



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Cybersickness e Dados Fisiológicos: Um Estudo Exploratório

Danillo Neves Souza

Monografia apresentada como requisito parcial
para conclusão do Curso de Engenharia da Computação

Orientadora
Prof.a Dr.a Carla Denise Castanho

Brasília
2020

Dedicatória

Dedico este trabalho a todos que verdadeiramente me ajudaram a terminá-lo.

Agradecimentos

Agradeço a participação da minha orientadora Carla Castanho, assim como dos professores Tiago Barros e Maurício Sarmet, na banca avaliadora deste estudo. Agradeço verdadeiramente a todos que participaram dos experimentos, inevitavelmente presenciais. Agradeço ao Gabriel Tollini pela parceria na condução dos experimentos. Agradeço ao Vitor Hugo, por ter demonstrado apoio e preocupação com o andamento do trabalho. Agradeço minha evolução pessoal e profissional, pois sem ela não teria força de vontade suficiente para finalizar a monografia.

Resumo

A indústria dos jogos é a que mais cresce atualmente. O mercado global de jogos foi avaliado em 151,55 bilhões de dólares em 2019 e deve atingir um valor de 256,97 bilhões em 2025. Entre o grande número de gêneros, tecnologias e inovações que a permeiam, uma das mais relevantes é a realidade virtual. No entanto, há grandes dificuldades que os desenvolvedores enfrentam quando tentam entrar nesse mercado. Um dos mais proeminentes é que muitas pessoas experienciam desconforto ou mal-estar quando expostos à realidade virtual, termo conhecido como *cybersickness*. Até agora, todavia sem sucesso, muitos pesquisadores tentaram identificar padrões que ajudem a prever a severidade dos enjoos. Com o objetivo de expandir o conhecimento atual sobre a *cybersickness* e minimizar a subjetividade dos quantificadores mais utilizados na área, foram conduzidos experimentos práticos com voluntários submetidos à realidade virtual e suas atividades cardíacas e dérmicas foram capturadas (BVP e EDA). Os jogadores foram categorizados (*a posteriori*) entre aqueles que completaram o experimento (Grupo 1) e aqueles que o interromperam (Grupo 2), visto que a média do score SSQ do segundo grupo foi mais de cinco vezes maior. O Coeficiente de Correlação de Spearman indicou uma correlação positiva relativamente forte ($s=0.61$, $p<0.05$) entre o score SSQ do Grupo 2 e o nível tônico (SCL) dos jogadores. A correlação foi fraca ($s=0.44$, $p>0.05$) entre os scores e o número de respostas fásicas (picos SCR). As correlações entre os scores SSQ e os demais dados afetivos coletados foram muito próximas de zero ($\text{abs}(s)<0.2$, $p>0.05$).

Palavras-chave: cybersickness, realidade virtual, jogos eletrônicos, computação afetiva, dados afetivos

Abstract

The gaming industry is the fastest growing industry in the world. The global gaming market was valued at USD 151.55 billion in 2019 and is expected to reach a value of USD 256.97 billion by 2025. Among the countless genres, technologies and innovations that permeate it, one of the most relevant today is virtual reality. However, there are major difficulties that developers face when trying to enter this market. One of the most prominent is that many people experience discomfort or malaise when exposed to virtual reality, a term known as cybersickness. So far, however unsuccessfully, many researchers have tried to identify patterns that help to predict the severity of motion sickness. In order to expand current knowledge about cybersickness and minimize the subjectivity of the quantifiers most used in the area, practical experiments were conducted with volunteers submitted to virtual reality, where their physiological data was captured. The players were separated between those who completed the experiment (Group 1) and those who interrupted it (Group 2), since the average SSQ score of the second group was more than five times higher. Spearman's correlation coefficient indicated a relatively strong positive correlation ($s=0.61$, $p<0.05$) between the Group 2 SSQ score and the players' tonic level (SCL). The correlation was weak ($s=0.44$, $p>0.05$) between the scores and the number of phasic responses (SCR peaks), which indicates that cybersickness is more linked to the user's overall level of arousal. No correlation was found between the SSQ scores and the other affective data collected ($\text{abs}(s)<0.2$, $p>0.05$).

Keywords: cybersickness, virtual reality, game analytics, affective computing, affective data

Sumário

1	Introdução	1
1.1	Definição do Problema	1
1.2	Objetivos	1
1.3	Estruturação	2
2	Fundamentação Teórica	3
2.1	Emoções	3
2.2	Reconhecimento de Emoções	4
2.2.1	Computação Afetiva	4
2.2.2	Classificação de Emoções	5
2.3	Dados Afetivos	9
2.3.1	Dados Fisiológicos	10
2.4	EDA	16
2.4.1	Processamento e Análise do EDA	16
2.4.2	Qualidade do Sinal EDA	18
2.4.3	Algoritmos de Processamento do EDA	20
2.5	Empatica E4	23
2.6	Metodologia de Coleta de Dados Afetivos	24
2.7	Realidade Virtual	25
2.8	Imersão	26
2.9	Cybersickness	28
3	Desenvolvimento do Trabalho	30
3.1	Estudos Preliminares	30
3.2	Visão Geral da Proposta	34
3.2.1	Roteiro do Experimento	34
3.2.2	Simulação Escolhida	36
3.2.3	Processamento e Análise dos Sinais	37
3.2.4	NeuroKit2	37

4	Resultados	39
4.1	Testes Piloto	39
4.2	Testes Definitivos	40
4.3	Resultados	41
4.3.1	Scores SSQ	41
4.3.2	Dados Afetivos	44
4.4	Cálculos de Correlação	47
4.4.1	Correlação de Pearson	47
4.4.2	Correlação de Spearman	49
4.5	Coleta do EDA	50
4.5.1	Sessão de Relaxamento	51
4.5.2	Sessão de Triagem	51
4.6	Empatica E4	51
4.6.1	Confiabilidade do Sensor	52
4.6.2	Local de Coleta dos Dados	54
4.6.3	Taxa de Amostragem (EDA)	55
4.7	Processamento e Análise do EDA	56
5	Conclusão	58
5.1	Trabalhos Futuros	59
	Apêndice	70
	A Termo de Consentimento Livre e Esclarecido	70
	B Roteiro do Experimento	72
	C Simulator Sickness Questionnaire (SSQ)	76

Lista de Figuras

2.1	Sensor Empatica E4, usado neste trabalho.	4
2.2	Modelo Circumplexo de Afeto (RUSSELL, 1980)	7
2.3	Gráfico de variabilidade da frequência cardíaca média com base na idade. O HRV diminui abruptamente à medida que as pessoas envelhecem. Indivíduos de 20-25 anos geralmente têm HRV médio na faixa de 55-105, enquanto 60-65 anos tendem a estar entre 25-45. Retirado de https://www.whoop.com/thelocker/heart-rate-variability-hrv . Acessado em 08/12/2020.	12
2.4	Forma característica do BVP para um adulto saudável. A altura P1 do pulso define a medida do fluxo sanguíneo relativo. O intervalo ou distância B entre os picos (pontos 2 e 4) define a frequência cardíaca que pode ser expressa como IBI (ms) ou BPM (batidas por minuto). Retirado de https://www.biofeedback-tech.com/articles/2016/3/24/the-blood-volume-pulse-biofeedback-basics . Acessado em 08/12/2020.	13
2.5	Extração com base em filtro mediano dos componentes tônicos e fásicos de um sinal EDA. Retirado de [53].	17
2.6	Forma de onda de um Non Specific Skin Conductance Response (NS-SCR) ou de um Event Related Skin Conductance Response (ES-SCR). Adaptado de [54].	17
2.7	Distribuições Assimétricas. Fonte: https://www.slideshare.net/ranilsonpaiva . Acessado em 08/12/2020.	19
2.8	Curtose. Fonte: portalaction.com.br/estatistica-basica/26-curtose . Acessado em 08/12/2020.	20
3.1	Em sentido horário: HMD do HTC Vive, controle de Xbox 360, fones de ouvido Audiotecnica m50x e abafadores de ruído.	35
3.2	Ponto de vista do usuário durante o jogo. Fonte: Google Imagens. Acessado em 2/12/2020.	36
4.1	Dados Demográficos	42
4.2	Scores SSQ	43

4.3	Número de interrupções por quartil.	45
4.4	Representação visual da divisão de uma sessão de jogo em quartis.	47
4.5	Representação visual da divisão de uma sessão de jogo em quartis (componente tônico extraído do Neurokit2).	48
4.6	Scores SSQ de cada integrante do Grupo 2.	49
4.7	Sinal do acelerômetro do Empatica E4 durante uma sessão de relaxamento (à esquerda da linha central vermelha) e um período entre sessões de coleta, em que o jogador podia se mover livremente (à direita da linha).	50
4.8	Sinal EDA bruto obtido durante as duas sessões de relaxamento. A primeira divisão representa a sessão de cinco minutos, enquanto a terceira divisão representa a sessão de triagem.	51
4.9	Arquivo corrompido obtido do Empatica E4.	52
4.10	Sinal BVP obtido pelo Empatica E4 durante uma sessão de relaxamento.	52
4.11	Sinal BVP obtido pelo Empatica E4 durante uma sessão de corrida.	52
4.12	Relação do BVP (parte vermelha) com o acelerômetro (parte roxa).	53
4.13	Imagem capturada durante o experimento.	53
4.14	Sinal BVP e sinal do acelerômetro do Empatica E4.	54
4.15	Erro emitido pelo NeuroKit2 quando a taxa de amostragem é muito baixa.	55
4.16	Componente tônico (nível SCL) do EDA pelo Neurokit2 com diferentes taxas de amostragem.	56
4.17	Componente fásico (amplitudes SCR) do EDA pelo Neurokit2 com diferentes taxas de amostragem.	57
4.18	Resultados dos algoritmos de decomposição do EDA mais utilizados atualmente. Retirado de [48].	57

Lista de Tabelas

4.1	Valores Médios (e desvios padrão) dos dados fisiológicos obtidos pelo Empatica E4 dos jogadores que completaram a sessão de corrida. SCL - Nível de Condutância da Pele. BVP - Pulso de Volume Sanguíneo. IBI - Intervalo entre Batimentos.	44
4.2	Valores Médios (e desvios padrão) dos dados fisiológicos obtidos pelo Empatica E4 pelo grupo que interrompeu a simulação. Os valores IBI do segundo, terceiro e quarto quartis não foram captados pelo Empatica E4. SCL - Nível de Condutância da Pele. BVP - Pulso de Volume Sanguíneo. IBI - Intervalo entre Batimentos.	45
4.3	Valores Médios (e desvios padrão) do componente fásico obtido pelo Empatica E4 do grupo de jogadores que completaram o experimento.	46
4.4	Valores Médios (e desvios padrão) do componente fásico obtido pelo Empatica E4 do grupo de jogadores que não completaram o experimento.	46
4.5	Média Picos SCR por segundo por sessão de cada grupo.	47
4.6	Relação entre os scores SSQ e valores médios (e desvios padrão) do Nível SCL e picos SCR do Grupo 2.	48

Lista de Abreviaturas e Siglas

BVP Blood Volume Pulse.

CS Cybersickness.

ES-SCR Event Related Skin Conductance Response.

HRV Heart Rate Variability.

NS-SCR Non Specific Skin Conductance Response.

PAD Pleasure, Arousal and Dominance.

PPG Fotopletismografia.

RSP Respiração.

SKT Temperatura da Pele.

SNP Sistema Nervoso Parassimpático.

SNS Sistema Nervoso Simpático.

SVM Support Vector Machine.

Capítulo 1

Introdução

1.1 Definição do Problema

A Cybersickness (CS) é um dos grandes problemas atualmente enfrentados pela tecnologia de realidade virtual (VR). Apesar das extensas documentações sobre as origens da CS, quantificar e qualificar suas nuances individuais ainda é um problema. Não se sabe ao certo quais os fatores que explicam o fato de certas pessoas sentirem ou não o mal-estar. Em adição, o quantificador de *cybersickness* mais utilizado em pesquisas é o *Simulation Sickness Questionnaire* (SSQ) [87], constantemente condenado pela sua parcialidade e pela sua impraticabilidade em realidade virtual [88].

A anexação da computação afetiva ao campo da CS foi o próximo passo. Os dados afetivos proporcionam imparcialidade na identificação do estado biofisiológico de um indivíduo. A atividade eletrodérmica (EDA), por exemplo, representa uma medida funcional quantitativa da atividade sudomotora e, conseqüentemente, avalia a função do sistema nervoso autônomo (ANS) e do nível de excitação cognitiva [48].

O desafio de se utilizar dados fisiológicos se encontra principalmente na variabilidade intra-individual, nas tecnologias de captação dos dados e na detecção e correção de artefatos nos sinais. Não há definição clara, consenso literário ou sistemas preditivos universalmente aceitos. As inovações sensoriais necessitam de validações detalhadas, e as relações entre a CS e os dados fisiológicos requerem esclarecimento.

1.2 Objetivos

Este trabalho visa explorar as potenciais relações entre os dados fisiológicos, com foco no EDA, de um indivíduo e o mal-estar relacionado à imersão em realidade virtual. Para isso, experimentos em realidade virtual serão conduzidos no Departamento de Ciência da Computação, onde os participantes serão instruídos de acordo com um roteiro experimental

criado exclusivamente para esse propósito. Nos testes, serão coletados dados fisiológicos (atividade cardíaca e dérmica) dos jogadores expostos à realidade virtual utilizando o dispositivo portátil Empatica E4.

1.3 Estruturação

O restante deste trabalho está organizado da seguinte forma:

- **Capítulo 2:** explicita os conceitos necessários para a compreensão total do trabalho;
- **Capítulo 3:** detalha a proposta e as direções tomadas ao decorrer do estudo;
- **Capítulo 4:** relata os testes conduzidos e os resultados obtidos;
- **Capítulo 5:** concluiu o trabalho e apresenta sugestões para pesquisas futuras.

Capítulo 2

Fundamentação Teórica

Este capítulo apresentará conceitos principais para o entendimento deste trabalho. Serão discutidos diversos conceitos, tais como computação afetiva, o reconhecimento e classificação de emoções, os principais dados afetivos e como são calculados e interpretados, além dos conceitos sobre a realidade virtual e a *cybersickness*.

2.1 Emoções

Não existe consenso na definição de emoção. Lindsley, D. B. (1951) define emoção como um fenômeno comportamental complexo que envolve variados níveis de integração química e neural [1]. Teorias mais recentes, como as da *cognição incorporada*, já consideram que emoções dependem diretamente de outras partes do corpo - como seus músculos e vísceras - e que áreas específicas do cérebro conseguem representar um grande número de estados emocionais [2].

A dificuldade de se encontrar uma definição exata vem da natureza intrínseca das emoções. Elas estão frequentemente interligadas com humor, temperamento, personalidade, disposição, criatividade e motivação [3, 4]. Além disso, podem ser interpretadas de várias maneiras diferentes: são refletidas na fala, nas mãos e nos gestos do corpo, nas expressões faciais, corporais, etc.

As inconsistências nas definições, no entanto, não foram impeditivos na tentativa de enquadramento emocional. As incontáveis teorias que tentaram explicar a origem, neurobiologia e função das emoções apenas fomentaram ainda mais pesquisas sobre o tópico. Nas últimas duas décadas, a contribuição aumentou significativamente pela psicologia, neurociência, neurociência afetiva, endocrinologia, medicina, história, sociologia e ciência da computação. É inegável a relevância do **reconhecimento de emoções**, nome dado a essa área de estudo.

2.2 Reconhecimento de Emoções

O reconhecimento de emoções é o processo de identificação de emoções humanas, tanto por humanos quanto por máquinas. Em geral, as pessoas não são consistentes no reconhecimento das emoções de terceiros. Por isso, a tecnologia é uma ferramenta frequentemente utilizada no auxílio da detecção emocional. Até então, a maior parte do emprego da computação no reconhecimento de emoções foi realizado na automatização do reconhecimento de expressões faciais, fala e texto. Um exemplo é o *Affectiva*, um software de inteligência artificial do MIT que torna mais eficiente tarefas antes feitas manualmente por pessoas [6]. Outra aplicação comum é a automatização na medição de dados fisiológicos. O *wearable Empatica E4* (Figura 2.1), por exemplo, é um equipamento leve e intuitivo que adquire uma série de dados vitais do usuário sem necessitar de cabos ou conexão com a Internet. Esses dados são a matéria-prima para detecção e reconhecimento de emoções, objeto de estudo da computação afetiva, e são denominados dados afetivos [7].



Figura 2.1: Sensor Empatica E4, usado neste trabalho.

2.2.1 Computação Afetiva

A Computação Afetiva estuda e desenvolve sistemas e dispositivos que podem reconhecer, interpretar, processar e até simular as emoções humanas a partir de uma série de dados afetivos. É um campo interdisciplinar que abrange ciência da computação, psicologia e ciências cognitivas.

Pode-se dizer que Rosalind Picard [8] popularizou a computação afetiva, em 1995. Picard notou a possibilidade de se dar inteligência emocional às máquinas, fazendo-a melhor interpretar o estado emocional dos humanos e adaptar seu comportamento a eles. Se máquinas tivessem habilidades emocionais, seria possível construir robôs e objetos sintéticos capazes de emular a vida humana e animal, construir máquinas inteligentes, entender emoções a partir de sua modelização e, finalmente, construir máquinas que

sejam menos frustrantes na interação homem-computador. Os sistemas tornariam-se mais acessíveis, possibilitando uma relação mais amigável do usuário com a máquina. Para isso, seria necessário que a máquina possuísse o que Picard chamou de *emotional experience*, isto é, a forma como a máquina percebe os estados emocionais do usuário.

A tecnologia atual ainda não atingiu o nível que Picard imaginou, mas também não está tão longe. A *CAER NET*, de Jiyong Lee et al. [9], explora não apenas a expressão humana na detecção de emoções, mas também informações do contexto em que ela está inserida. A famosa robô humanoide Sophia utiliza *machine learning*, processamento de linguagem natural e robótica animada para criar relações com as pessoas [10]. Nestes e em muitos outros casos de estudo, utilizar múltiplos dados afetivos na identificação emocional trouxe resultados mais precisos.

2.2.2 Classificação de Emoções

É cada vez mais frequente a necessidade de sistemas emocionalmente inteligentes, ou seja, sistemas que reconheçam as diferentes emoções humanas. O reconhecimento da emoção humana desempenha um papel importante no relacionamento interpessoal. Portanto, a extração e a compreensão da emoção têm um grande potencial no desenvolvimento da interação entre a comunicação humana e a máquina. Uma inteligência artificial (IA) que não possua habilidades emocionais seria incapaz de entender verdadeiramente seus usuários, criar um forte relacionamento com eles ou ser eficaz em provocar mudanças de comportamento. Isso é especialmente verdadeiro para um sistema que precisa interagir, se comunicar e coexistir com seres humanos, como um robô social ou uma enfermeira virtual.

Uma das grandes duas vertentes teóricas da classificação emocional é a da emoção discreta. Nela, pensa-se que todos os humanos têm um conjunto inato de emoções básicas que são reconhecíveis entre diferentes culturas pela expressão facial e processos biológicos de um indivíduo (e por isso são chamadas discretas). Cada emoção atua como uma categoria distinta, em vez de um estado emocional individual. Ekman et al. (1992) concluiu que as seis emoções básicas são raiva, repulsa ou nojo, medo, felicidade, tristeza e surpresa [11]. Ekman explica que existem características particulares associadas a cada uma dessas emoções, permitindo que sejam expressas em vários graus. Estudos mais recentes sugerem que a distinção entre raiva/repulsa e entre surpresa/medo tem base social, não biológica, o que resulta em apenas quatro emoções-base [12]. No estudo publicado na *Proceedings of National Academy of Sciences*, pesquisadores identificaram 27 categorias diferentes de emoção, experimentadas ao longo de um gradiente [13].

A outra grande vertente é conhecida como a emoção dimensional, a qual sugere que o sistema neurofisiológico humano é responsável por todos os estados afetivos. Esses

modelos contrastam as abordagens da emoção básica, que propõem que emoções diferentes surgem de sistemas neurais separados. As emoções são geralmente categorizadas ao longo de duas ou três dimensões:

- **Valência:** a positividade ou negatividade de uma emoção. A felicidade tem uma valência positiva, enquanto o medo tem uma valência negativa.
- **Excitação:** varia de excitação ao relaxamento. A raiva é uma emoção de alta excitação, enquanto a tristeza é de baixa excitação.
- **Dominância:** varia da submissão a sentir-se em controle. O medo é de baixa dominância, enquanto a admiração é de alta dominância.

Vários modelos dimensionais de emoção foram desenvolvidos. A maioria é bidimensional e incorpora os eixos de valência e excitação, visto que o eixo de dominância não é tão impactante no reconhecimento de emoções. Os mais conhecidos são: o modelo circumplexo [14], o modelo vetorial [15] e o modelo de Ativação Positiva - Ativação Negativa (PANA) [16]. Um dos modelos tridimensionais mais relevantes é o modelo *Pleasure, Arousal and Dominance* (PAD) [17].

- O modelo circumplexo de emoção foi desenvolvido por James Russell (Figura 2.2). Este modelo sugere que as emoções são distribuídas em um espaço circular bidimensional, contendo as dimensões de excitação e valência. Os estados emocionais podem ser representados em qualquer nível de valência e excitação, ou em um nível neutro de um ou de ambos os fatores. Uma variação dele é o modelo vetorial, que pressupõe que há sempre uma dimensão de excitação subjacente e que a valência determina a direção em que reside uma emoção particular. Os estados de alta excitação são diferenciados por sua valência, enquanto os estados de baixa excitação são mais neutros e são representados próximos ao ponto de encontro dos vetores. Os modelos vetoriais têm sido mais amplamente usados no teste de estímulos de palavras e imagens.
- A ativação positiva - ativação negativa (PANA) ou modelo consensual de emoção sugere que o afeto positivo e o negativo são dois sistemas separados [16]. Semelhante ao modelo vetorial, estados de maior excitação tendem a ser definidos por sua valência enquanto estados de menor excitação tendem a ser mais neutros. No modelo PANA, o eixo vertical representa a intensidade do efeito positivo e o eixo horizontal representa a intensidade do efeito negativo.
- O *Pleasure, Arousal and Dominance* (PAD) ou Valência, Excitação e Dominância representa todas as emoções a partir dos eixos que dão nome ao modelo. A parte PA do PAD é chamado de *core affect* (afeto central).

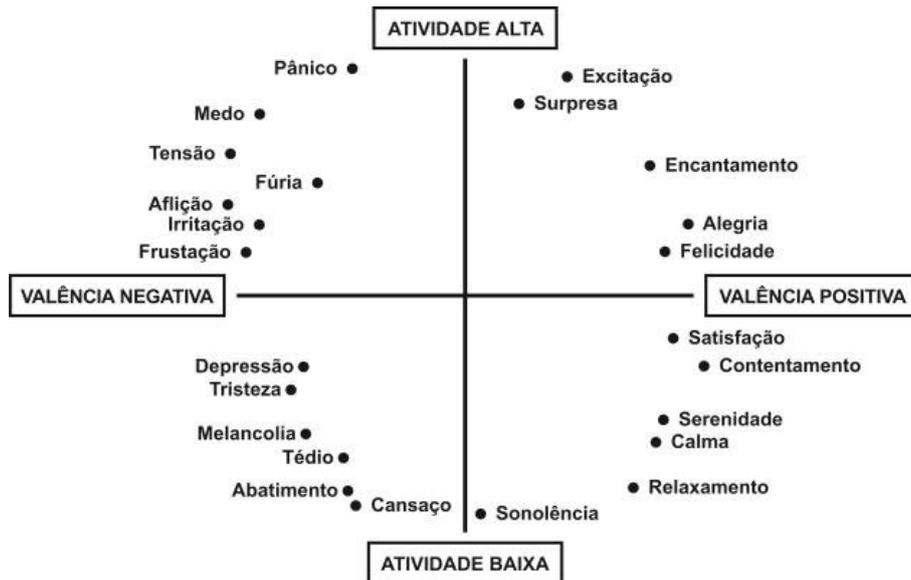


Figura 2.2: Modelo Circumplexo de Afeto (RUSSELL, 1980)

Vantagens e Desvantagens das Emoções Discretas

A questão de qual das abordagens é mais adequada para ser usada no reconhecimento de emoções ainda não foi respondida. A aplicabilidade de cada uma pode diferir consideravelmente. A seguir, serão explicadas as suas vantagens e desvantagens.

