

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA BIOMÉDICA

DÉBORA PEREIRA SALGADO

**CARGA COGNITIVA E SUA CORRELAÇÃO COM PISCADAS DURANTE O USO
DE UM SIMULADOR DE CADEIRA DE RODAS**

UBERLÂNDIA – MG

2020

DÉBORA PEREIRA SALGADO

**CARGA COGNITIVA E SUA CORRELAÇÃO COM PISCADAS DURANTE O USO
DE UM SIMULADOR DE CADEIRA DE RODAS**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Biomédica da Universidade Federal de Uberlândia, como requisito parcial para obtenção do título de Mestre em Ciências.

Área de Concentração: Engenharia Biomédica

Linha de Pesquisa: Engenharia de Reabilitação e Tecnologias Assistivas

Orientador: Prof. Eduardo Lázaro Martins Naves, Dr. (UFU)

Banca examinadora:

Prof. Dr. Eduardo Lázaro Martins Naves – Orientador (UFU)

Prof. Dr. Adriano Alves Pereira – Membro Interno (UFU)

Prof. Dr. Ederaldo José Lopes – Membro Interno (UFU)

Prof. Dr. Thiago Bruno Caparelli – Membro Externo (IFTM)

UBERLÂNDIA – MG

2020

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU com dados informados pelo(a) próprio(a) autor(a).

S164 Salgado, Débora Pereira, 1992-
2020 Carga cognitiva e sua correlação com piscadas durante
o uso de um simulador de cadeira de rodas [recurso
eletrônico] / Débora Pereira Salgado. - 2020.

Orientador: Eduardo Lázaro Martins Naves.
Dissertação (Mestrado) - Universidade Federal de
Uberlândia, Pós-graduação em Engenharia Biomédica.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2020.626>
Inclui bibliografia.
Inclui ilustrações.

1. Engenharia biomédica. I. Naves, Eduardo Lázaro
Martins, 1970-, (Orient.). II. Universidade Federal de
Uberlândia. Pós-graduação em Engenharia Biomédica. III.
Título.

CDU: 62:61

Bibliotecários responsáveis pela estrutura de acordo com o AACR2: Gizele

Cristine Nunes do Couto - CRB6/2091



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Coordenação do Programa de Pós-Graduação em Engenharia Biomédica
Av. João Naves de Ávila, 2121, Bloco 3N, Sala 115 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4761 - www.ppgeb.feelt.ufu.br - ppgeb@feelt.ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Biomédica			
Defesa de:	Dissertação de Mestrado Acadêmico, 062, PPGEB			
Data:	vinte e nove de julho de dois mil e vinte	Hora de início:	14:00	Hora de encerramento:
Matrícula do Discente:	11712EBI007			
Nome do Discente:	Débora Pereira Salgado			
Título do Trabalho:	Carga cognitiva e sua correlação com piscadas durante o uso de um simulador de cadeira de rodas			
Área de concentração:	Engenharia Biomédica			
Linha de pesquisa:	Engenharia de Reabilitação e Tecnologias Assistivas			
Projeto de Pesquisa de vinculação:	Sistema multimodal para treinamento a distância em ambiente virtual ou de realidade aumentada destinado aos usuários de cadeiras de rodas motorizadas			

Reuniu-se em web conferência pela plataforma Google Meet em conformidade com a PORTARIA Nº 36, DE 19 DE MARÇO DE 2020 da COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR - CAPES, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Biomédica, assim composta: Professores Doutores: Thiago Bruno Caparelli - IFTM; Adriano Alves Pereira - FEELT/UFU; Ederaldo José Lopes IP/UFU; Eduardo Lázaro Martins Naves - FEELT/UFU orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Eduardo Lázaro Martins Naves, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(as) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovada.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Eduardo Lazaro Martins Naves, Professor(a) do Magistério Superior**, em 29/07/2020, às 17:20, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Adriano Alves Pereira, Professor(a) do Magistério Superior**, em 29/07/2020, às 17:21, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Thiago Bruno Caparelli, Usuário Externo**, em 29/07/2020, às 17:21, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Ederaldo José Lopes, Coordenador(a)**, em 29/07/2020, às 17:22, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2162388** e o código CRC **05D5E8C8**.

Este trabalho é dedicado aos meus colegas, meu noivo, meus irmãos, aos meus queridos pais.

AGRADECIMENTOS

Considero o mestrado extremamente importante para o desenvolvimento profissional e pessoal, uma jornada marcada de superação, resiliência e autoconhecimento. Os questionamentos são frequentes e as respostas são alcançadas através do esforço e dedicação. As barreiras que enfrentamos muitas vezes são invisíveis, durante esse processo aprendi o quanto que a mente é poderosa, temos que ter controle para utiliza-la ao nosso favor e não contra nós. O encerramento desta etapa traz a motivação de desbravar novos caminhos e de compartilhar todo conhecimento adquirido.

Agradeço primeiramente a Deus que muitas vezes recorremos a Ele na procura de respostas e que sempre me mostrou que as perguntas que fazemos nem sempre são as que precisam ser respondidas naquele momento, as respostas sempre vêm com o tempo.

Agradeço aos meus pais, meus irmãos e ao meu noivo, pelo apoio incondicional e pelos valores compartilhados. Nossa conversa diária é parte importante do meu dia e que mesmo morando longe sinto vocês perto de mim.

Agradeço a participação de muitos amigos e colegas, aos quais me apoiaram e ajudaram ao longo da minha carreira acadêmica. Este trabalho não teria sido desenvolvido sem vocês.

Agradeço ao meu orientador Professor Dr. Eduardo Lázaro Martins Naves, por todo o apoio, dedicação e orientação, sendo fundamentais para realização de não somente dessa dissertação, mas também de inúmeros projetos e oportunidades de crescimento dentro e fora da Universidade Federal de Uberlândia.

Agradeço ao corpo docente da Universidade Federal de Uberlândia, os quais foram responsáveis pela a minha formação, e que os ensinamentos repassados foram cruciais para execução desse mestrado.

Agradeço à Coordenação do Programa de Pós-Graduação em Engenharia Biomédica da Universidade Federal de Uberlândia, pelos serviços prestados e apoio durante o desenvolvimento do trabalho.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro durante a graduação e a pós-graduação. Sem esse apoio não seria possível o desenvolvimento desse trabalho.

“What men and women need is encouragement. Their natural resisting powers should be strengthened, not weakened ... Instead of always harping on a man's faults, tell him of his virtues. Try to pull him out of his rut ... Hold up to him his better self, his real self that can dare and do and win out! ... People radiate what is in their minds and in their hearts.” (HODGMAN PORTER, 1913)

RESUMO

SALGADO, Débora Pereira. **CARGA COGNITIVA E SUA CORRELAÇÃO COM PISCADAS DURANTE O USO DE UM SIMULADOR DE CADEIRA DE RODAS.** 2020. 77p. Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia Biomédica da Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia-MG, 2020.

Dirigir uma cadeira de rodas é difícil e complexo, pois requer capacidade motora, visual e cognitiva. Existem simuladores voltados principalmente para usuários sem ou com pouca experiência do uso desta tecnologia assistiva como forma de treinamento segura. Porém, a cadeira de rodas real e o simulador precisam ser desenvolvidos ou aprimorados de forma que os usuários se sintam adaptados e motivados a utilizá-los. Desenvolver tecnologias para assistir as pessoas sem utilizar avaliações que ajudem a entender e aproximar os usuários dos desenvolvedores levam a desistência e desmotivação na utilização dessas ferramentas. Muitas dessas avaliações são feitas por meio de métricas subjetivas como, por exemplo, entrevistas ou aplicação de questionários; e as métricas objetivas muitas vezes são relacionadas ao desempenho do usuário. Geralmente essas métricas subjetivas são classificadas como explícitas, ou seja, são respostas que o usuário tem conhecimento. Essas métricas podem frequentemente não refletirem opiniões e sentimentos genuínos das pessoas. Em contraste, medidas implícitas inferem conteúdos mentais baseadas em tarefas que podem avaliar de forma contínua enquanto ocorre a execução da tarefa. Nesta perspectiva, o presente trabalho teve como objetivo de avaliar a carga cognitiva durante o uso da ferramenta de treinamento em realidade virtual, simulador de cadeira de rodas. O estudo propõe a utilização de métricas implícitas, especificamente, os índices das piscadas derivadas de um sinal de EEG. De modo geral, o estudo analisou métricas implícitas e explícitas de uma tarefa realizada em um simulador de cadeira de rodas em ambiente virtual. As métricas explícitas foram derivadas de questionário de usabilidade, presença, carga cognitiva e desempenho. Já as métricas implícitas foram as características de frequência, duração e velocidade das piscadas coletadas ao longo do experimento. Os resultados foram condizentes aos relatos encontrados na literatura, onde, a carga cognitiva que requer uma demanda mental-visual tem relação com a diminuição da frequência, duração e velocidade da piscada. Os resultados das análises de correlação entre os escores do NASA-TLX (avaliação da carga cognitiva subjetiva) e os parâmetros da piscada foram moderados e fracos. Assim, no presente momento, para a aplicação apresentada, conclui-se que avaliar apenas as piscadas não é o suficiente para inferir a carga cognitiva, mas é uma métrica que aplicada em conjunto com questionários pode mostrar mais informações relevantes do que somente a aplicação de questionários.

Palavras-chave: Carga Cognitiva. Simulador de Cadeira de Rodas, NASA-TLX, Piscadas, EEG, Tecnologia Assistiva, Usabilidade, Presença, Realidade Virtual.

ABSTRACT

SALGADO, Débora Pereira. **COGNITIVE LOAD AND ITS CORRELATION WITH BLINKS DURING THE USE OF A WHEELCHAIR SIMULATOR.** 2020. 79p. Dissertation (Master's) – Program in Biomedical Engineering, Faculty of Electrical Engineering, Federal University of Uberlândia, Uberlândia-MG, 2020.

Driving a wheelchair is difficult and complex, where it requires motor, visual and cognitive skills. There are simulators aimed mainly at users without or with little experience of using this assistive technology as a form of safe training. However, the real wheelchair and the simulator need to be developed or improved so that the users feel adapted and motivated to use them. Developing technologies to assist people without using assessments that help to understand and bring users closer to developers leads to giving up and discouraging the use of these tools. Many of these assessments are made using subjective metrics, such as interviews or questionnaires; and objective metrics are often related to user performance. Generally, these subjective metrics are classified as explicit, which are the responses that the user is aware of. These metrics can often not reflect people's genuine opinions and feelings. In contrast, implicit measures infer mental content based on tasks, that can be evaluated continuously while the task is being performed. In this perspective, the present work aimed to assess the cognitive load during the use of the virtual reality training tool, the wheelchair simulator. The study proposes the use of implicit metrics, specifically, the blink indexes derived from an EEG signal. In summary, the study analysed implicit and explicit metrics of a task performed in the wheelchair simulator in a virtual environment. The explicit metrics were derived from questionnaires on usability, presence, cognitive load and performance. The implicit metrics, on the other hand, were the characteristics of the frequency, duration and velocity of the blinks collected during the experiment. The results were consistent with the reports found in the literature, where, the cognitive load that requires a mental-visual demand is related to the decrease in the frequency, duration and velocity of the blink. The results of the correlation analyses between the NASA-TLX scores (subjective cognitive load assessment) and the blink parameters were moderate and weak. Thus, at the moment, for the application presented, it is concluded that evaluating only the blinks is not enough to infer the cognitive load, however it is a metric that applied in conjunction with questionnaires can show more relevant information than just the application of questionnaires.

Keywords: Cognitive Load. Wheelchair Simulator. NASA-TLX, Blinks, EEG, Assistive Technology, Usability, Presence, Virtual Reality.

LISTA DE FIGURAS

Figura 1 – Uma apresentação visual dos pressupostos subjacentes ao desenvolvimento mais recente da teoria da carga cognitiva adaptado de [34].	26
Figura 2 - Estrutura do trabalho.....	38
Figura 3 - Ambientes do Simulador: (a) pista de obstáculos; (b) percurso na rampa; (c) entrada e saída dos elevadores;.....	40
Figura 4 - Posição dos eletrodos utilizados na coleta do sinal de EEG com o dispositivo Mindwave. Os eletrodos estão representados de acordo com sistema internacional de colocação de eletrodos 10-20 [75]. Imagem adaptada de [76].	41
Figura 5 - Sistema de marcação de eventos e tempo para sincronização dos dados do simulador e do EEG.....	42
Figura 6 - Os pontos de referência da piscada “ <i>landmarks</i> ” da piscada. A figura tem o sinal de EEG no domínio do tempo na escala em segundos na porção que corresponde a piscada.....	44
Figura 7 - Exemplo de arquivo salvo das variáveis <i>leftZero</i> , <i>rightZero</i> , <i>leftBase</i> e <i>rightBase</i>	45
Figura 8 - Pontos do sinal da piscada relacionados ao movimento da pálpebra. Adaptado do [80].	46
Figura 9 - Distribuições das amplitudes máximas das piscadas retiradas de um sinal de EEG.....	48
Figura 10 - Pontos de referência da piscada da piscada relacionados ao tempo de duração.....	49
Figura 11 - Relatório de desempenho salvo pelo simulador. (a) Eventos salvos durante o percurso;(b) Evento de colisão; (c) Resumo do desempenho (tempo total, números de comandos e erros/colisões);	52
Figura 12 – Alternativas do NASA-TLX para determinar o peso dos seis fatores utilizados para o cálculo da carga cognitiva da tarefa.	54
Figura 13- Escalas do NASA-TLX [13].....	54
Figura 14 - Medidas do NASA-TLX: (a) medidas brutas; (b) peso dos fatores; (c) medida ajustada; (d) medida geral ajustada; (e) medida geral bruta;	56
Figura 15 - Etapas do experimento	59

Figura 16 - Percurso da tarefa de teste (ambiente da rampa): (a) percurso da execução da tarefa (representação 3D); (b) protocolo de direção para conclusão do percurso da rampa.	59
Figura 17 - Valores médios das métricas de desempenho.....	61
Figura 18 - Métricas de desempenho de todos os participantes. A linha pontilhada representa os valores médios.	62
Figura 19 - Valores médios de usabilidade e presença de todos os participantes.	63
Figura 20 - Médias dos escores do NASA-TLX. Valores brutos em azul e os ajustados (corrigidos) em laranja.	64
Figura 21 – Distribuição da carga cognitiva bruta (azul) e ajustada/corrigida (vermelho).....	65
Figura 22 - Frequência média das piscadas.	66
Figura 23 - Gráfico boxplot da frequência das piscadas durante baseline e o teste. .	68
Figura 24 - Distribuição da duração das piscadas. Imagem do topo é um gráfico quantil em relação às distribuições normais da duração média de <i>durationHZ</i> para cada conjunto de dados. As linhas pontilhadas vermelhas representam as distribuições normais mais adequadas. O gráfico do meio é a representação da duração média do conjunto de dados considerados “melhores”. O gráfico inferior mostra o histograma <i>durationHZ</i>	71
Figura 25 - Gráfico de dispersão (Scatter plot) com uma linha de tendência de PAVRZ por Demanda Temporal Ajustada.	74

LISTA DE TABELAS

Tabela 1 - Perguntas relacionadas a deficiência utilizadas no Censo 2010. Fonte IBGE,2010 [6].....	18
Tabela 2- Classificação de métodos para medir a carga cognitiva com base na objetividade e relacionamento causal. Adaptado de [28]	27
Tabela 3 - Tabela comparativa dos trabalhos correlatos encontrados.	31
Tabela 4 – Relação dos pontos de referência (landmarks) com o movimento das pálpebras.	46
Tabela 5 - Resumo das medidas da piscada utilizados para análise	51
Tabela 6 - Perguntas presentes no questionário sobre usabilidade e presença.	57
Tabela 7 - Resultados das perguntas sobre usabilidade e presença	62
Tabela 8 - Estatística descritiva da frequência das piscadas durante o baseline e o teste e da diferença entre eles (Teste – Baseline).....	67
Tabela 9 - Resultado do teste medidas repetidas para frequência de piscada.....	67
Tabela 10 - Estatística descritiva da duração da piscada (valores em ms)	69
Tabela 11 - Resultado da análise de diferença entre a duração das piscadas (s) entre o período baseline e o teste.	70
Tabela 12 - Estatística descritiva da razão amplitude-velocidade da piscada em 10msec(AVRZ).....	72
Tabela 13 - Diferença estatística dos valores PAVRZ e NAVRZ entre os períodos de baseline e teste.	72
Tabela 14 - Resumo dos resultados significativos do teste de correlação Spearman para os índices da piscada	74
Tabela 15 - Resumo dos resultados significativos do teste de correlação Spearman para as métricas de desempenho	75
Tabela 16 - Resumo dos resultados significativos do teste de correlação Spearman para as questões de presença (IPQ).....	76

LISTA DE ABREVIATURAS E SIGLAS

- CIF Classificação Internacional de Funcionalidade, Incapacidade e Saúde
- CSO *Central Statistics Office*
- DCU Design Centrado no Usuário
- ECG Eletrocardiograma
- EEG Eletroencefalografia
- EOG Eletro-oculografia
- GSR *Galvanic Skin Response*
- HMD *Head Mounted Display*
- IBGE Instituto Brasileiro de Geografia e Estatística
- IHM/IHC Interação Homem-Máquina/Computador
- IPQ *Igroup Presence Questionnaire*
- MS *Motion Sickness*
- NASA *National Aeronautics and Space Administration*
- OMS Organização Mundial da Saúde
- PMRT *Power Mobility Road Test*
- PNS Pesquisa Nacional de Saúde
- RV Realidade Virtual
- SUS *System Usability Scale*
- TA Tecnologia Assistiva
- TLX *Task Load Index*
- UX Experiência do Usuário

SUMÁRIO

1.	INTRODUÇÃO.....	17
1.1.	Contextualização	17
1.2.	Objetivos e Hipóteses.....	22
1.2.1.	Objetivo Geral.....	22
1.2.2.	Objetivos Específicos.....	22
1.2.3.	Hipóteses	22
1.2.3.1.	<i>Hipóteses sobre taxa da piscada.....</i>	22
1.2.3.2.	<i>Hipóteses sobre desempenho</i>	22
1.3.	Organização da Dissertação	24
2.	FUNDAMENTAÇÃO TEÓRICA E TRABALHOS CORRELATOS	25
2.1.	Carga Cognitiva	25
2.2.	Métodos Para Avaliar a Carga Cognitiva	27
2.2.1.	Sinais Fisiológicos Para Avaliar Carga Cognitiva	28
2.3.	Vantagem do Simulador de Cadeira de Rodas de RV do Tipo Desktop	29
2.4.	Trabalhos Correlatos	31
2.4.1.	Artigo 1 – “Assessment of Wheelchair Driving Performance in a Virtual Reality-Based Simulator”	32
2.4.2.	Artigo 2 – “Design of an Immersive Simulator for Assisted Power Wheelchair Driving” ..	33
2.4.3.	Artigo 3 – “Stability and Workload of the Virtual Reality-Based Simulator-2” ..	34
2.4.4.	Trabalho 4 – “Workload Detection Based on EEG Device for Teleoperation of a Mobile Robot”	35
2.4.5.	Trabalho 5 – “Mental Workload of Young Drivers during Curve Negotiation” ..	36
2.4.6.	Resumo Dos Trabalhos Avaliados	36
3.	MATERIAIS E MÉTODOS	38
3.1.	Simulador de Cadeira de Rodas	39

3.2.	Módulo de Coleta de Dados da Piscada.....	40
3.2.1.	Sistema de Sincronização.....	41
3.3.	Processamento dos Dados	42
3.3.1.	Processamento dos Sinais de EEG para Detecção das Piscadas	42
3.3.2.	Processamento das Métricas de Desempenho.....	52
3.3.3.	Processamento das Métricas Subjetivas de Esforço Cognitivo (NASA-TLX).....	53
3.3.4.	Processamento das Métricas Subjetivas de Usabilidade e de Presença	57
3.4.	Protocolo Experimental.....	58
3.5.	Análise Estatística	60
4.	RESULTADOS E DISCUSSÃO	61
4.1.	Variáveis de Desempenho.....	61
4.2.	Métricas de Usabilidade e Presença.....	62
4.3.	NASA-TLX.....	64
4.4.	Métricas da Piscada.....	66
4.4.1.	Frequência da Piscada	66
4.4.2.	Duração da Piscada	68
4.4.3.	Velocidade da Piscada	71
4.5.	Correlação do NASA-TLX com Índices da Piscada.....	73
4.6.	Correlação do NASA-TLX com Métricas de desempenho	75
4.7.	Correlação do NASA-TLX com as questões de presença (IPQ)	76
4.8.	Discussão	77
5.	CONCLUSÃO	80
5.1.	Conclusões	80
5.2.	Trabalhos Futuros	81
	REFERÊNCIAS.....	82
	ANEXO A – Formulario de Consentimento.....	91
	ANEXO B – Informações sobre a pesquisa	92

ANEXO C – Questionários aplicados.....95

1. INTRODUÇÃO

1.1. Contextualização

Determinar e mensurar o que é deficiência é uma tarefa complexa [1]. Pesquisadores de diversas áreas buscam aplicar questionários que variam conceitos e terminologias devido a impossibilidade de homogeneizar o conceito de deficiência [1]. Por isso, foi apresentado uma proposta de adotar um padrão, por exemplo, *Washington Group Disabilities*, ou se basear na Classificação Internacional de Funcionalidade, Incapacidade e Saúde (CIF). A Organização Mundial da Saúde (OMS) adotou a definição elaborada pela CIF. Portanto, OMS define a deficiência como “qualquer restrição ou falta (resultante de um dano, perda e/ou anormalidade) da capacidade de realizar uma atividade de uma maneira ou dentro da faixa considerada normal para um ser humano” [1], [2].

Pessoas portadoras de deficiência têm dificuldades de natureza física, mental, intelectual ou sensorial. Esses impedimentos são de longo prazo, os quais em interação com diversas barreiras podem levar a obstrução destas pessoas na participação plena e efetiva na sociedade em igualdades de condições com as demais. A OMS divulgou em 2011 que mais de 1 bilhão de pessoas têm algum tipo de deficiência. Isso representa aproximadamente 15% da população mundial [2].

Na Irlanda, o Central Statistics Office (CSO), “Escritório Central de Estatísticas”, constatou em abril de 2016 um total de 643.131 pessoas que declararam ter uma deficiência, 311.275 (48,4%) pessoas com deficiência eram do sexo masculino, enquanto 331.551 (51,6 %) eram do sexo feminino, o valor total representa 13,5% da população. Isso representou um aumento de 47.796 pessoas (8%) em relação a 2011 onde o número de pessoas portadoras de deficiência era de 595.335 [3].

No Brasil, segundo o Censo em 2010, 23,9 % (45,6 milhões) da população brasileira é portadora de pelo menos um tipo de deficiência [4]. Um relatório apresentado pela Pesquisa Nacional de Saúde (PNS) em 2013 estimou 200,6 milhões de pessoas residentes no território nacional. Deste total, 6,2% possuía pelo menos uma das quatro deficiências (visual, auditiva, motora e mental ou intelectual), e 1,3% era portadora de deficiência física. Sendo que 0,3% da população nasceu com deficiência física e 1% adquiriu em decorrência de doença ou acidente. Dentro desta parcela, 46,8% possuía grau intenso ou muito intenso de limitações, ou ainda não conseguia realizar as atividades habituais. 18,4% da população com deficiência física frequentava algum serviço de reabilitação [5].

Em 2018, o Instituto Brasileiro de Geografia e Estatística (IBGE) realizou uma releitura dos dados de pessoas com deficiência do Censo 2010. Na pesquisa do Censo 2010,

foi aplicado um quadro de perguntas curtas sobre as dificuldades que as pessoas possuem por conta da deficiência. A Tabela 1 apresenta um modelo de questionário utilizado com essas perguntas [6]. No censo de 2010, optou-se por identificar as pessoas com deficiência os indivíduos que responderem ter pelo menos alguma dificuldade (opção 3 da Tabela 1) em uma ou mais questões. Na releitura realizada em 2018, o IBGE classificou a pessoa com deficiência as que responderam afirmativamente que tem pelo menos muita dificuldade em uma ou mais das questões apresentadas (opção 2 da Tabela 1). Sendo assim, foi constatado que 6,7% (12,7 milhões) da população responderam afirmativamente que tem muita dificuldade (opção 1 da Tabela 1) ou não consegue de modo algum (opção 2 da Tabela 1).

Tabela 1 - Perguntas relacionadas a deficiência utilizadas no Censo 2010. Fonte IBGE,2010 [6]

6.14 Tem dificuldade permanente de enxergar? (Se utiliza óculos ou lentes de contato, faça sua avaliação quando os estiver utilizando)			
• 1 – Sim, não consegue de modo algum	• 2 – Sim, grande dificuldade	• 3 – Sim, alguma dificuldade	• 4 – Não, nenhuma dificuldade
6.15 Tem dificuldade permanente de ouvir? (Se utiliza aparelho auditivo, faça sua avaliação quando o estiver utilizando)			
• 1 – Sim, não consegue de modo algum	• 2 – Sim, grande dificuldade	• 3 – Sim, alguma dificuldade	• 4 – Não, nenhuma dificuldade
6.16 Tem dificuldade permanente de caminhar ou subir degraus? (Se utiliza prótese, bengala ou aparelho auxiliar, faça sua avaliação quando o estiver utilizando)			
• 1 – Sim, não consegue de modo algum	• 2 – Sim, grande dificuldade	• 3 – Sim, alguma dificuldade	• 4 – Não, nenhuma dificuldade
6.17 Tem alguma deficiência mental/intelectual permanente que limite as suas atividades habituais, como trabalhar, ir à escola, brincar, etc.?			
• 1 – Sim, não consegue de modo algum	• 2 – Sim, grande dificuldade	• 3 – Sim, alguma dificuldade	• 4 – Não, nenhuma dificuldade

Estima-se que nos próximos anos o número de pessoas com deficiência aumentará, devido ao fato de que a expectativa de vida é maior em comparação à última década e a probabilidade de adquirir deficiência com idade avançada é maior. Diante dessa realidade, há uma demanda crescente por tecnologias que ajudam as pessoas com alguma deficiência a ter

uma melhor qualidade de vida. A área de Tecnologia Assistiva (TA) é a responsável pelo desenvolvimento dessas tecnologias. De acordo com a OMS, dispositivos e tecnologias de assistência são aqueles cujo principal objetivo é manter ou melhorar a capacidade fundamental e a independência de um indivíduo para facilitar a participação e melhorar o bem-estar geral [2]. Elas também podem ajudar a prevenir deficiências e condições secundárias de saúde. Os diferentes tipos de TA são classificados como de comunicação alternativa, mobilidade, prótese e órteses, entre outros. Esses dispositivos foram projetados para ajudar as pessoas com deficiência a realizar tarefas diárias que não podem executar naturalmente, como se locomover ou se comunicar. Dessa forma, a TA pode proporcionar acessibilidade, autonomia, reintegração à sociedade e melhoria de sua autoestima e qualidade de vida.

Os dispositivos assistivos são parte integrante da assistência médica e fornecidos pelo sistema nacional de saúde. Em outros países, TA são fornecidas pelos governos através dos serviços de reabilitação, de reabilitação vocacional ou centros/agências de educação especial, organizações filantrópicas e não governamentais [1]. As pessoas que necessitam desses dispositivos precisam passar por um acompanhamento especializado, por exemplo, serviços de reabilitação e/ou terapia acompanhados por fisioterapeutas ocupais, médicos e outros profissionais capacitados.

Esse profissionais prescrevem o dispositivo mais adequado, pois eles podem ser oferecidos em diferentes configurações. A prescrição correta do dispositivo é importante para que as necessidades dos usuários sejam atendidas. Quando as necessidades não são satisfeitas podem causar a desistências na utilização desses dispositivos [1], [7]. Outros possíveis motivos de desistência foram constatados através do estudo de Emel Cruz [7]. Foi feita uma pesquisa em um grupo de 91 participantes, onde 18% dos dispositivos de TA foram abandonados. Os motivos mais comuns que levaram o usuário a abandonar sua TA foram pela falta de envolvimento do mesmo no desenvolvimento e seleção de TA, baixo desempenho, alterações nas necessidades do usuário que o dispositivo não era capaz de suprir, falta de treinamento, baixa aceitação social, falta de motivação, dispositivos com operação complicada e dispositivos sem aparência estética.

