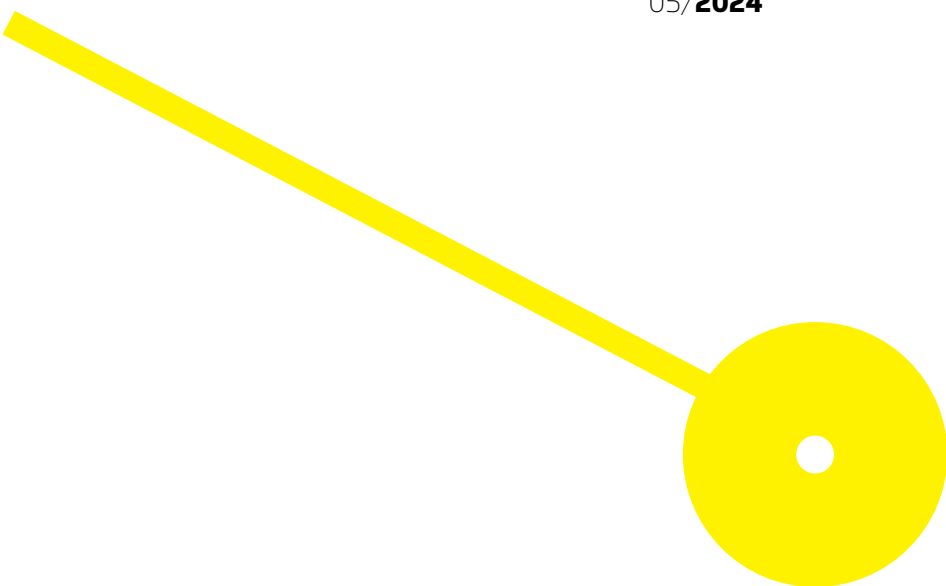




ProsoVR – desenvolvimento do protótipo de uma ferramenta de avaliação da prosódia mediada por realidade virtual

Diana Catarina Freitas Teixeira

05/2024





**ESCOLA
SUPERIOR
DE SAÚDE**



labrp |

LABORATÓRIO DE
REABILITAÇÃO PSICOSSOCIAL

**ProsoVR – desenvolvimento do protótipo de uma ferramenta de avaliação da prosódia mediada
por realidade virtual**

Autor

Diana Catarina Freitas Teixeira

Orientador

Professor Doutor Pedro André Guerreiro Martins de Araújo

Escola Superior de Saúde, Politécnico do Porto

Dissertação para cumprimento dos requisitos necessários à
obtenção do grau de Mestre em **Terapia da Fala** pela Escola Superior
de Saúde do Instituto Politécnico do Porto.

Agradecimentos

É estranhamente gratificante chegar ao ponto de escrever este capítulo. Escrever os agradecimentos àqueles que durante estes anos estiveram lá para mim. Engraçado ver a pessoa que eu era quando iniciei este percurso e pensar no que me tornei hoje. Este caminho não foi feito sozinha, até porque sozinha nunca teria chegado aqui. Por isso, agradeço:

Primeiramente ao Professor André Araújo, pela paciência, pelos anos de orientação, pelos desafios ao qual me propôs durante todo este tempo e por tudo o que me ensinou. É a minha referência no âmbito profissional.

Ao meu pai, por me ter ensinado a querer sempre mais e a ser ambiciosa. A não aceitar o pouco e a ser trabalhadora. À minha mãe, por me ter ensinado a ser uma boa pessoa e o quanto compensa sê-lo. A ser resiliente, como só ela sabe ser. Tenho os melhores pais do mundo e se hoje sou quem sou, deve-se aos meus pais e ao esforço que fizeram para que nunca me faltasse nada.

À minha irmã, por ser a minha companheira de todas as horas. Por ter sido mãe e pai, quando eles foram lutar por nós lá fora. Por ter feito o melhor que sabia para me proteger e por me ter dado a melhor prenda que podia ter tido, a minha sobrinha.

À minha Mariana, que é tão pequenina, mas que já me ensinou tanto. Que mudou a vida de todos nós e ainda nem imagina o quanto.

Ao Luís, por ser a calma da minha tempestade. Por ser o meu melhor amigo desde o primeiro minuto. Por ser o melhor namorado do mundo. Por me incentivar sempre a lutar pelos meus sonhos e por, sem nunca questionar, estar ao meu lado. Por acreditar mais em mim do que eu própria. Só podias ser Luís!

À Gete, Rogério, Margarida e Carolina por me terem acolhido sempre que precisei e por serem família no verdadeiro significado da palavra.

Resumo

Introdução: A prosódia apresenta um papel crucial nas interações sociais. Contudo, é especialmente difícil de avaliar, dado os protocolos tradicionais carecerem de validade ecológica. Para colmatar essa necessidade, surge o ProsoVR: um instrumento de avaliação da prosódia emocional expressiva, que recorre a ambientes virtuais para simular situações comunicativas.

Objetivo: Construção da versão Alpha do ProsoVR, concebida para realizar uma recolha de amostras de fala representativas de diferentes estados emocionais, e teste da validade ecológica e usabilidade do produto.

Métodos: Seguindo um modelo *Agile* de desenvolvimento, realizou-se o teste-piloto onde 6 utilizadores experienciaram a imersão virtual. De seguida, recorreu-se às escalas *Slater-Usoh-Steed Questionnaire* e à Escala de Usabilidade do Sistema para avaliar a validade ecológica e usabilidade do protótipo. Posteriormente, realizaram-se entrevistas semiestruturadas para identificar possíveis pontos de melhoria.

Resultados: Os resultados indicam que o protótipo possui elevados níveis de validade ecológica (80,56%) e usabilidade (81,25%), possibilitando uma imersão em ambientes virtuais ecologicamente válidos. Os utilizadores apontaram o Contexto 3 como o mais eficaz na indução da emoção-alvo.

Conclusão: Demonstrada a usabilidade e validade ecológica, o ProsoVR apresenta um protótipo funcional com potencial para avançar desenvolvimento da sua versão final e estudos em contexto clínico.

Palavras-chave: prosódia, realidade virtual, terapia da fala, avaliação, usabilidade, validade ecológica

Abstract

Introduction: Prosody plays a crucial role in social interactions. However, it is especially difficult to assess due to traditional protocols lacking ecological validity. To address this need, ProsoVR emerges: an instrument for evaluating expressive emotional prosody that uses virtual environments to simulate communicative situations.

Objective: Development of the Alpha version of ProsoVR, designed to collect speech samples representative of different emotional states, and to test the ecological validity and usability of the product.

Methods: Following an Agile development model, a pilot test was conducted where six users experienced virtual immersion. Subsequently, the Slater-Usuh-Steed Questionnaire and the System Usability Scale were used to evaluate the ecological validity and usability of the prototype. Semi-structured interviews were later conducted to identify potential areas for improvement.

Results: The results suggest that the prototype has high levels of ecological validity (80.56%) and usability (81.25%), enabling immersion in ecologically valid virtual environments. Users identified Context 3 as the most effective in inducing the target emotion.

Conclusion: After usability and ecological validity already observed, ProsoVR achieved a functional prototype and shows potential to progress to a final production and applied research in clinical environment.

Keywords: prosody, virtual reality, speech therapy, assessment, usability, ecological validity

Índice

1	Introdução.....	1
2.1	Prosódia: definições e conceitos	2
2.2	Protocolos de avaliação da prosódia	5
2.2.1.	Validade Ecológica e Usabilidade	10
2.3	Realidade virtual em Saúde	12
2.4	Aplicações da RV na TF	13
2.5	ProsoVR.....	14
2.6	Questão de partida e Objetivos	15
3	Métodos.....	16
4	Resultados.....	22
5	Discussão.....	34
6	Conclusão.....	37
	Referências Bibliográficas.....	38
	Apendice	43

Índice de abreviaturas

dB – decibel

F0 – Frequência fundamental

TF – Terapia da Fala

RV – Realidade Virtual

Bateria MAC-BR – Bateria Montreal de Avaliação da Comunicação

FAB – *Florida Affect Battery*

PECS-C – *Profiling Elements of Prosodic System – Communication*

PE – Português Europeu

PVSP – *Prodody Voice Screening Profile*

T-TRIP – *Tennessee Test of Rhythm and Intonation Patterns*

Mhealth – Mobile Health

1 Introdução

A prosódia é a área de estudo integrada na linguística alusiva à melodia do discurso, isto é, às variações das características suprasegmentais (frequência fundamental, duração e intensidade) que determinado interlocutor apresenta durante o seu discurso (Bodolay & Forte, 2023; Wells et al., 2004). Deste modo, a prosódia assume um papel crucial no que concerne à comunicação não verbal, uma vez que permite adicionar propriedades não-linguísticas aos significados das palavras, como emoções, ênfases e intenções (Bodolay & Forte, 2023; M. F. G. da Silva, 2011). Estas propriedades permitem dar ao discurso as mais variadas entoações melódicas transmitidas ao ouvinte, para que este possa descodificar e interpretar a mensagem.

Existem diversas classificações da prosódia, sendo que a avaliação padronizada divide-a em emocional e linguística (Hoekert et al., 2010). No presente trabalho dar-se-á ênfase à primeira, a prosódia emocional, que se foca na forma como o interlocutor consegue, através das demais variações melódicas, transmitir o seu estado emocional através do discurso. A “emoção” surge como resposta a um estímulo, interno ou externo, e assume um papel fundamental nas relações intra e interpessoais. Segundo Ekman (1970), as emoções podem ser divididas em seis categorias básicas: raiva, tristeza, felicidade, surpresa, medo e nojo (Silva, 2011).

Estas características dão à prosódia papel de destaque nas interações sociais (Carvalho, 2013; Hoekert et al., 2010). Pela sua relevância na componente paralinguística da comunicação, a prosódia é fundamental para uma adequada participação no que respeita à comunicação humana, pelo que é comum várias perturbações da comunicação apresentarem, entre outros aspetos, limitações ao nível do domínio da prosódia.

Do conjunto das patologias da comunicação com manifestações ao nível da prosódia, fazem parte diversos quadros, nomeadamente: disartria (Carvalho, 2013), gaguez (Lopes & Lima, 2014), disfonia (Vianello, 2006), surdez (Vasconcelos et al., 2011), lesões neurológicas (Souza & Cardoso, 2013), autismo (Zuanetti et al., 2018) e demência (Lopes & Lima, 2014).

