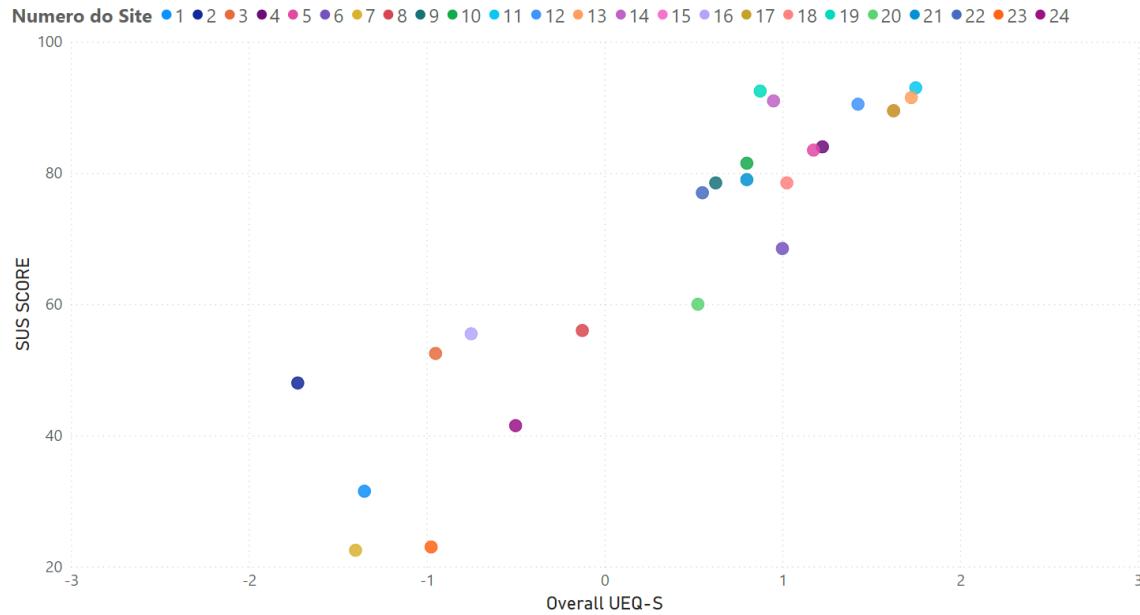


Figura 3.4: *Scatter Plot* Comparativo: Médias do UEQ-S com as Médias do SUS da Experiência 1.

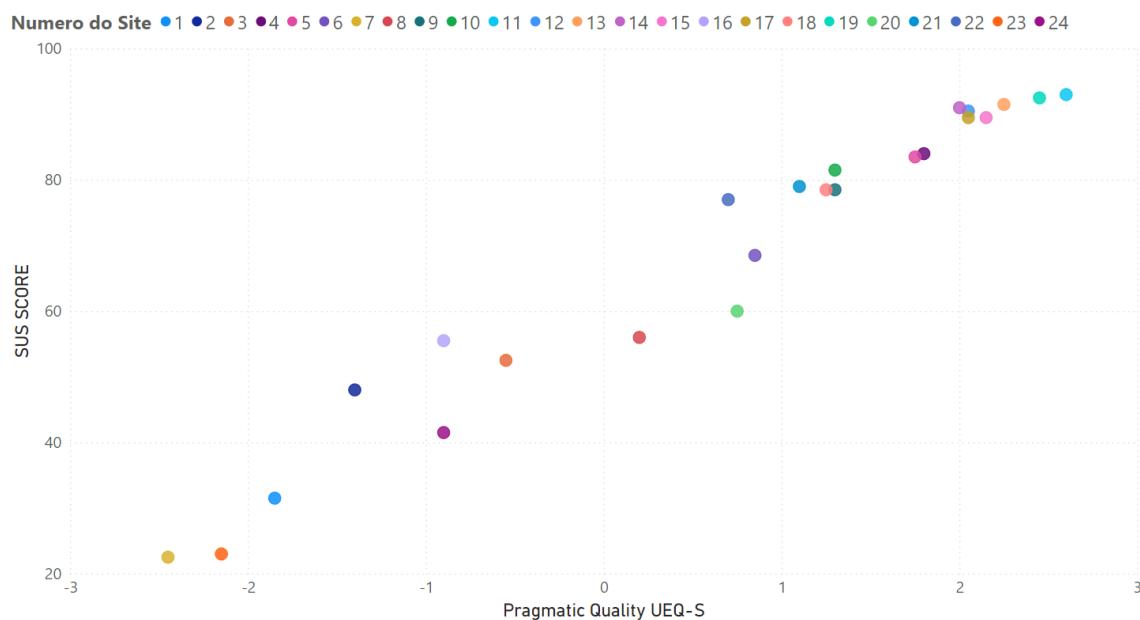


Figura 3.5: *Scatter Plot* Comparativo: Médias da qualidade pragmática com as Médias do SUS da Experiência 1.

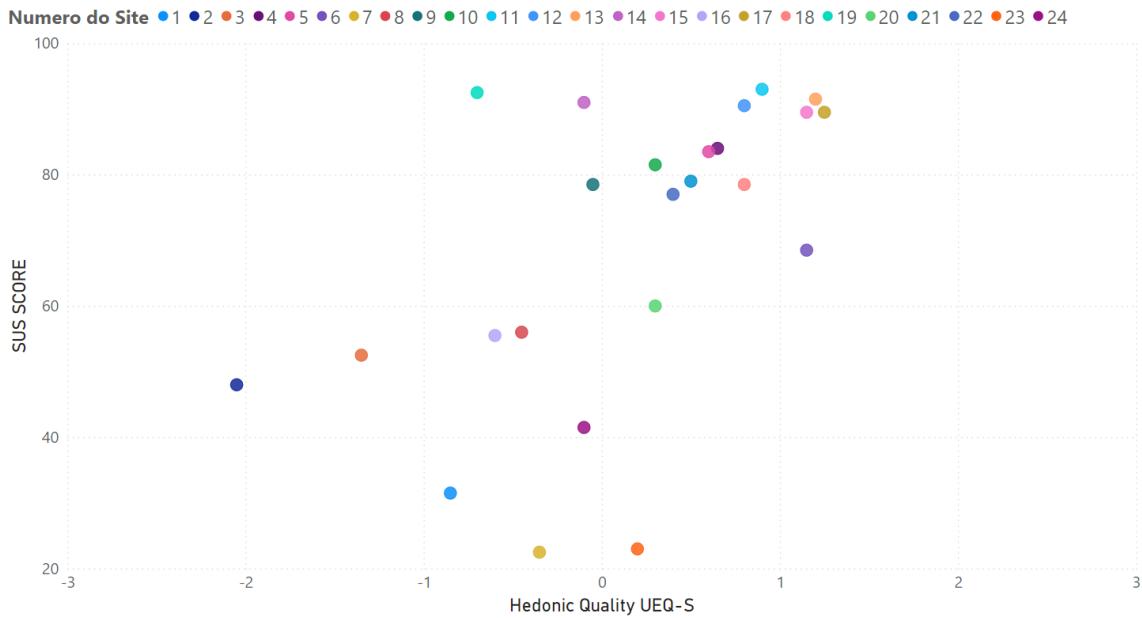


Figura 3.6: *Scatter Plot* Comparativo: Médias da qualidade hedônica com as Médias do SUS da Experiência 1.

### 3.2.2 Dados Recolhidos

Após a conclusão da fase de recolha de dados dos questionários, procedemos à análise dos mesmos com o objetivo de obter *insights* valiosos. Iniciamos a análise criando um gráfico de dispersão que relaciona a média do SUS com a média do UEQ-S de cada site (ver Figura 3.4). Este gráfico ofereceu uma visão preliminar das possíveis relações entre essas duas métricas consideradas essenciais para o nosso estudo.

Para aumentar a robustez das nossas conclusões, aprofundamos a análise com a criação de mais dois gráficos de dispersão. Desta vez, dividimos o UEQ-S nas suas duas dimensões: "Qualidade Pragmática" (ver Figura 3.5) e "Qualidade Hedônica" (ver Figura 3.6). Esta abordagem refinada permitiu-nos uma compreensão mais detalhada e pormenorizada da experiência de utilização em cada site.

Esta divisão surgiu porque a qualidade hedônica, no nosso caso específico, poderia gerar algumas interpretações equivocadas. Muitas vezes, em determinados sites, apesar dos valores da qualidade pragmática e do SUS serem bastante baixos, devido ao facto de o site ser substancialmente diferente, ao que os participantes estão habituados, estes acham a tarefa interessante e empolgante.

Um dos propósitos centrais desta experiência consistia em selecionar criteriosamente 12 sites a serem incluídos na etapa subsequente da pesquisa. Guiados pelos *insights* dos gráficos, adotamos uma abordagem estratégica. Escolhemos sites que abrangiam uma gama diversificada de posições no Gráfico 3.2 entre o SUS e o UEQ-S, sendo esses pontos na região mais negativa, na região central e na mais positiva. Esta abordagem multifacetada visava capturar as diferentes nuances que existiram durante a experiência.

É relevante destacar que identificar pontos de referência nos quadrantes de alto SUS e baixo UEQ-S, assim como alto UEQ-S e baixo SUS, apresentou-se como um desafio devido à interdependência subjacente entre o SUS e o UEQ-S nos diversos sites. Este problema acabou já por ser

referido mesmo antes da realização da tarefa na Secção 3.1.2.

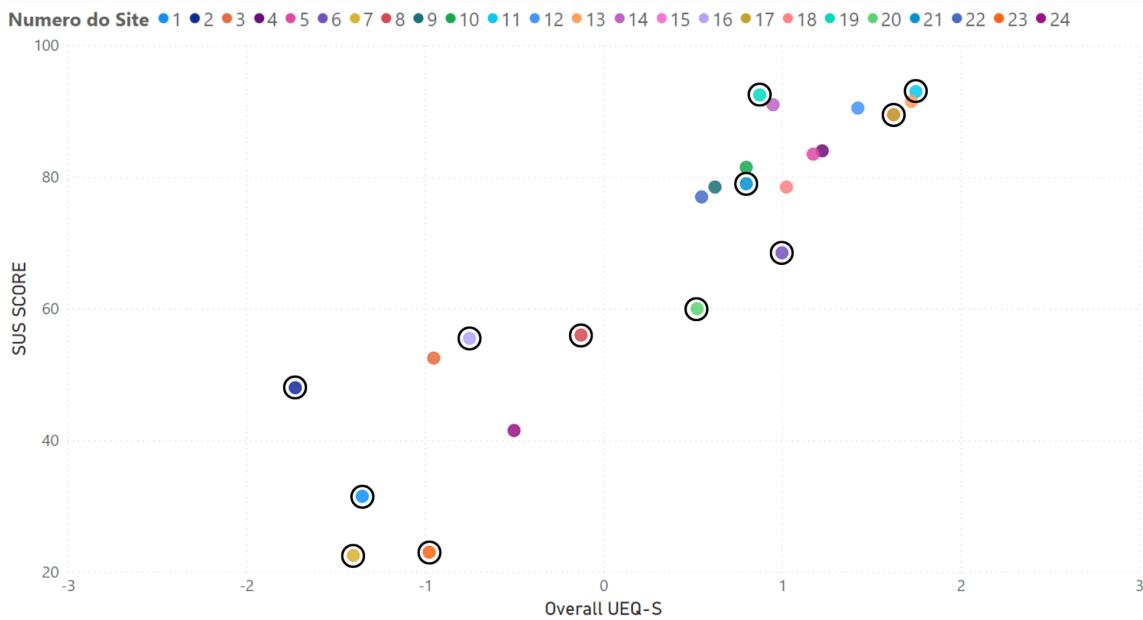


Figura 3.7: *Scatter Plot* com a escolha dos 12 sites.

Tendo em consideração todos os pontos mencionados anteriormente, foram selecionados os websites apresentados no gráfico de dispersão da Figura 3.7. Todos estes websites representam os indícios que desejávamos e mencionamos.

Todos os dados dos vários questionários foram registados para uma futura análise comparativa. Esta comparação foi essencial para a construção e validação do *dataset*.

### 3.3 Experiência 2

Nesta secção, exploramos em detalhe a segunda experiência realizada como parte desta investigação, a qual foi conduzida com o propósito de retirar os dados necessários para a construção do *dataset*.

Aprofundaremos esta experiência em três componentes principais: as Aplicações Tecnológicas utilizadas, o Procedimento Experimental e os Dados Recolhidos.

#### 3.3.1 Aplicações Tecnológicas

Nesta secção apresentamos as aplicações tecnológicas utilizadas na condução da experiência. Com o intuito de recolher os dados fisiológicos através do dispositivo Muse 2, utilizamos duas aplicações: o BlueMuse e o Muse LSL. A seguir, abordaremos o seu funcionamento e as adaptações necessárias para a sua integração no âmbito do nosso estudo.

##### BlueMuse

O BlueMuse [20] é um software de código aberto que suporta ondas EEG, PPG, acelerómetro, giroscópio e dados de telemetria. Foi concebido para ser utilizado no Muse 2 e no Muse S,

para transmitir dados através de uma ligação Bluetooth e fornecer visualização e análise de dados em tempo real. Além disso o BlueMuse suporta uma variedade de funcionalidades, tais como

<b>Muse 2</b>	<b>Signals</b>	<b>Rate (Hz)</b>	<b>Channels</b>
	EEG	256	TP9,AF7,AF8,TP10,RightAux
Accelerometer	52	X,Y,Z	
Gyroscope	52	X,Y,Z	
PPG	64	PPG1,PPG2,PPG3	

Figura 3.8: Sinais do Muse 2 e as suas configurações.

transmissão de dados em tempo real, análise de frequência, monitorização do estado mental e integração com aplicações de realidade virtual. Também permite a transmissão simultânea de dados de vários dispositivos Muse.

No nosso caso, usamos o BlueMuse como a peça-chave que torna a conexão entre o dispositivo Muse 2 e a ferramenta Muse LSL (explicada em seguida) possível. Esta conexão permite a transmissão contínua e precisa dos dados cerebrais e fisiológicos do Muse 2 para o Muse LSL. Esta sincronização é fundamental para garantir que as nossas análises sejam baseadas em informações confiáveis e em tempo real.

Na Figura 3.8, são apresentados os sinais que foram utilizados, bem como a taxa média de atualização de cada sinal e os canais contidos em cada um.

Já na Figura 3.9, é ilustrada a forma como o MUSE 2 é apresentado na aplicação Blue Muse, juntamente com as definições usadas na experiência em que este dispositivo foi utilizado.

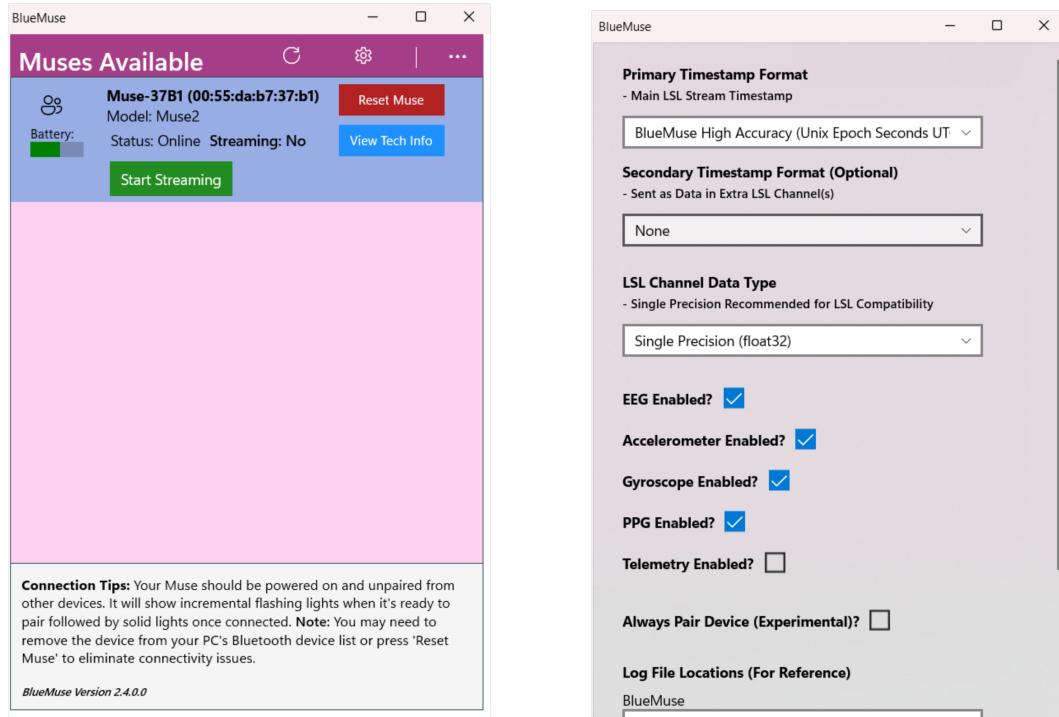


Figura 3.9: Aplicação BlueMuse e as definições usadas.

