

Marcelo Edmar Melo Nunes

Desenvolvimento de um Jogo Sérió para Apoio ao Tratamento de Transtornos de Ansiedade e Depressão

Uberlândia, MG

2026

Marcelo Edmar Melo Nunes

Desenvolvimento de um Jogo S rio para Apoio ao Tratamento de Transtornos de Ansiedade e Depress o

Trabalho de Conclus o de Curso da Engenharia Biom dica da Universidade Federal de Uberl ndia – UFU – Campus Santa M nica, como requisito para a obten o do t tulo de Gradua o em Engenharia Biom dica.

Universidade Federal de Uberl ndia - UFU
Faculdade de Engenharia El trica - FEELT

Orientador Prof. Dr. Eduardo L zaro Martins Naves

Uberl ndia, MG

2026

Agradecimentos

Primeiramente, agradeço à minha família, por ser minha base sólida e referência. Em especial aos meus pais, Marluce e Edmar: o amor, o apoio incondicional e os sacrifícios que fizeram foram fundamentais para que eu pudesse chegar até aqui e concluir a minha graduação em Engenharia Biomédica. Agradeço aos meus irmãos, que viram minha vida se resumir a visitas rápidas em datas comemorativas e curtos períodos de férias. Morar em outra cidade por todos estes anos foi o maior desafio da minha jornada. Em muitos momentos de solidão, cheguei a questionar se este diploma valia o preço de passar tanto tempo longe da companhia de vocês. Hoje, sei que cada segundo de ausência foi por um futuro melhor, mas o meu maior prêmio sempre será voltar para os braços de vocês. Obrigado por me esperarem e por me amarem, mesmo de longe.

À minha namorada, Giulia, expresso minha mais profunda gratidão. Obrigado por toda a paciência durante os momentos de estresse, pelo companheirismo diário e por me incentivar não apenas nos estudos, mas também nos meus hobbies e momentos de descontração, tornando essa jornada acadêmica, na medida do possível, muito mais leve e equilibrada. Agradeço por me acolher, muitas vezes na sua própria casa, e por me permitir participar da sua família e ser tão incluído e bem tratado que me sinto como um de vocês.

No âmbito acadêmico, agradeço à Universidade Federal de Uberlândia (UFU), pelos anos de ensino prestados. Estendo minha gratidão ao meu orientador, Professor Dr. Eduardo Lázaro Martins Naves, por todo o suporte, oportunidades, pela base científica transmitida e por me guiar com sabedoria ao longo deste projeto. Aos meus colegas e amigos do Núcleo de Tecnologias Assistivas (NTA), o meu muito obrigado pelo ambiente colaborativo. Um agradecimento especial aos colegas Andressa Rastrelo, Laura Foroni, Danilo Santos e Júlia Tannús, pelas trocas de ideias, pelas discussões valiosas na estruturação de diversos trabalhos, pelo auxílio com o aprimoramento da minha escrita e pela parceria de sempre.

No contexto profissional e prático deste desenvolvimento, gostaria de expressar minha profunda gratidão à Reabnet Tecnologia, onde realizei meu estágio, pela oportunidade de atuar em projetos que unem inovação e propósito. Agradeço à Isabela Alves Marques, pela liderança inspiradora, pela confiança depositada no meu trabalho e por todo o suporte.

Agradeço imensamente ao meu supervisor direto, Felipe Roque Martins, e a toda a equipe de desenvolvimento de jogos. Sua orientação técnica, paciência e disposição em compartilhar conhecimentos foram essenciais para o meu aprendizado, especialmente nos desafios de programação, design e arquitetura de software enfrentados neste protótipo.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), agra-

deço pelo incentivo e apoio financeiro, indispensáveis para a manutenção da minha dedicação à pesquisa e ao desenvolvimento tecnológico.

Por fim, agradeço a todos os amigos, professores e voluntários que contribuíram direta ou indiretamente para a conclusão desta etapa. Levo comigo não apenas a evolução técnica em desenvolvimento ou melhorias na minha escrita científica, mas também os espólios adquiridos durante todo esse período de graduação, que incluem lições valiosas sobre empatia, coragem, trabalho em equipe e o impacto transformador que a tecnologia pode ter na vida das pessoas.

Resumo

Os transtornos de ansiedade e depressão representam um dos maiores desafios globais de saúde pública, frequentemente esbarrando em barreiras de acesso e altas taxas de evasão no tratamento tradicional. Como alternativa complementar, os jogos sérios emergem como ferramentas promissoras. O objetivo deste trabalho foi desenvolver e avaliar preliminarmente a usabilidade e a imersão de um jogo sério de baixo custo e alta acessibilidade, integrado a tecnologias de Inteligência Artificial, atuando como uma ferramenta de apoio à regulação emocional. O protótipo foi desenvolvido na plataforma Unity (WebGL), focado em interações sem contato físico (contactless) via webcam e microfone utilizando as bibliotecas MediaPipe e Web Speech API. A intervenção foi estruturada em três mecânicas lúdicas baseadas em evidências: respiração diafragmática, afirmações positivas e organização ambiental. O estudo adotou um protocolo de avaliação preliminar com uma amostra não clínica de 10 jovens adultos saudáveis. Para a coleta de dados, utilizaram-se os instrumentos validados System Usability Scale (SUS) e a subescala de imersão do questionário EGameFlow. Os resultados quantitativos indicaram uma usabilidade classificada como excelente, com o sistema alcançando uma pontuação final composta de 85,8 no SUS. No EGameFlow, registrou-se uma média geral de 4,93 (em uma escala de 1 a 7), evidenciando um nível médio de engajamento imersivo, com destaque para a alta absorção cognitiva e o esquecimento de preocupações diárias ($M = 5,70$). As limitações do estudo incluem o tamanho reduzido e o perfil não clínico da amostra, além da sensibilidade da ferramenta a variações de iluminação e dependência de hardware adequado. Conclui-se que o protótipo apresenta viabilidade técnica e aceitabilidade como ferramenta de suporte acessível. Contudo, os achados refletem estritamente a qualidade de engenharia da interface, não permitindo inferir eficácia terapêutica, o que demandará futuros ensaios clínicos controlados com pacientes diagnosticados.

Palavras-chave: Jogos Sérios, Saúde Mental, Inteligência Artificial, Usabilidade.

Abstract

Anxiety and depression disorders represent one of the major global public health challenges, frequently facing access barriers and high dropout rates in traditional treatment. As a complementary alternative, serious games emerge as promising tools. The objective of this study was to develop and preliminarily evaluate the usability and immersion of a low-cost, highly accessible serious game, integrated with Artificial Intelligence technologies, acting as a complementary support tool for emotional regulation. The prototype was developed on the Unity engine (WebGL), focusing on contactless interactions via webcam and microphone using the MediaPipe and Web Speech API libraries. The intervention was structured around three evidence-based ludic mechanics: diaphragmatic breathing, positive affirmations, and environmental organization. The study adopted a preliminary evaluation protocol with a non-clinical sample of 10 healthy young adults. Data collection employed the validated System Usability Scale (SUS) and the immersion subscale of the EGameFlow questionnaire. Quantitative results indicated excellent usability, with the system achieving a final composite SUS score of 85.8. The EGameFlow registered an overall average of 4.93 (on a 1 to 7 scale), demonstrating a medium level of immersive engagement, highlighting high cognitive absorption and distraction from daily worries ($M = 5.70$). Study limitations include the small size and non-clinical profile of the sample, as well as the tool's sensitivity to lighting variations and reliance on adequate hardware. It is concluded that the prototype presents technical viability and acceptability as an accessible support tool. However, the findings strictly reflect the engineering quality of the interface and do not allow for the inference of therapeutic efficacy, which will require future randomized controlled trials with diagnosed patients.

Keywords: Serious Games, Mental Health, Artificial Intelligence, Usability Evaluation.

Lista de ilustrações

Figura 1 – Todos os 33 landmarks rastreados pelo MediaPipe.	23
Figura 2 – Interface do Menu Inicial.	35
Figura 3 – Interface de Autorrelato.	35
Figura 4 – Primeira Fase: Respiração.	36
Figura 5 – Segunda Fase: Afirmações.	37
Figura 6 – Terceira Fase: Organização.	37
Figura 7 – Sistema de Tutorial.	39

Lista de tabelas

Tabela 1 – Características Demográficas dos Participantes (Tamanho Amostral = 10)	30
Tabela 2 – Estatísticas Descritivas dos Itens do Questionário SUS (Tamanho Amostral = 10)	40
Tabela 3 – Estatísticas Descritivas da Subescala de Imersão do EGameFlow (Tamanho Amostral = 10)	41

Lista de abreviaturas e siglas

3D	Three-dimensional
API	Application Programming Interface
CAAE	Certificado de Apresentação de Apreciação Ética
CEP	Comitê de Ética em Pesquisa
DP	Desvio Padrão
FEELT	Faculdade de Engenharia Elétrica
IA	Inteligência Artificial
INSS	Instituto Nacional do Seguro Social
LGPD	Lei Geral de Proteção de Dados
M	Média
OMS	Organização Mundial da Saúde
PANAS	Escala de Afeto Positivo e Negativo
PHQ-9	Questionário de Saúde do Paciente
RGB	Red, Green, Blue
SUDS	Escala de Unidades Subjetivas de Angústia
SUS	System Usability Scale
TCC	Terapia Cognitivo-Comportamental
TCLE	Termo de Consentimento Livre e Esclarecido
UFTM	Universidade Federal do Triângulo Mineiro
UFU	Universidade Federal de Uberlândia
UI	Interface de Usuário
VR	Virtual Reality
WebGL	Web Graphics Library

Sumário

1	INTRODUÇÃO	11
1.1	Justificativa	12
1.2	Objetivo	13
1.2.1	Objetivos Específicos	13
1.3	Produção Científica e Contribuições	14
2	REFERENCIAL TEÓRICO	15
2.1	Transtornos de Ansiedade e Depressão	15
2.2	Jogos Sérios e a Reabilitação	16
2.3	Fundamentação das Mecânicas Lúdicas	18
2.4	Arquitetura e Integração de Tecnologias Digitais	18
2.5	Avaliação de Usabilidade e Imersão	20
3	METODOLOGIA	22
3.1	Idealização e Estruturação Terapêutica	22
3.2	Ferramentas de Desenvolvimento e Arquitetura	22
3.2.1	Menu Principal e Interface de Autorrelato	23
3.2.2	Fase 1: Exercício de Respiração Quadrada	25
3.2.3	Fase 2: Exercício de Afirmações Positivas	27
3.2.4	Fase 3: Organização Ambiental e Metáfora Comportamental	28
3.3	Protocolo de Avaliação Preliminar	29
3.3.1	Participantes e Aspectos Éticos	29
3.3.2	Protocolo Experimental e Equipamentos	30
3.3.3	Instrumentos de Avaliação	31
3.3.4	Análise de Dados	32
3.4	Privacidade, Segurança de Dados e LGPD	32
3.5	Assistência de Inteligência Artificial na Redação	33
4	RESULTADOS E DISCUSSÕES	34
4.1	Interface de Interação com o Usuário (UI)	34
4.2	Interface de Jogabilidade	36
4.3	Avaliação de Usabilidade (SUS)	39
4.4	Avaliação de Imersão (EGameFlow)	40
4.4.1	Feedback Qualitativo e Percepções dos Usuários	42
5	CONCLUSÃO	44

5.1	Limitações do Estudo	45
5.2	Implicações e Direções Futuras	46
	REFERÊNCIAS BIBLIOGRÁFICAS	48
6	PÁGINA DO ARTIGO (IEEE SEGAH 2026)	52

1 Introdução

No cenário global contemporâneo, os transtornos mentais representam um dos maiores desafios de saúde pública global do século XXI. Segundo a Organização Mundial da Saúde (OMS), aproximadamente uma em cada oito pessoas no mundo, cerca de 970 milhões de indivíduos, vive com alguma condição de saúde mental, sendo os transtornos de ansiedade e depressão os mais prevalentes. Somente a depressão afeta mais de 280 milhões de pessoas mundialmente, enquanto os transtornos de ansiedade impactam cerca de 301 milhões. Este quadro foi drasticamente agravado pela pandemia de COVID-19, que provocou um aumento de 25% na prevalência global dessas condições logo em seu primeiro ano ([World Health Organization, 2022](#)).

A incidência dessas condições apresenta variações importantes de acordo com a faixa etária e o contexto socioeconômico. Os jovens adultos com idade entre 18 e 29 anos são os que apresentam as maiores taxas de ansiedade, já a depressão aparece significativamente tanto para os jovens adultos quanto para grupos de idade mais avançada ([KESSLER et al., 2005](#)). Além disso, observa-se uma grave desigualdade no acesso aos cuidados em saúde mental. Em países de baixa e média renda, estima-se que mais de 75% das pessoas com transtornos mentais não recebem nenhum tipo de intervenção ou tratamento ([World Health Organization, 2022](#)).

A farmacoterapia, a psicoterapia, como a terapia cognitivo-comportamental (TCC), ou a combinação de ambas, são abordagens tradicionais de tratamento para ansiedade e depressão. Contudo, a aplicação desses métodos convencionais enfrenta desafios significativos ([CUIJPERS et al., 2013](#)). Entre as principais barreiras destacam-se as altas taxas de abandono do tratamento, que variam de 20% a 60%, a acessibilidade limitada devido a custos, as barreiras geográficas e o estigma social frequentemente associado à busca por cuidados em saúde mental ([SWIFT; GREENBERG, 2012; ANDRADE et al., 2014](#)). Visando contornar tais desafios, a gamificação e os jogos sérios emergiram como estratégias complementares promissoras, capazes de aumentar a motivação, melhorar a adesão aos protocolos terapêuticos e fornecer um ambiente livre de julgamentos ([FLEMING et al., 2017; LAU et al., 2017](#)).

Revisões sistemáticas recentes demonstram que intervenções de saúde mental gamificadas apresentam melhorias significativas nos sintomas de ansiedade e depressão, com efeitos comparáveis aos tratamentos tradicionais, ao mesmo tempo em que elas alcançam maiores taxas de engajamento ([LI; THENG; FOO, 2014; SCHOLTEN; GRANIC, 2019](#)). Os jogos sérios ultrapassam o mero entretenimento ao aliar elementos lúdicos e narrativos com objetivos reais de saúde, educação ou treinamento profissional, criando

ambientes interativos onde o usuário, por meio dos desafios constantes e feedback, pode desenvolver habilidades cognitivas, emocionais e comportamentais (FORMIGA et al., 2025; CAMILO et al., 2016).