Apesar da importância da TA para pessoas com deficiências físicas, existem dificuldades na utilização da TA de forma adequada e os treinamentos em situações reais apresentam altos níveis de riscos de segurança. Assim, é possível perceber a importância e a necessidade das ferramentas de treinamento para ensinar as pessoas a usar os dispositivos de TA com segurança. Dentre as ferramentas para treinamento de dispositivos de TA, destaca-se as ferramentas desenvolvidas para treinamento da utilização de cadeiras de rodas que serão

objeto de estudo nesse trabalho e para as quais foram desenvolvidos ambientes para treinamento em realidade virtual (RV).

RV veio como proposta para fornecer um serviço de treinamento personalizado que possam ser oferecidos em vários lugares (descentralização da oferta de serviços de treinamento) e com a possibilidade de oferecer um acompanhamento remoto. Além disso, RV oferece baixo/ nenhum risco e o ambiente desenvolvido é controlado com possibilidades de ajustes de acordo com as necessidades do usuário [8]. Entretanto, os simuladores em RV ainda é considerado uma tecnologia emergente, ou seja, ainda possuem alguns obstáculos técnicos. Por exemplo, ainda é comum ocorrem problemas de experiência do usuário na fase design e desenvolvimento do sistema/ aplicativo. Em outras palavras, designers ainda encontram desafios para entenderam quais restrições importantes de serem impostas para que o sistema seja aceito pelos usuários [9]. Sendo assim, o simulador de cadeira de rodas em RV enfrenta desafios como qualquer outro sistema em RV, precisa de estudos de avaliação do sistema para promover um nível mais alto de aceitação da tecnologia entre os usuários. Para isso, o simulador de cadeira de rodas precisa ser avaliado para evitar o abandono por motivos de falta de confiança, interesse e motivação no uso do sistema. Pesquisas mostraram que mais de 30% de todos os dispositivos TA são abandonados pelos usuários em um período de cinco anos de uso, e que alguns não são usados sequer uma vez [10], [11].

Diante dessa realidade, é possível perceber a necessidade de métodos para aproximar o desenvolvedor e o usuário do sistema, por exemplo, aplicando design centrado no usuário (DCU) e interação homem-máquina/computador (IHM/IHC), para que resulte em um aplicativo válido [12]. O DCU utiliza uma combinação de métodos subjetivos explícitos (e.g. questionários e entrevistas). Os dados coletados são analisados de forma que ampliem o entendimento das necessidades dos usuários e assim podem apresentar novas soluções para aprimorar o produto, sistema ou aplicação. Além dos estudos subjetivos, são aplicados métodos de análise objetiva implícita e/ou explícita, por exemplo, métricas de performance (número de acertos, erros, tempo de aprendizado, etc.). Outra forma de aprimoramento do sistema é investigando qual a carga cognitiva do usuário durante a utilização do sistema. Pois, uma sobrecarga cognitiva pode afetar a qualidade da experiência do usuário e o aumento da incidência de erros [13].

Sendo assim, pode-se dizer que a carga cognitiva é um fator importante ao avaliar o sistema, e é considerada parte do DCU, onde é fundamental que o sistema seja mais intuitivo.

Para isso é necessário seguir as orientações apresentadas na “Teoria da Carga Cognitiva” [13], onde o objetivo é minimizar a atividade cognitiva que não contribui para a aprendizagem, trabalhando com objetivos específicos de aprendizagem em mente e de relação multifacetada entre aprendizagem e avaliação. As principais questões relacionadas aos tipos de carga cognitiva incluem o contexto em que a aprendizagem ocorre, o uso contínuo de classificações de esforço mental e da carga cognitiva; e das medidas dos resultados da aprendizagem. Nesse contexto, é complexo desenvolver estratégias ideais para melhorar a aplicação, no caso o simulador de cadeira de rodas, sem primeiro saber a demanda de carga cognitiva do usuário durante a interação.

A carga cognitiva é constantemente avaliada de forma subjetiva por meio de questionários como “*Paas Cognitive Load Scale*” (Paas Scale) [14] e NASA “*Task Load Index*” (TLX) [15]. Entretanto, a análise cognitiva baseada somente em métodos subjetivos é bastante limitada. Pois os escores calculados por análise puramente subjetiva não refletem a carga cognitiva total [16]. Uma forma de reduzir o erro ao estimar a carga cognitiva é utilizar métricas objetivas implícitas que refletem a carga cognitiva. Por exemplo, recentes estudos mostraram que as características derivadas da piscada do usuário podem ser utilizadas para inferir a carga cognitiva de maneira implícita [17]. O principal benefício de conseguir inferir a carga cognitiva de forma implícita é a capacidade de monitorar a carga de trabalho mental do usuário de forma simultânea à tarefa e isso pode ajudar na adequação da interface e interações ao status cognitivo atual do usuário e alterá-las adequadamente quando necessário.

Em resumo, essa dissertação de mestrado apresenta um estudo de carga cognitiva implícita e explícita de sistema de treinamento. O sistema de treinamento utilizado neste estudo é o simulador de cadeira de rodas desenvolvido com tecnologia de realidade virtual (RV), o EWATS (*Electric Wheelchair Assistive Technology System*) [18]. Foram coletados sinais das piscadas enquanto participantes completavam uma tarefa e foi aplicado o questionário TLX (conhecido na literatura para avaliação da carga cognitiva em simuladores [19]) com o objetivo de inferir a carga cognitiva do sistema tanto de forma implícita quanto explícita.

1.2. Objetivos e Hipóteses

1.2.1. Objetivo Geral

Em vista da problemática e fatos apresentados, o presente trabalho tem como objetivo avaliar a carga mental de forma implícita e explícita associada ao uso da ferramenta de treinamento em realidade virtual.

1.2.2. Objetivos Específicos

A fim de se alcançar o objetivo geral proposto, alguns objetivos específicos foram estipulados. São eles:

- Ajustar o simulador para coletar dados fisiológicos de forma síncrona
- Ajustar o simulador para coletar o desempenho dos participantes ao utiliza-lo
- Realizar avaliações explícitas por meio de questionários (NASA-TLX e usabilidade).
- Realizar avaliação implícita por meio de coleta dos sinais da piscada
- Analisar e avaliar os resultados obtidos na utilização do simulador segundo os dados fisiológicos e os questionários aplicados.

1.2.3. Hipóteses

Foram elaboradas hipóteses sobre parâmetros da piscada e métricas de desempenho dos participantes com relação a carga cognitiva exigida durante a condução de uma cadeira de rodas virtual. As hipóteses estão descritas nas subseções abaixo.

1.2.3.1. Hipóteses sobre taxa da piscada

Alta taxa de piscada indica poucos itens e baixa taxa de piscada indica um grande número de itens que precisam ser processados em um único tipo de tarefa verificado em estudos anteriores [20]–[25]. Sendo assim, espera-se observar que:

H01: Os parâmetros da piscada (taxa, duração e razão amplitude-velocidade) aumentam durante tarefas de baixa carga cognitiva (baseline) e diminuem durante tarefas de alta cognitiva (teste).

H02: Os parâmetros da piscada estão correlacionadas com as medidas de carga cognitiva do NASA-TLX.

1.2.3.2. Hipóteses sobre desempenho

Estudos apresentados na literatura argumentam que a sobrecarga cognitiva pode afetar o desempenho da tarefa (e.g. conclusão da tarefa) [14], [26]–[28]. Onde os valores de carga

cognitiva são diferentes de acordo com o tipo de interação e características dos usuários [14], [29], [30]. Sendo assim, espera-se investigar a seguinte hipótese:

H0₃: As métricas de desempenho (número de comandos, erros (colisões) e tempo de execução da tarefa) estão correlacionadas com as medidas do NASA-TLX.

1.3. Organização da Dissertação

A metodologia utilizada para alcançar o objetivo principal deste trabalho, conforme citado anteriormente, será descrita nos capítulos subsequentes, os quais seguem a seguinte estrutura:

Capítulo 2 – este capítulo reúne um conjunto de trabalhos correlatos para descrever a fundamentação teórica aplicada acerca dos temas envolvidos neste estudo. Sendo eles, estudos em torno da carga cognitiva e para simuladores de realidade virtual aplicados em treinamento.

Capítulo 3 – apresenta os materiais, ferramentas e tecnologias utilizadas para o desenvolvimento do trabalho, incluindo a metodologia adotada para coletar dados dos participantes.

Capítulo 4 – apresenta os resultados e discussão do estudo realizado com 25 participantes que experimentaram o simulador.

Capítulo 5 – apresenta as considerações finais concluindo o estudo. No qual o objetivo foi avaliar a carga cognitiva associada ao uso do simulador de maneira implícita e explícita. Abordando-se ainda os resultados obtidos e possíveis ajustes para trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA E TRABALHOS CORRELATOS

Neste capítulo, serão abordados tópicos que apresentam a fundamentação da teoria que serve de embasamento para o tema desta pesquisa, sendo eles:

- I. Carga cognitiva;
- II. Métodos para avaliar a carga cognitiva de forma explícita (questionários subjetivos) e implícitas (com ênfase sinais biomédicos).
- III. Vantagem do uso do simulador de cadeira de rodas não imersivo

O outro objetivo do capítulo é apresentar pesquisas relacionadas a este tema. O foco principal desta busca foram trabalhos que apresentam aplicações de treinamento de cadeira de rodas usando simuladores em RV, quais métodos de avaliação foram aplicados e se foi avaliada carga cognitiva da tarefa.

2.1. Carga Cognitiva

Estudos relacionados com IHC e DCU não são as únicas áreas de interesse em medidas de carga cognitiva. Área de estudos em usabilidade e experiência do usuário (UX) são outras áreas em que o conceito de carga cognitiva é aplicável [27]. Além disso, a detecção de sobrecarga cognitiva é importante para manter a saúde mental do usuário durante a interação sustentada [31]. À medida que esses requisitos aumentam, a quantificação da carga cognitiva do usuário se torna uma etapa essencial para o progresso dessas áreas de pesquisa.

Na psicologia cognitiva, a carga cognitiva refere-se à quantidade total de esforço mental que está sendo usado na memória de trabalho. A teoria da carga cognitiva foi desenvolvida a partir do estudo de resolução de problemas de John Sweller no começo dos anos 90 [32], [33]. Sweller argumentou que o design instrucional pode ser usado para reduzir a carga cognitiva nos alunos. A teoria da carga cognitiva diferencia a carga cognitiva em três tipos: intrínseca, estranha e pertinente. A Figura 1 mostra as cargas cognitivas que constituem a carga total. Carga cognitiva intrínseca é o esforço associado a um tópico específico. Carga cognitiva estranha ou ineficaz refere-se à maneira como as informações ou tarefas são apresentadas a um aluno. E a carga cognitiva pertinente ou efetiva refere-se ao trabalho dedicado à criação de um estoque permanente de conhecimento, ou um esquema.

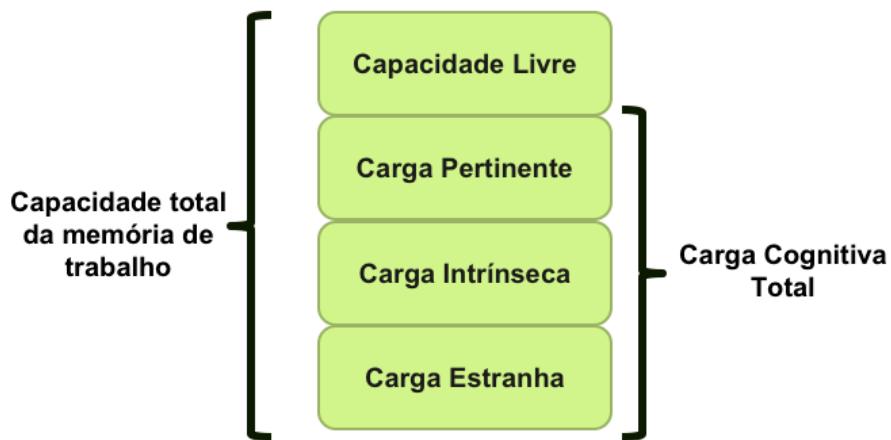


Figura 1 – Uma apresentação visual dos pressupostos subjacentes ao desenvolvimento mais recente da teoria da carga cognitiva adaptado de [34].

Os pesquisadores Paas e Van Merriënboer desenvolveram uma maneira de medir o esforço mental percebido, que é indicativo de carga cognitiva [35]. Os autores descrevem que o esforço mental pode ser medido com técnicas subjetivas, que usam escalas de classificação, e com técnicas objetivas que usam parâmetros fisiológicos. Além disso, os autores comentam que medidas de desempenho juntamente com essas informações de esforço mental pode fornecer informações sobre os gastos cognitivos para atingir o desempenho e informações sobre eficiência relativa das condições instrucionais da tarefa [35], [36].

A exploração de características fisiológicas como resposta galvânica da pele (GSR) [37], eletro-oculografia (EOG) [24], [38] e rastreamento do globo ocular (eye tracking) [23], eletroencefalografia (EEG) [39], [40] e parâmetros como diâmetro da pupila e da frequência de piscada tem sido aplicadas para inferir a carga cognitiva em vários domínios e aplicações. Por exemplo, essas medidas foram utilizadas para inferir a carga cognitiva durante tarefas de voo e controle de tráfego aéreo para monitorar o operador em situações críticas, outro exemplo, essas medidas foram estudadas para análise de tarefas para melhorar a interação humano-computador e para construção de treinamento eficaz e personalizado pelo usuário [17], [20], [22], [23].

Uma sobrecarga cognitiva pode ter efeitos negativos na conclusão da tarefa, e é importante observar que a experiência da carga cognitiva não é a mesma para as pessoas [14]. Os valores de carga cognitiva são diferentes de acordo com interação e características dos usuários [14]. Por exemplo, um estudo apresentou diferentes valores de carga cognitiva de acordo com à idade e formato da tarefa (instrução) [30]. Sendo assim, carga mental pode ser

definida como um aspecto da carga cognitiva que se origina da interação entre características da tarefa e do sujeito [14], [29], [30].

2.2. Métodos Para Avaliar a Carga Cognitiva

A carga cognitiva pode ser tratada como um construtor teórico, descrevendo os processos internos de processamento de informações que não podem ser observados diretamente [28]. No entanto, para usar essa teoria para esclarecer o design da instrução, são necessários instrumentos válidos e confiáveis para avaliar a carga cognitiva.

Os métodos de avaliação da carga cognitiva de acordo com Brunken and Plass podem ser classificados em duas dimensões, objetividade (subjetiva ou objetiva) e relação causal (direta ou indireta) [28]. A dimensão objetividade descreve se o método usa dados subjetivos e/ou objetivos, por exemplo, informações auto-relatadas, observações de comportamento, condições fisiológicas e desempenho. A dimensão da relação causal classifica métodos com base no tipo de relação do fenômeno observado pela medida (Tabela 2).

Tabela 2- Classificação de métodos para medir a carga cognitiva com base na objetividade e relacionamento causal. Adaptado de [28]

Objetividade	Relacionamento Causal	
	Indireto	Direto
Subjetivo	<ul style="list-style-type: none"> • Esforço mental investido auto relatado 	<ul style="list-style-type: none"> • Nível de estresse auto relatado • Dificuldade auto referida dos materiais
Objetivo	<ul style="list-style-type: none"> • Medidas Fisiológicas • Medidas comportamentais • Medidas de resultados de aprendizagem 	<ul style="list-style-type: none"> • Medidas da atividade cerebral (e.g., FMRI) • Desempenho de dupla tarefa

Sendo assim, as descobertas empíricas que apoiam os princípios da teoria da carga cognitiva podem ser resumidas em quatro tipos de medidas:

- I. **Medida indireta subjetiva**, usa questionários após a tarefa, nos quais os participantes são solicitados a relatar a quantidade de esforço mental investido para compreender e/ou concluir a tarefa [26]. NASA-TLX [15] é um exemplo de um método de avaliação indireta e subjetiva da carga de trabalho mental que tem suas origens em pesquisas sobre carga de trabalho mental.
- II. **Medida direta subjetiva**, envolvem classificação da dificuldade do material/tarefa diretamente relacionado à carga cognitiva imposta [28], [41].
- III. **Medida objetiva indireta**, método mais comum de investigar os efeitos da carga cognitiva é analisar medidas de resultados de desempenho na área de aprendizagem com multimídia [26], [27]. Outra forma é por medidas fisiológicas como frequência cardíaca, resposta galvânica da pele e variáveis de rastreamento ocular [42]

IV. **Medida objetiva direta**, desempenho de dupla tarefa (*dual-task*) [28]. Outra forma é por medidas de imagens médicas, tais como ressonância magnética funcional (fMRI) [43], ou eletroencefalográfico específico (EEG) [44], para medir a ativação cerebral durante execução de tarefas.

Outro termo de classificação é implícito e explícito. Medidas implícitas avaliam o comportamento do participante na ausência da consciência dos entrevistados sobre essa influência. Aplicação de medição explícita é o oposto, pedem diretamente às pessoas que descrevam suas atitudes e são representadas por várias medidas tradicionais de auto relato (entrevistas e questionários) que são comumente usadas em pesquisa.

2.2.1. Sinais Fisiológicos Para Avaliar Carga Cognitiva

Vários sinais fisiológicos têm sido utilizados para a medição da carga cognitiva: eletrocardiograma (ECG) [35], [41], [45], EOG [24], [38], EEG [39], [40] e GSR [37]. O GSR é a condutância elétrica da pele, um sinal fisiológico robusto e de baixo custo, facilmente capturado. Estudos já usaram a condutância da pele na detecção de emoções ou na diferenciação entre estresse e condições de carga cognitiva [46], [47], e alguns encontraram relações entre as características do GSR e a carga mental de trabalho [48]. Alguns outros tentaram, mas não obtiveram resultados satisfatórios para detectar a carga cognitiva a partir do GSR [49], [50].

Sinais que refletem comportamento como fala [51] e movimentos oculares [23] podem ser usados para medição da carga cognitiva. A atividade ocular pode revelar informações sobre a carga de trabalho mental. Informações como dilatação da pupila por meio de rastreador ocular podem ser captados por câmeras ou com rastreadores que são de alto custo. A piscada pode ser obtida com uma precisão aceitável por meio de uma câmera convencional ou EEG de canal único de baixo custo.

Portanto, o sinal da piscada é mais fácil de se obter quando comparado com sinal de EEG e pode ser captado a baixo custo, e pode ser usado para medição de carga cognitiva. Alguns trabalhos anteriores estudaram variações de piscada em relação à modalidade visual e acústica [52], localização central e periférica [21]. Outra pesquisa mediu a taxa de piscada nas condições de repouso, leitura e conversação [25]. Alguns estudos examinaram os recursos de piscada em dois níveis de carga cognitiva e os encontraram relacionados ao nível de carga mental [22], [53].

Em alguns domínios, como área de mobilidade e TA, é essencial poder medir carga cognitiva a baixo custo, com pouco tempo de preparação e pouca restrição aos movimentos

dos usuários. Considerando que o piscar de olhos é uma métrica que atende esses requisitos, nesta dissertação foram exploradas algumas características do sinal da piscada capturado durante a fase de descanso (baseline) e durante o uso do simulador. Essas fases foram separadas dessa forma para mostrar uma distinção entre os dois tipos demanda cognitiva (alta (teste) e baixa (descanso/baseline)).

Além disso, foi aplicado o questionário NASA-TLX (também conhecido como TLX), como medida subjetiva indireta explícita. O questionário NASA-TLX foi aplicado como forma de validar o uso de parâmetros do sinal de piscada (objetiva indireta implícita) como medida de carga cognitiva, uma medida objetiva indireta implícita, durante a utilização do simulador de cadeira de rodas no formato RV desktop. O principal motivo da escolha do NASA-TLX em vez de outros métodos existentes foi devido ao fato de ser a medida de classificação subjetiva usada mais amplamente do que outros questionários, como escala de “*Paas Cognitive Load Scale*” (Paas Scale) [14], [54].

2.3. Vantagem do Simulador de Cadeira de Rodas de RV do Tipo Desktop

Realidade virtual (RV) pode proporcionar que os usuários fiquem imersos em ambientes 3D usando headsets de RV também conhecidos em inglês como head mounted displays (HMD), onde sua aplicação vai além de entretenimento. Por exemplo, RV tem sido aplicada em diversas áreas para treinamento na utilização de máquinas industriais, procedimentos cirúrgicos, reabilitação e ensino na área acadêmica [55]–[60]. No entanto, a RV imersiva ainda possui limitações devido ao cybersickness e, por um tempo, o alto custo mantiveram a RV fora da maioria das áreas de interesse. Atualmente, a tecnologia de RV tornou-se mais acessível em termos de custos e possíveis soluções para o cybersickness têm sido constantemente investigadas.

Cybersickness refere-se a doença de movimento (motion sickness (MS)) enquanto experimenta uma aplicação de tecnologia imersiva, também é conhecido como doença de RV. Cybersickness é classificado como um subtipo de MS e recentemente ganhou muita atenção por causa do surgimento de novos headsets RV [61], [62]. De acordo com a teoria da “incompatibilidade sensorial”, a doença pode ser causada por disparidade entre estímulos visuais e feedback proprioceptivo [62]. Por exemplo, a exposição a um ambiente virtual com cenas em movimento pode causar cybersickness, onde o usuário está recebendo estímulo visual de movimento, porém o sistema vestibular, responsável pela detecção de movimentos do corpo, recebe informações conflitantes de que o corpo não está em movimento. No entanto, outros

tipos de aplicativos de RV podem não ser afetados pela doença, onde os usuários de RV estão em posição estacionária (em pé ou sentados) e o ambiente virtual envia o mesmo tipo estímulo.

Sistemas de ambientes virtuais para computadores de mesa (desktop) e laptops são conhecidos como RV não imersivos (de baixa imersão) ou RV para desktop [63]. A maior vantagem do sistema RV não imersivo é ausência de sintomas de cybersickness. RV para desktop tem sido amplamente pesquisado e utilizado em muitos campos, isso inclui educação, treinamento, exporte, planejamento urbano, exploração espacial, design industrial, medicina e reabilitação, entretenimento, construção de modelos e pesquisa em muitos campos da ciência.

A proposta de desenvolver ambientes virtuais como ferramenta de treinamento e reabilitação para usuários de cadeira de rodas tem sido apresentada em outros estudos. Por exemplo, o projeto de simulador para treinar crianças com deficiências para operar cadeiras de rodas elétricas [64]. O objetivo do trabalho foi avaliar a capacidade de um simulador em treinar e avaliar crianças para operar cadeiras de rodas. Os resultados encontrados foram que os usuários inexperientes aumentaram significativamente seu desempenho usando o simulador. No entanto, o desempenho foi menor para os usuários experientes. No geral, pode-se confirmar que um simulador pode ajudar no desenvolvimento e avaliação das habilidades necessárias para operar uma cadeira de rodas elétrica.

Os simuladores para treinamento mais recentes concentram-se em proporcionar melhores experiências de interação e imersão do usuário, com simulações mais realistas e com sensores para monitoramento de comportamento físico dos usuários [65], [66]. No entanto, o cybersickness é o efeito colateral desses simuladores que precisa ser resolvido para que a tecnologia seja completamente disseminada [60].

Conclui-se que RV imersivo ou não tem potencial para ser uma ferramenta de treinamento na área de TA e reabilitação. A criação de ambientes controlados permite estudos com diversos estímulos (feedback), para que as respostas comportamentais possam ser registradas e medidas [67]. Por exemplo, métricas de desempenho do usuário e as respostas fisiológicas podem ser coletadas ao longo de toda a experiência do usuário no simulador [68].

2.4. Trabalhos Correlatos

Dentre os trabalhos avaliados disponíveis na literatura, alguns deles apresentam, conceitos e referência teóricas relacionadas ao tema em questão. Entre todos os trabalhos encontrados, foram selecionados 5 trabalhos que apresentam as contribuições relevantes para a pesquisa. As palavras-chave para pesquisa de trabalhos correlatos foram feitas em inglês da seguinte forma:

- Wheelchair Simulator or Driving (simulador de cadeira de rodas ou simulador de direção)
- Virtual Reality (realidade virtual)
- Cognitive Load (carga cognitiva)

A Tabela 3 apresenta um estudo comparativo entre as principais características presentes nos trabalhos correlatos. Foram analisados a presença dos seguintes aspectos: (i) presença de RV; (ii) tipo de display utilizado; (iii) presença de sintomas de cybersickness; (iv) quais os métodos de avaliação subjetivo e/ou objetivos da carga cognitiva; (v) presença de outras avaliações.

Tabela 3 - Tabela comparativa dos trabalhos correlatos encontrados.

Trabalhos	Tem RV	Tem display	Tem avaliação de carga	Tem avaliação cognitiva	Outras avaliações	Sintomas de cybersickness
1. [69]	Sim	Desktop e HMD	Não	Não mencionado	PMRT	Não mencionado
2. [70]	Sim	Desktop e HMD	Sim	NASA-TLX	Desempenho, IPQ e QoE	Sim
3. [19]	Sim	CAVE	Sim	NASA-TLX	PRMT e desempenho	Não mencionado
4. [71]	Sim	Desktop	Sim	NASA-TLX E EEG	Desempenho	Não mencionado
5. [72]	Sim	NA	SIM	NASA-TLX e HR	NA	Não mencionado
Presente Trabalho	Sim	Desktop	Sim	Piscada e NASA-TLX	Desempenho, SUS e IPQ	Não

2.4.1.Artigo 1 – “Assessment of Wheelchair Driving Performance in a Virtual Reality-Based Simulator”

Esse trabalho teve o objetivo de avaliar o desempenho dos usuários durante a utilização do simulador de cadeira de rodas em realidade virtual [69]. O simulador foi desenvolvido para auxiliar profissionais durante os procedimentos de avaliação de forma padronizada. Os participantes controlaram a cadeira de rodas virtual ao longo de um circuito pré-definido de acordo com a metodologia de avaliação definida conhecida como *Power Mobility Road Test (PMRT)*. O simulador foi utilizado no modo desktop (não imersivo) e no modo imersivo com HMDs. Os participantes interagiram com sistema com um joystick de cadeira de rodas adaptado para o sistema.

O desempenho de condução dos participantes foi estimado e comparado usando-se métricas quantitativas do simulador. Avaliações qualitativas de dois clínicos experientes foram usadas para estimar confiabilidade entre avaliadores. Participaram do estudo 11 usuários regulares de cadeira de rodas, dentro deste número um participante foi excluído da análise por não conseguir concluir nenhum dos testes de direção virtual por sentir fadiga e tontura (sintomas de cybersickness).

Os resultados apresentados mostraram uma diferença significativa no tempo gasto para completar a tarefa entre os modos imersivo e o não imersivo do simulador. Os participantes levaram cerca de 111,97 segundos a mais para concluir os testes no modo imersivo (RV). Em relação a carga cognitiva, o simulador no modo imersivo tende a exigir uma carga cognitiva maior, os autores utilizaram esse argumento para justificar o resultado da direção mais lenta. Foi mencionado a necessidade de mais pesquisa para explorar preferências do usuário (usabilidade) e eficácia dos dois métodos do simulador.

Em modo geral, o estudo levanta pontos importantes sobre o desempenho e carga cognitiva no modo RV e desktop. Além disso o estudo seguiu um protocolo acompanhado de profissionais para avaliação. Para avaliar o desempenho foi utilizado um software validado chamado de VPMRT (virtual power mobility road test) para comparar a trajetória feita pelos os usuários com a trajetória ideal. Porém, os resultados relacionados aos valores de carga cognitiva não foram apresentados de forma clara, onde apenas foi mencionado qual modo teve uma carga cognitiva maior. Os autores mencionam que no final de cada teste foi perguntado sobre a preferências dos usuários a respeito do simulador sem a utilização de questionários conhecidos na literatura para inferir carga cognitiva, cybersickness, usabilidade, etc.

2.4.2.Artigo 2 – “Design of an Immersive Simulator for Assisted Power Wheelchair Driving”

O artigo apresenta um simulador imersivo para treinamento de condução assistida de uma cadeira de rodas elétrica [70]. O objetivo apresentado neste estudo foi validar o uso do simulador em termos de qualidade de experiência (QoE) do usuário. QoE é considerado um método de avaliação usado para determinar quão bem esse sistema está atendendo aos requisitos do usuário final. O estudo apresentou métricas de QoE em termos de presença, desempenho e carga cognitiva.