## Muse LSL

O MuseLSL é uma ferramenta que integra a tecnologia do dispositivo Muse com o protocolo de comunicação Lab Streaming Layer (LSL). Graças a esta junção, tanto investigadores como profissionais da neurociência podem tanto capturar como analisar dados cerebrais em tempo real com facilidade e precisão.

Esta ferramenta permite que os investigadores possam monitorar e registar a atividade elétrica do cérebro de forma não invasiva. Isto revela o funcionamento do cérebro em estudos de cognição e *neurofeedback*, entre outras pesquisas. Através da ligação entre o Muse e o LSL, o dispositivo EEG converte-se numa fonte confiável de dados cerebrais de elevada qualidade, que podem ser transmitidos e integrados com outras ferramentas de análise.

A integração com o protocolo LSL representa um dos principais benefícios do MuseLSL. O LSL é uma biblioteca de software que simplifica a comunicação e sincronização de dados em tempo real entre diferentes dispositivos e softwares de pesquisa. Com o MuseLSL, os pesquisadores podem transmitir os sinais EEG capturados pelo Muse diretamente para as suas aplicações e ambientes de pesquisa, permitindo análises em tempo real e a perfeita integração com outras tecnologias e sistemas de registo. Além disso, o MuseLSL oferece uma variedade de funcionalidades e opções de configuração que atendem às necessidades específicas de cada estudo ou pesquisa. Os utilizadores têm controlo total sobre a seleção dos canais de EEG, taxa de amostragem, filtros, entre outros. Essa flexibilidade permite que os investigadores personalizem as definições do MuseLSL de acordo com os requisitos da sua investigação, assegurando a captura e análise de dados cerebrais precisos e relevantes.

Outra vantagem do MuseLSL é a sua facilidade de utilização. O software MuseLSL, cuja configuração inicial é fácil de usar, permite que os pesquisadores se concentrem no seu próprio estudo em vez de se preocuparem com a complexidade técnica da aquisição de dados cerebrais.

## Alterações ao Muse LSL

No desenvolvimento da nossa pesquisa, encontramos um desafio essencial: a necessidade de utilizar as funções disponíveis no Muse LSL de maneira mais coesa e integrada. Embora o Muse LSL contivesse as funcionalidades que desejávamos, percebemos que estavam dispostas de forma fragmentada, o que dificultava o nosso trabalho.

Foi então necessário realizar várias adaptações em algumas das funções existentes para que pudessem ser aplicadas ao nosso estudo. Além disso, foi necessário incorporar novas funções para conseguirmos produzir todo o conteúdo da nossa investigação. Um dos aspectos cruciais foi a implementação de funcionalidades que permitissem a gravação simultânea de diversos sinais, como os sinais PPG, EEG, Acelerómetro e Giroscópio. Esta integração foi vital para garantir que obtinhamos um conjunto completo dos dados, permitindo uma análise mais abrangente e precisa.

Outro ponto relevante foi assegurar que a execução da experiência ocorresse sem contratempos. Para tal, desenvolvemos uma funcionalidade para exibir gráficos em tempo real ao longo de cada tarefa. Este recurso permitiu-nos não apenas monitorizar o progresso da experiência, mas também identificar anomalias que pudessem surgir durante a gravação.

Vale a pena destacar que a criação destes gráficos foi uma etapa cuidadosamente planeada. Optamos por gerá-los após o início da gravação dos ficheiros, de forma a garantir a consistência e uniformidade entre eles. Isso eliminou qualquer possibilidade de diferenças significativas que poderiam prejudicar a integridade dos nossos resultados.

### 3.3.2 Procedimento Experimental

O segundo estudo envolveu um total de 51 participantes, sendo 20 do sexo feminino e 31 do sexo masculino, para obter os dados finais para criar o *dataset*. A execução seguiu um processo semelhante ao do estudo anterior, porém com algumas variações importantes para garantir resultados imparciais e evitar viés de aprendizagem.

Grupo 1	Grupo 2	Grupo 3
Site 6 Airbnb <a href="http://www.airbnb.pt">www.airbnb.pt</a>	Site 1 Blinkee <a href="http://www.blinkee.com">www.blinkee.com</a>	Site 7 Roverp6parts <a href="http://www.roverp6parts.com">www.roverp6parts.com</a>
Site 17 Eneba <a href="http://www.eneba.com">www.eneba.com</a>	Site 2 Toronto Cupcake <a href="http://www.torontocupcake.com">www.torontocupcake.com</a>	Site 8 Connection <a href="http://www.connection.com">www.connection.com</a>
Site 19 Kiabi <a href="http://www.Kiabi.pt">www.Kiabi.pt</a>	Site 16 Kinguin <a href="http://www.kinguin.net">www.kinguin.net</a>	Site 11 Allbirds <a href="http://www.allbirds.eu">www.allbirds.eu</a>
Site 23 Lings Cars <a href="http://www.lingscars.com">www.lingscars.com</a>	Site 20 Pure Vehicle Leasing <a href="http://www.purevehicleleasing.co.uk">www.purevehicleleasing.co.uk</a>	Site 21 WayFair <a href="http://www.wayfair.com">www.wayfair.com</a>

Figura 3.10: Grupo de sites da segunda experiência. Cada site possui um número único.

Um dos pontos foi que todos os participantes que foram selecionados para a experiência anterior não foram chamados para esta, a fim de eliminar qualquer efeito de aprendizagem que pudesse existir.

Mais uma vez, tal como na primeira experiência, o estudo foi conduzido num ambiente confortável, isento de ruídos, com iluminação excelente e totalmente livre de distrações. Nesta etapa do estudo, utilizamos um subconjunto dos sites usados no estudo anterior, de acordo com a seleção na Secção 3.2.2. Estes sites foram escolhidos por estarem posicionados nos extremos de um espaço bidimensional que considera a usabilidade e a experiência de utilização (ver Figura 3.7). Estes sites foram agrupados em três conjuntos, cada um contendo quatro sites (ver Figura 3.10). A novidade desta fase foi a redução do número de sites atribuídos a cada participante. Graças a esta alteração podemos realizar a experiência num único dia, evitando complicações devido ao grande número de participantes.

Antes de implementar a experiência com todos os utilizadores, realizaram-se testes piloto com 5 utilizadores para assegurar a adequação do procedimento. O feedback obtido durante os testes piloto foi utilizado para melhorar várias etapas do processo. Com base nesse feedback, foram efetuadas algumas alterações ao desenho inicial da experiência. Por exemplo, optou-se por utilizar dois computadores, um para a gravação dos sinais e outro para a realização das tarefas nos sites. Isto deveu-se ao facto de que, a utilização do mesmo computador distrai os utilizadores. Além disso, a verificação de dados tornava-se bastante mais complexa sem a adição de mais um computador. Com tudo preparado, ao iniciar cada experiência, fornecíamos uma explicação detalhada sobre o propósito da pesquisa, os seus objetivos e os procedimentos envolvidos. Além disso, apresentavamos o dispositivo Muse 2, utilizado para recolher dados fisiológicos durante a

interação com os sites. Explicámos aos utilizadores que não deveriam efetuar movimentos excessivos ou bruscos, uma vez que o dispositivo era sensível. Era fundamental que mantivessem a concentração e evitassem movimentos abruptos durante o processo. Assim como no estudo anterior, enfatizavamos que o nosso foco estava na recolha de dados e não na avaliação do desempenho dos participantes. Garantimos também que todas as informações seriam mantidas anónimas para proteger a privacidade dos utilizadores.

Não existiu limite de tempo e deixamos claro aos participantes que poderiam encerrar a participação a qualquer momento, caso assim desejassem. Depois da introdução, todos os participantes completaram um questionário demográfico. Este questionário continha a mesma informação que o questionário da experiência anterior, mas incluía uma pergunta adicional sobre se os utilizadores já tinham utilizado o dispositivo Muse2. No entanto, não foi realizada nenhuma análise entre esta pergunta e os resultados dos dados, uma vez que nenhum dos utilizadores tinha utilizado o dispositivo até então. Em seguida, solicitamos que identificassem qualquer familiaridade com os sites atribuídos. Se algum dos sites fosse reconhecido, substituímos por outro de um grupo diferente, mantendo o processo da pesquisa anterior. Após as etapas iniciais, conduzimos um teste, onde pedimos aos utilizadores para se manterem descontraídos para verificarmos a correta transmissão dos sinais fisiológicos pelo dispositivo Muse 2 e para os mesmos iniciarem os testes em estado neutro. Após a conclusão de cada tarefa, pedímos aos utilizadores que se descontraíssem ao máximo. Caso notássemos que os sinais ou o próprio utilizador indicavam que não estavam relaxados, aguardávamos um curto período de tempo até que voltassem a um estado neutro.

Depois deste passo concluído, cada participante realizou uma tarefa em cada um dos quatro sites designados. Durante a execução da tarefa, os dados fisiológicos dos participantes foram recolhidos. Após cada tarefa, os participantes responderam aos questionários standard mencionados na Secção 3.1.2. Este processo foi repetido para os quatro sites atribuídos a cada participante. Esta abordagem permitiu recolher os diversos dados dos sites analisados como por exemplo os dados sobre a experiência de utilização e a usabilidade. Com estes dados e os dados fisiológicos gravados podemos então produzir o *dataset*.

### 3.3.3 Dados Recolhidos

A estrutura da recolha de dados desta experiência foi segmentada em duas partes distintas: a obtenção dos dados através do dispositivo Muse 2, responsável pelos dados fisiológicos, e a recolha dos dados por meio dos questionários standard.

Iniciando com os dados provenientes dos questionários standard, adotámos a abordagem feita na recolha anterior para compreender e interpretar os resultados obtidos. Para esta finalidade, criámos um gráfico de dispersão que estabelece uma relação entre a média das pontuações do SUS de cada site e a média dos resultados do UEQ-S (ver Figura 3.11). Mais uma vez, como realizado anteriormente, dividimos o UEQ-S nas suas duas dimensões: "Qualidade Pragmática" (consultar Figura 3.12) e "Qualidade Hedónica" (ver Figura 3.13), permitindo uma análise mais aprofundada.

Uma observação relevante é a constatação de que a dispersão das características dos sites se manteve consistente. Em outras palavras, os sites que inicialmente demonstraram pontuações negativas tanto em usabilidade quanto em experiência de utilização mantiveram esses padrões nesta experiência, assim como aqueles posicionados no meio do gráfico e na secção positiva.

Além disso, conseguimos compreender que os utilizadores da segunda experiência partilharam a mesma opinião dos que realizaram a primeira experiência. Apesar de considerarem que o site estava muito aquém do que seria esperado tanto em termos de usabilidade quanto em termos de qualidade pragmática, devido a ser um site diferente do habitual, os utilizadores acharam a tarefa entusiasmante e interessante, aumentando os valores da qualidade hedónica.

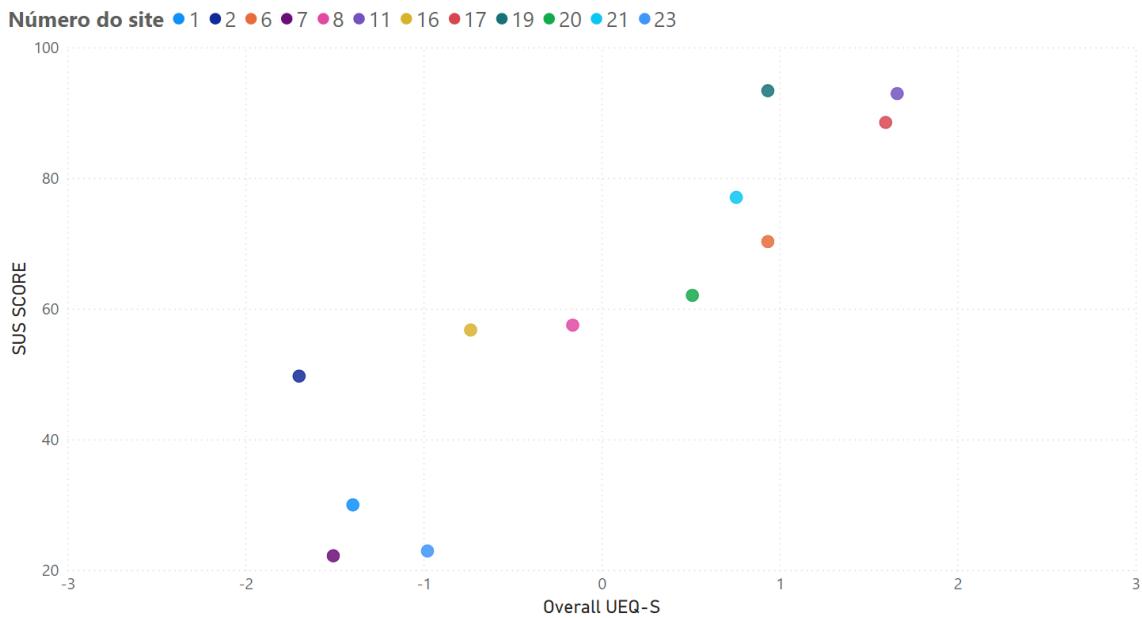


Figura 3.11: *Scatter Plot* Comparativo: Médias do UEQ-S com as Médias do SUS da Experiência 2.

Concluindo a análise, voltando agora a nossa atenção para os dados fisiológicos capturados por intermédio do dispositivo Muse 2, a recolha desses dados resultou na geração de quatro ficheiros CSV para cada tarefa executada num site. Cada um desses ficheiros, conforme detalhado na configuração do BlueMuse, contém os canais específicos de cada sinal, acompanhados pelo Timestamp. Através dos valores do Timestamp, fomos capazes de discernir que os sites que exhibiram, em média, tempos mais alargados estavam posicionados na secção do gráfico onde tanto a usabilidade quanto a experiência de utilização obtiveram avaliações negativas. Este padrão repete-se também para os casos opostos, onde as avaliações foram positivas.

### 3.4 *Dataset*

O *dataset* é o ponto fulcral do nosso estudo e para a produção do mesmo, com qualidade e com bastante detalhe foi necessário produzir um conjunto de passos específicos. Para este propósito, recorremos a uma base já estabelecida, inspirando-nos em *datasets* previamente reconhecidos e validados, como os conjuntos AMIGOS, DEAP, DREAMER e GAMEEMO, que já demonstraram resultados positivos.

A conceção final do nosso *dataset* foi moldada por uma abordagem iterativa, na qual a primeira etapa, como já referido anteriormente, envolveu uma experiência que forneceu os dados necessários para a seleção do conteúdo a ser incluído para a produção do nosso *dataset*. A partir de um conjunto inicial de 24 sites que foram submetidos a testes e onde foram recolhidos questionários, refinamos a seleção para os 12 que melhor se alinhavam com os nossos objetivos de pesquisa. Esta seleção foi fundamental para garantir que os sites escolhidos representassem adequadamente a variedade de dados que queríamos abranger no nosso *dataset*.

A próxima etapa envolveu a recolha dos dados para o nosso *dataset*. Durante esta fase, tanto os sinais fisiológicos quanto os dados dos questionários foram recolhidos dos participantes que interagiram com os sites selecionados. Todo este processo abrangeu os passos 1 a 7 que estão ilustrados na Figura 3.14.