O terapeuta da fala é o profissional responsável pelo estudo e reabilitação da comunicação humana e as suas perturbações, devendo por isso ter especial atenção à prosódia na sua avaliação e intervenção. O estudo da prosódia tem um papel importante dadas as fortes consequências que as pessoas com padrões prosódicos atípicos e/ou perturbados sofrem na sua comunicação. Assim, torna-se importante avaliar para perceber por que razão estas alterações acontecem e, assim, identificar a intervenção mais adequada (Filipe & Vicente, 2011; Hawthorne & Fischer, 2020).

Todavia, existe carência de protocolos de avaliação da prosódia validados para o português europeu (PE) e que apresentem adequados níveis de validade ecológica, uma vez que as amostras de fala são recolhidas em ambientes pouco realistas. Melhores protocolos de avaliação permitem uma

recolha de dados mais fiável, nomeadamente amostras de discurso mais válidas, permitindo diagnósticos mais precisos e um planos terapêuticos mais ajustados.

À medida que a tecnologia se desenvolve, as potencialidades da realidade virtual (RV) têm vindo a ser aplicadas em diversas áreas da saúde, já que permitem simular a presença em locais reais ou imaginários e proporcionam experiências sensoriais controladas (Fernandes, 2017).

Assim, com base no pressuposto que as novas tecnologias contribuem para colmatar a falta de realismo associado às avaliações formais existentes, foi proposta em estudos prévios (M. I. R. da Silva, 2021; M. C. Dias, 2022; Ferreira, 2022; D. Teixeira, 2021) a criação de um método de avaliação da prosódia mediado por RV: o ProsoVR. No presente estudo pretende-se apresentar o processo de desenvolvimento do protótipo do ProsoVR, bem como estudar a sua validade ecológica e a sua usabilidade para uso em contexto clínico e de investigação.

O presente estudo encontra-se dividido em seis capítulos principais, designadamente pela introdução e pelo enquadramento teórico, onde consta a revisão bibliográfica realizada. É também neste capítulo que surge o levantamento da questão de partida e os respetivos objetivos do estudo. Segue-se o terceiro capítulo, onde é apresentada a metodologia e todos os procedimentos inerentes à construção do ProsoVR. No quarto e quinto capítulos são apresentados, respetivamente, os resultados e a discussão. A dissertação termina com a conclusão do estudo, onde se apresenta uma reflexão sobre o desenvolvimento do ProsoVR, indicando as maiores entraves e desafios durante a realização do protótipo e onde é apresentado o levantamento das pistas para trabalhos futuros.

2 Enquadramento teórico

2.1 Prosódia: definições e conceitos

A comunicação humana vai muito além dos sentidos literais que os interlocutores verbalizam. Além disso, em bebé e numa fase pré-verbal, o ser humano faz uso dos padrões da entoação para comunicar intencionalmente (Tatit, 2020 citado por Lopes & Lima, 2014). Essas variações das características suprasegmentais (Lazzari & Petry, 2020), tais como a frequência fundamental (FO)¹,

¹ Frequência fundamental – parâmetro acústico medido em Hz, permite caracterizar um som de acordo com o número de ciclos vibratórios que as pregas vocais realizam aquando da passagem de ar pelas mesmas (Freitag & Lucente, 2017; Gramuglia, 2013)

duração², intensidade³ e velocidade⁴ de determinado material verbal ajudam a que um interlocutor consiga, através de uma mesma mensagem, transmitir um estado de alegria ou de tristeza (Galdino & Oliveira JR, 2023; Wells et al., 2004; Lopes & Lima, 2014).

Através da análise das informações não-linguísticas que determinado discurso apresenta, conseguimos auferir um significado que vai muito mais além do sentido literal das palavras. Estas informações dão à prosódia papel de destaque nas interações sociais (Carvalho, 2013; Hoekert et al., 2010; Lopes & Lima, 2014).

Embora a importância da prosódia nas participações sociais seja reconhecida, não é do conhecimento da autora a existência de uma definição específica que seja amplamente e/ou uniformemente empregue pelos terapeutas da fala.

O facto de não existir uma definição uniformemente utilizada durante as práticas clínicas, pode atrasar, prejudicar ou criar ambiguidades durante processo de avaliação e na tomada de decisão pela intervenção terapêutica mais assertiva (Carvalho, 2013).

A variação das definições da prosódia poderá decorrer das diferenças culturais, linguísticas e da abordagem metodológica que os autores priorizam (Ladd et al., 2019).

O termo prosódia surgiu no século VI a.C. para definir a acentuação das palavras. Mais tarde, passou a ser referente aos “eventos fonológicos que ocorrem num ponto específico da palavra” (Ladd et al., 2019). Desde então, várias são as tentativas de definir este termo ao longo dos tempos. Ladd *et al* (2019) realçaram a importância de diferenciar os traços prosódicos de outras características fonéticas e fonológicas para se conseguir chegar a uma definição linguística sólida (Ladd et al., 2019).

Em 1997, Fujisaki definiu prosódia como sendo referente às características segmentais e suprasegmentais das unidades linguísticas presentes no discurso, capazes de atribuir significado linguístico ou paralinguístico à fala (Fujisaki citado por Galdino & Oliveira JR, 2023)

Posteriormente já nos anos 2000, a prosódia é apresentada por Boutsen (2003) como “(...) uma ferramenta de expressão humana transmitida de forma acústica através de pistas de duração, intensidade e frequência. A essas pistas convencionais, poder-se-ia adicionar a linearidade (por exemplo, mudanças abruptas vs suaves de tom, volume ou duração) como uma possível quarta dimensão”⁵.

² Duração – “tempo de articulação” (Souza & Cardoso, 2013b) , geralmente medido em segundos.

³ Intensidade – parâmetro de avaliação vocal subjetivo que permite classificar um som como adequado, forte e/ou fraco (Behlau, 2004).

⁴ Velocidade – número de sílabas ou palavras que um interlocutor produz por minuto, permite caracterizar um discurso como lento ou rápido (Behlau, 2004).

⁵ “*Prosody is a tool of human expression that is conveyed acoustically by way of durational, intensity, and frequency cues. To these conventional cues, one could add linearity (e.g., abrupt vs. smooth changes in pitch, loudness, or duration) as a possible fourth dimension*” (Boutsen, 2003).

Por sua vez, mais recentemente, Bodolay e Forte (2023) apresentam o conceito de prosódia como sendo um elemento relativo às características suprasegmentais do discurso, como a melodia, entoação, acentuação e ritmo (Bodolay & Forte, 2023)

Assumindo a existência de diversas classificações, é comum no estudo da prosódia a divisão entre prosódia em linguística e emocional (Hoekert et al., 2010).

A prosódia linguística é relativa à melodia da fala que permite construir (quando na vertente expressiva) e/ou perceber (quando na vertente compreensiva) adequadamente, por exemplo, os enunciados exclamativos e interrogativos (Jorge, 2020).

Por sua vez, a prosódia emocional é referente à capacidade que o interlocutor tem em utilizar variações da prosódia (p.e. curvas melódicas) para exprimir emoções ou entender o estado emocional que o outro está a experienciar durante a comunicação verbal (Jorge, 2020).

Na vertente expressiva, a análise dos parâmetros vocais permite recolher dados concretos sobre as alterações dos valores de F0, duração ou de intensidade aquando variações das emoções (Oliveira Vassoler & Raposo de Medeiros, 2013).

Segundo Ekman (1970), as emoções podem ser divididas em seis categorias básicas: raiva, tristeza, felicidade, surpresa, medo e nojo (Silva, 2011), sendo que o presente estudo se foca somente na alegria⁶, raiva⁷ e tristeza⁸. Esta seleção será abordada e explicada mais à frente.

A título de exemplo apresentamos a fala de informantes do sexo e género masculino, que usualmente varia entre 100 e 150 Hz (Margarida & Nunes, 2009). No que concerne à alegria, os valores de F0 rondam os 160Hz e podem apresentar maior oscilação comparativamente às restantes emoções, com pausas no discurso que tendem a ser menos frequentes. A emoção de tristeza apresenta valores médios para F0 mais baixos do que as restantes emoções, rondando os 100Hz (Paulmann, 2008). As pausas podem ocorrer mais vezes e o ritmo de fala tende a ser frequentemente mais lento (Makarova e Petrushin, 2003 citado por Margarida & Nunes, 2009) já que as estruturas envolvidas se encontram tipicamente pouco ativas e flácidas (Oliveira Vassoler & Raposo de Medeiros, 2013). Por sua vez, o valor normativo médio de F0 para a raiva é de 256Hz (Paulmann, 2008), com uma intensidade vocal forte, uma velocidade de produção rápida e com poucas pausas, embora a bibliografia ressalve que as variações na raiva tendem a apresentar características mais individualizadas (Oliveira Vassoler & Raposo de Medeiros, 2013).

⁶ Alegria: emoção positiva que se manifesta como resposta a um estímulo de agrado ou sensação prazerosa, por exemplo, quando uma pessoa experiencia momentos agradáveis ou de conquistas desejadas (Oliveira Vassoler & Raposo de Medeiros, 2013).

⁷ Raiva: emoção que surge como resposta a situações que desafiam os valores ou necessidades do ser humano. Geralmente manifesta-se através de um aumento da tensão física e frequência cardíaca (Oliveira Vassoler & Raposo de Medeiros, 2013)

⁸ Tristeza: emoção associada a sentimentos de melancolia e desânimo que se manifestam como resposta a momentos de dor ou luto (Oliveira Vassoler & Raposo de Medeiros, 2013)

Pela sua relevância na componente paralingüística⁹ da comunicação (Serra, 2021), a prosódia é fundamental para que os interlocutores apresentem uma adequada participação no que respeita à comunicação humana, pelo que é comum várias perturbações da comunicação apresentarem, entre outros aspetos, limitações ao nível do domínio da prosódia. Do conjunto das patologias da comunicação com manifestações ao nível da prosódia, fazem parte os quadros de: disartria (Carvalho, 2013), gaguez (Lopes & Lima, 2014), lesões neurológicas (Souza & Cardoso, 2013a), disfonia (Vianello, 2006), surdez (Vasconcelos et al., 2011), autismo (Zuanetti et al., 2018) e demência (Lopes & Lima, 2014). Estes quadros são explorados mais adiante neste trabalho.

O terapeuta da fala é o profissional responsável pelo estudo e reabilitação da comunicação humana suas perturbações (C. O. Dias, 2019), devendo por isso ter especial atenção à prosódia na sua avaliação e intervenção.