Nesse contexto, este trabalho propõe o desenvolvimento de um jogo sério voltado ao apoio à regulação emocional de jovens adultos com sintomas leves a moderados de ansiedade e depressão. O jogo foi desenvolvido com foco na alta acessibilidade e usabilidade, não necessitando de sensores ou elementos externos a um computador comum para seu uso. Dessa forma, o usuário poderá interagir com a interface de forma simples, garantindo uma experiência interativa, segura e eficiente. Com uma temática envolvente e estimulante, integrando tecnologias de Inteligência Artificial para reconhecimento de gestos (*MediaPipe*) e processamento de fala (*Web Speech API*), o objetivo foi aliar mecânicas lúdicas a técnicas de relaxamento baseadas em evidências, incluindo respiração diafragmática e afirmações positivas, para que seja possível oferecer uma ferramenta complementar de baixo custo e alto potencial de engajamento.

Além do desenvolvimento técnico, é imprescindível garantir que a ferramenta proposta também seja bem recebida e de fácil utilização. Portanto, este trabalho também engloba uma etapa de avaliação preliminar com usuários, com o intuito de mensurar a usabilidade do sistema e o grau de imersão proporcionado pelas mecânicas lúdicas. Para tal, foram aplicados instrumentos validados na literatura, como a *System Usability Scale* (SUS) (BROOKE, 1996) e a subescala de imersão do *EGameFlow* (FU; SU; YU, 2009), visando avaliar preliminarmente a usabilidade e a imersão proporcionadas pelo protótipo, sem investigar, nesta etapa, sua eficácia clínica.

1.1 Justificativa

A justificativa para o desenvolvimento desta solução é baseada principalmente na urgente necessidade de democratizar e expandir o acesso a intervenções complementares de apoio à saúde mental. Como evidenciado anteriormente, a busca por métodos convencionais frequentemente esbarra em desafios socioeconômicos e altas taxas de evasão (SWIFT; GREENBERG, 2012; ANDRADE et al., 2014). Os jogos sérios têm o potencial de atenuar essas barreiras ao transformar práticas recomendadas, como a respiração diafragmática, a reestruturação cognitiva e a atenção plena, também conhecida como *mindfulness*, em mecânicas lúdicas, interativas e recompensadoras para o usuário (FLEMING et al., 2017; LAU et al., 2017).

Adicionalmente, a integração de Inteligência Artificial nativa em navegadores *web* traz um diferencial estratégico: elimina a necessidade de instalação de *softwares* pesados ou compra de equipamentos caros dedicados, permitindo interações naturais por meio da voz e do movimento corporal utilizando apenas a *webcam* e o microfone do próprio dispositivo

utilizado. Isso não apenas reduz barreiras tecnológicas, mas também transforma o jogo em uma ferramenta de interação multimodal, gerando dados objetivos e métricas quantificáveis, como o tempo de engajamento, que poderão, futuramente, auxiliar psicólogos e psiquiatras no acompanhamento clínico, promovendo uma ponte contínua entre as sessões de terapia e o cotidiano do usuário.

1.2 Objetivo

O objetivo geral deste trabalho foi desenvolver e avaliar preliminarmente a usabilidade e a imersão de um jogo sério de baixo custo e alta acessibilidade, integrado a tecnologias de Inteligência Artificial, utilizando visão computacional e reconhecimento de voz, como ferramenta complementar de apoio e promoção do bem-estar para indivíduos com sintomas de ansiedade e depressão. O jogo foi projetado com a incorporação lúdica de técnicas de relaxamento baseadas em evidências científicas, incluindo respiração diafragmática, afirmações positivas e estratégias de regulação emocional.

1.2.1 Objetivos Específicos

- Realizar um levantamento pontual de requisitos e práticas de relaxamento por meio de consultas e orientação com profissionais da área de psicologia;
- Desenvolver o ambiente virtual interativo e as lógicas de jogabilidade na plataforma Unity, com otimização arquitetural para execução nativa em navegadores (WebGL);
- Integrar e implementar as bibliotecas MediaPipe (para rastreamento corporal e reconhecimento de gestos) e Web Speech API (para comandos e síntese de voz);
- Implementar mecânicas lúdicas de apoio, incluindo dinâmicas de respiração quadrada, afirmações positivas, organização de ambiente e *mindfulness*;
- Estruturar, em caráter de prova de conceito, a arquitetura de *software* para a futura coleta e transmissão de dados do jogo (como escalas de autorrelato e capturas de tela), estabelecendo a base técnica para posterior acompanhamento por profissionais de saúde;
- Realizar testes preliminares com uma amostra não clínica para a avaliação da usabilidade do sistema (utilizando a escala SUS) e do grau de imersão do usuário (utilizando a subescala *EGameFlow*).

1.3 Produção Científica e Contribuições

Os resultados parciais e a arquitetura tecnológica desenvolvida no escopo deste trabalho de conclusão de curso renderam a submissão e aprovação de um artigo científico internacional. O trabalho intitulado *"Beacon of Hope": Usability and Immersion Evaluation of an AI-Enhanced Serious Game for Anxiety and Depression* foi apresentado e aceito para publicação nos anais da *14th International Conference on Serious Games and Applications for Health* (IEEE SeGAH 2026), evento ocorrido no Rio de Janeiro. No momento da finalização deste documento, o artigo encontra-se aceito, apresentado e aguardando publicação oficial pela base de dados do IEEE.

2 Referencial Teórico

2.1 Transtornos de Ansiedade e Depressão

A ansiedade e a depressão figuram entre as condições psiquiátricas mais comuns e incapacitantes da atualidade. A depressão é caracterizada por uma tristeza persistente, perda de interesse em atividades cotidianas, alterações no sono e no apetite, e uma forte sensação de desesperança. Já os transtornos de ansiedade englobam um conjunto de condições marcadas por preocupação excessiva, medo irracional e hiperatividade autonômica, que frequentemente paralisam o indivíduo perante situações normais do dia a dia. Segundo o relatório da Organização Mundial da Saúde (OMS), essas condições afetam, em conjunto, quase um bilhão de pessoas globalmente ([World Health Organization, 2022](#)).

O impacto socioeconômico e laboral dessas condições é profundo e quantificável. No contexto brasileiro, dados oficiais do Ministério da Previdência Social e do Instituto Nacional do Seguro Social (INSS) revelam uma crise aguda de saúde mental no ambiente de trabalho. Apenas no ano de 2024, o Brasil registrou mais de 472 mil concessões de benefícios por incapacidade temporária motivados por transtornos mentais e comportamentais. Esse volume representou um aumento expressivo de aproximadamente 68% em relação ao ano de 2023, consolidando o maior índice da série histórica da última década ([Ministério da Previdência Social, 2024](#)). Liderando as causas dessas licenças médicas encontram-se os transtornos ansiosos e os episódios depressivos, evidenciando o quão limitantes essas condições são para a funcionalidade diária, produtividade e qualidade de vida dos indivíduos.

Para o ano de 2025, a Previdência Social concedeu 546.254 benefícios por transtornos mentais e comportamentais, o que representou um crescimento de 15,66% em comparação a 2024. Estimativas indicam que a tendência é que o número continue crescendo em 2026. ([Ministério da Previdência Social, 2026](#))

Apesar da alta prevalência e da gravidade dos impactos, o enfrentamento dessas condições esbarra no que a literatura científica denomina como "lacuna de tratamento". A lacuna de tratamento representa a diferença absoluta entre o número de indivíduos afetados por um transtorno e a proporção que efetivamente recebe cuidados clínicos adequados ([KOHN et al., 2004](#)). O protocolo clínico padrão geralmente envolve a psicoterapia, comumente inclinada para a Terapia Cognitivo-Comportamental (TCC), associadas, quando necessário, à farmacoterapia. No entanto, estudos globais indicam que cerca de 72,4% das pessoas com transtornos de ansiedade não recebem qualquer tipo de tratamento

(ALONSO et al., 2018). Em relação à depressão, as lacunas na cobertura dos serviços são amplas mesmo em países de alta renda, onde apenas cerca de um terço das pessoas com transtornos depressivos maiores recebe atendimento formal em saúde mental. Em países de baixa e média renda, como o Brasil, a estimativa é de que mais de 75% das pessoas com transtornos mentais em geral permaneçam desassistidas (World Health Organization, 2022).

Fatores como a escassez de profissionais qualificados no sistema público, os altos custos das terapias na rede privada e o forte estigma social ligado ao diagnóstico psiquiátrico atuam como barreiras primárias. Mesmo entre os pacientes que conseguem iniciar o acompanhamento, o processo de reabilitação exige tempo, consistência e alto engajamento, o que frequentemente resulta em taxas de abandono que variam entre 20% e 60% nos tratamentos tradicionais (SWIFT; GREENBERG, 2012; ANDRADE et al., 2014).

Diante dessa lacuna assistencial e da urgência por intervenções escaláveis e contínuas, a tecnologia tem assumido um papel de protagonismo. O desenvolvimento de ferramentas digitais complementares busca não apenas reduzir as barreiras de acesso ao suporte psicológico, mas também oferecer ao usuário um acesso facilitado, acessível financeiramente, engajante e que ele possa utilizar no conforto e na privacidade do seu próprio ambiente, contornando os principais estigmas enfrentados.

2.2 Jogos Sérios e a Reabilitação

Os jogos sérios consistem em aplicações digitais desenvolvidas com propósitos que vão além de apenas o entretenimento puro. Enquanto a gamificação aplica elementos de jogos, como pontuações e medalhas, em contextos não lúdicos, e a telerreabilitação foca na prestação de serviços à distância, os jogos sérios são *softwares* autônomos que integram mecânicas interativas a objetivos terapêuticos, educacionais ou de treinamento. No contexto da saúde mental, essas ferramentas digitais de intervenção são comumente empregadas para guiar a execução de dinâmicas clínicas, fornecer *feedback* em tempo real e aumentar a motivação dos participantes. Além do forte componente lúdico, eles oferecem a vantagem de padronizar intervenções e registrar métricas objetivas de desempenho, o que apresenta um enorme potencial para reduzir os custos operacionais e ampliar o acesso ao cuidado em saúde como um formato de apoio (LOHSE et al., 2013).

A literatura científica já documenta algumas iniciativas diversas que utilizam dos jogos sérios para o tratamento de transtornos psiquiátricos. Um exemplo notório é o *SPARX*, um jogo imersivo de fantasia desenvolvido especificamente para entregar intervenções baseadas na Terapia Cognitivo-Comportamental (TCC) a adolescentes com depressão, demonstrando eficácia clínica na redução de sintomas comparável ao tratamento presencial tradicional (MERRY et al., 2012). Outras abordagens incluem o *MindLight*,

que utiliza mecânicas de *neurofeedback* para auxiliar crianças e jovens na regulação da ansiedade (SCHONEVELD et al., 2016), e o *SuperBetter*, voltado para o aumento da resiliência psicológica e o combate a sintomas depressivos por meio de desafios gamificados no cotidiano (ROEPKE et al., 2015).

Contudo, mesmo com alguns dados que atestem e defendam a eficácia, a implementação tecnológica estrutural dessas soluções ainda enfrenta barreiras significativas. Na literatura atual, são predominantes três principais formatos de jogos sérios em reabilitação. O primeiro formato baseia-se em soluções que exigem sensores de movimento dedicados, como o *Kinect*, ou equipamentos vestíveis para mapear gestos e respostas fisiológicas. Embora estes promovam melhoras funcionais, exigem a aquisição de *hardware* específico, o que acaba limitando sua escalabilidade, aplicação domiciliar e acessibilidade (FLEMING et al., 2017).

O segundo formato envolve os ambientes de Realidade Virtual (VR) totalmente imersivos, acessados por meio de óculos específicos para essa funcionalidade. Essas plataformas permitem um alto grau de presença e são eficientes para promover o isolamento sensorial e a distração de estímulos ansiogênicos, o que acaba oferecendo uma progressão de desafios recompensadora. Contudo, o custo de implementação e aquisição desses equipamentos é significativamente alto, o que compromete muito a democratização do acesso. Adicionalmente, o uso prolongado de VR pode induzir efeitos adversos indesejados, como já relatado em testes com estes equipamentos. Um desses efeitos é a cinetose digital, que também é conhecida como *cybersickness*, gerando desconforto visual e risco de quedas ou desorientação (LI et al., 2023; PORCINO et al., 2021)

Muito devido às lacunas de acessibilidade e custo deixadas pelas abordagens anteriores, o terceiro formato, que é focado em jogos digitais com visão computacional nativa, desponta como sendo um dos caminhos mais promissores. Esse modelo dispensa o uso de sensores externos, e acaba operando exclusivamente por meio da câmera embutida de computadores, tablets ou smartphones. A tecnologia processa algoritmos de reconhecimento corporal e facial em tempo real, garantindo *feedback* imediato das ações do jogador de forma totalmente sem contato.

Estudos recentes ilustram a viabilidade prática dessa abordagem. Trabalhos como o de Zhang et al. (2024) e Gonçalves et al. (2021) relataram o desenvolvimento de sistemas que integram o *MediaPipe* para analisar a postura e guiar exercícios de reabilitação física usando apenas uma câmera comum (ZHANG et al., 2024; GONÇALVES et al., 2021). No entanto, a literatura atual concentra o uso dessas ferramentas visuais principalmente na reabilitação motora e musculoesquelética. A transposição e adaptação dessa arquitetura para o escopo da saúde mental, realizando a união do rastreamento corporal e facial a dinâmicas de regulação emocional, processamento de voz e à atenção plena, representa um campo ainda pouco explorado.

Um exemplo prático de algoritmo capaz de processar essas posturas em tempo real é o *MediaPipe*, um projeto de código aberto desenvolvido pelo Google, amplamente utilizado para visão computacional rápida, como o rastreamento facial e de mãos (Google Developers, 2026). Revisões recentes apontam que eliminar a dependência de *hardwares* periféricos caros é a estratégia mais eficaz para romper barreiras socioeconômicas, ampliando o alcance e promovendo o engajamento autônomo (ZHENG et al., 2021; ACHARYA et al., 2025). É exatamente nessa lacuna estrutural e metodológica que o presente projeto se insere, buscando unir a fundamentação clínica de intervenções cognitivas à acessibilidade e escalabilidade da visão computacional moderna, em uma arquitetura sem uso de sensores vestíveis aplicada à saúde mental.

2.3 Fundamentação das Mecânicas Lúdicas

A transposição de protocolos clínicos para o ambiente virtual exige o respaldo de evidências científicas e a orientação técnica adequada. Nesse sentido, todas as mecânicas lúdicas implementadas neste projeto foram desenhadas e refinadas sob a consultoria e validação de profissionais de psicologia atuantes, garantindo que a gamificação respeitasse os princípios fundamentais das terapias base.

A técnica da "respiração quadrada", ou respiração cadenciada estruturada, por exemplo, atua na modulação do sistema nervoso autônomo. Práticas de respiração com controle inspiratório e expiratório demonstraram eficácia na redução da excitação fisiológica e na melhora do estado de humor em curtos períodos de aplicação (BALBAN et al., 2023).