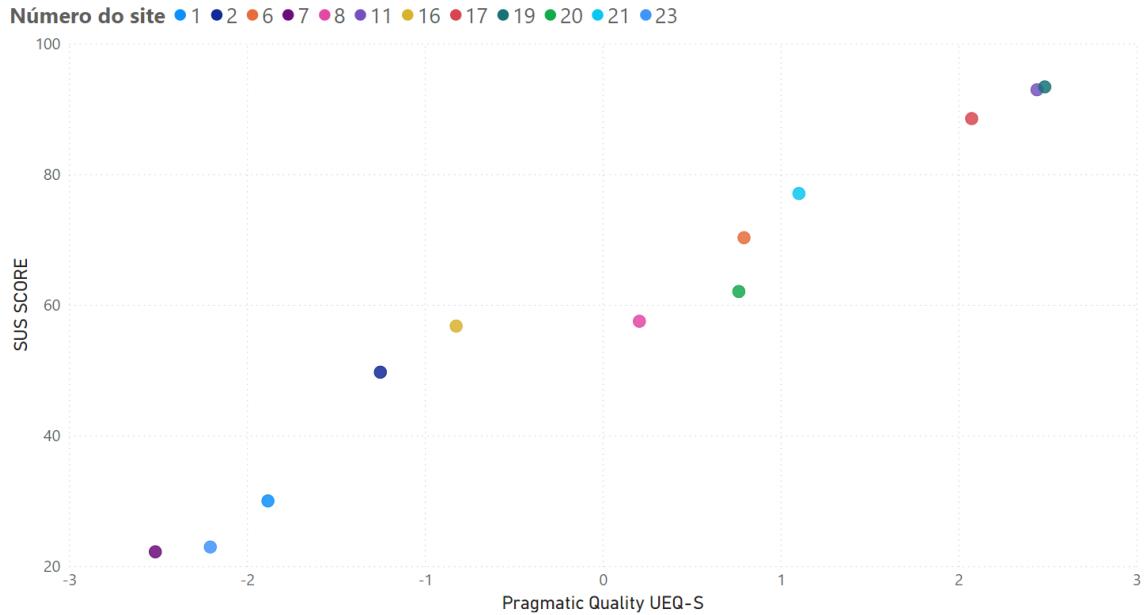


Figura 3.12: *Scatter Plot* Comparativo: Médias da qualidade pragmática com as Médias do SUS da Experiência 2.

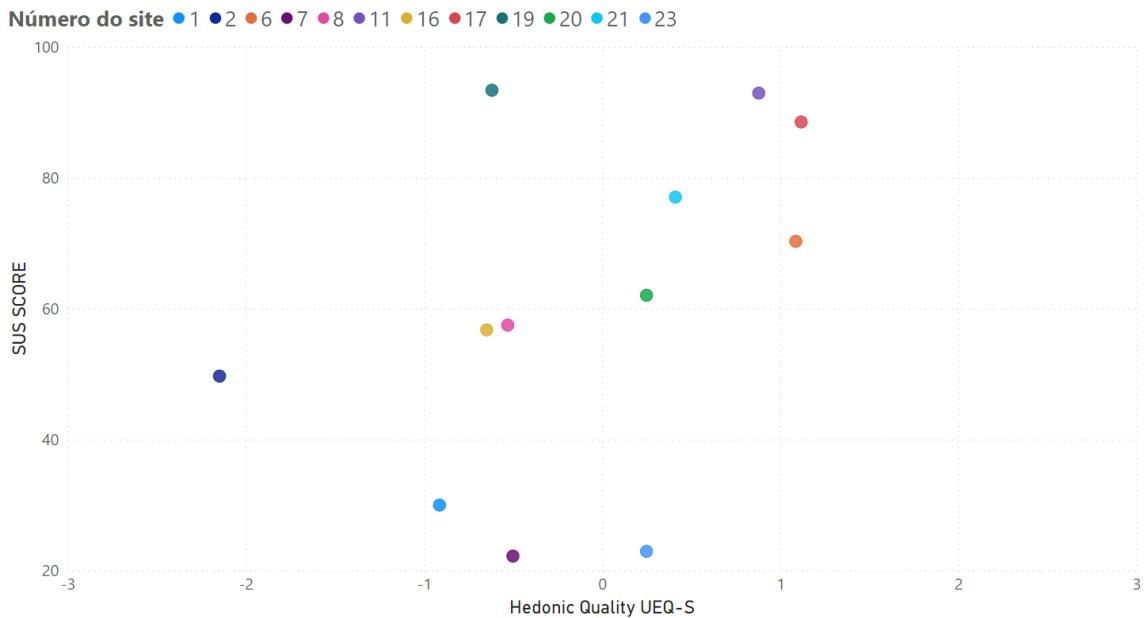
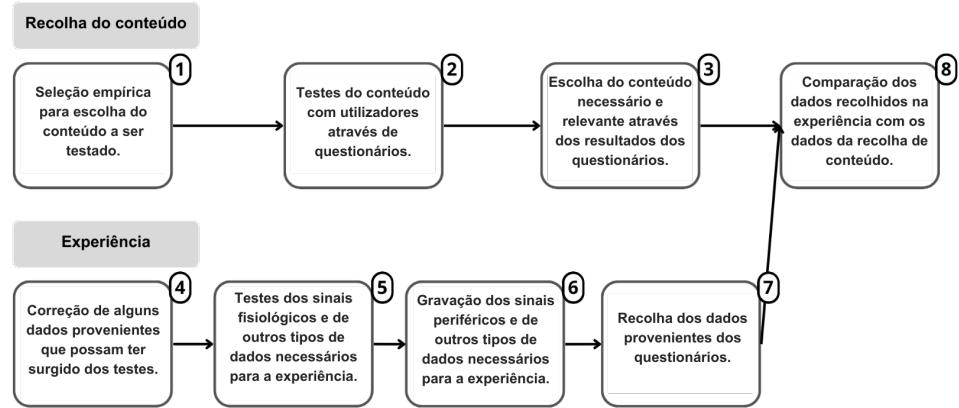
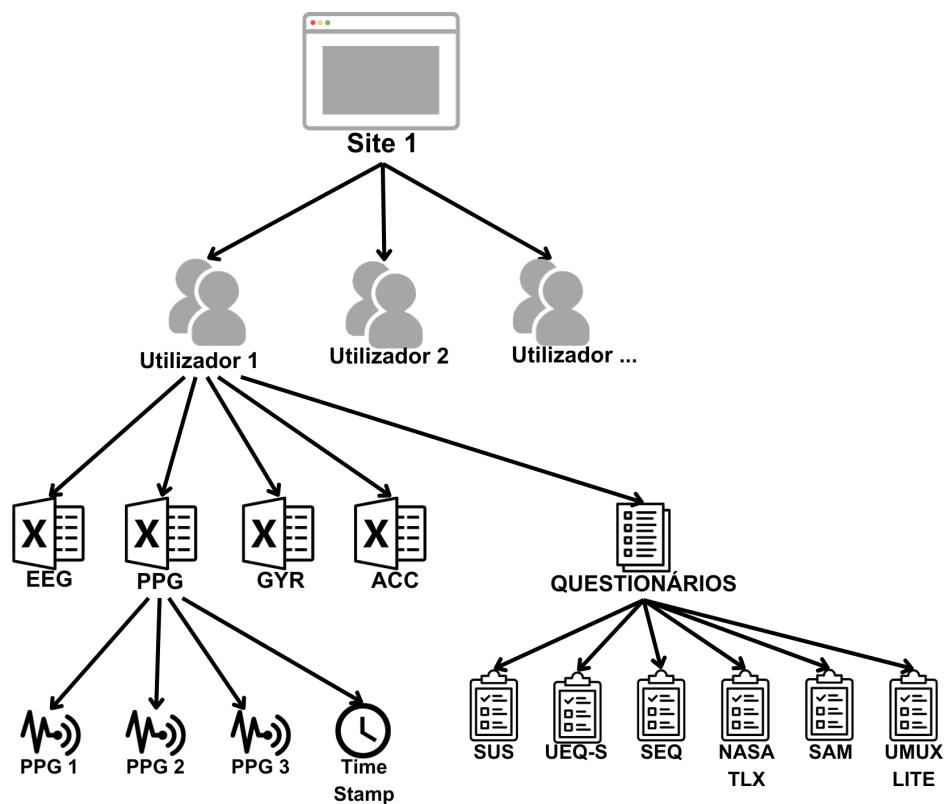


Figura 3.13: *Scatter Plot* Comparativo: Médias da qualidade hedônica com as Médias do SUS da Experiência 2.

Ao projetar a estrutura do nosso *dataset*, optamos por uma abordagem hierárquica para organizar os sinais fisiológicos. A Figura 3.15 que representa esta estrutura mostra que, no nível mais alto (nível 0), cada site é considerado uma entidade individual. Em seguida, no nível 1, estão os utilizadores que participaram em tarefas dentro do site em questão. Por fim, no nível 2, encontramos os quatro conjuntos de dados fisiológicos, junto com o Timestamp e as respostas dos questionários resultantes das atividades realizadas pelos utilizadores em cada site. O objetivo subjacente a esta organização é facilitar a análise dos resultados específicos de cada site, permitindo-nos examinar os padrões e as diferenças entre as experiências dos utilizadores em contextos diversos. Esta

Figura 3.14: Procedimento para a produção do *Dataset*.Figura 3.15: Abordagem hierárquica para organizar o *Dataset*.

abordagem estruturada e hierárquica não só assegura a consistência dos dados, como facilita a exploração das relações entre os diferentes níveis de interação dentro do *dataset* e ajuda a vida daqueles que utilizarão o *dataset* no futuro. Esta distribuição estruturada permite uma pesquisa mais eficiente e específica.

### 3.5 Síntese

O Capítulo 3 foi dedicado à criação do *dataset*. Este começa com uma introdução sobre a importância da criação de um *dataset* para este tipo de estudo e apresenta a metodologia utilizada para a construção do *dataset*, incluindo a aplicação interativa usada, a seleção dos questionários, a seleção dos participantes e os equipamentos utilizados para a recolha dos sinais fisiológicos.

Em seguida, o capítulo descreve as experiências realizadas para a recolha de dados, onde na primeira experiência, 20 participantes, realizaram tarefas em seis sites de compras selecionados de quatro grupos, evitando familiaridade prévia. Os participantes responderam a questionários após cada tarefa, incluindo um questionário demográfico que foi respondido antes de qualquer tarefa. Foram produzidos testes piloto que ajudaram a ajustar o procedimento, e algumas mudanças foram feitas. As tarefas foram divididas em dois dias para evitar fadiga, com um intervalo mínimo de cinco dias. O objetivo foi reduzir a lista de 24 sites para 12.

No segundo estudo, participaram 51 indivíduos, o procedimento foi semelhante ao estudo anterior. Um subconjunto de sites foi escolhido com base na primeira experiência. Cada participante avaliou quatro sites e em cada site foram recolhidos as respostas aos questionários e os sinais fisiológicos. Testes piloto com cinco participantes foram realizados para aprimorar o processo. Os participantes receberam explicações detalhadas e preencheram o questionário demográfico antes da experiência. O objetivo desta experiência foi recolher os dados necessários para a construção do *dataset*.

Por fim, o capítulo demonstra como foi feita a organização dos dados e destaca a importância da distribuição estruturada dos dados para facilitar a exploração dentro do *dataset*.

# Capítulo 4

## Validação do *Dataset*

Neste capítulo, realizámos uma análise cuidadosa da qualidade e consistência dos dados que recolhemos. Neste processo, começámos por descrever como foi efetuada a validação do *dataset*.

Em seguida, apresentámos os resultados obtidos durante o processo. Esta etapa foi crucial para avaliar a confiabilidade dos dados e identificar eventuais problemas que pudessem ter afetado a sua utilidade.

Por fim, procedemos à análise dos resultados, discutindo as implicações práticas das descobertas feitas durante a validação do *dataset*. Isto incluiu a identificação de potenciais melhorias e ajustes necessários para garantir que os dados fossem adequados ao propósito do estudo.

### 4.1 Descrição da validação

A etapa de validação do conjunto de dados corresponde ao passo oito ilustrado na Figura 3.14. Com o propósito de avaliar o nosso conjunto de dados, realizamos uma análise comparativa entre os dados obtidos na primeira e na segunda experiência. O objetivo era verificar se a informação recolhida dos utilizadores através dos questionários se mantinha similar. Na primeira experiência, participaram um total de 20 participantes, enquanto na segunda experiência, o número total de participantes foi de 51.

Importa salientar que esta comparação se circunscreverá exclusivamente aos questionários standard usados, não abrangendo os dados relacionados com os sinais fisiológicos, pois estes só foram recolhidos na segunda experiência.

Este processo de validação assume, assim, um papel fundamental no delineamento do nosso estudo, assegurando a integridade e a consistência dos resultados obtidos.

### 4.2 Resultados

Foi realizada uma análise estatística detalhada para comparar os dados dos questionários nas experiências 1 e 2. A escolha do teste estatístico desempenhou um papel crucial na interpretação dos resultados. Neste contexto, pelas duas amostras serem independentes e pelas variáveis serem ordinais tivemos de utilizar o teste *Mann-Whitney U*.

Inicialmente, apresentamos os resultados obtidos, incluindo o valor de p calculado no teste de *Mann-Whitney U* e as médias das observações do SUS e do UEQ-S em cada site usado na produção do *dataset*. Estes dados são detalhados na Figura 4.1, ordenados pelo número dado a cada site.

	System Usability Scale			User Experience Questionnaire Short		
	Média Exp 1	Média Exp 2	P-value	Média Exp 1	Média Exp 2	P-value
Site 1	31,50	30,00	0,87	-1,350	-1,397	0,75
Site 2	48,00	49,71	0,94	-1,730	-1,699	0,91
Site 6	68,50	70,29	0,75	1,000	0,934	0,61
Site 7	22,50	22,21	1,00	-1,400	-1,507	0,78
Site 8	56,00	57,50	0,91	-0,125	-0,162	0,97
Site 11	93,00	92,94	0,30	1,750	1,661	0,75
Site 16	55,50	56,77	1,00	-0,750	-0,736	0,75
Site 17	89,50	88,53	0,81	1,625	1,596	0,81
Site 19	94,50	93,38	0,97	1,150	1,200	0,81
Site 20	60,00	62,06	0,75	0,525	0,510	1,00
Site 21	79,00	77,05	0,61	0,800	0,757	0,21
Site 23	23,00	22,94	0,94	-1,100	-1,200	0,88

Figura 4.1: Resultados dos testes de *Mann-Whitney U* e médias do SUS e UEQ-S em cada site e em cada experiência.

O nosso primeiro objetivo era obter evidência estatística que confirmasse que não existia diferença estatisticamente significativa entre os dois grupos (Hipótese Nula). Para isso, fizemos o cálculo do valor de *p*. Neste caso, em todos os pontos da análise, verificamos que os resultados não atingem significância estatística, com  $p = 0,05$ .

Além das avaliações realizadas com os questionários SUS e UEQ-S, expandimos a nossa análise abrangendo todos os questionários standard presentes no nosso estudo, que incluíram o SEQ, o SAM, o NASA-TLX e o UMX-LITE.

Aplicamos o teste estatístico de *Mann-Whitney U* a cada um desses questionários, com o mesmo rigor e metodologia usados anteriormente. Os resultados destas análises encontram-se na Figura 4.2. Os resultados destes questionários também não alcançam significância estatística.

### 4.3 Análise dos Resultados e Discussão

Em todos os pontos da nossa análise, abrangendo os diferentes questionários e métricas, os resultados revelaram que não alcançaram significância estatística, com valores de  $p > 0,05$ . Isto indica que não foram encontradas diferenças estatisticamente significativas nas respostas dos utilizadores entre as duas experiências avaliadas.

Devido a este facto, podemos avançar para a próxima etapa da análise, que envolve a comparação dos diversos dados. Esta comparação é realizada através da análise de valores e de gráficos visuais, permitindo-nos visualizar a semelhança entre os dados dos questionários.

Ao recolher informações da Figura 4.1 e da Figura 4.3, conseguimos perceber que as médias dos vários questionários são bastante semelhantes entre a primeira e segunda experiência. Esta afirmação pode ser melhor visualizada na Figura 4.4. Além disso, observamos que em situações

	<i>P-Value</i>					
	NASA-TLX	UMUX-LITE	SEQ	SAM (Valence, Arousal, Dominance)		
				0,83	0,81	0,61
Site 1	0,46	0,90	0,80	0,83	0,81	0,61
Site 2	0,51	0,53	0,68	0,87	0,94	0,87
Site 6	0,88	0,50	0,93	0,66	0,97	0,97
Site 7	0,58	0,75	1,00	0,85	1,00	0,90
Site 8	0,91	0,39	0,97	0,97	0,97	0,69
Site 11	0,66	1,00	0,95	0,63	0,78	1,00
Site 16	0,72	0,84	0,83	0,52	0,87	0,94
Site 17	0,64	0,80	0,37	0,97	1,00	0,92
Site 19	0,78	0,73	0,74	0,96	0,78	0,87
Site 20	0,64	0,77	0,56	0,84	0,87	0,84
Site 21	0,67	0,50	0,39	0,84	0,87	0,93
Site 23	0,91	0,963	0,82	0,90	0,84	0,89

Figura 4.2: Resultados dos testes de *Mann-Whitney U* do SEQ, NASA-TLX, UMX-LITE e SAM em cada site e em cada experiência.

em que o SUS apresentou valores mais baixos, o UEQ-S, o SEQ e o UMX-LITE também registraram valores mais baixos. De maneira semelhante, o NASA-TLX seguiu a mesma tendência, apresentando altos índices de carga de trabalho quando estes questionários registraram valores mais baixos. O inverso também se verificou, mostrando resultados semelhantes.