Foi realizado experimentos com nove participantes saudáveis que executaram tarefas de navegação no simulador com ou sem assistência de direção. As métricas de desempenho utilizadas foram o (i) números de comandos, (ii) posição e velocidade da cadeira de rodas virtual, (iii) colisões, dados de sensor virtual e (iv) conclusão do tempo. Depois de completar os percursos de obstáculos no simulador, foi pedido aos participantes para completar dois tipos de questionários, o Igroup Presence Questionnaire (IPQ) [73] e o NASA-TLX [15].

IPQ é utilizado para classificar a sensação de presença experimentada em um ambiente virtual, também conhecido como nível de presença ou sensação de presença. Sensação/nível de presença é compreendido como a sensação subjetiva de estar em um ambiente virtual. Além disso, é importante ressaltar que a sensação de presença pode ser separada da capacidade de uma tecnologia de envolver o usuário. Embora essa imersão seja uma variável da tecnologia e possa ser descrita objetivamente, a presença é uma variável da experiência do usuário. Portanto, as medidas da sensação de presença a partir de escalas de avaliação subjetivas. Em outras palavras, IPQ é uma escala para medir a sensação de presença experimentada em um ambiente virtual. Este teste consiste 4 categorias: (i) presença espacial, (ii) realismo, (iii) envolvimento e (iv) impressão geral avaliada usando uma escala Likert de 7 pontos.

O segundo questionário aplicado foi o NASA -TLX [15] utilizado para inferir (classificar) a carga cognitiva de forma subjetiva. Esta classificação é utilizada para avaliar a eficácia e outros aspectos do desempenho de uma tarefa, sistema ou equipe. NASA-TLX consiste em duas partes, a carga de trabalho total é dividida em seis sub-escalas subjetivas (demanda mental, física, temporal, desempenho, esforço e frustração). A segunda parte cria uma ponderação individual dessas sub-escalas, permitindo que os participantes as comparem aos pares com base em sua importância percebida. Isso requer que o usuário escolha qual medição é mais relevante para carga de trabalho. O número de vezes que cada um é escolhido é a pontuação ponderada. Cada sub-escala tem uma pontuação ponderada que é utilizado para

obter uma carga de trabalho total ajustada também chamado de “TLX ajustado/ponderado”. O valor da carga de trabalho total sem levar em consideração essa pontuação ponderada é chamado de “TLX bruto”.

O nível de presença foi mensurado usando os valores de média e o desvio padrão para todos escores das quatro categorias, nos quais, todos tiveram uma média acima de 3. As pontuações do NASA-TLX mostraram pontuações com média de $27,2 \pm 18,2$ DP sendo que dos 9 participantes, 3 tiveram valores acima da média. O autor não menciona quais valores do NASA-TLX foram usados, bruto ou ajustado. A explicação sobre as pontuações do NASA-TLX não entrou em detalhes nas sub-escalas e o autor argumenta que os valores encontrados foram baixos utilizando a afirmação que os resultados foram abaixo de 50 como suporte. Entretanto, essa afirmação não teve apoio de referências.

Foi relatado que 6 dos 9 participantes apresentaram sintomas de cybersickness durante e / ou após o experimento, sendo que 4 deles tiveram dores de cabeça. Foi relatado que os sintomas de cybersickness não impediu que os usuários completassem as tarefas. Os resultados dos experimentos mostram que os participantes colidiram um total de 25 vezes sem assistência ao usuário, enquanto apenas um participante colidiu uma vez com um obstáculo quando a assistência de navegação foi ativada.

Em suma, o estudo apresentou resultados em termos de carga cognitiva utilizando um questionário o NASA-TLX e para identificar os níveis de presença “imersão” foi utilizando um questionário também conhecido. A diferença significativa reportada foi apenas em termos de desempenho onde o sistema no modo de assistência ativado apresentou melhores resultados. Foi relatado sintomas de cybersickness porém não foi aplicado nenhum questionário ou métrica para inferir os níveis desses sintomas.

2.4.3. Artigo 3 – “Stability and Workload of the Virtual Reality-Based Simulator-2”

Objetivo do trabalho foi avaliar quatro interfaces homem-máquina diferentes para o simulador de cadeira de rodas em RV chamado de VRSIM-2. Foram utilizadas telas imersivas do tipo caverna digital CAVE.

A avaliação do sistema foi focada em termos de carga cognitiva e estabilidade do simulador [19]. Os participantes deste estudo foram compostos por 21 usuários em cadeira de rodas elétrica. Todos eles controlaram o simulador com quatro diferentes interfaces e fizeram uma tarefa na cadeira de rodas real, no total eles experimentaram cinco condições de condução (quatro virtuais e uma real). A Metodologia adotada foi o de medidas repetidas dentro dos

sujeitos. Foi aplicado o PMRT (o mesmo aplicado no artigo 1) e para avaliar a carga cognitiva o NASA-TLX (o mesmo aplicado no artigo 2).

Resultados encontrados foram que houve estabilidade moderada (coeficiente de correlação intraclasse (CCI) entre 0,5 - 0,75) nas pontuações totais do PMRT ($p < 0,001$) e nas pontuações de desempenho auto referidas pelos usuários ($p < 0,001$) entre as cinco condições de condução. Houve uma diferença significativa na carga de trabalho entre as cinco condições de condução diferentes, refletidas pelo NASA-TLX ($p = 0,009$). As sub análises revelaram que essa diferença se devia à diferença nas sub escalas de demanda mental ($p = 0,007$) e frustração ($p = 0,007$). As análises post hoc revelaram que essas diferenças nas pontuações da sub escalas do NASA-TLX eram devidas às diferenças entre as pontuações na condução real e as virtuais, ou seja, não houve diferença entre os quatro modos de condução virtual.

Em síntese, o estudo aplicou o PMRT e o NASA-TLX para avaliar o simulador em termos de estabilidade e carga cognitiva. Os usuários não foram avaliados em termos de cybersickness e imersão, onde não foi relatado nenhum dos sintomas de cybersickness e aspectos “imersão” presença dos usuários o estudo.

2.4.4. Trabalho 4 – “Workload Detection Based on EEG Device for Teleoperation of a Mobile Robot”

O artigo apresenta um método de detecção de carga de trabalho baseada no sinal de EEG durante a tarefa de tele operação de um robô móvel na presença de atraso de tempo [71]. Os participantes telem operaram o robô em quatro diferentes combinações, dois graus de dificuldade e dois níveis de atraso adicionados pelo canal de comunicação, o sistema foi utilizado na configuração desktop. O teste foi realizado com 20 pessoas (14 mulheres e 6 homens) com a faixa etária dos participantes entre 24 e 40 anos. Cada participante executou testes para todas as combinações de dificuldades e níveis de atraso.

O sinal EEG coletado foi processado aplicando transformação wavelet, filtragem de ruídos e decomposição do sinal em diferentes frequências e resoluções. Foi analisada a energia relativa nas bandas de frequências Delta, Theta, Alpha, Beta e Gamma do sinal para detectar os estados de cargas cognitivas para as tarefas. Os índices obtidos foram validados usando as métricas do NASA-TLX, tempo para concluir a tarefa e o erro do rastreamento de caminho.

De acordo com os resultados encontrados foram constatados que o atraso de tempo e o nível de dificuldade aumentam a demanda de carga cognitiva. Esses resultados estão de acordo com a literatura. No artigo foi relatado como trabalho futuro a investigação da carga cognitiva utilizando o NASA-TLX e EEG durante a tele operação de cadeira de rodas.

2.4.5. Trabalho 5 – “Mental Workload of Young Drivers during Curve Negotiation”

O artigo apresenta um estudo sobre carga de trabalho mental de jovens motoristas durante a negociação de curvas [72]. Um experimento foi conduzido em um simulador para observar as alterações fisiológicas, frequência cardíaca, dos jovens condutores durante uma alta demanda de carga de trabalho mental. Durante a condução os participantes tiveram que fazer outras tarefas mentais (contagem regressiva e de sons). Participaram dos testes 20 motoristas, sendo que 9 deles com menos de 25 anos (faixa 21-24) e 11 com idade superior a 25 (faixa 29-33). Eles foram instruídos a dirigir a uma velocidade constante.

Neste estudo, os resultados apontaram que as pontuações da NASA-TLX geral foram maiores nível para os jovens condutores. Para todos os participantes, houve alterações significativas na frequência cardíaca ao dirigir sob alta carga de trabalho, em comparação com a tarefa única de dirigir. Os valores de carga cognitiva e frequência cardíaca são significativamente maiores durante a negociação de curvas e para os participantes mais jovens os maiores valores.

De modo geral, o estudo procurou inferir a carga cognitiva de forma subjetiva (NASA-TLX) e objetiva (frequência cardíaca) em participantes considerados jovens. Porém, faltou explicar um possível viés no estudo relacionado com o tempo de experiência em dirigir dos participantes.

2.4.6. Resumo Dos Trabalhos Avaliados

Este capítulo apresentou estudos relacionados a carga cognitiva e os tipos de medidas utilizados para medir a mesma. Foi abordado um conjunto de 5 trabalhos que mais se aproximam ao tema proposto na pesquisa atual dentro do cenário de tecnologia assistiva, carga cognitiva e realidade virtual.

Os trabalhos apresentam o enfoque de avaliar os simuladores em função de adequá-los para os usuários, e verificou-se que não existe, até onde a busca na literatura se estendeu, um trabalho que apresente apenas os parâmetros da piscada (sem uso dos questionários para classificação subjetiva) para inferir a carga cognitiva em simulador de cadeira de rodas em realidade virtual. Entretanto, sabe-se que a piscada tem sido utilizada em conjunto com outras métricas fisiológicas como GSR para inferir carga cognitiva, mas em outros tipos de aplicações.

Em suma, a análise dos conteúdos presentes nos trabalhos selecionados indica que é viável investigar sinais fisiológicos como métrica para inferir a carga cognitiva do usuário. A

vantagem de utilizar sinais fisiológicos para identificar o esforço mental é a possibilidade de monitorar o usuário durante o uso do simulador de forma contínua. Essa análise contínua pode ajudar a compreender melhor as condições cognitivas e na identificação da manifestação de cybersickness em tempo real.

3. MATERIAIS E MÉTODOS

Este capítulo descreve os componentes, ferramentas, tecnologias e métodos utilizados para o processamento dos dados coletados durante o desenvolvimento do estudo, sendo eles:

- I. Simulador de cadeira de rodas;
- II. Módulo de coleta dos dados da piscada;
 - a. Sistema de sincronização dos dados do simulador com o módulo EEG;
- III. Processamento dos dados;
 - a. Implícitos – métricas objetivas (sinais fisiológicos e métricas de desempenho);
 - b. Explícitos – métricas subjetivas (NASA-TLX, usabilidade e presença);
- IV. Protocolo Experimental;

A estrutura do trabalho está representada conforme a Figura 2. As seções a seguir descrevem os tópicos mencionadas acima. Onde, a seção 3.1 descreve os componentes do simulador. A seção 3.2 a presenta os dispositivos e ferramentas utilizados para a coleta das picadas. Na subseção 3.2.1 é explicado como foi feito a sincronização dos dados coletados no simulador com o sinal EEG. Na seção 3.3 é explicado como foi feito o processamento dos dados implícitos objetivos (subseção 3.3.1 referente ao sinal de EEG e subseção 3.3.2 sobre as métricas de desempenho) e explícitos subjetivos (subseção 3.3.3 referente ao questionário NASA-TLX e subseção 3.3.4 refere-se ao questionário sobre usabilidade e presença). A última parte do capítulo explica as etapas do protocolo de teste (seção 3.4).

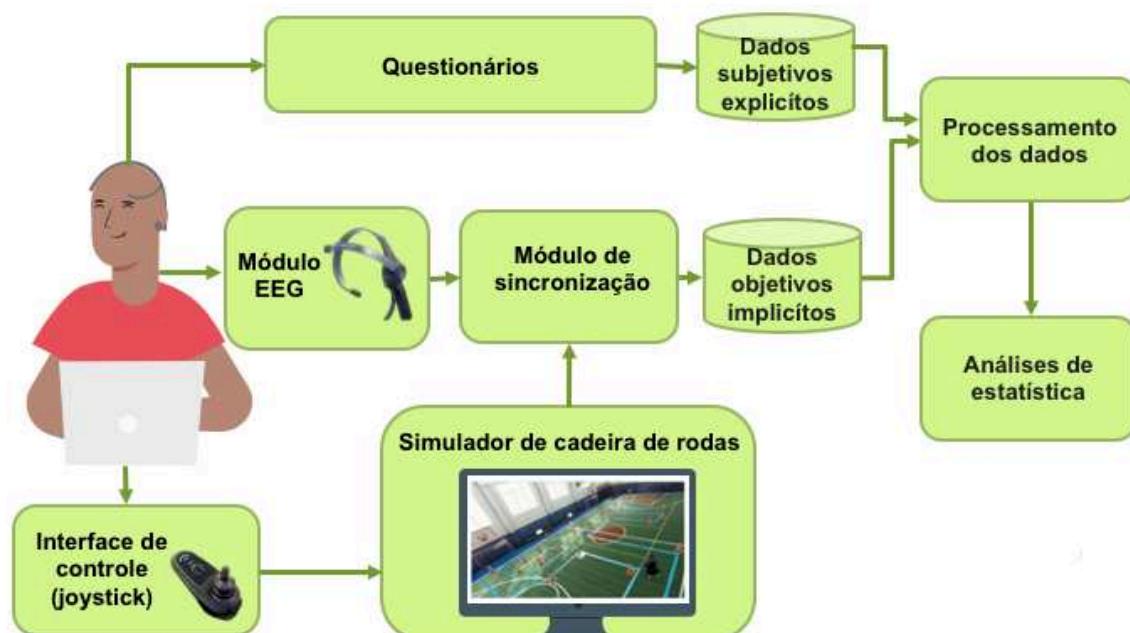


Figura 2 - Estrutura do trabalho

3.1. Simulador de Cadeira de Rodas

O EWATS (*Electric Wheelchair Assistive Technology System*) [18] foi o simulador de treinamento imersivo usado nesta pesquisa, conforme Figura 2. Ele foi desenvolvido por [18] utilizando a plataforma Unity (Versão 2017.2.0f3). O PC usado foi o Windows 10 Enterprise, CPU Intel CoreTM i7-8700 @ 3,20 GHz, 16 GB de RAM com uma placa gráfica NVIDIA GeForce GTX 1080 de 8 GB. O monitor utilizado tem o tamanho de 22 polegadas e a cena virtual é reproduzida com uma resolução de 1600x900. A simulação foi feita em primeira pessoa onde dentro do ambiente virtual tem a representação da cadeira de rodas e um avatar sentado na cadeira onde o usuário enxerga como se fosse o avatar. Desta forma, a simulação visa fornecer a usuários inexperientes de cadeira de rodas elétrica um método para aprender habilidades de operação em um ambiente seguro.

Na versão atual, três cenários diferentes estão disponíveis para os usuários (Figura 3): (a) pista de obstáculos; b) rampas de navegação; (c) manobras dentro de elevadores. Eles foram projetados para reproduzir simulações comumente encontradas por usuários de cadeira de rodas durante uma navegação normal durante o dia. O simulador da cadeira de rodas foi controlado com joystick da cadeira de rodas modelo VR2 com feedback haptico, que foi adaptado seguindo as instruções em [74]. O joystick adaptado para o computador envia 5 comandos referentes às condições de direção de forma discreta pela porta serial que, conectada ao computador, interpreta o comando e move a cadeira virtual. Os comandos utilizados na condução são: (i)mover para frente, (ii)mover para trás, (iii)girar para o lado direito, (iv)girar para o lado esquerdo e (v) parada, sendo que simulador só conta como estado de parada se o usuário demorar mais de 1 segundo para enviar um comando.

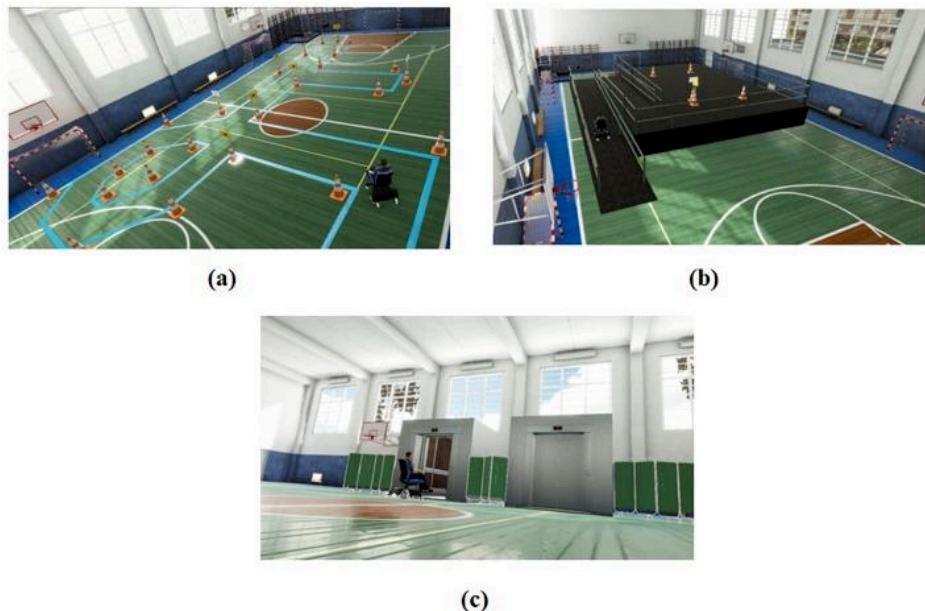


Figura 3 - Ambientes do Simulador: (a) pista de obstáculos; (b) percurso na rampa; (c) entrada e saída dos elevadores;

3.2. Módulo de Coleta de Dados da Piscada

O módulo de coleta de dados da piscada consiste no conjunto de dispositivos e ferramentas utilizados para capturar informações sobre as piscadas durante experimento. Os parâmetros da piscada foram derivados do sinal de eletroencefalografia (EEG). Para capturar os dados do EEG, foi utilizado o sensor Mindwave, um dispositivo de EEG de canal único, eletrodo seco, largura de banda de 3 a 100 Hertz (Hz), 12 bits de resolução, taxa de amostragem de 512 Hz e transmissão via Bluetooth. De acordo com o sistema internacional de colocação de eletrodo 10-20 [75], o principal eletrodo do dispositivo está localizado na região Fp1, os eletrodos de referência e terra estão localizados na região A1 e em T4, respectivamente (Figura 4).

A plataforma OpenVIBE foi utilizada para salvar e visualizar os sinais em tempo real [34]. Essa plataforma pode se comunicar com vários dispositivos de aquisição de sinal EEG. O motivo de escolher essa ferramenta foi devido ao suporte que é oferecido para utilizar o protocolo de sincronização de dados entre aplicativos chamado Lab Streaming Layer. O sistema de sincronização é explicado na seção 3.2.1.

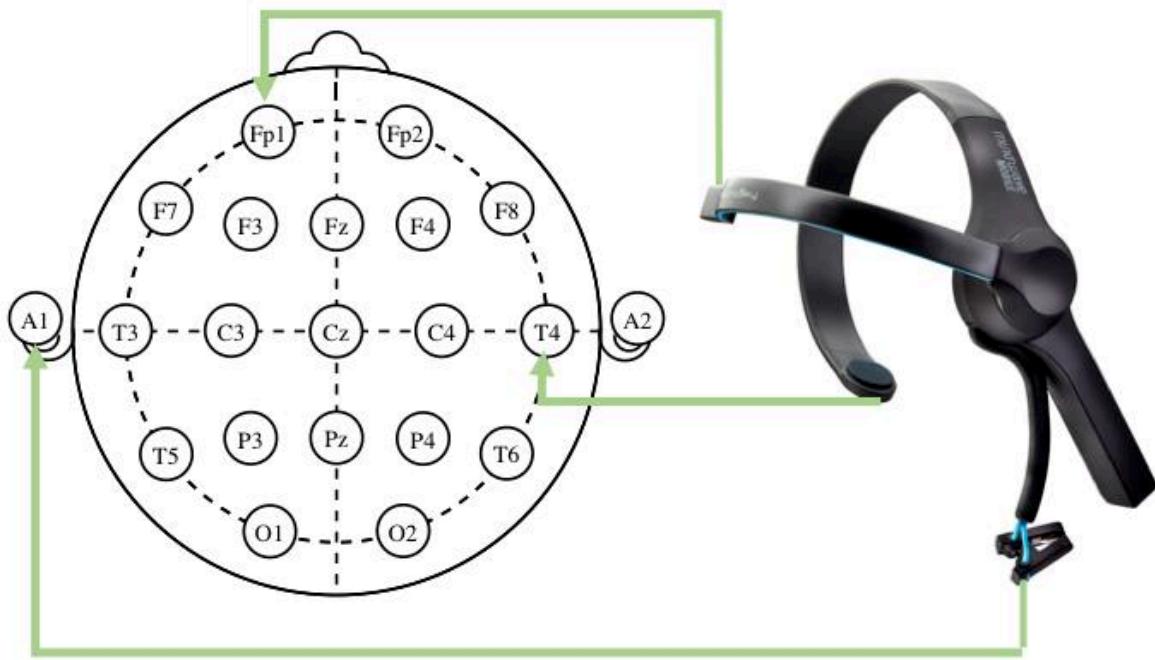


Figura 4 - Posição dos eletrodos utilizados na coleta do sinal de EEG com o dispositivo Mindwave. Os eletrodos estão representados de acordo com sistema internacional de colocação de eletrodos 10-20 [75]. Imagem adaptada de [76].

3.2.1. Sistema de Sincronização

Os dados do EEG e do simulador foram sincronizados utilizando o *Lab Streaming Layer* (LSL), que é um software de código aberto para aquisição de dados e sincronização e acesso em tempo real. O LSL pode enviar vários fluxos de dados de sensores para aplicativos, além disso, pode registrar e manipular os dados na plataforma MATLAB [77].

A Figura 5 mostra o sistema de sincronização utilizado neste estudo. Durante o experimento, são coletados sinais EEG do participante enquanto controla a cadeira de rodas virtual usando o joystick. As métricas de desempenho capturadas pelo simulador e os eventos do Joystick são gravados simultaneamente com os dados EEG pelo LSL em um mesmo arquivo.

Os dados salvos foram posteriormente processados e analisados utilizando o MATLAB. O dispositivo EEG foi usado com a plataforma OpenVibe, que permite se comunicar com o LSL. O simulador foi desenvolvido em Unity, no qual o LSL possui compatibilidade de comunicação. Foi desenvolvido um script que se comunica com LSL e envia marcadores dos eventos detectados pelo simulador. No geral, foram enviados os seguintes eventos para o LSL:

- Início do experimento;
- Final do experimento;
- Início do teste;

- Final do teste;
- Início da coleta do baseline;
- Final da coleta do baseline;
- Botão pressionado;
- Botão liberado;
- Colisão (erro do usuário).

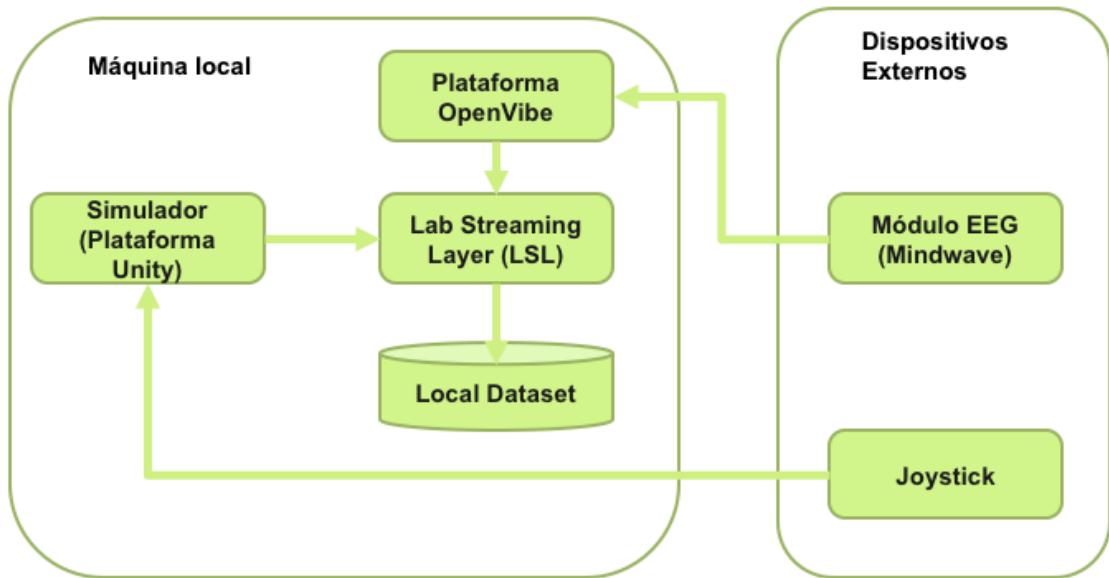


Figura 5 - Sistema de marcação de eventos e tempo para sincronização dos dados do simulador e do EEG.

3.3. Processamento dos Dados

Nesta seção é explicado como foi feito o processamento dos dados coletados durante os experimentos. Os dados são do tipo implícito objetivo e explícito subjetivo. As métricas implícitas objetivas são o sinal de EEG, utilizado para detecção de características da piscada, e medidas de desempenho (tempo da tarefa, número de comandos e colisões). Os dados explícitos subjetivos são provenientes dos questionários respondidos pelos participantes depois do experimento. O questionário NASA-TLX foi utilizado para inferir a carga cognitiva da tarefa (esforço mental). Além disso, os participantes responderam questões relacionadas a usabilidade e presença, essas questões foram baseadas no questionários SUS [78] e IPQ [73].

3.3.1. Processamento dos Sinais de EEG para Detecção das Piscadas

O sinal foi processado utilizando MATLAB onde foi utilizado o EEGLAB [79], uma ferramenta de processamento de sinais EEG. O sinal primeiramente foi filtrado para remover

artefatos de 60 Hz da rede elétrica. Em seguida ele foi particionado de acordo com os marcadores de eventos mencionadas anteriormente na seção 3.2.1.

Em seguida, foi utilizado uma biblioteca aberta existente dentro do EEGLab chamada Blinker [79], [80]. Blinker é um detector automatizado de piscadas em sinais de EEG. Entretanto, antes de utilizar a biblioteca, foi feita um pré-processamento (análise visual e corte do sinal). As etapas de processamento foram resumidas em 5 partes:

- I. Pré-processamento e detecção das regiões das possíveis piscadas no sinal EEG.
 - II. Identificação dos pontos de referência das piscadas (*landmarks*).
 - III. Eliminação das piscadas com muito ruído (relação sinal-ruído (SNR) baixo).
 - IV. Eliminação de *outliers* (piscadas falsas) e seleciona-se as “melhores” piscadas como as piscadas “usadas”.
- V. Calcula-se os índices oculares do sinal “usado”.

Etapa I, foi feita uma análise visual do sinal utilizando o EEGLAB, em seguida o sinal foi cortado (eliminando partes com ruídos). Em seguida foi aplicado ao sinal um filtro FIR de passa banda com janelamento de Hamming no intervalo 1- 20 Hz (função `pop_eeg-filtnew()` encontrada no EEGLAB). Depois foi determinado intervalos onde o valor sinal é maior que a média geral do sinal mais 1,5 do desvio padrão. Esses intervalos são regiões das possíveis piscadas. Consideram-se que as piscadas tenham um tempo maior que 50 ms e que estão pelo menos 50 ms separadas entre si. De acordo com [80], esses critérios eliminam pequenos movimentos rápidos dos olhos sem parecer eliminar os piscadas reais.

Etapa II, depois de identificar os intervalos das possíveis piscadas, foi aplicado um processo de ajuste “*fitting*” para encontrar pontos de referência “*landmarks*” para cada piscada. As informações dos *landmarks* foram salvas em um vetor. Figura 6 foi gerada pelo BLINKER onde mostra os *landmarks* de uma piscada em um sinal de EEG de um participante.

A Figura 6 mostra o sinal de EEG no domínio do tempo (escala em segundos) na porção que corresponde a piscada. As cruzes marcam a base esquerda (*leftBase*) e a base direita (*rightBase*), respectivamente. A linha horizontal verde marca a metade da duração de uma piscada a partir do ponto *leftZero* até o pico, chamada de *half zero duration*. As linhas pontilhadas representam os melhores ajustes lineares para os movimentos ascendente (*upStroke*) e descendente (*downStroke*) da piscada, respectivamente. A linha vertical é o comprimento da amplitude do pico ajustado “*Tent-peak*” em relação ao ponto de valor zero do sinal.