Adicionalmente, a mecânica de organização ambiental é baseada na forte correlação entre espaços físicos desorganizados e a sobrecarga cognitiva. Estudos em psicologia ambiental comprovam que o acúmulo visual desordenado aumenta significativamente os níveis de estresse e o esgotamento emocional do indivíduo (ROSTER; FERRARI; JURKAT, 2016). Ao exigir que o usuário ordene o espaço virtual, o jogo fornece um exercício prático de foco e de retomada de controle, reduzindo a hiperestimulação. Por fim, o uso vocal de afirmações positivas atua como uma técnica derivada da Teoria da Autoafirmação. Verbalizar valores e afirmações otimistas tem se mostrado uma intervenção psicológica eficaz para romper ciclos de ruminação, mitigar o estresse defensivo e promover a reestruturação cognitiva, elementos cruciais no suporte a indivíduos com sintomas ansiosos e depressivos (COHEN; SHERMAN, 2014).

2.4 Arquitetura e Integração de Tecnologias Digitais

O motor gráfico conhecido como Unity é consolidado como uma das principais plataformas para o desenvolvimento de jogos sérios e aplicações imersivas na área da saúde.

Essa proeminência decorre de sua capacidade multiplataforma, rápida prototipagem e de um ecossistema robusto para a construção de Interfaces de Usuário (UI) e sistemas de animação complexos, além de contar com amplo suporte a pacotes externos, *assets*, entre outros. Na engenharia desses softwares quando aplicados à saúde, essas características facilitam a tradução direta de objetivos clínicos e terapêuticos em mecânicas lúdicas de jogo. Em algumas intervenções voltadas à saúde mental encontradas, o Unity destaca-se pela simplicidade e facilidade da gestão de cenas, boa integração e adequação dos conteúdos, e a capacidade de implementação de telemetria de uso, que são alguns dos elementos essenciais para garantir a escalabilidade e a avaliação de eficácia do sistema (RAHMANI-KATIGARI; MOHAMMADIAN; SHAHMORADI, 2023; RAHANI; VARD; NAJAFI, 2018). Ademais, a capacidade de exportação da plataforma em formatos baseados em WebGL viabiliza que esses sistemas complexos rodem de maneira nativa e leve em navegadores web convencionais, isentando o usuário da necessidade de efetuar instalações locais de softwares pesados.

Para promover a interatividade do participante com o ambiente virtual de forma totalmente sem contato, sistemas modernos de saúde digital têm integrado algoritmos de visão computacional on-device. O *MediaPipe* é um *framework* de percepção multimodal de código aberto que fornece pipelines otimizados de aprendizado de máquina com baixa latência para diferentes plataformas (Google Developers, 2026). O uso de seus módulos de rastreamento de poses e marcos corporais, chamados de *landmarks*, elimina a dependência de sensores periféricos proprietários de alto custo. A extração em tempo real das coordenadas articulares permite mapear não apenas movimentos corporais amplos, mas também gestos finos de membros superiores, dinâmicas das mãos e do rosto (Google AI Edge, 2025). Estudos demonstram que, a partir do pós-processamento e normalização desses marcos de contorno, torna-se viável classificar padrões de gestos com alta fidelidade e estabilidade operacional, permitindo que os landmarks extraídos atuem diretamente como mecanismos eficientes de controle de jogo e navegação em menus interativos (MENG et al., 2024; AHIRE et al., 2025). O monitoramento contínuo desses marcadores cinemáticos abre precedentes importantes para a análise automatizada de gestos corporais associados a estados emocionais (NOROOZI et al., 2018).

Complementarmente às dinâmicas cinemáticas, o processamento de linguagem natural e as interações baseadas em voz têm ganhado espaço como ferramentas de interação multimodal e apoio à regulação cognitiva. A avaliação automatizada e o suporte a transtornos psiquiátricos por meio da fala encontram respaldo na literatura, que aponta o processamento computacional de áudio como um método promissor, seguro e não invasivo para monitorar alterações e engajar o participante em tarefas terapêuticas baseadas em respostas verbais (LOW et al., 2020). A *Web Speech API*, uma Interface de Programação de Aplicativos voltada para voz, surge como uma solução viável para implementar essas funcionalidades diretamente no navegador, dividindo-se em duas interfaces principais: a síntese de voz, que é responsável por converter texto em áudio, e o reconhecimento de

fala, que atua transcrevendo áudio em texto, transformando as vocalizações do jogador em tempo real (SHIRES; WEN, 2022).

A literatura e a prática de engenharia de software indicam que a união dessas três frentes tecnológicas, motor gráfico *Unity*, visão computacional *MediaPipe* e processamento vocal *Web Speech API*, tem o potencial de estabelecer uma arquitetura híbrida, síncrona e integrada. Por meio de uma compilação *WebGL* unificada, é possível utilizar linguagens de programação para gerenciar o fluxo de dados bidirecional entre o navegador e o motor gráfico. Nesse paradigma, as coordenadas espaciais corporais capturadas por uma câmera convencional e as expressões verbais registradas por microfone podem ser interpretadas instantaneamente pela lógica de um sistema. Teoricamente, essa configuração permite que ações conscientes de um usuário modifiquem cenários e promovam progressões lúdicas em tempo real, consolidando interfaces naturais baseadas em gestos e voz com baixo custo e alta acessibilidade.

2.5 Avaliação de Usabilidade e Imersão

O desenvolvimento de soluções tecnológicas em saúde exige métodos rigorosos de validação para garantir que a interface seja intuitiva e não adicione uma carga cognitiva extra. A usabilidade é um atributo de qualidade que define a facilidade com que interfaces de usuário são utilizadas. Na literatura, a *System Usability Scale* (SUS) é um dos instrumentos mais consolidados para essa avaliação. Composta por dez afirmações avaliadas em uma escala Likert, a SUS gera um escore de 0 a 100 que permite classificar a aceitabilidade e a facilidade de uso do sistema de forma rápida e confiável (BROOKE, 1996).

A aplicação da escala SUS em jogos voltados à saúde permite extrair *insights* valiosos sobre a curva de aprendizado do participante e a clareza dos comandos. Trabalhos que avaliaram a usabilidade em terapias digitais gamificadas, como o de Maskeliūnas *et al.* (2022), evidenciam que identificar falhas na interface por meio dessas métricas é essencial para o aprimoramento do *software* (MASKELIŪNAS *et al.*, 2022). No contexto de saúde mental, escores elevados na SUS indicam uma menor taxa de frustração tecnológica, o que é fundamental para evitar que a complexidade da própria interface acabe se tornando um fator ansiogênico adicional para o usuário.

No contexto específico dos jogos sérios, além da usabilidade, o engajamento do usuário é uma métrica crucial para o sucesso da intervenção. Esse engajamento é frequentemente mensurado pelo grau de imersão, estado no qual o jogador perde a noção do tempo e do mundo real, focando inteiramente na tarefa lúdica. Para quantificar essa experiência, ferramentas como o *EGameFlow* são amplamente adotadas. O instrumento avalia dimensões como concentração, desafio, autonomia e imersão, fornecendo um panorama claro sobre o potencial do jogo em manter o paciente motivado durante as sessões de reabilitação

(FU; SU; YU, 2009). Na prática, a análise dessas dimensões ajuda os desenvolvedores a calibrarem a progressão do jogo. Pesquisas indicam que níveis equilibrados de desafio e autonomia são preditores de adesão a longo prazo, sinalizando se os elementos lúdicos estão efetivamente distraindo o usuário de ruminações negativas e facilitando o estado de *flow* terapêutico (SWEETSER; WYETH, 2005; LOHSE et al., 2013).

3 Metodologia

3.1 Idealização e Estruturação Terapêutica

O processo de desenvolvimento do protótipo do jogo sério para apoio ao tratamento de ansiedade e depressão começou com uma pesquisa sobre estratégias utilizadas como propostas de intervenção terapêutica que estivessem documentadas na literatura. A partir da identificação de algumas técnicas baseadas em evidências para o alívio sintomático, como práticas de relaxamento, meditação guiada e Terapia Cognitivo-Comportamental (TCC), as propostas iniciais contaram com a consultoria técnica informal de dois profissionais de psicologia. Cabe ressaltar que a participação desses profissionais ocorreu estritamente em caráter de orientação teórica e apoio pontual ao desenvolvimento das mecânicas, não configurando um processo de validação clínica estruturada ou oficial. Essa colaboração, aliada ao suporte contínuo da equipe de desenvolvimento de jogos, teve como objetivo apenas assegurar que o núcleo do jogo estivesse alinhado às práticas clínicas documentadas, garantindo uma proposta pertinente antes de sua gamificação e implementação no ambiente digital. Ao mesmo tempo, almejou-se uma proposta executável e condizente com os conhecimentos e a disponibilidade da equipe.

Com a estrutura terapêutica estabelecida, a etapa seguinte consistiu na gamificação das atividades. A temática visual foi cuidadosamente projetada em torno de ambientes naturais, incorporando elementos de paisagens calmas, de cunho relaxante. Essa estética visual foi combinada a um design de áudio focado em elementos biofílicos, incorporando sons de ondas do mar, fluxo suave de água e cantos de pássaros, com o intuito de minimizar a sobrecarga cognitiva e oferecer um espaço digital mais acolhedor. A lógica de jogabilidade foi estruturada para tentar traduzir os exercícios clínicos em interações mais lúdicas e de baixa complexidade. O planejamento do jogo prevê que a progressão ocorra em conjunto com a correta execução das mecânicas propostas, mas sem limitar ou travar o jogador em cenários em que ele encontre dificuldades, buscando transformar a prática terapêutica em uma experiência interativa e atrativa, com o intuito de aumentar a imersão e adesão do usuário à ferramenta proposta.

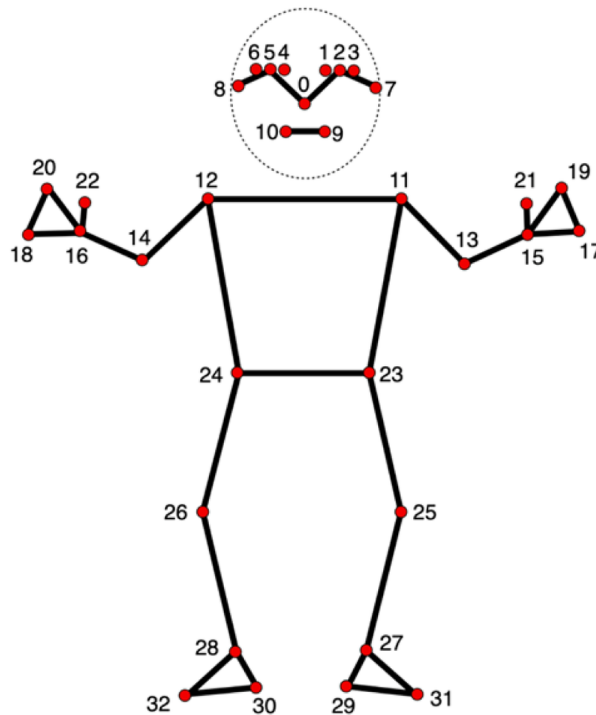
3.2 Ferramentas de Desenvolvimento e Arquitetura

A construção do ambiente virtual interativo e a orquestração do sistema foram realizadas no motor gráfico Unity, utilizando a versão 2022.3.41f1 ([Unity Technologies, 2022](#)) e a linguagem de programação C#. Em consonância com a arquitetura descrita na Seção 2.4, o projeto foi configurado e otimizado para compilação no formato WebGL,

assegurando a execução nativa e direta em navegadores de internet convencionais, sem a necessidade de processos de instalação pelo usuário final.

Para efetivar a interação sem o uso de sensores vestíveis, a comunicação bidirecional entre a *engine* e as APIs do navegador foi estruturada por meio de *scripts* de interoperabilidade codificados em JavaScript e C#. O rastreamento corporal cinemático foi implementado utilizando o *MediaPipe* (Google Developers, 2026). A aplicação extrai, em tempo real, os 33 pontos articulares (*landmarks*) a partir da captação de uma câmera RGB padrão (Fig. 1), convertendo essas coordenadas espaciais em variáveis de controle direto para as dinâmicas espaciais do jogo.

Figura 1 – Todos os 33 landmarks rastreados pelo MediaPipe.



De forma síncrona, a *Web Speech API* (SHIRES; WEN, 2022) foi integrada ao fluxo de processamento para capturar e validar as vocalizações do jogador via microfone. Os dados de áudio capturados foram processados e vinculados à lógica do Unity, estabelecendo essa arquitetura híbrida de gestos e voz como os controladores exclusivos das interações no ambiente virtual para a condução das mecânicas do jogo.

3.2.1 Menu Principal e Interface de Autorrelato

A jornada do usuário no sistema tem início no Menu Principal, onde a interface foi projetada baseada nos mesmos princípios de redução de carga cognitiva e acolhimento visual mencionados na idealização do projeto. O plano de fundo apresenta uma animação suave de uma paisagem praiana com um farol ao fundo, buscando estabelecer a temática relaxante. As movimentações dos elementos presentes são calmas e simulam o vento e as

ondas do mar, ao mesmo tempo que uma trilha sonora é reproduzida durante a transição do usuário pelas telas do menu. A navegação inicial é simplificada em três opções centrais: "Iniciar", "Configurações" e "Sair". A função de saída encerra a aplicação, fechando a página da *web* e finalizando a sessão do usuário de forma segura.

Visando a acessibilidade e o conforto, o submenu de "Configurações" oferece ajustes de volume independentes para a música de fundo e para os efeitos sonoros, permitindo que o usuário adeque a experiência auditiva à sua sensibilidade. Além disso, foram implementadas opções de alternância, botões *toggles*, para desativar a rotação lateral automática da câmera, que é utilizada para vislumbres da paisagem, e o balanço de câmera durante a caminhada (também conhecido como *headbobbing*), que são utilizados nas fases subsequentes. Esses recursos metodológicos são fundamentais na engenharia do jogo para prevenir sintomas de cinetose digital e desorientação espacial, garantindo que os usuários que são mais sensíveis possam interagir de uma forma confortável.

Ao selecionar "Iniciar", o sistema exibe primeiramente um aviso ético e clínico. Esta mensagem tem a função de esclarecer que o jogo é uma ferramenta interativa de apoio ao bem-estar emocional, e que não deve substituir o tratamento primário ou então o acompanhamento psicológico e psiquiátrico profissional. Adicionalmente, a interface também emite uma recomendação para que seja feito o uso de fones de ouvido, atuando como uma estratégia de modulação sensorial voltada para tentar isolar estímulos externos e maximizar o estado de imersão do jogador.