Agora, ao analisar os valores médios do SAM, procuramos entender se conseguíamos relacionar estes dados com os valores médios do UEQ-S e do SUS. Estes dados estão representados na Figura 4.5, onde o eixo x representa a valência e o eixo y representa a excitação.

Dentro deste referencial, para tornar esta representação mais clara e informativa, utilizamos figuras para representar os valores do SUS e cores para refletir os resultados do UEQ-S.

As figuras que representam o SUS têm tamanhos diferentes, indicando a divisão dos valores em categorias específicas. Estas divisões foram retiradas do estudo [2]. O Rank F foi representado por figuras de pentágonos, e o tamanho do pentágono indicava a classificação dentro do Rank F, que podia variar entre F-, F e F+. Os outros "Ranks" seguiram a mesma lógica, com figuras diferentes para cada categoria: "Rank D" representado por losangos, "Rank C" por círculos, "Rank B" por retângulos e "Rank A" por triângulos.

As cores utilizadas para representar os resultados do UEQ-S variaram de acordo com os seus valores, indo de -3 a -2 em vermelho, de -2 a -1 em verde, de -1 a 0 em azul, de 0 a 1 em amarelo, de 1 a 2 em roxo e de 2 a 3 em laranja. Esta paleta de cores facilitou a visualização das avaliações dos sites em termos da experiência de utilização.

Esta representação ajudou a identificar tendências e padrões importantes em relação às respostas emocionais, à usabilidade percebida e à experiência de utilização dos participantes.

Assim torna-se mais fácil perceber que os sites que obtiveram melhores pontuações, como o site 11, site 17 e site 19, tanto no SUS quanto no UEQ-S, transmitiram sentimentos mais positivos

Média													
	NASA-TLX		UMUX-LITE		SEQ		SAM (Valence, Arousal, Dominance)						
	Exp 1	Exp 2	Exp 1	Exp 2	Exp 1	Exp 2	Exp 1	Exp 2	Exp 1	Exp 2	Exp 1	Exp 2	Exp 1
Site 1	63,83	60,83	17,50	17,65	1,80	1,76	1,80	1,94	6,40	6,24	1,60	1,88	
Site 2	45,50	40,88	27,50	33,09	2,60	2,71	3,40	3,29	5,00	5,10	4,00	3,94	
Site 6	39,17	36,23	62,50	61,76	3,60	3,70	5,80	5,18	5,00	4,90	6,80	6,35	
Site 7	65,00	62,65	7,50	5,15	1,60	1,59	1,40	1,53	7,00	6,94	1,60	1,76	
Site 8	37,00	35,54	47,50	53,68	3,20	3,35	5,20	5,12	4,40	3,94	5,40	5,94	
Site 11	10,17	11,91	95,00	95,59	5,00	4,82	7,80	7,65	7,00	6,76	8,60	8,71	
Site 16	42,17	41,42	40,00	44,12	3,00	3,18	4,60	4,12	4,20	4,06	6,00	5,47	
Site 17	19,17	17,16	87,50	88,97	4,40	4,53	8,20	8,12	6,40	6,47	8,60	8,35	
Site 19	11,83	12,55	95,00	93,38	5,00	4,88	8,40	8,06	6,80	6,65	9,00	8,82	
Site 20	29,50	33,43	67,50	64,71	3,40	3,47	5,80	5,53	4,60	4,41	6,80	6,53	
Site 21	27,20	28,43	72,50	75,00	4,00	4,10	5,60	5,94	5,00	4,88	7,40	7,35	
Site 23	68,83	67,50	5,00	6,18	1,60	1,53	1,60	1,82	7,20	7,18	1,60	1,47	

Figura 4.3: Resultados das médias do SEQ, NASA-TLX, UMUX-LITE e SAM em cada site e em cada experiência.



Figura 4.4: *Scatter Plots* Comparativos: Valores dos vários questionários da primeira e segunda experiência.

aos participantes. Por exemplo, os participantes ficaram encantados com estes sites. Por outro lado, ao analisar os sites que obtiveram as piores pontuações no SUS e no UEQ-S, como o site 1,

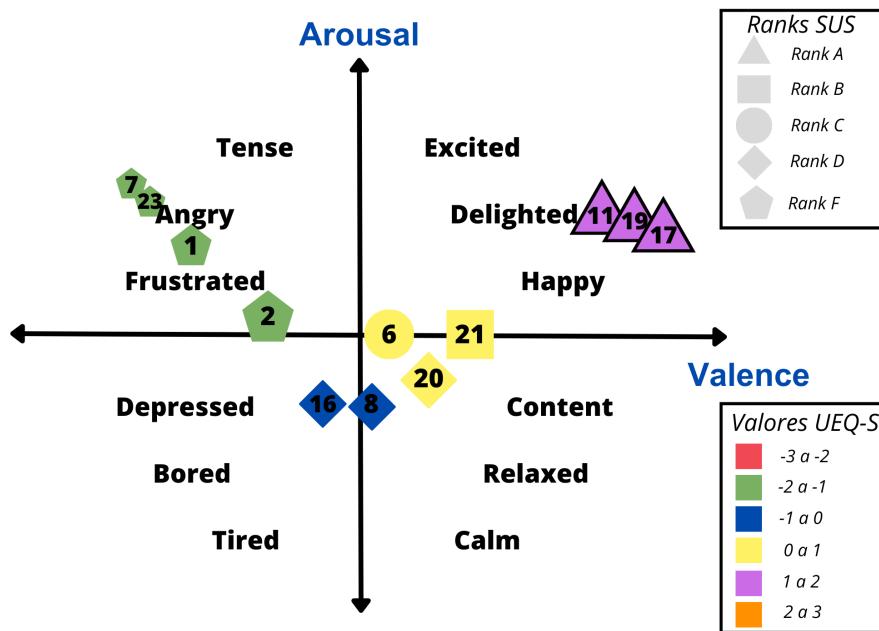


Figura 4.5: Referencial que representa os dados do SAM, SUS e UEQ-S. Os dados do SUS são representados por figuras geométricas e podem variar de tamanho conforme a categoria (exemplo: "F-", "F" e "F+"). Os dados do SAM são representados pelos valores no referencial, enquanto os do UEQ-S são identificados por cores.

site 7 e site 23, transmitiram emoções mais negativas, como frustração e raiva. Os demais sites estão mais próximos da zona neutra.

No caso do site 2, que possui um valor SUS mais alto do que os três mencionados anteriormente, mas proporcionou uma experiência de utilização bastante negativa, ele se encontra mais próximo da zona de frustração mas muito longe da zona de raiva. No caso do site 16, embora tenha obtido valores mais elevados no UEQ-S e no SUS em comparação com o site 2, esses valores ainda permanecem relativamente baixos.

Como resultado, o site 16 ainda se encontra mais próximo das emoções negativas do que das positivas, apesar de não ter gerado tanta frustração nos participantes. No entanto, ele também não conseguiu criar uma sensação de excitação entre os participantes. Este site, apesar de superar os do quadrante LVLA (Baixo Valence, Baixo Arousal), ainda não atinge o patamar de ser considerado com boa usabilidade e experiência de utilização. Todos os outros sites seguem mais ou menos a mesma tendência referida anteriormente.

Por fim, para obter uma visão clara da distribuição dos conjuntos de dados, procedemos à criação de *boxplots*. Estes *boxplots* estão divididos nas imagens da seguinte forma: SUS (Figura 4.6), UEQ-S e SEQ (Figura 4.7), e UMUX-LITE e NASA-TLX (Figura 4.8). Ao analisar os resultados destas avaliações, notamos uma tendência notável de distribuições semelhantes entre as duas experiências em estudo. Uma das principais indicações desta semelhança reside na proximidade das medianas em ambos os conjuntos de dados. As medianas, que representam a tendência central, foram observadas como sendo bastante próximas entre as duas experiências. Isto sugere que, em termos de tendência central, as avaliações realizadas nas duas experiências partilham características semelhantes. Além disso, é visível uma notável semelhança nos boxplots entre os dois conjuntos de dados para os vários questionários. Os boxplots indicam uma consistência na distribuição dos dados entre as experiências, reforçando a ideia de que as avaliações têm padrões semelhantes em ambas as situações.

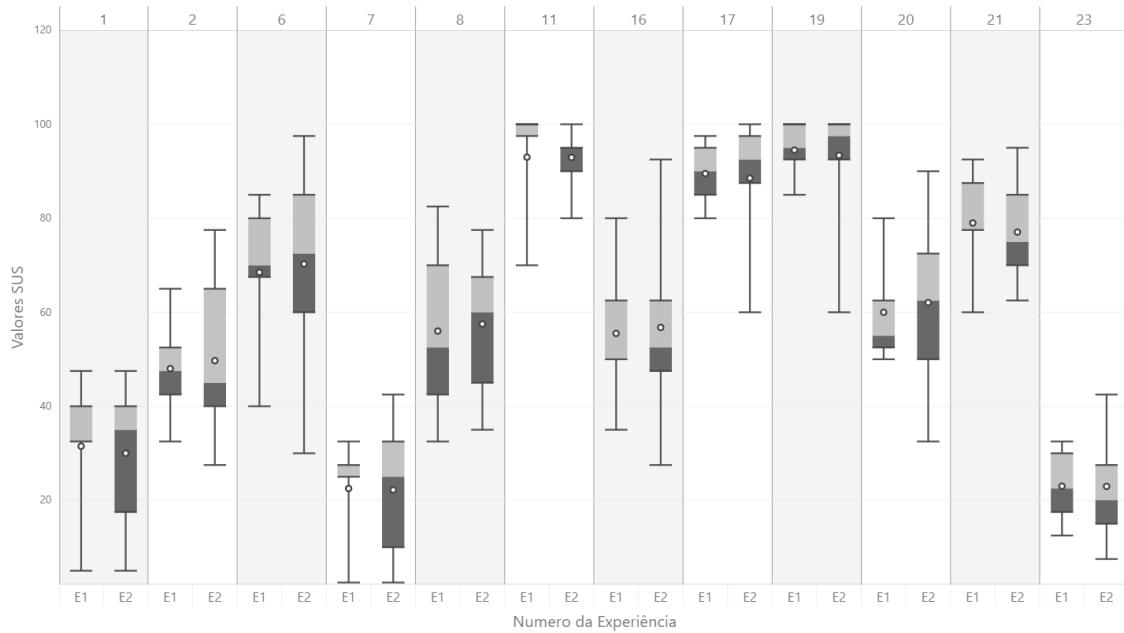


Figura 4.6: *Boxplot* com os valores do SUS da primeira e segunda experiência.

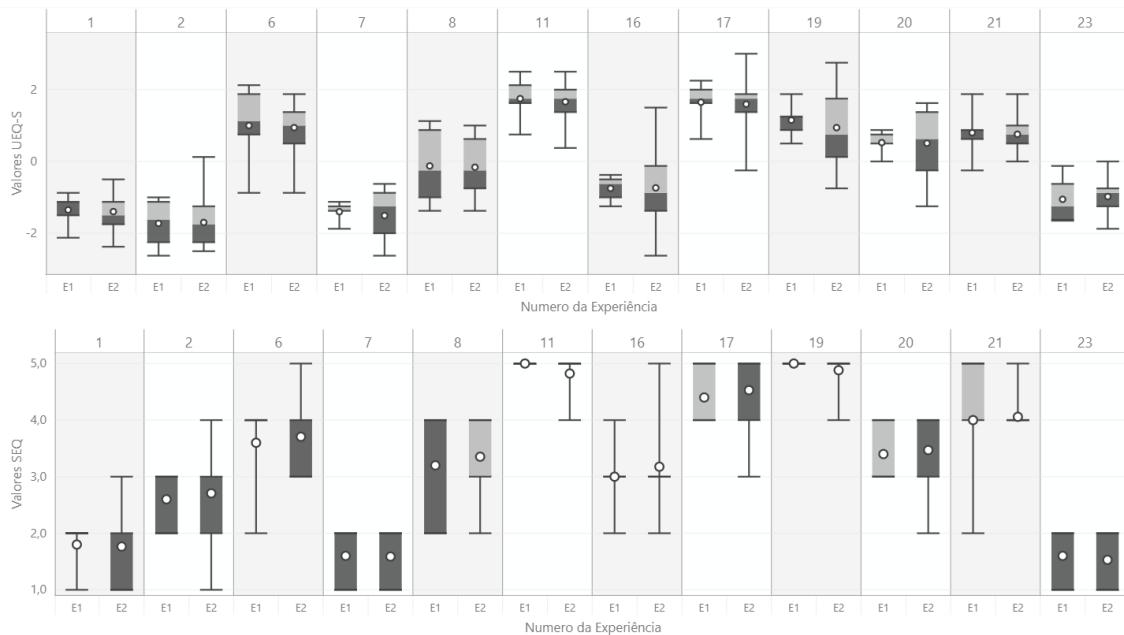


Figura 4.7: *Boxplots* com os valores do UEQ-S e SEQ da primeira e segunda experiência.

É importante notar que a segunda experiência incluiu um número significativamente maior de participantes por site em comparação com a primeira experiência. Esta diferença no tamanho da amostra pode explicar a presença de mais *outliers* nos dados da segunda experiência. Como seria de esperar, uma amostra maior pode revelar mais variações nos resultados. No entanto, é interessante notar que, apesar desta diferença, ainda encontramos alguns *outliers* que são comuns entre as duas experiências.

Estes resultados sugerem que, apesar das variações no tamanho da amostra entre as duas experiências, existem semelhanças notáveis nas avaliações dos cinco questionários. A consistência

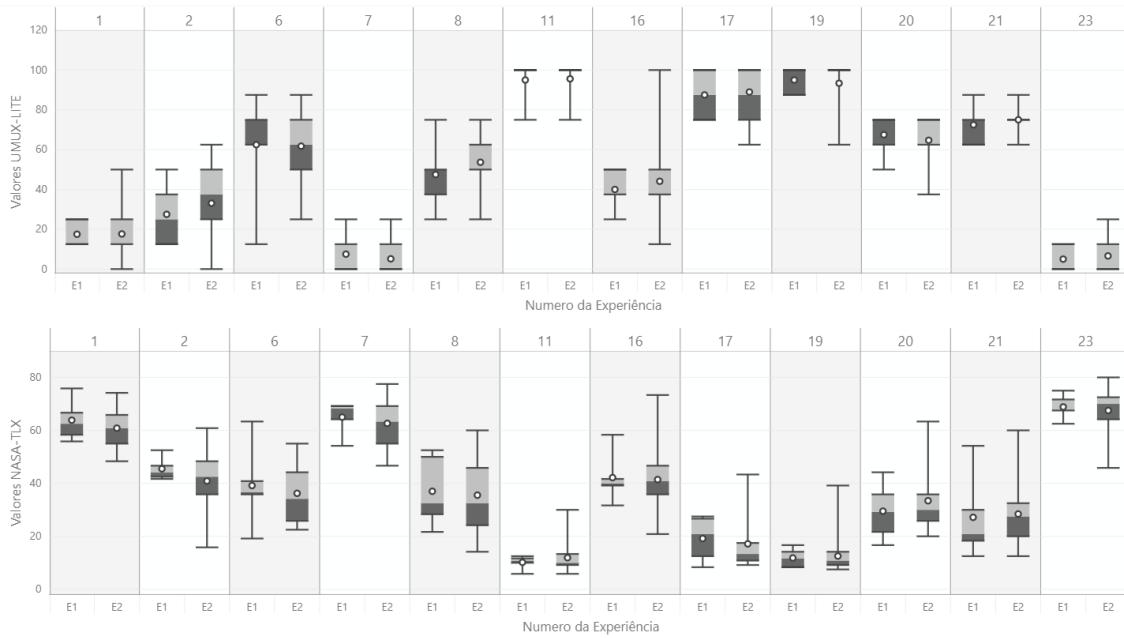


Figura 4.8: *Boxplots* com os valores do UMUX-LITE e NASA-TLX da primeira e segunda experiência.

nas medianas, a ocorrência de *outliers* partilhados e a semelhança nos boxplots indicam uma certa estabilidade nas respostas dos participantes, independentemente da variação no tamanho da amostra.