2.2 Protocolos de avaliação da prosódia

O estudo da prosódia tem um papel importante dadas as fortes consequências que as pessoas com padrões prosódicos atípicos e/ou perturbados sofrem na sua comunicação. Por isso, torna-se importante avaliar para perceber por que razão estas alterações acontecem e, dessa forma, identificar a intervenção mais adequada (Filipe & Vicente, 2011; Hawthorne & Fischer, 2020).

As alterações da prosódia podem ser manifestadas ao nível da expressão, que é referente à forma como os interlocutores fazem uso das variações melódicas para transmitir emoções ou intenções linguísticas, como por exemplo, para fazer uma pergunta ou expressar sarcasmo (M. F. G. da Silva, 2011; Galdino & Oliveira JR, 2023; Ladd et al., 2019; Lopes & Lima, 2014).

Alterações ao nível da compreensão afetam a interpretação dos padrões prosódicos que o outro está a tentar transmitir durante a conversação, o que afeta a eficácia da comunicação já que pode levar a uma interpretação errada da intenção do falante (M. F. G. da Silva, 2011; Lopes & Lima, 2014). Nesse sentido, segue o conjunto dos quadros que padecem de alterações da comunicação com manifestações ao nível da prosódia mais comuns:

Tabela 1 - Caracterização das patologias da comunicação com manifestações ao nível da prosódia

Patologia da comunicação	Descrição
Disartria	Pessoas com disartria apresentam falta de coordenação pneumofonoarticulatória, amplitude e/ou precisão de movimentos. Geralmente a vertente expressiva da prosódia encontra-se alterada, com um discurso que apresenta variações desadequadas seja em excesso ou em insuficiência (Carvalho, 2013).

⁹ Informações suplementares não verbais, como ritmo, intensidade, frequência e duração, que complementam o significado do material verbal oral (Serra, 2021).

Gaguez	Interlocutores que apresentam alterações na fluência do discurso apresentam, também, prejuízo na vertente expressiva da prosódia. Seja através de repetições, prolongamentos ou bloqueios, por exemplo, a velocidade e as variações do discurso ficam alteradas (Lopes & Lima, 2014).
Lesões neurológicas	A literatura refere as dificuldades ao nível da prosódia causadas por lesão neurológica são classificadas como aprosódia ou hiperprosódia, quando existe falta ou excesso de variações prosódicas, respetivamente. Pessoas com lesão localizada no hemisfério direito apresentam frequentemente limitações ao nível dos extremos de FO, aliada a uma dificuldade em manter a velocidade de fala adequada (Souza & Cardoso, 2013).
Disfonia	A falta de qualidade vocal causada pela disfonia impossibilita a pessoa de realizar as adequadas variações de frequência fundamental, duração e intensidade, o que por sua vez afeta a prosódia na vertente expressiva (Vianello, 2006). Deste grupo destacam-se os laringectomizados totais, dado a complexidade do processo de adaptação a uma fala alaríngea, onde por vezes a recuperação das componentes prosódicas do discurso fica condicionada (Delft, 2018).
Surdez	Existe uma relação direta entre a perda auditiva e a diminuição da qualidade vocal. Comumente, as pessoas que apresentam perda auditiva tendem a apresentar alterações na emissão vocal, nomeadamente na presença de nasalidade, e nos traços suprasegmentais da fala, ou seja, dificuldades na adequação prosódia (Vasconcelos et al., 2011).
Autismo	A pessoa com autismo habitualmente apresenta um discurso com acentuação desadequada das palavras, podendo apresentar alguma ecolália. Tendencialmente apresenta simultaneamente dificuldade em descodificar as variações emitidas pelo interlocutor com quem comunica, o que leva a uma descodificação errada da mensagem (Zuanetti et al., 2018).
Demência	Num quadro demencial existe perda progressiva de competências, nomeadamente das competências cognitivas que, por sua vez, interferem no funcionamento normativo da comunicação. Tal sucede uma vez que existe uma dificuldade acrescida em retirar a intenção da conversação com o outro ou até mesmo em expressar aquilo que estão a sentir (Lopes & Lima, 2014).

A avaliação da prosódia expressiva pode ser feita através de avaliação áudio-percetiva, análise acústica, ou de forma combinada. Quanto à avaliação perceptiva, esta requer que o avaliador realize uma recolha da amostra de fala (*corpus* de fala) do indivíduo em avaliação por meio de uma tarefa (leitura, discurso conversacional, emissão de sons isolados,...), para proceder à respetiva análise auditiva por comparação àquilo que o ouvinte (pessoa que está a realizar a avaliação) considera normal. Este procedimento é, portanto, um método subjetivo de avaliação (Alves, 2015).

Por sua vez, a análise acústica requer igualmente uma escolha, por parte do terapeuta da fala, da tarefa e/ou material que o indivíduo em análise vai utilizar durante o procedimento, mediante os parâmetros acústicos que se procuram avaliar (Alves, 2015). A ferramenta *Prosogram* é utilizada neste tipo de avaliação uma vez que permite analisar os contornos prosódicos das amostras de voz (Mertens, 2004) e as tarefas mais utilizadas são: o discurso conversacional e a imitação de enunciados (Jorge, 2020).

Num estudo prévio da atual linha de investigação realizou-se um estudo que procurou reunir e resumir os protocolos da prosódia existentes. Verificou-se que a maioria dos protocolos descritos avaliam a prosódia apenas parcialmente, sendo esta, na sua maioria, incluída nos protocolos apenas como um parâmetro da qualidade vocal (Franco, 2022).

Além disso, constatou-se que apenas três dos protocolos estudados foram construídos/revistos a nível internacional há menos de 10 anos, nomeadamente o Protocolo da Avaliação da Qualidade Vocal da Universidade de Aveiro (Miguel Teixeira de Jesus et al., 2015), a Bateria Montreal de Avaliação da Comunicação (MAC) (Kerr, 2012) e o protocolo Profiling Elements of Prosody in Speech – Communication (PEPS-C) (Peppé & McCann, 2015), o que demonstra uma carência no investimento científico na área da prosódia (Freitas, 2020).

Na realidade internacional há diversos protocolos desenvolvidos para avaliar a competência prosódica, ainda que apenas parcialmente (Hawthorne & Fischer, 2020), os quais passamos a descrever sumariamente.

Frenchay Dysarthria Assessment

A segunda versão deste protocolo está adaptada para PE e avalia percetivamente os seguintes componentes prosódicos: FO, intensidade e velocidade da amostra de fala recolhida, que no caso será uma contagem de números e discurso espontâneo. O avaliador, atribui uma classificação qualitativa mediante as descrições presentes no manual de aplicação. Está apto para ser aplicado em adultos (Enderby & Palmer, 2008)

Florida Affect Battery

O protocolo Florida Affect Battery (FAB) construído por Bowers (1990), permite avaliar percetivamente a prosódia emocional (emoções de alegria, raiva e tristeza) na vertente compreensiva, e a prosódia linguística na vertente compreensiva também. Para a avaliação, é solicitado à pessoa que nomeie o material previamente gravado e estabeleça a correspondência dos padrões prosódicos às respetivas expressões faciais. Posteriormente é feita a atribuição de uma pontuação binária (correto/incorrecto) a cada uma das respostas fornecidas.

Profiling Elements of Prosody in Speech – Communication

O protocolo Profiling Elements of Prosody in Speech – Communication (PEPS-C) avalia a prosódia nas vertentes de expressão e compreensão. Primeiramente é feita uma avaliação inicial que tem como objetivo estudar as competências vocais do indivíduo para, posteriormente, avaliar as alterações de prosódia afim de perceber o impacto que estas têm na sua comunicação. Não se encontra traduzido para português, porém encontra-se traduzido e validado para inglês, neerlandês, francês, norueguês e espanhol (Peppé & McCann, 2015).

Prosody Voice Screening Profile

O protocolo *Prosody Voice Screening Profile* (PVSP) permite avaliar a prosódia na vertente expressiva através da análise perceptiva, em que é atribuída uma cotação de 0 (uso inapropriado) a 100 (uso apropriado), sendo que valores entre 0-90 são correspondentes a “não adequado” e valores superiores a 90 são considerados “adequados”. O avaliador introduz um tema para que seja avaliado o discurso espontâneo da pessoa, exceptuando palavras isoladas, imitações e interjeições (Shriberg, Kwiatkowski & Rasmussen, 1990).

Prosody Profile

O protocolo *Prosody Profile* construído por Crystal encontra-se disponível em inglês (britânico) e passível de ser utilizado com uma população adulta. A avaliação decorre através da recolha do discurso espontâneo, com posterior análise dos parâmetros vocais (*pitch/altura tonal*, velocidade, pausas e ritmo) para determinar o padrão prosódico e, assim, definir se o mesmo se encontra, ou não alterado (Crystal, 1979).

Swedish Prosodic Assessment

O protocolo *Swedish Prosodic Assessment* avalia a prosódia na vertente expressiva através da análise perceptiva de palavras, frases e discurso, onde é atribuída uma pontuação por cada produção correta. Destaca-se dos restantes protocolos por ser passível de ser aplicado na população juvenil (Samuelsson, Scooco & Nettelbladt, 2003).

Tennessee Test of Rhythm and Intonation Patterns

O protocolo *Tennessee Test of Rhythm and Intonation Patterns* (T-TRIP) encontra-se apto para ser utilizado em inglês (norte americano). Na aplicação deste protocolo é solicitado ao indivíduo que faça diversas repetições da sílaba /ma/, correspondentes a diferentes variações prosódicas. O avaliador, executa a avaliação acústica com recurso ao software *Praat*, e faz a avaliação perceptiva através da atribuição de uma classificação binária (correto/incorrecto) (Koike & Asp, 1981).

Protocolo da Avaliação da Qualidade Vocal da Universidade de Aveiro

Tal como o nome indica, este protocolo propõe-se a avaliar parâmetros de qualidade vocal e, por esse motivo, apenas é passível de avaliar parcialmente a prosódia. Trata-se de uma avaliação perceptiva do discurso espontâneo, em que o avaliador atribui uma classificação de “adequado” ou “diminuído” aos parâmetros de débito de fala, acentuação e às inflexões vocais são atribuídas classificações como “adequadas”, “agudas”, “graves”, “tremor” e “alteradas”. Destaca-se por ser dos poucos construído e validado em PE (Miguel Teixeira de Jesus et al., 2015)

Bateria Montreal de Avaliação da Comunicação (MAC)

Este instrumento caracteriza-se por ser dos raros na língua portuguesa e que se foca totalmente na prosódia. Avalia a vertente linguística – compreensão e repetição para a interrogativa, afirmativa e imperativa; e a vertente emocional – compreensão, repetição e produção de indicadores de alegria, raiva e tristeza (Knerr, 2012). A componente da produção da prosódia linguística é a única que, com este instrumento, não é possível avaliar. Durante a aplicação deste protocolo, para a componente da prosódia, são apresentados estímulos auditivos, sendo que o paciente deve identificar, repetir ou produzir (no caso da prosódia emocional apenas) consoante o conteúdo que é apresentado. Existe, para esta avaliação, o protocolo de registo pré-definido onde o examinador deve colocar um X no parâmetro correspondente à resposta do indivíduo. Para obtenção de resultados, devem ser feitos os somatórios de forma individual a cada parâmetro e comparar o resultado à norma. É de destacar que este protocolo, se encontra traduzido e adaptado para PE, com base da sua tradução a versão Brasileira Bateria MAC-BR (Knerr, 2012).