A vantagem da abordagem das emoções discretas é que ela rotula as emoções de forma simples, automática e intuitiva. A única maneira pela qual os modelos que aplicam-na podem se desviar diz respeito a quantas classes eles estão listando. Cada categoria diferente consiste em uma grande variedade de emoções. Além disso, os mesmos estados emocionais podem ser expressos por meio de diferentes categorias emocionais devido a diferenças culturais, ambientais, linguísticas ou de personalidade, que dificultariam a fixação da verdadeira categoria emocional a que pertencem.

Por outro lado, estudos etnográficos e transculturais das emoções mostraram a variedade de maneiras pelas quais as emoções diferem entre as culturas, fazendo que muitos psicólogos e antropólogos transculturais desafiassem completamente a ideia de classificações universais de emoções. Diferenças culturais foram observadas na maneira como as emoções são avaliadas, expressas e reguladas. As normas sociais para emoções, como a frequência ou as circunstâncias em que são expressas, também variam drasticamente. Por exemplo, os ocidentais experimentam emoções de alta excitação mais do que emoções de baixa excitação. Em contraste, na cultura oriental ou coletivista, as emoções de baixa excitação são mais valorizadas do que as emoções de alta excitação [18].

A maior evidência que contesta a universalidade das emoções é a linguagem ¹. As diferenças dentro das línguas estão diretamente relacionadas às diferenças na taxonomia das emoções. As línguas diferem porque categorizam as emoções com base em componentes diferentes. Alguns podem categorizar por tipos de evento, enquanto outros categorizam por prontidão para ação. Além disso, as taxonomias das emoções variam devido às diferentes implicações que as emoções têm em diferentes idiomas. Dito isso, há palavras para emoções presentes em alguns idiomas, mas não em outros.

Por último, a categorização ampla da discretização emocional pode resultar em detecção de emoção inadequada. Em vez de identificar uma emoção específica, a máquina (ou indivíduo) é forçada a selecionar alguma das categorias base (raiva, nojo, etc.) embora exista o caso de emoções mais neutras/brandas.

Vantagens e Desvantagens das Emoções Dimensionais

Os modelos dimensionais podem capturar conceitos emocionais sutis, ou seja, fornecem uma identificação única e uma ampla gama de conceitos de emoção. Além disso, uma descrição dimensional é uma abordagem favorável para medir os distintos estados emocionais. Os estados estão relacionados entre si em um espaço dimensional, o que é uma abordagem significativamente diferente dos modelos discretos.

A principal vantagem da abordagem dimensional diz respeito à resolução mais rica na caracterização de exemplos emocionalmente ambíguos em comparação à abordagem discreta. No estudo comparativo de Eerola et al. [19], as técnicas de mapeamento linear entre a abordagem discreta e a abordagem dimensional revelaram ter uma alta correspondência ao longo das emoções centrais (valência e excitação). Lichtenstein et al. [20] confirma isso em seu estudo: as taxas de reconhecimento de emoções atingiram até 72% para uma emoção específica e até 82% para uma dimensão subjacente.

Comparação entre os Modelos da Teoria Dimensional

Rubin et al. [21] compara a qualidade de detecção de emoções de múltiplos modelos dimensionais; O modelo vetorial forneceu uma descrição consistentemente melhor dos dados do que um modelo circunflexo, tanto em geral quanto na ausência de estímulos de valência neutros e de alta intensidade. O modelo PANA pôde ser testado em altos níveis de ativação, onde se mostrou idêntico ao modelo vetorial. As avaliações de valência neutra e alta excitação resultaram da média de avaliações individuais de valência positiva e negativa. Assim, em comparação com um modelo circunflexo, os modelos vetoriais e PANA forneceram ajustes gerais melhores.

¹ncbi.nlm.nih.gov/pmc/articles/PMC4396134/

2.3 Dados Afetivos

Não existe uma definição clara para dados afetivos. Para o escopo desse projeto, contudo, não é inadequado considerar um dado afetivo como aquele que possui potencial de auxiliar na análise de sentimentos e que, portanto, é utilizado na computação afetiva. Segundo Shoumy et al. [22] e Wioleta [23], os principais tipos de dados afetivos são:

Texto

Maciçamente utilizados em pesquisas de detecção de emoções e pode ser encontrado em críticas de produtos, filmes, chats e comentários de redes sociais, e-mails e muitos outros. Casos de estudo das aplicações do texto na computação afetiva podem ser encontrados em [22].

Imagem

Outro dado afetivo abundantemente utilizado, especialmente no reconhecimento facial. Extrair dados implícitos de imagens tem o potencial de trazer conhecimento inestimável para a computação afetiva. A fonte mais óbvia é a Internet, banco de imagens praticamente infinito. Desde que o Instagram foi lançado em 2010, estima-se que mais de 40 bilhões de fotos foram enviadas à rede social. Atualmente, mais de 95 milhões de fotos são postadas por dia ².

Áudio

Geralmente disponíveis em *podcasts*, músicas, mensagens de voz, aulas, televisão e rádio. Cai L. et al. [24] utiliza redes neurais para otimizar a performance de um sistema de reconhecimento de emoções a partir de um dos maiores bancos de dados emocionais disponíveis, o *IEMOCAP*, que contém mais de 12 horas de dados audiovisuais. Já Sayedelahl et al. [25] utiliza o banco de dados *SEMAINE*, populado por várias conversas emocionalmente carregadas, para propor um sistema de reconhecimento de emoções que seja capaz de lidar com emoções em grandes volumes não segmentados, não prototípicos e não pré-selecionados.

Vídeo

Uma combinação entre imagens e áudio, mas pode também conter texto. É amplamente utilizado em aplicações de segurança e vigilância, entretenimento, medicina e em programas educação e esportes. Do vídeo, podem ser retirados dados emocionais muito valiosos,

²<https://99firms.com/blog/instagram-marketing-statistics/#gref>

principalmente quando se combina a linguagem corporal com as expressões faciais de um indivíduo. O tédio, por exemplo, foi associado com o colapso da parte superior do corpo e a cabeça inclinada para trás.

Dados Fisiológicos

Representam respostas estruturais de um indivíduo quando este é submetido a determinados estímulos. Eles desempenham um papel importante na análise de sentimentos, visto que podem inferir emoções específicas do indivíduo mesmo que outros dados digam o contrário. Os mais utilizados são: eletromiograma, eletroencefalograma, eletrocardiografia, resposta galvânica da pele, pulso de volume sanguíneo, volume respiratório, temperatura da pele e frequência cardíaca.

Medidas Subjetivas

Existem relatórios de autoavaliação emocional, como o *Self-Assessment Manikin* (SAM), comumente usados para avaliar o estado mental de uma pessoa medindo as três dimensões independentes (PAD) [26]. Este método fornece uma alternativa para avaliações psicológicas de um paciente feitas por uma profissão médica, por exemplo, cuja realização exige treinamento e experiência completos para compreender as condições de saúde mental do paciente. No entanto, a validade e corroboração das informações fornecidas pelo indivíduo usando o relatório SAM nem sempre são confiáveis, uma vez que muitas pessoas têm dificuldade em se expressar honestamente por falta de conhecimento ou compreensão de seu estado mental. Os sinais fisiológicos, por outro lado, podem fornecer informações valiosas do ser vivo diretamente para os profissionais, o que os permite avaliar sua condição quase que imediatamente. Contudo, há desafios no tratamento e interpretação dos dados.

2.3.1 Dados Fisiológicos

Variabilidade da Frequência Cardíaca

A Heart Rate Variability (HRV) ou Variabilidade da Frequência Cardíaca resulta diretamente da interação entre o Sistema Nervoso Simpático (SNS) e o Sistema Nervoso Parassimpático (SNP). É um dos sinais mais relevantes no monitoramento do coração. Geralmente, a atividade cardíaca é registrada a partir de dois eletrodos no peito. A cada batida do coração, eles captam a onda de pulso gerada pelo sangue viajando pelo corpo. Níveis mais altos do HRV são associados com uma maior regulação emocional (incluindo níveis mais baixos de preocupação, ruminação e ansiedade mais baixa), habilidades sociais e habilidades cognitivas [27, 28, 29, 30, 31]. O mais interessante é que isso ocorre mesmo durante o controle de variáveis independentes conhecidas por influenciarem o HRV: sexo,

índice de massa corporal, hábitos de fumar, níveis de atividade física, depressão, ansiedade e estresse.

De uma forma geral, o HRV é essencialmente um indicativo de atividade autonômica, e por si só não é uma ferramenta eficiente na detecção de emoções. Embora os processos autonômicos sejam essenciais na regulação da emoção, também existem vários outros sistemas e processos envolvidos. Choi et al. [32] avaliou a validade do HRV como ferramenta objetiva para medir a emoção humana. Os achados deste estudo sugeriram que é possível usar uma avaliação baseada no HRV apenas quando um alto nível de emoção é induzido por estímulos visuais.

Wu et al. [33] estudou a influência de emoções específicas no padrão de atividades cardíacas e descobriu que a diversão leva a uma diminuição da frequência cardíaca. Considerando a diminuição, Wu inferiu que a diversão ativa seletivamente o sistema nervoso parassimpático. Mais importante, seu estudo mostrou que o HRV era maior na diversão do que na raiva, medo e condições neutras, e que o HRV foi maior na raiva do que na condição de medo. Embora a raiva e o medo pertençam a emoções negativas, a raiva muitas vezes está ligada à ação, enquanto o medo está ligado à evasão, sugerindo que o aumento no HRV reflita na motivação do indivíduo de agir. Em emoções como diversão e relaxamento, portanto, a motivação de ação diminui.

O HRV pode variar muito ao longo do dia, de um dia para o outro e de uma pessoa para outra, sem contar as muitas outras variáveis que são conhecidas por influenciar no HRV. Segundo Shaffer et al. [34], o HRV diminui com a idade (Figura 2.3). Quanto ao sexo, mulheres tiveram médias maiores de batidas por minuto (BPM) o que diminui o HRV: um maior BPM diminui o tempo entre batidas sucessivas e conseqüentemente, a chance de variar o intervalo entre batidas. Dessa forma, menores BPMs mais lentos aumenta o HRV. Este fenômeno é chamado de *cycle length dependence*. Quanto à saúde, o HRV aumenta de acordo com o nível de aptidão física.

Pulso de Volume Sanguíneo

O Blood Volume Pulse (BVP) ou Pulso de Volume Sanguíneo mede as mudanças no volume de sangue nas artérias e capilares as quais indicam as mudanças na frequência cardíaca e no fluxo sanguíneo. As pontas dos dedos ou os lóbulos das orelhas são comumente usados nas medições pelo uso de um sensor de Fotopletismografia (PPG). O padrão BVP bruto é influenciado pelo local do sensor, a ejeção ventricular esquerda do coração e a elasticidade/rigidez da aorta e artérias. Por esses motivos, o BVP pode variar bastante entre os indivíduos.

A amplitude do BVP exhibe o HRV momento a momento e pode oferecer uma visão significativa das respostas emocionais individuais [35]. A amplitude do BVP também pode

Middle 50% of HRV Values by Age

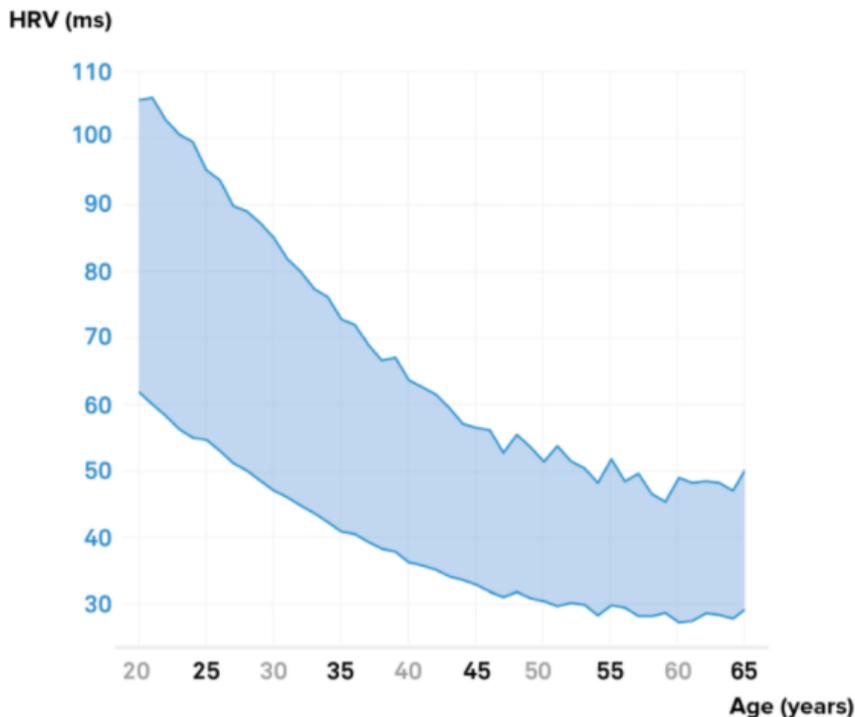


Figura 2.3: Gráfico de variabilidade da frequência cardíaca média com base na idade. O HRV diminui abruptamente à medida que as pessoas envelhecem. Indivíduos de 20-25 anos geralmente têm HRV médio na faixa de 55-105, enquanto 60-65 anos tendem a estar entre 25-45. Retirado de <https://www.whoop.com/thelocker/heart-rate-variability-hrv>. Acessado em 08/12/2020.

indicar até que ponto a pessoa foi capturada por suas emoções; quanto mais capturado, mais tempo leva para o sinal BVP retornar a algum nível de linha de base. O BVP também pode ser interpretado em termos de mudanças de temperatura periférica porque o sinal reflete as mudanças na vasoconstrição e vasodilatação. As alterações no BVP representam alterações no volume de sangue. Se a amplitude da BVP aumenta, ela reflete um aumento na vasodilatação, que leva a um aumento na temperatura periférica; Se a amplitude do BVP diminui, ela reflete uma diminuição na circulação periférica e uma diminuição na temperatura periférica. As mudanças de amplitude do BVP são muito rápidas e precedem as mudanças de temperatura, pois a temperatura é um sinal de média lenta.

O BVP é frequentemente usado em situações de treinamento de biofeedback, em vez do ECG, que é provavelmente o preferido em algumas situações clínicas ou situações em que o sujeito está se movendo muito. O sensor BVP é mais fácil de aplicar e suficiente para aplicações de treinamento, mas menos específico do que o ECG e está sujeito a alguns erros de medição em potencial que precisam ser evitados.

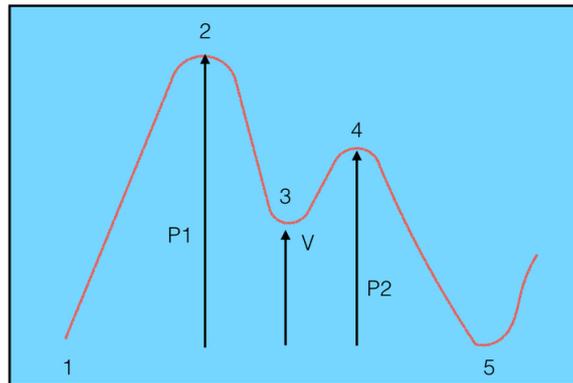


Figura 2.4: Forma característica do BVP para um adulto saudável. A altura P1 do pulso define a medida do fluxo sanguíneo relativo. O intervalo ou distância B entre os picos (pontos 2 e 4) define a frequência cardíaca que pode ser expressa como IBI (ms) ou BPM (batidas por minuto). Retirado de <https://www.biofeedback-tech.com/articles/2016/3/24/the-blood-volume-pulse-biofeedback-basics>. Acessado em 08/12/2020.

Temperatura da Pele

A Temperatura da Pele (SKT) é um sinal fisiológico muito simples, mas bastante útil. Depende do lugar da medição no corpo, a hora do dia, e se a pessoa está em repouso ou se exercitando. Ela desempenha um papel importante em termos de reconhecimento de emoções humanas e tem sido usada principalmente para estimativa de emoções agradáveis e desagradáveis (valência). Rimm-Kaufman et al. [37] notou que quando usuários eram expostos a cliques relaxantes, suas mãos esquentavam, enquanto perguntas pessoais produziram esfriamento nas mãos e um leve aquecimento no rosto, o que confirmou o resultado de outros pesquisadores e implicam que a medição da temperatura do dedo pode diferenciar estados com diferentes valências emocionais. Segundo Rimm-Kaufman, isso é menos verdadeiro para a frequência cardíaca, BVP ou atividade eletrodérmica.

A temperatura da pele facial foi utilizada por Mustafa et al. [38] para propor um novo método de detecção de emoções humanas. Conjuntamente com o HRV, o método proposto alcançou a melhor precisão de 88,75% na estimativa de emoções de estado em repouso, medo, alegria e tristeza. O estudo de Park et al. [39] tentou diferenciar felicidade e tristeza usando SKT e BVP. Como resultado, o método multimodal baseado em Support Vector Machine (SVM) proporcionou melhor precisão de classificação (92,41%) do que métodos baseados em um único recurso, como usar apenas a amplitude SKT (89,29%).

Atividade Elétrica Muscular

A atividade elétrica muscular (EMG) ocorre durante os ciclos de contração e relaxamento de um músculo. A grande maioria das pesquisas da computação afetiva usa o EMG para detectar movimentos faciais. Sem dúvidas, pode ser muito útil para o reconhecimento de emoções, mas os eletrodos fixados no rosto são muito desconfortáveis e intrusivos para um usuário e inadequados em várias situações. Contudo, outras áreas do corpo foram exploradas, como no estudo de Muhammad et al. [40], que utiliza o sensor EMG *MyoSensor* no antebraço para atingir precisão de 88,1% no eixo de excitação. Isso mostra oportunidades significativas em vários campos da ciência da computação, como a área de jogos e ferramentas de *e-learning*, onde os sinais EMG podem ser usados para detectar emoções humanas e fazer com que o sistema forneça *feedback* com base nelas (campo conhecido como *biofeedback*).

No mundo dos jogos, o EMG tem sido usado no biofeedback para controlar todo tipo de sistema. Hernandez et al. mostrou que é possível controlar um sistema protético usando um sinal EMG conectado diretamente à pele [41]. Zhang et al. mostrou que é possível fazer uso de EMG e dados do acelerômetro para classificar qual gesto uma pessoa faz e usar esse gesto para controlar um videogame [42].

Respiração

A Respiração (RSP), como dado afetivo, pode ser medida de várias formas, mas é mais comumente definida como a mudança de tensão da caixa torácica de um indivíduo. O sensor se estende ao redor da cavidade torácica, coletando (a partir da tensão de um pequeno elástico do sensor) a profundidade e a taxa de respiração do usuário. Esses dados são indicativos diretos relacionados à taxa de excitação. Quando em um nível muito elevado de excitação, a tendência é que a tensão da respiração aumente, conseqüentemente aumentando a taxa de respiração e sua profundidade. Jerath et al. [43] afirma que o método mais poderoso de regulação da emoção que pode ser praticado pelo leigo pertence aos aspectos da resposta mente-corpo, sendo o método mais eficaz as técnicas de respiração, como o *pranayama*. A respiração é especial porque não é apenas uma influência poderosa na psicologia e na fisiologia, mas também no ritmo corporal que é controlado voluntariamente. Isso o coloca na vanguarda das técnicas mente-corpo.

A respiração rápida e superficial está relacionada a emoções negativas como medo, raiva e ansiedade [44, 45]. A repulsa/nojo leva à supressão e à interrupção da respiração, provavelmente uma reação natural de evitar a inalação de conteúdos nocivos. Já emoções positivas produzem aumentos na variabilidade do padrão respiratório e diminuições no volume corrente e no tempo de inspiração. Contudo, os efeitos das emoções positivas

na respiração dependem do quão estimulantes elas são. Emoções de maior excitação aumentam a taxa de respiração. Outro ponto interessante é que padrões lentos e profundos de respiração podem fortalecer emoções positivas quando as negativas prevalecem.

Outro estudo de Zhang et al.[46], em 2017, concluiu que os sinais de respiração, por si só, têm um grande potencial para reconhecer emoções e que o modelo valência/excitação pode melhorar sua precisão. Em adição, Mirmohamadsadeghi et al. [47] atingiu maior precisão no reconhecimento de emoções do que tecnologias mais conhecidas, usando apenas sinais do coração e de respiração, sendo um bom candidato para sistemas de monitoramento portáteis.

Atividade Elétrica do Cérebro

A atividade elétrica do cérebro é captada pelo eletroencefalograma (EEG), que amplifica os impulsos elétricos cerebrais e os registra de forma gráfica. Esses impulsos são os responsáveis pelas atividades realizadas pelo corpo humano, transmitidos como comandos cerebrais através de células chamadas neurônios. Como a atividade elétrica espontânea está presente desde o nascimento, o EEG pode ser útil em todas as idades, desde recém-nascidos até pacientes idosos.

Em geral, a atividade cerebral é um dos sinais mais poderosos de extração emocional. O sinal de ondas cerebrais de um ser humano produz uma série de sinais de neurônios que gerenciam todas as funcionalidades do corpo. O cérebro humano armazena as experiências emocionais que são reunidas ao longo de sua vida. Explorando diretamente os sinais das ondas cerebrais, podemos examinar as respostas emocionais de uma pessoa quando exposta a certos ambientes.

O EEG é usado para registrar tais atividades e é confiável para o reconhecimento de emoções devido à sua avaliação relativamente objetiva da emoção em comparação com pistas não fisiológicas (expressão facial, gesto, etc.) [68, 69]. Trabalhos descrevendo que o EEG contém os recursos mais abrangentes, como as bandas espectrais de potência, podem ser utilizados para classificações emocionais básicas [69]. Existem três estruturas no sistema límbico, onde o cérebro implica fortemente a emoção e a memória: o hipotálamo, a amígdala e o hipocampo. O hipotálamo lida com a reação emocional enquanto a amígdala lida com estímulos externos que processam as informações emocionais a partir do reconhecimento de situações, bem como da análise de ameaças potenciais. Estudos sugeriram que a amígdala é a base biológica das emoções que armazenam o medo e a ansiedade [70, 71, 72]. Finalmente, o hipocampo integra a experiência emocional com a cognição.

Atividade Elétrica da Pele

Também conhecido como *Electrodermal Activity* (EDA) ou atividade eletrodérmica, e como *Galvanic Skin Response* (GSR) ou resposta galvânica da pele, é uma manifestação elétrica do sistema simpático na inervação das glândulas sudoríparas [48]. O EDA de um indivíduo variará por vários motivos, incluindo sexo, dieta, tipo de pele e humor. Ele carrega uma história psicofisiológica (incluindo emocional ou estresse cognitivo) desde 1879, por muitos anos apenas retirado dele a média dos valores para avaliar a excitação. Muito mais tarde, os pesquisadores descobriram que o EDA contém informações não apenas nas mudanças lentas (componente tônico) que a média representa, mas também nas mudanças rápidas ou fásicas do sinal (componente fásico). Dessa forma, as medições de condutância tonal da pele indicam a ausência de qualquer evento ambiental específico ou estímulo externo, variando em dezenas de segundos a minutos e podendo diferir significativamente entre os indivíduos. As medições de condutância fásica, por outro lado, são associadas a eventos de curto prazo e ocorrem na presença de estímulos ambientais discretos. De todos os sinais psicofisiológicos supracitados, este trabalho se concentrará no EDA e, por isso, a próxima seção é dedicada exclusivamente à ele.