A clara semelhança nas respostas dos participantes (evidenciada pelos gráficos visuais previamente apresentados e pela proximidade dos valores, como por exemplo a média), bem como a ausência de diferenças estatisticamente significativas entre as experiências, demonstram que este conjunto de dados pode servir como uma base sólida para futuras pesquisas e aplicações relacionadas ao sistema ou produto em análise. A robustez das métricas de usabilidade e experiência de utilização utilizadas, juntamente com o tamanho da amostra recolhida, torna este conjunto de dados uma ferramenta valiosa para compreender a satisfação e o desempenho dos utilizadores.

## 4.4 Síntese

Este capítulo descreve o processo de validação do conjunto de dados recolhido para o estudo. A validação envolveu a análise dos dados recolhidos para garantir que os resultados obtidos eram bastante semelhantes.

O processo de validação incluiu o teste não paramétrico *Mann-Whitney U*, a comparação das médias entre os diversos questionários standard e a utilização de representações visuais para compreender a correlação entre eles. Todos estes dados mostraram-se consistentes entre si, indicando que o conjunto de dados era similar, e que não havia erros ou discrepâncias significativas.

A validação do conjunto de dados é um passo crucial antes de concluir o projeto, assegurando que os objetivos iniciais foram cumpridos e que o conjunto de dados tem qualidade.



## Capítulo 5

# Conclusões e Trabalho Futuro

Neste Capítulo, apresentamos um resumo da dissertação, as nossas conclusões finais e as contribuições do nosso trabalho. Também abordamos as novas questões que podem ser exploradas no futuro.

### 5.1 Resumo da dissertação

Neste trabalho, propusemos o Usabiliphy, um *dataset* composto por sinais fisiológicos e dados de questionários standard. No Capítulo 2, fazemos uma breve introdução sobre o tema dos questionários, explicando cada um e apontando as suas vantagens e desvantagens. Em seguida, o capítulo apresenta uma revisão teórica sobre os sinais fisiológicos, como os sinais EEG, PPG, ACC e GYR. O capítulo também discute as técnicas de recolha e os resultados obtidos pelos principais *datasets* que contêm sinais fisiológicos e dados de questionários standard. Além disso, aborda a importância da validação desses dados, um processo crítico para garantir a confiabilidade e a validade dos resultados. São apresentados diversos métodos de validação, essenciais para avaliar a qualidade dos dados recolhidos e garantir que sejam representativos e confiáveis para a análise estatística.

Por fim, apresentamos uma série de estudos que investigam a correlação entre sinais fisiológicos, como o EEG, e questionários standard ou questionários personalizados. Os estudos demonstram uma correlação significativa entre os dados do EEG e as respostas dos utilizadores aos questionários, sugerindo que os sinais fisiológicos podem ser indicadores valiosos da usabilidade, experiência de utilização, entre outros durante a interação com sistemas e dispositivos. Esta abordagem pode aprimorar a compreensão e o design de sistemas mais eficazmente.

No Capítulo 3, descrevemos em detalhe o processo utilizado para criar o *dataset*. Começamos por descrever a metodologia utilizada para criar o *dataset*, incluindo a seleção dos participantes, as aplicações interativas usadas, os questionários standard usados e o equipamento necessário. Descrevemos as duas experiências que foram realizadas para recolher os dados para o *dataset*. Na primeira experiência, um conjunto de tarefas foi realizado com 20 utilizadores, recolhendo os dados exclusivamente com questionários standard (SUS, UEQ-S, NASA-TLX, SEQ,SAM e UMUX-LITE). O objetivo desta experiência foi escolher o conteúdo a ser utilizado na segunda experiência.

Na segunda experiência, o processo foi o mesmo, mas além da recolha dos questionários, também houve a recolha dos dados fisiológicos (sinais EEG, PPG, GYR e ACC) a partir do dispositivo Muse2. Esta experiência foi realizada por 51 participantes e teve como objetivo recolher os dados para produzir o *dataset*. Sempre antes de cada uma das duas experiências, foram realizados testes piloto para a solução, a fim de definir alguns parâmetros e corrigir eventuais problemas.

Finalmente, este capítulo descreve os passos usados para produzir o *dataset* a partir dos dados recolhidos nas experiências, para uma pesquisa eficiente e prática.

No Capítulo 4, descrevemos como realizamos o processo de validação para o conjunto de dados recolhido. Esta validação foi realizada por meio de testes estatísticos nos resultados dos questionários entre a experiência 1 e 2. Através do teste não paramétrico *Mann-Whitney U*, percebemos que o conjunto de dados usado no *dataset* não possuia diferença estatisticamente significativa entre os dois grupos. Através de gráficos visuais e valores como a média, pudemos constatar que os dados se mantiveram bastante semelhantes entre as experiências. Além disso, pudemos notar que todos os questionários apresentam uma boa consistência nos resultados; quando um demonstra dados negativos, todos os outros refletem esta tendência, evidenciando uma convergência nos dados obtidos.

## 5.2 Contribuições e Limitações

Nesta secção, abordaremos as principais contribuições e limitações do presente estudo. Avaliar esses aspectos é fundamental para compreender a relevância dos nossos resultados e a solidez das conclusões.

Contribuições:

- **Base de Dados:** Destacamos que uma das principais contribuições do estudo é a criação de um *dataset*, estruturado e confiável que contém métricas de usabilidade, experiência de utilização, entre outras, assim como sinais fisiológicos.
- **Validação dos Dados:** O estudo incluiu etapas de validação dos dados, contribuindo para a confiabilidade e a validade dos resultados e aumentando a credibilidade do *dataset*.
- **Potencial para Uso Futuro:** Este *dataset* validado está disponível como um recurso para investigadores, designers e desenvolvedores interessados em utilizar o *dataset* e as métricas levantadas por meio dos diversos questionários standard.

Limitações:

- **Tamanho da Amostra:** Reconhecemos que o tamanho da amostra pode ser limitado, o que pode afetar a generalização dos resultados para uma população maior.
- **Contexto Específico:** Admitimos que os resultados são específicos para o contexto do estudo e podem não ser diretamente aplicáveis a outras situações ou produtos.
- **Permanente Evolução do Design:** Estamos cientes de que o design e o desenvolvimento de sistemas ou produtos estão em constante evolução, e os resultados podem não permanecer relevantes a longo prazo.

## 5.3 Trabalho Futuro

Como trabalho futuro, seria de grande relevância aproveitar a base de dados criada, que inclui tanto sinais fisiológicos como respostas a vários questionários standard. Este *dataset* foi validado, tornando-se uma ferramenta poderosa para análises futuras.

Uma das direções promissoras para investigações subsequentes seria o desenvolvimento de modelos capazes de prever a usabilidade com base nos sinais fisiológicos registados. Estes modelos poderiam complementar as avaliações convencionais por ser uma abordagem mais objetiva e detalhada.

Outra hipótese, seria o uso de machine learning e análise de dados para identificar padrões específicos nos sinais fisiológicos que estão correlacionados com a usabilidade. Isto abriria novas possibilidades para o design e aprimoramento de produtos e sistemas com base em evidências concretas sobre como eles afetam os utilizadores no nível fisiológico.

# Bibliografia

- [1] Talha Burak Alakus, Murat Gonen, and Ibrahim Türkoglu. Database for an emotion recognition system based on EEG signals and various computer games - GAMEEMO. *Biomed. Signal Process. Control.*, 60:101951, 2020.
- [2] Jonas Blattgerste, Jan Behrends, and Thies Pfeiffer. A web-based analysis toolkit for the system usability scale. In *PETRA '22: The 15th International Conference on PErvasive Technologies Related to Assistive Environments, Corfu, Greece, 29 June 2022 - 1 July 2022*, pages 237–246. ACM, 2022.
- [3] Katarzyna Blinowska and Piotr Durka. *Electroencephalography (EEG)*. John Wiley Sons, Ltd, 2006.
- [4] Teah-Marie Bynion and Matthew T. Feldner. *Self-Assessment Manikin*, pages 4654–4656. Springer International Publishing, Cham, 2020.
- [5] Sandra P. Cano, Nelson Araujo, Cristian Guzmán, Cristian Rusu, and Sergio Albiol-Pérez. Low-cost assessment of user experience through EEG signals. *IEEE Access*, 8:158475–158487, 2020.
- [6] Sandra P. Cano, Jonathan Soto, Laura Acosta, Victor Peñeyory, and Fernando Moreira. Electroencephalography as an alternative for evaluating user experience in interactive systems. In Álvaro Rocha, Hojjat Adeli, Gintautas Dzemyda, Fernando Moreira, and Ana Maria Rama-Iho Correia, editors, *Trends and Applications in Information Systems and Technologies - Volume 1, WorldCIST 2021, Terceira Island, Azores, Portugal, 30 March - 2 April, 2021*, volume 1365 of *Advances in Intelligent Systems and Computing*, pages 435–444. Springer, 2021.
- [7] Sandra P. Cano, Jonathan Soto, Laura Acosta, Victor M. Peñeyory, and Fernando Moreira. Using brain -computer interface to evaluate the user experience in interactive systems. *Comput. methods Biomech. Biomed. Eng. Imaging Vis.*, 11(3):378–386, 2023.
- [8] Yi Ding, Yaqin Cao, Qing-Xing Qu, and Vincent G. Duffy. An exploratory study using electroencephalography (EEG) to measure the smartphone user experience in the short term. *Int. J. Hum. Comput. Interact.*, 36(11):1008–1021, 2020.
- [9] Paul Ekman. Basic emotions. *Handbook of cognition and emotion*, 98(45-60):16, 1999.
- [10] Anastasiia Evmenenko and Diogo S. Teixeira. The circumplex model of affect in physical activity contexts: a systematic review. *International Journal of Sport and Exercise Psychology*, 20(1):168–201, 2022.
- [11] I Arun Faisal, T Waluyo Purboyo, and A Siswo Raharjo Ansori. A review of accelerometer sensor and gyroscope sensor in imu sensors on motion capture. *J. Eng. Appl. Sci.*, 15(3):826–829, 2019.

- [12] Kraig Finstad. The usability metric for user experience. *Interact. Comput.*, 22(5):323–327, 2010.
- [13] Jérémie Frey, Maxime Daniel, Julien Castet, Martin Hachet, and Fabien Lotte. Framework for electroencephalography-based evaluation of user experience. In Jofish Kaye, Allison Druin, Cliff Lampe, Dan Morris, and Juan Pablo Hourcade, editors, *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016*, pages 2283–2294. ACM, 2016.
- [14] Meiyazi Gao, Philip Kortum, and Frederick Oswald. Psychometric evaluation of the use (usefulness, satisfaction, and ease of use) questionnaire for reliability and validity. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 62, pages 1414–1418. SAGE Publications Sage CA: Los Angeles, CA, 2018.
- [15] Ben D Harper and Kent L Norman. Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. In *Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference*, volume 224, page 228. Citeseer, 1993.
- [16] Peter Hoonakker, Pascale Carayon, Ayse P Gurses, Roger Brown, Adjhaporn Khunlertkit, Kerry McGuire, and James M Walker. Measuring workload of icu nurses with a questionnaire survey: the nasa task load index (tlx). *IIE transactions on healthcare systems engineering*, 1(2):131–143, 2011.
- [17] Stamos Katsigiannis and Naeem Ramzan. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE J. Biomed. Health Informatics*, 22(1):98–107, 2018.
- [18] Brandy Klug. An overview of the system usability scale in library website and system usability testing. *Weave: Journal of Library User Experience*, 1(6), 2017.
- [19] Sander Koelstra, Christian Mühl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. DEAP: A database for emotion analysis using physiological signals. *IEEE Trans. Affect. Comput.*, 3(1):18–31, 2012.
- [20] J BlueMuse Kowaleski. Available online: <https://github.com/kowalej/BlueMuse> (accessed on 25 May 2022), 2019.
- [21] James R. Lewis. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM SIGCHI Bull.*, 23(1):78–81, 1991.
- [22] James R. Lewis, Brian S. Utesch, and Deborah E. Maher. Measuring perceived usability: The sus, umux-lite, and altusability. *Int. J. Hum. Comput. Interact.*, 31(8):496–505, 2015.
- [23] Nikola Marangunić and Andrina Granić. Technology acceptance model: a literature review from 1986 to 2013. *Universal access in the information society*, 14(1):81–95, 2015.
- [24] Ana Isabel Martins, Ana Filipa Rosa, Alexandra Queirós, Anabela Silva, and Nelson Pa-checo Rocha. European portuguese validation of the system usability scale (sus). *Procedia computer science*, 67:293–300, 2015.
- [25] Hitoshi Masaki, Masao Ohira, Hidetake Uwano, and Ken-ichi Matsumoto. A quantitative evaluation on the software use experience with electroencephalogram. In *International Conference of Design, User Experience, and Usability*, pages 469–477. Springer, 2011.

- [26] Juan Abdon Miranda-Correa, Mojtaba Khomami Abadi, Nicu Sebe, and Ioannis Patras. Amigos: A dataset for affect, personality and mood research on individuals and groups. *IEEE Transactions on Affective Computing*, 12(2):479–493, 2018.
- [27] Christina Orphanidou. Quality assessment for the photoplethysmogram (ppg). In *Signal Quality Assessment in Physiological Monitoring*, pages 41–63. Springer, 2018.
- [28] Robert Plutchik. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist*, 89(4):344–350, 2001.
- [29] Ana Filipa Rosa, Ana Isabel Martins, Victor Costa, Alexandra Queirós, Anabela Silva, and Nelson Pacheco Rocha. European portuguese validation of the post-study system usability questionnaire (pssuq). In *2015 10th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–5. IEEE, 2015.
- [30] Jeff Sauro and Joseph S Dumas. Comparison of three one-question, post-task usability questionnaires. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1599–1608, 2009.
- [31] Thomas Schmidt, Miriam Schlindwein, Katharina Lichtner, and Christian Wolff. Investigating the relationship between emotion recognition software and usability metrics. *i-com*, 19(2):139–151, 2020.
- [32] M Schrepp. User experience questionnaire handbook: All you need to know to apply the ueq successfully in your projects (2015). URL <http://www.ueq.online.org>, 2022.
- [33] Martin Schrepp. User experience questionnaire handbook. *All you need to know to apply the UEQ successfully in your project*, 2015.
- [34] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. Design and evaluation of a short version of the user experience questionnaire (ueq-s). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4 (6), 103-108., 2017.
- [35] Pekka Siirtola, Satu Tamminen, Gunjan Chandra, Anusha Ihalapathirana, and Juha Röning. Predicting emotion with biosignals: A comparison of classification and regression models for estimating valence and arousal level using wearable sensors. *Sensors*, 23(3):1598, 2023.
- [36] Priyanka Surendran et al. Technology acceptance model: A survey of literature. *International Journal of Business and Social Research*, 2(4):175–178, 2012.
- [37] Xiao Zhang, Yongqiang Lyu, Xin Hu, Ziyue Hu, Yuanchun Shi, and Hao Yin. Evaluating photoplethysmogram as a real-time cognitive load assessment during game playing. *International Journal of Human–Computer Interaction*, 34(8):695–706, 2018.



# Apêndice A

## Comissão de Ética

### A.0.1 Formulário de Pedido



Ciências  
ULisboa

Comissão de Ética de Ciências

#### **Lista de documentos para solicitação de parecer à Comissão de Ética de Ciências (CEC)**

V1, 31-05-2022

Para submeter um processo à CEC, o responsável do estudo deve submeter os seguintes documentos (disponíveis em <https://ciencias.ulisboa.pt/pt/comissao-etica-ciencias>):

1. Formulário (em inglês)
2. Protocolo do ou dos Estudos
3. Proposta de Folheto Informativo a ser explicado a cada participante
  - a. Está disponível um modelo-tipo, que explicita algumas das perguntas para as quais devem ser construídas explicações escritas.
4. Proposta de formulário de Consentimento Informado, a ser assinado por cada participante
  - a. Está disponível um modelo-tipo. É admissível a eliminação criteriosa de items, ou a inserção de novos items, em função das especificidades do estudo.
5. *Curriculum Vitae* dos membros da equipa
6. Outros documentos relevantes, caso existam
  - a. Por exemplo, requerimentos a, ou autorizações de, outras entidades, se aplicável.

Estes ficheiros devem ser submetidos por correio electrónico para o Secretariado da Comissão ([direccao@ciencias.ulisboa.pt](mailto:direccao@ciencias.ulisboa.pt)).

Os membros da Comissão estão obrigados ao dever de sigilo, e, com exceção do título do estudo e do conteúdo da Secção 13 (**Informação Pública**), nenhuma informação recebida dos proponentes será, de alguma forma, divulgada, tornada pública ou usada indevidamente.