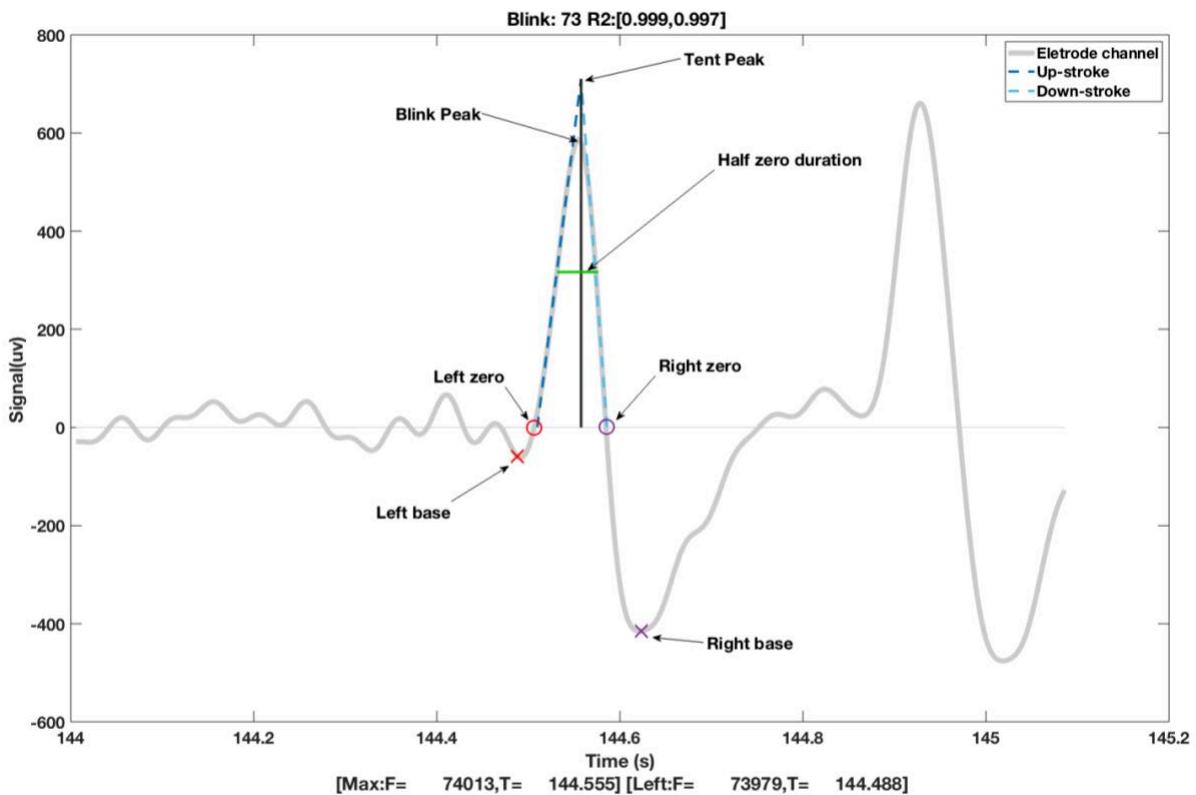


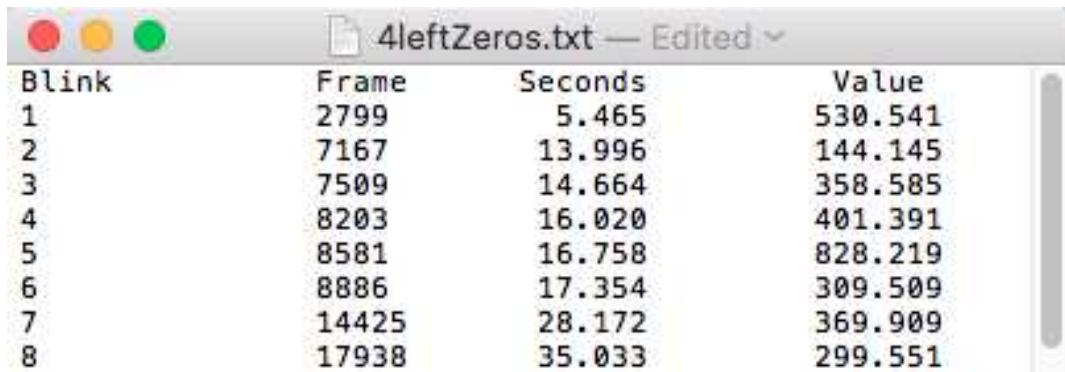
Figura 6 - Os pontos de referência da piscada “*landmarks*” da piscada. A figura tem o sinal de EEG no domínio do tempo na escala em segundos na porção que corresponde a piscada.

Caffier et al. 2003 observaram que o sinal da piscada tem o formato de tenda (*tent*) arredondada, e aplicou a estratégia de *tent-fitting* para caracterizá-la. Calcula-se os melhores ajustes lineares para 80% da parte interna do *upStroke* e *downStroke*, respectivamente, para cada possível piscada encontrada no sinal de EEG [81]. A qualidade da correlação (*r₂*) dessas linhas são representadas como *R₂* [*upStroke*, *downStroke*]. Os valores para *R₂* são mostrados no topo da Figura 6, as correlações são 0,999 e 0,997. Além disso, foi calculado a interseção dessas duas linhas ajustadas, denominada como o *tentPeak*, a amplitude do *tentPeak* é exibida como uma fina linha preta (Figura 6). O *tentPeak* está um pouco depois e acima do valor real do pico da piscada (*blinkPeak*). Os valores das correlações do *upStroke* e *downStroke* (*R₂*), bem como a posição relativa do *tentPeak* ao *blinkPeak* são parâmetros que ajudam a inferir o quão perto o sinal coletado da piscada se assemelha a um estereotípico do sinal de piscada ideal.

Ainda no gráfico da Figura 6, é possível ver que o valor máximo dessa piscada ocorre no frame (janela) de número 74013, denominado como “*maxFrame*” (tamanho máximo da janela), esse valor de amplitude máxima ocorre no tempo 144,555s. No caso, se dois ou mais pontos dentro deste frame (janela) atingirem o valor máximo, é considerado o primeiro valor como *maxFrame*. Depois de determinar o *maxFrame*, é calculado os demais *landmarks* da piscada da seguinte forma:

- **LeftZero** – É o valor do sinal quando cruza no ponto zero do eixo-y antes de atingir o *maxFrame* (Figura 6). Se o sinal não cruzar o zero entre essa piscada e a piscada anterior, *leftZero* será o frame da menor amplitude entre as piscadas.
- **RightZero** – Tem uma definição similar a do *leftZero*. É o valor do sinal quando cruza no ponto zero depois de atingir o *maxFrame*.
- **UpStroke** – É o intervalo entre *leftBase* e o *maxFrame* (parte ascendente da maior curva).
- **DownStroke** – É o intervalo entre *maxframe* e o *rightBase* (parte descendente da maior curva).
- **LeftBase** – É a região que se encontra o primeiro valor mínimo da piscada à esquerda do *maxFrame* no *upStroke*. Na Figura 6 o *leftBase* está no frame 73979 no tempo 144,488s.
- **RightBase** – É a região que se encontra o primeiro valor mínimo local depois do *maxFrame* no *downStroke*.
- **Left-outer** – É o intervalo do *rightZero* da piscada anterior ao *leftZero* da piscada atual (ou início do sinal da primeira piscada). as piscadas.
- **Right-outer** – É o intervalo do *rightZero* da piscada atual para o *leftZero* da próxima piscada (ou no final do sinal da última piscada).

Figura 7 mostra um exemplo das informações salva dos *leftZero* do sinal coletado do participante de ID 4. Essa estrutura contém informações sobre número da piscada, frame, tempo em segundos e o valor do sinal. As variáveis *rightZero*, *leftBase* e *rightBase* foram salvos neste mesmo tipo de estrutura.



Blink	Frame	Seconds	Value
1	2799	5.465	530.541
2	7167	13.996	144.145
3	7509	14.664	358.585
4	8203	16.020	401.391
5	8581	16.758	828.219
6	8886	17.354	309.509
7	14425	28.172	369.909
8	17938	35.033	299.551

Figura 7 - Exemplo de arquivo salvo das variáveis *leftZero*, *rightZero*, *leftBase* e *rightBase*.

Esses *landmarks* foram escolhidos devido aos estudos anteriores apontarem forte relação com os pestanejos das pálpebras ao abrir/fechar. A Figura 8 e a Tabela 4 mostram os resultados encontrados num estudo onde foi coletado o sinal EOG em conjunto com o rastreio do movimento ocular com câmeras. Percebe-se que a metade da parte ascendente e descendente da curva representa a região onde a pálpebra encosta no topo da pupila. A região ascendente está relacionada ao movimento de fechamento das pálpebras, a descendente está relacionada

ao movimento de abertura. O pico da curva representa o momento que as se fecham por completo.

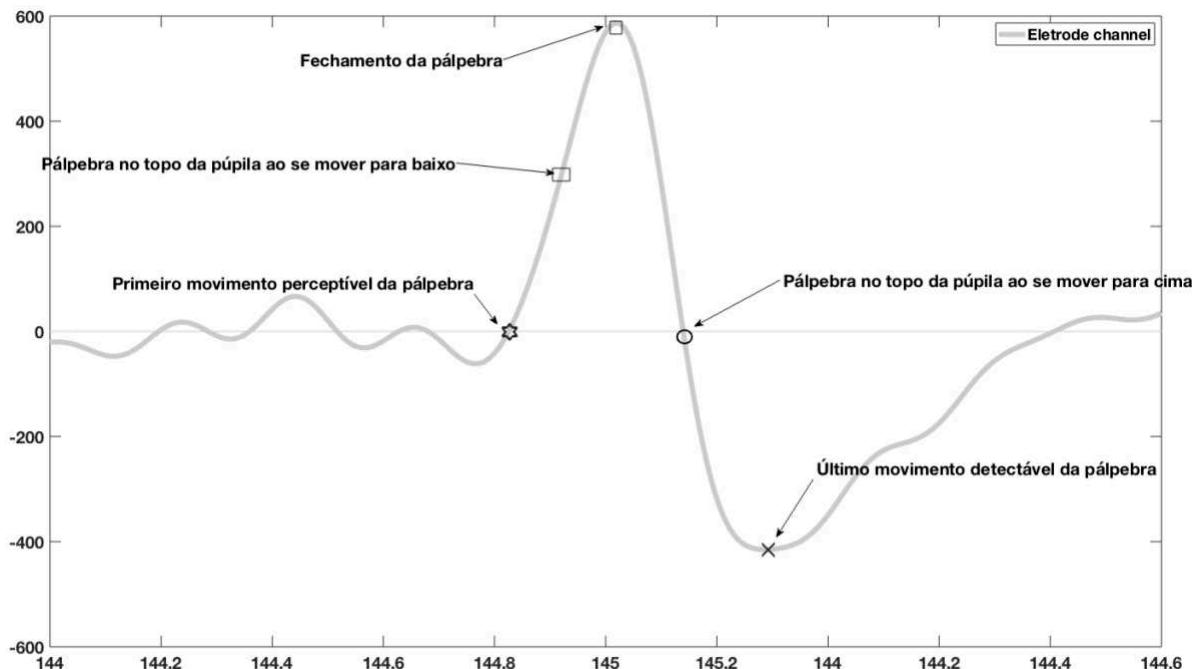


Figura 8 - Pontos do sinal da piscada relacionados ao movimento da pálpebra. Adaptado do [80].

Tabela 4 – Relação dos pontos de referência (landmarks) com o movimento das pálpebras.

Landmarks	Relação com movimento das pálpebras ao piscar
<i>LeftZero</i>	Primeiro momento perceptível das pálpebras ao se mover para baixo.
<i>LeftBase</i>	
<i>RightZero</i>	Último momento perceptível das pálpebras (pálpebras retraídas).
<i>RightBase</i>	
<i>Blink Peak</i>	Momento do fechamento completo das pálpebras
<i>UpStroke</i>	Movimento de fechamento das pálpebras
<i>DownStroke</i>	Movimento de abertura das pálpebras
<i>Metade da porção do upStroke e dopwnStroke</i>	Representa uma marcação próxima de quando a pálpebra encosta no topo da pupila.

Neste estudo, como foi mencionado anteriormente na seção 3.2., o canal Fp1 é utilizado como entrada para calcular o vetor de *landmarks* da piscada. Porém, existem outras técnicas no EEGLAB e no BLINLKER para determinar o melhor sinal candidato quando se é utilizado um equipamento com vários canais de EEG e EOG.

Etapa III, após determinar os *landmarks* para todas as possíveis piscadas encontradas no sinal de EEG, é então calculada a razão da amplitude média das piscadas para o resto do sinal EEG (R_{PS}). A qual é dada pelo quociente entre a amplitude média do sinal das piscadas entre os cruzamentos *leftZero* e *rightZero* (Ap) e amplitude média da fração positiva do resto do sinal que não constitui um sinal de piscada (A_s) de acordo com a equação 3.1.

$$R_{PS} = \frac{A_p}{A_s} \quad (3.1)$$

Foram rejeitados os sinais que possuem um R_{PS} fora de um intervalo padrão de 3 a 50. Estudos anteriores descobriram de forma empírica que sinais com valores de R_{PS} no intervalo 5 a 20 geralmente capturam a piscada de forma aceitável [80], [82]. Pode-se dizer que R_{PS} é uma medida da razão sinal-ruído (SNR) considerando a piscada o sinal alvo e o ruído o que não é a piscada (EEG).

Etapa IV, foi aplicado um outro tipo de filtragem para eliminar *outliers* considerados como os possíveis falsos positivos (piscadas falsas). Como foi mencionado anteriormente, a piscada tem uma morfologia do tipo de tenda arredondada e possui uma amplitude alta em relação ao sinal de EEG. Infelizmente, pequenos movimentos dos globos oculares e artefatos com picos altos de amplitude podem confundir a detecção da piscada.

Um estudo sobre amplitudes de piscadas foi usado para exclusão de possíveis *outliers* [80]. Sendo assim, seguindo essas observações [80], foi adotado um critério com três classificações para as piscadas: “melhores” ($R2 \geq 0,98$), “ótimas” ($R2 \geq 0,95$) e “boas” ($R2 \geq 0,90$). Além disso, as piscadas normais têm uma distribuição de amplitude máxima no formato de “sino” em torno da mediana das “melhores” piscadas. Esses critérios eliminam as piscadas cujo $R2$ é baixo ($R2 \leq 0,90$) e cujas amplitudes estão longe da mediana das piscadas classificadas como “melhores”. Por padrão, foram eliminadas as “melhores” piscadas com valor de desvio padrão maior que 5 da mediana das “melhores” piscadas e das piscadas “boas” foram eliminadas as piscadas com valores maiores de 2 da mediana das “melhores” piscadas.

Figura 9 mostra um exemplo das distribuições de amplitudes máximas para três casos típicos. A linha verde mostra a distribuição das amplitudes máximas de todas as possíveis piscadas e a linha azul mostra a distribuição das amplitudes máximas das piscadas consideradas “boas” ($R2 \geq 0,90$). A linha laranja mostra a distribuição das piscadas consideradas “ótimas” ($R2 \geq 0,95$), e a linha preta mostra a distribuição das amplitudes máximas das piscadas consideradas como as “melhores” ($R2 \geq 0,98$). A linha magenta mostra as piscadas usadas para análises ($R2 \geq 0,90$). Ainda na Figura 9 é exibido a mediana das “melhores” piscadas com uma linha vertical cinza e os desvio padrão da mediana com linhas cinza tracejadas. Para análise

estatística foram usadas as piscadas classificadas como “boas” ($R^2 \geq 0,90$) e as consideradas “melhores” ($R^2 \geq 0,98$) e as piscadas com $R^2 < 90$ foram descartadas.

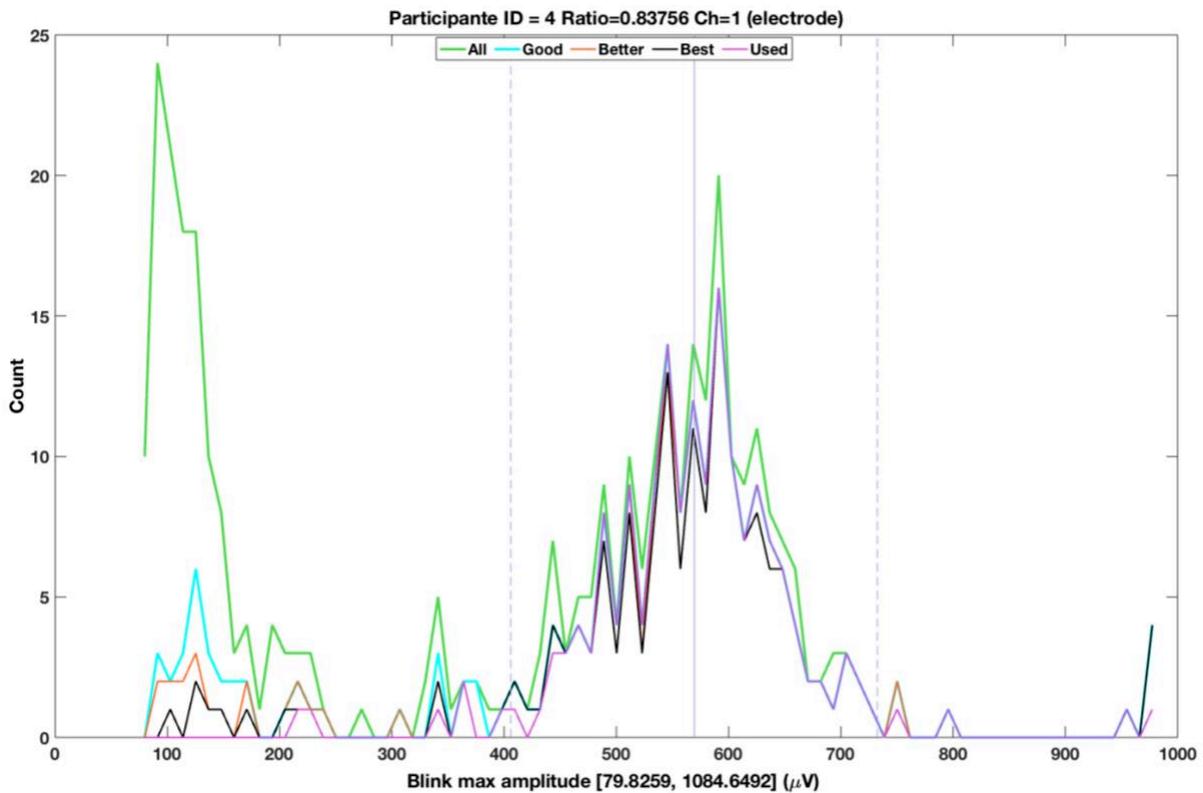


Figura 9 - Distribuições das amplitudes máximas das piscadas retiradas de um sinal de EEG.

Etapa V, foi calculado os índices oculares coletados de cada participante. Sendo eles, a frequência (taxa), duração e razão amplitude-velocidade das piscadas. A Frequência da piscada é o número de piscadas por minuto ou ciclos por minuto (cpm). Foi calculada a frequência média das piscadas como o número de piscadas encontrado no sinal dividido pelo comprimento total do sinal em minutos. Para o cálculo da duração das piscadas, foram usados os seguintes métodos:

- **Half-zero duration** – é a largura da piscada na metade da sua amplitude positiva (linha horizontal verde na Figura 6 e na Figura 10). Devido ao seu baixo desvio padrão, *half-zero duration* é menos sensível aos efeitos de piscadas complexas e movimentos oculares.
- **Half-base duration** – é a largura na metade da amplitude da piscada medida a partir do *leftBase*, representado como linha horizontal vermelha na Figura 10). A duração do *half-base* é um pouco mais longa que a duração do *half-zero*, embora essas medidas sejam altamente correlacionadas.
- **Base duration** – é a subtração entre *rightBase* e *leftBase* (*rightBase* – *leftBase*) representado como linha horizontal bege na Figura 10.
- **Zero duration** – é a subtração entre *rightZero* e *leftZero* (*rightZero* – *leftZero*) representado como linha horizontal roxa na Figura 10.

- **Tent duration** – é a diferença entre as interseções das linhas de ajuste linear *downstroke* e *upStroke* com a linha do eixo zero, representado como linha horizontal rosa na Figura 10.

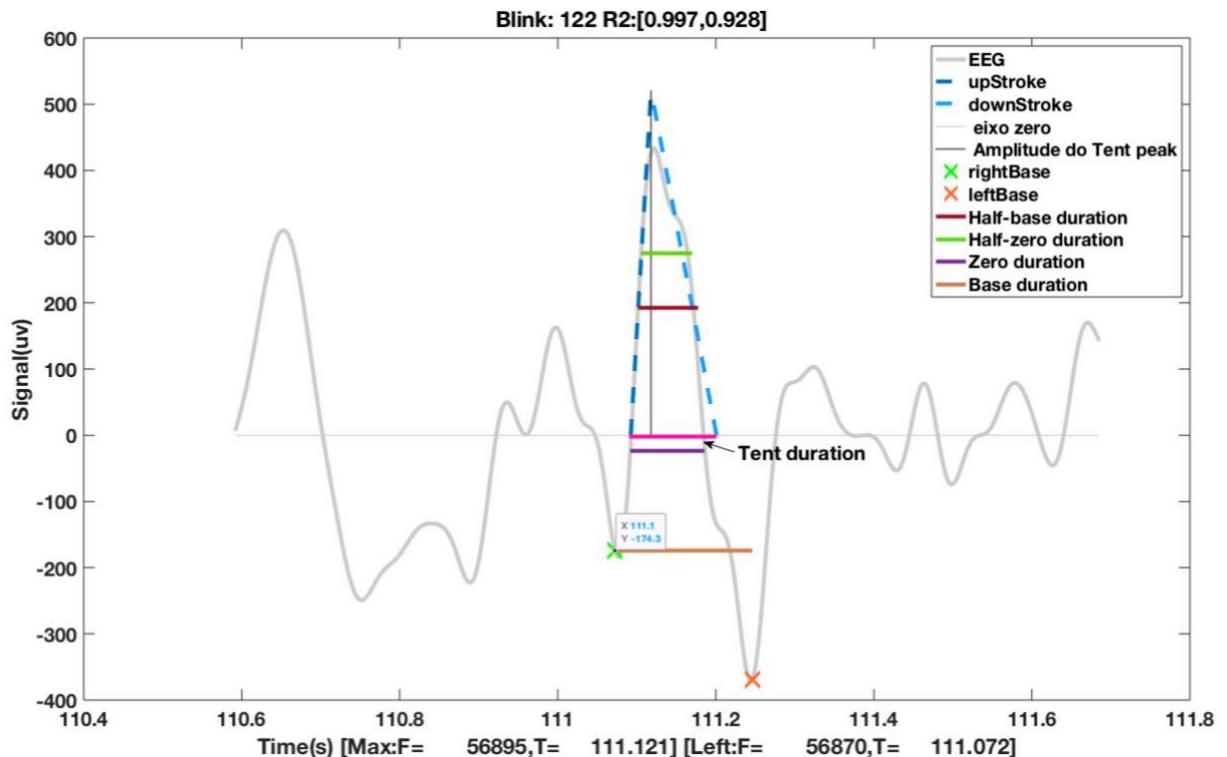


Figura 10 - Pontos de referência da piscada da piscada relacionados ao tempo de duração.

Estudos em [39], [82]–[84] propuseram duas medidas de velocidade como potenciais indicadores de fadiga e estado de alerta: a taxa de amplitude-velocidade positiva, “*positive amplitude velocity ratio*” (PAVR) e a taxa de amplitude-velocidade negativa, “*negative amplitude velocity ratio*” (NAVR). Essas medidas têm unidades de centisegundos. Nessa dissertação foram utilizadas essas métricas para identificar se os valores encontrados estão na faixa considerada de alerta, valores acima de 6 10msec está correlacionado com o estado de sonolência). Sabe-se que a velocidade (taxa de variação) do sinal pode ser calculada através da derivada do sinal em relação ao tempo de acordo com a equação 3.2.

$$v(t) = \frac{dx}{dt} \quad (3.2)$$

Na qual:

x – é o sinal da piscada no domínio do tempo em segundos.

PAVR é a razão da amplitude máxima da piscada sobre a velocidade máxima encontrada durante o período de *upStroke*, de acordo com equação a 3.3. NAVR é a razão da amplitude máxima da piscada sobre a velocidade máxima encontrada durante o período de *downStroke*, de acordo com a equação 3.4.

$$pAVR = 100 \left(\frac{\text{abs}\left(\frac{A_{\max_{upStroke}}}{V_{\max_{upStroke}}}\right)}{srate} \right) \quad (3.3)$$

$$nAVR = 100 \left(\frac{\text{abs}\left(\frac{A_{\max_{downStroke}}}{V_{\max_{downStroke}}}\right)}{srate} \right) \quad (3.4)$$

$A_{\max_{upStroke}}$ – é a amplitude máxima encontrada no período *upStroke*.

$V_{\max_{upStroke}}$ – é a velocidade máxima encontrada no período *upStroke*.

$A_{\max_{downStroke}}$ – é a amplitude máxima encontrada no período *downStroke*.

$V_{\max_{downStroke}}$ – é a velocidade máxima encontrada no período *downStroke*.

$srate$ – é a taxa de amostragem do sinal.

Outro critério utilizado para eliminar falsas piscadas foi o critério da razão amplitude-velocidade positiva das piscadas (PAVR). De acordo com a literatura [80] as possíveis piscadas com $PAVR \leq 3$ não correspondem as piscadas normais, mas correspondem aos movimentos oculares curtos e rápidos.

Além disso, foi adotado como método de exclusão do sinal para análise, se o sinal coletado não tiver um número suficiente de piscadas (pelo menos 20), pois não é possível obter uma boa estimativa da distribuição e os resultados não serão confiáveis [80]. Todas as características/índices das piscadas estão resumidamente descritas na Tabela 5.

Tabela 5 - Resumo das medidas da piscada utilizados para análise

Variáveis	Definição
blinksPerMin	Número de piscadas por minuto para o conjunto de dados (cpm).
srate	Taxa de amostragem do sinal (512 Hz)
maxFrame	O valor máximo no intervalo para ocorrer uma piscada. Ou seja, o número do frame da primeira amplitude máxima de uma piscada.
Tent peak	É interseção das linhas <i>upStroke</i> e <i>downStroke</i> ajustados.
leftZero	O cruzamento no lado esquerdo no eixo zero antes da amplitude positiva máxima piscada.
rightZero	O cruzamento no lado direito no eixo zero depois da amplitude positiva máxima piscada.
leftBase	O valor mínimo valor da piscada do lado esquerdo na parte ascendente da curva principal.
rightBase	O valor mínimo da piscada do lado direito na parte descendente da curva principal
durationB	<i>Base duration</i> , duração das piscadas em segundos, calculadas de acordo com a equação 3.5.
	$\frac{(rightBase - leftBase)}{srate} \quad (3.5)$
durationZ	<i>Zero duration</i> , duração da piscada calculadas de acordo com equação 3.6.
	$\frac{(rightZero - leftZero)}{srate} \quad (3.6)$
durationT	<i>Tent duration</i> , duração da piscada calculada pela largura da interseção das linhas de ajustes (<i>upStroke</i> e <i>downStroke</i>) com eixo zero, calculadas de acordo com a equação 3.7.
	$\frac{(IntersectZero_{downStroke} - IntersectZero_{upStroke})}{srate} \quad (3.7)$
durationHB	<i>Half-base duration</i> , duração das piscadas calculadas pela largura do piscada na metade da amplitude entre <i>leftBase</i> e o <i>maxFrame</i> .
durationHZ	<i>Half-zero duration</i> , durações das piscadas calculadas pela largura das piscadas na metade da amplitude máxima.
numberBlinks	Número total de piscadas encontrados no sinal.
numberGoodBlinks	Número de piscadas considerados “melhores” piscadas ($R^2 \geq 0,98$) encontrados no sinal.
NAVRZ	Razão amplitude-velocidade negativa, calculada a partir da divisão amplitude máxima ($A_{max_{downStroke}}$) pela velocidade máxima ($V_{max_{downStroke}}$) encontrada entre <i>maxFrame</i> e a <i>rightZero</i> (parte descendente da curva, <i>downStroke</i>), de acordo com equação a 3.4.
PAVRZ	Razão amplitude-velocidade positiva, calculada a partir da divisão amplitude máxima ($A_{max_{upStroke}}$) pela velocidade máxima encontrada ($V_{max_{upStroke}}$) no intervalo <i>leftZero</i> para o <i>maxFrame</i> (parte ascendente da curva, <i>upStroke</i>), de acordo com equação a 3.3.