Os protocolos de avaliação vocal frequentemente avaliam segmentos de fala *standard* (como textos ou frases pré-definidos) os quais servem, à semelhança de outros protocolos de terapia da fala (TF), para basear um julgamento audioperceptivo (de natureza subjetiva) e/ou uma análise acústica do sinal de fala (Vasconcelos et al., 2011). Note-se que, para além da subjetividade associada ao avaliador, é importante considerar o contexto em que o *corpus* de fala é recolhido, nomeadamente quando tal acontece em contexto clínico.

Apesar disso, um estudo recente nos EUA indica que existe falta de conhecimento dos métodos de intervenção e avaliação mais adequados a cada caso e salientam que, na amostra recolhida, os terapeutas da fala relatam baixa experiência na prática clínica e baixa disponibilidade de recursos relativa a casos de alteração de prosódia (Hawthorne & Fischer, 2020). Em Portugal, resultados similares foram reportados, indicando que os terapeutas da fala salientam a importância de criar

maior oferta formativa na área e a falta de protocolos de avaliação formais da prosódia que sejam facilmente aplicados durante as sessões terapêuticas (Freitas, 2020).

Uma das críticas apresentadas a este tipo de protocolos é a sua expectável validade ecológica reduzida, uma vez que os resultados obtidos em contexto clínico são pouco naturais ou pouco próximos de uma situação real. À semelhança de outros instrumentos de avaliação, nomeadamente nas neurociências, todos os protocolos de avaliação da prosódia explorados implicam recolha de discurso em condições de comunicação artificiais, quer ao nível da compreensão, quer da expressão (Kourtesis et al., 2020).

No que concerne à avaliação perceptiva como a avaliação acústica da prosódia, torna-se necessário investir na construção de melhores protocolos para uma avaliação que demonstre maior validade e fiabilidade, à semelhança do que tem vindo a ser feito na área da avaliação da disфонia, das perturbações dos sons da fala ou da gaguez.

Ao recolher amostras de fala de forma mais sistemática será possível desenvolver métodos mais fiáveis de análise, o que permite uma avaliação diagnóstica mais precisa e um planeamento da intervenção mais adequado, bem como a possibilidade de medir e comparar os resultados das intervenções terapêuticas.

No que respeita à primeira problemática exposta, a reduzida validade ecológica associada às avaliações em contexto clínico, uma das soluções exploradas em neurociências tem sido o recurso à RV, já que permite a imersão controlada em ambientes “reais” e, assim, aproximar o mais possível a avaliação/intervenção ao mundo real (Dores et al., 2012; Manuel & Pinto, 2015).

2.2.1. Validade Ecológica e Usabilidade

À semelhança das provas padronizadas no âmbito da neuropsicologia, as provas de avaliação de TF recorrem comumente a estímulos pouco próximos de contextos reais, que podem condicionar a performance do indivíduo em estudo (Holmes, 2018).

Uma vez que a criação do ProsoVR visa proporcionar aos utilizadores ambientes virtuais que sejam representativos de ambientes do dia-a-dia, torna-se essencial avaliar a validade ecológica do protótipo.

Uma avaliação com validade ecológica visa aproximar o contexto o mais possível a situações reais, tendo em vista três critérios: realismo – o ambiente virtual deve ser o mais parecido possível com uma representação real (Massigli et al., 2011); união – ponto de encontro entre o estudo e a validade interna, isto é, o ambiente deve manter o nível de realismo que se aproxime de representações reais mas, ao mesmo tempo, ser capaz de demonstrar causalidade (Massigli et al., 2011); e análise eclética – análise de pelo menos um dos comportamentos em contexto real (Massigli et al., 2011), ou seja, o

tipo de tarefas que são solicitadas ao participante durante a imersão devem ser tarefas que facilmente sejam representativas do seu quotidiano (Pasquali, 2017). No presente estudo iremos focar o realismo e a análise eclética, uma vez que o objetivo do ProsoVR passa por simular a presença do utilizador num contexto real, para que as amostras de fala sejam recolhidas num ambiente próximo da realidade.

Para proceder à avaliação desses mesmos componentes, existe a escala de preenchimento autónomo de avaliação subjetiva *Slater-Usoh-Steed Questionnaire*, construída por Slater, Usoh e Steed no ano de 1994, comumente utilizada no âmbito de estudos que visam comprovar se efetivamente o utilizador experiêcia uma sensação de presença representativa da realidade. A escala permite quantificar o grau de realismo atribuído pelo utilizador aquando da imersão no ambiente virtual em estudo mediante a “sensação de estar lá” e no facto de a imersão virtual ser ou não capaz de transportar (metaforicamente falando) o utilizador para um local conhecido.

Para a validação da primeira versão da presente escala, os investigadores aplicaram-na num estudo empírico com o intuito de descobrir os aspetos que poderiam ser melhorados para potenciar a “sensação de estar lá” (Slater et al., 1994).

Diversos investigadores (Obana & Tori, 2021; A. C. O. da S. Ribeiro, 2013; A. Ribeiro & Monteiro, 2015) têm recorrido esta escala de avaliação da sensação de presença como uma ferramenta que possibilita estudar com mais detalhe a experiência do utilizador durante a imersão virtual, o que contribui para a análise da eficácia dos recursos em estudo.

Assim, a aplicação da versão traduzida para PE da escala *Slater-Usoh-Steed Questionnaire* torna-se fundamental em estudos que procurem estudar a experiência do utilizador em imersão ambientes virtuais.

Por sua vez, uma vez que o ProsoVR surge de um conceito ainda pouco explorado na TF, considerou-se importante testar e avaliar a usabilidade do mesmo.

A usabilidade de um produto é caracterizada com base na satisfação do utilizador durante a uso do produto, mediante a eficiência, eficácia e potencial do mesmo (Catecati et al., 2018; Scanferla, 2022). Estes parâmetros podem ser avaliados com base numa avaliação padronizada, nomeadamente recorrendo à Escala de Usabilidade do Sistema, em inglês *System Usability Scale* (SUS) construída por John Brooke em 1986, posteriormente traduzida e validada para o PE (Martins et al., 2015), comumente aplicada na avaliação de ferramentas inerentes à área da saúde (Scanferla, 2022). A pontuação total possível varia entre 0 e 100, estando valores acima de 68, segundo a bibliografia, associados a uma classificação de usabilidade considerada alta (Fonsêca RO, et al, 2015 & World Health Organization, 2011 citado por Scanferla, 2022).

Para a validação da versão traduzida e adaptada para PE da escala, realizou-se um estudo empírico junto de uma amostra que foi submetida a um processo de uso e teste de uma aplicação de agenda *online*. Os resultados obtidos sugerem que a versão traduzida e adaptada apresenta validade de constructo, com valores de concordância significativos (76,67%), o que sugere que esta versão se encontra passível de ser aplicada em estudos que pretendam determinar a usabilidade de determinado produto (Martins et al., 2015).

2.3 Realidade virtual em Saúde

Entende-se por RV a simulação eletrónica de ambientes através da utilização de dispositivos de saída (compostos por óculos, auscultadores e roupa) e dispositivos de entrada (compostos por microfones e controlo remoto) ligado a cabos onde o utilizador experimenta uma interação realista com um ambiente tridimensional (Manuel & Pinto, 2015).

Comparativamente a outros meios de comunicação ou sistemas de comunicação, a RV destaca-se pelas seguintes características (Burdea & Coiffet, 2003; Dores et al., 2012): i) Imersão – “sensação de estar dentro do ambiente, alcançada através de tecnologia específica, tais como óculos de projeção estereoscópica, telas especiais e luvas” (Burdea & Coiffet, 2003); ii) Interação – capacidade que o computador tem em detetar e modificar, de forma instantânea, as ações que o utilizador tem no ambiente virtual (Manuel & Pinto, 2015); iii) Envolvimento – envolvimento do utilizador com o mundo virtual, tendo por base o grau de motivação para a realização das atividades dentro do mesmo (Dores et al., 2012); iv) Presença – “sensação de estar lá”, ou seja, a presença física dentro do mundo virtual (Ribeiro & Monteiro, 2015).

Quanto à criação dos ambientes, de modo a perceber qual se enquadra melhor mediante o que é pretendido, os mesmos podem ser divididos como: realismo baixo, realismo intermédio, realismo elevado, consoante as ações forem mais ou menos parecidas com o mundo real (Manuel & Pinto, 2015).

A utilização da RV na reabilitação e avaliação clínica possui um papel importante e inovador, já que permite explorar as ferramentas tradicionais com um estímulo controlado, potenciando os níveis de motivação do utilizador e por isso revela-se útil para a investigação, formação, avaliação e intervenção terapêutica (Dores et al., 2012).

Existem atualmente aplicações da RV desenhadas para casos de paralisia cerebral, perturbações do espectro do autismo, síndrome de Down, doença de Parkinson e tratamento de várias perturbações como: alimentares, pânico e perturbações da imagem corporal (Dores et al., 2012).

Na área da saúde a aplicação da RV permite que, por exemplo, no caso do tratamento da ansiedade social, o paciente experiencie um ambiente onde o estímulo é dado de forma controlada e gradual (R. M. D. B. T. Teixeira, 2014).

Os equipamentos de RV viáveis têm um custo relativamente baixo, o que significa acesso ao mesmo para um maior número de pessoas. Torna-se igualmente importante o sistema ser bem concebido para que exista uma maior sensação de presença real do utilizador durante a experiência nos ambientes virtuais (Manuel & Pinto, 2015).