2.4 EDA

A condutância da pele oferece ideias tremendamente valiosas sobre a excitação subconsciente de um indivíduo quando é confrontado com material de estímulo emocionalmente carregado. No entanto, o sinal de EDA tem uma grande variabilidade devido aos diferentes contextos em que o sinal pode ser adquirido e à variabilidade fisiológica entre os indivíduos. Essa variabilidade torna difícil encontrar o significado dos resultados da classificação e relacionar os resultados na resposta do sinal de EDA a um tipo específico de estímulo. Apenas com base na resposta galvânica, não é possível identificar se a excitação foi devido a conteúdo de estímulo negativo ou positivo, visto que ambos podem resultar em um aumento na excitação, desencadeando picos EDA idênticos em formato. Em outras palavras, o EDA não é capaz de revelar a qualidade das emoções. O verdadeiro poder do EDA desdobra-se quando é combinado com outras fontes de dados para se medir variáveis complexamente interligadas de forma a pintar o quadro completo do comportamento emocional.

2.4.1 Processamento e Análise do EDA

As características mais frequentemente usadas no domínio do tempo são estatísticas descritivas, como valor médio do sinal, desvio padrão, curtose, assimetria e variância [53].

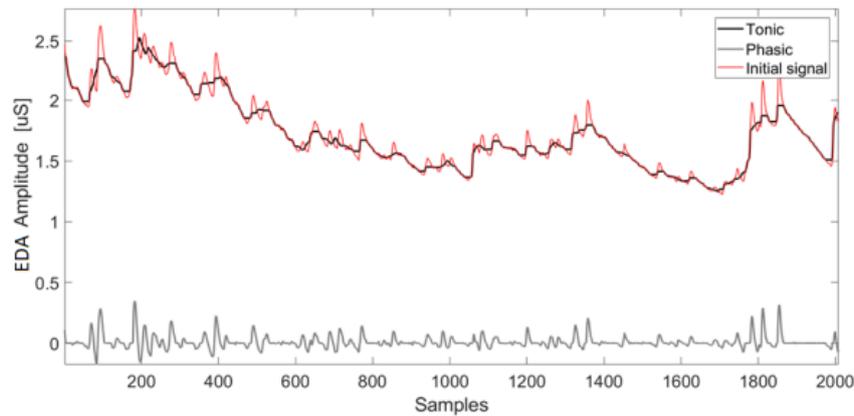


Figura 2.5: Extração com base em filtro mediano dos componentes tônicos e fásicos de um sinal EDA. Retirado de [53].

Outras métricas frequentemente encontradas na literatura estão relacionadas às alterações morfológicas do sinal, como a área sob a curva. Algumas pesquisas se concentraram em recursos no domínio da frequência, mas como as características transitórias do sinal EDA afetam principalmente o domínio do tempo, os recursos de computação do sinal pela Transformada Rápida de Fourier são considerados uma abordagem promissora. Para o componente fásico, é preferível se concentrar principalmente sobre os SCRs diretamente relacionados a eventos (ER-SCR), visto que eles podem ser interpretados como medidas diretas de excitação e engajamento do indivíduo. SCRs individuais podem ser caracterizados pela Figura 2.6.

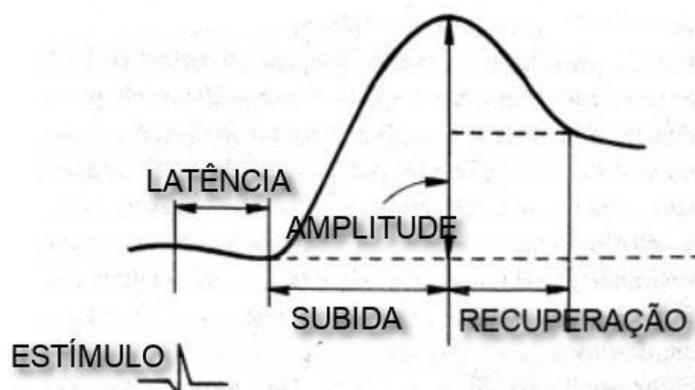


Figura 2.6: Forma de onda de um Non Specific Skin Conductance Response (NS-SCR) ou de um Event Related Skin Conductance Response (ES-SCR). Adaptado de [54].

O ER-SCR varia de acordo com as diferenças individuais no nível tônico ou devido ao ruído decorrente do movimento ou artefatos de respiração. Vários picos do EDA podem

ocorrer em sucessão direta, resultando em um acréscimo do sinal em vez de voltar aos níveis base. Esse fenômeno faz a análise do tempo de recuperação após picos muito mais difícil em comparação com a análise de medidas de latência e amplitude do sinal. A análise de picos e magnitudes das respostas fásicas são importantes, pois permitem o pesquisador analisar qualquer uma das seguintes métricas em relação às diferenças entre estímulos, ou apenas compará-los ao longo do tempo. Os dados mais comumente retirados das respostas fásicas são:

- Números de picos GSR: informa quantos picos de GSR ocorreram durante uma condição de gravação. Pode-se obter o número para cada entrevistado e calcular um número médio de picos GSR entre os entrevistados.
- Amplitude e magnitude média do pico GSR: se o experimento apresentará o mesmo estímulo várias vezes, pode-se calcular a média das amplitude GSR. Para esta medida, respostas nulas (onde o entrevistado não mostrou o pico GSR) são simplesmente ignorados. Para refletir também estes, alguns autores informam a magnitude GSR, que inclui respostas nulas.

2.4.2 Qualidade do Sinal EDA

Uma das principais preocupações sobre o EDA diz respeito à qualidade do sinal. O SCL e os NS-SCRs são altamente variáveis entre indivíduos, e deve haver uma separação clara entre respostas espontâneas da condutância da pele e respostas induzidas por componentes específicos do experimento. Tradicionalmente, a obtenção de NS-SCRs exigia um observador para contar os picos, o que era difícil se o EDA fora modificado por artefatos.

O principal causador de erros de coleta no EDA são artefatos de movimento. Outros conhecidos por causarem artefatos são flutuações na temperatura, ruído do sensor, respiração irregular, fala ou qualquer tipo de evento que possa causar respostas no sinal sem o conhecimento do estudo. Tais artefatos podem ser facilmente confundidos com respostas galvânicas legítimas. A incapacidade de distinguir esses artefatos de SCR reais limita o alcance da análise e interpretação de dados EDA.

Detecção e Correção de Erros

As abordagens tradicionais para lidar automaticamente com ruído e artefatos de movimento em EDA incluem filtragem passa-baixa (com frequência de corte típica de 1 Hz), suavização exponencial e remoção dos segmentos de dados corrompidos. Estudos recentes já propõem soluções mais robustas, como a utilização de *machine learning* para detectar artefatos no EDA.

Normalização

O SCL e os SCRs podem ser não normalmente distribuídos e podem exibir sinais de assimetria e curtose (Figuras 2.7 e 2.8). Por isso, o sinal EDA geralmente é normalizado. As medições de SCL podem ser corrigidas usando o log dos valores. O log seria calculado após as amplitudes foram determinadas (não aplicando o log aos valores brutos). Eles também podem ser transformados através de transformações de raiz quadrada (square-root transform). As medições de SCR podem ser corrigidas aplicando uma transformação de raiz quadrada às amplitudes do SCR, o que não requer a inclusão de uma constante. Pode-se também tomar o log dos valores - novamente, aplicado após a determinação das amplitudes (não aplicando o log às pontuações brutas). Para medições onde respostas nulas estão incluídas, um 1 é adicionado a todas as amplitudes SCR antes do log ser realizado ($\text{Log}(\text{SCR} + 1)$) porque o logaritmo de zero é indefinido.

Distribuição Assimétrica (Enviesada)

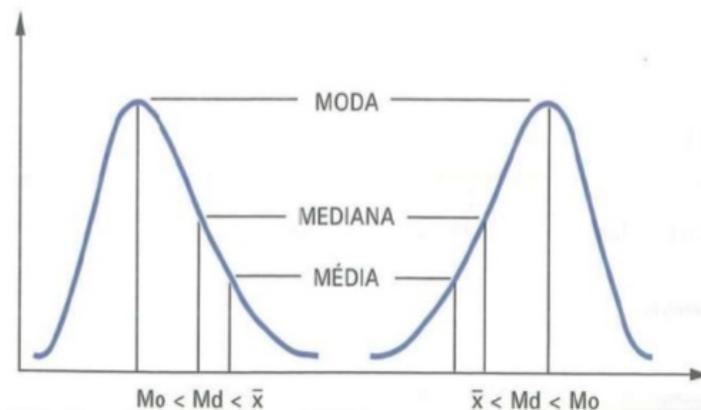


Figura 2.7: Distribuições Assimétricas. Fonte: <https://www.slideshare.net/ranilsonpaiva>. Acessado em 08/12/2020.

Padronização

A padronização, por sua vez, refere-se à correção dos dados para que possam ser direta e significativamente comparados entre indivíduos. É prudente transformar quaisquer SCRs de um indivíduo em um sistema coordenado que reflete as limitações de sua própria capacidade de responder à estímulos. O objetivo da padronização é facilitar comparações individualmente diferentes onde todas as SCRs foram transformadas em relação à sua

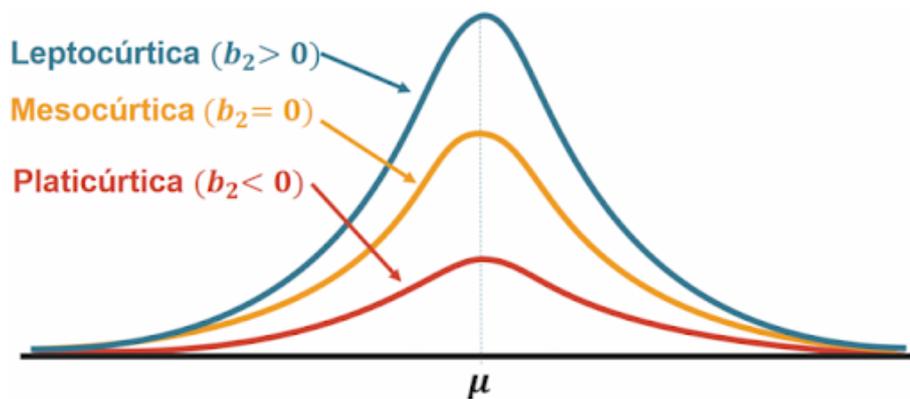


Figura 2.8: Curtose. Fonte: portalaction.com.br/estatistica-basica/26-curtose. Acessado em 08/12/2020.

responsividade fisiológica, e assim para excluir questões como espessura da pele, etc. Uma série de sugestões foram feitas e usadas na literatura e sua utilidade ainda está sendo debatida. Muitos anos atrás, os mais utilizados eram:

- **Pontuações corrigidas por intervalo:** calcula-se o SCL mínimo durante um período de base ou descanso e um SCL máximo durante o período mais excitante. Em qualquer outro período de tempo no estudo, o SCL do participante pode então ser delineado como uma proporção de sua faixa máxima individual de resposta psicofisiológica através da fórmula $(SCL - SCL_{\min}) / (SCL_{\max} - SCL_{\min})$ [49].
- **Proporção de Resposta Máxima:** no caso de dados SCR, assume-se o SCR mínimo zero e o máximo para ser o resultado de um estímulo de medo. A partir disso, pode-se então padronizar os valores para o indivíduo, dividindo cada SCR individual pelo SCR máximo do participante.

O problema com essas abordagens é que o que conta como “SCL/ SCR mínimo” pode ser artificialmente estrangido pela tecnologia, não necessariamente representando o verdadeiro mínimo do indivíduo. Além disso, as SCRs não são consistentes, mesmo dentro do mesmo individual, e procedimentos de correção de intervalo são problemáticos se diferentes grupos divergem muito nos valores. Como consequência, os métodos descritos acima tornaram-se cada vez mais controversos.

2.4.3 Algoritmos de Processamento do EDA

Existem muitos estudos [50, 51, 52] que examinaram formas automáticas para detectar NS-SCRs, extrair amplitude ou outras medidas de um único SCR e gerenciar os componentes do sinal. A figura 2.5, por exemplo, ilustra um algoritmo de decomposição do EDA em

seus respectivos componentes. Separando-os com precisão, pode-se usá-los para medir objetivamente as reações dos sujeitos a estímulos específicos ou situações significativas.

Extração do Componente Tônico

A decomposição tônica/fásica é uma das tarefas mais demandadas no processamento do EDA. Uma simples média não é adequada como medição do SCL já que provavelmente vai superestimar o seu verdadeiro valor. São necessários algoritmos de decomposição, como:

- Subtrair as amplitudes do SCR do sinal tônico ou desconsiderar os períodos SCRs do sinal;
- Utilizar a frequência de NS-SCRs e a amplitude desses picos. Valores típicos são de 1-3 por minuto em relaxamento e mais de 20 em situações de alto estresse [54]. É importante, contudo, utilizar um período fixo de tempo ao longo de todas as medições quando se foca nesses picos. Isso está baseado em observações que o SCL pode cair quando a frequência de NS-SCRs aumenta. Analisar a amplitude dos NS-SCRS e seus desvios padrão podem ser indícios adicionais de excitação tônica;
- Calcular a média de todos os valores da condutância da pele anteriores a ER-SCRs. É recomendado para situações em que esses valores SCL ocorressem fora de SCRs anteriores e que a frequência de ER-SCRs seja próxima ou igual a 100%. Isso é importante porque cada evento em um determinado sinal forneceria um valor SCL a partir do qual calcular uma média geral. Como tal, é preciso garantir que ocorram observações de SCL individuais aproximadamente iguais para que as médias sejam baseadas em um número semelhante de observações. Se a frequência de SCRs for muito diferente entre os indivíduos, esta abordagem não é aconselhável porque o número de observações que contribuem para os valores finais da SCL seria incomparável;
- Especificar uma janela de tempo antes de cada SCR, pois isso também dá uma medida da SCL tônica que provavelmente estará fora do SCR. Novamente, um problema aqui é que o número de observações subjacentes a qualquer média seria ditado pelo número de SCRs presentes no sinal. Isso pode não ser correspondido entre condições ou participantes.

Extração do Componente Fásico

Um atrito geral que tem que ser equilibrado em qualquer experimento que busque medir o EDA é determinar a priori se um determinado SCR é relacionado a eventos (ER-SRCs)

ou não específico (NS-SCRs). Se os critérios são muito vagos, há um risco de incluir respostas não específicas na análise de respostas específicas, levando à conclusão errônea que esses valores estão ligados com o experimento; muito restritos, esses critérios podem levar muitos ER-SCRs a serem descartados, alterando significativamente os valores calculados. Algumas propostas da literatura são:

- Realizar experimentos piloto para verificar as características responsivas do EDA em certas condições experimentais para ajudar a definir esses critérios [54];
- Para cada amostra do sinal, computar a média do valor EDA dos valores ao redor da amostra (mais ou menos 4 segundos). Então, se subtrai a média desses valores da amostra em questão. O resultado é o componente fásico. Para um sinal normalizado e corrigido, 10Hz de taxa amostral são suficientes [55];
- Para detecção de picos, 1) encontrar os inícios de pico ($> 0,01 \mu\text{S}$) e deslocamentos ($< 0 \mu\text{S}$) nos dados fásicos; 2) nos dados EDA não filtrados, encontrar o valor máximo de EDA dentro de cada par de inícios e deslocamentos. Estes são os picos. A amplitude de pico EDA é o amplitude no pico menos a amplitude no início ³.

Em termos fásicos, amplitudes normalmente podem variar de limiar a um máximo de cerca de $2-3\mu\text{S}$ ($X = 0,30\mu\text{S} - 1,30\mu\text{S}$: crus, não normalizados) e em torno de $0,20 - 0,60\mu\text{S}$ em média quando logaritimizado. Em experimentos onde estímulos altamente aversivos / ameaçadores / temerosos são usados, este máximo a resposta pode aumentar para cerca de $8\mu\text{S}$ (não normalizado), mas isso é raro. Latência: Normalmente, um ER-SCR é definido como um SCR cujo período de latência entre início de estímulo e o primeiro desvio significativo no sinal é entre 1-3seg. Alterações no sinal que ocorrem antes deste período são tipicamente definidos como NS-SCRs e não são visto diretamente de manipulações experimentais. Historicamente, o limiar mais comum foi fixado em $0,05\mu\text{S}$ (aproximadamente o menor mudança visível nos gravadores de gráficos de papel). No entanto, avanços mais recentes em tecnologia e precisão levaram a limiares de $0,04\mu\text{S}$, $0,03\mu\text{S}$ e $0,01\mu\text{S}$ tornando-se mais comum na literatura. Deflexões no sinal que não satisfazem o limiar os critérios não são contados como SCR ou NS-SCRs. Estimativas para a frequência de NS-SCRs é difícil, mas uma média de 1-3 por/min (com cerca de 10 por/min sendo possível durante períodos de silêncio) tem sido sugerido. Alta excitação está associada com cerca de 20-25 por/min [49].

³<https://imotions.com/guides/eda-gsr/>

2.5 Empatica E4

A pulseira Empatica E4 é um dispositivo de pesquisa vestível que oferece aquisição de dados fisiológicos em tempo real e ferramentas para análise e visualização deles a partir dos seguintes sensores ⁴:

- Sensor de Atividade Eletrodérmica;
- Sensor PPG, que mede BVP, no qual a HRV pode ser derivada;
- Termopilha Infravermelha, a qual registra a temperatura da pele periférica;
- Marcação de eventos (relógio interno de alta precisão) e vinculação com sinais fisiológicos;
- Acelerômetro de 3 eixos, capturando atividades baseadas em movimento.

A principal justificativa de se utilizar o Empatica E4 para se obter os dados fisiológicos neste estudos foi que o uso de sensores mais tradicionais podem impactar na aceitação e no conforto dos usuários, o que pode causar uma mudança no estado emocional da pessoa, atrapalhando o resultado final da pesquisa [75]. De fato, isso faz bastante sentido no contexto de jogos, onde a maioria requer que o jogador utilize as duas mãos. Outra forte justificativa é sobre a qualidade do sinal. O sensor de EDA do Empatica E4 possui uma faixa dinâmica que permite uma faixa de valores excepcionalmente alta [0, 100] μS com uma resolução de 900 pS. Segundo Dawson [49], essa faixa e resolução são mais do que satisfatórias para registrar o SCL e o SCR. Os dispositivos endossomáticos, como o Empatica E4, são considerados mais simples pois exigem apenas um amplificador de alta impedância de entrada ($> 10 \text{ MOhms}$). No entanto, eles exigem um ganho de amplificador e um referência para medir a diferença de potencial entre os dois eletrodos. O sinal obtido com este método é denominado resposta potencial da pele, cuja relação com SCRs é direta. A resposta potencial da pele pode ser monofásica positiva, monofásica negativa, bifásica ou trifásica, complexidade que dificulta a pontuação e interpretação do sinal. Isso limitou o uso de registros endossomáticos em estudos recentes [48].

O Empatica E4 já foi comparado com o sensor de laboratório Biopac MP150 [56], por exemplo, e no sentido de reconhecimento de emoções demonstrou precisão semelhante. Esses resultados suportam a confiabilidade do reconhecimento de emoções fora do laboratório. Em termos de discriminação de estresse, o EDA da pulseira E4 mostrou-se mais discriminante do que os eletrodos de dedo, sugerindo um bom desempenho para detecção de estresse, apesar da grande diferença no nível do sinal [57]. Milstein et al. [58] concluiu que a pulseira E4 é uma ferramenta confiável para adquirir dados médios de IBI e HR

⁴<https://www.empatica.com/research/e4/>

médios. Além disso, após a edição manual, a série temporal IBI, derivada do sensor PPG do E4, era semelhante à série IBI derivada dos dados de ECG obtidos pelo dispositivo móvel MW por meio de eletrodos conectados ao torso.

Todavia, a literatura é ambígua quanto ao EDA coletado pelo Empatica E4. Milstein encontrou, no mesmo estudo, duas grandes desvantagens utilizando o Empatica E4: a falha em produzir dados confiáveis de EDA e a quantidade de dados IBI ausentes, especialmente quando o participante está sendo mais dinâmico. O E4 foi menos preciso para HRV, obtido pelo software externo Kubios Premium, e que falhou em produzir dados EDA confiáveis em sua amostra. Esses resultados estão de acordo com outros relatórios preliminares (mais detalhes em [58]). Da mesma forma, a versão anterior do E4, o E3, havia se demonstrado um dispositivo de medição de EDA de alta precisão, quando dois dedos eram conectados por eletrodos à pulseira (Empatica, 2017). Pode ser que, para fins científicos, as opções atuais para uma medida de pulso de EDA não sejam satisfatórias. Judith et al. [77] utiliza a frequência cardíaca, EDA e atividade cerebral para tentar induzir no jogador um estado de relaxamento a partir de biofeedback. O EDA, como sinal de biofeedback, mostrou ter um tempo de reação muito lento para ser percebido conscientemente pelo jogador. Borrego et al. [59] tentou determinar a confiabilidade do E4 para registrar mudanças na resposta galvânica da pele ao enfrentar imagens emocionalmente valiosas em comparação com um sistema de nível de laboratório. Seus dados mostraram que a taxa de amostragem do E4 pode ser muito baixa para registrar o componente fásico do EDA, que codifica as mudanças relacionadas à excitação. Este componente do EDA foi sugerido para variar de 0,05-1,5 Hz ou mesmo 0,05-5 Hz, o que levaria a uma taxa de amostragem mínima de 3 Hz ou 10 Hz, respectivamente. Neste último caso, a taxa de amostragem do E4 (4 Hz) não seria suficiente para captar o componente fásico do EDA, o que restringiria o E4 para medir mudanças no componente tônico.

2.6 Metodologia de Coleta de Dados Afetivos

As recomendações da literatura disponível [49, 54] são:

- Inalação excessiva e exalação (ou prender a respiração) resultam em fortes derivadas em os dados GSR. Estes poderiam facilmente ser mal interpretados como mudanças na excitação emocional. É recomendado registrar atividade respiratória junto com GSR usando um cinto respiratório ou dispositivos similares;
- Movimentos, particularmente das partes do corpo onde os sensores EDA estão conectados, podem afetar fortemente a qualidade dos dados;

- Falar deve ser evitado durante sessões de coleta, pois provoca variações lentas no sinal que não reflete excitação emocional;
- O período de relaxamento não deve apresentar nenhum tipo de estímulo. Alguns dos estímulos são efeitos sonoros de ambiente, paisagens, objetos que podem ser carregados e manipulados, entre outros. Caso se opte por utilizar cenas relaxantes, é recomendável que também sejam mostradas cenas igualmente estressantes;
- Como os picos de GSR ocorrem dentro de 1-5 segundos após o início de um estímulo, deve-se apresentar qualquer material tempo suficiente para os entrevistados processarem seu conteúdo. É recomendável apresentar estímulos funcionais tempo suficiente e usar estímulos intermediários para que o sinal GSR para todos os respondentes possa recuperar e começar a partir de níveis de linha de base para o próximo estímulo.

2.7 Realidade Virtual

Segundo Burdea [60], a realidade virtual (VR) é uma simulação criada com gráficos computadorizados para criar um mundo que parece real. Um elemento que ajuda bastante nesta simulação é a possibilidade de interação com este mundo virtual, onde o jogador pode utilizar alguma forma de input para modificar o mundo virtual prontamente. Em vez de localizar uma câmera real em um ambiente físico (como a realidade aumentada faz), a posição dos olhos do usuário está localizada no ambiente simulado. Se a cabeça do usuário virar, os gráficos reagem de acordo. A realidade virtual não constrói objetos virtuais em uma cena real. Ela cria um mundo interativo para o usuário, ou seja, um ambiente artificial completamente habitável.