3.3.2.Processamento das Métricas de Desempenho

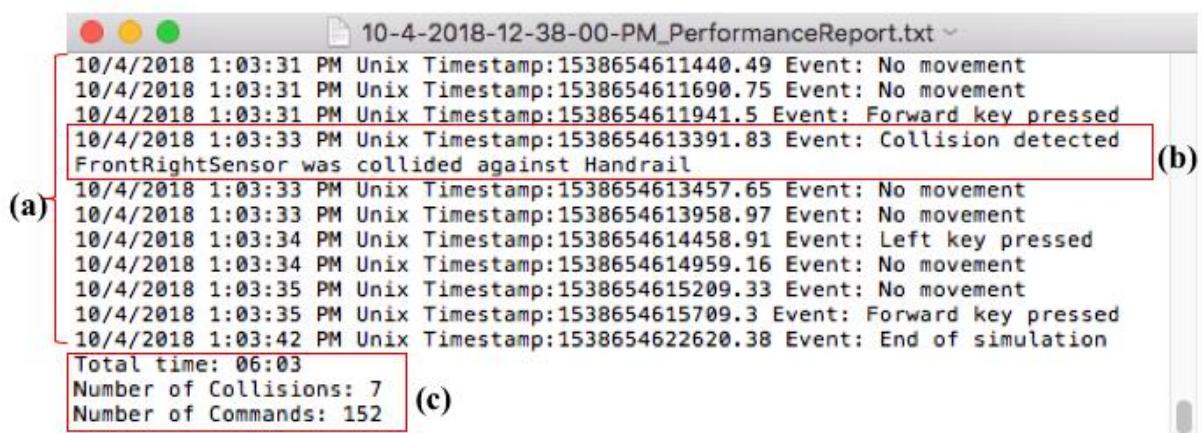
Conforme observado nos trabalhos correlatos [19], [69]–[71], os parâmetros principais para avaliação de desempenho na condução de uma cadeira de rodas na literatura são:

- Tempo gasto;
- Número de colisões realizadas
- Número de comandos feitos com o joystick;

Dessa forma, esses três parâmetros foram calculados durante o percurso proposto como tarefa para os participantes. Ao fim da tarefa, o relatório de desempenho é salvo em um arquivo no computador. A Figura 11 mostra um exemplo das informações de desempenho salvas pelo simulador sobre um participante. Esse arquivo contém informações sobre o tempo e o tipo de cada evento que ocorreu durante o teste. Os eventos salvos foram:

- Começo do percurso(tarefa);
- Final do percurso (tarefa);
- Eventos do Joystick;
- Direcionar a cadeira para frente;
- Cadeira de rodas virtual parada (sem movimento);
- Girar a cadeira para o lado direito;
- Girar a cadeira para o lado esquerdo;
- Cadeira andar para atrás (ré);
- Evento de colisão (local da colisão e qual parte da cadeira de rodas virtual foi atingida);

A última linha do arquivo contém o resumo do teste. Mostrando o tempo total da colisão, número de colisões e números de comandos. Para o cálculo do número de comandos, foi contabilizado apenas a mudança de estado de um movimento para o outro.



10-4-2018-12-38-00-PM_PerformanceReport.txt ~

10/4/2018 1:03:31 PM Unix Timestamp:1538654611440.49 Event: No movement
10/4/2018 1:03:31 PM Unix Timestamp:1538654611690.75 Event: No movement
10/4/2018 1:03:31 PM Unix Timestamp:1538654611941.5 Event: Forward key pressed
10/4/2018 1:03:33 PM Unix Timestamp:1538654613391.83 Event: Collision detected
FrontRightSensor was collided against Handrail
(b)

10/4/2018 1:03:33 PM Unix Timestamp:1538654613457.65 Event: No movement
10/4/2018 1:03:33 PM Unix Timestamp:1538654613958.97 Event: No movement
10/4/2018 1:03:34 PM Unix Timestamp:1538654614458.91 Event: Left key pressed
10/4/2018 1:03:34 PM Unix Timestamp:1538654614959.16 Event: No movement
10/4/2018 1:03:35 PM Unix Timestamp:1538654615209.33 Event: No movement
10/4/2018 1:03:35 PM Unix Timestamp:1538654615709.3 Event: Forward key pressed
10/4/2018 1:03:42 PM Unix Timestamp:1538654622620.38 Event: End of simulation
Total time: 06:03
Number of Collisions: 7
Number of Commands: 152
(c)

Figura 11 - Relatório de desempenho salvo pelo simulador. (a) Eventos salvos durante o percurso;(b) Evento de colisão; (c) Resumo do desempenho (tempo total, números de comandos e erros/collisões);

3.3.3.Processamento das Métricas Subjetivas de Esforço Cognitivo (NASA-TLX)

Para analisar a quantidade de carga cognitiva exigida para controlar o simulador foi aplicado o questionário NASA “*Total Load Task*” (NASA-TLX) que mede subjetivamente essa carga [15], [85]. O questionário tem uma classificação multidimensional que deriva uma pontuação geral da carga de trabalho com base em uma média de seis fatores (sub escalas), sendo eles:

- I. Demanda mental (Quanto foi exigido de atividade mental e perceptiva para concluir a tarefa?);
- II. Demanda física (Quanto foi exigido fisicamente para concluir a tarefa?);
- III. Demanda temporal (Quanto você se sentiu apressado para concluir a tarefa / Quanto de pressão de tempo você sentiu para concluir esta tarefa?);
- IV. Desempenho (Quão bem-sucedido você foi? / Você acha que alcançou os objetivos da tarefa?);
- V. Esforço (O quanto você se esforçou/trabalhou para atingir seu nível de desempenho?);
- VI. Frustração (Quão inseguro, desanimado, irritado e estressado você esteve durante a tarefa?);

O NASA-TLX consiste em dois tipos de escores (carga total bruta e ajustada). Para o cálculo, os participantes primeiramente responderam 15 alternativas que tem o objetivo de determinar as prioridades (peso) de cada fator, as alternativas estão representadas na Figura 12. Em seguida, os fatores recebem classificações de intensidade. Figura 13 mostra as escalas utilizadas para determinar a intensidade percebida de cada fator.

For each pair, choose the factor that was more important to your experience of the workload in the task that you recently performed:

1	<input type="checkbox"/> Temporal Demand	<input type="checkbox"/> Mental Demand
2	<input type="checkbox"/> Performance	<input type="checkbox"/> Mental Demand
3	<input type="checkbox"/> Mental Demand	<input type="checkbox"/> Effort
4	<input type="checkbox"/> Temporal Demand	<input type="checkbox"/> Effort
5	<input type="checkbox"/> Physical Demand	<input type="checkbox"/> Performance
6	<input type="checkbox"/> Performance	<input type="checkbox"/> Temporal Demand
7	<input type="checkbox"/> Effort	<input type="checkbox"/> Physical Demand
8	<input type="checkbox"/> Mental Demand	<input type="checkbox"/> Physical Demand
9	<input type="checkbox"/> Performance	<input type="checkbox"/> Frustration
10	<input type="checkbox"/> Effort	<input type="checkbox"/> Performance
11	<input type="checkbox"/> Frustration	<input type="checkbox"/> Effort
12	<input type="checkbox"/> Frustration	<input type="checkbox"/> Mental Demand
13	<input type="checkbox"/> Physical Demand	<input type="checkbox"/> Temporal Demand
14	<input type="checkbox"/> Physical Demand	<input type="checkbox"/> Frustration
15	<input type="checkbox"/> Temporal Demand	<input type="checkbox"/> Frustration

Figura 12 – Alternativas do NASA-TLX para determinar o peso dos seis fatores utilizados para o cálculo da carga cognitiva da tarefa.

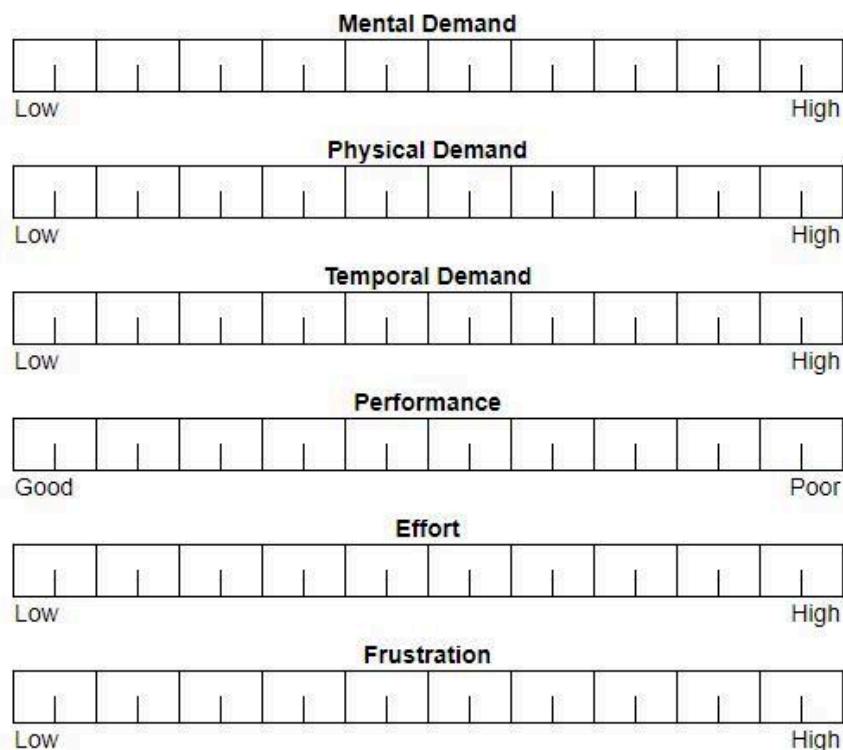


Figura 13- Escalas do NASA-TLX [13]

As médias ponderadas dos valores nessas seis dimensões são calculadas para fornecer uma medida geral da carga de trabalho do tipo bruta (sem levar em consideração o peso dos fatores). Sendo assim, a carga total bruta é calculada de acordo com a equação 3.8.

$$TLX_{Bruto} = \left(\frac{D_{Mental} + D_{Física} + D_{Temporal} + D_{Desempenho} + D_{Esforço} + D_{Frustação}}{6} \right) \quad (3.8)$$

TLX_{Bruto} – é a medida geral da carga bruta

D – é a representação dos escores de cada sub-escala

A medida da carga total ajustada (considerando o peso dos fatores) é calculada de acordo com a equação 3.9. A Figura 14 mostra um exemplo de resultado final do NASA-TLX. A aplicação do NASA-TLX nesse estudo foram realizadas na versão em papel e seguiram as diretrizes descritas em [15], [85], [86].

$$TLX_{ajustado} = \left(\frac{((D*P)_{Mental} + (D*P)_{Física} + (D*P)_{Temporal} + (D*P)_{Desempenho} + (D*P)_{Esforço} + (D*P)_{Frustação})}{15} \right) \quad (3.9)$$

$TLX_{ajustada}$ – é a medida geral da carga ajustada

D – é a representação dos escores de cada sub-escala

P – é a representação do peso de cada sub-escala

	(a)	(b)	(c)
	Medida Bruta	Peso	Medida Ajustada
Demandas Mental	15	2	30
Demandas Físicas	5	2	10
Demandas Temporais	15	5	75
Desempenho	15	4	60
Esforço	15	2	30
Frustração	5	0	0
(d) Medida Geral Ajustada: 13.67 (e) Medida Geral Bruta: 11.66			

Figura 14 - Medidas do NASA-TLX: (a) medidas brutas; (b) peso dos fatores; (c) medida ajustada; (d) medida geral ajustada; (e) medida geral bruta;

De acordo com a literatura, o NASA-TLX é aplicado em vários sistemas de interface homem-máquina. Portanto, pode mostrar correlações com várias variáveis, como desempenho em simuladores [15]. Por exemplo, alta carga cognitiva pode ter efeito negativo na conclusão de uma tarefa.

3.3.4. Processamento das Métricas Subjetivas de Usabilidade e de Presença

Além do NASA-TLX, os participantes responderam perguntas sobre presença e a usabilidade do sistema. A Tabela 6 apresenta as 10 questões aplicadas no estudo. Essas questões foram baseadas no *Igroup Presence Questionnaire* (IPQ) e no *System Usability Scale* (SUS) [73], [78]. Foi solicitado aos participantes que classificassem cada pergunta usando o sistema de classificação absoluta de categoria “*absolute category rating*” (ACR) descrito na norma ITU-TP.913 [87]. O sistema de classificação usou uma escala de *Likert* de cinco pontos para determinar se o usuário concordou ou discordou das declarações.

As questões de usabilidade tiveram o objetivo de coletar opiniões dos participantes sobre o simulador. Por exemplo, se foi fácil usar o simulador, se eles acreditam que outras pessoas podem aprender a usar uma cadeira de rodas usando o simulador, sobre a consistência do simulador e se o simulador pode ser usado sem nenhum conhecimento prévio. As questões de presença avaliaram quatro aspectos, sendo eles: I. Presença geral “sensação de estar no ambiente virtual (AV)”, II. Presença espacial “a sensação de estar fisicamente presente na AV”, III. Envolvimento “medindo a atenção dedicada à experiência no AV” e IV. Realismo “medindo a experiência subjetiva do realismo no AV”.

Tabela 6 - Perguntas presentes no questionário sobre usabilidade e presença.

Pergunta	Referência
I. Achei o sistema desnecessariamente complexo.	SUS [66]
II. Eu pensei que o sistema era fácil de usar.	
III. Eu imaginaria que a maioria das pessoas aprenderia a usar o sistema muito rapidamente.	
IV. Achei o sistema muito complicado de usar.	
V. Preciso aprender muitas coisas antes de poder continuar com esse sistema.	
VI. No mundo gerado por computador, eu tinha a sensação de "estar lá".	IPQ [74]
VII. Tive a sensação de atuar no espaço virtual, em vez de operar algo livre lá fora.	
VIII. Eu me senti presente no espaço virtual.	
IX. Quão consciente você estava do ambiente do mundo real enquanto navegava no mundo virtual? (ou seja, sons, temperatura ambiente, outras pessoas, etc.)	
X. Quanto sua experiência no ambiente virtual pareceu consistente com sua experiência no mundo real.	

3.4. Protocolo Experimental

O projeto teve a aprovação do comitê de ética do Athlone Institute of Technology para avaliar o sistema com participantes saudáveis. Foram recrutados (amostragem por conveniência) 29 participantes, sendo que os dados coletados de 4 participantes foram considerados inelegíveis para análise devido a qualidade do sinal da piscada não foi considerado satisfatório de acordo com protocolo de processamento do sinal discutido na seção 3.3.1, a relação sinal-ruído (SNR) não foi satisfatório ($R^2 < 0.90$) e número de piscadas encontrados no sinal foi menor que 20). No total foram analisados os dados de 25 participantes (14 mulheres e 11 homens) com idade $26,50 \pm 5,7$ anos (média \pm DP).

O protocolo experimental do estudo seguiu a norma ITU-TP.913 [87]. O experimento com cada participante durou por volta de 35 a 45 minutos, de acordo com a norma ITU-TP.913 [87] o experimento não deve durar acima de 60 minutos. Deste modo, o experimento foi dividido em seis fases conforme mostrado na Figura 15, são eles:

- I. Explicação sobre o experimento, foram explicadas as informações sobre o experimento. Após a explicação, todos os participantes assinaram um termo de consentimento.
- II. Triagem, foram perguntados aos participantes se eles sofriam de epilepsia, falta de sono (se dormiram pelo menos de seis horas nas últimas 24 horas), suspeita de gravidez e se consumiram álcool nas últimas 24 horas. Nesta mesma etapa, os participantes foram avaliados em relação a acuidade visual e percepção de cores. Para acuidade visual, foi realizado o teste de Snellen [88]. Para avaliação de percepção das cores foi aplicado o teste de Ishihara [89].
- III. Coleta de baseline, foi coletado os valores de EEG dos participantes por 5 minutos, neste período eles ficaram isolados na sala de teste sem fazer nenhuma atividade.
- IV. Treinamento, foi apresentado o simulador e como se controla a cadeira virtual utilizando o joystick. O treinamento foi feito em um espaço virtual livre de obstáculos. Também foi instruído que os participantes fizessem alguns movimentos sob a orientação do investigador principal (responsável de supervisionar os experimentos).
- V. Teste, primeiramente, foi explicado a tarefa necessária a ser realizada durante o teste. Como parte da avaliação, o participante teve que concluir um percurso predefinido em um dos três cenários usando o joystick. A tarefa escolhida foi a navegação em rampa (Figura 16). Durante essa etapa, os participantes interagiram com o simulador, os dados do EEG assim como os parâmetros de desempenho foram gravados ao longo da tarefa.
- VI. Aplicação dos questionários, após a conclusão dos testes, os participantes responderam o NASA-TLX, e as perguntas de usabilidade (SUS) e presença (IPQ).

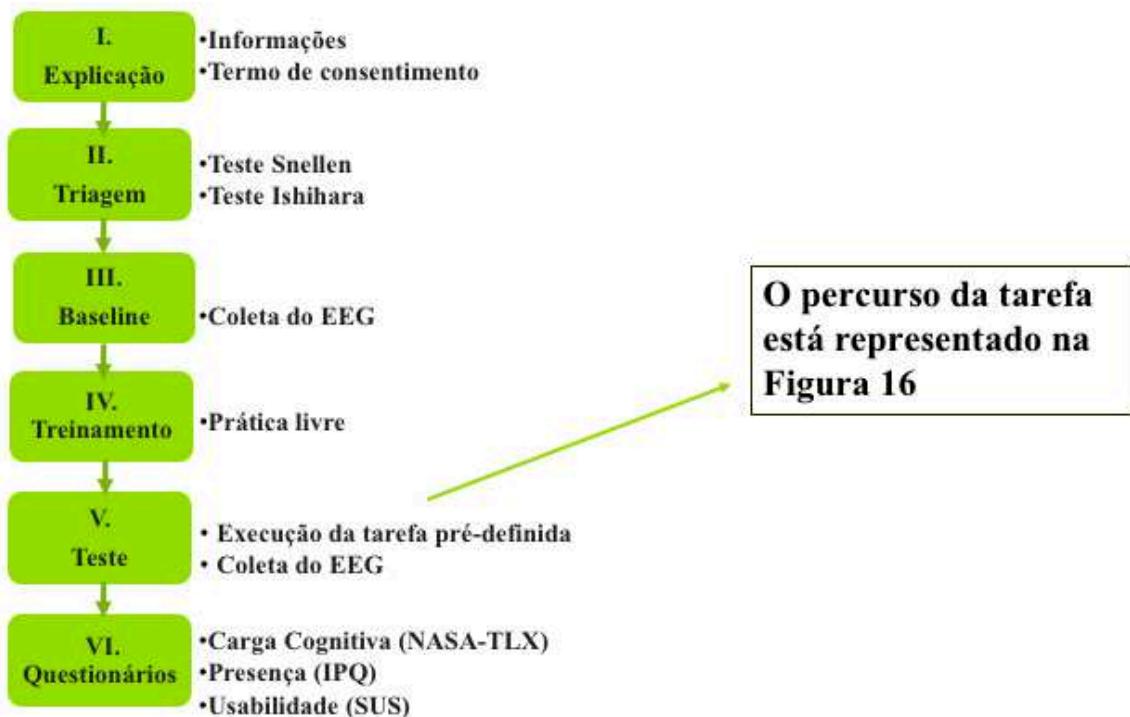


Figura 15 - Etapas do experimento

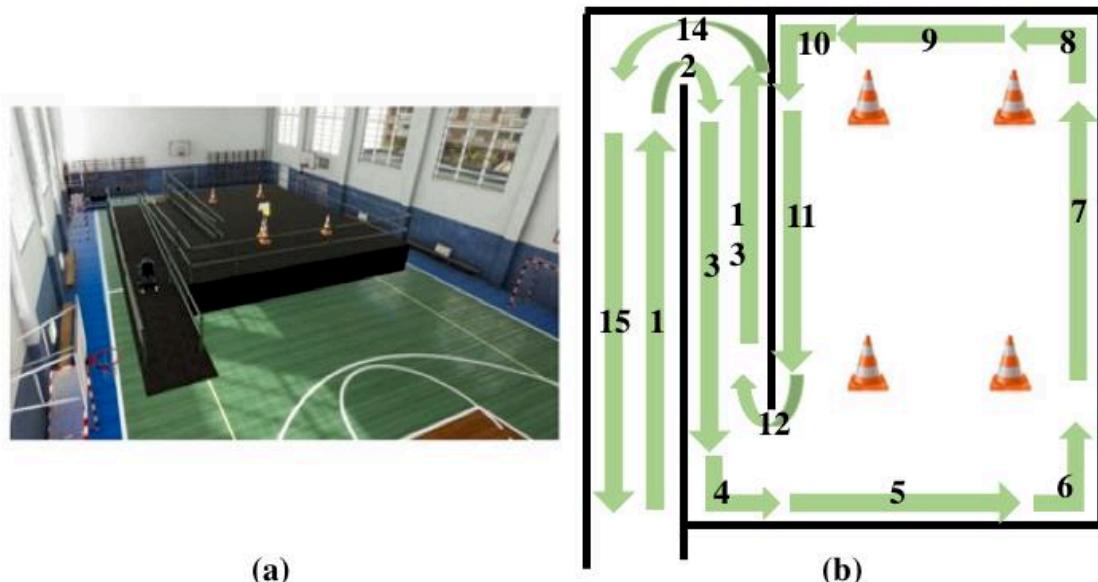


Figura 16 - Percurso da tarefa de teste (ambiente da rampa): (a) percurso da execução da tarefa (representação 3D); (b) protocolo de direção para conclusão do percurso da rampa.

3.5. Análise Estatística

Primeiramente, foi realizado o teste de normalidade Shapiro-Wilk, o qual indicou que os dados referentes ao sinal da piscada vem de uma distribuição normal, entretanto os valores de desempenho e dos questionários não foram considerados normais. Devido a esses resultados, foi realizado uma abordagem de análise paramétrica (teste t para comparação de medidas repetidas) para os dados referentes a piscada e uma análise não paramétrica para os demais (teste de correlação Spearman). Todas as análises foram feitas utilizando as plataformas IBM SPSS e o Matlab.

4. RESULTADOS E DISCUSSÃO

Nesta seção, são apresentados os resultados referentes aos dados coletados durante os experimentos, incluindo uma análise estatística dos resultados.

4.1. Variáveis de Desempenho

Todos os participantes concluíram a tarefa. A Figura 17 mostra os valores médios do número de comandos (81,32), número de colisões (6,64) e tempo gasto para completar a tarefa (192,76s). Figura 18 apresenta a distribuição dos resultados de desempenho, a mediana do número de comandos é 75,5, erros (colisões) 6 e do tempo da execução da tarefa é 179 segundos. A moda para os números de comandos, erros e tempo são: 66, 6 e 166s, respectivamente.

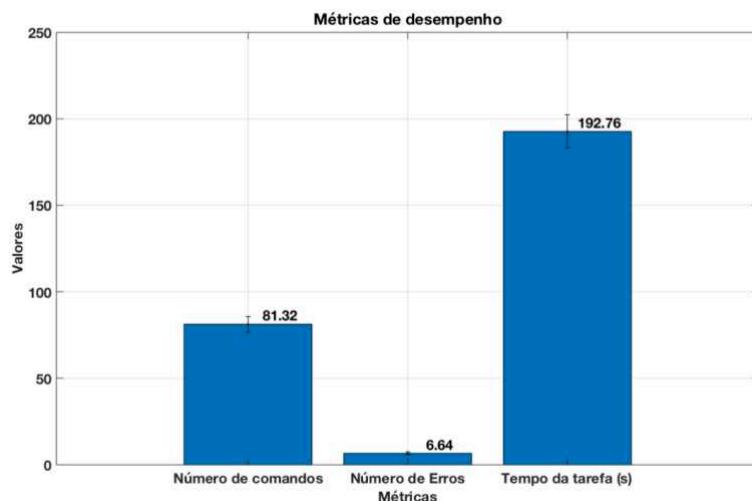


Figura 17 - Valores médios das métricas de desempenho.

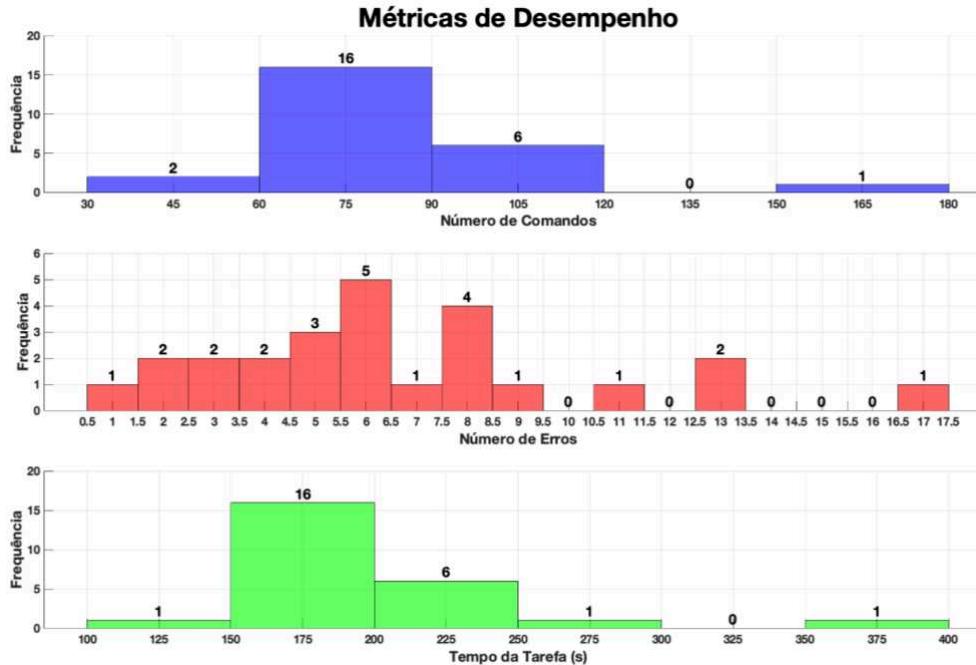


Figura 18 - Métricas de desempenho de todos os participantes. A linha pontilhada representa os valores médios.

4.2. Métricas de Usabilidade e Presença

A Tabela 7 mostra os valores médios dos resultados coletados sobre usabilidade e presença. Foi aplicado uma escala Likert de 5 (discordo completamente a concordo completamente). O valor máximo que pode atingir é de 5 para cada pergunta e o mínimo é um. Ao marcar no meio da escala, referente a 3, representa um resultado neutro.

Tabela 7 - Resultados das perguntas sobre usabilidade e presença

	Média	Desvio Padrão	Erro Padrão
Questão 1 (usabilidade)	4,38	0,62	0,12
Questão 2 (usabilidade)	3,89	0,63	0,12
Questão 3 (usabilidade)	4,04	0,79	0,15
Questão 4 (usabilidade)	4,32	0,61	0,12
Questão 5 (usabilidade)	4,32	0,61	0,12
Score final (usabilidade)	4,19	0,39	0,07
Questão 6 (immersion)	3,14	1,11	0,21
Questão 7 (presença)	3,50	0,88	0,17
Questão 8 (presença)	3,43	1,10	0,21
Questão 9 (presença)	2,75	1,08	0,20
Questão 10 (presença)	3,18	0,77	0,15
Score final (presença)	3,20	0,63	0,12

A Figura 19 mostra o histograma dos valores médios encontrados para usabilidade e presença. O primeiro gráfico em azul mostra os valores de usabilidade. A mediana para

usabilidade foi 4,15 e a moda 4, isso representam que a maioria dos participantes deram notas de boa para ótima ao simulador em termos usabilidade. Valores próximos de 3 correspondem aos participantes que têm uma opinião neutra a respeito do sistema.