O uso da RV para fins terapêuticos traz motivação e interatividade para os pacientes. Isto é, a RV proporciona a oportunidade de fazer a terapia de forma divertida e diferente e reduz a vergonha e frustração que uma situação real de conversação pode causar aos pacientes, proveniente das suas limitações (Lee, 2019). O facto de ser possível aceder ao feedback imediato da avaliação ou da prestação do paciente é também uma mais-valia comparativamente com outros métodos convencionais (Lee, 2019). O controlo de estímulos que o paciente pode experienciar no ambiente virtual é também um dos motivos para que a RV seja importante para a intervenção terapêutica, já que estes podem ser ajustados consoante a fase em que o paciente se encontra (Lee, 2019). Estes fatores são diferenciadores e, por isso, a terapia associada à RV tem vindo a apresentar resultados evidentemente superiores, quando comparada à terapia que não faz uso das potencialidades da RV (Miranda et al., 2021; Lee, 2019)

2.4 Aplicações da RV na TF

Na TF, a utilidade das potencialidades da RV para o tratamento da gaguez já foi demonstrada (Walkom, 2016). O uso de RV durante sessões de pacientes com gaguez e ansiedade social diminuíram os níveis de ansiedade e melhoria ao nível da fluência em geral, possivelmente pelo facto de os pacientes se sentirem mais confortáveis naquele ambiente, o que os ajuda a preparar para uma situação de vida real (Walkom, 2016).

Ao nível da comunicação oral em público, a RV apresenta-se como uma boa opção para o treino de comunicação nas mais variadas situações, dado a sua versatilidade e possibilidade de simulação de uma infinidade de situações do quotidiano, difíceis de simular em contexto clínico.

Em Portugal, existe o projeto *“SPEAK UP VR Social Phobia - Virtual Reality for Public Speaking Training”*, desenvolvido pelo Laboratório de Reabilitação Psicossocial. Este projeto visa auxiliar os indivíduos no treino de competências de comunicação em público, através da imersão num ambiente que simula a apresentação oral perante um auditório com plateia. Nesse ambiente, o utilizador consegue receber feedback imediato da sua prestação com base nas reações da audiência (Donga et al., 2020).

Destaca-se também a nível nacional o projeto "*Speech Immersion Hub*", desenvolvido pela Speech Care, (mais concretamente pela equipa do iStutter Center/ Centro de Tratamento de Gaguez) em conjunto com a Sherlock GmbH e a *VideoSensing.studio*, que tem o objetivo de aprimorar a avaliação e intervenção em gaguez. Trata-se de um protótipo de RV que, apesar de ser desenvolvido a pensar nas disfluências do discurso, torna-se passível de ser aplicado às restantes áreas da comunicação, assim como de contribuir para a formação dos terapeutas da fala (Carvalheda, 2021).

Além dos mencionados, salienta-se também o projeto "*VIRTUS*", financiado pelo programa ERASMUS+. Este projeto tem como objetivo auxiliar jovens com autismo a potenciar as suas competências, de forma a facilitar o ingresso e permanência no mercado de trabalho através de um programa de formação profissional que usa a RV para o treino da performance profissional destes jovens (VIRTUS, 2023)

Apesar de estar bem desenvolvida noutras áreas, as imagens da RV na intervenção em TF ainda aparentam ser bastante irrealistas e desenvolvidas há já mais de quatro anos (Lee, 2019) (Bryant et al., 2020). Embora os estudos existentes sejam motivadores para os terapeutas da fala implementarem a RV na sua prática clínica, ainda não existe uma orientação de como o devam fazer, uma vez que as opções de RV existentes com aplicabilidade à TF são tendencialmente protótipos (Lee, 2019) (Bryant et al., 2020). Contudo, já é possível encontrar algumas aplicações clínicas em RV já estudadas em TF, nomeadamente nas áreas de: défice auditivo, disartria, afasia, perturbação específica da linguagem e perturbações dos sons da fala (Chen et al., 2016). Isto mostra que esta é uma área ainda muito incipiente na TF, mas também com muito potencial de desenvolvimento.

2.5 ProsoVR

O presente estudo está integrado numa linha de investigação dedicada ao estudo da prosódia. Nesta linha procedeu-se à construção de um protocolo de avaliação da prosódia emocional centrado em RV: o ProsoVR. Assim, o ProsoVR está integrado num projeto de investigação alargado, liderado pelo Prof. Doutor André Araújo, no Centro de Investigação em Reabilitação – Laboratório de Reabilitação Psicossocial (CIR-LabRP) do qual faz parte uma equipa que integra estudantes-investigadores, em curso desde 2018. Neste sentido, a autora do presente estudo desenvolveu a sua investigação de final de licenciatura no sentido de criar e validar o conteúdo de um instrumento de recolha de amostras de fala com recurso à RV, de forma a induzir emoções através da imersão em ambientes virtuais controlados e previamente definidos e que se assemelhem o mais possível ao mundo real. Neste estudo desenvolveram-se todos os guiões necessários para a construção do protótipo do ProsoVR, tendo os mesmos sido validados ao nível do seu conteúdo (D. Teixeira, 2021) e posteriormente testados com um conjunto de participantes normofalantes com e sem treino vocal

(Guimarães, 2023) e com perturbações da comunicação: disartria (Ferreira, 2022), fala alaríngea (M. C. Dias, 2022).

Para a construção do protótipo físico, a equipa de investigação seguiu recomendações da literatura (Kourtesis et al., 2020; Ribeiro & Monteiro, 2015; Teixeira, 2021) e optou pela gravação 360º pelo facto de ser possível recorrer a atores e cenários reais, o que atribui um maior nível de realismo aos ambientes virtuais e, por consequência, uma maior proximidade da experiência com as representações do mundo real. A imersão em RV deve ter uma duração total inferior a 60 minutos (Kourtesis et al., 2020).

O protótipo (em desenvolvimento) conta com quatro ambientes virtuais, cada um alusivo a uma componente/emoção específica. As emoções representadas no presente projeto foram selecionadas com base na Bateria MAC-PT, as quais são também referenciadas no protocolo *Florida Affect Battery*. Foram criadas frases “chave” que surgem em todos os contextos, de modo a existirem excertos homogéneos para comparação de resultados.

Além dos contextos alusivos às emoções, foi acrescentado um Contexto Zero que permite ao utilizador experienciar a imersão virtual antes de iniciar o momento de avaliação.

Tabela 2 – Caracterização dos ambientes virtuais do ProsoVR

Ambiente	Emoção-alvo	Descrição
Zero	Neutra	Contexto: Recepção Interação: Telefonista/rececionista Conteúdo: Pedido de orientações Objetivo: Interações básicas para experiência e teste do utilizador
Um	Alegria	Contexto: Cafeteria Interação: Empregado de balcão Conteúdo: Celebrar o sorteio e recebimento de um prémio Objetivo: Vivência de uma experiência indutora de alegria
Dois	Tristeza	Contexto: Consultório Interação: Médico Conteúdo: Transmissão de uma má notícia Objetivo: Vivência de uma experiência indutora de tristeza
Três	Raiva	Contexto: Estabelecimento Interação: Comerciante Conteúdo: Atendimento rude e com falta de educação Objetivo: Vivência de uma experiência indutora de raiva

Terminada a preparação dos guiões do protótipo no estudo realizado previamente (Teixeira, 2021), o desenvolvimento do protótipo prosseguiu para a fase de prototipagem, a qual será alvo da presente investigação e descrita mais adiante.

2.6 Questão de partida e Objetivos

Para a concretização do presente estudo, enquadrado na fase de prototipagem do projeto, definiu-se como objetivo principal “Desenvolver e testar a versão Alpha do protótipo do ProsoVR”.

De modo a orientar o estudo, estabeleceram-se os seguintes objetivos específicos:

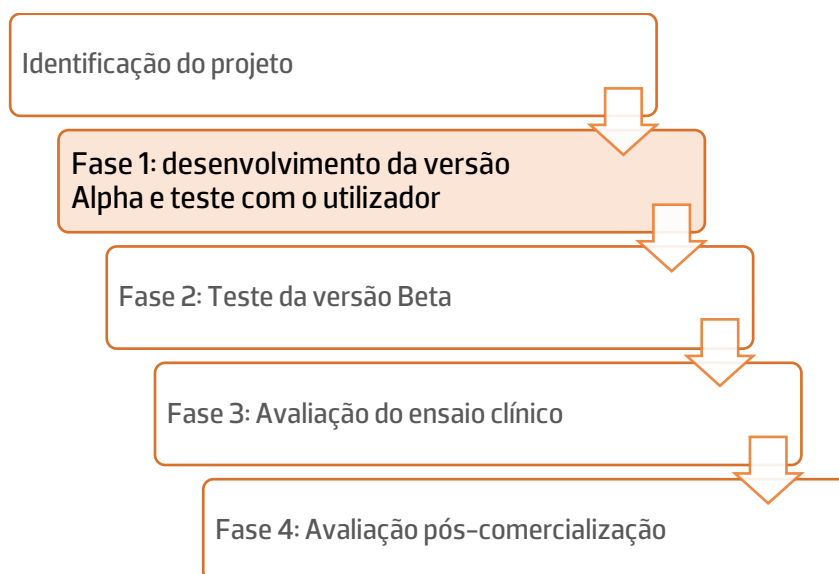
1. Atualizar e construir o protótipo físico do ProsoVR com recurso aos materiais previamente desenvolvidos e validados;
2. Avaliar a usabilidade do ProsoVR;
3. Avaliar a validade ecológica do ProsoVR;
4. Descrever, na ótica dos utilizadores, a qualidade da experiência do ProsoVR na recolha de amostras de fala representativa de diferentes estados emocionais.

3 Métodos

3.1 Tipo e Desenho de estudo

O presente estudo enquadra-se num processo de construção e validação de um instrumento, o ProsoVR, recorre a duas abordagens de forma complementar: abordagem qualitativa e abordagem quantitativa (Morse, JM 1991 citado por Alexandre & Coluci, 2009).

O processo de desenvolvimento de um protótipo com as características do ProsoVR vai ao encontro ao ciclo de desenvolvimento de ferramentas *mobile health*, commumente mencionado como *mhealth*. Os instrumentos enquadrados na área da *mhealth* procuram aprimorar os serviços de saúde através do recurso a dispositivos tecnológicos inovadores que auxiliam no processo de diagnóstico e intervenção (Carlos et al., 2016; Duque, 2017; Wilson et al., 2018). Este processo de construção encontra-se subdividido igualmente em fases, sendo a fase o presente estudo à Fase 1 da construção de uma ferramenta *mhealth* (Wilson et al., 2018).