Durante o passar dos anos, foram propostas várias implementações de hardware para tentar simular esse mundo. O mais utilizado é o Head Mount Display (HMD), muito provavelmente pelo seu custo-benefício. O HMD costuma ser a maior diferença entre os sistemas de realidade virtual imersivos e as interfaces de usuário tradicionais. Estes óculos possuem um isolamento externo, de modo que pouca ou nenhuma luz exterior consiga entrar. A imagem é separada para as duas lentes do óculos. Desta forma, a imagem parece mais fidedigna aos olhos humanos. Para que a aplicação saiba para onde o usuário está olhando, o óculos possui um acelerômetro integrado. Alguns exemplos de óculos de realidade virtual disponíveis no mercado são o *Oculus Rift* e *HTC Vive*. O futuro dos HMDs está se desenrolando, mas ainda é desconhecido. Existem também grupos de modelos desses óculos, que utilizam a mesma tecnologia mas são fabricados por

empresas diferentes. Exemplos destes grupos são o Windows Mixed Reality e o Google Cardboard.

2.8 Imersão

Mesmo sendo um conceito muito popular (principalmente na área de jogos em geral), a imersão é raramente estruturada ou explicada. Ela parece ser entendida pela comunidade de jogos, mas a definição de imersão é subjetiva e suas origens incertas. Por exemplo, existem jogos com um mundo realista e sons atmosféricos (características comuns dos jogos que muitas vezes causam o efeito da imersão) e, mesmo assim, a imersão não é alcançada ⁵. É muito difícil descobrir o que exatamente se entende por imersão e até mesmo se as diferentes pesquisas sobre imersão estão falando sobre o mesmo conceito. É fundamental que a imersão seja definida claramente, pois só assim será possível quantificá-la e qualificá-la adequadamente.

Originária da área cinematográfica, a palavra é usada para se referir ao ato de se imergir completamente ou mergulhar em outro mundo (artificial). Esse conceito contrasta com a metáfora de uma janela, onde se observa o que está acontecendo de fora. No contexto da realidade virtual, imersão é a condição na qual o usuário perde a consciência do fato de estar realmente em um mundo artificial. Ele ou ela experimenta o mundo virtual com todos os seus sentidos e é capaz de, ao contrário da imersão cinematográfica, interagir com o ambiente virtual.

Estudos de Cairns et al. [62, 63] explicam que a imersão se comporta ao longo de um espectro. Ao entrevistar um grupo de jogadores sobre seus jogos favoritos, descobriu uma série de características que induzem a imersão, como o realismo do mundo do jogo, sons atmosféricos e profundidade. O estudo também descobriu que os jogadores não esperavam estar totalmente imersos (presentes) no jogo o tempo todo. Segundo Cairns, a imersão é estruturada em três estágios: engajamento, absorção e presença. O maior nível (presença) é igualado à imersão. Contudo, pode ser muito difícil de alcançar, já que cada nível de engajamento é controlado por barreiras, tanto da parte humana (como a concentração) quanto da parte digital (como a construção do mundo do jogo). Cada nível só é realmente alcançado quando suas respectivas barreiras são removidas. De qualquer forma, remover tais barreiras não garante a experiência, apenas permite que ela ocorra. Os três estágios da imersão de Cairns serão explicados a seguir.

- **Engajamento:** primeiro estágio da imersão. Isso significa que ele deve necessariamente ocorrer antes dos outros dois estágios. O engajamento está diretamente

⁵Taylor J. - Review of Circle of Blood

ligado com a preferência do jogador, isto é, se eles gostam ou não de um certo tipo de jogo e se lhes prende a atenção. Cairns explica que, para que esse nível seja atingido, o jogador deve investir tempo, esforço e atenção.

- **Absorção:** o segundo estágio da imersão. Um jogador está absorvido pelo jogo quando suas emoções são diretamente afetadas por ele. Para atingir esse estágio, um jogador (já engajado com o jogo) deve se conectar com o mundo virtual apresentado a ele. Assim, a barreira para esse estágio é a construção do jogo, isto é, a combinação de gráficos, trama e sons que criam uma atmosfera específica.
- **Presença:** o terceiro e último estágio da imersão. A presença se iguala à imersão total. Para os jogadores, esse estágio significa que o jogo é tudo que importa no momento, ou seja, um sentimento de desapego total da realidade, mesmo que a maioria das vezes seja só por alguns instantes. Em comparação com o estágio de absorção, a presença significa que o jogo é o único que interfere nas emoções do jogador.

Cairns descobriu que os gêneros mais imersivos foram *First Person Shooters*, isto é, jogos em primeira pessoa, e *Role Playing Games* (RPG), em que os participantes assumiam um personagem que escolhiam ser. Isso significa que a empatia (a capacidade de entender e compartilhar os sentimentos dos outros) é um importante fator imersivo. Outro fator crucial é a atenção do jogador. A maioria dos jogos lida com três estágios da atenção: visual, auditivo e mental. O nível de imersão sentido pelos jogadores parece correlacionar-se com o número de fontes de atenção necessário, bem como a quantidade de cada tipo de atenção. Se os jogadores precisam prestar atenção às suas ações e ao que acontece ao redor deles, maior a energia e investimentos necessários para se jogar e, conseqüentemente, mais imerso o jogador pode se sentir. A imersão total não é possível se houver falhas nos controles.

A imersão é um conceito compartilhado entre os jogadores, mas ainda é altamente subjetivo. Não é uma característica necessária para o entretenimento pois os jogadores escolhem os jogos dependendo do humor. É vista majoritariamente como uma experiência positiva, levando em consideração que ninguém quer sentir que está perdendo tempo em um jogo. A verdade é que compreender de fato o que torna uma experiência imersiva ainda é incerto. De toda a forma, o que se tem certeza é necessário para o prosseguimento desse estudo. A imersão é um sentimento universal e poderoso. Espera-se que esse estudo facilite o entendimento e *design* focado na imersão no futuro dos jogos.

2.9 Cybersickness

A *cybersickness* (CS) é o nome dado ao desconforto corporal associado à exposição ao conteúdo de realidade virtual. Existem algumas teorias comuns que explicam a CS, com a mais prevalente sendo incompatibilidade sensorial: se houver uma incompatibilidade entre o que um indivíduo sente (ouvido interno, olhos) e o que ele experiencia, esse indivíduo poderá sentir enjoo [81].

As principais fontes da CS foram exploradas nas últimas décadas. Em termos gerais, as fontes da CS podem ser categorizadas em hardware, software e usuário. Hardware inclui resolução e tamanho da tela, monitores, etc. Software se refere à aplicação em si, como a velocidade de movimento do personagem, gráficos e interface com o usuário. A terceira categoria diz respeito a fatores individuais: idade, sexo, experiência com realidade virtual, entre outros.

Embora avançada, a tecnologia de realidade virtual ainda peca na credibilidade da simulação do ambiente virtual. Desde sua invenção, desenvolvedores tentam gerar a sensação consistente e resiliente de presença nos usuários de VR. A outra segunda barreira proeminente é a *cybersickness*, previsivelmente. Liberar o potencial de VR dependerá de compreender e então solucionar esses obstáculos. Um grande corpo de pesquisa surgiu de tentativas de identificar se a presença e a CS são deterministicamente ligados. Os resultados são altamente discordantes, pois não existe nenhum consenso atualmente sobre a natureza do relação entre SC e presença [82].

As causas da *cybersickness* são bem conhecidas, porém ainda há muitas dúvidas que permeiam suas ramificações individuais. Enquanto muitos indivíduos experimentam CS em VR, outros parecem ser robustos aos sintomas. Uma vez que não há um bom modelo de estimativa da CS, se um desenvolvedor quer saber se um aplicativo e sistema pode causar CS, ele precisa realmente implementá-lo e testar os efeitos dele nos jogadores. De fato, a *cybersickness* é difícil de medir e quantificar e os resultados de diferentes laboratórios são difíceis de reproduzir e comparar, o que com certeza dificulta o desenvolvimento de um modelo preditivo da *cybersickness*.

Mais ainda, cerca de 75% do campo utiliza uma forma subjetiva de quantificação da *cybersickness*: o *Simulation Sickness Questionnaire* (SSQ) [87], criticado por suas qualidades psicométricas e aplicabilidade no contexto da realidade virtual. Recentemente, duas variantes do SSQ foram oferecidas para medir a CS: o *Cybersickness Questionnaire* (CSQ) [88] e o *Virtual Reality Sickness Questionnaire* (VRSQ) [89]. Ao comparar o SSQ com suas variantes, Sevinc et al. concluiu que ambos VRSQ e CSQ foram medidas válidas e confiáveis de CS, revelando melhores indicadores de validade que o questionário original [88].

Com o objetivo de se remover (ou pelo menos diminuir) a subjetividade da quantificação de *cybersickness*, várias pesquisas incorporaram os dados fisiológicos, relacionando a CS ao aumento da frequência cardíaca e na respiração, aumento da condutância da pele, aumento da taquiarritmia gástrica, a mudanças na atividade do córtex cerebral humano e à frequência do piscar dos olhos [82, 83]. Esses resultados sugerem que a CS acompanha as mudanças de padrão nas atividades dos sistemas nervoso central e autônomo.

Capítulo 3

Desenvolvimento do Trabalho

Neste capítulo, será inicialmente apresentada uma breve discussão sobre alguns estudos preliminares que foram fundamentais para as decisões tomadas neste trabalho. Em seguida é apresentada uma visão geral da proposta, passando por todas as etapas do desenvolvimento, delineamento do jogo e da pesquisa, além de justificar as decisões tomadas neste contexto. Adicionalmente, discorre-se acerca das ferramentas necessárias para a concepção do trabalho.

3.1 Estudos Preliminares

Vários estudos de graduação anteriores a este no âmbito do Departamento de Computação da UnB utilizaram o sensor vestível Empatica E4 e o EDA para inferir estados emocionais nos usuários. A seguir, serão destacados alguns destes:

Cayres [73] investigou um método de estimação da dificuldade de níveis de jogos do gênero plataforma e se um mecanismo híbrido do Ajuste Dinâmico de Dificuldade (ADD) possibilita adequar a dificuldade ao jogador e mantê-lo em estado de fluxo, além de comparar sua eficiência com os outros modelos. Para isso, um jogo foi desenvolvido para se adaptar com base aos dados extraídos por algoritmos de análise de desempenho do jogador correlacionados aos obtidos por um sensor de captura de dados fisiológicos, mais especificamente pela Atividade Eletrodérmica (EDA). Além de jogar com os distintos modelos de ADD, cada participante respondeu questionários e teve seus dados coletados para averiguação dos objetos de pesquisa. O modelo híbrido demonstrou-se capaz de manter o jogador em estado de fluxo e adequar a dificuldade ao jogador, com resultados superiores aos demais modelos.

Lopes et al. [74] conduziu uma análise comparativa entre modelos de ADD em um jogo do gênero Plataforma/Shooter, com o objetivo de averiguar se um sistema Híbrido é mais eficiente em proporcionar desafios compatíveis às habilidades dos jogadores e em

mantê-los em estado de imersão que sistemas não adaptativos ou que consideram apenas o desempenho do jogador. Para tanto, foi adaptado um jogo para coletar dados de desempenho, de questionários e fisiológicos dos jogadores, além da implementação de um modelo de ADD por Desempenho e um Híbrido a este. Foram realizados testes com os modelos, coletando diversos dados dos participantes. Após análise, o modelo Híbrido demonstrouse mais eficiente em proporcionar desafios compatíveis às habilidades dos jogadores e em mantê-los em estado de imersão.

Bertanini [75] propôs uma ferramenta para auxiliar na compreensão das reações dos jogadores por meio da inferência de estados afetivos a partir dos dados fisiológicos obtidos do usuário, que permite que os pesquisadores utilizem a ferramenta em qualquer tipo de jogo mesmo que seja necessário a utilização de um óculos, utilizado nos jogos de realidade virtual. Esta ferramenta pode ser treinada utilizando parâmetros do pesquisador, podendo se adaptar a diversos estudos a respeito da emoção, alcançando um uso diversas situações. Por meio de arquivos .srt o pesquisador pode utilizar as emoções reconhecidas como legendas em vídeos, facilitando o uso a qualquer tipo de vídeo, permitindo que o pesquisador escolha como deseja visualizar os dados adquiridos pela ferramenta. Além disso, salienta-se que a ferramenta não está limitada a um jogo, permitindo seu uso a diversas situações como reconhecer a emoção em uma conversa, durante uma sessão de terapia, durante a realização de uma prova, durante o sono, cabendo ao pesquisador utilizar da forma que lhe atenda melhor. Foram realizados, ainda, testes para exemplificar como a ferramenta pode vir a ser utilizada e são demonstradas algumas possíveis análises feitas com a ferramenta. Os testes demonstraram o uso de jogos e vídeos para análises qualitativas visuais além de análises numéricas por meio dos dados de um arquivo .SRT.

Mota [76] implementou e avaliou estas três abordagens de forma qualitativa em um jogo de realidade virtual com *biofeedback*. Foram conduzidos experimentos práticos com voluntários que jogaram um jogo desenvolvido com a implementação destas abordagens. Concluiu-se que a abordagem de mecânica foi a que mais afetou os jogadores, e demonstrou um potencial mediano para controle emocional do jogador. A abordagem de ambiente foi a que menos afetou os jogadores, mas demonstrou o maior potencial para controle emocional. A abordagem de recompensa afetou os jogadores de forma mediana e foi a pior abordagem para controle emocional.

O aspecto multidisciplinar deste trabalho, tais como a metodologia seguida por cada um dos estudos supracitados, acabaram por levantar várias questões no contexto da computação afetiva:

1. Os modelos de classificação dimensional se mostraram mais precisos do que os modelos de classificação discreta. Contudo, não foram encontradas comparações entre os modelos dimensionais. Seria interessante determinar qual dos modelos é o mais

apropriado para o contexto de computação afetiva, jogos, biofeedback e realidade virtual.

2. O consenso é que apenas o componente tônico do EDA não é tão informativo. Utilizar apenas o componente fásico traria quais tipos de resultados? E se fossem utilizados juntos?
3. A grosso modo, mais dados fisiológicos geralmente significam mais precisão na extração emocional. Existiria alguma combinação mais precisa do que outra? Ou talvez um número máximo de dados fisiológicos que aumentariam a precisão? É possível conduzir um teste que utiliza a mesma tecnologia de detecção, mas usando diferentes dados fisiológicos.
4. A literatura recomenda que o sinal da condutância da pele seja coletado seguindo algumas regras. Quanto à isso:
 - Inalar ou expirar excessivamente/profundamente, assim como prender a respiração ou falar, resulta em anomalias no EDA. Quanto a respiração impacta nos valores? Quanto esse artefato poderia atrapalhar um sistema de detecção de emoções, por exemplo? Quanto seria a diminuição da precisão?
 - A mesma pergunta se mantém para movimentos em geral. Artefatos de movimentos conseguem ser retirados sem impactar na precisão de reconhecimento emocional?
 - O período de relaxamento não deve apresentar nenhum tipo de estímulo. Caso tenha algum estímulo, é recomendável que se mostre um estímulo de valência de mesma magnitude, mas oposta. Qual é a diferença no EDA entre essas duas abordagens? Ainda mais, elas trazem algum benefício quando em comparação com um ambiente com estímulos?
 - Para experimentos que utilizem eventos manipulados para se induzir picos no EDA, é recomendável apresentar estímulos funcionais tempo suficiente e usar estímulos intermediários para que o sinal GSR para todos os respondentes possa “esfriar” e começar a partir da linha de base para o próximo estímulo. De quanto tempo estamos falando? Um segundo? Cinco? Talvez um minuto? Isso traria alguma diferença nos dados?
5. A literatura recomenda que o sinal da condutância da pele seja processado e interpretado de uma certa forma. A razão pela qual é dessa forma, no entanto, não está exatamente clara, principalmente quanto à parte de processamento do sinal.
 - Das normalizações encontradas, existe alguma mais adequada?

- Das padronizações encontradas, existe alguma mais adequada?
 - Como o RAT separa os componentes tônicos e fásicos do EDA? Como ele retorna um nível de excitação de 0 a 100? Esse método de cálculo condiz com a literatura a respeito?
6. Foram encontrados vários artigos que validavam o Empatica E4 como ferramenta eficiente para detecção emocional. Durante o desenvolvimento deste trabalho, no entanto, surgiram várias dúvidas quanto às limitações do sensor utilizado.
- Por que a recomendação de amostragem mínima da resposta galvânica é entre 200Hz-400Hz (Figner et al), mas o Empatica E4 suporta uma taxa amostral de apenas 4Hz? Algoritmos modernos para delinear e separar os componentes tônicos e fásicos (se ocorrerem em sucessão próxima) são melhor satisfeitos por uma maior amostragem do sinal (ou seja, métodos de desconvolução/desconvolução negativa), visto que são cálculos para áreas da curva. Muitas vezes é necessário suavizar um sinal devido a fatores como ruído, deflexões súbitas, etc. No entanto, algum sinal algoritmos de suavização requerem down-sampling. Com altas taxas de amostra, pode-se remover uma série de artefatos sem alterar a forma do sinal consideravelmente e ainda têm muitas remanescentes para uma análise adequada (o cenário ideal). Como regra geral, 500Hz - 2000Hz são mais do que suficientes para estudos de laboratório e facilmente alcançáveis para dispositivos contemporâneos como o MP36R. Em contrapartida, esses exemplos são destinados a ambientes controlados. Mas e em ambientes não controlados, como no monitoramento de um paciente? E em situações em que o usuário se movimenta mais, como em um contexto de jogos? Os 4Hz seriam suficientes?
 - A maior parte da literatura não dá certeza quanto ao EDA coletado pelo Empatica E4.
 - O Empatica E4 é realmente confiável na aquisição dos dados? Durante o experimento, percebeu-se erros de coleta na atividade cardíaca mesmo quando o sensor estava firmemente preso ao pulso do jogador. Como o BVP possui uma forma de onda bem característica, foi possível notar as falhas na coleta apenas observando a forma de onda crua do sensor. Como confiar que os outros dados coletados são confiáveis quando se sabe com certeza que um não foi coletado corretamente? Sabe-se que pessoas com temperatura corporal muito baixa ou sob uso de certas medicações que suprimem as respostas do sistema nervoso simpático não conseguem ter seus dados captados pelos eletrodos EDA. Contudo, essas limitações não são precisas o suficiente para que o Empatica E4

seja utilizado de forma que se consiga minimizar as variáveis que causem erros na coleta. É necessário um estudo que explicita e estude a fundo as limitações desse sensor.

É possível notar o nível de complexidade em responder tais questões. Cada uma delas merece um estudo devidamente projetado e controlado. Muitas das questões, além do mais, requerem comparações entre diferentes conjuntos de variáveis em diferentes contextos, o que está claramente fora do escopo deste projeto. De toda forma, estes levantamentos nortearam a realização do presente trabalho, principalmente aqueles que tangem o EDA e o Empatica E4. A visão geral será detalhada na seção a seguir.

3.2 Visão Geral da Proposta

Este trabalho tem o objetivo de investigar potenciais relações entre a atividade cardíaca (BVP) e dérmica (EDA) de um indivíduo e o mal-estar relacionado à realidade virtual. Existem vários estudos que correlacionam dados fisiológicos, como por exemplo Golding [85] e Gavagni et al. [86], os quais procuraram quantificar a CS utilizando a condutância da pele, obtendo correlação positiva. Essa relação é todavia inespecífica: Dennison et al. [84] relatou que o aumento da condutância da pele encontrado no estudo foi em grande parte devido ao aumento da excitação causada pela interação com VR, em vez da CS. Além disso, estudos que exploram o uso de dados fisiológicos obtidos por dispositivos portáteis no contexto de *cybersickness* não foram encontrados.

Para este trabalho, foi selecionado o SSQ como ferramentas de pontuação da reação à interação com o ambiente virtual. O SSQ foi selecionado por ser amplamente utilizado na área há anos, apesar das críticas que recebe. Os dados fisiológicos foram selecionados para complementar os resultados do SSQ e observar as correspondências entre ambos (os sinais levados em consideração serão explicados na seção seguinte). A simulação podia ser vista através *Head Mounted Display* (HMD) do HTC Vive (Figura 3.1). Buscou-se calcular a força das interações entre os sinais biológicos e os scores SSQ através do Coeficiente de Correlação de Pearson e o Coeficiente de Correlação de Spearman.

3.2.1 Roteiro do Experimento

De modo geral, o experimento consistiu em coletar os dados fisiológicos do participante durante três períodos de tempo, as quais serão chamadas de sessões. Para cada sessão, são coletados o BVP/IBI/HR, EDA, temperatura da pele e dados de movimento do jogador utilizando o Empatica E4. Os dados coletados são salvos no *E4 Connect*, onde podem ser

recuperados em formato CSV. As sessões serão explicadas a seguir, na ordem que foram realizadas no experimento.

1. **Relaxamento:** o jogador permanece sentado, imóvel, com o braço do sensor apoiado sobre uma mesa, usando um abafador de ruídos e com os olhos fechados durante cinco minutos;
2. **Triagem:** imediatamente antes da simulação virtual, os dados vitais são coletado por mais um minuto com o jogador na posição em que vai jogar;
3. **Corrida:** Os dados vitais são coletados durante os 10 minutos da simulação virtual.

Para a realização do experimento, foram necessários, além do sensor Empatica E4 (Figura 3.1):

1. Head Mount Display (HMD) do HTC Vive
2. Controle de Xbox
3. Abafador de Ruído
4. Fones de Ouvido



Figura 3.1: Em sentido horário: HMD do HTC Vive, controle de Xbox 360, fones de ouvido Audio Technica m50x e abafadores de ruído.

Um roteiro experimental foi criado de modo a maximizar o número de participantes e, ao mesmo tempo, proteger o protocolo de coleta, garantir a qualidade do sinal fisiológico

obtido e evitar possíveis enviesamentos. Para participar dele, o jogador assinou o Termo de Consentimento Livre e Esclarecido (TCLE), que se encontra no Apêndice A. No plano do experimento, estabeleceu-se uma ordem de passos para o aplicador do teste, desde as frases que serão faladas aos participantes até a apresentação dos objetos utilizados. Os experimentos foram realizados em conjunto com outro graduando da Engenharia da Computação, Gabriel Tollini, que utilizou a mesma base de dados em seu estudo. O roteiro do experimento está disponível no Apêndice B. Imediatamente antes e imediatamente depois da simulação virtual, o participante preenche o SSQ (Apêndice C).

3.2.2 Simulação Escolhida

A *cybersickness* é diretamente relacionada à ilusão de movimento. Por esse motivo, foi escolhido o *RaceRoom Racing Experience* (Figura 3.2), um simulador de corrida gratuito para Microsoft Windows desenvolvido pela *Sector3 Studios* e publicado pela *RaceRoom Entertainment AG*. Seu objetivo é fornecer uma experiência de corrida autêntica por meio de modelos detalhados de carros e pistas, bem como comportamentos e sons realistas dos carros. O jogo oferece uma seleção de carros de corrida *free-to-play* e pistas para dirigir nos modos *multiplayer* e *singleplayer*.



Figura 3.2: Ponto de vista do usuário durante o jogo. Fonte: Google Imagens. Acessado em 2/12/2020.

Durante a corrida, o jogador é instruído a se mexer e a falar o mínimo possível, a fim de evitar artefatos nos sinais biológicos, principalmente no EDA.