Após a confirmação do aviso o usuário é direcionado a uma interface de autorrelato. Esta etapa é crucial para a coleta de dados sobre o estado emocional pré-intervenção, o que acaba fornecendo aos profissionais de saúde métricas importantes para o acompanhamento do usuário por meio de um ambiente diferente, que almeja ser percebido como mais confortável e seguro. A avaliação inicial é realizada em quatro etapas fundamentadas na literatura psiquiátrica: Primeiro, o usuário seleciona termos que melhor descrevem suas emoções no momento, com base na Escala de Afeto Positivo e Negativo (PANAS). Em sequência, ele quantifica o seu nível atual de ansiedade através de uma régua interativa com escala de 0 a 10, que foi baseada na Escala de Unidades Subjetivas de Angústia (SUDS). O sistema também disponibiliza um campo de texto livre, configurando um canal aberto para que o usuário externar pensamentos da forma como preferir e caso sinta vontade. Por fim, o Questionário de Saúde do Paciente (PHQ-9) é aplicado em um de seus formatos padronizados para rastreamento de sintomas depressivos, gerando um *score* automatizado. A conclusão deste fluxo de autorrelato encerra as interações de menu e aciona o carregamento do próximo ambiente virtual do jogo. É importante mencionar que parte dessa interface de autorrelato será repetida no final do uso da ferramenta, após a finalização das 3 fases propostas, com intuito de coletar dados comparativos antes e depois do uso da ferramenta. Mais especificamente, são repetidos os registros baseados

nas escalas PANAS e SUDS, enquanto o questionário PHQ-9 permanece restrito apenas à avaliação inicial no menu. Vale ressaltar também que, no escopo deste trabalho, a implementação destas escalas clínicas (PANAS, SUDS e PHQ-9) tem caráter arquitetural e de prova de conceito. O objetivo dessa funcionalidade é estruturar a base técnica para a futura integração e o envio desses dados de telemetria e acompanhamento diretamente aos profissionais de saúde, não sendo esses dados objeto de análise da avaliação preliminar deste estudo.

Com o intuito de minimizar as dificuldades do usuário durante os primeiros contatos com as mecânicas do sistema, foi desenvolvido um tutorial interativo e dinâmico. Estruturado para ser acionado automaticamente no início de cada uma das três fases principais, o tutorial tem como objetivo preparar o usuário para as exigências específicas da intervenção subsequente. Ele é apresentado por meio de uma interface baseada em cartões instrucionais sequenciais, onde cada tela é composta por três elementos de informação centrais: um título descritivo, uma imagem ilustrativa e um texto explicativo detalhando as orientações e os objetivos daquela etapa. Embora a arquitetura visual da interface do tutorial seja padronizada para manter a consistência, o conteúdo é modular e se adapta às atividades propostas em cada uma das fases.

Um aspecto importante deste tutorial é a sua integração com a mecânica principal de visão computacional. A navegação pelas instruções utiliza do mesmo sistema de reconhecimento de gestos rastreados pelo *MediaPipe* que será utilizado na jogabilidade. Para interagir com a interface, o usuário realiza movimentos corporais simples capturados pela câmera, como levantar o braço direito avança a tela para o próximo cartão de instruções, enquanto o levantamento do braço esquerdo permite pular o tutorial completamente, encerrando as instruções e iniciando a fase de forma imediata.

Essa decisão de *design* foi implementada com um duplo propósito. Em primeiro lugar, ela garante que os usuários adquiram experiência prática com a interface e compreendam a responsividade do sistema em um ambiente inicial, antes das práticas reais começarem. Em segundo lugar, atua como uma etapa implícita de calibração motora, permitindo que o jogador ajuste seu posicionamento frente à câmera e sinta-se confortável com o rastreamento antes de engajar-se profundamente com o conteúdo proposto. Essa pauta de calibração motora segue debate com a equipe, que pretende implementar futuras melhorias nessa etapa para tornar a experiência mais fluida e prevenir erros de calibração ou oclusão de partes importantes para a execução do jogo na visão da câmera.

3.2.2 Fase 1: Exercício de Respiração Quadrada

A primeira fase proposta foca na regulação autonômica por meio da técnica de respiração diafragmática, popularmente conhecida como respiração quadrada. Esta prática é uma intervenção fundamentada e normalmente reconhecida por sua capacidade de ativar

o sistema nervoso parassimpático e reduzir sintomas agudos de ansiedade (BALBAN et al., 2023). Embora a literatura técnica frequentemente referencie o padrão de quatro tempos (4-4-4-4), o sistema do jogo foi adaptado e configurado para um ciclo de três segundos (3-3-3-3): três segundos de inspiração, três segundos de retenção de ar, três segundos de expiração e três segundos de pausa pós-expiratória. Essa calibração metodológica visa proporcionar uma progressão mais acessível, reduzindo o risco de hiperventilação ou desconforto causado por longas retenções respiratórias em usuários ansiosos ou iniciantes. Visando a segurança, o sistema disponibiliza a função de pausa durante as atividades e prevê a opção de o usuário interromper ou sair da fase a qualquer momento, evitando que a mecânica atue como um gatilho ansiogênico e retenha o usuário. Para as próximas etapas do projeto, planeja-se que a definição do tempo de intervalo seja totalmente parametrizável, permitindo a customização do ciclo respiratório de acordo com as necessidades e limitações de cada pessoa.

O suporte para guiar a execução correta do exercício é feito de múltiplas formas. O jogador é instruído visualmente por um círculo animado no centro da tela, que se expande e se contrai em sincronia exata com o ritmo respiratório prescrito. Complementarmente à interface gráfica, efeitos sonoros específicos acompanham a animação temporal do círculo, ditando o compasso da respiração e auxiliando o usuário a manter o ritmo e imersão mesmo que desvie momentaneamente o olhar ou tenha outras eventuais distrações. A ambiência geral do cenário reforça a imersão por meio dos elementos biofílicos detalhados na concepção do *design*, englobando a emissão contínua de sons de ondas do mar e cantos de pássaros para facilitar o relaxamento (ULRICH et al., 1991; KAPLAN, 1995; ALVARSSON; WIENS; NILSSON, 2010).

Para garantir o engajamento contínuo e ancorar a atenção do usuário no momento presente, a fase incorpora uma mecânica de interação física viabilizada pelo rastreamento corporal do *MediaPipe*. Durante o fluxo do exercício, elementos de névoa animada surgem de forma gradual e começam a se acumular nas bordas da tela, cobrindo parcialmente a visão do cenário litorâneo. O usuário deve então utilizar o movimento de suas mãos para limpar e dissipar essa névoa em tempo real. Essa abordagem acopla o exercício respiratório a uma leve exigência motora, caracterizando uma técnica de *mindfulness* ativo.

O sistema monitora o desempenho da interação e fornece um *feedback* textual dinâmico diretamente na tela. As mensagens são programadas para alternar entre lembretes instrucionais, alertando o usuário para afastar a névoa caso ela preencha demasiadamente o campo de visão, e inserções de frases reconfortantes, estimulantes e de apoio emocional, que buscam validar os sentimentos do usuário durante a prática.

A sessão desta primeira fase foi projetada para ter uma duração contínua de um pouco mais de três minutos. No término desse período, a mecânica da névoa é encerrada, e o sistema emite um efeito sonoro indicativo de conclusão. Uma mensagem final de

parabenização surge, reconhecendo o esforço dedicado pelo jogador àquela atividade. Logo depois dessa validação positiva, o motor gráfico faz a transição automática para o tutorial da segunda fase.

3.2.3 Fase 2: Exercício de Afirmações Positivas

A segunda fase do jogo integra o uso de afirmações positivas, uma técnica frequentemente utilizada para promover o autoencorajamento e direcionar a atenção do usuário para pensamentos mais adaptativos. O usuário é colocado em um cenário de caminhada por uma praia virtual, onde se depara com obstáculos físicos instanciados de forma aleatória, como pedras, troncos caídos e vespeiros. Esses elementos de cenário atuam como representações metafóricas de barreiras emocionais. Para concluir a fase com sucesso, o jogador deve superar uma sequência de cinco obstáculos distintos, cada um exigindo uma interação específica baseada em fala.

Para progredir e transpor cada obstáculo, o usuário precisa vocalizar frases de afirmação positiva que são exibidas na tela. O sistema utiliza a *Web Speech API* para captar a entrada de voz pelo microfone, convertendo o áudio em texto mediante algoritmos nativos de correspondência de caracteres (*string matching*) para verificar a acurácia da pronúncia. Atualmente, a validação exige uma similaridade mínima de 85% entre a fala do jogador e a frase esperada. Esse limite de tolerância foi definido empiricamente durante testes técnicos prévios da equipe para acomodar variações normais de sotaque, velocidade de fala e ruídos de fundo menores, reduzindo a frustração do usuário com eventuais falhas de transcrição pontuais da API. O *design* dessa mecânica ampara-se na premissa de que a verbalização ativa exige um maior nível de engajamento, auxiliando o usuário a manter o foco nas mensagens de apoio.

Caso o sistema não reconheça a frase com a similaridade necessária, o usuário é instruído a repetir a afirmação. Contudo, para evitar que limitações de *hardware* do usuário, como microfones de baixa qualidade, ou variações na fala gerem ansiedade e frustração, o projeto prevê a futura consolidação de um mecanismo de dificuldade adaptativa. Nessa lógica, a cada tentativa malsucedida em um mesmo obstáculo, a taxa de similaridade exigida pelo sistema é gradativamente reduzida, até atingir um limite inferior aceitável que permita o avanço do jogador.

A imersão durante esta fase é reforçada por escolhas específicas de *design* de câmera e ambiente. Enquanto os elementos sonoros biofílicos e relaxantes da fase anterior são mantidos, a perspectiva visual introduz um leve movimento vertical contínuo, simulando o balanço natural de uma caminhada humana. Em momentos programados entre os obstáculos, a câmera também realiza rotações laterais suaves, proporcionando ao usuário vislumbres intencionais da paisagem 3D desenvolvida, com o intuito de maximizar a apreciação estética do cenário e imersão do usuário. Conforme estipulado na engenharia

do menu inicial, tais movimentações podem ser previamente desativadas caso o usuário possua sensibilidades relacionadas e prefira jogar sem essas movimentações.

A dinâmica é enriquecida pela inserção de efeitos sonoros pontuais associados à superação das barreiras e pela exibição de mensagens de reforço positivo ao longo do percurso. Após a transposição do quinto e último obstáculo, a caminhada é concluída. O sistema emite um efeito sonoro de sucesso acompanhado de uma mensagem de parabenização pelo esforço demonstrado. Em seguida, o motor gráfico inicia o carregamento da próxima etapa, efetuando a transição para o tutorial da Fase 3.

3.2.4 Fase 3: Organização Ambiental e Metáfora Comportamental

A terceira e última fase da intervenção lúdica emprega uma abordagem metafórica voltada à regulação emocional, fundamentando-se em pesquisas que relacionam a organização do ambiente físico ao bem-estar psicológico. Nesta etapa, o usuário é posicionado diante de um cenário virtual inicialmente desorganizado e repleto de elementos dispersos. O objetivo principal é restaurar a ordem do ambiente utilizando duas ferramentas simbólicas interativas, cujas ações são controladas por meio do reconhecimento de gestos processado pelo *MediaPipe*.

A primeira ferramenta disponibilizada é uma "borracha", que representa o conceito de desapego cognitivo. Ao selecioná-la e utilizá-la por meio de movimentos das mãos em frente à câmera, o usuário consegue apagar e remover elementos indesejados do cenário. Essa ação atua como uma dinâmica metafórica para a liberação e o afastamento de pensamentos intrusivos e emoções negativas. A segunda ferramenta é uma "varinha mágica", simbolizando a reconstrução positiva. Com ela, o jogador adquire a capacidade de introduzir novos elementos ou transformar os já existentes no ambiente, personalizando o espaço.

A implementação dessa mecânica dupla foi pensada para reforçar o conceito terapêutico de agência ativa na regulação emocional. Através da interação prática e da exigência motora suave, a dinâmica busca instigar no usuário a percepção de que ele possui controle sobre sua própria paisagem emocional interna e que tem a capacidade de modificá-la por meio de ações deliberadas e conscientes.

Ao finalizar a organização do cenário de acordo com sua preferência pessoal, o sistema oferece ao usuário a opção de realizar uma captura de tela do ambiente restaurado. Esse registro visual funciona como uma métrica qualitativa de conclusão e pode ser exportado e compartilhado posteriormente com os profissionais de saúde responsáveis. Essa funcionalidade de registro não apenas salva o *screenshot*, como também encerra a jornada do jogador pela fase 3 e aciona a segunda etapa da interface de autorrelato. Essa etapa final repete as coletas de informação baseadas nas escalas SUDS e PANAS, da mesma

forma que foi feito no menu antes do início da fase 1, estabelecendo a base arquitetural que viabilizará a futura comparação e avaliação do estado emocional do usuário pelo profissional de saúde no contexto de acompanhamento clínico.

3.3 Protocolo de Avaliação Preliminar

Para conduzir uma validação inicial da usabilidade e da capacidade de imersão do jogo antes de avançar para ensaios clínicos com pacientes reais, foi realizado um estudo de avaliação preliminar com uma amostra não clínica. O estudo concentrou-se em jovens adultos, um dos grupos demográficos mais afetados pela ansiedade e depressão de acordo com dados epidemiológicos, com o objetivo de avaliar se a interface e as modalidades de interação sem contato seriam intuitivas e engajadoras para a faixa etária alvo (KESSLER et al., 2005).

3.3.1 Participantes e Aspectos Éticos

A seleção dos participantes ocorreu por meio de amostragem por conveniência, recrutando indivíduos do ambiente acadêmico dispostos a participar do estudo. O único critério de exclusão rigoroso estabelecido foi a participação direta no desenvolvimento tecnológico do projeto, visando mitigar vieses de avaliação. A amostra final consistiu em dez adultos saudáveis voluntários, sem histórico relatado de transtornos de ansiedade ou depressão.

O protocolo do estudo foi submetido e aprovado pelo Comitê de Ética em Pesquisa (CEP) da Universidade Federal do Triângulo Mineiro (UFTM), sob o Certificado de Apresentação de Apreciação Ética (CAAE) nº 83857524.7.3001.5154, aprovado em 19 de maio de 2025. Todos os indivíduos forneceram o Termo de Consentimento Livre e Esclarecido (TCLE) de forma virtual, cujo aceite foi integrado como etapa obrigatória no próprio formulário eletrônico, Google Forms, utilizado para a coleta das respostas. Os dados coletados foram armazenados de forma anônima e segura em conta do Google Drive restrita aos pesquisadores e protegida com senha, garantindo a confidencialidade das informações e a conformidade com as diretrizes éticas para a pesquisa.