Ciências  
ULisboa

## Comissão de Ética de Ciências

### APPLICATION FOR ETHICAL APPROVAL OF A RESEARCH PROJECT

This application form is to be used by STAFF seeking ethical approval for research projects/studies that involve human subjects.

Research must NOT begin until approval has been received from CEC.

### INDEX

SECTION 1. APPLICANT.....	3
SECTION 2. PROJECT.....	3
SECTION 3. TYPE OF PROJECT .....	3
SECTION 4. PROJECT DETAILS .....	4
4.1 - PROJECT OUTLINE & AIMS .....	4
4.2 - PROPOSED RESEARCH METHODS.....	4
SECTION 5. PARTICIPANT DETAILS.....	6
SECTION 6. PARTICIPANT INFORMATION .....	7
SECTION 7. PARTICIPANT CONSENT.....	7
SECTION 8. PARTICIPANT DEBRIEFING.....	8
SECTION 9. PROTECTION OF PERSONAL DATA OF PARTICIPANTS .....	8
SECTION 10. RISK CONSIDERATIONS.....	8
10.1 - POTENTIAL RISK TO PARTICIPANTS AND RISK MANAGEMENT PROCEDURES.....	8
10.2 - POTENTIAL RISK TO RESEARCHERS AND RISK MANAGEMENT PROCEDURES .....	8
SECTION 11. IDENTIFICATION JOINT ACTIVITIES WITH THE MEMBERS OF THE COMMISSION.....	8
SECTION 12. SUMMARY OF CRITICAL ISSUES .....	9
SECTION 13. PUBLIC INFORMATION .....	9
SECTION 14. DECLARATION .....	10



**Ciências  
ULisboa**

## Comissão de Ética de Ciências

### Section 1. Applicant

Name of Principal Researcher (Applicant):	Manuel J. Fonseca
Institution (if different from FCUL, FCIências.ID):	Click here to enter text.
Email address:	mjfONSECA@ciencias.ulisboa.pt
Contact Address:	Click here to enter text.
Telephone Number:	Click here to enter text.
Research Unit:	LASIGE
CENSUS number:	Click here to enter text.

### Section 2. Project

Project / Study Title (public):	UsabiliPhy: Estudo da relação entre a Usabilidade Percecionada e os Sinais Fisiológicos
Team members and institutions:	<ul style="list-style-type: none"> <li>• André Nilo Monteiro- FCUL</li> <li>• Manuel J. Fonseca – FCUL</li> </ul>

### Section 3. Type(s) of Study

<b>Questionnaire/Survey</b> e.g. surveys of members of particular groups / organizations; mail out questionnaires, street surveys	<input type="checkbox"/>
<b>Experiments</b> e.g. participants completing tasks under controlled conditions, use of tasks/method other than or in addition to questionnaires/surveys	<input checked="" type="checkbox"/>
<b>Observational</b> e.g. observing how people behave in a natural setting or in a laboratory	<input type="checkbox"/>
<b>Data-based</b> e.g. the use of official statistics where individuals could be identified	<input type="checkbox"/>
<b>Other</b>	<input type="checkbox"/>



**Ciências  
ULisboa**

## Comissão de Ética de Ciências

If 'Other', please describe.	Click here to enter text.
------------------------------	---------------------------

### Section 4. Project Details

Proposed date on which the project/study will begin <i>(assuming, if applicable, that funding has already been granted by a funding agency):</i>	As soon as approved
Proposed date on which the project/study will end:	26-07-2023

#### 4.1 - Project Outline & Aims

Briefly describe:

- The aims of this research
- The main tasks (or tests) that participants will be required to complete
- What use will be made of sensitive economic, social or personal data.

This description must be in everyday language, free from jargon, technical terms or discipline-specific phrases.

*(No more than 300 words)*

This research is composed of two experiments. The first aims at collecting data about the usability and user experience of a set of shopping websites. To do this, users will perform site tasks and answer standard usability questionnaires for each site.

The participants will be required to perform tasks on six shopping websites (from a set of 24), randomly selected from those they are not familiar with. After completing these tasks, the participants will be asked to respond to several questionnaires pertaining to their tasks.

The information gathered from this study will enable us to pinpoint the 12 websites that have a broader range concerning usability and user experience. In other words, sites nearer to the extremes of the two-dimensional space represented by the usability and user experience axes.

The second research experiment aims to collect data on the usability and user experience of the 12 shopping websites selected in the previous experiment and physiological data from users while performing tasks on these websites.

Participants will perform an experiment like the first one but with the addition of recording physiological signals using a Muse2 device.

The ultimate goal of this study is to have a dataset composed of the physiological signals and the data from standard questionnaires that allow studies to be conducted to see if there are relationships between both types of data.

#### 4.2 - Proposed Research Methods

Please provide an outline of the proposal research methods, in layman's terms, avoiding using jargon and technical terms as much as possible. Do include:

- Present an outline of the method in a step-by-step chronological order;
- All tasks that participants will be asked to complete;



- If the research will take place outside of Portugal or in collaboration with internationally-based partners, and/or if research will take place using the Internet;
- Where and how data will be collected and stored.

*(No more than 700 words)*

The first experiment will consist of a set of tasks that users will have to perform and then give feedback on through standard usability questionnaires.

The outline of the method step-by-step:

1. Inform the participant about the experiment and why we are doing it;
2. Ask participants to take a questionnaire to understand which websites they do not know or have not made any purchases;
3. Selection of 6 websites randomly from the above;
4. Perform the tasks on one of the websites from step 3);
5. Answer the standard questionnaires;
6. Take a 5-minute break;
7. Repeat step 4), 5) and 6).

The second experiment will be like the previous one except that we will add the recording of physiological signals, using the Muse2 device.

Before performing this experiment, we will make pilot tests to make sure the collection of physiological signals is working properly.

The outline of the method step-by-step:

1. Inform the participant about the experience and why we are doing it;
2. Ask participants to take a questionnaire to understand which websites they do not know or have not made any purchases;
3. Place the Muse2 device on the participant;
4. Test the collection of physiological signals and other data needed for the experiment;
5. Choose 4 websites randomly from step 2);
6. Perform the tasks on one of the websites from step 5);
7. Record the physiological signals of the user while performing the task;
8. Ask users to answer the standard questionnaires provided;
9. Take a 5-minute break;
10. Repeat step 6), 7), 8) and 9).



**Ciências  
ULisboa**

## Comissão de Ética de Ciências

### Section 5. Participant Details

Does this research specifically target (select all that apply):

Students or staff of this institution	<input type="checkbox"/>	
Adults (over the age of 18 years and competent to give consent)	<input checked="" type="checkbox"/>	
Children/legal minors (anyone under the age of 18 years)	<input type="checkbox"/>	
The elderly	<input type="checkbox"/>	
People with intellectual or communication difficulties	<input type="checkbox"/>	
People in custody	<input type="checkbox"/>	
People engaged in illegal activities (e.g., drug-taking)	<input type="checkbox"/>	
Number of participants:	Expected at least 20 for the first experiment and at least 50 for the second experiment	
Age from:	18	
Age to:	No limit	
Target populations to recruit participants, and means to select participants:	Click here to enter text.	
Reasons to select the required populations:	Click here to enter text.	
Does this project require approval from an external authority (e.g., CNPD, schools, governing body)?	<input type="checkbox"/>	<b>YES</b>
Has approval already been granted?	<input type="checkbox"/>	<b>NO</b>



## Section 6. Participant Information

YES      NO

Will you inform participants that their participation is <b>voluntary</b> ?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Will you inform participants that they may <b>withdraw</b> from the research at any time and for any reason?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Will you inform participants that their data will be treated with full <b>confidentiality</b> and that, if published, it will not be identifiable as theirs?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Will you provide an <b>information sheet</b> that will include the contact details of the researcher/team?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Will you obtain <b>written consent</b> for participation?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Will you debrief participants at the end of their participation (i.e., give them an explanation of the study and its aims and hypotheses)?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Will you provide participants with <b>written debriefing</b> (i.e., a sheet that they can keep that shows your contact details and explanations of the study)?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
If using a <b>questionnaire</b> , will you give participants the option of omitting questions that they do not want to answer?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
If an <b>experiment</b> , will you describe the main experimental procedures to participants in advance, so that they are informed about what to expect?	<input checked="" type="checkbox"/>	<input type="checkbox"/>
If the research is <b>observational</b> , will you ask participants for their consent to being observed?	<input type="checkbox"/>	<input type="checkbox"/>

## Section 7. Participant Consent

Please describe the arrangements you are making to inform participants, before providing consent, of what is involved in participating in your study:

We will give a brief description of the experiment and that any data collected will be done anonymously. In addition, we will stress that we are not evaluating them, and that any result is a good result.  
We will also explain that they can withdraw at any time for any reason they please, and that they do not need to give any explanation.

Please describe the arrangements you are making for participants to provide their full consent before data collection begins. Note that you can adapt, minimally, the template of the "Formulário de Consentimento Informado", to take into account the specificities of your studies:

We will use an adapted version of the "Informed Consent Form" when asking participants for their consent. This form will also be available, if necessary online, so that people can give their consent online.

Participants should be able to provide written consent. If you think gaining consent in this way is inappropriate for your project, then please explain how consent will be obtained and recorded.

**Section 8. Participant Debriefing**

Please describe the debriefing that participants will receive following the study and the exact point at which they will receive the debriefing:

Participants will be thanked for their participation and will be told what the future plans are. We also ensure that their participation is confidential, and that the data is completely anonymous.

It is a researcher's obligation to ensure that all participants are fully informed of the aims and methodology of the project, and to ensure that participants do not experience any levels of stress, discomfort, or unease following a research session. Also describe any particular provisions or debriefing procedures that will be in place to ensure participants feel respected and appreciated after they leave the study. Please attach the written debriefing sheet that you will give to participants. If you do not plan to provide a written debriefing sheet, please explain why.

**Section 9. Protection of personal data of participants**

Describe, in some detail, the types of personal data that will be requested to participants and how this data is going to be organized, protected, shared, and eventually backed-up. In particular, describe the anonymity procedures of the responsibility of the principal investigator ensuring that the members of the research team have no access to personal data which is irrelevant for research purposes. Address the use of internet or public / private information systems throughout the research. State how you are taking actions to follow the legal requirements related to GDPR.

*(No more than 300 words)*

The data from the experiment and from the questionnaire are all anonymous. Each participant will be anonymized through a unique id. No identifiable data will be stored/distributed online. It is expected that only the main investigator will access all the data. The information collected is personal and private.

**Section 10. Risk Considerations****10.1 - Potential risk to participants and risk management procedures**

Identify, as far as possible, all potential risks (small and large) to participants (e.g. physical, psychological, etc.) that are associated with the proposed research. Please explain any risk management procedures that will be put in place and attach any risk assessments or other supporting documents.

The only foreseeable risk is that participants may experience a bit of fatigue after the procedure. To minimize this problem, participants will be given a rest time for each task they complete.

**10.2 - Potential risk to researchers and risk management procedures**

What are the potential risks to researchers themselves? For example, personal safety issues such as lone or out of normal hours working or visiting participants in their homes; travel arrangements, including overseas travel; and working in unfamiliar environments. Please explain any risk management procedures that will be put in place and attach any risk assessments or other supporting documents.

No risks are foreseeable.

**Section 11. Identification joint activities with the members of the Commission**



Did you participate in common research projects or did you co-authored scientific papers with any of the members of the Comissão de Ética para a Recolha e Protecção de Dados Pessoais (See [Composição da Comissão](#)) in the last 24 months?

No.

#### Section 12. Summary of critical issues

If all the answers are NO, the Commission will follow a fast evaluation procedure.

		YES	NO
1	The project involves <a href="#">children</a> or other <a href="#">vulnerable groups</a> ?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
2	The project requires the co-operation of a <a href="#">gatekeeper</a> (defined as someone who can exert undue influence) for initial access to the groups or individuals to be recruited?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
3	Is it necessary for participants to take part in the project without their knowledge and consent e.g. covert observation of people in non-public places?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
4	The project includes deliberately misleading participants in any way?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
5	The project includes discussion of sensitive topics e.g. sexual activity or drug use?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
6	The project may cause psychological stress, anxiety, harm or negative consequences, beyond that encountered in normal life?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
7	The project requires prolonged or repetitive testing i.e. more than 4 hours commitment or attendance on more than two occasions?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
8	Are there financial inducements due to participants (other than <a href="#">reasonable expenses and compensation for time</a> )?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
9	The project causes pain or more than mild discomfort?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
10	The project collects and stores personal or sensitive data from participants?	<input type="checkbox"/>	<input checked="" type="checkbox"/>
11	The project plans to transfer participants' personal or sensitive data to other institutions somehow participating in the studies?	<input type="checkbox"/>	<input checked="" type="checkbox"/>

#### Section 13. Public information

Describe the study and the populations associated to it, for public release in the site of the Commission, enabling others to contact you in case of similar activities.  
(No more than 100 words)

Title: UsabiliPhy: Studying the relationship between Perceived Usability and Physiological Signals

The main goal of this research study is to collect data to create a database (dataset) consisting of physiological signals and responses to standard usability questionnaires. The ultimate goal is to have a



**Ciências  
ULisboa**

## Comissão de Ética de Ciências

dataset that will allow studies to be conducted to verify if there are relationships between both types of data collected.

The study will be carried out in two phases. In the first phase, we will identify a set of websites with good usability and user experience coverage. To do this, participants will perform website tasks and report their perception of usability and user experience through standard questionnaires.

During the second phase, we will utilize the websites identified in the initial phase to gather not only perceived usability and user experience data but also physiological signals from the participants as they carry out tasks on the websites.

### Section 14. Declaration

I certify that the information contained in this application is accurate. I have attempted to identify the risks that may arise in conducting this research and acknowledge my obligations and the rights of the participants.	
Name of Principal Investigator:	Manuel J. Fonseca
Signature:	
Date	02/03/2023

## A.1 Folheto Informativo Experiência 1



**PROJECTO: UsabiliPhy: Estudo da relação entre a Usabilidade Percecionada e os Sinais Fisiológicos**

**INVESTIGADOR RESPONSÁVEL: Manuel João da Fonseca**

Vimos desta forma convidá-la/o a participar no nosso estudo de investigação focado em recolher dados para criar uma base de dados (*dataset*) constituída por sinais fisiológicos e respostas a questionários standard de usabilidade. O objetivo final deste estudo é ter um *dataset* que permita a realização de estudos para verificar se existem relações entre ambos os tipos de dados recolhidos.

Antes de decidir, gostaríamos de lhe apresentar os detalhes desta investigação, a sua razão de ser, a sua utilidade potencial e as implicações da sua participação. Um membro da equipa de investigação irá acompanhá-lo na leitura deste folheto e responderá a quaisquer perguntas que queira fazer.

### 1 - Em que consiste o estudo “UsabiliPhy”?

Este estudo tem como objetivo recolher dados sobre a usabilidade e experiência de utilização de um conjunto de sites de compras. Para isso, os utilizadores irão realizar tarefas nos sites e responder a questionários standard de usabilidade para cada um dos sites. Esta informação sobre a usabilidade dos sites irá nos ajudar a escolher os sites a incluir no próximo estudo em que iremos recolher os sinais fisiológicos.

### 2 - Tenho de participar neste estudo?

A participação no estudo é totalmente voluntária. Vamos descrever o estudo e apresentar o conteúdo deste folheto informativo, incluindo os detalhes da sua participação. Se concordar em participar, irá assinar um Formulário de Consentimento.

### 3 - E se eu desejar desistir do estudo?

É livre de desistir, em qualquer altura, sem ter que fornecer quaisquer razões ou explicações.

### 4 - O que terei de fazer no âmbito do estudo?

No âmbito do estudo, irá realizar tarefas em seis sites de compras, que serão escolhidos aleatoriamente de entre os sites que não conhece e no final dessas tarefas irá responder a vários questionários com perguntas sobre as tarefas que realizou.

A sessão não tem tempo máximo, nem tempo mínimo, porém para não ser exaustivo, espera-se que dure entre 25 a 30 minutos.

O objetivo da investigação é recolher as respostas aos questionários standard, não tem como objetivo avaliá-la/o a si.

### 5 - Quais as desvantagens e riscos de participar?

Não estão previstos quaisquer riscos associados e a expectativa da equipa de investigação é de que as sessões em que participar sejam uma experiência agradável para os participantes.

O horário das sessões será combinado com cada participante mediante a sua disponibilidade.