Segundo a literatura, esta fase corresponde ao momento de construção e avaliação do produto numa perspetiva do utilizador. Esta fase é crucial para que as necessidades da população, seja profissionais ou utentes, possam ser suprimidas. Além disso, é nesta fase que a primeira versão do produto (versão Alpha) é submetida a teste que permite determinar se o protótipo serve o propósito para o qual foi construído e, igualmente importante, determinar se o mesmo cumpre as normas relacionadas com a segurança e privacidade (Wilson et al., 2018)

Concretizando, para a criação do protótipo do ProsoVR foi necessária a construção dos ambientes virtuais, cujos guiões já estavam previamente concluídos e validados ao nível do conteúdo por um painel de peritos (Teixeira, 2021).

Assim, numa primeira fase (fase 1), recorreu-se à criação do protótipo físico, incluindo a gravação e edição dos diferentes ambientes virtuais. Apesar de ter se partir de guiões previamente validados, durante o processo realizaram-se ajustes numa dinâmica exploratória de construção dinâmica, essencial para o processo de construção da versão Alpha do protótipo.

Na segunda fase (fase 2) deste estudo, realizou-se um teste de usabilidade e um teste alusivo à sensação de presença experienciada pelo utilizador, no qual se pretendeu recolher dados sobre a validade ecológica do instrumento. Esta fase tem uma natureza mista: quantitativa, na medida em que se recolheram dados parametrizados para determinar a validade ecológica e a usabilidade do protótipo apresentado; qualitativa, pois em paralelo foram realizadas entrevistas que permitiram receber feedback sobre a experiência de utilização do ProsoVR e identificar pontos fortes e pontos de melhoria.

3.2 Participantes

No presente estudo consideraram-se diversos participantes com diferentes tipos de contributo em cada uma das fases já apresentadas.

Na fase 1, para a construção do protótipo, contou-se com a participação de voluntários, uma vez que o mesmo implicava a gravação de cenários. Os participantes foram recrutados a partir dos contactos pessoais da equipa de investigação e obedeceram aos seguintes critérios: inclusão – ter mais de 18 anos, sem alterações autorrelatadas ao nível da comunicação, linguagem, fala e/ou audição e ter experiência (ainda que amadora) na área da representação. Não foram identificados critérios de exclusão relevantes.

Na fase de teste o grupo de participantes contou com 6 pessoas no total, através de uma amostragem de conveniência, obtida através da rede informal de contactos da equipa de investigação. Considerando que o Proso VR foi desenhado para ser usado em TF pelos clientes e pelos próprios terapeutas da fala, enquanto avaliadores, existiram 2 sub-grupos: a) terapeutas da fala; b) população em geral. No grupo A incluíram-se terapeutas da fala com pelo menos 2 anos de experiência profissional e, preferencialmente, com experiência em intervenção clínica. Os participantes do grupo B obedeceram aos seguintes critérios: critérios de inclusão – ter mais de 18 anos, sem alterações autorrelatadas ao nível da comunicação, linguagem, fala e/ou audição; critérios de exclusão – terapeutas da fala, atores ou pessoas experientes em representação, ter problemas de visão (ser cego) e limitações da fluência da leitura (nomeadamente pessoas analfabetas).

3.3 Instrumentos

Na Fase 1, para as filmagens usou-se uma câmara de gravação em 360º, modelo GoPRO FUSION. As imagens foram transferidas para o computador do centro de investigação onde foram submetidas ao processo de edição, no qual se adicionaram as instruções (*voice over*), previamente gravadas com recurso a um microfone modelo Audio-Technica MB1k|c e ao *software* Audacity versão 3.3.3.

Na Fase 2, para a imersão nos ambientes virtuais foram necessários óculos de RV, modelo Oculus QUEST com sistema de audio integrado, e um gravador de voz portátil para recolha das amostras de fala, modelo ZOOM F2-BT.

Para avaliar se o instrumento é efetivamente realista, a equipa de investigação recorreu à escala *Slater-Usuh-Steed Questionnaire (SUS)* construída por Slater, Usuh e Steed no ano de 1994, que permite qualificar o quão próximo da realidade se encontram os contextos mediante o nível de sensação de presença que o utilizador experiêcia (Ribeiro & Monteiro, 2015). Para isso é solicitado ao utilizador que atribua uma classificação pontuável de 1 a 7, conforme o afastamento ou aproximação da realidade, respetivamente, a seis questões relacionadas com a imersão virtual (Ribeiro & Monteiro, 2015).

Tabela 3 – “*Slater-Usuh-Steed Questionnaire*” (*SUS*) (A. Ribeiro & Monteiro, 2015)

1. Até que ponto se sentiu presente no cenário apresentado?

1 a 7, em que 1 corresponde a uma experiência normal, e 7 o sentido de estar presente no cenário.

2. Até que ponto o cenário apresentado se tornou uma realidade e quase se esqueceu do contexto laboratorial em que decorre a experiência?

1 a 7, em que 1 corresponde nunca, e 7 sempre.

3. Até que ponto ao recordar-se do cenário, o vê como uma imagem, ou como um sítio onde esteve presente?

1 a 7, em que 1 apenas como imagem, e 7 como um sítio em que esteve presente.

4. Durante o tempo da experiência, foi mais forte o sentido de estar no cenário ou de estar em outro lugar?

1 a 7, em que 1 corresponde a estar em outro lugar, e 7 estar no cenário.

5. Em que medida as condições estruturais (cor, objetos) do cenário o recordam de outras situações semelhantes em que já esteve?

1 a 7, em que 1 equivale pouco, e 7 muito.

6. Durante algum momento da experiência pensou que estivesse realmente no cenário?

1 a 7, em que 1 equivale a pouco, e 7 muito.

Para analisar a usabilidade do produto, recorreu-se à Escala de Usabilidade do Sistema, em inglês *System Usability Scale* (SUS) construída por John Brooke em 1986, posteriormente traduzida e validada para o PE (Martins et al., 2015) Trata-se de uma escala de preenchimento autónomo, em que o utilizador deve cotar cada uma das afirmações numa escala de likert, sendo que 1 corresponde a “discordo totalmente” e 5 “concordo totalmente” (Martins et al., 2015).

Os *scores* totais possíveis variam de 10 a 50, com uma variação de 40 pontos entre o mínimo e o máximo pontuável possível. Em termos percentuais, a usabilidade varia entre 0% (correspondendo ao total mínimo possível de 10 pontos) a 100% (correspondendo ao total máximo possível de 50 pontos).

Tabela 4 – *Escala da Usabilidade do Sistema* (Martins et al., 2015)

Escala de Usabilidade do Sistema

Acho que gostaria de utilizar este produto com frequência.

Considereei o produto mais complexo do que necessário.

Achei o produto fácil de utilizar.

Acho que necessitaria de ajuda de um técnico para conseguir utilizar este produto.

Considereei que as várias funcionalidades deste produto estavam bem integradas.

Achei que este produto tinha muitas inconsistências.

Suponho que a maioria das pessoas aprenderia a utilizar rapidamente este produto.

Considereei o produto muito complicado de utilizar.

Senti-me muito confiante a utilizar este produto.

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

Além das escalas supramencionadas, aplicaram-se questões de resposta aberta adicionais (Apêndice 2) em modo de entrevista semiestruturada, com o objetivo de recolher *feedback* complementar no que respeita à proximidade dos contextos com a situações reais e no que concerne à indução da emoção *vs* emoção efetivamente experienciada. Aqui, é feito o levantamento de qual dos cenários estava mais próximo de uma representação real e de que forma as personagens, sons e legendas se encontram alinhadas com o objetivo proposto.

3.4 Procedimentos

Para a concretização do protótipo do ProsoVR foi necessária a formalização da colaboração com o LabRP - Laboratório de Reabilitação Psicossocial, para que fosse facultada consultoria e formação, acesso às instalações e aos recursos materiais e humanos necessários para a realização do projeto. O presente estudo foi submetido à aprovação pela Comissão de Ética da Escola Superior de Saúde do Politécnico do Porto (CE0020D).

O desenvolvimento do projeto decorreu em duas fases distintas, previamente identificadas como: 1. construção do protótipo, e 2. recolha de dados. Segue-se a descrição dos procedimentos inerentes a cada uma das respetivas fases.

Fase 1 – Construção do protótipo

Tal como previsto, no processo de produção do protótipo, apesar de existir uma validação prévia dos guiões (Teixeira, 2021), constatou-se ser necessário proceder a ajustes no sentido de alcançar um produto mais realista e funcional.

Para isso, a equipa de investigação construiu um guião que foi analisado por um painel de peritos e posteriormente revisto e construída a versão final do mesmo para ser utilizado nas gravações dos cenários do ProsoVR.

Durante o decorrer da imersão, o utilizador encontra frases “chave”, com a finalidade de proporcionar excertos de fala homogéneos. A equipa selecionou e validou junto de um painel de peritos (em estudos anteriores) frases que retratassem expressões do dia-a-dia que permitissem transmitir as três emoções selecionadas, sendo elas “se eu soubesse...”, “uns com tudo, outros sem nada” e “obrigado”.

Numa fase inicial do presente trabalho, aquando da preparação das gravações, procedeu-se a uma nova reformulação das falas do guião do Contexto 2 (por ter sido considerado o mais sensível devido à simulação de uma situação clínica) com o auxílio de uma especialista em comunicação clínica, que

forneceu orientações claras sobre a estrutura da comunicação de uma má notícia em ambiente clínico.

Após essa reformulação, solicitou-se a autorização à Escola Superior de Saúde do Politécnico do Porto para usufruir das instalações, com o objetivo de realizar as gravações nas mesmas. Contudo, na impossibilidade de encontrar um ambiente que simulasse o que inicialmente se pretendia para o Contexto 3, adaptou-se o guião a um cenário diferente. O que inicialmente seria gravado num quiosque, passou a ser numa loja de comércio local e a interação manteve-se entre um cliente, que será o utilizador, e a funcionária desse mesmo estabelecimento.

Quanto à recolha de imagem, a mesma aconteceu num só *take* para que não fossem necessários cortes, uma vez que esse tipo de montagem poderia prejudicar a qualidade da imagem e interferir negativamente na experiência do utilizador.

Uma das limitações ao explorar um ambiente virtual prende-se com o facto de a área do ambiente ser restrito, isto é, o indivíduo que experiencie o ambiente virtual não pode circular livremente pelo espaço (Kourtesis et al., 2020). Por esse motivo, nas filmagens a câmara permaneceu fixa numa distância de entre 50cm e 100cm do ator, ao nível do olhar do ator, para que se possa simular de forma realista a manutenção do contacto ocular com o mesmo.