As características demográficas da amostra participante estão detalhadas na Tabela 1. A amostra foi composta por jovens adultos com média de idade de 22,3 anos (Desvio Padrão = 2,3; faixa: 20 a 27 anos), com uma distribuição de gênero equilibrada (60% feminino, 40% masculino). Em relação ao nível de escolaridade, a maioria dos participantes relatou estar cursando ou possuir pós-graduação incompleta (70%), enquanto os demais declararam estar com o ensino superior incompleto ou em andamento (30%). Além disso, a maior parte da amostra não possuía formação profissional na área da saúde (80%). Em relação à experiência com jogos, os participantes demonstraram níveis variados de

engajamento com videogames, variando de nenhuma experiência a jogo muito frequente, com 60% relatando não ter experiência prévia com jogos sérios. Nenhum participante relatou limitações físicas ou cognitivas que pudessem interferir na jogabilidade ou no rastreamento corporal.

Tabela 1 – Características Demográficas dos Participantes (Tamanho Amostral = 10)

Variável	Categoria	n (%)
Idade (anos)	Média $\pm$ DP	22,3 $\pm$ 2,3
	Faixa	20–27
Gênero	Feminino	6 (60%)
	Masculino	4 (40%)
Nível de Escolaridade	Pós-graduação (incompleta/cursando)	7 (70%)
	Ensino superior (incompleto/cursando)	3 (30%)
Área da Saúde	Sim	2 (20%)
	Não	8 (80%)
Frequência de Jogo	Nunca	1 (10%)
	Raramente	3 (30%)
	Às vezes	3 (30%)
	Frequentemente	2 (20%)
	Muito frequentemente	1 (10%)
Experiência com jogos sérios	Sim	4 (40%)
	Não	6 (60%)
Limitações Físicas	Nenhuma	10 (100%)

3.3.2 Protocolo Experimental e Equipamentos

As sessões de teste foram conduzidas de forma presencial e assistida, sendo executadas nos computadores portáteis (*notebooks*) dos próprios pesquisadores, equipados com *webcam* RGB integrada e fones de ouvido para o isolamento e *feedback* de áudio. A maior parte das avaliações ocorreu no ambiente do laboratório, enquanto uma parcela menor foi realizada em ambiente domiciliar. Durante todas as sessões, o pesquisador esteve presente atuando como um observador silencioso, encarregado de registrar eventuais incidentes técnicos ou erros do sistema, sem interferir na jogabilidade. O protocolo de posicionamento exigiu que os participantes permanecessem sentados, com o computador apoiado sobre uma mesa à frente, a aproximadamente a extensão de um braço esticado da tela, garantindo o enquadramento correto para o uso do *MediaPipe*. Embora a postura e a distância tenham sido padronizadas, a iluminação e o ruído não sofreram controle laboratorial estrito, configurando uma avaliação que reflete as condições mais próximas das reais de uso doméstico. Cabe ressaltar que não foram registrados incidentes técnicos graves e a taxa de conclusão das tarefas (*completion rate*) foi de 100% entre os voluntários.

Os testes foram executados utilizando o navegador Google Chrome, o que assegurou a total compatibilidade e a disponibilidade nativa da *Web Speech API*. O sistema foi configurado para processar as interações de voz tanto em português brasileiro quanto em inglês, adequando-se ao idioma de resposta escolhido por cada participante durante a execução. A duração das sessões não foi registrada como variável quantitativa estrita do estudo; todavia, o tempo médio estimado para a conclusão da intervenção, englobando a passagem pelas três fases do jogo, foi de aproximadamente 15 minutos por participante.

Cada sessão experimental seguiu um protocolo estruturado. Primeiro, os participantes receberam uma breve apresentação descrevendo os objetivos do projeto e a fundamentação terapêutica referente ao *design* do jogo. Em seguida, os usuários interagiram com as fases propostas, com o sistema de tutorial integrado fornecendo orientação inicial sobre as interações baseadas em gestos e voz. Não foram impostas restrições de tempo ou requisitos de desempenho, permitindo que os participantes explorassem as interações naturalmente e progredissem em seu próprio ritmo, a fim de avaliar a fluidez do sistema. Também não foram fornecidas explicações ou auxílios extras além daqueles que estão integrados na ferramenta e dos presentes nas instruções dos documentos iniciais, com objetivo de refletir de maneira mais precisa os dados de usabilidade e imersão.

3.3.3 Instrumentos de Avaliação

Após a conclusão da sessão de jogo, os participantes preencheram dois instrumentos psicométricos em suas versões traduzidas e validadas para o português brasileiro e inglês, aplicados por meio de uma plataforma de questionário *online* (Google Forms):

- ***System Usability Scale (SUS)***: Conforme fundamentado na Seção 2.5, o instrumento foi aplicado contendo 10 itens avaliados em uma escala Likert de 5 pontos, variando de 1 (discordo totalmente) a 5 (concordo totalmente). O objetivo foi aferir a usabilidade percebida do protótipo e gerar a pontuação composta para classificação de aceitabilidade do sistema (BROOKE, 1996).
- **Subescala de Imersão *EGameFlow***: Para quantificar a sensação de presença e o engajamento dos voluntários durante a jogabilidade, utilizou-se a dimensão de imersão deste questionário (FU; SU; YU, 2009), mediante a aplicação de seus 7 itens medidos em uma escala Likert de 7 pontos.

É fundamental esclarecer que, embora o protótipo desenvolvido possua uma interface funcional para o registro de autorrelatos clínicos do jogador (PHQ-9, PANAS e SUDS), os dados gerados por essas mecânicas internas não foram objeto de coleta, armazenamento ou análise metodológica nesta etapa de avaliação. O protocolo experimental do presente estudo limitou-se estritamente à aplicação das escalas externas de avaliação técnica (SUS

e EGameFlow), respeitando o escopo não clínico do grupo de voluntários e focando exclusivamente na validação da usabilidade e da experiência de uso da ferramenta.

3.3.4 Análise de Dados

Os dados coletados foram tabulados e analisados por meio de estatística descritiva, extraindo-se as médias (M) e os desvios-padrão (DP) para sumarizar o desempenho da amostra. Não houve tratamento de dados ausentes (*missing data*), visto que a configuração do formulário eletrônico exigia o preenchimento obrigatório de todos os itens.

O cálculo da *System Usability Scale* (SUS) seguiu o protocolo padrão de Brooke (1996). Para os itens de formulação positiva (ímpares), a pontuação foi calculada subtraindo-se 1 do valor da resposta do usuário. Para os itens de formulação negativa (pares), a pontuação foi obtida subtraindo-se o valor da resposta de 5, invertendo assim a escala. A soma de todas as pontuações ajustadas foi multiplicada por 2,5, resultando no *score* final de usabilidade, que varia de 0 a 100 (BROOKE, 1996).

Paralelamente, o cálculo da subescala de imersão do *EGameFlow* foi realizado através da média aritmética simples das respostas válidas (escala Likert de 1 a 7), sem a necessidade de inversão de itens nesta dimensão específica, com os resultados agrupados para identificar o nível geral de engajamento proporcionado.

3.4 Privacidade, Segurança de Dados e LGPD

O tratamento de dados sensíveis relacionados à saúde mental exige conformidade com a Lei Geral de Proteção de Dados (LGPD). No escopo deste protótipo, a arquitetura do sistema foi pensada para minimizar a exposição do participante e priorizar o processamento local (*on-device*). A captura e a análise de imagem pelo *MediaPipe* ocorrem de forma totalmente local, rodando no navegador do usuário, o que garante que o *feed* de vídeo da *webcam* não seja gravado, armazenado ou enviado para servidores externos em nenhum momento da intervenção.

Em contrapartida, o uso da *Web Speech API* para o reconhecimento de voz apresenta um ponto de atenção arquitetural. Dependendo do navegador utilizado, os pacotes de áudio podem ser enviados para servidores em nuvem de terceiros para realizar o processamento e a transcrição do texto. Para o uso do sistema em um ambiente clínico real, essa característica exigirá a formulação de um Termo de Consentimento Livre e Esclarecido (TCLE) específico e transparente, autorizando o tráfego desse dado, ou a futura substituição dessa API por um modelo de reconhecimento de voz treinado para rodar de forma totalmente *offline* no dispositivo.

Quanto à coleta de dados do autorrelato (resultados do PHQ-9, PANAS, SUDS,

campos de texto livre e eventuais capturas de tela da Fase 3), ressalta-se que o presente protótipo atua apenas como prova de conceito. Na versão final voltada para uso clínico contínuo, estabelece-se como diretriz que esses dados não serão armazenados localmente de forma desprotegida no dispositivo do usuário. Em vez disso, essas informações deverão ser criptografadas e transmitidas de ponta a ponta para servidores seguros ou destino adequado, garantindo que o acesso a essas métricas seja restrito única e exclusivamente aos profissionais de saúde autorizados e diretamente responsáveis pelo acompanhamento do usuário.

3.5 Assistência de Inteligência Artificial na Redação

Para garantir a clareza textual e a precisão da formatação técnica deste trabalho, foi utilizada a ferramenta de Inteligência Artificial Gemini 3.1 Pro ([GOOGLE, 2026](#)). A ferramenta foi aplicada exclusivamente como instrumento de suporte com o objetivo de: realizar a tradução e o refinamento gramatical de referenciais teóricos consultados em língua inglesa; efetuar a revisão ortográfica, correção gramatical, coerência e coesão do texto em língua portuguesa; e auxiliar na estruturação e codificação do texto na linguagem LaTeX para compilação na plataforma Overleaf. É importante ressaltar que a IA mencionada não foi utilizada para a geração de conteúdos teóricos, análises de dados ou outras tarefas não mencionadas. O texto foi revisado e validado pelo autor, o qual se responsabiliza pelo seu conteúdo.

4 Resultados e Discussões

4.1 Interface de Interação com o Usuário (UI)

A etapa de desenvolvimento resultou na construção de um protótipo funcional do jogo sério, estruturado para atender aos objetivos propostos no apoio ao tratamento de ansiedade e depressão. A arquitetura do sistema foi organizada em cenas sequenciais, que podem ser iniciadas e configuradas a partir do menu principal. O fluxo de navegação do usuário foi projetado para seguir uma ordem linear e contínua: inicia-se no menu (Figura 2), passa pela primeira interface de autorrelato (Figura 3), avança pelas três fases terapêuticas, retorna para a segunda coleta de autorrelato e chega ao encerramento. Toda essa transição de telas e o gerenciamento de estados do jogo foram viabilizados pela conexão direta de *scripts* desenvolvidos no projeto e pela integração com as APIs de rastreamento corporal e de reconhecimento de voz.

A interface gráfica foi projetada com base nos princípios de *design* centrado no usuário, buscando oferecer telas simples e intuitivas. Como o público-alvo pode estar vivenciando momentos de vulnerabilidade ou sobrecarga emocional, o direcionamento visual e lógico dentro do jogo foi pensado para evitar ambiguidades, diminuindo as chances de o usuário cometer erros de navegação que poderiam gerar frustrações indesejadas. A ambientação sonora e visual manteve-se fiel à temática relaxante da natureza durante toda a experiência. Esse cuidado se estendeu a todos os elementos da interface: caixas de mensagens, alertas e instruções que aparecem no decorrer do processo foram desenhados para carregar e transmitir seu conteúdo mantendo o tom calmo e acolhedor do projeto.

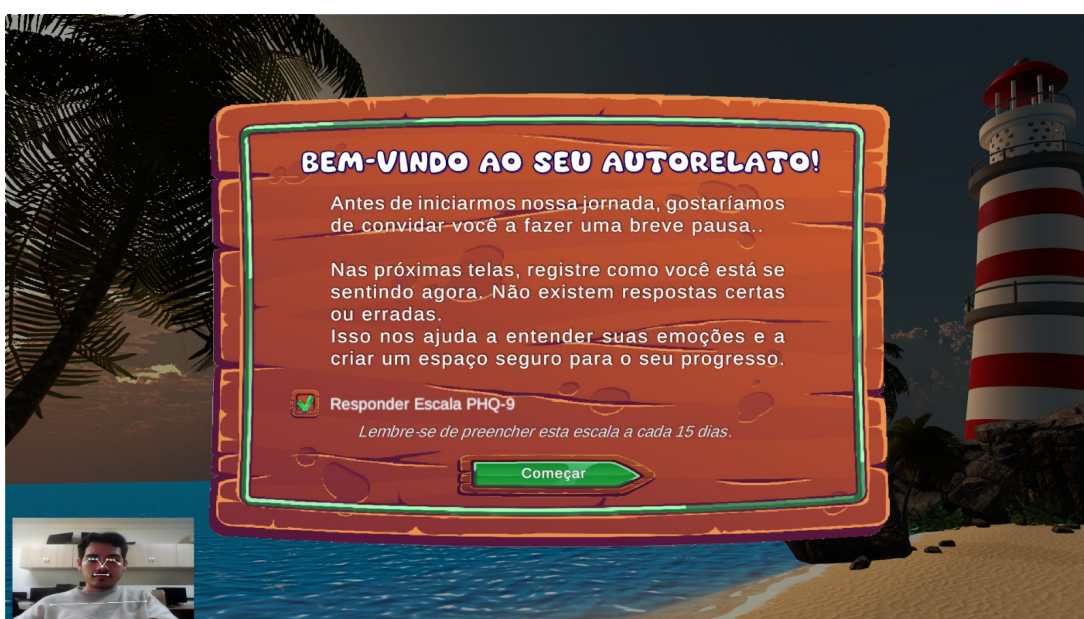
Uma preocupação especial durante o desenvolvimento foi garantir a fluidez da imersão. Tentou-se ao máximo otimizar o sistema para evitar travamentos ou tempos de carregamento longos entre as cenas. Nas transições onde o tempo de espera não pôde ser completamente eliminado devido ao carregamento de recursos tridimensionais, foram implementadas telas de carregamento temáticas. Nessas telas, além da indicação visual clara de progresso, como a porcentagem de carregamento, o jogador visualiza animações e frases de cunho motivacional e de apoio. Essa estratégia foi adotada para informar o usuário de que o sistema continua operando, evitando que ele se sinta perdido e mantendo o engajamento até que a próxima fase comece.

A estabilidade e a usabilidade dessa interface foram refinadas de forma empírica. O jogo foi validado inicialmente de forma local, dentro do próprio ambiente de desenvolvimento do Unity (modo *Play*), recebendo *feedback* contínuo da equipe. Em um segundo momento, a aplicação foi exportada e testada diversas vezes em ambiente WebGL, rodando por meio

Figura 2 – Interface do Menu Inicial.



Figura 3 – Interface de Autorrelato.



de um *live server* no computador do desenvolvedor. Essa etapa de testes diretamente no navegador web foi fundamental para aproximar o protótipo da realidade do usuário final. Durante esse processo, foi possível identificar e corrigir erros, travamentos ou comportamentos inesperados e problemas de redimensionamento de interface (UI) em diferentes proporções de tela, garantindo que o protótipo final estivesse de acordo com a proposta, rodasse de maneira mais estável e não prejudicasse a experiência do usuário.