3.2.3 Processamento e Análise dos Sinais

Obtidos os sinais dos participantes, dá-se início a parte de processamento e análise dos sinais. Como explicado no Capítulo 2, não há um método universalmente acordado para padronizações e correções dos dados afetivos (principalmente no caso do EDA), mas o que se sabe é que algoritmos que utilizam valores mínimos e máximos propostos por Dawson [49] para dados fisiológicos são métodos tidos como cada vez mais controversos. Várias soluções para a fórmula de Dawson foram propostos, como a transformação de SRCs em Z-Scores (utilizando média e desvio padrão em vez do máximo) de Boucsein [49], ou a proporção padronizada de Ben-Shakar [90]. Atualmente, existem ferramentas mais modernas disponíveis, entre elas a decomposição convexa (utilizada pelo *NeuroKit2*), a decomposição esparsa (utilizada pelo *sparsEDA*), a decomposição contínua (utilizada pela *CDA-Ledlab*), a análise de decomposição discreta disponível em *DDA-Ledlab* e a modelagem causal dinâmica disponível em *psPM (DCM-psPM)*. As principais restrições práticas de tais métodos são a viabilidade de implementação dos cálculos matemáticos em sensores portáteis e a necessidade de calibração para cada indivíduo [48]. Levando tais pontos em consideração, o *NeuroKit2* foi escolhido para processamento do sinal. A biblioteca pode processar todos os dados biológicos obtidos pelo *Empatica E4* sem necessidade de adaptações no código, o que aconteceria na maioria das outras ferramentas. Mais detalhes sobre o *NeuroKit2* a seguir.

3.2.4 NeuroKit2

O *NeuroKit2* é uma biblioteca Python para processamento neurofisiológico de sinais. Ela permite que os dados fisiológicos sejam facilmente processados a partir da rotina *bio_process()*, desde que esse sinal esteja em formato de lista ou vetor. Um exemplo seria o processamento do EDA, que retornaria dois elementos: um dataframe contendo os diferentes sinais (por exemplo, o sinal bruto, sinal limpo, amostras de SCR marcando os diferentes recursos, etc.) e um dicionário contendo informações sobre os picos de Resposta de Condutância da Pele (SCR) (por exemplo, inícios, amplitude de pico, etc.)¹. A função de processamento faz duas coisas importantes: em primeiro lugar, ela limpa o sinal. Em segundo lugar, ele detecta a localização de 1) início de pico, 2) amplitude de pico e 3) tempo de meia recuperação.

O *neurokit* utiliza otimização convexa [78] para calcular automaticamente o EDA. O modelo é baseado em estatísticas Bayesianas e uma representação simples, mas fisiologicamente sólida, da condutância da pele observada como a soma de três componentes: um componente tônico lento, a saída da convolução entre uma função de resposta de

¹<https://neurokit2.readthedocs.io/en/latest/examples/eda.html>

impulso (IRF) e a atividade do nervo sudomotor (SMNA) fásico não negativo esparsos (compacto, intermitente) e um componente ruidoso. Greco et al. [78], responsáveis por proporem a otimização, confirmaram que a otimização convexa fornece uma decomposição do EDA que é robusta ao ruído, supera o problema de SCRs sobrepostos e fornece uma janela da atividade do sistema nervoso autônomo (ANS). Além disso, a solução incorpora o aspecto fisiológico dos componentes fásicos e tônicos, sem exigir etapas de pré ou pós-processamento.

Capítulo 4

Resultados

Neste capítulo, serão apresentadas as análises exploratórias a partir dos dados obtidos nos experimentos descritos no Capítulo 3. Tais análises buscam sanar as questões levantadas neste estudo, especialmente as que envolvem o Empatica E4 e os dados fisiológicos obtidos por ele na quantificação da *cybersickness*, com foco no EDA.

4.1 Testes Piloto

Foram conduzidos duas coletas de dados para detectar potenciais problemas com os protocolos do experimento. Elas seguiram o mesmo roteiro que os testes definitivos. A partir deles, pôde-se observar que:

1. Notou-se que o material esponjoso do HMD do HTC Vive absorvia o álcool em gel, causando irritação na pele ao redor dos olhos de quem estava jogando. A solução foi utilizar álcool em spray para higienizar os equipamentos imediatamente depois do uso.
2. Os fones de ouvido disponíveis estavam defeituosos. Eles reproduziam o som em apenas um dos dois reprodutores, o que poderia causar distrações consideráveis no jogador. Foi providenciado um outro fone de ouvido que funcionasse.
3. Percebeu-se uma oportunidade de agilizar as sessões dos experimento ao inicializar de antemão os programas necessários para a coleta de dados.
4. É notável a presença de movimento durante o jogo. Contudo, os acelerômetros indicam que o movimento é mínimo, já que é muito próximo dos valores das sessões de relaxamento (em que a pessoa pode ficar completamente parada). Os acelerômetros podem detectar movimento e conseqüentemente possíveis artefatos durante a coleta dos dados.

5. É notável a diferença de valores absolutos no EDA dos dois testes. O primeiro não passava de 0.2 uS, enquanto o segundo chegava a 0.7uS. Isso confirma a necessidade de se padronizar os valores do EDA para comparação tanto entre os valores do mesmo participante quanto quando comparando-os com os demais.
6. Mesmo que a magnitude dos valores tenha variado bastante entre as duas coletas, ambos apresentaram valores extremamente estáveis durante a sessão de relaxamento. Com excessão do EDA de um dos testes, que apresentou uma descida linear lenta, apesar de também estável. Isso é um indicador de que tais sessões apresentam período adequado para uma boa linha de base.
7. Durante a corrida, os dois EDAs apresentaram leve descida seguida de leve subida mais ao final da sessão. Se isso acontecer em outros participantes, seria interessante descobrir o porquê.
8. O impacto de dos outros valores (BVP, BPM e temperatura da pele) no EDA é incerto. Pelo menos na primeira análise dos dados (majoritariamente visual), não se vê relação direta desses valores com o BVP. Contudo, a literatura diz que outros dados são importantes para confirmar respostas GSR e retirar artefatos, assim como facilitar a detecção o estado emocional do jogador.

4.2 Testes Definitivos

Os testes finais foram conduzidos no Laboratório (LAICO) ao decorrer de 4 dias. No total, foram obtidos dados vitais de 52 pessoas, além de idade e sexo de cada jogador (retirados dos SSQs). A seguir, algumas observações feitas nos dias das coletas.

1. De forma geral, o Empatica E4 facilitou a aquisição dos dados. Contudo, algumas sessões foram perdidas simplesmente porque o sensor perdia a conexão bluetooth com o aplicativo. É uma falha grave na arquitetura de persistência dos dados. O ideal é que o sensor continue obtendo os dados mesmo sem conexão, visto que ele possui uma memória interna.
2. Os participantes tiveram muita dificuldade durante o jogo de saber a direção correta da pista. Isso foi facilmente resolvido explicando a eles que existia um marcador na esquerda superior da tela mostrando se o jogador estava ou não na contra-mão.
3. Outra dificuldade clara foi quanto à ré. O carro demorava alguns segundos para engatar a ré e isso confundia os participantes, que provavelmente estavam acostumados com jogos que respondiam imediatamente.

4. Muitos jogadores saíam da pista nas curvas mais fechadas, pois o jogo não indicava quando ou para onde virar. Isso também foi fácil resolver explicando que existia um mini-mapa na região inferior esquerda da tela indicando o layout da pista.
5. Os SSQs permitiam que a pessoa deixasse sintomas em branco, o que foi prontamente corrigido no Google Forms.
6. Observou-se que a maioria das pessoas não se sentiram mal após a simulação, mas quase sempre apresentavam algum sintoma leve após o uso do VR. O mais comum, na verdade, foi agravar algum sintoma já existente (exemplo: uma pessoa com dor de cabeça reclamava da dor de cabeça ter piorado após a sessão do experimento).
7. Em vários participantes, os dados afetivos coletados pelo Empatica E4 apresentavam artefatos claros, mesmo com a pulseira firmemente presa ao pulso. A causa mais provável para isso é a sensibilidade do sensor a movimentos. Outra explicação pode ser o que já tinha sido observado em estudos anteriores: pessoas com temperatura corporal muito baixa ou sob uso de certas medicações que suprimem as respostas do sistema nervoso não conseguem ter seus dados captados.

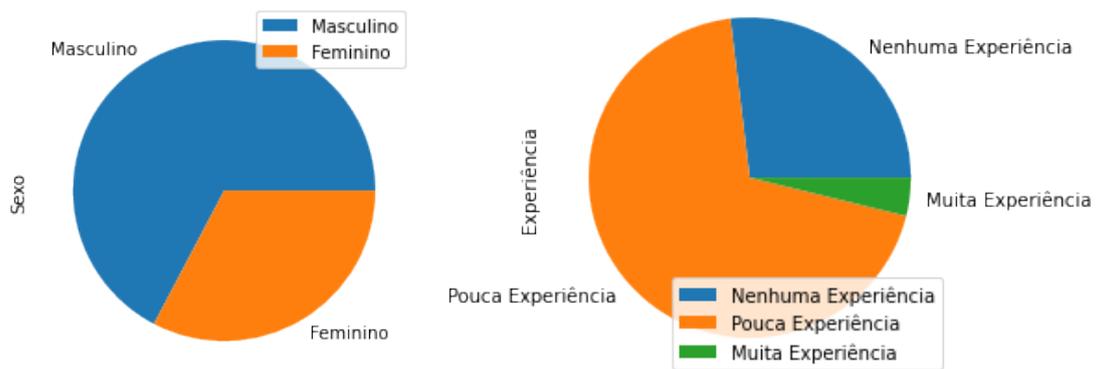
4.3 Resultados

Em resumo, a média de idade de todo o grupo foi de 26,96 anos. A maioria dos participantes foram do sexo masculino (67.3%), totalizando 35 participantes, e apenas 17 (32.7%) do sexo feminino; 96% (50 participantes) dos participantes relataram possuir nenhuma ou pouca experiência com VR. Por fim, apenas 12 (23.1%) interromperam o experimento. A Figura 4.1 ilustra a distribuição de participantes por idade, sexo e nível de experiência com a realidade virtual.

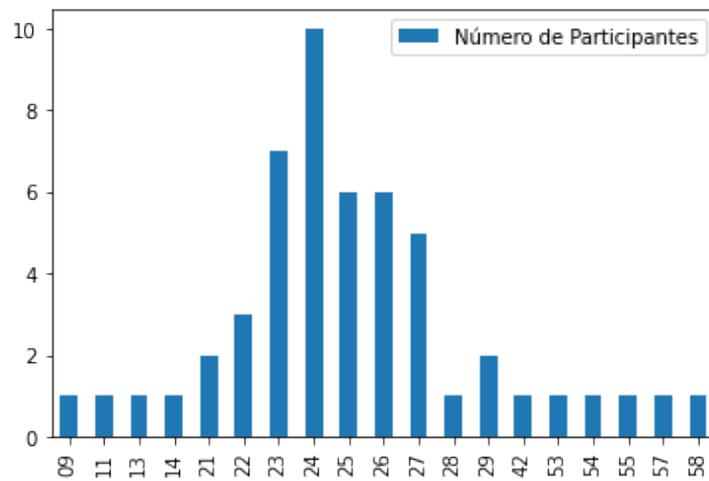
4.3.1 Scores SSQ

Os *scores* dos SSQs foram calculados a partir da fórmula proposta por Kennedy et al [87]. A Figura 4.2 ilustra o número de participantes para cada resultado do segundo SSQ subtraído do primeiro SSQ. Para uniformidade, foi adicionada uma coluna nula aos gráficos de distribuição. Como se pode perceber, alguns *scores* foram negativos, ou seja, alguns jogadores relataram se sentirem melhor após o experimento.

Pelas distribuições, é possível calcular a média dos *scores* de cada um dos questionários. Levando todos os participantes em consideração, a média do SSQ foi de 16.97 com desvio padrão 28.56. Separando o grupo entre aqueles jogadores que completaram o experimento e aqueles que não, é possível notar que a média dos valores do segundo grupo é 5.4

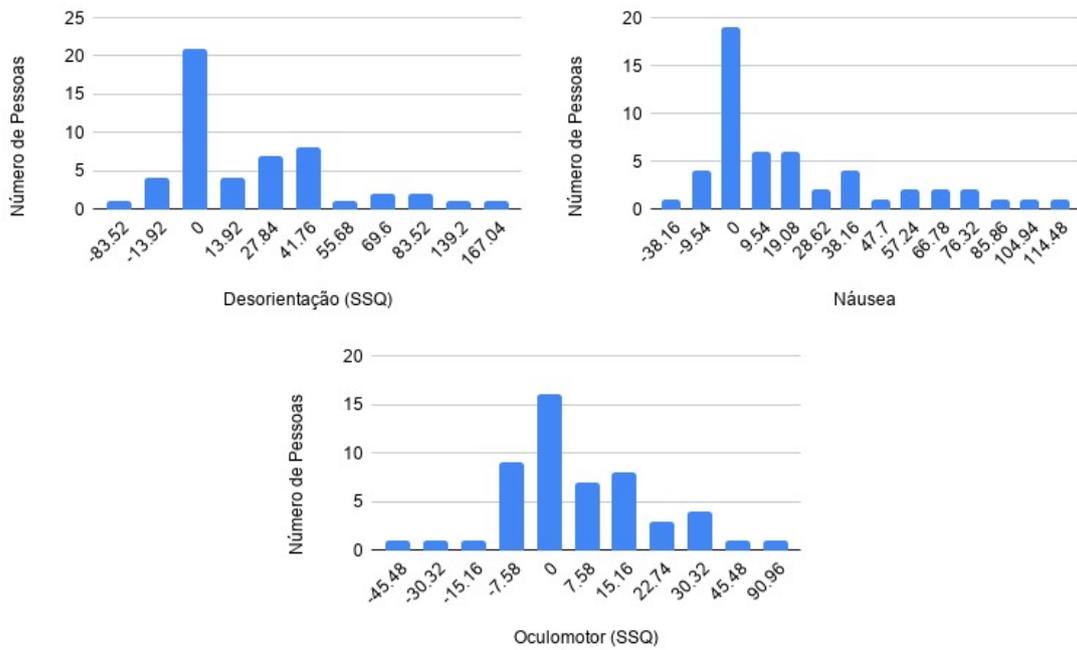


(a) Distribuição dos participantes por sexo e experiência com VR.

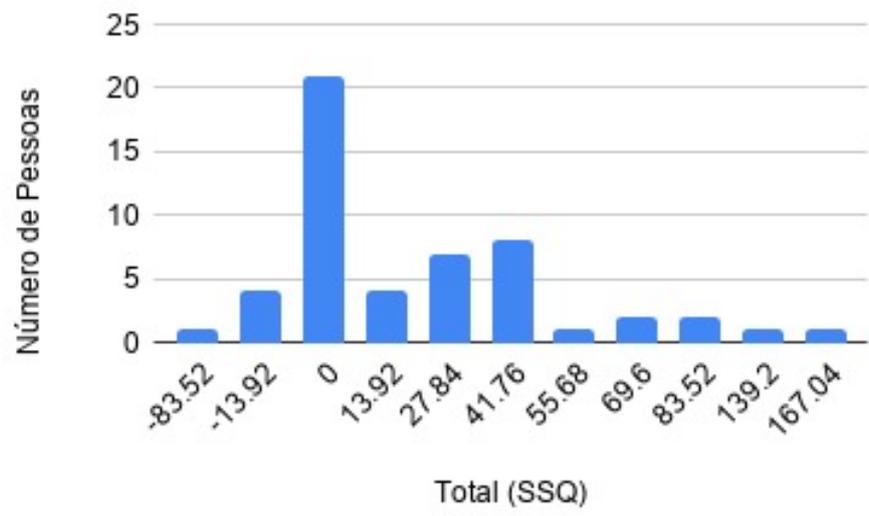


(b) Distribuição dos participantes por idade.

Figura 4.1: Dados Demográficos



(a) Resultado de cada grupo do SSQ.



(b) Distribuição do SSQ Total entre os jogadores.

Figura 4.2: Scores SSQ

Variáveis ANS	5 Min	1 Min	10 min	1º Quartil	2º Quartil	3º Quartil	4º Quartil
SCL (uS)	2.30 (0.38)	2.95 (0.14)	4.26 (0.50)	3.71 (0.52)	4.21 (0.39)	4.44 (0.32)	4.70 (0.27)
BVP (-)	13.16 (19.68)	11.19 (21.19)	15.90 (28.84)	15.55 (28.13)	16.14 (28.18)	15.33 (27.24)	16.61 (29.08)
IBI (s)	0.72 (0.05)	0.65 (0.05)	0.64 (0.07)	0.6 (0.07)	0.63 (0.06)	0.60 (0.06)	0.55 (0.42)
Freq. Cardíaca (Hz)	78.78 (11.26)	77.21 (16.69)	79.2 (22.80)	79.51 (23.28)	79.00 (22.85)	78.60 (22.10)	79.80 (21.64)
Temperatura da Pele (°C)	32.14 (0.18)	32.40 (0.03)	32.46 (0.12)	32.44 (0.04)	32.45 (0.04)	32.46 (0.04)	32.49 (0.04)

Tabela 4.1: Valores Médios (e desvios padrão) dos dados fisiológicos obtidos pelo Empatica E4 dos jogadores que completaram a sessão de corrida. SCL - Nível de Condutância da Pele. BVP - Pulso de Volume Sanguíneo. IBI - Intervalo entre Batimentos.

vezes maior: 45.5 com desvio padrão 30.76, enquanto o grupo dos que interromperam o experimento obtiveram média 8.415 com desvio padrão 21.43. Tendo em vista estes resultados, os participantes foram categorizados entre aqueles que não interromperam o experimento (chamado de Grupo 1) e aqueles que interromperam (Grupo 2), visando encontrar potenciais padrões nos dados fisiológicos de cada grupo. O Grupo 1 incluiu 40 participantes; o Grupo 2, 12 participantes.

4.3.2 Dados Afetivos

A Tabela 4.1 ilustra os valores médios e desvios padrão de cada sinal obtido pelo Empatica E4 dos jogadores que completaram a sessão. Nesta tabela e nas demais, cada coluna representa um período de tempo do experimento; a sessão de relaxamento está representada como “5Min”; a sessão de triagem, “1Min”; a sessão de corrida, “10Min”, que também foi dividida em quatro quartis, de forma que o primeiro quartil representa os primeiros 25% da sessão de corrida (0-2.5min), o segundo representa de 2.5-5min, e assim por diante. Uma representação visual dos quartis pode ser vista nas Figuras 4.4 e 4.5. Todos os valores foram calculados a partir do NeuroKit2, com exceção da temperatura da pele e do IBI, os quais não necessitavam de filtragem ou processamento.

A Tabela 4.2 apresenta os valores do grupo que interrompeu a simulação virtual. Todas as 12 sessões foram levadas em consideração. Vale ressaltar que todos os participantes descontinuaram especificamente a sessão de corrida, período em que o HTC Vive era utilizado. No momento da interrupção, um evento era marcado no Empatica E4. A maioria das interrupções ocorreu a partir do terceiro quartil, como pode ser visto na Figura 4.3, indicando que quanto mais tempo passa, maior é a chance da simulação ser interrompida. Algumas pessoas interromperam o experimento depois dos dez minutos de corrida. Esse número foi adicionado às interrupções do quarto quartil.

Como cada quartil tem a duração exata de 2.5 minutos, tal resultado pode sugerir um certo período de tempo para que a pessoa comece a realmente a sentir os sintomas, e que esses sintomas provavelmente se agravam à medida que o tempo passa. Contudo, como

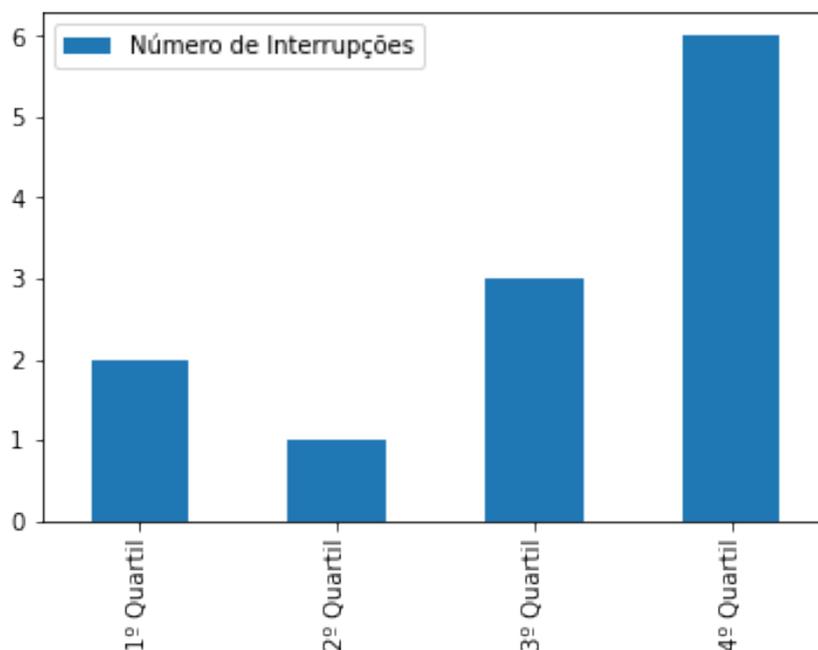


Figura 4.3: Número de interrupções por quartil.

Variáveis ANS	5 Min	1 Min	10 min	1º Quartil	2º Quartil	3º Quartil	4º Quartil
SCL (uS)	1.96 (0.31)	3.02 (0.10)	5.28 (1.61)	3.37 (0.46)	4.42 (0.78)	6.3 (0.33)	10.03 (0.43)
BVP (-)	12.96 (17.47)	17.41 (28.83)	15.46 (28.46)	13.83 (24.90)	14.69 (25.92)	16.76 (32.05)	19.39 (33.18)
IBI (s)	0.85 (0.06)	0.82 (0.06)	0.78 (0.07)	0.78 (0.07)	-	-	-
Freq. Cardíaca (Hz)	72.30 (10.17)	72.48 (15.99)	74.53 (20.92)	72.82 (20.64)	74.89 (21.26)	77.63 (21.36)	80.21 (19.36)
Temperatura da Pele (°C)	32.10 (0.16)	32.42 (0.03)	32.43 (0.22)	32.43 (0.03)	32.51 (0.11)	32.44 (0.37)	32.96 (0.03)

Tabela 4.2: Valores Médios (e desvios padrão) dos dados fisiológicos obtidos pelo Empatica E4 pelo grupo que interrompeu a simulação. Os valores IBI do segundo, terceiro e quarto quartis não foram captados pelo Empatica E4. SCL - Nível de Condutância da Pele. BVP - Pulso de Volume Sanguíneo. IBI - Intervalo entre Batimentos.

o grupo de pessoas que interromperam o experimento é pequeno, é difícil confirmar um padrão.

Os valores médios do componente tônico (que se relacionam com os componentes de ação mais lenta e características de plano de fundo do sinal) de ambos os grupos aumentou com o passar do tempo, o que é um indicativo de aumento no nível de excitação dos jogadores ao decorrer do experimento. A provável explicação é a imersão e a novidade da experiência, que muito provavelmente causaram respostas intensas nos jogadores. A atividade cardíaca (BVP, IBI e frequência cardíaca) dos dois grupos também aumentou com o passar dos quartis, mas a diferença nos valores é pequena. Quanto ao EDA, as tabelas indicam uma correlação positiva entre o CS e o SCL (os cálculos de correlação serão apresentados na seção seguinte), visto que as amplitudes do componente sobem muito mais rapidamente no Grupo 2 (que sentiu muito mais o mal-estar do que o Grupo 1) na sessão de 10 minutos.