Após as gravações concluídas, os vídeos brutos foram alvo de análise e edição no LabRP - Laboratório de Reabilitação Psicossocial. Uma vez que os ambientes são gravados num só take e as edições (como cortes e adições) não são possíveis dado a natureza das ferramentas, nesta fase de edição foi necessária a adaptação e reformulação de alguns dos guiões na fase de edição, de forma a ajustar o texto que o utilizador tem de produzir aos tempos de resposta que o ator fornece. Para aumentar o nível de realismo, autores realçam a importância de adicionar sons específicos que permitam aumentar a sensação de presença real no espaço (por exemplo: sons ambientes) ou até mesmo para dar *feedback* ao indivíduo durante a imersão (Kourtesis et al., 2020). Assim, o som dos momentos de silêncio, isto é, o momento em que o utilizador vai ler as legendas, foi reajustado para que não existisse quebras em relação ao restante som ambiente do próprio ambiente em que foram realizadas as gravações.

De forma a uniformizar a experiência a todos os utilizadores, os excertos correspondentes a cada um dos cenários foram condensados num vídeo único e adicionadas em voz *off* orientações e comandos verbais previamente gravados.

As orientações fornecidas vão desde a receção do participante, em que são dadas as boas-vindas e é brevemente apresentado o protótipo, e também sobre como o equipamento deve ser ajustado para um maior conforto. Além disso, também são fornecidas as instruções sobre o guião que o participante deve seguir (legendas) e sobre o posicionamento físico que o participante deve adotar (ex: quando é que deve estar em pé, quando é que deve sentar,...). As orientações são fornecidas

numa linguagem corrente e clara, de forma a não gerar ambigüidades, tornando assim o processo claro e uniforme.

Finalizado o processo de gravação, montagem e edição dos ambientes virtuais, nasce então o ProsoVR, que é a primeira proposta de ferramenta que faz uso da RV para obter amostras de fala pretensamente ecologicamente válidas, com o objetivo de avaliar a prosódia na componente expressiva.

Fase 2 - Recolha de dados

A fase de teste decorreu de forma contínua e individualizada com cada um dos participantes anteriormente descritos no subcapítulo "Participantes", respetivamente na fase 2.

A imersão no ProsoVR decorreu de forma contínua. Os óculos de RV foram colocados pelo próprio com auxílio da autora no início da experiência e só se retiraram após o término da imersão em todos os ambientes. Através de um computador, a equipa de investigação conseguia visualizar simultaneamente a imagem apresentada nos óculos.

O gravador foi acoplado ao bolso do participante e o microfone do mesmo acoplado aos óculos de RV, para garantir maior estabilidade e a mesma distância com fonte durante todo o processo.

Para a recolha de dados propriamente dita, recorreu-se à escala *Slater-Usoh-Steed Questionnaire (SUS)* (Tabela 3), construída por Slater, Usoh e Steed no ano de 1994, Para a análise da validade ecológica (Massigli et al., 2011). Por sua vez, para testar a usabilidade do produto aplicou-se a Escala de Usabilidade do Sistema (Tabela 4) (Martins et al., 2015),

Ambas as escalas foram reunidas num só documento nomeado de Questionário (Apêndice X), e entregues ao participante para preenchimento autónomo e individual. Após o término do preenchimento do mesmo, a autora realizou a entrevista semiestruturada, onde os participantes puderam, também de forma individual, partilhar as suas opiniões sobre a imersão, bem como destacar os pontos fortes e fracos da experiência e as respetivas sugestões de melhoria. As respostas foram registadas pela investigadora através de transcrição manual.

Os dados recolhidos foram aglomerados num só documento, a fim de facilitar a comparação das respostas de todos os peritos no momento da análise.

4 Resultados

O presente capítulo encontra-se subdividido em duas partes essenciais: a) Protótipo do ProsoVR – onde é descrito o protótipo final; b) Usabilidade e Validade Ecológica – onde são apresentados os resultados da aplicação das escalas supracitadas e da entrevista semiestruturada.

4.1 Protótipo do ProsoVR

Como já descrito, o presente projeto surgiu da necessidade de criar as condições necessárias para que as amostras de fala, que serão alvo de avaliação da prosódia emocional, sejam recolhidas num ambiente ecologicamente válido.

O primeiro protótipo do ProsoVR é então composto por quatro ambientes virtuais e cada um é alusivo a uma emoção muito específica, à exceção do primeiro. A imersão virtual com o ProsoVR dura 13:08 minutos.

O primeiro cenário, denominado como “neutro” não tem o objetivo de induzir qualquer tipo de emoção. Serve apenas para a pessoa se ambientar e experienciar toda a dinâmica de uma imersão em ambientes virtuais. Já o segundo, contexto da “alegria”, decorre num café e simula o discurso entre pessoa-alvo e funcionária, onde o cliente acaba por receber um café grátis e mais uns prémios. O contexto três, alusivo à tristeza, simula o ambiente hospitalar onde a pessoa-alvo se encontra a aguardar notícias de um colega que sofreu um acidente, notícias essas que não são positivas, bem pelo contrário. O último contexto, o da raiva, retrata toda uma cena que acontece num quiosque, em que a funcionária que está a atender a pessoa-alvo se torna rude e pouco educada.

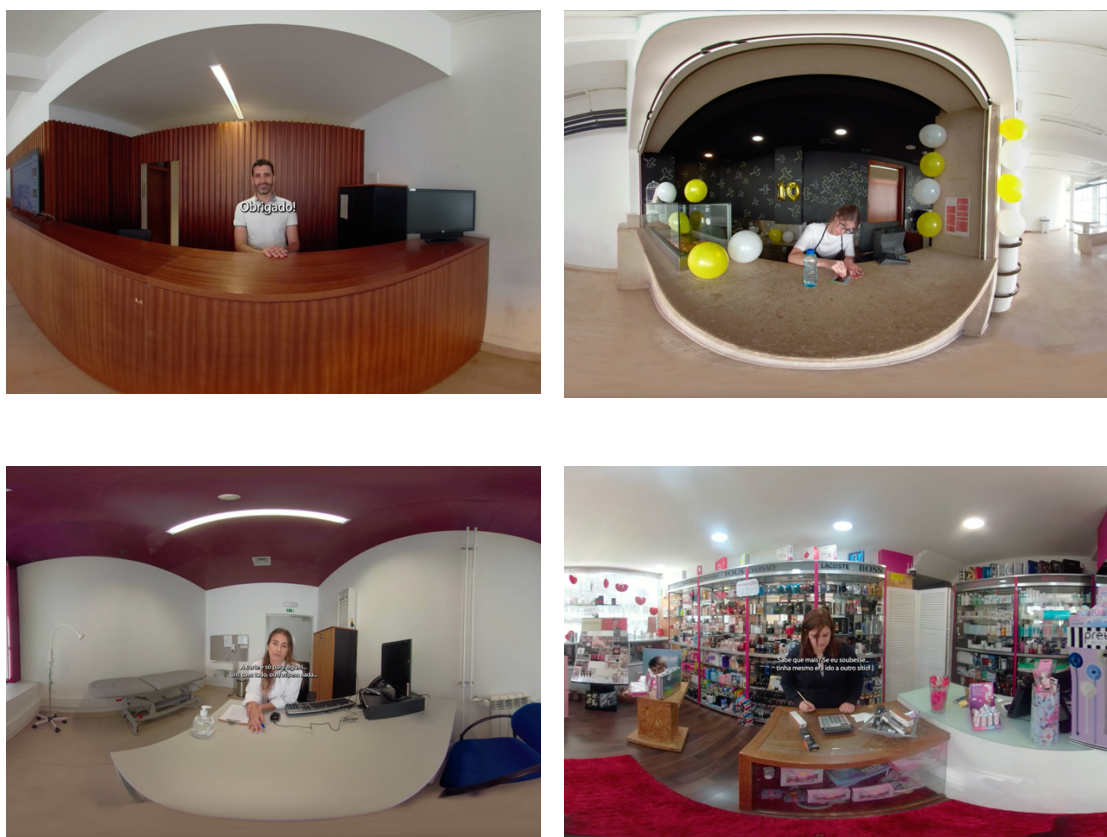


Figura 2 - Ambientes virtuais do ProsoVR

Para que não exista uma fuga do objetivo da imersão virtual, a pessoa em avaliação tem em rodapé aquilo que deve dizer, ou seja, o guião que ela própria deve seguir enquanto experiencia cada um dos contextos.

Com o objetivo de introduzir a emoção de cada cenário, antes do início da imersão em cada ambiente e de forma sincronizada com a narração, surge a palavra (Alegria / Tristeza / Raiva) e o *emoticon* correspondente, para ajudar a antecipar a emoção pretendida.

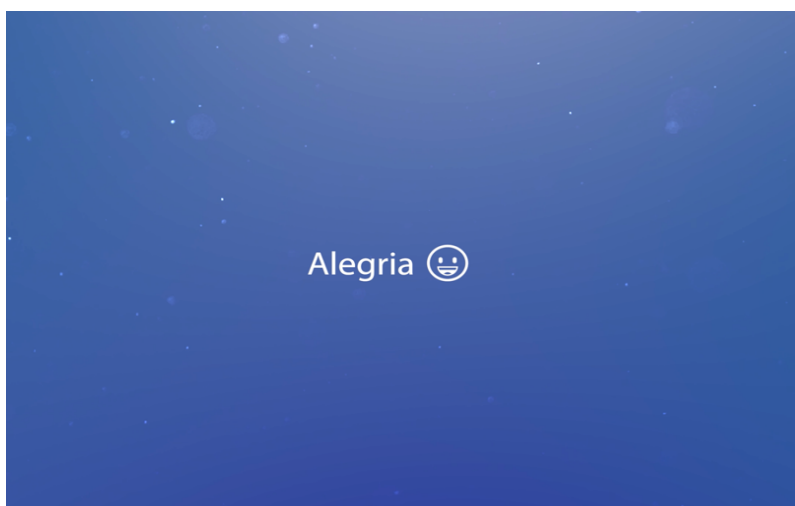


Figura 3 – Exemplo da pista visual para a antecipação do Ambiente 1 "Alegria"

As legendas que o utilizador deve seguir, foram colocadas durante o processo de edição na zona do pescoço do ator, para que a manutenção do contacto ocular com o mesmo seja facilitada e o processo de leitura se torne mais facilitado.