4.2 Interface de Jogabilidade

A jogabilidade é o que forma o núcleo funcional do projeto, onde os conceitos terapêuticos foram efetivamente traduzidos nas nossas mecânicas interativas. A sequência das fases foi estruturada para construir uma leve narrativa de evolução do jogador. A interface de autorrelato incentiva o jogador a reconhecer o seu estado emocional atual e a trazer mais transparência. A Fase 1 (Figura 4) foi concebida com o propósito de favorecer o relaxamento e a atenção à respiração, preparando o terreno para a interação; entretanto, a confirmação desse efeito clínico na atenuação da tensão inicial requer avaliação em estudos futuros. Em seguida, a Fase 2 (Figura 5) passa a exigir uma interação mais ativa do usuário, que vai ter que vocalizar as frases de afirmação positiva, se concentrando no exercício de autoencorajamento. Por fim, a Fase 3 (Figura 6) atua como uma intervenção lúdica, voltada à regulação emocional do usuário, que se dá através da tomada de decisão trabalhada nesta fase. Essa estrutura de progressão foi projetada intencionalmente para reforçar a motivação do usuário, associando um *feedback* positivo à execução correta das mecânicas e ao progresso pelas fases, criando um ciclo de recompensa contínuo.

Figura 4 – Primeira Fase: Respiração.

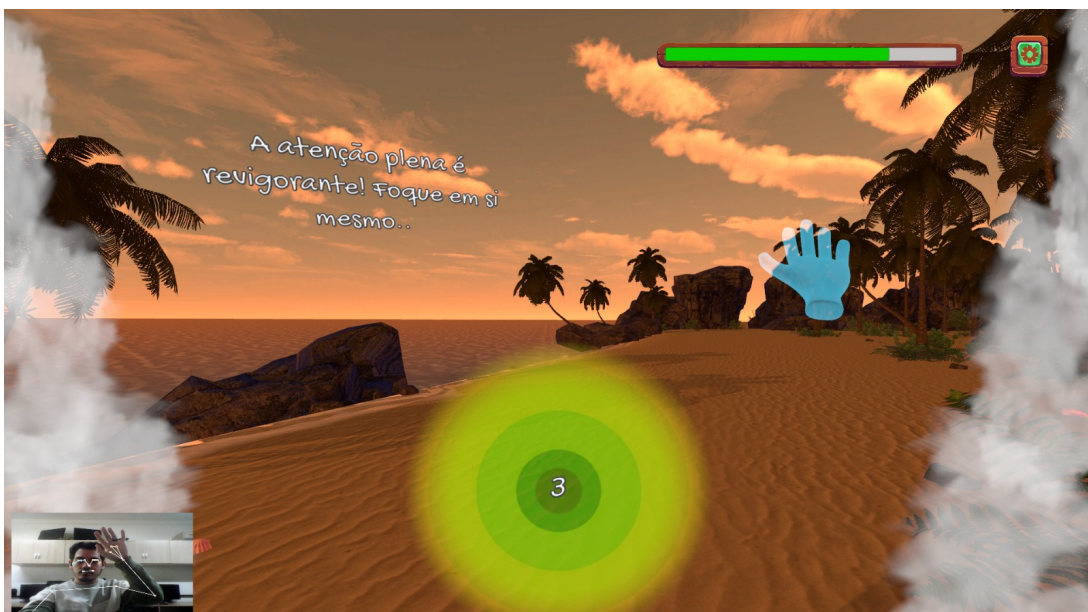


Figura 5 – Segunda Fase: Afirmações.

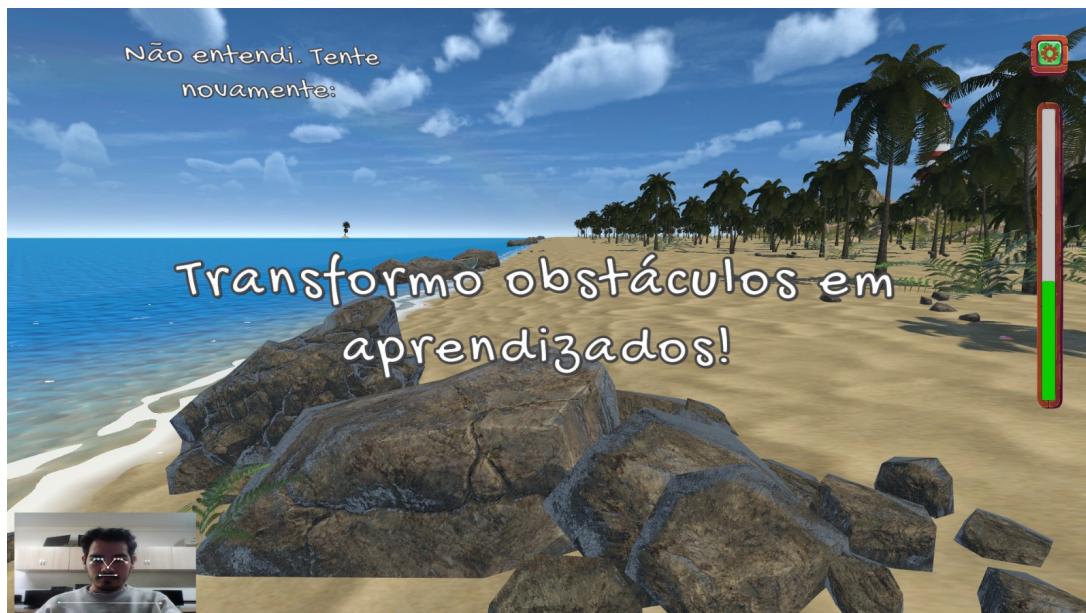


Figura 6 – Terceira Fase: Organização.



O desempenho técnico do jogo durante os testes demonstrou estabilidade operacional. O sistema integrado com o *MediaPipe* conseguiu reconhecer e atualizar a postura do usuário com latência adequada, permitindo que os jogadores interagissem e concluíssem as tarefas propostas nos cenários. A captação de voz pela *Web Speech API* também apresentou responsividade dentro dos parâmetros esperados. Algumas limitações encontradas durante os testes iniciais na segunda fase, que podem ocorrer devido a microfones de baixa qualidade, ambientes barulhentos ou sotaques muito fortes, foram mitigadas pela implementação da lógica de similaridade de texto. Como mencionado na metodologia, a aceitação de uma correspondência parcial provou ser uma solução de contorno funcional. Além disso, a futura implementação de uma mecânica de redução dinâmica desse limiar de similaridade será fundamental para facilitar o avanço e evitar gargalos de interação, modificação esta que já integra o conjunto de melhorias previstas.

Elementos de interface de usuário (UI) foram posicionados estrategicamente durante a *gameplay* para auxiliar a navegação. Barras de progresso foram inseridas para guiar e informar o jogador quanto ao tempo decorrido, número de repetições restantes ou ao desempenho esperado na atividade. Além disso, botões de pausa e de acesso rápido às configurações foram mantidos visíveis durante as fases, permitindo ajustes de volume ou o encerramento da sessão a qualquer momento, o que reforça a sensação de controle do usuário sobre a ferramenta. A intenção foi sustentar a imersão pelo uso de cenários temáticos coerentes e pela execução de efeitos sonoros de conclusão ao término de cada desafio.

O tutorial interativo (Figura 7) cumpriu seu papel de familiarizar o usuário com a tecnologia de rastreamento antes do início real das atividades, instruindo sobre o que deveria ser feito. No entanto, os testes revelaram oportunidades de melhoria. Na Fase 3, por exemplo, foi possível notar uma dificuldade recorrente. Alguns usuários se distraíam ou se perdiam na tarefa de limpar e organizar o cenário e acabavam se esquecendo das instruções passadas no tutorial. Consequentemente, eles não sabiam que precisavam clicar no botão de capturar a tela (*screenshot*) para sinalizar que estavam satisfeitos com a organização e, assim, finalizar a fase. Essa situação gerou dúvidas momentâneas para a conclusão da etapa. Como medida de contorno, as instruções do tutorial foram reforçadas, mas a equipe já prevê trabalhos futuros para melhorar essa sinalização dentro do próprio cenário da fase. Além disso, existe o planejamento de uma etapa formal de calibração, prevendo posicionamento e ajuste correto do corpo logo no início do jogo, o que vai ajudar a evitar problemas de oclusão de membros importantes perante a câmera ou webcam, o que pode acabar prejudicando o trabalho do *MediaPipe*.

Apesar dos contratempos pontuais que foram observados, e que são esperados em protótipos de desenvolvimento acadêmico, o sistema respondeu de forma consistente aos comandos gerais. A ferramenta atendeu ao objetivo técnico primário de atuar como uma

Figura 7 – Sistema de Tutorial.



opção de *software* sem sensor vestível viável, apresentando taxas de conclusão de tarefa satisfatórias e estabelecendo a base arquitetural para as futuras validações.

4.3 Avaliação de Usabilidade (SUS)

A avaliação quantitativa da usabilidade resultou em uma pontuação final composta de 85,8 em uma escala de 0 a 100. Analisando esse valor com base na escala de adjetivos consolidada por Bangor, Kortum e Miller (2008), pontuações acima de 85 inserem o sistema na categoria de usabilidade "Excelente" (BANGOR; KORTUM; MILLER, 2008; BROOKE, 1996). A mediana dos escores individuais foi de 85, com as notas variando entre o mínimo de 77,5 e o máximo de 92,5. Essa distribuição indica uma percepção de facilidade de uso consistente na amostra testada, sem grandes discrepâncias ou notas baixas que pudessem indicar falhas críticas na interface usada. É imprescindível destacar, contudo, que esse índice atesta exclusivamente a qualidade técnica e de engenharia da interface, avaliando a curva de aprendizado da ferramenta.

Para entender melhor onde o jogo acertou e onde ainda pode melhorar, a Tabela 2 apresenta a estatística descritiva de cada uma das dez perguntas presentes no questionário respondido pelos participantes.

Observando os dados detalhados, os itens relacionados à facilidade de aprendizado e à simplicidade do sistema foram os que mais se destacaram positivamente. A afirmativa 7 ("Aprenderia a usar rapidamente") alcançou a maior média geral ($M = 4,8$; $DP = 0,42$). Em contrapartida, a afirmativa 2 ("Achei o sistema desnecessariamente complexo") recebeu a nota mínima de todos os voluntários ($M = 1,0$; $DP = 0,0$), mostrando um consenso de

Tabela 2 – Estatísticas Descritivas dos Itens do Questionário SUS (Tamanho Amostral = 10)

Item do SUS	Média	DP	Faixa
Gostaria de usar frequentemente	3,1	0,57	2–4
Achei desnecessariamente complexo	1,0	0,00	1–1
Achei fácil de usar	4,5	0,53	4–5
Precisaria de suporte técnico	1,9	0,74	1–3
As funções estão bem integradas	4,4	0,70	3–5
Achei muita inconsistência	1,5	0,53	1–2
Aprenderia a usar rapidamente	4,8	0,42	4–5
Achei muito trabalhoso de usar	1,2	0,42	1–2
Me senti confiante ao usar o sistema	4,2	0,63	3–5
Precisei aprender muito antes de usar	1,1	0,32	1–2
Pontuação Final do SUS	85,8 / 100		

que a interface não é confusa. Esse dado, somado à baixa pontuação do item 10 (“Precisei aprender muito antes de usar”), ajuda a confirmar que a escolha de usar os gestos naturais e comandos de voz deixaram a ferramenta intuitiva o suficiente, mesmo para as pessoas que relataram não ter o costume de jogar videogames frequentemente.

Os participantes também relataram se sentir no controle durante a experiência. As afirmativas 3 (“Achei fácil de usar”) e 9 (“Me senti confiante ao usar o sistema”) mantiveram médias altas de 4,5 e 4,2, respectivamente. No entanto, um ponto de atenção interessante surgiu no item 4 (“Acho que precisaria do suporte de uma pessoa técnica para conseguir usar esse sistema”). Embora a média tenha sido baixa (1,9), o desvio padrão de 0,74 indica que houve uma variação moderada nas respostas.

Essa variação faz bastante sentido quando cruzamos esse dado com as observações práticas feitas durante os testes da Fase 3. Como alguns usuários tiveram dúvidas pontuais sobre como finalizar a organização do ambiente e salvar o *screenshot*, é provável que essa ligeira insegurança tenha refletido na pontuação do item 4. Isso reforça a conclusão que encontramos na análise de jogabilidade: mesmo com o sistema funcionando bem e a usabilidade geral sendo considerada excelente, o tutorial interativo e as sinalizações visuais dentro do jogo, especialmente na Fase 3, ainda precisam ser aprimorados para garantir que todos os usuários consigam navegar sem nenhuma ajuda externa adicional.

4.4 Avaliação de Imersão (EGameFlow)

Para mensurar o nível de engajamento cognitivo e a sensação de presença dos usuários durante a interação com o jogo, aplicamos a subescala de imersão do questionário *EGameFlow*. Diferente do SUS, esse instrumento utiliza uma escala Likert de 7 pontos, uma escolha metodológica adotada por oferecer maior sensibilidade nas respostas e evitar

problemas de interpolação na percepção dos usuários (FINSTAD, 2010). O resultado geral obtido pela ferramenta indicou uma média de imersão de 4,93 (Desvio Padrão = 1,33). De acordo com as diretrizes de interpretação adotadas para essa escala, médias abaixo de 4,0 apontam para uma baixa imersão, valores entre 4,0 e 5,4 indicam imersão média, e pontuações acima de 5,5 configuram alta imersão. Portanto, o protótipo atingiu um nível médio de experiência imersiva nesse primeiro teste realizado, o que demonstra que o ambiente conseguiu capturar e manter a atenção dos participantes. É evidente que foram encontradas diversas oportunidades para correções e eventuais melhorias que poderão ser abordadas e, futuramente, implementadas, almejando elevar o nível da usabilidade e do engajamento dos usuários.

A Tabela 3 detalha a estatística descritiva para cada um dos sete itens avaliados, que buscam cobrir diferentes espectros do estado de fluxo durante a jogabilidade.