No segundo histograma (cor vermelha) é mostrado a distribuição dos resultados de presença que estão relacionados ao nível de “imersão”, o quanto os usuários sentiram-se presentes no ambiente virtual ao ponto de esquecer o que está ocorrendo em sua volta no mundo real. Ambos os valores da mediana e moda é 3,20. Valores próximos de 3, similar a classificação da escala de usabilidade, estão relacionados a opinião neutra. Em outras palavras, isso representa pouco nível de presença, o que de certa maneira era esperado considerando que o monitor não é um dispositivo de RV imersivo. Apenas 3 participantes (1, 14 e 20) relataram um nível de presença alto (acima de 4,8).

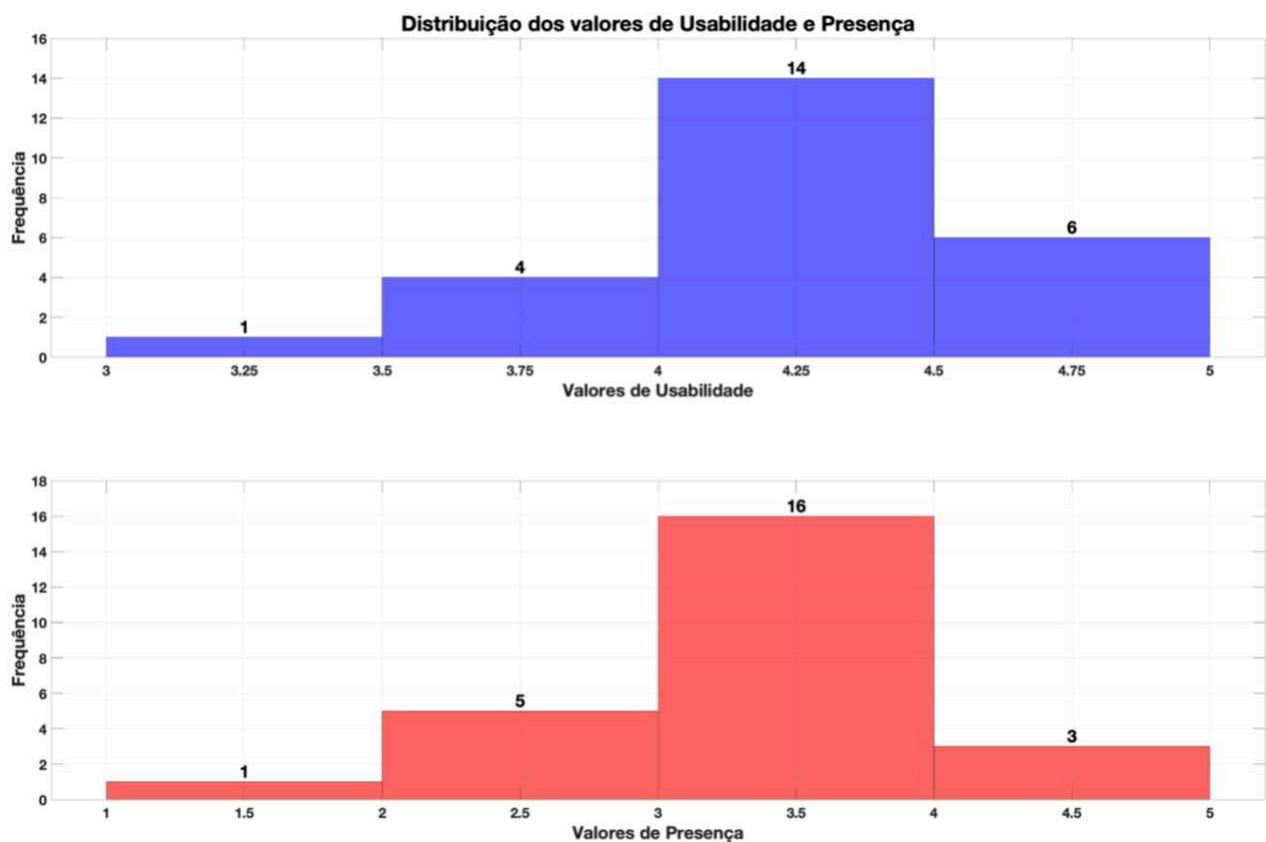


Figura 19 - Valores médios de usabilidade e presença de todos os participantes.

4.3. NASA-TLX

A carga de trabalho foi avaliada usando o NASA-TLX e as pontuações médias estão exibidas na Figura 20. Os resultados mostram que ambos os valores finais brutos e ajustados apresentam médias de $32,6 \pm 2,81$ e $40,56 \pm 3,71$ respectivamente. A demanda mental teve o maior valor ajustado e bruto, isso demonstra que os participantes consideraram que a tarefa exige mais carga mental do que os demais fatores. De forma geral, pode-se dizer que a percepção dos participantes quanto aos fatores avaliados pelo NASA-TLX numa escala decrescente foi: carga mental, esforço, desempenho, temporal, físico e frustração.

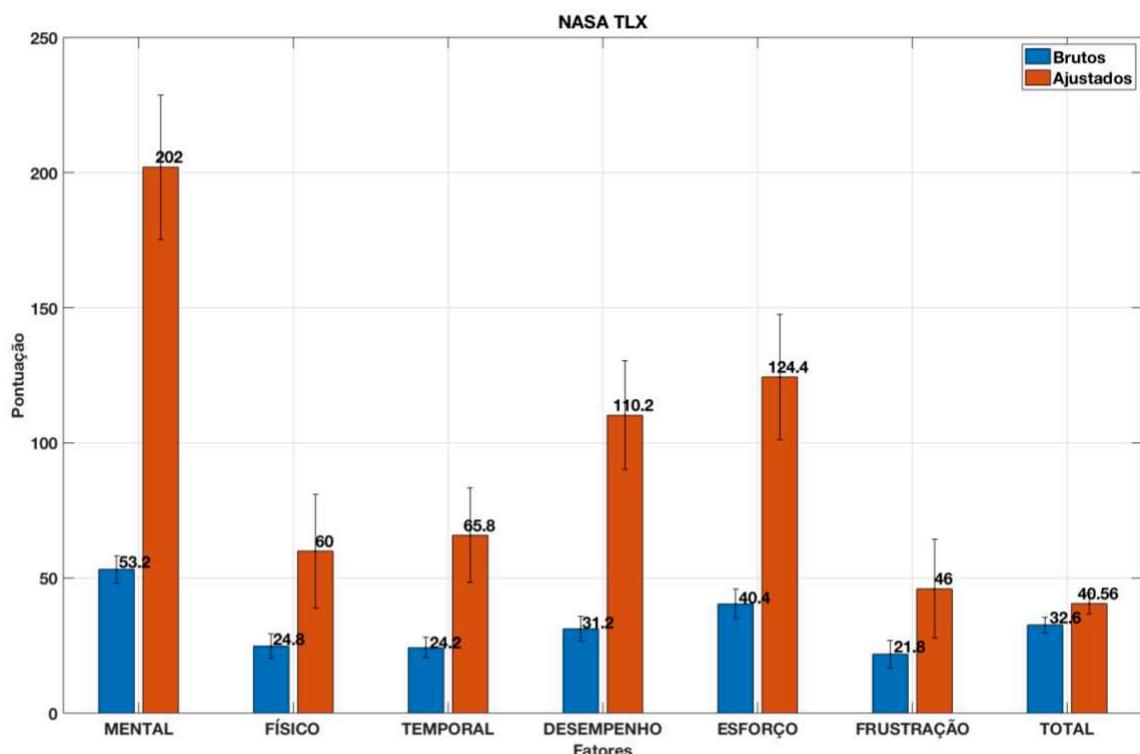


Figura 20 - Médias dos escores do NASA-TLX. Valores brutos em azul e os ajustados (corrigidos) em laranja.

A Figura 21 mostra a distribuição dos valores brutos da carga cognitiva em azul e em vermelho os ajustados. Pode-se observar que a maioria dos (12) participantes tiveram uma carga de trabalho cognitivo bruta na faixa entre 20-40 e para os valores ajustados a maior incidência ficou na faixa entre 40-60.

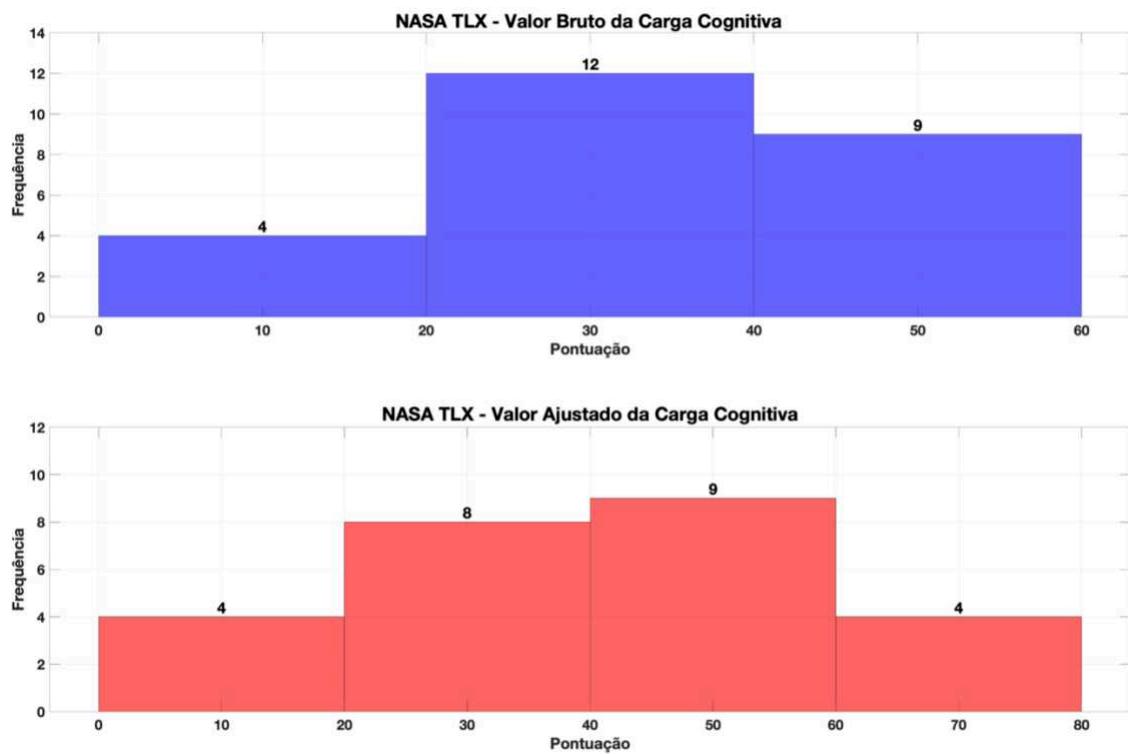


Figura 21 – Distribuição da carga cognitiva bruta (azul) e ajustada/corrigida (vermelho).

4.4. Métricas da Piscada

4.4.1. Frequência da Piscada

Na **Error! Reference source not found.** e Tabela 8 pode-se observar que as frequências médias das piscadas por minuto foram $18,37 \pm 1,42$ (baseline) e $13,13 \pm 1,16$ (teste), as quais estão na faixa dos valores de frequência de piscada espontânea relatados na literatura [25], [81], [84], [90], que estão entre 10 a 20 piscadas por minuto para adultos entre 20 a 30 anos. A qualidade média dos sinais coletados (SNR) apresentou 0,60. Esse valor representa uma proporção menor de piscadas boas e possíveis piscadas ao comparar com valores SNR de análises de um canal EOG. Isso é esperado, pois canais EEG são posicionados a uma distância maior dos olhos e devem ter ocorrência de outros fenômenos além da piscada. A **Error! Reference source not found.** mostra os valores da frequência média de piscadas boas durante o baseline e o teste.

Foi aplicado um teste t de medidas repetidas para verificar se houve diferença estatística significativa entre as frequências durante o baseline e o teste. A frequência da piscada apresentou uma diferença significativa, onde o valor médio durante o teste foi menor do que apresentado durante o baseline (Tabela 8).

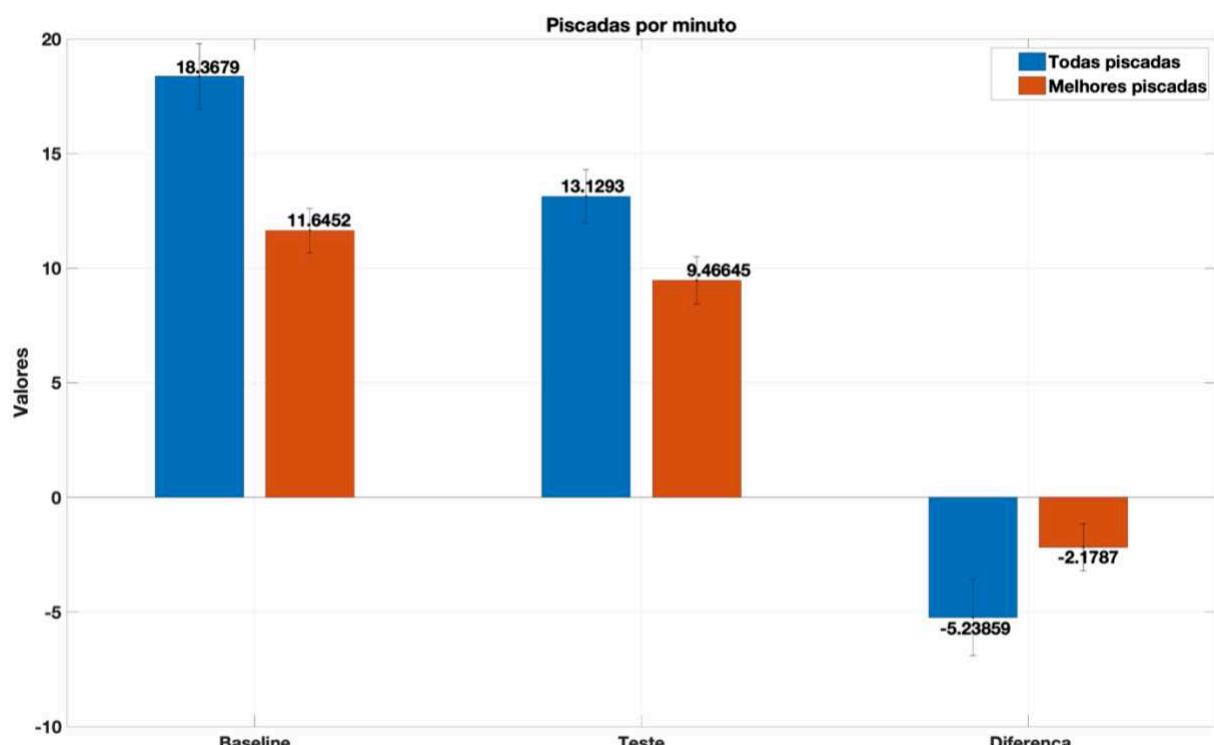


Figura 22 - Frequência média das piscadas.

Tabela 8 - Estatística descritiva da frequência das piscadas durante o baseline e o teste e da diferença entre eles (Teste – Baseline)

	Faixa	Min.	Máx.	Soma	Média	Erro Padrão	Desvio Padrão	Variância
Baseline	30,63	6,85	37,48	459,20	18,37	1,42	7,11	50,56
Teste	20,70	3,76	24,45	328,23	13,13	1,16	5,82	33,84
Baseline (melhores*)	20,52	3,85	24,38	291,13	11,65	0,96	4,82	23,23
Teste (melhores*)	19,76	3,05	22,81	236,66	9,47	1,04	5,19	26,93
Teste - Baseline	38,22	-9,59	28,63	130,96	-5,24	1,67	8,35	69,74
Teste - Baseline (melhores*)	26,28	-9,41	16,87	54,47	-2,18	1,02	5,10	26,03

* Valores encontrados nas piscadas com $R^2 \geq 0.98$ (melhores piscadas).

A Tabela 9 apresenta o resultado do teste estatístico, a primeira linha é referente a métrica *BlinkMin* (quantidade de piscadas encontradas no sinal por minuto), e na segunda linha da tabela é em relação a métrica *GoodBlinkMin* (quantidades de piscadas classificadas como “melhores” ($R^2 \geq 0,98$) encontradas no sinal por minuto). Ambas as métricas de frequência (taxa de piscada) *BlinkMin* e *GoodBlinkMin* foram estatisticamente significativas.

A Figura 23 compara as distribuições das frequências médias da piscada, agrupando nos períodos de baseline e teste, a mediana do conjunto de dados para baseline é maior que a mediana da frequência média de piscada do conjunto de dados para o teste. Esse resultado está de acordo com a literatura, pois, espera-se que a frequência diminua um pouco durante a execução de uma tarefa que exige uma demanda de carga mental-visual [90].

Tabela 9 - Resultado do teste medidas repetidas para frequência de piscada

Teste t de Medidas Repetidas

Baseline - Teste	Diferenças					t	df	Sig. (bicau- dal)		
	Médi- a	Desvio Padrão	Erro Padrão	95% IC da diferença						
				o	Inferior	Superior				
BlinkMin	5,24	8,35	1,67	1,79	8,69	3,14	24,00	0,00		
GoodBlinkMin*	2,18	5,10	1,02	0,07	4,28	2,14	24,00	0,04		

* Valores encontrados nas piscadas com $R^2 \geq 0.98$ (melhores piscadas).

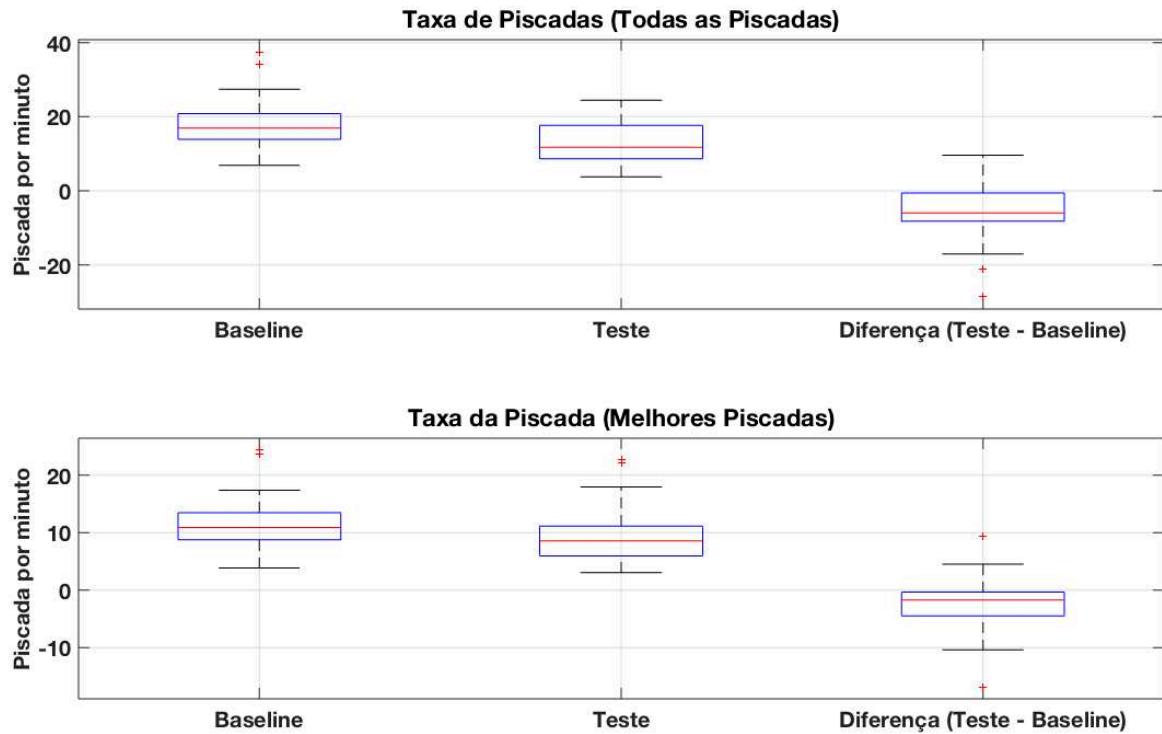


Figura 23 - Gráfico boxplot da frequência das piscadas durante baseline e o teste.

4.4.2. Duração da Piscada

A Tabela 10 apresenta um resumo da duração média da piscada do conjunto de dados do sinal usando *landmarks* (pontos de referência) diferentes, sendo eles: *durationZ*, *durationHZ*, *durationB* e *durationbHB*. Como esperado, o valor médio da variável *durationHZ* (duração Half-zero) teve o menor desvio padrão e uma média mais consistente com os valores relatados. Conforme discutido na seção sobre processamento do sinal EEG (Seção 3.3.1), a variável *durationHZ* corresponde aproximadamente à medição do período de quando a pálpebra passa pela pupila pela primeira vez. Os *eye-trackers* que dependem da detecção da pupila também costumam usar esse ponto como sendo o ponto de partida da piscada [80]. *DurationHZ* também pode ser definido como a duração média da piscada na metade da amplitude a partir da linha zero. Esse valor teve média $105,6 \pm 1,66\text{ms}$ (baseline) e $55,1 \pm 0,87\text{ms}$ (teste).

Tabela 10 - Estatística descritiva da duração da piscada (valores em ms)

Métricas de duração		Min.	Máx.	Média	Erro Padrão	Desvio Padrão	Variância
DurationZ	Baseline	104,0 0	204,0 0	142,92	5,547	27,734	0,769
	Baseline (melhores*)	96,00 0	173,0 0	120,40	3,144	15,719	0,247
	Teste	93,00 0	183,0 0	122,72	4,160	20,800	0,433
	Teste (melhores*)	87,00 0	158,0 0	110,52	3,318	16,588	0,275
DurationHZ	Baseline	88,00 0	198,0	129,60	5,781	28,906	0,836
	Base (melhores*)	86,00 0	159,0	105,56	3,318	16,588	0,275
	Teste	46,00	96,00	60,52	2,465	12,326	0,152
	Teste (melhores*)	43,00	80,00	55,12	1,750	8,748	0,077
DurationB	Baseline	52,00 0	671,0	125,48	25,38	126,87 7	16,098
	Baseline (melhores*)	52,00 0	212,0	79,16	6,62	33,075	1,094
	Teste	141,0 0	223,0 0	173,96	5,08	25,376	0,644
	Teste (melhores*)	138,0 0	225,0 0	166,16	4,84	24,200	0,586
DurationHB	Baseline	50,00 0	103,0	69,40	2,67	13,367	0,179
	Base (melhores*)	50,00	91,00	61,16	1,82	9,114	0,083
	Teste	46,00 0	276,0	82,08	10,16	50,778	2,578
	Teste (melhores*)	46,00 0	359,0	76,16	12,33	61,636	3,799

* Valores encontrados nas piscadas com $R^2 \geq 0.98$ (melhores piscadas).

Tabela 11 mostra os resultados estatísticos. Foi aplicado o teste estatístico de amostras repetidas, de acordo com os resultados, o único valor que não apresentou diferença significativa foi o da duração da piscada na metade da amplitude em relação a base da curva do sinal (*durationHB* mostrado na Tabela 11 em cor vermelha).

No geral, a duração das piscadas durante o teste foi menor comparada a do baseline. De acordo com a literatura, na maioria das piscadas durante estado de alerta, as pálpebras não ficam fechadas por mais que alguns ms e a duração da piscada é entre 100-300ms [82], [84]. No estado de sonolência, algumas piscadas duram mais de 500ms, devido ao fato das pálpebras

permanecerem fechadas por um tempo mais longo, onde a contração dos músculos levantadores das pálpebras superiores são inibidos por um tempo maior que o normal [82], [84]. No geral piscadas com longa duração ($>300\text{ms}$) estão relacionadas com movimentos oculares mais complexos ou estão relacionados à fadiga e sonolência [82], [84].

Os resultados apresentados estão de acordo com a literatura, onde os valores encontrados durante o teste foram menores que os encontrados durante o baseline (período de descanso). Em ambos os períodos, baseline e teste, os valores médios ficaram na faixa que é considerado como “estado em alerta”. A Figura 24 exibe as durações médias das piscadas durante o teste. A duração média das piscadas entre 50-60 ms foram as que tiveram maior frequência na amostra.

Tabela 11 - Resultado da análise de diferença entre a duração das piscadas (s) entre o período baseline e o teste.
Teste t de Medidas Repetidas

Baseline - Teste	Diferenças					t	df	Sig. (bica udal)			
	Média	Desv io Padr ão	Erro Padrã o Médi o	95% IC da diferença							
				Inferio r	Superi or						
Zero	0,02	0,03	0,01	0,01	0,03	3,41	24,00	0,00			
Zero <i>(melhores*)</i>	0,01	0,02	0,00	0,00	0,02	2,56	24,00	0,02			
Tent	0,09	0,03	0,01	0,07	0,10	14,13	24,00	0,00			
Tent <i>(melhores*s)</i>	0,09	0,02	0,00	0,08	0,10	18,81	24,00	0,00			
Half-zero	0,07	0,03	0,01	0,06	0,08	12,42	24,00	0,00			
Half-zero <i>(melhores*)</i>	0,05	0,02	0,00	0,04	0,06	15,66	24,00	0,00			
Half-Base	-0,01	0,05	0,01	-0,03	0,01	-1,26	24,00	0,22			
Half-Base <i>(melhores*)</i>	-0,02	0,06	0,01	-0,04	0,01	-1,22	24,00	0,23			
Base	-0,05	0,13	0,03	-0,10	0,01	-1,83	24,00	0,08			
Base <i>(melhores*)</i>	-0,09	0,04	0,01	-0,10	-0,07	-10,32	24,00	0,00			

* Valores encontrados nas piscadas com $R^2 \geq 0.98$ (melhores piscadas).

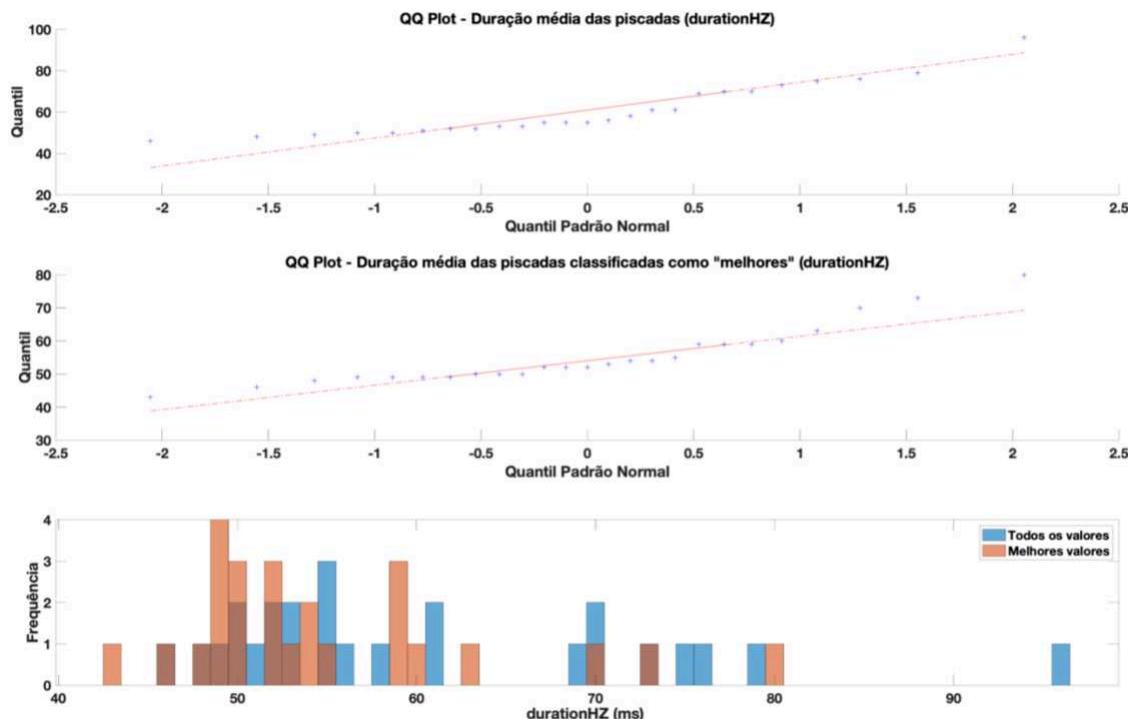


Figura 24 - Distribuição da duração das piscadas. Imagem do topo é um gráfico quantil em relação às distribuições normais da duração média de *durationHZ* para cada conjunto de dados. As linhas pontilhadas vermelhas representam as distribuições normais mais adequadas. O gráfico do meio é a representação da duração média do conjunto de dados considerados “melhores”. O gráfico inferior mostra o histograma *durationHZ*.