Variáveis ANS	5 Min	1 Min	10 min	1º Quartil	2º Quartil	3º Quartil	4º Quartil
Nível SCR (uS)	0.001 (0.05)	-0.001 (0.07)	0.000 (0.13)	0.000 (0.16)	-0.001 (0.13)	-0.001 (0.13)	0.001 (0.13)
Picos SCR (qnt.)	56.91 (52.95)	13.46 (11.84)	77.08 (53.92)	17.86 (13.71)	19.54 (14.74)	20.14 (14.80)	19.54 (13.80)

Tabela 4.3: Valores Médios (e desvios padrão) do componente fásico obtido pelo Empatica E4 do grupo de jogadores que completaram o experimento.

Variáveis ANS	5 Min	1 Min	10 min	1º Quartil	2º Quartil	3º Quartil	4º Quartil
Nível SCR (uS)	0.001 (0.05)	-0.014 (0.1)	0.000 (0.17)	0.001 (0.14)	-0.001 (0.12)	-0.003 (0.17)	0.01 (0.18)
Picos SCR (qnt.)	73.58 (54.08)	13.66 (7.29)	39.33 (39.84)	11.3 (9.06)	11.9 (10)	16.22 (12.07)	15.66 (16.36)

Tabela 4.4: Valores Médios (e desvios padrão) do componente fásico obtido pelo Empatica E4 do grupo de jogadores que não completaram o experimento.

As Tabelas 4.3 e 4.4 foram construídas a partir dos componentes fásicos do EDA de cada grupo. As amplitudes fásicas não são muito informativas e bem mais baixas do que valores reportados em estudos anteriores. Quanto ao número de picos SCR, é possível perceber números maiores de picos para o grupo daqueles que interromperam o experimento na sessão de linha base (5min). Esses picos são provavelmente respostas espontâneas da pele (NS-SCRs), já que os estímulos durante esses períodos eram mantidos no mínimo. Esse fato pode indicar certa propensão de pessoas que possuem mais NS-SCRs com a *cybersickness*. No entanto, não há informações sobre os eventos de cada sessão de corrida, o que inviabiliza a separação dos picos SCR entre NS-SCRs e ES-SCRs. Com a exceção da sessão de cinco minutos, o número de picos é consideravelmente menor para o Grupo 2. O cálculo que gerou esse resultado, no entanto, não levou em consideração a duração reduzida dos quartis do Grupo 2, o que pode diminuir artificialmente o número de picos.

Para retirar o problema da duração dos quartis, calculou-se a frequência dos picos SCR por segundo por quartil. Os resultados, no entanto, não mudam muito, como pode ser visto na Tabela 4.5. Após uma nova inspeção, foi descoberto que o número de picos detectados em várias instâncias do grupo 2 foi zero, o que, da mesma forma que o problema da duração dos quartis, reduz os valores. É difícil dizer se esses estão corretos ou não. É muito provável que a aparente incoerência ocorreu devido a erros na leitura do sinal ou no algoritmo de detecção de picos do NeuroKit2. Porém, mesmo retirando tais instâncias manualmente, os valores não mudaram significativamente. É possível que o Grupo 2 realmente tenha apresentado respostas fásicas reduzidas em comparação com o Grupo 1, o que teoricamente significaria que os jogadores presenciaram tanto menos NS-SCRs quanto ES-SCRs (por batidas, sair da pista, etc). Não se sabe, contudo, a proporção NS/ES dessa redução. De toda forma, caso essa hipótese seja verdadeira, a *cybersickness* teria uma relação mais forte com o componente tônico do EDA comparado ao componente fásico.

Variáveis ANS	5 Min	1 Min	10 min	1º Quartil	2º Quartil	3º Quartil	4º Quartil
Picos SCR/s - Grupo 1	0.19	0.21	0.13	0.12	0.13	0.13	0.13
Picos SCR/s - Grupo 2	0.25	0.23	0.02	0.02	0.02	0.03	0.03

Tabela 4.5: Média Picos SCR por segundo por sessão de cada grupo.

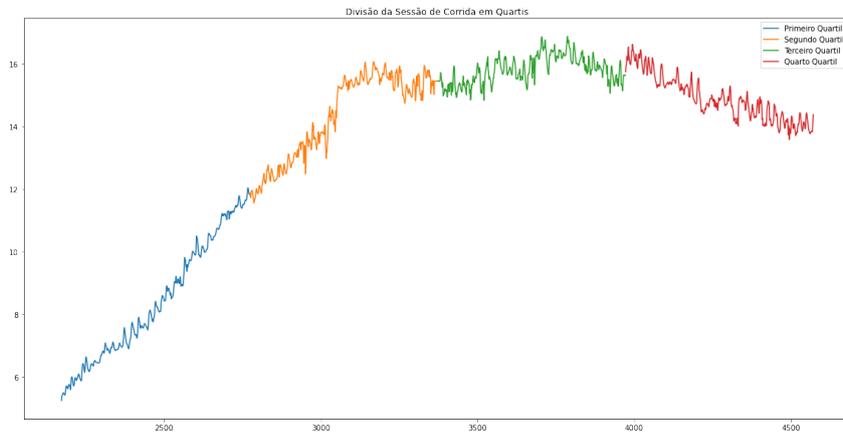


Figura 4.4: Representação visual da divisão de uma sessão de jogo em quartis.

Para diferenciar as sessões uma das outras, foi utilizado o sistema de eventos do Empatica E4, que permitia guardar o momento em que o botão era pressionado. Apertar fisicamente o botão foi um erro, pois causou artefatos intensos de movimento. Além disso, o sistema não possui detalhe o suficiente para distinguir um evento do outro, fato que implicava correção manual quando havia um número inesperado de eventos marcados nas sessões.

4.4 Cálculos de Correlação

Visando confirmar a potencial relação entre a *cybersickness*, a atividade cardíaca e o EDA encontrado na seção anterior, o Grupo 2 foi analisado com mais detalhe. Desses valores, foram calculados os coeficientes de correlação de Pearson e de Spearman entre cada relação. O motivo de utilizar os dois coeficientes serão explicitados a seguir. A Tabela 4.6 resume os valores encontrados para os doze jogadores do Grupo 2.

4.4.1 Correlação de Pearson

O coeficiente de correlação de Pearson (nomeado em homenagem a Karl Pearson) pode ser usado para resumir a força da relação linear entre duas amostras de dados. Ele é calculado como a covariância das duas variáveis dividida pelo produto do desvio padrão (stdv) de cada amostra de dados. É a normalização da covariância (cov) entre as duas

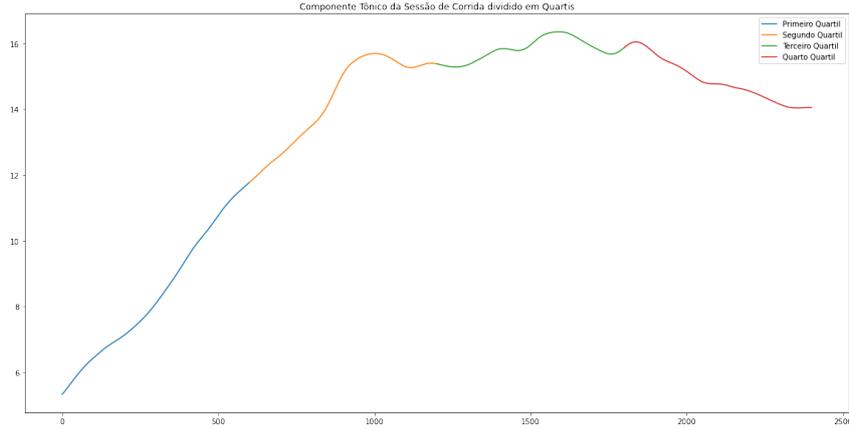


Figura 4.5: Representação visual da divisão de uma sessão de jogo em quartis (componente tônico extraído do Neurokit2).

Jogador	Score SSQ	Nível SCL (uS)					Picos SCR (qnt.)				
		Total	1º Quartil	2º Quartil	3º Quartil	4º Quartil	Total	1º Quartil	2º Quartil	3º Quartil	4º Quartil
1002106	52.36	0.51 (0.26)	0.42 (0.04)	0.62 (0.36)	-	-	7.00	2.00	5.00	-	-
1002077	41.14	0.35 (0.23)	0.11 (0.01)	0.26 (0.16)	0.58 (0.08)	0.58 (0.06)	12.00	0.00	1.00	8.00	3.00
1004413	37.4	0.31 (0.04)	0.30 (0.02)	0.30 (0.01)	0.33 (0.06)	-	13.00	5.00	4.00	4.00	-
1002245	22.44	0.17 (0.11)	0.10 (0.01)	0.21 (0.13)	0.21 (0.13)	-	8.00	0.00	3.00	5.00	-
1005326	41.14	15.80 (2.83)	15.80 (2.83)	-	-	-	29.00	29.00	-	-	-
1003328	29.92	0.38 (0.07)	0.38 (0.07)	-	-	-	2.00	2.00	-	-	-
1007039	22.44	0.27 (0.08)	0.23 (0.01)	0.24 (0.00)	0.32 (0.12)	0.28 (0.06)	21.00	4.00	3.00	9.00	5.00
1004120	0	2.45 (0.28)	2.17 (0.13)	2.73 (0.12)	2.44 (0.23)	2.37 (0.01)	60.00	9.00	26.00	24.00	1.00
1005507	48.62	2.00 (0.64)	1.38 (0.35)	2.24 (0.32)	2.66 (0.47)	-	28.00	9.00	12.00	7.00	-
1007738	52.36	24.99 (8.88)	12.01 (1.43)	22.76 (4.25)	31.31 (1.33)	33.88 (0.93)	129.00	22.00	26.00	37.00	44.00
1007108	67.32	7.95 (2.98)	3.38 (0.24)	7.73 (1.59)	9.93 (0.22)	10.76 (0.25)	81.00	15.00	20.00	25.00	21.00
1007728	130.9	8.16 (3.00)	4.34 (0.41)	7.13 (0.84)	8.89 (0.35)	12.29 (1.29)	82.00	16.00	19.00	27.00	20.00

Tabela 4.6: Relação entre os scores SSQ e valores médios (e desvios padrão) do Nível SCL e picos SCR do Grupo 2.

variáveis (X,Y) para dar uma pontuação interpretável. No caso deste estudo, a primeira amostra é o conjunto de scores SSQ do Grupo 2, e a segunda mostra é a média nível tônico ou o número de picos SCR da sessão de 10 minutos, seguindo a fórmula:

$$r = cov(X, Y) / (stdv(X) * stdv(Y)) \quad (4.1)$$

O coeficiente retorna um valor entre -1 e 1 que representa os limites de correlação de uma correlação totalmente negativa para uma correlação totalmente positiva. Um valor de 0 significa nenhuma correlação. O valor deve ser interpretado, onde frequentemente um valor abaixo de -0,5 ou acima de 0,5 indica uma correlação notável e valores abaixo desses valores sugerem uma correlação menos notável.

O coeficiente de Pearson entre os scores SSQ e o nível SCL no Grupo 2 é de 0.31 ($p > 0.05$), o que sugere uma correlação positiva fraca. Da mesma forma, obteve-se $r = 0.43$ ($p > 0.05$) entre os scores SSQ e número de picos SCR, o que contradiz os resultados encontrados na seção anterior. Uma possível causa para esses resultados é, novamente, erros na detecção de picos do algoritmo causado pelos artefatos do sinal coletado. No

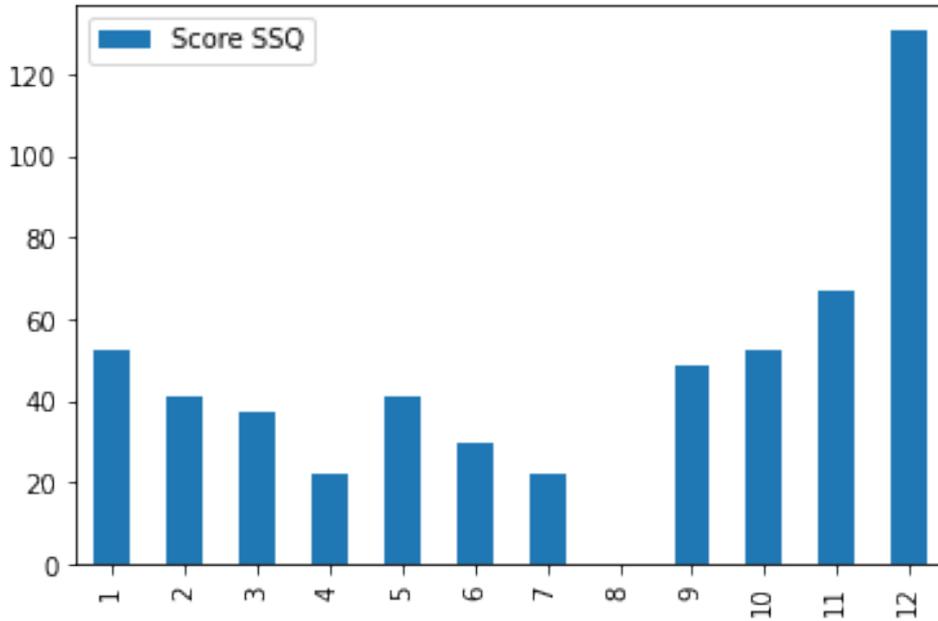


Figura 4.6: Scores SSQ de cada integrante do Grupo 2.

entanto, foi encontrada correlação forte ($r=0.79$, $0.001 < p < 0.01$) para o nível SCL e o número de picos SCR, o que indica que as respostas fásicas tendem a aumentar o nível geral de excitação do jogador. Não houve indicação de correlação entre os scores SSQ e os demais dados afetivos coletados ($\text{abs}(r) < 0.2$, $p > 0.05$).

4.4.2 Correlação de Spearman

Duas variáveis podem estar relacionadas por uma relação não linear, de forma que a relação seja mais forte ou mais fraca na distribuição das variáveis. Neste caso, o coeficiente de correlação de Spearman pode ser usado para resumir a força entre as duas amostras de dados. Este teste de relacionamento também pode ser usado se houver uma relação linear entre as variáveis, mas terá um pouco menos poder (por exemplo, pode resultar em pontuações de coeficiente mais baixas). Tal como acontece com o coeficiente de correlação de Pearson, as pontuações estão entre -1 e 1 para variáveis perfeitamente correlacionadas negativamente e perfeitamente correlacionadas positivamente, respectivamente. Em vez de calcular o coeficiente usando covariância e desvios padrão nas próprias amostras, calcula-se a classificação relativa de valores em cada amostra (rank). Esta é uma abordagem comum usada em estatísticas não paramétricas, por exemplo métodos estatísticos em que não assumimos uma distribuição dos dados como gaussiana. O coeficiente de Spearman é calculado da seguinte forma:

$$s = \text{cov}(\text{rank}(X), \text{rank}(Y)) / (\text{stdv}(\text{rank}(X)) * \text{stdv}(\text{rank}(Y))) \quad (4.2)$$

Os resultados para o coeficiente de Spearman foram, na ordem: $s=0.61$ ($p<0.05$), $s=0.44$ ($p>0.05$) e $s=0.76$ ($0.001<p<0.01$). Os resultados condizem com os resultados das tabelas 4.1 - 4.5, enquanto confirmando as forças das demais relações. Portanto, é possível que exista uma relação não linear entre o SCL e a CS. Não houve indicação de correlação entre os scores SSQ e os demais dados afetivos coletados ($\text{abs}(s)<0.2$, $p>0.05$).

4.5 Coleta do EDA

Vários pontos foram levantados no que diz respeito à metodologia de coleta do EDA, principalmente quanto a variáveis que são conhecidas por afetarem artificialmente os valores da condutância da pele e que porventura aparecem nos experimentos, como prender a respiração, fala, movimentos em geral, entre outros. Muitos podem ser controlados durante o experimento. No caso deste estudo, por exemplo, as pessoas eram instruídas a respirarem normalmente, não falarem e evitarem ao máximo se movimentarem durante as sessões de coleta. Isso pode ser visto na Figura 4.7, a qual ilustra os dados do acelerômetro do Empatica E4 obtidos durante uma sessão de relaxamento (à esquerda da linha vermelha central) e quando a pessoa podia se mover livremente (à direita da linha vermelha central), o que ocorria entre as sessões de coleta. É notável a diferença de movimento entre as duas sessões.



Figura 4.7: Sinal do acelerômetro do Empatica E4 durante uma sessão de relaxamento (à esquerda da linha central vermelha) e um período entre sessões de coleta, em que o jogador podia se mover livremente (à direita da linha).

É complexo, no entanto, quantificar o impacto de cada um desses aspectos isoladamente, ainda mais sabendo que a condutância da pele pode variar por idade, sexo, temperatura ambiente e humor - para citar apenas alguns - tanto entre o mesmo indivíduo quanto entre diferentes grupos de pessoas. Seria necessária uma série de estudos destinados apenas para esse propósito.

4.5.1 Sessão de Relaxamento

Neste estudo, tentou-se retirar qualquer tipo de estímulo durante as sessões de relaxamento (5 Min). Os jogadores usavam um abafador de ruídos, permaneciam imóveis e de olhos fechados ao decorrer de toda a sessão. Na teoria, isso aumentaria as chances de retratar a verdadeira linha de base do EDA da pessoa naquele momento. A maioria dos participantes apresentou queda da condutância da pele durante esse período ou condutância estável, o que sugere a que essas sessões indiquem boas estimativas para o nível mínimo do EDA do indivíduo. Infelizmente, não é possível confirmar esse fato, tanto pela natureza do EDA quanto pela falta de uma comparação dessas sessões com sessões de linha de base com estímulos.

4.5.2 Sessão de Triagem

Outra diferença na metodologia de coleta que pode ser comparada a estudos anteriores foi a adição de uma sessão de triagem imediatamente antes da sessão de corrida, como explicado no Capítulo 3. Em poucas palavras, a triagem serviria como mais um período de calibração, na teoria auxiliando no controle de aumentos artificiais do EDA (artefato conhecido como deriva).

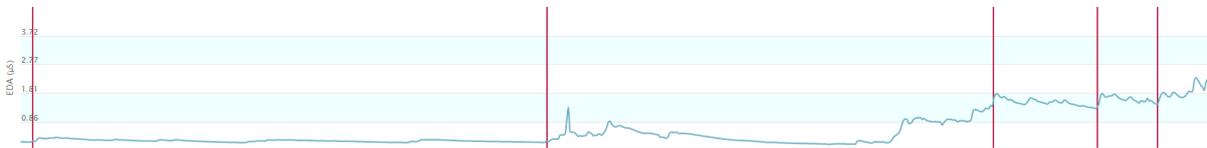


Figura 4.8: Sinal EDA bruto obtido durante as duas sessões de relaxamento. A primeira divisão representa a sessão de cinco minutos, enquanto a terceira divisão representa a sessão de triagem.

A Figura 4.8 mostra o gráfico do sinal bruto do EDA durante uma sessão de relaxamento (primeira divisão) e uma sessão de triagem (terceira divisão). Como é possível perceber, os valores do EDA mudam consideravelmente em um curto período de tempo. Seria impossível perceber que, na realidade, o jogador iniciou o experimento com condutância da pele maior do que a coletada na primeira sessão. Com as Tabelas 4.1 e 4.2, ainda, pode-se notar uma diferença considerável entre os valores da primeira sessão e a segunda. Com isso, pode-se confirmar a utilidade da sessão de triagem.

4.6 Empatica E4

Algumas das principais dúvidas que surgiram durante a concepção e realização deste experimento foi em relação à confiabilidade do sensor Empatica E4 disponível para a

captura dos dados biológicos, o local de coleta dos dados e, mais especificamente em relação ao EDA, a taxa de amostragem da condutância da pele. Cada um desses pontos serão analisados a seguir.

4.6.1 Confiabilidade do Sensor

Como explicitado nos capítulos anteriores, a qualidade do sinal do Empatica E4 se mostrou bastante adequada. No entanto, eram frequentes relatos de erros de coleta, muitos dos quais sem causa aparente. Neste experimento, foram identificados os mesmos problemas. Dos 40 participantes que completaram o experimento, 33% (12) dos arquivos obtidos sobre a atividade cardíaca estavam corrompidos (Figura 4.9).

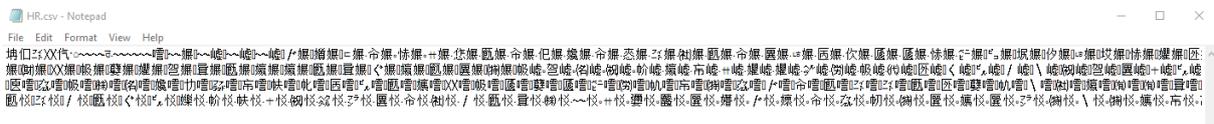


Figura 4.9: Arquivo corrompido obtido do Empatica E4.

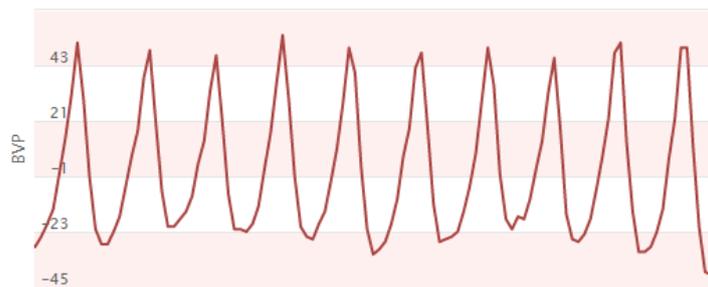


Figura 4.10: Sinal BVP obtido pelo Empatica E4 durante uma sessão de relaxamento.



Figura 4.11: Sinal BVP obtido pelo Empatica E4 durante uma sessão de corrida.

As Figuras 4.10 e 4.11 representam o sinal BVP por tempo de um mesmo indivíduo. A primeira figura é uma parte do gráfico obtido durante a sessão de cinco minutos (relaxamento), enquanto a segunda é uma parte do gráfico obtido durante a sessão de corrida. Pode-se perceber a semelhança do sinal obtido na primeira figura com a forma de onda da Figura 2.4 (retirada da documentação da Empatica) e, por consequência, o erro na

coleta do BVP na segunda figura. Erros como os da Figura 4.11 ocorreram em todas os sinais BVP coletados.

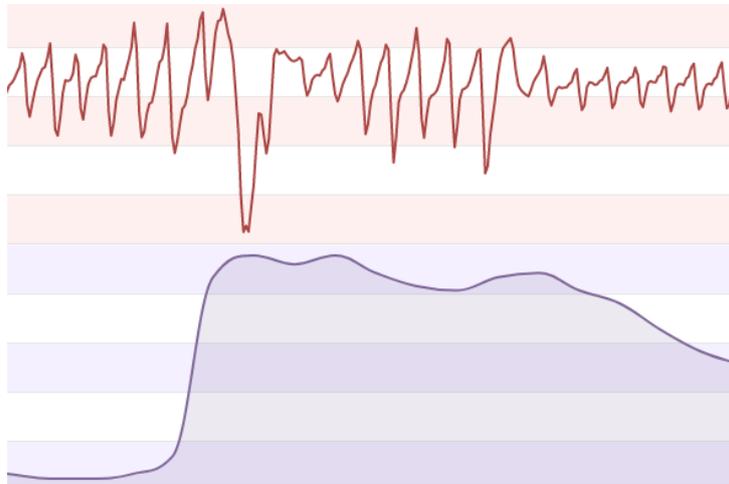


Figura 4.12: Relação do BVP (parte vermelha) com o acelerômetro (parte roxa).

É possível notar pela Figura 4.12 a relação direta entre o movimento do jogador e o sinal BVP obtido pelo sensor. Após o aumento dos valores do acelerômetro (representado pela linha inferior), é possível perceber os artefatos gerados no sinal BVP (linha superior), os quais ocorrem imediatamente depois. Isso condiz com estudos que utilizaram o Empatica E4 em contextos onde os pacientes se movimentavam mais. O que chama a atenção aqui é que neste experimento os jogadores não precisavam se mover (Figura 4.13).