Tabela 3 – Estatísticas Descritivas da Subescala de Imersão do EGameFlow (Tamanho Amostral = 10)

Descrição do Item	Média	DP	Faixa
I1. Esquecimento da passagem do tempo	4,90	1,29	2–6
I2. Perda de noção do ambiente ao redor	4,90	1,45	2–7
I3. Esquecimento das preocupações diárias	5,70	1,06	4–7
I4. Sensação de tempo alterada	4,60	1,07	3–6
I5. Envolvimento geral	5,50	1,08	4–7
I6. Envolvimento emocional	4,60	1,51	2–7
I7. Envolvimento visceral	4,30	1,89	2–7
Imersão Geral	4,93	1,33	

Ao analisarmos os dados de forma individualizada, fica evidente que o sistema teve um impacto muito mais significativo nas dimensões cognitivas da imersão do que nos aspectos puramente físicos ou instintivos. O item I3 ("Esquecimento das preocupações diárias") foi o grande destaque de todo o questionário, alcançando a maior nota média da avaliação ($M = 5,70$; $DP = 1,06$). O item I5 ("Envolvimento geral na tarefa", que reflete o grau de concentração profunda do usuário) também obteve uma excelente pontuação ($M = 5,50$; $DP = 1,08$).

Esses achados sugerem que as mecânicas desenvolvidas exigiram uma carga de atenção elevada por parte do usuário. A necessidade de sincronizar a respiração com a animação da tela (Fase 1) ou então de focar ativamente na pronúncia e validação das afirmações (Fase 2) demandou foco cognitivo contínuo nas tarefas lúdicas. Isso explica os altos índices de absorção na tarefa relatados pelos voluntários, corroborando a premissa técnica de que interfaces multimodais baseadas em voz e gestos requerem maior presença atencional do usuário para serem operadas adequadamente.

Em contrapartida, o item I7 ("Envolvimento visceral" ou absorção total no jogo) obteve a menor média ($M = 4,30$) e, principalmente, apresentou a maior variabilidade

de respostas ($DP = 1,89$). Essa oscilação mostra que a experiência de se sentir imerso e presente dentro do mundo virtual foi bastante heterogênea entre os usuários participantes.

Uma justificativa para esse comportamento reside na própria arquitetura sem uso de sensores vestíveis da aplicação. Como os usuários precisavam interagir através do rastreamento corporal (*MediaPipe*) e por comandos de voz, era necessário manter um nível de consciência contínua sobre a própria postura, a distância da câmera e a clareza da fala. Essa exigência de autocontrole físico e vocal pode acabar impedindo uma imersão visceral absoluta. Adicionalmente, as pausas programadas entre os minijogos para o carregamento do sistema e leitura do tutorial também atuam como interrupções do estado de imersão contínuo. Apesar disso, essas características não desvalidam a ferramenta, elas fornecem um direcionamento claro para possíveis implementações futuras, mostrando onde as melhorias de calibração motora e de *design* de transição devem ser focadas.

4.4.1 Feedback Qualitativo e Percepções dos Usuários

Além da coleta de dados quantitativa, os participantes forneceram *feedbacks* qualitativos por meio de um campo de texto livre inserido ao final do formulário eletrônico de avaliação. Para a análise e apresentação, as respostas abertas foram lidas e agrupadas tematicamente em três categorias descritivas principais: (1) percepção estética e sensorial, (2) dificuldades operacionais e (3) sugestões de melhoria. Na primeira categoria, os relatos destacaram a atmosfera do jogo, o *design* de som biofílico e o uso das frases motivacionais como elementos engajadores. Alguns participantes relataram percepções de conforto e relaxamento durante a sessão, sugerindo uma boa receptividade à ambientação simulada.

As respostas abertas foram fundamentais para identificar pontos cegos e oportunidades de melhoria no protótipo. Em relação à Fase 1, por exemplo, houve sugestões para que o ritmo da orientação respiratória, atualmente fixado em ciclos de três segundos, permitisse uma adaptação mais gradual, já que alguns usuários sentiram uma leve dificuldade em sincronizar a respiração logo no início. Na Fase 2, alguns voluntários descreveram a experiência como um pouco estressante em determinados momentos, citando que elementos auditivos, como o som repetitivo dos passos na areia, poderiam ser ajustados para tornar a caminhada mais agradável.

As observações qualitativas também reforçaram os dados colhidos sobre a Fase 3, onde os relatos confirmaram que as instruções iniciais poderiam ser mais claras e que o mecanismo de conclusão gerou uma sensação de incompletude para quem não entendeu exatamente como finalizar a etapa. Condições técnicas externas também foram frequentemente mencionadas como fatores que influenciaram a experiência, em especial a necessidade de uma boa iluminação ambiente e do posicionamento correto da câmera para que o rastreamento via *MediaPipe* funcionasse sem falhas.

Por fim, um ponto de discussão muito relevante levantado por alguns participantes foi a sugestão de integrar o jogo a dispositivos de Realidade Virtual (VR) para maximizar a imersão. Embora a implementação de VR tenha, de fato, um grande potencial para aumentar o isolamento cognitivo e a sensação de presença, essa abordagem exigiria o uso de *hardware* especializado e custoso. Essa funcionalidade entraria em conflito direto com uma das propostas deste trabalho, que priorizou o uso de tecnologias de acesso fácil e já difundidas, como *webcams* e câmeras embutidas e o processamento via navegador (WebGL), visando garantir que a ferramenta fosse barata e acessível para conseguir atender usuários em tratamento domiciliar, por exemplo.

5 Conclusão

O presente trabalho apresentou o desenvolvimento e a avaliação preliminar de um jogo sério projetado como potencial ferramenta tecnológica de apoio ao tratamento de ansiedade e depressão. Utilizando o motor gráfico Unity para a construção do ambiente virtual e exportação no formato WebGL, o sistema integrou o *MediaPipe* e a *Web Speech API* para viabilizar interações multimodais baseadas em rastreamento corporal e comandos de voz, priorizando o processamento local e explorando uma arquitetura sem sensores vestíveis. Essa abordagem, que dispensa o uso de sensores vestíveis ou equipamentos específicos adicionais, reforçou o caráter acessível e de baixo custo desse projeto, permitindo que o jogo seja executado diretamente em um navegador na web e facilitando o seu uso.

A execução do projeto permitiu que os objetivos específicos delineados fossem alcançados satisfatoriamente. Inicialmente, o levantamento de requisitos foi realizado de forma pontual por meio de consultas e orientação com profissionais da área de psicologia, fundamentando as dinâmicas adotadas. Em seguida, o ambiente virtual e as lógicas de jogabilidade foram desenvolvidos e otimizados para execução nativa em navegadores, integrando com sucesso as bibliotecas de visão computacional e de processamento de fala. As mecânicas de jogo baseadas em respiração quadrada, afirmações positivas e organização ambiental foram totalmente implementadas, bem como a estrutura de *software*, em caráter de prova de conceito, para a futura captura e transmissão de dados do autorrelato clínico. Por fim, a etapa de avaliação preliminar de usabilidade e imersão foi devidamente conduzida com o grupo de voluntários estabelecido.

Os resultados dessa avaliação, realizada com dez participantes, indicaram que o protótipo apresentou um funcionamento satisfatório nas condições testadas e uma elevada usabilidade percebida pela amostra. A aplicação do questionário SUS resultou em uma pontuação final de 85,8 de 100, sugerindo que a interface e os comandos de interação natural foram considerados claros e intuitivos pelos usuários avaliados. Complementarmente, a subescala do *EGameFlow* apontou um nível médio de imersão (4,93 de 7), sugerindo que as mecânicas foram capazes de gerar um nível adequado de engajamento cognitivo durante a sessão de teste.

Apesar da viabilidade técnica atestada, a avaliação também expôs que o sistema ainda tem espaço para aprimoramentos. Observou-se que a estabilidade das interações depende de fatores ambientais e de *hardware*, como a qualidade da iluminação, câmera e microfone do usuário. Além disso, foram mapeadas necessidades de refinamento na sinalização visual dos tutoriais, na melhoria da clareza de algumas mecânicas, como a interface de conclusão e salvamento da Fase 3, e na cadência das orientações respiratórias

da Fase 1. Esses ajustes deverão ser priorizados nos próximos ciclos de desenvolvimento.

Para contornar essas barreiras e continuar evoluindo a ferramenta, os próximos passos do projeto incluem a implementação de um sistema adaptativo de dificuldade para o reconhecimento de voz, uma etapa formal de calibração corporal no início do jogo, o aprimoramento geral dos tutoriais e melhorias pontuais em cada uma das fases.

Conclui-se, portanto, que o estudo atingiu o seu propósito de demonstrar a viabilidade técnica da construção de um jogo sério assistido por Inteligência Artificial rodando em navegadores comuns. Contudo, é imprescindível ressaltar que os dados atuais refletem exclusivamente a aceitabilidade tecnológica em um grupo não clínico. Qualquer afirmação acerca de benefícios clínicos reais, impacto na qualidade de vida, eficácia terapêutica ou efetividade da aplicação autônoma em ambiente domiciliar requer a futura condução de ensaios clínicos controlados com pacientes diagnosticados, operando sob o acompanhamento e a validação de profissionais de saúde mental. A ferramenta não tem a intenção de substituir o tratamento psicológico ou psiquiátrico profissional, mas sim de oferecer um suporte acessível e mais engajador. Com as melhorias mapeadas e as futuras validações clínicas, espera-se que essa solução possa contribuir para a redução das barreiras de acesso, usando a tecnologia como uma verdadeira aliada no cuidado em saúde.

5.1 Limitações do Estudo

Apesar dos resultados quantitativos favoráveis obtidos nas avaliações de usabilidade e imersão, é importante pontuar algumas limitações metodológicas inerentes a esta etapa de avaliação preliminar. Primeiro, o tamanho reduzido da amostra (tamanho amostral = 10) e a experiência heterogênea prévia dos voluntários com videogames indicam que os dados refletem uma percepção inicial, não capturando necessariamente toda a amplitude e variabilidade dos perfis dos usuários finais.

Além disso, como os testes foram realizados com uma amostra não clínica, composta por jovens adultos saudáveis e sem histórico de transtornos de ansiedade ou depressão, não é possível, neste momento, tirar conclusões definitivas sobre a eficácia da ferramenta. Indivíduos com diagnósticos clínicos confirmados podem apresentar diferentes padrões de interação, níveis de tolerância à frustração e desafios de usabilidade que não foram observados neste primeiro grupo de teste.

Outro fator a ser considerado é a dinâmica assistida do teste. Embora as sessões tenham tolerado variações naturalísticas de iluminação e ruído para simular o ambiente doméstico, a presença de um pesquisador, mesmo atuando apenas como observador silencioso, e a padronização inicial da postura física diferem de um contexto de uso domiciliar totalmente autônomo. Em um cenário de adoção real, a ausência completa de suporte presencial pode revelar desafios adicionais de calibração do usuário com a câmera

e com os comandos iniciais. Por fim, a própria natureza de protótipo da versão avaliada, que ainda não conta com todas as otimizações de *software* implementadas, como os ajustes dinâmicos de sensibilidade na captação de voz, pode ter gerado pequenos gargalos técnicos que impactaram a fluidez da experiência.

Do ponto de vista técnico e de *design*, uma limitação inerente ao protótipo atual reside na sua dependência de condições adequadas de *hardware* e ambiente físico. Embora o uso do *MediaPipe* e da *Web Speech API* favoreça a acessibilidade ao dispensar sensores vestíveis, a acurácia dessas ferramentas se mostra sensível a fatores externos. O reconhecimento corporal contínuo pode sofrer oclusões ou atrasos caso a iluminação do local seja inadequada ou se o computador utilizado apresentar limitações de processamento para executar a aplicação WebGL com uma taxa de quadros estável. De modo análogo, a captação de comandos de voz está sujeita a interferências de ruídos de fundo e à variação na qualidade dos microfones embutidos nos dispositivos de uso. Por fim, o volume de conteúdo desenvolvido até o momento, restrito a três fases curtas, é apropriado para a avaliação preliminar e prova de conceito, mas também pode se tornar repetitivo e limitar o engajamento a longo prazo caso a ferramenta seja adotada para o uso contínuo.

5.2 Implicações e Direções Futuras

Apesar das limitações pontuadas, os resultados obtidos com este protótipo reforçam a viabilidade do uso de inteligência artificial aplicada, especificamente visão computacional e reconhecimento de voz, no desenvolvimento de ferramentas acessíveis para a saúde mental. Os testes mostraram que é possível criar interações sem uso de sensores vestíveis que alcancem uma boa aceitação e fácil adoção pelo usuário, sem a necessidade de *hardware* especializado, quebrando barreiras tecnológicas e financeiras de acesso ([ANDRADE et al., 2014](#)).

Pensando nos próximos passos do desenvolvimento de *software* e *design* de interface, o *feedback* dos usuários e as observações feitas durante a etapa de jogabilidade traçaram um caminho claro para o refinamento do jogo. Está no escopo de trabalhos futuros a consolidação de um sistema adaptativo de dificuldade mais robusto para a *Web Speech API*, reduzindo as exigências de pronúncia de forma dinâmica para evitar frustrações na Fase 2. Adicionalmente, planeja-se a criação de um processo de calibração motora mais estruturado para o *MediaPipe* antes da Fase 1, além de melhorias nos mecanismos de *feedback* visual, como indicativos mais claros de conclusão e salvamento de tarefas na Fase 3.

No âmbito clínico e da pesquisa, as próximas etapas devem priorizar ensaios clínicos controlados e randomizados com o público-alvo real, pacientes com diagnóstico clínico, utilizando métricas validadas de desfecho terapêutico. Outras vertentes de investigação

necessárias incluem: avaliações longitudinais para medir a manutenção do engajamento e os efeitos do jogo a longo prazo; a integração da ferramenta com biossensores ou *wearables* para a coleta de métricas fisiológicas objetivas simultâneas; e a avaliação das melhores formas de integrar os dados de autorrelato extraídos do jogo diretamente aos prontuários e sistemas de telemonitoramento clínico existentes.

Por fim, é essencial reafirmar o escopo e o propósito fundamental desta ferramenta. O jogo sério desenvolvido neste trabalho atua estritamente como um instrumento complementar. Seu objetivo é apoiar a adesão ao tratamento, promover a regulação emocional e oferecer um suporte de autogerenciamento mais acessível ao usuário, não se propondo a substituir o acompanhamento profissional e os tratamentos convencionais que são aplicados em saúde mental (LAU et al., 2017).

Referências Bibliográficas

ACHARYA, V. et al. Beyond the joystick: deep learning games for hand movement recovery. *Frontiers in Rehabilitation Sciences*, v. 6, 2025. Disponível em: <<https://www.frontiersin.org/journals/rehabilitation-sciences/articles/10.3389/fresc.2025.1653302>>. Citado na página 18.

AHIRE, K. et al. Air canvas using media pipe for computer vision in unity 3d hand tracking. *International Journal of Advanced Research in Science, Communication and Technology*, p. 444–448, 4 2025. Citado na página 19.

ALONSO, J. et al. Treatment gap for anxiety disorders is global: results of the world mental health surveys in 21 countries. *Depression and anxiety*, Wiley Online Library, v. 35, n. 3, p. 195–208, 2018. Citado na página 16.