4.4.3. Velocidade da Piscada

Outro parâmetro analisado da piscada foi a razão entre amplitude-velocidade da piscada. Como foi discutido na Seção 3.3.1, *PAVRZ* representa essa razão na parte ascendente da curva principal da piscada que corresponde ao fechamento das pálpebras. A métrica *NAVRZ* representa a razão entre amplitude-velocidade na parte descendente que corresponde ao movimento de abertura dos olhos (retração da pálpebra). Essas métricas foram apresentadas como indicadores de fadiga e sonolência em trabalhos anteriores apresentados na literatura [82]–[84].

A Tabela 12 apresenta os valores de estatística descritiva da razão amplitude-velocidade positiva (fechamento dos olhos) e negativa (abertura dos olhos). Tabela 13 mostra os resultados da diferença estatística. Pode-se observar que somente os valores relacionados a abertura dos olhos tiveram uma diferença significativa (*NAVRZ*) entre os períodos de baseline e teste, onde o valor médio durante o teste foi menor, porém ambos os períodos apresentaram as médias na faixa de estado de alerta. Geralmente, valores acima de 6 10msec tem uma relação com o estado de sonolência.

Tabela 12 - Estatística descritiva da razão amplitude-velocidade da piscada em 10msec(AVRZ)

Razão Amplitude- Velocidade da piscada		Min.	Máx.	Média	Erro Padrão	Desvio Padrão	Variância
PAVRZ	Baseline	2,86	4,14	3,54	0,059	0,296	0,087
	Baseline (melhores*))	2,95	4,13	3,48	0,053	0,265	0,070
	Teste	2,84	4,46	3,45	0,082	0,412	0,170
	Teste (melhores*))	2,79	4,44	3,40	0,084	0,422	0,178
NAVRZ	Baseline	2,51	5,48	3,60	0,152	0,762	0,581
	Base (melhores*))	2,49	5,42	3,38	0,131	0,656	0,430
	Teste	2,32	5,27	3,03	0,124	0,620	0,385
	Teste (melhores*))	2,31	6,35	3,01	0,158	0,788	0,621

* Valores encontrados nas piscadas com $R^2 \geq 0.98$ (melhores piscadas).

Tabela 13 - Diferença estatística dos valores PAVRZ e NAVRZ entre os períodos de baseline e teste.

Paired Samples Test

Baseline - Test	Diferença emparelhadas					t	df	Sig. (2- tailed)		
	Médi a	Desvio Padrã o	Error Padrã o	95% Intervalo de confiança da diferença						
				Médio	Inferior	Superior				
PAVRZ	0,09	0,37	0,07	-0,06	0,24	1,19	24,0	0,24		
PAVRZ (melhores*)	0,08	0,37	0,07	-0,07	0,23	1,12	24,0	0,27		
NAVRZ	0,57	0,67	0,13	0,30	0,85	4,29	24,0	0,00		
NAVRZ (melhores*)	0,37	0,78	0,16	0,05	0,69	2,39	24,0	0,02		

* Valores encontrados nas piscadas com $R^2 \geq 0.98$ (melhores piscadas).

4.5. Correlação do NASA-TLX com Índices da Piscada

Foi analisado os coeficientes de correlação entre os índices das piscadas e os escores do NASA-TLX, para isso foi aplicado o teste de correlação linear não paramétrica Spearman. O motivo dessa análise é avaliar se as características das piscadas podem ser utilizadas como análise da carga cognitiva durante o uso do simulador de cadeira de rodas virtual. O coeficiente de correlação de Spearman foi usado para examinar a força da relação linear entre o NASA-TLX e os parâmetros de piscada. Os principais resultados obtidos nessa análise são mostrados na Tabela 14.

Alguns estudos já examinaram alguns parâmetros de piscada em dois níveis de carga cognitiva e foi encontrado correlações com o nível de carga mental [22], [53]. Entretanto, essas correlações com carga mental não foram encontradas com os resultados obtidos no presente estudo. As correlações encontradas foram com as métricas de razão amplitude-velocidade e duração da piscada com a demanda temporal e de esforço, os valores estão apresentados na Tabela 14.

As métricas da duração da piscada (duration Zero e *duration Base*) tiveram correlações negativas fracas a medida ajustada da demanda temporal. A correlação entre razão amplitude-velocidade durante a fechadura da pálpebra (*PAVRZ*) e a demanda temporal é uma relação forte negativa ($r=-0,721$), essa relação está presentada no gráfico de dispersão (Figura 25), pode-se observar que mesmo apresentando uma correlação forte é possível ver vários pontos afastados da linha de tendência. Além disso, *PAVRZ* tem correlação negativa moderada com o peso ($r=-0,515$) e a medida bruta ($r=-0,506$) da demanda temporal, esses valores estão apresentados na Tabela 14. Outra correlação encontrada foi entre razão amplitude-velocidade durante a abertura da pálpebra (*NAVRZ*) e o peso da demanda de esforço ($r=0,399$). Porém, essa correlação não teve efeito na medida ajustada.

Tabela 14 - Resumo dos resultados significativos do teste de correlação Spearman para os índices da piscada

		Demanda Temporal		Demanda de Esforço	
		Peso	Medida	Medida	Peso
		bruta	ajustad	bruta	Medid
		a			
Razão	PAVR	r	-0,515	-0,506	-0,721
amplitude	Z	Sig.	0,008	0,010	0,000
-	NAVR	r			0,399
velocidade	Z	Sig.			0,048
Duração	Durati	r	-0,429		-0,468
da piscada	onZero	Sig.	0,033		0,018
	Durati	r	-0,525		-0,469
	onBase	Sig.	0,007		0,018

Gráfico de Dispersão (Scatter plot)

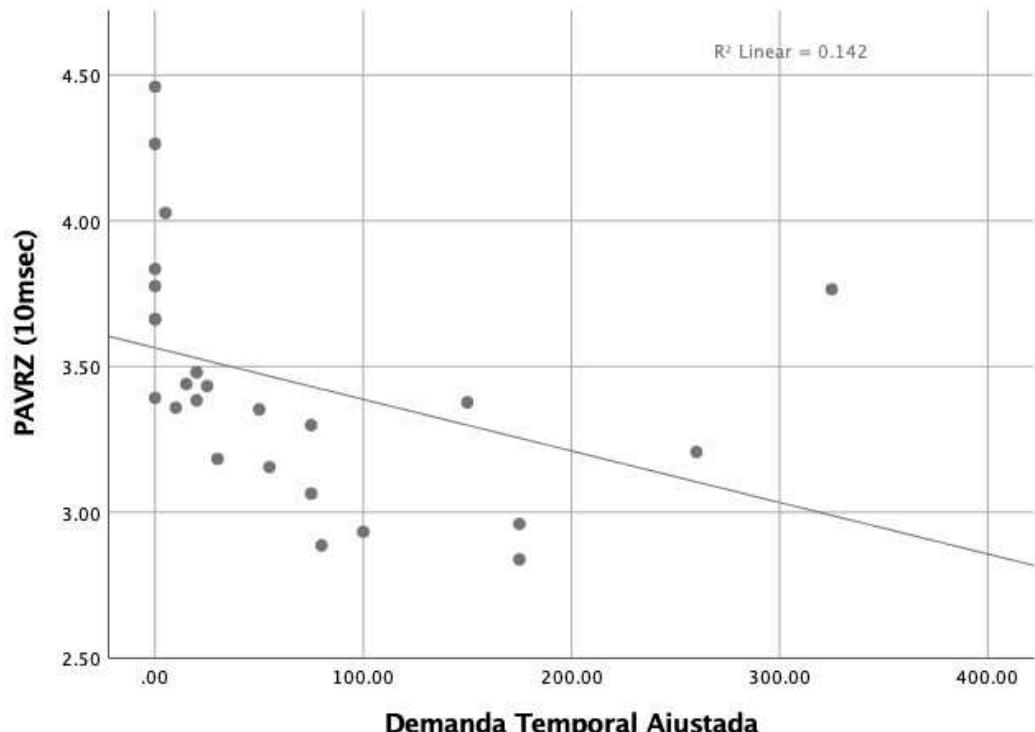


Figura 25 - Gráfico de dispersão (Scatter plot) com uma linha de tendência de PAVRZ por Demanda Temporal Ajustada.

4.6. Correlação do NASA-TLX com Métricas de desempenho

Na literatura é mencionado que as métricas de desempenho podem ser utilizadas como uma forma de inferir carga cognitiva [26]–[28]. Entretanto, alguns autores infere a carga cognitiva com uso de escalas subjetivas de classificação através da aplicação de questionários (e.g. NASA-TLX) e as métricas de desempenho são reportadas juntamente com essas informações para descrever a eficiência relativa das condições instrucionais da tarefa [35], [36].

No presente estudo, foram reportados os valores relativos à carga cognitiva utilizando o NASA-TLX e o desempenho dos participantes utilizando três tipos de métricas (tempo de completar a tarefa, número de comandos e colisões (erros)). O resultado da análise de correlação entre essas métricas está na Tabela 15, onde foram encontradas duas correlações negativas fracas. Na primeira linha da Tabela 15 é mostrado que o número de colisões tem correlação negativa fraca com a medida bruta da demanda física ($r=-0,444$). Na segunda linha é apresentado a correlação negativa fraca entre o tempo de completar a tarefa com a medida ajustada de esforço ($r=-0,426$).

Tabela 15 - Resumo dos resultados significativos do teste de correlação Spearman para as métricas de desempenho

	Demanda Física			Esforço		
	Peso	Medida	Medida	Peso	Medida	Medida
		bruta	ajustada		bruta	ajustada
Número de colisões	r		-0,444			
	Sig.		0,018			
Tempo de completar a tarefa	r				-0,426	
	Sig.					0,024

4.7. Correlação do NASA-TLX com as questões de presença (IPQ)

Foi encontrado uma correlação entre duas questões sobre presença no ambiente virtual (questão 9 e 10) com as variáveis do questionário NASA-TLX, os resultados estão representados na Tabela 16. A questão 9 é uma pergunta sobre envolvimento (atenção dedicada a experiência no ambiente virtual). A questão 10 é uma pergunta sobre realismo (o quanto a experiência no ambiente virtual pareceu consistente com a experiência no mundo real).

Na primeira linha da Tabela 16 é mostrado que a questão 9 tem correlação negativa fraca com a medida bruta da demanda de desempenho ($r=-0,463$) e com as medidas gerais da carga cognitiva bruta ($r = -0,465$) e ajustada ($r = -0,464$). Na segunda linha é apresentado que a questão 10 tem correlação negativa moderada com a medida bruta de desempenho ($r= -0,546$) e negativa fraca com a medida ajustada ($r=-0,487$).

Tabela 16 - Resumo dos resultados significativos do teste de correlação Spearman para as questões de presença (IPQ)

	Desempenho			Medida geral da Carga Cognitiva	
	Peso	Medida	Medida	Medida	Medida
		bruta	ajustada	bruta	ajustada
Questão 9 (envolvimento)	r	-0,463		-.465	-0,464
	Sig.	0,013		0,013	0,013
Questão 10 (realismo)	r	-0,546	-0,487		
	Sig.	0,003	0,009		

4.8. Discussão

O simulador aparenta-se aceito no ponto de vista de usabilidade de acordo com resultados apresentados nos questionários. Porém, no ponto de vista de nível de presença, o simulador apresentou ser neutro, com a moda e a mediana próximo de 3, o que era esperado até certo ponto para a configuração não imersiva sob análise. Essa é uma desvantagem do simulador nesta configuração, por outro lado, a maior vantagem é que não apresenta sintomas de *cybersickness*. É importante que esses sintomas não estejam presentes durante o treinamento dos usuários. Os resultados relacionados ao desempenho foram reportados. Os participantes desta pesquisa fizeram o teste sem nenhuma experiência anterior com cadeira de rodas e com simulador.

Seguindo a literatura, a piscada foi analisada através de parâmetros como amplitude, velocidade e valores médios e máximos das fases de abertura e fechamento das pálpebras. Baseado na Tabela 9, Tabela 11, Tabela 13 e Tabela 14 apresentamos nossos resultados medidos diretamente a partir dos dados experimentais. Os dados foram interpretados como valores médios \pm desvio padrão.

Pode-se perceber que os resultados obtidos, de modo geral, são compatíveis com a literatura, uma vez que a discrepância está dentro do erro estatístico. Valores com alto desvio padrão também foram relatados pela maioria dos autores. Este fato deve-se à média entre diferentes sujeitos. Quando as piscadas são analisadas de apenas um participante, o desvio padrão na maioria das vezes é menor [23], [25], [81], [82], [91].

Estudos apontaram que a relação entre a frequência da piscada e diferentes tipos de tarefa não são simples de justificar [13], [24], [92]. Existem vários fatores que podem afetar esse valor como, por exemplo, problemas de visão, nível de estresse, e outros [30]. Por isso, é importante investigar de forma fracionada e isolada os diferentes efeitos que podem influenciar não somente a frequência, mas outros índices oculares em diversas condições e tipos de tarefa.

A maioria dos argumentos encontrados na literatura sobre carga cognitiva argumentam que a carga mental das tarefas cognitivas tem relação com aumento da taxa de piscadas, mas se tiver uma demanda visual para concluir a tarefa ocorre o oposto, ou seja, a taxa das piscadas diminui [80], [93], [94]. A frequência de piscadas durante a execução da tarefa em um ambiente virtual foi menor quando comparada ao período do baseline (descanso), o que está de acordo com a literatura. Além disso, foram encontradas diferenças estatísticas também na duração e na razão Amplitude-Velocidade das piscadas. Essas métricas são frequentemente usadas para identificar o estado de alerta da pessoa [39], [82]–[84]. Os resultados apresentados mostraram que tanto no baseline quanto na tarefa os valores médios se

mantiveram na faixa fisiológica; e os valores do baseline foram maiores àqueles apurados durante a execução da tarefa.

Os resultados relacionados a correlação entre o NASA-TLX e os índices de piscadas foram relativamente coerentes com a literatura, onde as correlações são apresentadas como moderadas e fracas em termos de força, mas são significantes. Não foi encontrado correlação entre os índices da piscada com a medida geral do NASA-TLX, porém foi encontrado correlação com outras medidas do NASA-TLX (medida bruta e ajustada da demanda temporal) com índices da piscada relacionados a razão amplitude-velocidade (*PAVRZ* e *NAVRZ*) e duração da piscada (*durationZero* e *durationBase*). De modo geral, o estudo pôde apresentar que existem vários pontos de referência da piscada derivados de um sinal EEG que têm o potencial para serem aplicados como métricas implícitas da carga cognitiva, porém com os nossos resultados não foi possível mostrar que são suficientes para determinar a carga cognitiva de forma isolada. Em outras palavras, é necessário aplicar o questionário de avaliação de carga cognitiva em conjunto com a análise do sinal da piscada.

Os resultados relacionados a correlação entre o NASA-TLX e as métricas de desempenho mostraram coerentes, onde desempenho em termos de números de erros (colisões) e tempo da conclusão da tarefa tem correlação negativa fraca com as medidas do NASA-TLX sobre a demanda física e esforço. Isso é esperado, pois essas duas medidas são sobre desempenho.

As questões de presença (IPQ) sobre envolvimento e realismo do sistema tiveram correlações negativas fracas com a demanda de desempenho e a com medida geral de carga cognitiva do NASA-TLX. Isso representa que o nível de “imersão” que o simulador fornece pode afetar a carga cognitiva do usuário. Entretanto, a força dessa correlação é fraca, necessita-se uma melhor investigação para analisar de forma mais aprofundada a relação entre nível de presença e carga cognitiva.

De modo geral, foi possível responder as hipóteses apresentadas na (Seção 1.2.3). As hipóteses nulas relacionadas a diferença de valores dos índices das piscadas entre o baseline e o teste foram retidas. A hipótese sobre as correlações entre o valor da taxa de piscada e as medidas de carga cognitiva do NASA-TLX foi rejeitada. Entretanto, foram encontradas outras correlações com submedidas do questionário NASA-TLX com os índices de duração e razão amplitude-velocidade da piscada. A hipótese sobre a correlação entre as métricas de desempenho e as medidas de carga cognitiva bruta e ajustada foi rejeitada. Entretanto, foi encontrada uma correlação entre o tempo de execução da tarefa com a submedida de demanda de esforço do NASA-TLX.

Em suma, neste estudo, foi empregado a teoria da carga cognitiva para se avaliar a demanda cognitiva a partir das características da piscada e do questionário NASA-TLX. As contribuições presentes foram as seguintes: demonstração dos efeitos da carga cognitiva nas características da piscada, estendendo o uso dessa medida para medição da carga cognitiva em estudos na área de tecnologia assistiva e confirmação as possibilidades de utilizar características da piscada para medição da carga cognitiva.

5. CONCLUSÃO

Neste capítulo, são mostradas as conclusões e limitações da pesquisa apresentada nesta dissertação. Além disso, em função do conhecimento adquirido no decorrer do trabalho, também são apresentadas sugestões para trabalhos futuros para aprimorar a proposta.

5.1. Conclusões

Este trabalho proporcionou investigar a viabilidade do uso do sinal de piscada durante o uso do simulador de cadeira de rodas para detecção da carga cognitiva utilizando um sensor vestível de baixo custo, o qual pode ser utilizado em outras aplicações. Além disso, foi possível responder as hipóteses apresentadas na Seção 1.2.3.

De acordo com os resultados apresentados, foi possível identificar diferença estatística na taxa de piscada onde esse valor apresentou ser maior durante o baseline e menor durante a execução da tarefa (período de teste), sendo assim, a hipótese nula H01 foi retida. Esses resultados da taxa de piscada estão em acordo com a literatura [20]–[25], onde a taxa de piscada foi utilizada como indicador na identificação de demanda cognitiva. Entretanto, os resultados da hipótese sobre as correlações entre os índices da piscada com a medida geral (escore final) do NASA-TLX foi rejeitada. Devido ao fato dessa hipótese ser rejeitada, os índices das piscadas não podem ser aplicados como uma medida exclusiva para inferir a carga cognitiva durante a utilização do simulador.

Os resultados das hipóteses sobre métricas de desempenho terem uma correlação com o escore final do NASA-TLX também foram rejeitadas, porém essas variáveis tiveram correlações com submedida do NASA-TLX. Esse resultado representa que a análise do desempenho dos participantes tem um certo nível de correlação com a demanda cognitiva. Pode ser utilizada em conjunto com NASA-TLX fornecer mais informações sobre a carga cognitiva em diferentes tipos de tarefas como é argumentado na literatura [14], [26]–[30].

Espera-se, em longo prazo, que seja possível implementar outras formas implícitas de identificar a carga cognitiva no intuito de atenuar erros provenientes de análises subjetivas. Do ponto de vista científico, as principais contribuições deste trabalho são:

- Disponibilização de alternativas como avaliar a carga cognitiva por meio de análise do sinal da piscada utilizando apenas um canal de EEG.
- Inclusão de formas de avaliar o simulador de cadeira de rodas além das avaliações convencionais por questionário e métricas de desempenho.

5.2. Trabalhos Futuros

Ainda que os resultados foram coerentes com a literatura, foi possível perceber que há um grande potencial de refinar a análise e investigar outros protocolos de experimento para coletar a carga cognitiva em diferentes tipos de tarefa, do simples ao complexo. Na literatura, existem várias opções disponíveis para avaliar carga cognitiva, com uso frequente de questionários (abordagem subjetiva) em conjunto com uma ou duas métricas objetivas como, por exemplo, a resposta galvânica da pele e medidas de desempenho.

Existe uma demanda por estudos que apresentem resultados sobre a relação entre diferentes tarefas em realidade virtual com carga cognitiva. Esses estudos servem para a tomada de decisão de quais abordagens podem ajudar resolver problemas como *cybersickness*. Além disso, existe uma outra demanda de investigação referente ao entre o nível de presença no ambiente virtual com a carga cognitiva. Os resultados apresentados na Seção 4.7 apresentou indícios de correlações das questões do questionário de presença com a medida geral da carga cognitiva.

Em relação ao design do estudo, faz-se necessário averiguar diferentes tipos de tarefas no ambiente virtual e quais serão os valores de carga cognitiva. Sendo os tipos de tarefas únicos ou duais (*dual-task*) [28].

Por fim, sugere-se investigar o uso de técnicas de inteligência artificial (ex: redes neurais artificiais, algoritmos de aprendizagem de máquina, etc.) no intuito de conceber um modelo computacional para reconhecimento de padrões. Tal modelo seria aplicado no monitoramento contínuo da carga cognitiva visando garantir o bem-estar e atenção do usuário na realização de tarefas diversas.

REFERÊNCIAS

- [1] M. B. Lenzi, “Os dados sobre Deficiência nos Censos Demográficos Brasileiros,” *XVIII Encontro Nac. Estud. Popul. - Transform. na Popul. Bras. complexidades, Incert. e Perspect.*, no. Xviii, pp. 1–20, 2012 [Online].
- [2] W. Health Organization, “WORLD REPORT ON DISABILITY,” 2011 [Online].
- [3] Central Statistics Office, “Disability - CSO - Central Statistics Office,” 2016. [Online].
- [4] CENSO, “Cartilha do Censo 2010: Pessoas com Deficiência.” 2010 [Online].
- [5] IBGE, *Pesquisa Nacional de Saúde 2013*. 2014 [Online].
- [6] IBGE, “Releitura dos dados de pessoas com deficiência no Censo Demográfico 2010 à luz das recomendações do Grupo de Washington - NotaTécnica 01/2018,” pp. 1–8, 2018 [Online].
- [7] D. M. Cruz *et al.*, “Assistive Technology Accessibility and Abandonment: Challenges for Occupational Therapists,” *Open J. Occup. Ther.*, vol. 4, pp. 1–1, 2016, doi: [10.15453/2168-6408.1166](https://doi.org/10.15453/2168-6408.1166). [Online].
- [8] A. Harrison Derwent, G., Enticknap, A., Rose, F.D., and Attree, A., “The role of virtual reality technology in the assessment and training of inexperienced powered wheelchair users,” *Disabil. Rehabil.*, vol. 24, no. 11–12, pp. 599–606, 2002.
- [9] K. T. Manis and D. Choi, “The virtual reality hardware acceptance model (VR-HAM): Extending and individualizing the technology acceptance model (TAM) for virtual reality hardware,” *J. Bus. Res.*, 2019, doi: [10.1016/j.jbusres.2018.10.021](https://doi.org/10.1016/j.jbusres.2018.10.021).
- [10] R. Verza, M. L. L. Carvalho, M. a Battaglia, and M. M. Uccelli, “An interdisciplinary approach to evaluating the need for assistive technology reduces equipment abandonment.,” *Mult. Scler.*, vol. 12, no. May 2005, pp. 88–93, 2006, doi: [10.1191/1352458506ms1233oa](https://doi.org/10.1191/1352458506ms1233oa).
- [11] P. B., Z. H., B. Phillips, and H. Zhao, “Predictors of assistive technology abandonment,” *Assist. Technol.*, vol. 5, no. 1, pp. 35–45, 1993.
- [12] L. M. P. Bracciali, *Qualidade de Vida e Novas Tecnologias / FEF*. Campinas: Ipes Editorial, 2007 [Online].
- [13] J. Leppink, “Cognitive load theory: Practical implications and an important challenge,” *J. Taibah Univ. Med. Sci.*, vol. 12, no. 5, pp. 385–391, 2017, doi: [10.1016/j.jtumed.2017.05.003](https://doi.org/10.1016/j.jtumed.2017.05.003). [Online].

- [14] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. Van Gerven, "Cognitive load measurement as a means to advance cognitive load theory," *Educ. Psychol.*, vol. 38, no. 1, pp. 63–71, 2003, doi: [10.1207/S15326985EP3801_8](https://doi.org/10.1207/S15326985EP3801_8).
- [15] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Adv. Psychol.*, 1988, doi: [10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9).
- [16] L. M. Naismith, J. J. H. Cheung, C. Ringsted, and R. B. Cavalcanti, "Limitations of subjective cognitive load measures in simulation-based procedural training," *Med. Educ.*, vol. 49, no. 8, pp. 805–814, 2015, doi: [10.1111/medu.12732](https://doi.org/10.1111/medu.12732).
- [17] S. Chen and J. Epps, "Using task-induced pupil diameter and blink rate to infer cognitive load," *Human-Computer Interact.*, vol. 29, no. 4, pp. 390–413, 2014, doi: [10.1080/07370024.2014.892428](https://doi.org/10.1080/07370024.2014.892428).
- [18] F. Roque Martins, "Simulador Para Treinamento de Cadeirantes em Ambinete Virtual Acionado Por Comandos Musculares e/ou Visuais," Federal University of Uberlândia, 2017.
- [19] D. C. Kamaraj, B. E. Dicianno, H. P. Mahajan, A. M. Buhari, and R. A. Cooper, "Stability and Workload of the Virtual Reality-Based Simulator-2," *Arch. Phys. Med. Rehabil.*, 2016, doi: [10.1016/j.apmr.2016.01.032](https://doi.org/10.1016/j.apmr.2016.01.032).
- [20] T. Čegovnik, K. Stojmenova, G. Jakus, and J. Sodnik, "An analysis of the suitability of a low-cost eye tracker for assessing the cognitive load of drivers," *Appl. Ergon.*, vol. 68, no. October 2017, pp. 1–11, 2018, doi: [10.1016/j.apergo.2017.10.011](https://doi.org/10.1016/j.apergo.2017.10.011).
- [21] C. Fogarty and J. A. Stern, "Eye movements and blinks: their relationship to higher cognitive processes," *Int. J. Psychophysiol.*, 1989, doi: [10.1016/0167-8760\(89\)90017-2](https://doi.org/10.1016/0167-8760(89)90017-2).
- [22] R. Martins and J. M. Carvalho, "Eye blinking as an indicator of fatigue and mental load—a systematic review," in *Occupational Safety and Hygiene III - Selected Extended and Revised Contributions from the International Symposium on Safety and Hygiene*, 2015, doi: [10.1201/b18042-48](https://doi.org/10.1201/b18042-48).
- [23] S. Chen, J. Epps, N. Ruiz, and F. Chen, "Eye activity as a measure of human mental effort in HCI," in *International Conference on Intelligent User Interfaces, Proceedings IUI*, 2011, doi: [10.1145/1943403.1943454](https://doi.org/10.1145/1943403.1943454).
- [24] F. Chen *et al.*, "Eye-Based Measures," 2016.
- [25] A. R. Bentivoglio, S. B. Bressman, E. Cassetta, D. Carretta, P. Tonali, and A.