### 6 - Quais os possíveis benefícios de participar?

**Folheto Informativo**

De acordo com a nossa experiência, as pessoas gostam de participar em estudos que promovem a comunicação com cientistas. O seu envolvimento irá ajudar no desenvolvimento de recursos essenciais para o próximo estudo de criação do dataset.

**7 - O que acontece quando o estudo terminar?**

A análise dos dados terminará quando for feito o estudo a pelo menos 20 utilizadores. Os resultados do estudo poderão ser publicados num artigo científico sobre o estudo da relação entre a usabilidade percecionada e os sinais fisiológicos. Se desejar saber detalhes sobre os resultados e implicações do estudo, poderemos enviar uma cópia do relatório do estudo, mas não antes de 30 de Setembro.

**8 - E se ocorrer algum problema?**

Se tiver alguma preocupação sobre qualquer aspecto deste estudo, deve falar com um dos investigadores responsáveis, Prof. Manuel J. Fonseca que fará o seu melhor para o elucidar e responder às suas dúvidas, por um dos emails abaixo referidos. Caso esteja descontente ou queira apresentar uma queixa formal, pode fazê-lo contactando o Diretor da Faculdade de Ciências da Universidade de Lisboa [email: [direccao@fc.ul.pt](mailto:direccao@fc.ul.pt)].

**9 - A minha informação será mantida confidencial?**

Sim. Seguiremos todas as práticas éticas e legais e toda a informação sobre si será tratada de forma absolutamente confidencial. Para garantir a anonimidade, os registos pessoais estarão apenas disponíveis na sua integralidade, para o investigador responsável, e os membros da equipa de investigação apenas terão acesso aos dados que necessitarem de conhecer. Se os seus dados forem usados para publicações ou apresentações, serão totalmente anónimizados, sem qualquer referência, direta ou indireta, à sua identidade. Se forem tiradas fotografias, e for nossa intenção usá-las em alguma apresentação, ser-lhe-á pedida autorização prévia. Se estiver disponível para que usemos fotografias ou vídeos para esse propósito, pedir-lhe-emos primeiro que assine autorizações específicas com esse objetivo.

**10 - O estudo passou por um processo de revisão?**

Sim. Com efeito, este estudo foi revisto pela Comissão de Ética de Ciências (<https://ciencias.ulisboa.pt/pt/comissao-etica-ciencias>). Esta comissão analisou a proposta de estudo, bem como todos os seus materiais e não levantou objeções do ponto de vista ético.

**11 - Quem posso contactar relacionado com este estudo?**

<b>Prof. Manuel J. Fonseca</b> Faculdade de Ciências da Universidade de Lisboa <a href="mailto:mifonseca@fc.ul.pt">mifonseca@fc.ul.pt</a>	<b>André Monteiro</b> Faculdade de Ciências da Universidade de Lisboa <a href="mailto:fc51718@alunos.fc.ul.pt">fc51718@alunos.fc.ul.pt</a>
--	---

Sinta-se à vontade para os contactar em qualquer matéria relacionada com este estudo.

**Muito Obrigado.**

Obrigado por nos ter dedicado este seu tempo e por considerar participar neste estudo.

O Investigador Responsável

\*\*\*\*

## A.2 Folheto Informativo Experiência 2



**PROJECTO: UsabiliPhy: Estudo da relação entre a Usabilidade Percecionada e os Sinais Fisiológicos**

**INVESTIGADOR RESPONSÁVEL: Manuel João da Fonseca**

Vimos desta forma convidá-la/o a participar no nosso estudo de investigação focado em recolher dados para criar uma base de dados (*dataset*) constituída por sinais fisiológicos e respostas a questionários standard de usabilidade. O objetivo final deste estudo é ter um *dataset* que permita a realização de estudos para verificar se existem relações entre ambos os tipos de dados recolhidos.

Antes de decidir, gostaríamos de lhe apresentar os detalhes desta investigação, a sua razão de ser, a sua utilidade potencial e as implicações da sua participação. Um membro da equipa da investigação irá acompanhá-lo na leitura deste folheto e responderá a quaisquer perguntas que queira fazer.

### 1 - Em que consiste o estudo “UsabiliPhy”?

Este estudo tem como objetivo recolher dados sobre a usabilidade e experiência de utilização de um conjunto de sites de compras e dados fisiológicos dos utilizadores enquanto realizam tarefas nesses sites. Para isso, os utilizadores irão realizar tarefas nos sites e responder a questionários standard de usabilidade para cada um dos sites. Adicionalmente, iremos recolher dados fisiológicos dos utilizadores, usando um dispositivo Muse2, enquanto estes realizam as tarefas nos sites.

### 2 - Tenho de participar neste estudo?

A participação no estudo é totalmente voluntária. Vamos descrever o estudo e apresentar o conteúdo deste folheto informativo, incluindo os detalhes da sua participação. Se concordar em participar, irá assinar um Formulário de Consentimento.

### 3 - E se eu desejar desistir do estudo?

É livre de desistir, em qualquer altura, sem ter que fornecer quaisquer razões ou explicações.

### 4 - O que terei de fazer no âmbito do estudo?

No âmbito do estudo, irá realizar tarefas em quatro sites de compras, que serão escolhidos aleatoriamente de entre os sites que não conhece, enquanto recolhemos os seus sinais fisiológicos usando um dispositivo Muse2. No final dessas tarefas irá responder a vários questionários com perguntas sobre as tarefas que realizou.

A sessão não tem tempo máximo, nem tempo mínimo, porém para não ser exaustivo, espera-se que dure entre 25 a 30 minutos.

O objetivo da investigação é recolher os sinais fisiológicos e as respostas aos questionários standard, não tem como objetivo avaliá-la/o a si.

### 5 - Quais as desvantagens e riscos de participar?

Não estão previstos quaisquer riscos associados e a expectativa da equipa de investigação é de que as sessões em que participar sejam uma experiência agradável para os participantes.

O horário das sessões será combinado com cada participante mediante a sua disponibilidade.

### 6 - Quais os possíveis benefícios de participar?



De acordo com a nossa experiência, as pessoas gostam de participar em estudos que promovem a comunicação com cientistas. O seu envolvimento irá ajudar no desenvolvimento de um dataset que vai conter não só dados fisiológicos como os resultados dos questionários standard e de usabilidade.

**7 - O que acontece quando o estudo terminar?**

A análise dos dados terminará quando for feito o estudo a pelo menos 50 utilizadores. Os resultados do estudo poderão ser publicados num artigo científico sobre o estudo da relação entre a usabilidade percecionada e os sinais fisiológicos. Se desejar saber detalhes sobre os resultados e implicações do estudo, poderemos enviar uma cópia do relatório do estudo, mas não antes de 30 de Setembro.

**8 - E se ocorrer algum problema?**

Se tiver alguma preocupação sobre qualquer aspecto deste estudo, deve falar com um dos investigadores responsáveis, Prof. Manuel J. Fonseca que fará o seu melhor para o elucidar e responder às suas dúvidas, por um dos emails abaixo referidos. Caso esteja descontente ou queira apresentar uma queixa formal, pode fazê-lo contactando o Diretor da Faculdade de Ciências da Universidade de Lisboa [email: [direccao@fc.ul.pt](mailto:direccao@fc.ul.pt)].

**9 - A minha informação será mantida confidencial?**

Sim. Seguiremos todas as práticas éticas e legais e toda a informação sobre si será tratada de forma absolutamente confidencial. Para garantir a anonimidade, os registos pessoais estarão apenas disponíveis na sua integralidade, para o investigador responsável, e os membros da equipa de investigação apenas terão acesso aos dados que necessitarem de conhecer. Se os seus dados forem usados para publicações ou apresentações, serão totalmente anônimizados, sem qualquer referência, direta ou indireta, à sua identidade. Se forem tiradas fotografias, e for nossa intenção usá-las em alguma apresentação, ser-lhe-á pedida autorização prévia. Se estiver disponível para que usemos fotografias ou vídeos para esse propósito, pedir-lhe-emos primeiro que assine autorizações específicas com esse objetivo.

**10 - O estudo passou por um processo de revisão?**

Sim. Com efeito, este estudo foi revisto pela Comissão de Ética de Ciências (<https://ciencias.ulisboa.pt/pt/comissao-etica-ciencias>). Esta comissão analisou a proposta de estudo, bem como todos os seus materiais e não levantou objeções do ponto de vista ético.

**11 - Quem posso contactar relacionado com este estudo?**

<b>Prof. Manuel J. Fonseca</b> Faculdade de Ciências da Universidade de Lisboa <a href="mailto:mifonseca@fc.ul.pt">mifonseca@fc.ul.pt</a>	<b>André Monteiro</b> Faculdade de Ciências da Universidade de Lisboa <a href="mailto:fc51718@alunos.fc.ul.pt">fc51718@alunos.fc.ul.pt</a>
--	---

Sinta-se à vontade para os contactar em qualquer matéria relacionada com este estudo.

**Muito Obrigado.**

Obrigado por nos ter dedicado este seu tempo e por considerar participar neste estudo.

O Investigador Responsável

\*\*\*\*

### A.3 Procedimento do Trabalho

## Protocolo dos Estudos

### Sistema

Os testes vão-se realizar num computador portátil (CPU 3,30GHz, 8GB RAM) a correr em Windows. O dispositivo utilizado para recolher os sinais fisiológicos será o Muse2.

### Experiência 1: Seleção do conteúdo

Esta experiência irá envolver um total de 20 participantes, de preferência o mesmo número de participantes do sexo feminino e masculino.

Os 24 sites de compras a usar no estudo foram escolhidos empiricamente de acordo com a sua usabilidade e experiência de utilização. A escolha foi feita de modo a estarem distribuídos pelos quatro quadrantes de um espaço bidimensional em que temos num eixo a usabilidade (dada pelo *System Usability Scale (SUS)*) e no outro a experiência de utilização (dada pelo *User Experience Questionnaire (UEQ)*).

Nesta experiência, teremos duas sessões por participante. As sessões serão realizadas em dias diferentes, para evitar os efeitos da aprendizagem, e em cada sessão o participante realizará um conjunto de tarefas em 3 sites de compras.

O primeiro passo da experiência será explicar a cada participante qual será a experiência, os seus objetivos, os vários passos que a compõem, que o que pretendemos é recolher os dados e não avaliar o desempenho dos participantes, e que todos os dados serão recolhidos de forma anónima.

Em seguida, faremos uma breve explicação de como irá funcionar a experiência. Esta não terá um tempo limite para cada sessão e iremos avisar os participantes que poderão desistir em qualquer momento.

Depois de todas as dúvidas serem esclarecidas, o utilizador irá preencher um questionário inicial, para identificar de entre os 24 sites fornecidos os que lhe são familiares. Com base nas respostas iremos selecionar de maneira aleatória 6 sites (3 para cada sessão), de entre os não familiares, para o mesmo realizar as tarefas.

A seguir, para cada um dos sites o participante faz os seguintes passos: i) realiza as tarefas atribuídas ao site; ii) responde aos seguintes questionários standard:

- System Usability Scale (SUS)
- User Experience Questionnaire – Short (UEQ-S)
- NASA Task Load Index (TLX)
- Single Ease Question (SEQ)
- Self-Assessment Manikin (SAM)

O participante terá tempo para descansar entre cada um dos sites. Depois de terminar os três sites o participante irá responder a um questionário sobre a sua informação demográfica. No final, iremos agradecer ao utilizador a sua participação.

## Experiência 2: Recolha de sinais fisiológicos

Esta experiência irá envolver um total de 50 participantes, de preferência o mesmo número de participantes do sexo feminino e masculino, e realizar-se-á de maneira bastante similar á Experiência 1. Estes serão distintos dos que participarem na Experiência 1, para evitar o efeito da aprendizagem, pois os sites serão um subconjunto dos usados na Experiência 1. Todo o processo que se descreve a seguir, será validade através de testes piloto com no mínimo 2 pessoas.

Nesta experiência iremos usar os 12 sites que tenham uma maior abrangência ao nível da usabilidade e experiência de utilização. Isto é, que estejam mais perto dos extremos do espaço bidimensional definido pelos eixos usabilidade e experiência de utilização. Para isso, usaremos os dados recolhidos na Experiência 1.

Começaremos por explicar a cada participante qual será a experiência, os seus objetivos, os vários passos que a compõem, que o que pretendemos é recolher os dados e não avaliar o desempenho dos participantes, e que todos os dados serão recolhidos de forma anónima.

Em seguida, faremos uma breve explicação de como funciona esta experiência com a adição do funcionamento do Muse2. Não existirá limite de tempo para cada sessão e iremos avisar os participantes que podem desistir em qualquer altura.

Depois de todas as dúvidas serem esclarecidas, o utilizador irá preencher um questionário inicial, para identificar de entre os 12 sites fornecidos os que lhe são familiares. Com base nas respostas iremos selecionar de maneira aleatória 4 sites, de entre os não familiares, para o mesmo realizar as tarefas. A seguir testaremos a recolha dos sinais fisiológicos e outros dados necessários para a experiência.

A seguir, para cada um dos sites o participante faz os seguintes passos: i) realiza as tarefas atribuídas ao site (enquanto realiza a tarefa os seus dados fisiológicos vão ser recolhidos através do dispositivo Muse2); ii) responde aos seguintes questionários standard:

- System Usability Scale (SUS)
- User Experience Questionnaire – Short (UEQ-S)
- NASA Task Load Index (TLX)
- Single Ease Question (SEQ)
- Self-Assessment Manikin (SAM)

O participante terá tempo para descansar entre cada um dos sites. Depois de terminar os 4 sites o participante irá responder a um questionário sobre a sua informação demográfica. No final, iremos agradecer ao utilizador a sua participação.

### Resumo das experiências

O esboço da primeira experiência passo a passo:

1. Informar o participante sobre a experiência e a razão pela qual a estamos a fazer;
2. Fazer um questionário para compreender em que sites os utilizadores não sabem ou não fizeram qualquer compra;
3. Escolher 6 sites web aleatoriamente dos anteriores;
4. Realizar as tarefas num dos sites web da etapa 3);
5. Responder aos questionários standard;
6. Fazer uma pausa de 5 minutos;
7. Repetir os passos 4), 5) e 6).

O esboço da segunda experiência passo a passo:

1. Informar o participante sobre a experiência e a razão pela qual a estamos a fazer;

2. Fazer um questionário para compreender em que sites os utilizadores não sabem ou não fizeram qualquer compra;
3. Colocar o dispositivo Muse2 no participante;
4. Testar a recolha dos sinais fisiológicos e outros dados necessários para a experiência;
5. Escolher 4 sites web aleatoriamente dos anteriores;
6. Realizar as tarefas num dos sites web da etapa 5);
7. Registar os sinais fisiológicos do utilizador durante a realização da tarefa;
8. Responder aos questionários standard fornecidos;
9. Fazer uma pausa de 5 minutos;
10. Repetir os passos 6), 7), 8) e 9).



## Apêndice B

# Questionários Experiência 2

### B.1 Questionário Demográfico

20/09/23, 22:44

Questionário Demográfico

#### Questionário Demográfico

\* Indica uma pergunta obrigatória

1. Número de utilizador? \*

---

2. Confirmo que concordo com todos os pontos do Formulário de Consentimento Informado. \*

Marcar apenas uma oval.

- Sim  
 Não

3. Qual é a sua idade? \*

---

4. Qual é o seu género? \*

Marcar apenas uma oval.

- Masculino  
 Feminino  
 Outros  
 Prefere não responder

20/09/23, 22:44

Questionário Demográfico

5. **Educação \****Marcar apenas uma oval.*

- Até ao ensino secundário
- Bacharelado
- Mestrado ou superior

## 6. Alguma vez utilizou o dispositivo Muse 2? \*

*Marcar apenas uma oval.*

- Sim
- Não

---

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários

## B.2 Questionário para Cada Tarefa

29/09/23, 18:04

Questionários

### Questionários

\* Indica uma pergunta obrigatória

1. Número de utilizador? \*

---

2. Número do site? \*

---

#### Single Ease Question Questionnaire

3. Em geral, quão difícil ou fácil foi a tarefa de concluir? \*

Marcar apenas uma oval.

1 2 3 4 5

---

Muil      Muito fácil

#### Usability Metric for User Experience - Lite

4. Este site é fácil de utilizar. \*

Marcar apenas uma oval.

1 2 3 4 5

---

Disc      Concorda fortemente

29/09/23, 18:04

Questionários

5. **Os recursos deste site atendem os meus requisitos. \****Marcar apenas uma oval.*

1 2 3 4 5

Disc      Concorda fortemente**User Experience Questionnaire- Short**

A fim de avaliar o produto, por favor preencha o seguinte questionário. É constituído por pares de opostos relativos às propriedades que o produto possa ter.