Figura 4.13: Imagem capturada durante o experimento.

No evento específico da figura, está muito claro a causa do artefato. Uma solução simples para casos como esse é simplesmente não levar os dados afetados em consideração. A

Figura 4.14, em contrapartida, mostra como isso rapidamente se complica: não é possível afirmar que a mudança no BVP não foi consideravelmente afetada pela movimentação do jogador. Além disso, como o BVP foi claramente obtido com erros, não se pode dizer que os outros dados foram coletados corretamente.

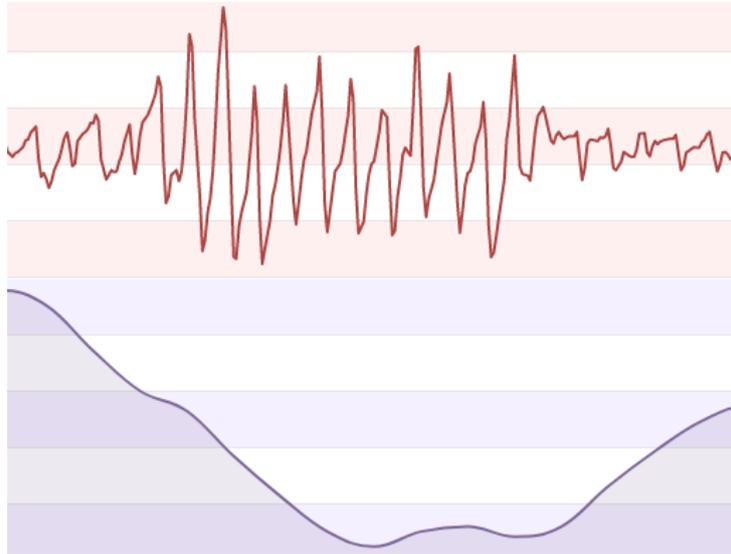


Figura 4.14: Sinal BVP e sinal do acelerômetro do Empatica E4.

É possível notar eventos como os ilustrados nas figuras acima em basicamente todas as sessões de corrida. Dessa forma, não se pode afirmar que o Empatica E4 é um sensor confiável para coleta de dados fisiológicos no contexto de jogos. A grande maioria induz movimentos e reações no jogador que dificilmente conseguem ser contidas, e estudos que incluem jogos devem levar isso em consideração. Mesmo contidas, como no caso deste experimento, o Empatica E4 produziu erros nas coletas dos dados. Erros como esses são inaceitáveis em estudos que levam respostas biológicas imediatas em consideração, como o HRV e ES-SCRs. É impossível criar um sensor perfeito, à prova de tudo. Contudo, devido ao teor extremamente complexo e interdependente dos sinais biológicos, principalmente quando se fala da resposta galvânica da pele, se faz necessário um sensor com uma taxa de sucesso de aquisição dos dados muito mais alta ou, pelo menos, que se saiba as limitações do Empatica E4 para que, durante o planejamento de futuros experimentos, tais limitações sejam devidamente levadas em consideração.

4.6.2 Local de Coleta dos Dados

Outra dúvida levantada neste estudo é pela região de coleta dos dados fisiológicos pelo Empatica E4, com mais foco na condutância da pele. Embora um sensor leve e compacto como o Empatica ofereça praticidade e portabilidade na coleta dos dados, captar todos

os dados a partir do pulso não é o consenso literário. O BVP, por exemplo, é geralmente coletado nas regiões periféricas do corpo, como na ponta dos dedos ou nos lóbulos. O EDA é geralmente coletado nas palmas das mãos ou dos pés. De toda forma, as pesquisas relacionadas à validade dos locais de coleta do EDA são raras. O estudo de Kushki et al. [36], por exemplo, observou que regiões não palmares podem ser locais de coleta de dados BVP e EDA, mas isso ainda há de ser confirmado. Por enquanto, ainda há evidências de que a palma das mãos e pés ainda são a melhor opção para a coleta da atividade eletrodérmica, já que essas regiões apresentam alta densidade de glândulas écrinas e são conhecidas por serem mais responsivas a estímulos psicológicos do que a estímulos térmicos [48].

4.6.3 Taxa de Amostragem (EDA)

A taxa de amostragem do EDA pelo sensor de pulso também provou ser fonte de dúvidas. O Empatica E4 calcula a condutância da pele quatro vezes por segundo (4Hz), valor extremamente baixo quando comparados a sensores de laboratório, com os quais se pode chegar à mais de 1500Hz de amostragem. Existem artigos que indicam uma taxa mínima de 300-400Hz de amostragem, enquanto outros indicam 10Hz. A questão que surge, portanto, é se a taxa do Empatica E4 para a resposta galvânica é adequada. Infelizmente, responder isso foge às capacidades deste trabalho. Seria necessário um estudo que ataca diretamente essa questão. Por exemplo, um que controlasse outras variáveis que impactam o EDA e variasse as taxas de amostragem, comparando o impacto das taxas de amostragem nos algoritmos de correção e normalização dos dados e, conseqüentemente, na interpretação deles. Vale relatar, contudo, que o NeuroKit2 não aceita 4Hz como taxa de amostragem, como pode ser visto na Figura 4.15. Esse e todas as outras questões já explicadas sobre a taxa de amostragem do Empatica E4 são indícios de que a amostragem é baixa demais para a captação condutância da pele, seja pela confiabilidade dos dados obtidos ou pela necessidade de uma frequência aquisitiva maior para que o sinal seja propriamente processado sem alterá-lo consideravelmente.

```
9: NeuroKitWarning: The sampling rate is too low. Sampling rate must exceed the Nyquist rate to avoid aliasing problem. In this analysis, the sampling rate has to be higher than 6 Hz
```

Figura 4.15: Erro emitido pelo NeuroKit2 quando a taxa de amostragem é muito baixa.

4.7 Processamento e Análise do EDA

Os principais desafios para estudos que utilizam o EDA são: artefatos de movimento, e decomposição tônica e fásica. Em ambas as frentes, há um longo histórico de pesquisas que propuseram diferentes tipos de soluções, a maioria automática. Não foi possível comparar diferentes algoritmos de detecção e correção de erros pois isso exigiria uma análise profunda de cada um, o que é impraticável com a janela de tempo disponível para este estudo. No entanto, foi possível realizar uma comparação do NeuroKit2 e seu algoritmo de decomposição tônica e fásica com o *cvxEDA*.

A Figura 4.16 é o resultado da decomposição tônica a partir de um único sinal EDA bruto. A Figura 4.17 representa o componente fásico do mesmo sinal. Em ambos os casos, apenas as taxas de amostragem informadas ao NeuroKit2 foram modificadas, a fins de ilustração. É notável a incongruência dos sinais. Pode-se perceber, também, que quanto maior a taxa de amostragem, mais espaçado e suave o sinal fica, o que é um indicativo que maiores taxas amostrais implicam melhor leitura dos dados, principalmente para o componente fásico.

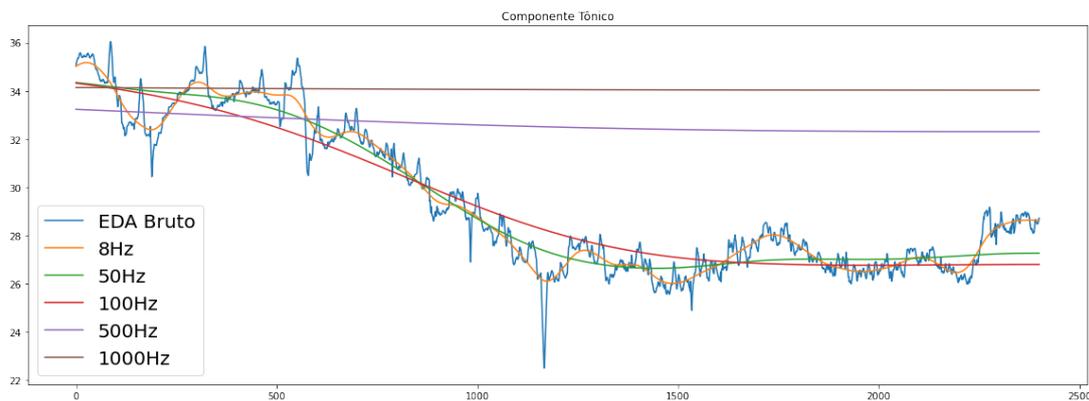


Figura 4.16: Componente tônico (nível SCL) do EDA pelo Neurokit2 com diferentes taxas de amostragem.

Seria natural assumir, portanto, que o mesmo aconteceria para diferentes algoritmos de decomposição. A Figura 4.18 confirma essa hipótese. Foram incluídas as decomposições tônicas de mesmo sinal EDA usando o *cvxEDA*, o *sparsEDA*, o *CDA-Ledalab*, o *DDA-Ledalab* e a modelagem causal *DCM-psPM*. Também não há congruência nos resultados entre os diferentes métodos de decomposição. É importante frisar que a incongruência em si não é indesejada, mas a falta da definição das utilidades e limitações de cada uma delas.

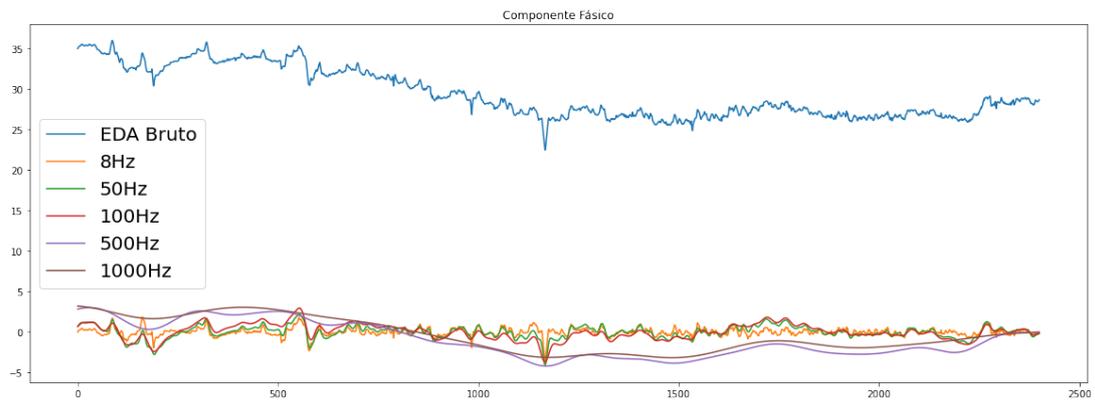


Figura 4.17: Componente fásico (amplitudes SCR) do EDA pelo Neurokit2 com diferentes taxas de amostragem.

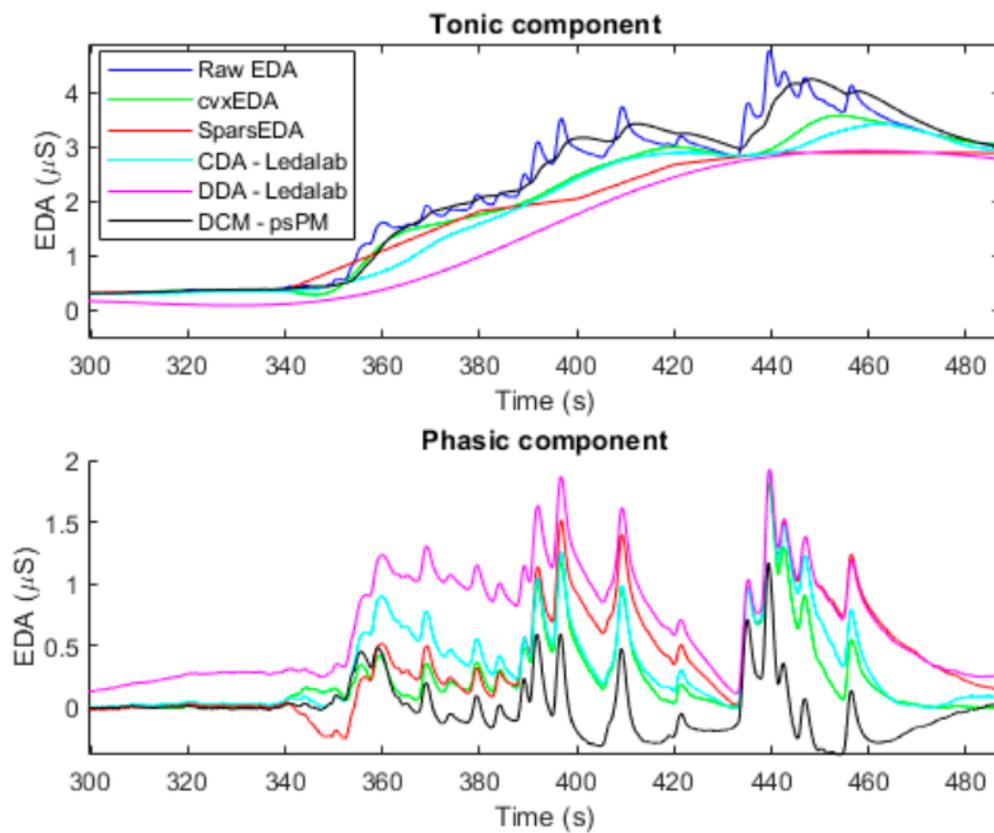


Figura 4.18: Resultados dos algoritmos de decomposição do EDA mais utilizados atualmente. Retirado de [48].

Capítulo 5

Conclusão

Este estudo explorou a potencial relação entre a severidade da *cybersickness* experienciada por um indivíduo e suas respostas vitais, em especial o EDA. Os jogadores foram categorizados entre aqueles que completaram o experimento (Grupo 1) e aqueles que o interromperam (Grupo 2), visto que a média do score SSQ do segundo grupo foi mais de cinco vezes maior. O Coeficiente de Correlação de Spearman indicou uma correlação positiva relativamente forte ($s=0.61$, $p<0.05$) entre o score SSQ do Grupo 2 e o nível tônico (SCL) dos jogadores. A correlação foi fraca ($s=0.44$, $p>0.05$) entre os scores e o número de respostas fásicas (picos SCR), o que indica que a CS está mais ligada ao nível geral de excitação do usuário.

As duas sessões de linha de base (5 min. e 1 min.) se mostraram bastante úteis na interpretação do nível de excitação do usuário. Em contrapartida, os erros de coleta, tanto da função cardíaca quanto da condutância da pele dificultaram enormemente o processamento e interpretação dos dados. Grande parte disso foi causado pela ineficiência do dispositivo vestível Empatica E4. Ele é limitado a funções como detectar se o usuário está usando o dispositivo, ou análise básica, como simplesmente computar o nível de condutância, não contendo a robustez necessária para garantir a qualidade dos sinais. O Empatica E4 apresentou alta sensibilidade a movimentos, contendo vários erros de coleta nos dados afetivos (principalmente na sessão de corrida, período em que os jogadores mais se movimentavam), tornando-o inadequado para aquisição de sinais fisiológicos no contexto de jogos. A viabilidade de dispositivos vestíveis para fornecer a cadeia de recursos necessária deve ser mais investigada no contexto de jogos. O Empatica E4 tornava-se confiável apenas nas sessões de relaxamento, onde os usuários ficavam completamente imóveis. Além disso, a taxa de amostragem da condutância da pele de 4Hz não é adequada para os algoritmos de processamento mais modernos, como *cvxEDA* utilizado pela biblioteca Python *NeuroKit2*.

Em relação aos algoritmos de processamento do EDA, se faz necessário uma definição clara das utilidades e limitações de cada um. Monografias de graduação passadas utilizaram o *Real-Time Arousal Asset* (RTA) ¹ para calcular o nível fásico e tônico da condutância da pele, além do nível geral de excitação do usuário. O RTA aplica os processamentos descritos por Dawson, que consideram valores mínimos e máximos no nível de condutância da pele, que é uma abordagem controversa pela natureza variável do EDA. Os algoritmos mais atuais (cvxEDA, sparsEDA, Ledalab, etc) são mais robustos na detecção e correção de erros, e na decomposição fásica/tônica. No entanto, acredita-se que esse problema permanecerá sem solução até que exista um conceito bem definido para o sinal da condutância da pele. A consistência das medidas derivadas da EDA ainda é motivo de preocupação, especificamente na presença de artefatos de movimento.

5.1 Trabalhos Futuros

Constituem objetos de estudo para trabalhos futuros:

- A maior falha deste estudo foi não ter capturado o contexto em que o jogador estava inserido. As informações contextuais são extremamente necessárias. Uma arquitetura e implementação de um sistema para adquirir, processar e visualizar EDA e outros sinais fisiológicos foi desenvolvida (*FEEL* foi o nome dado a este sistema), e seus resultados indicam que o sistema permitiu que os usuários anotassem informações contextuais a serem usadas na análise dos dados [79]. Trabalhos futuros devem, sem dúvidas, incorporar sistemas contextuais como o *FEEL* para melhor interpretar os resultados obtidos.
- Além da diferença no tempo computacional, os algoritmos de decomposição mostrados neste estudo fornecem uma estimativa muito diferente do componente tônico e fásico. *Machine learning* ou *Deep Learning* podem ajudar nessa tarefa. Se uma definição bem estabelecida de como os componentes tônicos e fásicos devem ser compostos existisse, um modelo poderia ser treinado para realizar tal decomposição.
- A sessão de triagem se mostrou útil para auxiliar no processamento e interpretação corretos da resposta galvânica da pele. Fica a dúvida de como seriam outras sessões de calibração, como uma imediatamente depois do experimento e até durante o experimento, se for o caso.
- Validar o sinal do Empatica E4 no contexto de jogos. O interessante seria testar o Empatica E4 em uma série de jogos diferentes, que requerem todo tipo de movimentação e geram diferentes tipos de reações.

¹<https://github.com/ddessy/RealTimeArousalDetectionUsingGSR>

- Um estudo que isola diferentes locais de coleta de EDA e comparam os resultados obtidos de cada região.
- Não existem estudos que exploram diferentes taxas de amostragem e o impacto que ela geraria em um sistema de classificação de emoções, por exemplo.
- Embora haja tremenda estabilidade na temperatura da pele periférica entre os indivíduos e que ela esteja relacionada diretamente com a temperatura ambiente, sugere-se investigar o impacto dela nos dados fisiológicos dos participantes. Um estudo nessa linha poderia delimitar melhor algumas variáveis que influenciam a aquisição de dados fisiológicos.
- Mais estudos que confirmassem a precisão dos dados da atividade cardíaca e da respiração com modelos valência/excitação no campo da computação afetiva.
- Um estudo que relacionasse a presença da tecnologia de RV e a Cybersickness.

Referências Bibliográficas

- [1] Lindsley, D. B. (1951). *Emotion*. In S. S. Stevens (Ed.), Handbook of experimental psychology (p. 473–516). Wiley. 3
- [2] M. P. & Niedenthal. (2007). *Embodying Emotion*. Science 18 May 2007: Vol. 316, Issue 5827, pp. 1002-1005 DOI: 10.1126/science.1136930. 3
- [3] Averill, J. R. (1999). *Individual Differences in Emotional Creativity: Structure and Correlates*. Journal of Personality, 67(2), 331–371. doi:10.1111/1467-6494.00058 . 3
- [4] Plutchik, R., & Kellerman, H. (Eds.). (2013). *Theories of emotion (Vol. 1)*. Academic Press. 3
- [5] Cacioppo, J. T. & Gardner, W. L. (1999). *Emotion*. Annual Review of Psychology 1999 50:1, 191-214.
- [6] McDuff & Daniel et al. (2013). *Affectiva-mit facial expression dataset (am-fed): Naturalistic and spontaneous facial expressions collected*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 4
- [7] Daily, S. B., James, M. T., Cherry, D., J. Porter, J., Darnell, S. S., Isaac, J., & Roy, T. (2017). *Affective Computing: Historical Foundations, Current Applications, and Future Trends. Emotions and Affect in Human Factors and Human-Computer Interaction*. 213–231. doi:10.1016/b978-0-12-801851-4.00009-4 4
- [8] Picard, R. (2000). *Affective computing*. MIT press. 4
- [9] Lee & Jiyoung et al. (2019). *Context-aware emotion recognition networks*. Proceedings of the IEEE International Conference on Computer Vision. 5
- [10] Retto, J. (2017). *Sophia, first citizen robot of the world*. ResearchGate. 5
- [11] Ekman, P. (1992). *An Argument for Basic Emotions*. Cognition and Emotion. 6 (3/4): 169–200. 5

- [12] Jack, R. E., Sun, W., Delis, I., Garrod, O. G. B., & Schyns, P. G. (2016). *Four not six: Revealing culturally common facial expressions of emotion*. *Journal of Experimental Psychology: General*, 145(6), 708–730. doi:10.1037/xge0000162 . 5
- [13] Cowen, A. S., & Keltner, D. (2017). *Self-report captures 27 distinct categories of emotion bridged by continuous gradients*. *Proceedings of the National Academy of Sciences*, 114(38), E7900–E7909. doi:10.1073/pnas.1702247114. 5
- [14] Posner, J., Russel, J. A., & Peterson, B. S. (2005). *The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology*. *Development and Psychopathology*, 17(03). doi:10.1017/s0954579405050340. 6
- [15] Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). *Remembering pictures: Pleasure and arousal in memory*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2), 379–390. <https://doi.org/10.1037/0278-7393.18.2.379>. 6
- [16] Watson, D. & Tellegen, A. (1985). *Toward a consensual structure of mood*. *Psychological Bulletin* 98.2: 219. 6
- [17] Bakker, I., van der Voordt, T., Vink, P. et al. (2014). *Pleasure, Arousal, Dominance: Mehrabian and Russell revisited*. *Curr Psychol* 33, 405–421. <https://doi.org/10.1007/s12144-014-9219-4>. 6
- [18] N., Lim. (2016). *Cultural differences in emotion: differences in emotional arousal level between the East and the West*. *Integrative Medicine Research*, Volume 5, Issue 2, 2016, Pages 105-109, ISSN 2213-4220. 7
- [19] E., Tuomas & V., Jonna. (2011). *A comparison of the discrete and dimensional models of emotion in music*. *Psychology of Music*. 10.1177/0305735610362821. 8
- [20] Lichtenstein, A., Oehme, A., Kupschick, S., & Jürgensohn, T. (n.d.). *Comparing Two Emotion Models for Deriving Affective States from Physiological Data*. *Lecture Notes in Computer Science*, 35–50. doi:10.1007/978-3-540-85099-1_4. 8
- [21] Rubin, D. & Talarico, J. (2009). *A Comparison of Dimensional Models of Emotion: Evidence from Emotions, Prototypical Events, Autobiographical Memories, and Words*. *Memory (Hove, England)*. 17. 802-8. 10.1080/09658210903130764. 8
- [22] Shoumy, N. J., Ang, L.-M., Seng, K. P., Rahaman, D. M. M., & Zia, T. (2019). *Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals*. *Journal of Network and Computer Applications*, 102447. doi:10.1016/j.jnca.2019.102447 . 9