- Albanese, "Analysis of blink rate patterns in normal subjects," *Mov. Disord.*, 1997, doi: [10.1002/mds.870120629](https://doi.org/10.1002/mds.870120629).
- [26] P. Ayres, "Using subjective measures to detect variations of intrinsic cognitive load within problems," *Learn. Instr.*, vol. 16, no. 5, pp. 389–400, 2006, doi: [10.1016/j.learninstruc.2006.09.001](https://doi.org/10.1016/j.learninstruc.2006.09.001).
- [27] N. Hollender, C. Hofmann, M. Deneke, and B. Schmitz, "Integrating cognitive load theory and concepts of human-computer interaction," *Comput. Human Behav.*, vol. 26, no. 6, pp. 1278–1288, 2010, doi: [10.1016/j.chb.2010.05.031](https://doi.org/10.1016/j.chb.2010.05.031). [Online].
- [28] R. Brünken, J. L. Plass, and D. Leutner, "Direct measurement of cognitive load in multimedia learning," *Educ. Psychol.*, vol. 38, no. 1, pp. 53–61, 2003, doi: [10.1207/S15326985EP3801_7](https://doi.org/10.1207/S15326985EP3801_7).
- [29] F. G. W. C. Paas and J. J. G. Van Merriënboer, "Instructional control of cognitive load in the training of complex cognitive tasks," *Educ. Psychol. Rev.*, 1994, doi: [10.1007/BF02213420](https://doi.org/10.1007/BF02213420).
- [30] F. Paas, G. Camp, and R. Rikers, "Instructional compensation for age-related cognitive declines: Effects of goal specificity in maze learning," *J. Educ. Psychol.*, 2001, doi: [10.1037/0022-0663.93.1.181](https://doi.org/10.1037/0022-0663.93.1.181).
- [31] B. Cinaz, R. La Marca, B. Arnrich, and G. Tröster, "Towards Continuous Monitoring of Mental Workload," *Int. Work. Ubiquitous Heal. Wellness*, no. September, 2010.
- [32] J. Sweller, "Cognitive load theory, learning difficulty, and instructional design," *Learn. Instr.*, vol. 4, no. 4, pp. 295–312, 1994, doi: [10.1016/0959-4752\(94\)90003-5](https://doi.org/10.1016/0959-4752(94)90003-5).
- [33] J. Sweller, *Cognitive Load Theory*, vol. 55. Elsevier Inc., 2011 [Online].
- [34] L. Plass Jan, R. Moreno, and B. Roland, *Cognitive Load Theory*. 2010.
- [35] F. G. W. C. Paas and J. J. G. Van Merriënboer, "Variability of Worked Examples and Transfer of Geometrical Problem-Solving Skills: A Cognitive-Load Approach," *J. Educ. Psychol.*, 1994, doi: [10.1037/0022-0663.86.1.122](https://doi.org/10.1037/0022-0663.86.1.122).
- [36] F. G. W. C. Paas and J. J. G. Van Merriënboer, "The efficiency of instructional conditions: An approach to combine mental effort and performance measures," *Hum. Factors*, 1993, doi: [10.1177/001872089303500412](https://doi.org/10.1177/001872089303500412).
- [37] N. Nourbakhsh, Y. Wang, F. Chen, and R. A. Calvo, "Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks," in

Proceedings of the 24th Australian Computer-Human Interaction Conference, OzCHI 2012, 2012, doi: [10.1145/2414536.2414602](https://doi.org/10.1145/2414536.2414602).

- [38] J. Xu, Y. Wang, F. Chen, and E. Choi, "Pupillary response based cognitive workload measurement under luminance changes," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, doi: [10.1007/978-3-642-23771-3_14](https://doi.org/10.1007/978-3-642-23771-3_14).
- [39] C. Berka *et al.*, "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviat. Sp. Environ. Med.*, 2007.
- [40] A. Knoll *et al.*, "Measuring cognitive workload with low-cost electroencephalograph," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, doi: [10.1007/978-3-642-23768-3_84](https://doi.org/10.1007/978-3-642-23768-3_84).
- [41] J. Leppink, F. Paas, C. P. M. Van der Vleuten, T. Van Gog, and J. J. G. Van Merriënboer, "Development of an instrument for measuring different types of cognitive load," *Behav. Res. Methods*, vol. 45, no. 4, pp. 1058–1072, 2013, doi: [10.3758/s13428-013-0334-1](https://doi.org/10.3758/s13428-013-0334-1).
- [42] T. Van Gog and H. Jarodzka, "Eye Tracking as a Tool to Study and Enhance Cognitive and Metacognitive Processes in Computer-Based Learning Environments," *Appl. Cogn. Psychol.*, vol. 21, no. 3, pp. 117–128, 2016, doi: [10.1007/978-1-4419-5546-3](https://doi.org/10.1007/978-1-4419-5546-3). [Online].
- [43] R. R. Whelan, "Neuroimaging of cognitive load in instructional multimedia," *Educ. Res. Rev.*, 2007, doi: [10.1016/j.edurev.2006.11.001](https://doi.org/10.1016/j.edurev.2006.11.001).
- [44] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, "Using Electroencephalography to Measure Cognitive Load," *Educational Psychology Review*, 2010.
- [45] A. M. Hughes, G. M. Hancock, S. L. Marlow, K. Stowers, and E. Salas, "Cardiac Measures of Cognitive Workload: A Meta-Analysis," *Hum. Factors*, 2019, doi: [10.1177/0018720819830553](https://doi.org/10.1177/0018720819830553).
- [46] A. Nakasone, H. Prendinger, and M. Ishizuka, "Emotion Recognition from Electromyography and Skin Conductance," *5th Int. Work. Biosignal Interpret.*, 2005.
- [47] A. Niculescu, Y. Cao, and A. Nijholt, "Manipulating Stress and Cognitive Load in Conversational Interactions with a Multimodal System for Crisis Management

- Support," 2010.
- [48] Y. Shi, N. Ruiz, R. Taib, E. Choi, and F. Chen, "Galvanic skin response (GSR) as an index of cognitive load," in *Conference on Human Factors in Computing Systems - Proceedings*, 2007, doi: [10.1145/1240866.1241057](https://doi.org/10.1145/1240866.1241057).
- [49] E. Haapalainen, S. Kim, J. F. Forlizzi, and A. K. Dey, "Psycho-physiological measures for assessing cognitive load," in *UbiComp'10 - Proceedings of the 2010 ACM Conference on Ubiquitous Computing*, 2010, doi: [10.1145/1864349.1864395](https://doi.org/10.1145/1864349.1864395).
- [50] C. S. Ikehara and M. E. Crosby, "Assessing cognitive load with physiological sensors," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2005, doi: [10.1109/hicss.2005.103](https://doi.org/10.1109/hicss.2005.103).
- [51] K. Huttunen, H. Keränen, E. Väyrynen, R. Pääkkönen, and T. Leino, "Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights," *Appl. Ergon.*, 2011, doi: [10.1016/j.apergo.2010.08.005](https://doi.org/10.1016/j.apergo.2010.08.005).
- [52] O. V. Lipp and D. L. Neumann, "Attentional blink reflex modulation in a continuous performance task is modality specific," *Psychophysiology*, 2004, doi: [10.1111/j.1469-8986.00165.x](https://doi.org/10.1111/j.1469-8986.00165.x).
- [53] M. K. Holland and G. Tarlow, "Blinking and mental load.,," *Psychol. Rep.*, 1972, doi: [10.2466/pr0.1972.31.1.119](https://doi.org/10.2466/pr0.1972.31.1.119).
- [54] J. C. F. de Winter, "Controversy in human factors constructs and the explosive use of the NASA-TLX: A measurement perspective," *Cogn. Technol. Work*, 2014, doi: [10.1007/s10111-014-0275-1](https://doi.org/10.1007/s10111-014-0275-1).
- [55] K. E. Laver, B. Lange, S. George, J. E. Deutsch, G. Saposnik, and M. Crotty, "Virtual reality for stroke rehabilitation," *Cochrane Database of Systematic Reviews*. 2017.
- [56] J. McComas, J. Pivik, and M. Laflamme, "Current uses of virtual reality for children with disabilities," *Stud. Health Technol. Inform.*, vol. 58, pp. 161–169, 1998, doi: [10.3233/978-1-60750-902-8-161](https://doi.org/10.3233/978-1-60750-902-8-161).
- [57] E. F. Damasceno, A. Cardoso, and E. A. Lamounier, "An middleware for motion capture devices applied to virtual rehab," in *Proceedings - IEEE Virtual Reality*, 2013, pp. 171–172, doi: [10.1109/VR.2013.6549417](https://doi.org/10.1109/VR.2013.6549417).
- [58] A. Faisal, "Computer science: Visionary of virtual reality," *Nature*, vol. 551, p. 298, Nov. 2017, doi: <http://dx.doi.org/10.1038/551298a>. [Online].
- [59] S. Nichols and H. Patel, "Health and safety implications of virtual reality: A

- review of empirical evidence," *Appl. Ergon.*, 2002, doi: [10.1016/S0003-6870\(02\)00020-0](https://doi.org/10.1016/S0003-6870(02)00020-0).
- [60] H. Rivera, K. A. Hernandez-ossa, B. Longo, E. Caldeira, and T. Bastos, "Evaluation of Cybersickness and Sense of Presence in a VR Simulator of Electric-Powered Wheelchairs," pp. 1–4.
- [61] C. Gartner; Pettey, "3 Reasons Why VR and AR Are Slow to Take Off - Smarter With Gartner," 2018. [Online].
- [62] A. Mazloumi Gavgani, F. R. Walker, D. M. Hodgson, and E. Nalivaiko, "A comparative study of cybersickness during exposure to virtual reality and 'classic' motion sickness: are they different?," *J. Appl. Physiol.*, vol. 125, no. 6, pp. 1670–1680, Dec. 2018, doi: [10.1152/japplphysiol.00338.2018](https://doi.org/10.1152/japplphysiol.00338.2018). [Online].
- [63] E. A. L. Lee and K. W. Wong, "Learning with desktop virtual reality: Low spatial ability learners are more positively affected," *Comput. Educ.*, 2014, doi: [10.1016/j.compedu.2014.07.010](https://doi.org/10.1016/j.compedu.2014.07.010).
- [64] A. Hasdai, A. S. Jesse, and L. Patrice, "Aya Hasdai, Adam S. Jesse!, Patrice L. Weiss," pp. 215–220, 1995.
- [65] A. N. Silva, Y. Morere, E. L. M. Naves, A. A. R. De Sa, and A. B. Soares, "Virtual electric wheelchair controlled by electromyographic signals," *ISSNIP Biosignals Biorobotics Conf. BRC*, pp. 1–5, 2013, doi: [10.1109/BRC.2013.6487517](https://doi.org/10.1109/BRC.2013.6487517).
- [66] L. Ribeiro Borges, F. Roque Martins, and E. Lázaro Martins Naves, "Electric-Powered Wheelchair Control Using Eye Tracking Techniques," *Int. J. Innov. Res. Comput. Commun. Eng. (An ISO, vol. 3297, no. 9, 2016, doi: 10.15680/IJIRCCE.2016*. [Online].
- [67] N. Pittaras, F. Markatopoulou, V. Mezaris, and I. Patras, "Comparison of fine-tuning and extension strategies for deep convolutional neural networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10132 LNCS, pp. 102–114, 2017, doi: [10.1007/978-3-319-51811-4_9](https://doi.org/10.1007/978-3-319-51811-4_9).
- [68] L. Chouinard-Thuly *et al.*, "Technical and conceptual considerations for using animated stimuli in studies of animal behavior.," *Curr. Zool.*, vol. 63, no. 1, pp. 5–19, Feb. 2017, doi: [10.1093/cz/zow104](https://doi.org/10.1093/cz/zow104). [Online].
- [69] H. P. Mahajan, B. E. Dicianno, R. A. Cooper, and D. Ding, "Assessment of wheelchair driving performance in a virtual reality-based simulator," *J. Spinal Cord Med.*, vol. 36, no. 4, pp. 322–332, 2013, doi:

- [10.1179/2045772313Y.0000000130](https://doi.org/10.1179/2045772313Y.0000000130).
- [70] L. Devigne, M. Babel, F. Nouviale, V. K. Narayanan, F. Pasteau, and P. Gallien, “Design of an immersive simulator for assisted power wheelchair driving,” *IEEE Int. Conf. Rehabil. Robot.*, pp. 995–1000, 2017, doi: [10.1109/ICORR.2017.8009379](https://doi.org/10.1109/ICORR.2017.8009379).
- [71] V. Moya, E. Slawiński, V. Mut, and E. H. M. Couto, “Workload detection based on EEG device for teleoperation of a mobile robot,” *2017 17th Work. Inf. Process. Control. RPIC 2017*, vol. 2017-Janua, pp. 1–6, 2017, doi: [10.23919/RPIC.2017.8211636](https://doi.org/10.23919/RPIC.2017.8211636).
- [72] C. X. Li *et al.*, “Mental workload of young drivers during curve negotiation,” *2014 Int. Conf. Connect. Veh. Expo, ICCVE 2014 - Proc.*, pp. 291–292, 2014, doi: [10.1109/ICCVE.2014.7297558](https://doi.org/10.1109/ICCVE.2014.7297558).
- [73] Igroup Project Consortium, “Igroup Presence Questionnaire,” *igroup.org*, 2015.
- .
- [74] E. Caetano, Daniel; Mattioli, Fernando; Lamounier, Edgard; Cardoso, Alexandre; Naves, “Adaptação de uma Interface USB para Joystick VR2 aplicada ao Treinamento de Usuários de Cadeira de Rodas.”
- [75] “American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature1,” *J. Clin. Neurophysiol.*, 1991, doi: [10.1097/00004691-199104000-00007](https://doi.org/10.1097/00004691-199104000-00007).
- [76] J. W. Matiko, S. Beeby, and J. Tudor, “Real time eye blink noise removal from EEG signals using morphological component analysis,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, no. February 2015, pp. 13–16, 2013, doi: [10.1109/EMBC.2013.6609425](https://doi.org/10.1109/EMBC.2013.6609425).
- [77] C. Kothe, “Lab streaming layer (lsl),” <https://github.com/sccn/labstreaminglayer>. Accessed on October, 2014. .
- [78] J. Brooke, “SUS - A quick and dirty usability scale,” in *Usability Evaluation in Industry*, 1996.
- [79] A. Delorme and S. Makeig, “EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” *J. Neurosci. Methods*, 2004, doi: [10.1016/j.jneumeth.2003.10.009](https://doi.org/10.1016/j.jneumeth.2003.10.009).
- [80] K. Kleifges, N. Bigdely-Shamlo, S. E. Kerick, and K. A. Robbins, “BLINKER: Automated extraction of ocular indices from EEG enabling large-scale analysis,” *Front. Neurosci.*, vol. 11, no. FEB, pp. 1–17, 2017, doi: [10.3389/fnins.2017.00011](https://doi.org/10.3389/fnins.2017.00011).

[10.3389/fnins.2017.00012](https://doi.org/10.3389/fnins.2017.00012).

- [81] P. P. Caffier, U. Erdmann, and P. Ullsperger, “Experimental evaluation of eye-blink parameters as a drowsiness measure,” *Eur. J. Appl. Physiol.*, vol. 89, no. 3–4, pp. 319–325, 2003, doi: [10.1007/s00421-003-0807-5](https://doi.org/10.1007/s00421-003-0807-5).
- [82] M. W. Johns, “The amplitude velocity ratio of blinks: A new method for monitoring drowsiness,” *Sleep*, vol. 26, pp. A51-52, 2003.
- [83] M. Ingre, T. Åkerstedt, B. Peters, A. Anund, and G. Kecklund, “Subjective sleepiness, simulated driving performance and blink duration: Examining individual differences,” *J. Sleep Res.*, vol. 15, no. 1, pp. 47–53, 2006, doi: [10.1111/j.1365-2869.2006.00504.x](https://doi.org/10.1111/j.1365-2869.2006.00504.x).
- [84] R. Schleicher, N. Galley, S. Briest, and L. Galley, “Blinks and saccades as indicators of fatigue in sleepiness warnings: Looking tired?,” *Ergonomics*, vol. 51, no. 7, pp. 982–1010, 2008, doi: [10.1080/00140130701817062](https://doi.org/10.1080/00140130701817062).
- [85] S. G. Hart, “NASA-task load index (NASA-TLX); 20 years later,” in *Proceedings of the Human Factors and Ergonomics Society*, 2006.
- [86] B. A. Myers, “A brief history of human-computer interaction technology,” *Interactions*, vol. 5, no. 2, pp. 44–54, 1998, doi: [10.1145/274430.274436](https://doi.org/10.1145/274430.274436).
- [87] ITU, “ITU-P.913 : Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment.” [Online].
- [88] S. Sue, “Test distance vision using a Snellen chart.,” *Community eye Heal.*, vol. 20, no. 63, p. 52, Sep. 2007 [Online].
- [89] A. Hoffmann and M. Menozzi, “Applying the Ishihara test to a PC-based screening system,” *Displays*, vol. 20, no. 1, pp. 39–47, 1999, doi: [10.1016/S0141-9382\(98\)00053-5](https://doi.org/10.1016/S0141-9382(98)00053-5).
- [90] A. A. V. Cruz, D. M. Garcia, C. T. Pinto, and S. P. Cechetti, “Spontaneous eyeblink activity,” *Ocul. Surf.*, vol. 9, no. 1, pp. 29–41, 2011, doi: [10.1016/S1542-0124\(11\)70007-6](https://doi.org/10.1016/S1542-0124(11)70007-6). [Online].
- [91] J. Perez, J. Espinosa, B. Domenech, D. Mas, and C. Illueca, “Blinking kinematics description through non-invasive measurement,” *J. Mod. Opt.*, vol. 58, no. 19–20, pp. 1857–1863, 2011, doi: [10.1080/09500340.2011.554898](https://doi.org/10.1080/09500340.2011.554898).
- [92] R. Brünken, J. L. Plass, and R. Moreno, “Current issues and open questions in cognitive load research,” *Cogn. Load Theory*, no. May, pp. 253–272, 2010, doi: [10.1017/CBO9780511844744.014](https://doi.org/10.1017/CBO9780511844744.014).

- [93] M. Á. Recarte, E. Pérez, Á. Conchillo, and L. M. Nunes, "Mental Workload and Visual Impairment: Differences between Pupil, Blink, and Subjective Rating," *Span. J. Psychol.*, 2008, doi: [10.1017/s1138741600004406](https://doi.org/10.1017/s1138741600004406).
- [94] J. B. Nezlek, "A multilevel framework for understanding relationships among traits, states, situations and behaviours," *Eur. J. Pers.*, 2007, doi: [10.1002/per.640](https://doi.org/10.1002/per.640).

ANEXO A – Formulario de Consentimento

Consent Form

Title of Project: A Quality of Experience Evaluation of a Wheelchair Training Simulator in Virtual Reality Environments.

Name of Researcher: Débora Pereira Salgado

Please Tick the Box

1. I am satisfied that I understand the information provided and have had enough time to consider the information.
2. I do not suffer from photosensitive epilepsy or any other form of epilepsy.
3. I'm not pregnant and/or I am not experiencing any symptoms of pregnancy.
4. I have not consumed alcohol beverages for the last 24 hours.
5. I slept at least 6 hours on last 24 hours.
6. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my legal rights being affected.
7. I understand that any data collected in the course of this study will be used for research purpose only and in the strictest confidence. Any information related to me will be discarded at the completion of this research.
8. I agree to take part in the above study.
9. I confirm that I have read the information sheet dated ____/____/2018 for the above study and have had the opportunity to ask questions.

10. Gender: Female Male

Name of Participant

Date

Signature

Name of Person taking consent
(if different from researcher)

Date

Signature

Researcher

Date

Signature

ANEXO B – Informações sobre a pesquisa

Information Sheet

Principal Investigator: Débora Pereira Salgado

Title of project: A Quality of Experience Evaluation of Wheelchair Simulator in Virtual Reality Environments

In this experiment, we aim to evaluate user quality of experience when using a traditional desktop setup and with a virtual reality head mounted display. Participants will be randomly divided into two groups (desktop and Headset group), a script will be used in the MATLAB software to randomly select which group the participants will be divided. Using the conventional screen allows users to view information and interact with virtual objects whilst being aware of their surroundings that is not relate with the simulator. Virtual reality head mounted displays project virtual objects to the wearers field of view, however they are not aware of their surroundings. We aim to capture data using both devices in order to determine if a user's quality of experience is enriched using an conventional desktop setup (common monitor) or virtual reality HMDs device.

Introduction

I am inviting you to take part in a research experiment to be carried out in the Software Research Institute in Athlone Institute of Technology. The aim of this document is to explain why the research is being carried out and what it will involve.

If you are not clear on any points, please do not hesitate ask questions. Thank you for reading this.

What is the purpose of the project?

Virtual Reality (VR) technology has improved significantly in recent years to the point where recreated virtual scenes now provide the user with a sense of realism. With the improvement of graphics cards, increased computational power, 360-degree 3D HD cameras and many more technological advances, these are rapidly changing the future potential of VR. In this experiment, we

aim to evaluate if a higher quality of experience is experienced using traditional desktop compared to VR Head mounted display device . The aim of this experiment is to evaluate quality of experience within PC monitor screen and HDM while using a wheelchair training simulator.

Do I have to take part?

It is entirely up to you to decide whether you wish to take part in this experiment. Refusal to take part is entirely at your discretion. If you decide to take part, you can keep this information sheet and will be required to sign a consent form.

What does the experiment involve?

This experiment should last a minimum of 25 mins and no more than 30 min. Participants will be seated in a laboratory in the AIT Engineering Building. The lab will consist of desktop computer, or VR headset, the non-invasive objective metric sensor (E4 and Mindwave)which the E4 will be attached to the participant's wrist and Mindwave will be put on the head. The test will involve exposing participants to a virtual scene using a conventional monitor or virtual reality head mounted display. The virtual scene consists of circuit with obstacles where the user must drive a wheelchair without commit collisions. The participant will be required to use USB joystick or keyboard to interact with the virtual environment while wearing non-intrusive equipment. The participant will be asked to fill out a questionnaire at the end of the experiment to give their thought on the quality of experience.

What do I have to do?

On the day of the test, participants will undergo a visual screening to ensure they are eligible for the test. The visual screening process involves testing the participant's visual clarity using a Snellen chart, and testing a participant's colour perception using the Ishihara test.

If you are pregnant or suspect that you may be pregnant, please let the administrator of the test know.

If you did not sleep at least 6 hours on the previous night, please let the administrator of test know.

If you consumed alcohol in the last 24 hours, please let the administrator of test know.

What are the possible disadvantages and risks of taking part?

Some users may feel some nausea when using a head mounted display (VR), but this soon goes away after removing the headset. Should a participant at any point feel a high level of nausea it is important to communicate this to the PI.

Will my taking part in this project be kept confidential?

Any information collected during this test will be strictly confidential. All data will be stored in a secure manner and it will not be possible to recognise you from this experiment.

What will happen to the results of the research project?

The results of this experiment will be used to produce a paper for publication as part of my research.

Thanks!

Just like to say, thank you very much for your time and help with this experiment.

ANEXO C – Questionários aplicados

Subject ID: ___ Date: ___/___/___ Gender: F M Environment: VR Screen

Questionnaire – Part 1

You will see some statements about experiences. Please indicate, whether or not each statement applies to your experience. There are no right or wrong answers, only your opinions counts. And please remember: Answer all these questions only referring to this one experience. Read through them to make sure you understand the statement. If you have any questions, please ask your administrator.

1. I found the system unnecessarily complex.

_____ | _____ | _____ | _____ | _____ |

Strongly Disagree Neither Agree Strongly
Disagree Agree/Disagree Agree

2. I thought the system was easy to use.

_____ | _____ | _____ | _____ | _____ |

Strongly Disagree Neither Agree Strongly
Disagree Agree/Disagree Agree

3. I would imagine that people would learn to use the wheelchair training simulator very quickly.

_____ | _____ | _____ | _____ | _____ |

Strongly Disagree Neither Agree Strongly
Disagree Agree/Disagree Agree

4. I found the system very difficult to use.

_____ | _____ | _____ | _____ | _____ |

Strongly Disagree Neither Agree Strongly
Disagree Agree/Disagree Agree

5. I needed to learn a lot of things before I could get going with this system.

_____ | _____ | _____ | _____ | _____ |

Strongly Disagree Neither Agree Strongly
Disagree Agree/Disagree Agree

Subject ID: ___ Date: ___/___/___ Gender: F M Environment: VR Screen

6. In the computer generated world I had a sense of "being there".

7. I had a sense of acting in the virtual space, rather than operating something from outside.

A horizontal scale with five vertical tick marks. Below the scale, the following labels are positioned: 'Strongly Disagree' under the first tick, 'Disagree' under the second, 'Neither' under the third, 'Agree' under the fourth, and 'Strongly Agree' under the fifth. The word 'Agree/Disagree' is centered below the 'Neither' and 'Agree' labels.

8. I felt present in the virtual space.

Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
Agree/Disagree				

9. How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?

A horizontal scale consisting of five vertical tick marks and four intervals. Below the scale, the first interval is labeled "Extremely Aware", the second is "Moderate", and the third is "Not aware at all".

10. How much did your experience in the virtual environment seem consistent with your real world experience ?

Subject ID: ___ Date: ___ / ___ / ___ Gender: F M Environment: VR Screen

Questionnaire – Part 2

The evaluation you are about to perform is a technique that has been developed by Nasa to assess the relative importance of six factors in determining how much workload you experienced while performing a task that you recently completed. These six factors are defined below on this page.

Read through them to make sure you understand what each factor means. If you have any questions, please ask your administrator.

Workload factors	Definition
Mental Demand Level (low/high)	How much mental and perceptual activity was required (for example, thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, forgiving or exacting?
Physical Demand Level (low/high)	How much physical activity was required (for example, pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal Demand Level (low/high)	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance Level(good/poor)	How successful do you think you were in accomplish the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplish these goals?
Effort Level (low/high)	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration Level (low/high)	How insecure, discouraged, irritated, stressed , and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

Subject ID: ___ Date: ___ / ___ / ___ Gender: F M Environment: VR Screen

For each pair, choose the factor that was more important to your experience of the workload in the task that you recently performed:		
1	<input type="checkbox"/> Temporal Demand	<input type="checkbox"/> Mental Demand
2	<input type="checkbox"/> Performance	<input type="checkbox"/> Mental Demand
3	<input type="checkbox"/> Mental Demand	<input type="checkbox"/> Effort
4	<input type="checkbox"/> Temporal Demand	<input type="checkbox"/> Effort
5	<input type="checkbox"/> Physical Demand	<input type="checkbox"/> Performance
6	<input type="checkbox"/> Performance	<input type="checkbox"/> Temporal Demand
7	<input type="checkbox"/> Effort	<input type="checkbox"/> Physical Demand
8	<input type="checkbox"/> Mental Demand	<input type="checkbox"/> Physical Demand
9	<input type="checkbox"/> Performance	<input type="checkbox"/> Frustration
10	<input type="checkbox"/> Effort	<input type="checkbox"/> Performance
11	<input type="checkbox"/> Frustration	<input type="checkbox"/> Effort
12	<input type="checkbox"/> Frustration	<input type="checkbox"/> Mental Demand
13	<input type="checkbox"/> Physical Demand	<input type="checkbox"/> Temporal Demand
14	<input type="checkbox"/> Physical Demand	<input type="checkbox"/> Frustration
15	<input type="checkbox"/> Temporal Demand	<input type="checkbox"/> Frustration

Subject ID: ___ Date: ___/___/___ Gender: F M Environment: VR Screen

You will now be presented with a Series of rating scales.

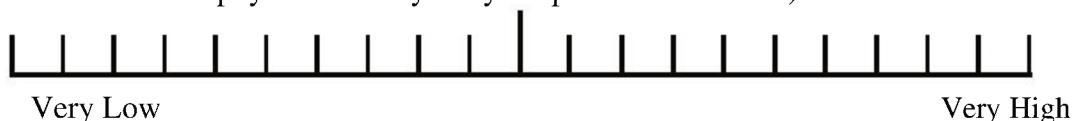
For each of the six scales, evaluate the task you recently performed by cross on the scale's location that matches your experience. Each line has two endpoint that describe the scale.

Consider your responses carefully in distinguishing among the different task conditions, and consider each individually.

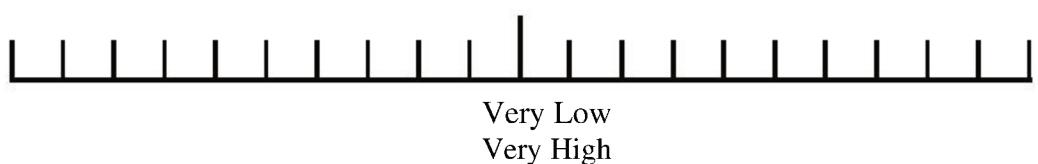
1. Mental Demand (How mentally demanding was the task?/ How much mental and perceptual activity did you spend for this task?)



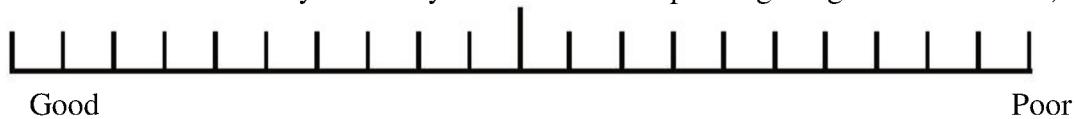
2. Physical Demand (How physically demanding was the task?/ How much physical activity did you spend for this task?)



3. Temporal Demand (How hurried or rushed was the pace of the task?/ How much time pressure did you feel in order to complete this task?)



4. Performance (How successful were you in accomplishing what you were asked to do?/ How successful do you think you were in accomplishing the goals of the task?)



5. Effort (How hard did you have to work to accomplish your level of performance?)



6. Frustration (How insecure, discouraged, irritated, stressed, and annoyed were you during this task?)