As graduações entre os opostos são representadas por círculos. Ao marcar um dos círculos, você pode expressar sua opinião sobre um conceito.

**Exemplo:**

Atraente	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	Feio				
----------	-----------------------	----------------------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	------

Esta resposta significa que avalia o produto mais atraente do que feio.

Marque a sua resposta da forma mais espontânea possível. É importante que não pense demasiado na resposta porque a sua avaliação imediata é que é importante. Por favor, assinale sempre uma resposta, mesmo que não tenha certezas sobre um par de termos ou que os termos não se enquadrem com o produto.

Não há respostas "certas" ou respostas "erradas". A sua opinião pessoal é que conta! Por favor, dê-nos a sua avaliação atual do produto em causa. Por favor, marque apenas um círculo por linha.

## 6. \*

*Marcar apenas uma oval.*

1 2 3 4 5 6 7

Obs        Condutor (Supportive)

29/09/23, 18:04

Questionários

7. \*

*Marcar apenas uma oval.*

1 2 3 4 5 6 7

Corr        Fácil

8. \*

*Marcar apenas uma oval.*

1 2 3 4 5 6 7

Inefi        Eficiente

9. \*

*Marcar apenas uma oval.*

1 2 3 4 5 6 7

Con        Evidente

10. \*

*Marcar apenas uma oval.*

1 2 3 4 5 6 7

Abo        Excitante

29/09/23, 18:04

Questionários

11. \*

*Marcar apenas uma oval.*

1 2 3 4 5 6 7

Desi        Interessante

12. \*

*Marcar apenas uma oval.*

1 2 3 4 5 6 7

Con        Original/Criativo

13. \*

*Marcar apenas uma oval.*

1 2 3 4 5 6 7

Cor        Vanguardista/ Inovador**The System Usability Scale Satandard Version**

29/09/23, 18:04		Questionários				
14. *						
		Marcar apenas uma oval por linha.				
		Discorda fortemente	Discorda	Neutro	Concorda	Concorda fortemente
<b>Acho que gostaria de utilizar este produto com frequência.</b>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Considerei o produto mais complexo do que necessário.</b>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Achei o produto fácil de utilizar.</b>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Acho que necessitaria de ajuda de um técnico para conseguir utilizar este produto</b>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Considerei que as várias funcionalidades deste produto estavam bem integradas</b>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Achei que este produto tinha muitas inconsistências</b>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>Suponho que a maioria das pessoas aprenderia a utilizar rapidamente este produto</b>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

29/09/23, 18:04

Questionários

**Considero o produto muito complicado de utilizar.**

**Senti-me muito confiante a utilizar este produto.**

**Tive que aprender muito antes de conseguir lidar com este produto.**

**NASA Task Load Index**

DIMENSÕES	DEFINIÇÃO
Exigência Mental	Exigências da tarefa em termos de atividade mental e perceptiva (pensamentos, decisão, cálculos, lembranças, contato visual, procurar, etc.).
Exigência Física	Exigências da tarefa em termos de atividade física (puxar, empurrar, girar, deslizar, etc.).
Exigência Temporal	Nível de pressão temporal sentida em função do tempo necessário e do que possui disponível.
Nível Esforço	Grau de esforço mental e/ou físico que o indivíduo precisa empenhar na realização para obter seu nível de desempenho.
Nível Desempenho	Até que ponto o indivíduo se sente satisfeito com o seu nível de rendimento e desempenho no trabalho que desenvolve.
Nível Frustração	Até que ponto o indivíduo apresenta sentimentos de insegurança, estresse, irritabilidade, descontentamento, etc., durante a realização da tarefa.

<https://docs.google.com/forms/d/1QEe1jd1Qgf--KKcuznKhvpdRMXPRUAzhTvp9mA9NxNM/edit>

6/15

29/09/23, 18:04

Questionários

15. Exigência Mental - Quão exigente mentalmente foi a tarefa? **\***  
(1- Muito baixa / 20- Muito alta)

*Marcar apenas uma oval.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20

29/09/23, 18:04

Questionários

16. Exigência Física - Quão exigente fisicamente foi a tarefa? \*

(1- Muito baixa / 20- Muito alta)

*Marcar apenas uma oval.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20

29/09/23, 18:04

Questionários

17. Exigência Temporal - Quão apressado foi o ritmo da tarefa? \*

(1- Muito baixa / 20- Muito alta)

*Marcar apenas uma oval.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20

29/09/23, 18:04

Questionários

18. Desempenho - Quão bem sucedido foi na realização do que lhe foi pedido? \*  
(Perfeito / Fracasso)

*Marcar apenas uma oval.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20

29/09/23, 18:04

Questionários

19. Esforço - Quão arduamente teve de trabalhar para atingir o seu nível de desempenho? \*  
(1- Muito baixo / 20- Muito alto)

*Marcar apenas uma oval.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20

29/09/23, 18:04

Questionários

20. Frustração - Quão inseguro, desencorajado, irritado e stressado estava você? \*  
(1- Muito baixa / 20- Muito alta)

*Marcar apenas uma oval.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20

#### **Self Assessment Manikin**

A fim de avaliar o produto, por favor preencha o seguinte questionário.

29/09/23, 18:04 Questionários

**Self Assessment Manikin**

The figure displays a 3x5 grid of manikin icons representing emotional states. The columns are labeled on the left as 'UnHappy' (top), 'Calm' (middle), and 'Controlled' (bottom). The columns are labeled on the right as 'Happy' (top), 'Excited' (middle), and 'Incontrol' (bottom). Each row contains five icons, numbered 1 through 9 from left to right. The icons show increasing levels of the respective emotion from left to right. For example, in the 'UnHappy' column, icon 1 shows a sad face, while icon 9 shows a smiling face. In the 'Excited' column, icon 9 shows a manikin with arms raised and stars around it, indicating the highest level of excitement.

21. Infeliz / Feliz \*

*Marcar apenas uma oval.*

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9

<https://docs.google.com/forms/d/1QEe1jd1Qgf--KKcuznKhvpdRMXPRUAzhTvp9mA9NxNM/edit> 13/15

29/09/23, 18:04

Questionários

## 22. Calmo / Entusiasmado \*

*Marcar apenas uma oval.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9

## 23. Controlado / No Controlo \*

*Marcar apenas uma oval.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9

---

Este conteúdo não foi criado nem aprovado pela Google.**Google Formulários**

## Apêndice C

# Tarefas das Experiências

### Guia de Avaliação Experimental

1. Explicação do que vai ser realizado nas sessões de testes (sendo diferente para a primeira e segunda sessão), qual a duração estimada, o que lhe vai ser pedido e quais os objetivos dos testes. Explicar ao utilizador também que não poderá usar a barra de pesquisa dos sites, terá de usar as várias ferramentas que o site oferece.
2. Preenchimento do Formulário de Consentimento Informado.
3. Preenchimento de um questionário para verificar se o utilizador conhece algum dos sites a ser avaliado.
4. Apenas na sessão 2: Explicação de como funciona o Muse2 e as funcionalidades que irão ser usadas nesta sessão.
5. Apenas na sessão 2: Pedir permissão para gravar os dados fisiológicos
6. Apenas na sessão 2: Teste do dispositivo muse2 para verificar se existe algum problema.
7. Realização do procedimento experimental.
8. Pedir ao utilizador que responda a um questionário demográfico (género / idade/ Educação / experiência com a tecnologia).

#### Site 1: <https://blinkee.com/>

Recentemente notou que anda a beber pouca água e, por isso, quer encontrar um copo com capacidade para pelo menos 295 ml de água (10 ounces). Por gostar bastante de leds e heavy metal, quer que o copo tenha vários leds e formato de uma caveira. Compre um produto com estas características e proceda ao checkout.

#### Site 2: <https://www.torontocupcakes.com>

O seu melhor amigo vai fazer anos daqui a uma semana e você quer encomendar uns cupcakes onde possam ser inseridas fotos personalizadas de vários momentos que vocês tiveram ao longo dos anos. Além desses cupcakes, quer encomendar de oreos, de baunilha e de baunilha e morango, pois são os sabores favoritos do seu melhor amigo. Compre cinco de cada produto com estas características e proceda ao checkout.

#### Site 3: <https://www.pennyjuice.com>

Você vai fazer uma festa para o seu primo mais novo e por isso quer comprar sumos concentrados de vários sabores (Juice Concentrates). Como o tema da festa é o Benfica quer fazer 3 sumos que sejam de cor vermelha. Compre os produtos com estas características e proceda ao checkout. (não é necessário introduzir os dados pessoais)

**Site 4: <https://www.ralphlauren.pt/>**

Você vai à festa do seu primo mais novo, cujo tema é o Benfica, mas como é sportinguista quer levar uma camisa formal, mas que seja da cor verde, Slim Fit, que lhe sirva (tamanho XL) e com o melhor preço. Compre a camisa e proceda ao checkout.

**Site 5: <https://www.frankandoak.com/>**

Você vai à festa do seu primo mais novo, cujo tema é o Benfica, mas como é sportinguista quer levar uma camisa formal, mas que seja da cor verde, de manga comprida, que lhe sirva (tamanho XL) e com o melhor preço. Compre a camisa e proceda ao checkout.

**Site 6: <https://www.airbnb.pt/>**

Você e três dos seus amigos querem ir de férias para a Austrália, mas não querem ir para um hotel ou casa normal, pretendem um sítio diferente, como uma casa flutuante. Reserve de 1 a 8 de agosto a casa com o valor mais acessível, de modo a terem um quarto para cada um e proceda ao checkout.

**Site 7: <http://www.roverp6parts.com/>**

Você foi de férias e o seu carro Rover P5 avariou precisando de trocar o disco de travão dianteiro. Compre o par de discos de travões dianteiros (Front brake disks) que sai mais em conta e proceda ao checkout (não é necessário introduzir os dados pessoais).

**Site 8: <https://www.connection.com/>**

Você vai de férias e como tal irá usar internet pública, o que é um pouco inseguro. Por isso pretende comprar um software VPN. Compre a VPN com o melhor preço visível possível e que esteja disponível e proceda ao checkout.

**Site 9: <https://www.dollartree.com/>**

Você vai à festa do seu primo mais novo e quer oferecer-lhe umas action figures de animais selvagens. Compre o pack de actions figures de animais selvagens que tem a melhor avaliação do site e proceda ao checkout.

**Site 10: <https://www.grainnemorton.co.uk/>**

Recentemente você foi de férias para diversas praias, mas não conseguiu comprar nenhum souvenir para a sua mãe. Compre um colar que contenha algumas características que lhe façam lembrar estas férias, e que esteja disponível para entrega, pois a sua mãe faz anos na próxima semana. Logo em seguida proceda ao checkout.

**Site 11: <https://www.allbirds.com/>**

Você e o seu amigo estão de férias e os ténis de corrida pretos que ele tanto gosta rasgaram-se. O seu amigo calça o tamanho 43 e quer uns ténis exatamente com as mesmas características dos que se rasgaram (leve e arejados (Light & Breezy Tree)). Compre uns ténis com essas características e o mais acessível. Logo em seguida proceda ao checkout.

**Site 12: <https://www.crateandbarrel.com/>**

Você gosta muito de batatas fritas, mas quer uma maneira mais saudável de as fazer, sem ser com óleo. Compre a AirFryer que tem o melhor preço e que esteja pronta a ser enviada, e proceda ao checkout.

**Site 13: <https://www.vitra.com/en-pt/living/product>**

Este mês recebeu um prémio do seu patrão, cujo valor lhe permite comprar aquela poltrona com um apoio de pés que sempre sonhou. Compre a poltrona mais em conta e personaliza a mesma com a base em carvalho escuro (dark oak) e a cor da capa em menta/malaquita (mint/malachite). Finalmente proceda ao checkout.

**Site 14: <https://www.zennioptical.com/>**

Você acabou de comprar uns óculos de sol para a sua melhor amiga, mas para evitar que eles estejam sempre a cair, gostaria de ter uma corrente para os pendurar ao pescoço. Compre uma corrente verde (para condizer com os óculos verdes) e que saia mais em conta. Além da corrente procura uma caixa para os óculos da mesma cor e que seja a mais popular. Depois de escolher o que foi pedido, proceda ao checkout.

**Site 15: <https://www.instant-gaming.com/>**

Você acabou de comprar um computador de gaming novo e está ansioso para o experimentar. Por ser fã do Tom Clancy's, um jogo de ação e aventura, quer comprar o jogo mais recente da saga que esteja em stock. Sabendo que o jogo é da plataforma Ubisoft Connect, compre o jogo com estas características e proceda ao checkout.

**Site 16: <https://www.kinguin.net/>**

Você acabou de comprar um computador de gaming novo e está ansioso para o experimentar. Por ser fã do Tom Clancy's, um jogo de ação e aventura, quer comprar o jogo mais recente da saga que esteja em stock. Sabendo que o jogo é da plataforma Ubisoft Connect, compre o jogo com estas características e proceda ao checkout.

**Site 17: <https://www.eneba.com/pt/>**

Você acabou de comprar um computador de gaming novo e está ansioso para o experimentar. Por ser fã do Tom Clancy's, um jogo de ação e aventura, quer comprar o jogo mais recente da saga que esteja em stock. Sabendo que o jogo é da plataforma Ubisoft Connect, compre o jogo com estas características e proceda ao checkout.

**Site 18: <https://www.aboutyou.pt>**

Você vai à festa do seu primo mais novo, cujo tema é o Benfica, mas como é sportingista quer levar uma camisa casual de manga comprida, mas que seja da cor verde, de algodão, que lhe sirva (tamanho M) e que seja a mais acessível. Compre a camisa com estas características e proceda ao checkout.

**Site 19: <https://www.kiabi.pt/>**

Você vai à festa do seu primo mais novo, cujo tema é o Benfica, mas como é sportingista quer levar uma camisa formal de manga comprida, mas que seja da cor verde, de algodão, que lhe sirva (tamanho M) e que seja a mais acessível. Compre a camisa com estas características e proceda ao checkout.

**Site 20: <https://www.purevehicleleasing.co.uk/>**

Recentemente o seu salário aumentou e queria fazer leasing de um carro. Faça o leasing de um Volkswagen Golf Hatchback com o motor 1.5 eTSI e a versão style. O carro tem de ser inferior a 474 euros (420 libras) por mês e durante 24 meses no máximo (sendo que desses 24, 3 têm de ser de entrada). Além destes fatores quer escolher o plano de 16094 km por ano (10000 "miles" por ano) e com manutenção. Não necessita de preencher os detalhes pessoais.

**Site 21: <https://www.wayfair.com/>**

Este mês recebeu um prémio do seu patrão, cujo valor lhe permite comprar uma cadeira reclinada, da mesma cor dos seus sofás (azul-marinho ("navy blue")) e com couro genuíno, que sempre sonhou. Compre a cadeira reclinada mais acessível e que tenha a opção entrega rápida. Finalmente proceda ao checkout.

**Site 22: <https://meshlabel.com/>**

Recentemente você foi de férias para diversas praias, mas não conseguiu comprar nenhum souvenir para a sua mãe. Compre um colar que contenha algumas características que lhe façam lembrar estas férias, e que esteja disponível para entrega, pois a sua mãe faz anos na próxima semana. Logo em seguida proceda ao checkout.

**Site 23: <https://www.lingscars.com/>**

Recentemente o seu salário aumentou e queria fazer leasing de um carro. Faça o leasing de um Peugeot 2008 Estate (não interessa o modelo) com um preço inferior a 650 euros por mês (571 libras) por mês e durante 26 meses no máximo (sendo que desses 26, 3 têm de ser de entrada). Além destes fatores quer escolher o plano de 16094 km por ano (10000 "miles" por ano) e com manutenção. Não necessita de preencher os detalhes pessoais.

Observação: Usar o e-mail: [UFexp@gmail.com](mailto:UFexp@gmail.com) e a password: uf123

**Site 24: <https://www.lexautolease.co.uk/>**

Recentemente o seu salário aumentou e queria fazer leasing de um carro. Faça o leasing de um Volkswagen Golf Hatchback 2.0 TDI com um preço inferior a 722 euros (640 libras) por mês e durante 24 meses no máximo (sendo que desses 24, 3 têm de ser de entrada). Além destes fatores quer escolher o plano de 24 140.16 km por ano (15000 "miles" por ano) e com manutenção e com a versão R-line. Não necessita de preencher os detalhes pessoais.