- [23] Wioleta, S. (2013). *Using physiological signals for emotion recognition*. 6th International Conference on Human System Interactions (HSI). doi:10.1109/hsi.2013.6577880. 9
- [24] C. Busso, M. Bulut, C.-C. Lee et al. (2008). *IEMOCAP: interactive emotional dyadic motion capture database*. Language Resources and Evaluation, vol. 42, no. 4, pp. 335–359, 2008. 9
- [25] Sayedelahl A., Fewzee P., Kamel M.S., Karray F. (2011). *Audio-based Emotion Recognition from Natural Conversations Based on Co-Occurrence Matrix and Frequency Domain Energy Distribution Features*. In: D’Mello S., Graesser A., Schuller B., Martin JC. (eds) Affective Computing and Intelligent Interaction. ACII 2011. Lecture Notes in Computer Science, vol 6975. Springer, Berlin, Heidelberg. 9
- [26] Bradley, Margaret M., & Peter J. Lang. (1994). *Measuring emotion: the self-assessment manikin and the semantic differential*. Journal of behavior therapy and experimental psychiatry 25.1 (1994): 49-59. 10
- [27] Fabes, R. A., Eisenberg, N., & Eisenbud, L. (1993). *Behavioral and physiological correlates of children’s reactions to others in distress*. Dev. Psychol. 29, 655–663. 10
- [28] Quintana, D., Guastella, A., Outhred, T., Hickie, I., & Kemp, A. (2012). *Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition*. International Journal Of Psychophysiology, 86(2), 168-172. doi: 10.1016/j.ijpsycho.2012.08.012. 10
- [29] Hansen, A., Johnsen, B., & Thayer, J. (2009). *Relationship between heart rate variability and cognitive function during threat of shock*. Anxiety, Stress & Coping, 22(1), 77-89. doi: 10.1080/10615800802272251. 10
- [30] Mather, M., & Thayer, J. F. (2018). *How heart rate variability affects emotion regulation brain networks*. Current Opinion in Behavioral Sciences, 19, 98–104. doi:10.1016/j.cobeha.2017.12.017. 10
- [31] Quintana, D. S., Guastella, A. J., Outhred, T., Hickie, I. B., & Kemp, A. H. (2012). *Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition*. International Journal of Psychophysiology, 86(2), 168–172. doi:10.1016/j.ijpsycho.2012.08.012 . 10
- [32] Choi, K., Kim, J., Kwon, O. S., Kim, M. J., Ryu, Y. H. & Park, J. (2017). *Is heart rate variability (HRV) an adequate tool for evaluating human emotions? – A focus on*

- the use of the International Affective Picture System (IAPS)*. *Psychiatry Research*, 251(), 192–196. doi:10.1016/j.psychres.2017.02.025 . 11
- [33] Yan, w., Ruolei, G., Qiwei, Y. & Yue-jia, L. (2019). *How Do Amusement, Anger and Fear Influence Heart Rate and Heart Rate Variability?* *Frontiers in Neuroscience*, Vol 13, 2019. Pages 1139, ISSN 1662-453X. 11
- [34] Shaffer, F. & Ginsberg, J. P. (2017). *An Overview of Heart Rate Variability Metrics and Norms*. *Frontiers in Public Health*, 5(), 258–. doi:10.3389/fpubh.2017.00258. 11
- [35] Peper, E. & Harvey, Ri. & Lin, I. & Tylova, H. & Moss, D. (2007). *Is There More to Blood Volume Pulse Than Heart Rate Variability, Respiratory Sinus Arrhythmia, and Cardiorespiratory Synchrony?*. *Biofeedback*. 35. 54-61. 11
- [36] Kushki, A. & Fairley, J. & Merja, S. & King, G. & Chau, T. (2011). *Comparison of blood volume pulse and skin conductance responses to mental and affective stimuli at different anatomical sites*. *Physiological measurement*. 32. 1529-39. 10.1088/0967-3334/32/10/002. 55
- [37] Rimm-Kaufman, S. E., & Kagan, J. (1996). *The psychological significance of changes in skin temperature*. *Motivation and Emotion*, 20(1), 63–78. doi:10.1007/bf02251007 . 13
- [38] Mustafa, K. & Kamiokam E. (n.d.). *Emotions detection scheme using facial skin temperature and heart rate variability*. Graduate School of Engineering and Science, Shibaura Institute of Technology, Tokyo, Japan. 13
- [39] Park, M. W., Kim, C. J., Hwang, M., & Lee, E. C. (2013). *Individual Emotion Classification between Happiness and Sadness by Analyzing Photoplethysmography and Skin Temperature*. 2013 Fourth World Congress on Software Engineering. doi:10.1109/wcse.2013.34. 13
- [40] Shihab Rashida, M., Zamanb, Z., Mahmudc, H. & Md. Kamrul Hasand. (n.d.). *Emotion Recognition with Forearm-based Electromyography* Department of Computer Science and Engineering, Islamic University of Technology, Boardbazar, Gazipur, Bangladesh 14
- [41] Hernandez, A., Katoh, R., Yokoi, H. & Wenweim Y. (2006). *Development of a multi-dof electromyography prosthetic system using the adaptive joint mechanism*. In *Applied Bionics and Biomechanics*, 3(2), pp. 101–111, 2006. 14

- [42] Zhang, X. Chen, W.-h. Wang, J.-h. Yang, V. Lantz, & K.-q.Wang. (2009). *Hand gesture recognition and virtual game control based on 3d accelerometer and emg sensors*. Proc. of the International Conference on Intelligent User Interfaces, pp. 401–406, 2009. 14
- [43] Jerath, R., Beveridge, C. (2020). *Respiratory Rhythm, Autonomic Modulation, and the Spectrum of Emotions: The Future of Emotion Recognition and Modulation*. Frontiers in Psychology, 11. doi:10.3389/fpsyg.2020.01980. 14
- [44] Boiten & Frans, A. (1998). *The effects of emotional behaviour on components of the respiratory cycle*. Biological psychology 49.1-2 (1998): 29-51. 14
- [45] Masaoka, Yuri, and Homma, I. (2001). *The effect of anticipatory anxiety on breathing and metabolism in humans*. Respiration physiology 128.2 (2001): 171-177. 14
- [46] Zhang Q., Chen X., Zhan Q., Yang T. & Xia S. (2017). *Respiration-based emotion recognition with deep learning*. Computers in Industry, 92-93, 84-90. 15
- [47] Mirmohamadsadeghi, L., Yazdani, A., Vesin, J.-M. (2016). *Using cardio-respiratory signals to recognize emotions elicited by watching music video clips*. 2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP). 15
- [48] Posada-Quintero, H.F. & Chon, K.H. (2020). *Innovations in Electrodermal Activity Data Collection and Signal Processing: A Systematic Review*. Sensors 2020, 20, 479. x, 1, 16, 23, 37, 55, 57
- [49] Roth, W. T., Dawson, M. & Filion, D. L. (2012). *Publication recommendations for electrodermal measurements*. Psychophysiology, vol. 49, pp. 1017–1034, 2012. 20, 22, 23, 24, 37
- [50] Bach, D.R., Flandin, G., Friston, K.J., Dolan, R.J. (n.d.). *Time-series analysis for rapid event-related skin conductance responses*. J. Neurosci. Methods 2009, 184, 224–234. 20
- [51] Bach, D.R., Friston, K.J. & Dolan, R.J. (2010). *Analytic measures for quantification of arousal from spontaneous skin conductance fluctuations*. Int. J. Psychophysiol. 2010, 76, 52–55. 20
- [52] Bach, D.R., Flandin, G., Friston, K.J. & Dolan, R.J. (n.d). *Modelling event-related skin conductance responses*. Int. J. Psychophysiol. 2010, 75, 349–356. 20

- [53] Cecchi, S., Piersanti, A., Poli, A., & Spinsante, S. (2020). *Physical Stimuli and Emotions: EDA Features Analysis from a Wrist-Worn Measurement Sensor*. 2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD). ix, 16, 17
- [54] Braithwaite, J. J., Watson, G., Jones, R. & Rowe, M. (2015). *A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments*. Technical Report, 2nd version: Selective Attention and Awareness Laboratory (SAAL). Behavioural Brain Sciences Centre, University of Birmingham, UK. ix, 17, 21, 22, 24
- [55] Kappeler-Setz, C, Gravenhorst, F., Schumm, J., Arnrich, B. & Tröster, G. (2013). *Towards long term monitoring of electrodermal activity in daily life*. Pers. Ubiquitous Comput., vol. 17, no. 2, pp. 261–271, 2013. 22
- [56] Ragot, M., Nicolas, M., Sonia, E, Pallamin, N. & Diverrez, J. (2017). *Emotion Recognition Using Physiological Signals: Laboratory vs. Wearable Sensors*. 10.1007/978-3-319-60639-2_2. 23
- [57] Ollander, S., Godin, C., Campagne, A., & Charbonnier, S. (2016). *A comparison of wearable and stationary sensors for stress detection*. IEEE International Conference on Systems, Man, and Cybernetics (SMC). doi:10.1109/smc.2016.7844917. 23
- [58] Milstein, N., & Gordon, I. (2020). *Validating Measures of Electrodermal Activity and Heart Rate Variability Derived From the Empatica E4 Utilized in Research Settings That Involve Interactive Dyadic States*. Frontiers in Behavioral Neuroscience, 14. doi:10.3389/fnbeh.2020.00148. 23, 24
- [59] Borrego, A., Latorre, J., Alcaniz, M., & Llorens, R. (2019). *Reliability of the Empatica E4 wristband to measure electrodermal activity to emotional stimuli*. 2019 International Conference on Virtual Rehabilitation (ICVR). doi:10.1109/icvr46560.2019.8994546. 24
- [60] Burdea, G. C., & Coiffet, P. (2003). *Virtual reality technology*. John Wiley & Sons. 25
- [61] Csizentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper Perennial
- [62] Brown, E. & Cairns, P. (2004). *A grounded investigation of game immersion*. In CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04). Association for Computing Machinery, New York, NY, USA, 1297–1300. DOI:<https://doi.org/10.1145/985921.986048>. 26

- [63] Cheng, K. & Cairns, P. (2005). *Behaviour, realism and immersion in games*. CHI '05 Extended Abstracts on Human Factors in Computing Systems - CHI '05. doi:10.1145/1056808.1056894. 26
- [64] A. A. Varghese, J. P. Cherian & J. J. Kizhakkethottam. (2015). *Overview on emotion recognition system*. 2015 International Conference on Soft-Computing and Networks Security (ICSNS), Coimbatore, 2015, pp. 1-5, doi: 10.1109/ICSNS.2015.7292443.
- [65] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W. & Taylor, J. G. (2001). *Emotion recognition in human-computer interaction*. IEEE Signal Processing Magazine, 18(1), 32–80. doi:10.1109/79.911197.
- [66] Sreeja, P. S. & Mahalakshmi, G. S. (2017). *Emotion Models: A Review*. International Journal of Control Theory and Applications. 10. 651-657.
- [67] Ferit, K. (2020). *Using a Chatbot, Replika, to Practice Writing Through Conversations in L2 English: A Case Study*. New Technological Applications for Foreign and Second Language Learning and Teaching. IGI Global, 2020. 222-239.
- [68] Ahern, G. L. & Schwartz, G. E. (1985). *Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis*. Neuropsychologia, vol. 23, no. 6, pp. 745–755, 1985. 15
- [69] Gunes, H & Piccardi, M. (2007). *Bi-modal emotion recognition from expressive face and body gestures*. Journal of Network and Computer Applications, vol. 30, no. 4, pp. 1334–1345, 2007. 15
- [70] Blackford, J. U, & Pine, D. S. (2012). *Neural substrates of childhood anxiety disorders*. Child and Adolescent Psychiatric Clinics of North America, vol. 21, no. 3, pp. 501–525, 2012. 15
- [71] Goosens, K. A. & Maren, S. (2002). *Long-term potentiation as a substrate for memory: evidence from studies of amygdaloid plasticity and pavlovian fear conditioning*. Hippocampus, vol. 12, no. 5, pp. 592–599, 2002. 15
- [72] Turner, M. R., Maren, S., Phan, K. L, & Liberzon, I. (2013). *The contextual brain: implications for fear conditioning, extinction and psychopathology*. Nature Reviews Neuroscience, vol. 14, no. 6, pp. 417–428, 2013. 15
- [73] Cayres, M. P. (2019). *Ajuste Dinâmico de Dificuldade híbrido em um jogo do gênero plataforma*. Universidade de Brasília, Instituto de Ciências Exatas, Departamento de Ciência da Computação. 30

- [74] Lopes et al. (2019). *Comparação entre modelos de Ajuste Dinâmico de Dificuldade em um jogo do gênero plataforma/shooter*. Universidade de Brasília, Instituto de Ciências Exatas, Departamento de Ciência da Computação. 30
- [75] Bertanini, P. (2019). *Ferramenta para análise do estado afetivo baseado em dados fisiológicos*. Universidade de Brasília, Instituto de Ciências Exatas, Departamento de Ciência da Computação. 23, 31
- [76] Mota, G. (2019). *Estudo sobre abordagens de uso de biofeedback para jogos em realidade virtual*. Universidade de Brasília, Instituto de Ciências Exatas, Departamento de Ciência da Computação. 31
- [77] Fernandez, J. A., Fusté, A., Richer, R. & Maes, P. (2019). *Deep reality: an underwater VR experience to promote relaxation by unconscious HR, EDA, and brain activity biofeedback*. In ACM SIGGRAPH 2019 Virtual, Augmented, and Mixed Reality (SIGGRAPH '19). Association for Computing Machinery, New York, NY, USA, Article 17, 1. DOI:<https://doi.org/10.1145/3306449.3328818>. 24
- [78] Greco, A., Valenza, G., Lanata, A., Scilingo, E., & Citi, L. (2016). *cvxEDA: a Convex Optimization Approach to Electrodermal Activity Processing*. IEEE Transactions on Biomedical Engineering, 1–1. doi:10.1109/tbme.2015.2474131 37, 38
- [79] Ayzenberg, Y. & Picard, R.W. (2014). *FEEEL: A system for frequent event & electrodermal activity labeling*. IEEE J. Biomed. Health Inform. 2014, 18, 266–277. 59
- [80] Gallego, H., Francisco, Luengo, David, Rodríguez, A. (2017). *Feature Extraction of Galvanic Skin Responses by Non-Negative Sparse Deconvolution*. IEEE Journal of Biomedical and Health Informatics. PP. 1-1.
- [81] Rebenitsch, L. (2015). *Managing cybersickness in virtual reality*. XRDS: Crossroads, The ACM Magazine for Students, 22(1), 46–51. doi:10.1145/2810054 28
- [82] Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). *Presence and Cybersickness in Virtual Reality Are Negatively Related: A Review*. Frontiers in Psychology, 10. 28, 29
- [83] Kim, Y. Y., Kim, H. J., Kim, E. N., Ko, H. D., & Kim, H. T. (2005). *Characteristic changes in the physiological components of cybersickness*. Psychophysiology, 0(0), 050826083901001 doi:10.1111/j.1469-8986.2005.00349.x . 29
- [84] Dennison, M. S., Wisti, A. Z. & D’Zmura, M. (2016). *Use of physiological signals to predict cybersickness*. Displays, Volume 44, 2016,Pages 42-52, ISSN 0141-9382, <https://doi.org/10.1016/j.displa.2016.07.00>. 34

- [85] Golding, J. F. (1992). *Phasic skin conductance activity and motion sickness*. *Aviation, Space, and Environmental Medicine*, 63(3), 165–171. 34
- [86] Gavvani, A. M., Nesbitt, K. V., Blackmore, K. L., & Nalivaiko, E. (2017). *Profiling subjective symptoms and autonomic changes associated with cybersickness*. *Autonomic Neuroscience*, 203, 41–50. doi:10.1016/j.autneu.2016.12.004. 34
- [87] Kennedy, R.S., Lane, N. E., Berbaum, K. S. & Lilienthal, M. G. (1993). *Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness*. *The International Journal of Aviation Psychology*, 3(3):203–220, 1993. https://doi.org/10.1207/s15327108ijap0303_3. 4, 5, 6, 7, 23. 1, 28, 41
- [88] Sevinca, V. & Ilkerb, B. M. (2020). *Psychometric evaluation of Simulator Sickness Questionnaire and its variants as a measure of cybersickness in consumer virtual environments*. *Applied Ergonomics* Volume 82, January 2020, 102958. 1, 28
- [89] Kim, Hyun, K., Park, J, Choi, Y. & Choe, M. (2018). *Virtual reality sickness questionnaire (vrsq): Motion sickness measurement index in a virtual reality environment*. *Applied Ergonomics*, 69:66 – 73, 2018, ISSN 0003-6870. <http://www.sciencedirect.com/science/article/pii/S000368701730282X>. 6, 10, 14, 16, 23. 28
- [90] Ben-Shakar, G. (1985). *Standardization within individuals: A simple method to neutralize individual differences in skin conductance*. *Psychophysiology*, 22, 292–299. 37

Apêndice A

Termo de Consentimento Livre e Esclarecido



TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

(De acordo com as normas da Resolução nº 466, do Conselho Nacional de Saúde de 12/12/2012)

Você está sendo convidado para participar da pesquisa **Avaliação de sintomas de cybersickness em realidade virtual utilizando dados subjetivos e fisiológicos**. A sua participação não é obrigatória e a qualquer momento você pode desistir de participar e retirar seu consentimento. Sua recusa não trará nenhum prejuízo em sua relação com o pesquisador ou com a instituição.

O objetivo deste estudo é **comparar métodos subjetivos de quantificação de cybersickness**.

Sua participação nesta pesquisa consistirá em interagir com um jogo, de modo que serão coletados dados fisiológicos através do uso de sensores. Além disso, você responderá dois questionários, um antes e um depois do experimento. O experimento dura em torno de 10 minutos em jogo, podendo ser abortado quando você quiser, mais em torno de 10 minutos para coleta de dados fisiológicos e resposta dos questionários.

Os riscos relacionados com sua participação são nulos, de modo que o ambiente do experimento e os equipamentos são seguros, não levando uso de produtos químicos ou equivalentes e sendo higienizados com álcool em gel.

O benefício relacionado a sua participação é colaborar diretamente com um projeto de pesquisa de graduação, de modo a fornecer dados experimentais fidedignos.

As informações obtidas através dessa pesquisa serão confidenciais e asseguramos o sigilo sobre sua participação. Os dados não serão divulgados de forma a possibilitar sua identificação. Ao final da análise dos dados, todas as informações pertinentes a sua participação serão excluídas.

Uma cópia deste Termo de Consentimento Livre e Esclarecido ficará com o senhor (a) ou será enviado por E-mail (digitalizado), podendo tirar suas dúvidas sobre o projeto e sua participação, agora ou a qualquer momento com os pesquisadores responsáveis:

- GABRIEL DE CARVALHO TOLLINI no e-mail **gabriel tollini@hotmail.com** ou no telefone (61) 98189-6777.
- DANILLO NEVES SOUZA no e-mail **danillo.neves97@gmail.com** ou no telefone (61) 9 8130-3010

Pesquisador Responsável

Declaro que entendi os objetivos, riscos e benefícios de minha participação na pesquisa e concordo em participar.

Brasília, ____ de _____ de 20__.

Sujeito da pesquisa

Pai / Mãe ou Responsável Legal (Caso o sujeito seja menor de idade)

Apêndice B

Roteiro do Experimento

Roteiro do Experimento

Apresentação e termo de consentimento:

- Fala: “Olá. Antes de começarmos, vou explicar rapidamente o experimento. Trata-se de um jogo de corrida em Realidade Virtual, onde você não será avaliado(a) pelo seu desempenho no jogo.”
- Fala: “Você vai correr por 10 minutos, e quando o tempo acabar você poderá finalizar a última volta na pista, se quiser. Saliento que você pode interromper o experimento a qualquer momento.” **Explicar os controles e sanar quaisquer dúvidas sobre o jogo antes de continuar.**
- Fala: “Antes e depois do experimento, serão passados formulários para que você preencha se está sentindo algum tipo de mal estar ou não”.
- Fala: “Também colocarei no seu pulso essa pulseira que mede os seus batimentos cardíacos e a condutância da pele durante a imersão na Realidade Virtual”.
- Mostrar a pulseira Empatica E4 e o HMD (*Head Mounted Display*) para o participante.
- Higienizar os equipamentos na frente do participante.
- Fala: “Antes de tudo, preciso que assine um termo de consentimento.”
- **Entregar o termo de consentimento.**
- Fala: “Leia e, se tiver qualquer dúvida sobre o termo, posso esclarecer.”
- Esperar o participante ler e assinar, e esclarecer qualquer eventual dúvida.

Início do experimento:

- Fala: “Bom, a primeira coisa que eu preciso é que você responda esse questionário dizendo se está sentindo algum desses sintomas ou não.” Entregar o formulário com o **SSQ** (*Simulator Sickness Questionnaire*).
- Fala: “Nele constam 16 sintomas, e para cada um você deve marcar como está sentindo aquele sintoma, dentre as seguintes opções: ‘Nada’, ‘Fraco’, ‘Moderado’ ou ‘Forte’.”

- Esperar o participante responder o questionário e sanar quaisquer dúvidas sobre ele, caso haja.
- Fala: “Vou colocar essa pulseira no seu pulso direito. Ela vai medir seus dados fisiológicos. Inicialmente iremos medir esses dados por 5 minutos com você em repouso. Durante esse tempo, o ideal é que você fique o mais confortável possível e de olhos fechados. Evite também falar e se mexer durante o experimento, principalmente o braço com o sensor. Quando o tempo acabar, nós te avisamos.”
- Quando o participante estiver parado, confortável e com os abafadores de ruído, **marcar evento** para início da sessão de relaxamento. Esperar 5 minutos. Ao final, **marcar outro evento**.
- Fala: “Pronto, temos a medida base. Agora podemos iniciar o experimento.”
- Colocar o HMD (*Head Mounted Display*) no participante e fazer ajustes de conforto, se necessário. Colocar o controle na mão do participante, explicar os controles e perguntar se está confortável.
- Fala: “Lembrando que você pode interromper o experimento a qualquer momento. Pedimos que tente mexer o pulso e os braços o mínimo possível, pois isso pode afetar na coleta dos dados. Quando estiver confortável, podemos começar a medição. Antes de você começar a jogar, vamos deixar o sensor calibrar por 1 minuto com você nessa posição.”
- Assim que o participante estiver com o controle em mãos, iniciar a coleta dos dados da pulseira Empatica E4 para linha de base. Depois que o minuto inicial acabar, liberar o jogador para iniciar o jogo. No começo e no final da sessão de 1 minuto, **marcar os respectivos eventos** no Empatica E4.
- Fala: “Certo, o sensor está calibrado. Quando você estiver pronto, pode começar.”
- Começar o jogo e **marcar um evento** no Empatica E4.

Fim do experimento:

- Fala: “Acabou o tempo do experimento. Você pode finalizar a última volta se desejar e irei retirar os equipamentos.”

- Marcar outro evento no sensor, remover o HMD, o controle e o Empatica E4 do participante. **Higienizar todos os equipamentos após o uso.**
- Fala: “Agora preciso que você responda mais um questionário, igual ao que você respondeu antes do experimento.” **Entregar outro formulário do SSQ.**
- Após o participante preencher o segundo SSQ, agradecê-lo pela participação.

Apêndice C

Simulator Sickness Questionnaire (SSQ)

SSQ

Respondido antes e depois da simulação

* Required

1. Nome: *

2. Idade *

3. Sexo *

Mark only one oval.

Masculino

Feminino

4. Qual a sua familiaridade com realidade virtual? *

Mark only one oval.

Nunca tive experiência

Tive pouca experiência

Já tive muitas experiências

Para cada um dos sintomas a seguir, marque a intensidade que você está sentindo agora:

5. Mal-estar generalizado *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

6. Cansaço *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

7. Dor de cabeça *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

8. Vista cansada *

Mark only one oval.

Nada

Fraco

Moderado

Forte

9. Dificuldade de manter o foco *

Mark only one oval.

Nada

Fraco

Moderado

Forte

10. Aumento de salivação *

Mark only one oval.

Nada

Fraco

Moderado

Forte

11. Sudorese *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

12. Náusea *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

13. Dificuldade de concentração *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

14. "Cabeça pesada" *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

15. Visão embaçada *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

16. Tontura com olhos abertos *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

17. Tontura com olhos fechados *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

18. Vertigem *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

19. Desconforto abdominal *

Mark only one oval.

- Nada
- Fraco
- Moderado
- Forte

20. Arroto *

Mark only one oval.

Nada

Fraco

Moderado

Forte

This content is neither created nor endorsed by Google.

Google Forms