ALVARSSON, J. J.; WIENS, S.; NILSSON, M. E. Stress recovery during exposure to nature sound and environmental noise. *International Journal of Environmental Research and Public Health*, Molecular Diversity Preservation International, v. 7, n. 3, p. 1036–1046, 2010. Citado na página 26.

ANDRADE, L. H. et al. Barriers to mental health treatment: results from the who world mental health surveys. *Psychological Medicine*, Cambridge University Press, v. 44, n. 6, p. 1303–1317, 2014. Citado 4 vezes nas páginas 11, 12, 16 e 46.

BALBAN, M. Y. et al. Brief structured respiration practices enhance mood and reduce physiological arousal. *Cell Reports Medicine*, Elsevier, v. 4, n. 1, p. 100895, 2023. Citado 2 vezes nas páginas 18 e 26.

BANGOR, A.; KORTUM, P. T.; MILLER, J. T. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, Taylor & Francis, v. 24, n. 6, p. 574–594, 2008. Citado na página 39.

BROOKE, J. SUS: A 'quick and dirty' usability scale. In: JORDAN, P. W. et al. (Ed.). *Usability evaluation in industry*. London: Taylor & Francis, 1996. p. 189–194. Citado 5 vezes nas páginas 12, 20, 31, 32 e 39.

CAMILO, M. d. S. et al. Jogos sérios para a saúde: uma revisão narrativa. In: *Anais do XV Congresso Brasileiro de Informática em Saúde*. Goiânia: SBIS, 2016. p. 211–218. Accessed: 2025-06-15. Citado na página 12.

COHEN, G. L.; SHERMAN, D. K. The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, Annual Reviews, v. 65, n. 1, p. 333–371, 2014. Citado na página 18.

CUIJPERS, P. et al. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *The Canadian Journal of Psychiatry*, v. 58, n. 7, p. 376–385, 2013. Citado na página 11.

FINSTAD, K. Response interpolation and scale sensitivity: Evidence against 5-point scales. *Journal of Usability Studies*, v. 5, n. 3, p. 104–110, 2010. Citado na página 41.

- FLEMING, T. M. et al. Serious games and gamification for mental health: current status and promising directions. *Frontiers in psychiatry*, Frontiers, v. 7, p. 215, 2017. Citado 3 vezes nas páginas 11, 12 e 17.
- FORMIGA, N. P. F. et al. Serious games no ensino em saúde: uma narrativa de conceitos, arquiteturas e metodologias de desenvolvimento. *Revista Praxis*, v. 16, n. 30, 2025. Disponível em: <<https://revistas.unifoa.edu.br/praxis/article/view/4987>>. Citado na página 12.
- FU, F.-L.; SU, R.-C.; YU, S.-C. EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, Elsevier, v. 52, n. 1, p. 101–112, 2009. Citado 3 vezes nas páginas 12, 21 e 31.
- GONÇALVES, R. S. et al. Development of serious games for the rehabilitation of the human vertebral spine for home care. *Journal of Health Informatics*, v. 12, n. 2, 2021. Citado na página 17.
- GOOGLE. *Gemini (versão 3.1 Pro): Ferramenta de inteligência artificial*. 2026. Acesso em: 25 jul. 2026. Assistência em tradução, revisão gramatical e formatação LaTeX. Disponível em: <<https://gemini.google.com/>>. Citado na página 33.
- Google AI Edge. *Pose landmark detection guide (MediaPipe Pose Landmarker)*. 2025. Acesso em: 15 maio 2026. Disponível em: <<https://ai.google.dev/edge/mediapipe/solutions/vision/poselandmarker>>. Citado na página 19.
- Google Developers. *MediaPipe: On-device machine learning for everyone*. 2026. Acesso em: 15 maio 2026. Disponível em: <<https://mediapipe.org/>>. Citado 3 vezes nas páginas 18, 19 e 23.
- KAPLAN, S. The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, Elsevier, v. 15, n. 3, p. 169–182, 1995. Citado na página 26.
- KESSLER, R. C. et al. Lifetime prevalence and age-of-onset distributions of dsm-iv disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, American Medical Association, v. 62, n. 6, p. 593–602, 2005. Citado 2 vezes nas páginas 11 e 29.
- KOHN, R. et al. The treatment gap in mental health care. *Bulletin of the World Health Organization*, SciELO Public Health, v. 82, p. 858–866, 2004. Citado na página 15.
- LAU, H. M. et al. Serious games for mental health: are they accessible, feasible, and effective? a systematic review and meta-analysis. *Frontiers in Psychiatry*, Frontiers Media SA, v. 7, p. 209, 2017. Citado 3 vezes nas páginas 11, 12 e 47.
- LI, J.; THENG, M. A.; FOO, S. Game-based digital interventions for depression therapy: A systematic review and meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, v. 17, n. 8, p. 519–527, 2014. Citado na página 11.
- LI, X. et al. Considering the consequences of cybersickness in immersive virtual reality rehabilitation: A systematic review and meta-analysis. *Applied Sciences*, MDPI, v. 13, n. 8, p. 5159, 2023. Citado na página 17.

- LOHSE, K. et al. Video games and rehabilitation: using design principles to enhance engagement in physical therapy. *Journal of Neurologic Physical Therapy*, LWW, v. 37, n. 4, p. 166–175, 2013. Citado 2 vezes nas páginas 16 e 21.
- LOW, D. M. et al. Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology*, Wiley Online Library, v. 5, n. 1, p. 96–116, 2020. Citado na página 19.
- MASKELIŪNAS, R. et al. Evaluation of myrelief serious game for better self-management of health behaviour strategies on chronic low-back pain. *Informatics*, MDPI, v. 9, n. 2, 2022. Citado na página 20.
- MENG, Y. et al. Real-time hand gesture monitoring model based on mediapipe's registerable system. *Sensors*, MDPI, v. 24, n. 19, p. 6262, 2024. Citado na página 19.
- MERRY, S. N. et al. The effectiveness of sparx, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *BMJ*, British Medical Journal Publishing Group, v. 344, p. e2598, 2012. Citado na página 16.
- Ministério da Previdência Social. *Previdência Social concede mais de 472 mil benefícios por incapacidade temporária por transtornos mentais em 2024*. [S.l.]: Governo Federal do Brasil / INSS, 2024. Acesso aos dados abertos e boletins estatísticos da Previdência Social. Citado na página 15.
- Ministério da Previdência Social. *Previdência Social concede 546.254 benefícios por incapacidade temporária por transtornos mentais e comportamentais*. 2026. Acesso em: 15 maio 2026. Disponível em: <<https://www.gov.br/previdencia/ptbr/noticias/2026/janeiro/previdencia-social-concede-546-254beneficios-por-incapacidade-temporaria-por-transtornosmentais-e-comportamentais>>. Citado na página 15.
- NOROOZI, F. et al. Survey on emotional body gesture recognition. *IEEE Transactions on Affective Computing*, IEEE, v. 12, n. 2, p. 505–523, 2018. Citado na página 19.
- PORCINO, T. et al. A cybersickness review: causes, strategies, and classification methods. *Journal on Interactive Systems*, SBC, v. 12, n. 1, 2021. Citado na página 17.
- RAHANI, V. K.; VARD, A.; NAJAFI, M. Claustrophobia game: Design and development of a new virtual reality game for treatment of claustrophobia. *Journal of Medical Signals and Sensors*, Medknow, v. 8, n. 4, p. 231–237, 2018. Citado na página 19.
- RAHMANI-KATIGARI, M.; MOHAMMADIAN, F.; SHAHMORADI, L. Development of a serious game-based cognitive rehabilitation system for patients with brain injury. *BMC Psychiatry*, Springer, v. 23, n. 1, p. 893, 2023. Citado na página 19.
- ROEPKE, A. M. et al. Randomized controlled trial of superbetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. *Games for Health Journal*, Mary Ann Liebert, Inc., v. 4, n. 3, p. 235–246, 2015. Citado na página 17.
- ROSTER, C. A.; FERRARI, J. R.; JURKAT, M. P. The dark side of home: Assessing possession 'clutter' on subjective well-being. *Journal of Environmental Psychology*, Elsevier, v. 46, p. 32–41, 2016. Citado na página 18.

- SCHOLTEN, H.; GRANIC, A. Use of the principles of design thinking to address limitations of digital mental health interventions for youth: viewpoint. *Journal of Medical Internet Research*, v. 18, n. 6, p. e155, 2019. Citado na página 11.
- SCHONEVELD, E. A. et al. A neurofeedback video game (mindlight) to prevent anxiety in children: A randomized controlled trial. *Computers in Human Behavior*, Elsevier, v. 63, p. 321–333, 2016. Citado na página 17.
- SHIRES, G.; WEN, H. *Web Speech API*. W3C Community Group Report, 2022. Acesso em: 15 maio 2026. Disponível em: <<https://wicg.github.io/speech-api/>>. Citado 2 vezes nas páginas 20 e 23.
- SWEETSER, P.; WYETH, P. Gameflow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, ACM New York, NY, USA, v. 3, n. 3, p. 3–3, 2005. Citado na página 21.
- SWIFT, J. K.; GREENBERG, R. P. Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, v. 80, n. 4, p. 547–559, 2012. Citado 3 vezes nas páginas 11, 12 e 16.
- ULRICH, R. S. et al. Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, Elsevier, v. 11, n. 3, p. 201–230, 1991. Citado na página 26.
- Unity Technologies. *Unity Game Engine (Version 2022.3.41f1)*. 2022. Acesso em: 15 maio 2026. Disponível em: <<https://unity.com/>>. Citado na página 22.
- World Health Organization. *World mental health report: Transforming mental health for all*. Geneva, 2022. Citado 3 vezes nas páginas 11, 15 e 16.
- ZHANG, W. et al. Combined mediapipe and yolov5 range of motion assessment system for spinal diseases and frozen shoulder. *Scientific Reports*, Nature Publishing Group UK London, v. 14, n. 1, p. 15879, 2024. Citado na página 17.
- ZHENG, H. et al. Gaming your mental health: a narrative review on mitigating symptoms of depression and anxiety using commercial video games. *JMIR Serious Games*, JMIR Publications Inc., v. 9, n. 2, p. e26575, 2021. Citado na página 18.

6 Página do Artigo (IEEE SeGAH 2026)

A figura a seguir apresenta a primeira página do artigo *Beacon of Hope: Usability and Immersion Evaluation of an AI-Enhanced Serious Game for Anxiety and Depression*, derivado direto dos resultados desta pesquisa, aceito e apresentado na *14th International Conference on Serious Games and Applications for Health* (IEEE SeGAH 2026).

“Beacon of Hope”: Usability and Immersion Evaluation of an AI-Enhanced Serious Game for Anxiety and Depression

Sarah Chameroy¹, Marcelo Melo², Julia Tannus²,
Danilo Rocha², Bruno De Paula², Gwladys Rubayita¹ and Eduardo Naves²

¹Polytech Lyon, Claude Bernard University Lyon 1, Villeurbanne, France

²Faculty of Electrical Engineering, Federal University of Uberlandia, Uberlandia, Brazil

Abstract—Anxiety and depression affect over 970 million people worldwide (2022), yet conventional treatments face challenges including high dropout rates (20-60%), limited accessibility, and stigma. While serious games show promise as complementary interventions, most existing solutions require specialized hardware, limiting widespread adoption. This paper presents “Beacon of Hope,” an accessible AI-enhanced serious game that leverages contactless gesture recognition (MediaPipe) and speech interaction (Web Speech API) to deliver evidence-based therapeutic techniques without specialized equipment. The game implements three mini-games targeting diaphragmatic breathing, cognitive restructuring through positive affirmations, and emotional regulation. A preliminary usability and immersion evaluation with ten healthy young adults (aged 20-27) employed the System Usability Scale (SUS) and EGameFlow immersion subscale. Results demonstrated excellent usability (SUS: 85.8/100) and medium immersion (EGameFlow: 4.93/7), with particularly strong cognitive engagement facilitating temporary detachment from worries ($M = 5.70/7$). Qualitative feedback identified specific refinement opportunities for breathing pacing and tutorial clarity. The findings validate the feasibility of AI-powered natural interactions for mental health interventions, supporting accessible, stigma-free complementary approaches that reduce technological barriers while maintaining evidence-based therapeutic content.

Index Terms—Serious Games, Mental Health, Artificial Intelligence, Gesture Recognition, Usability Evaluation.

I. INTRODUCTION

Mental health disorders represent one of the major global public health challenges of the 21st century. According to the 2022 report from the World Health Organization (WHO), approximately one in eight people worldwide, nearly 970 million individuals, live with a mental health condition, with anxiety and depressive disorders being the most prevalent. Depression alone affects over 280 million people globally, while anxiety disorders impact an estimated 301 million individuals. The COVID-19 pandemic dramatically exacerbated this scenario: in the first year of the pandemic alone, global prevalence of anxiety and depression increased by 25%, representing an unprecedented surge in mental health burden [1].

The distribution of mental health conditions varies significantly across demographic groups. Young adults aged 18-

29 experience the highest rates of anxiety disorders, while depression peaks in both young adulthood and older age groups [2]. Geographic disparities are also evident, with low- and middle-income countries facing substantial treatment gaps: more than 75% of people with mental disorders in these regions receive no treatment at all [1].

Traditional treatment approaches for anxiety and depression typically involve pharmacotherapy, psychotherapy, particularly cognitive-behavioral therapy (CBT), or a combination of both [3]. However, these conventional interventions can face several challenges: high dropout rates (ranging from 20% to 60%), limited accessibility due to cost and geographic barriers, stigma associated with seeking mental health care, and lack of patient engagement [4], [5]. In response to these limitations, gamification and serious games have emerged as promising complementary strategies, which can be especially popular among the young adults. Serious games, defined as games designed with primary purposes beyond pure entertainment, can enhance motivation, increase adherence to therapeutic protocols, and provide accessible, stigma-free environments for mental health intervention [6], [7]. Recent systematic reviews demonstrate that gamified mental health interventions show significant improvements in anxiety and depression symptoms, with effect sizes comparable to traditional treatments while achieving higher engagement rates [8], [9].

The integration of Artificial Intelligence (AI) technologies can further enhance the therapeutic potential of serious games. MediaPipe, developed by Google, is an open-source framework that enables real-time gesture and pose recognition through machine learning models, allowing natural human-computer interaction without specialized hardware [10]. The Web Speech API provides browser-based speech recognition and synthesis capabilities, enabling voice-controlled interactions and linguistic pattern analysis [11]. These AI-powered technologies facilitate remote, accessible therapeutic interventions by: (1) enabling natural, intuitive user interactions that reduce technological barriers; (2) providing objective, quantifiable data on patient engagement and progress; (3) personalizing therapeutic content based on user behavior; and (4) eliminating the need for expensive specialized equipment, thereby democratizing access to digital mental health interven-