Contexto 0 - Introdução

Como mencionado anteriormente, este contexto não tem como objetivo induzir qualquer emoção. Atendendo ao facto que os participantes possam nunca ter experienciado uma imersão em ambientes virtuais, a equipa de investigação decidiu criar um primeiro momento de imersão que servisse como tutorial. Aqui, os participantes têm tempo para olhar ao seu redor enquanto o rececionista fala ao telefone. Para dar a entender o funcionamento da interação, foram colocadas legendas, no caso para o participante obter informação de onde é o café, que acaba por antecipar e preparar o participante para o cenário seguinte.

Os ambientes virtuais foram gravados nas instalações da Escola Superior de Saúde do Porto, na entrada de um dos edifícios.

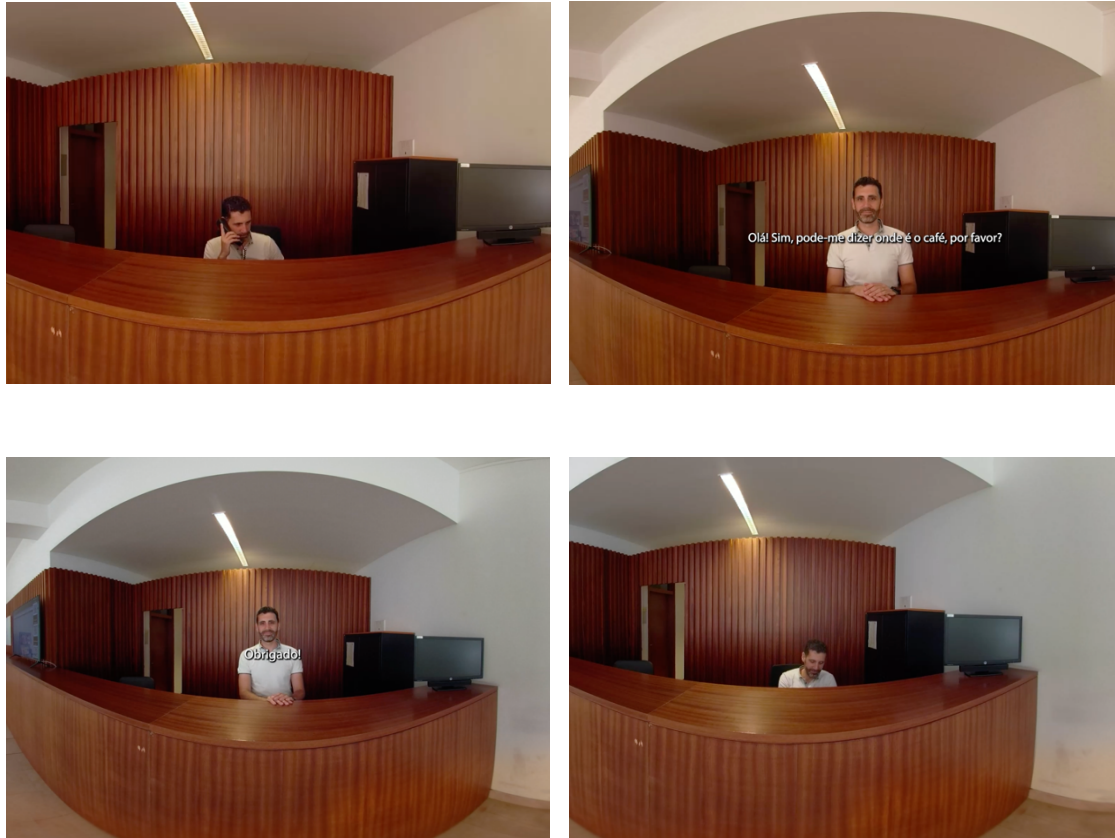


Figura 4 - Imagens representativas do Contexto 0 - Introdução

Contexto 1 - Alegria

Os ambientes virtuais deste contexto foram gravados nas instalações da Escola Superior de Saúde do Politécnico do Porto, no local do bar, por mais se assemelhar ao contexto real de um café já que possuía balcão, vitrines, máquina registadora.

Este contexto visa induzir a emoção de alegria. Decorre num café e pretende simular o discurso entre o participante e a funcionária de um café que está em celebrações e, por isso, todo o cenário foi decorado com balões alusivos à celebração. Logo na primeira interação, a funcionária recebe o participante com alegria, já que é o cliente número cem e isso lhe traz recompensas especiais, nomeadamente a oferta do seu pedido e uma raspadinha. Uma vez que na interação o participante não pode manipular objetos, a funcionária oferece-se para raspar a raspadinha e demonstra espanto e entusiasmo ao perceber que foi atribuído um valor monetário elevado. A interação desenrola-se neste enquadramento, sendo as frases alvo introduzidas num contexto emocional esperado de alegria.



Figura 5 - Imagens representativas do Contexto 1 - Alegria

Contexto 2 - Tristeza

Este contexto pretende induzir a emoção de tristeza. Decorre num hospital em que o paciente recebe a notícia de que o seu colega se encontra com um estado de saúde agravado após ter tido um acidente nessa manhã, após saírem juntos para caminhar.

Os ambientes virtuais foram gravados nas instalações da Escola Superior de Saúde do Politécnico do Porto, na clínica pedagógica, uma vez que já possuía as condições físicas e materiais ideais para o cenário pretendido.

De forma a antecipar o cenário e, de alguma forma, quebrar a alegria sentida no contexto anterior, o Contexto 2 inicia numa sala de espera em que, durante cerca de 30 segundos, o participante possa entrar na personagem e sentir o peso da espera. Posteriormente a médica vem chamá-lo à sala de espera e, dá-se uma transição de fade out/fade in para o consultório médico.

Já no consultório a médica transmite a notícia, em que através da expressão não-verbal atribui ainda mais carga negativa à informação que está a passar. Para este papel, a personagem foi criteriosamente escolhida de acordo com a sua experiência em comunicação clínica. Durante as filmagens o texto previamente validado foi ajustado de forma a aumentar a naturalidade e o realismo da situação retratada.



Figura 6 - Imagens representativas do Contexto 2 - Tristeza

Contexto 3 - Raiva

O Contexto 3 procura induzir raiva e pretende simular o discurso com uma funcionária de uma loja local de perfumes e cosmética, onde o participante tenta comprar um creme hidratante, sem sucesso.

O ambiente virtual foi gravado numa loja real para melhor simular o ambiente, já que tinha todo o cenário, balcão, produtos e outros adereços essenciais para que a imersão se assemelhasse o mais possível com a realidade.

Quanto à experiência neste contexto, logo no início da interação, a personagem responde de forma rude dizendo “diga lá, não vê que estou ocupada?” e por diversas vezes ignora a presença do interlocutor. A interação com a funcionária permanece difícil, com revirar de olhos por parte da mesma, até que entra um cliente habitual que recebe tratamento especial, o que pretende gerar desconforto e indignação.

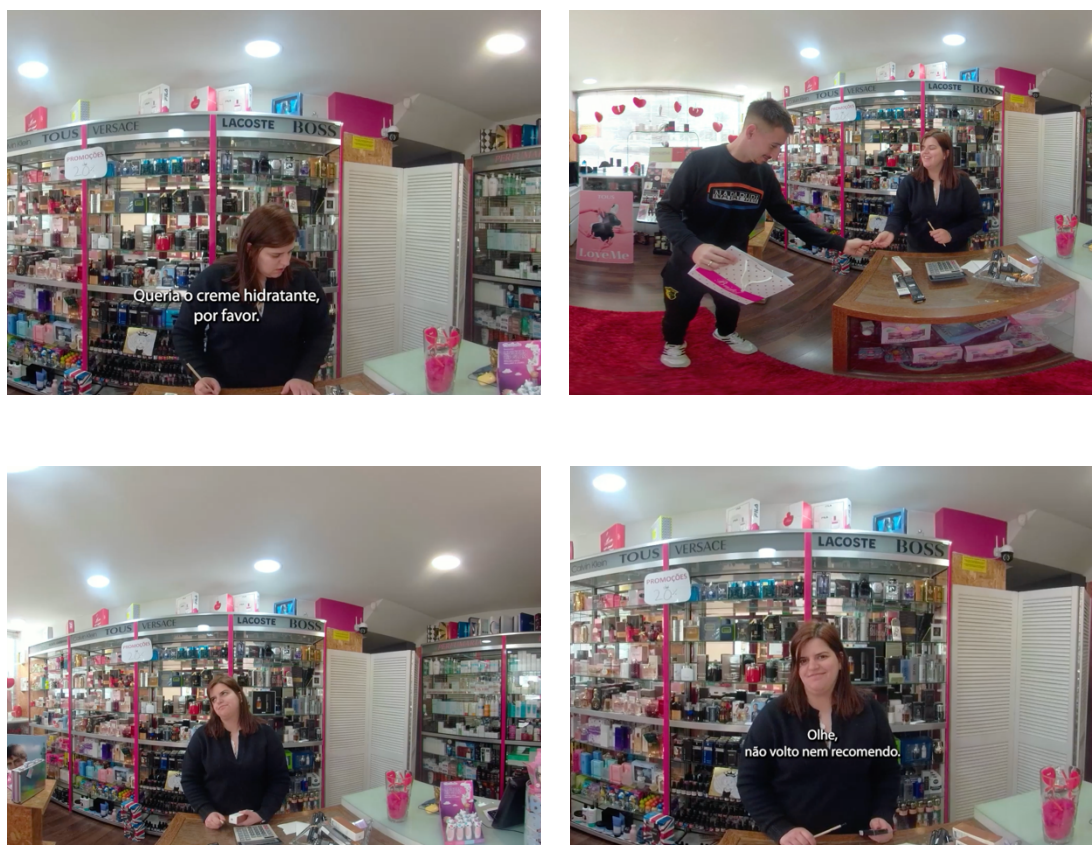


Figura 7 - Imagens representativas do Contexto 3 - Raiva

4.2 Usabilidade e Validade Ecológica

Na fase de teste do protótipo do ProsoVR participaram 6 pessoas, encontrando-se no quadro seguinte características do grupo. Doravante os participantes serão mencionados como TF1, TF2, TF3, UC1, UC2 e UC3, com o código "TF" correspondente a terapeuta da fala e "UC" correspondente a utilizador comum.

Tabela 5- Caracterização dos participantes

	UC1	UC2	UC3	TF1	TF2	TF3
Sexo	Fem	Fem	Masc	Fem	Fem	Fem
Idade	20-35	20-35	20-35	20-35	51-65	36-50

Os 6 participantes responderam a um Questionário (Apêndice 1) que engloba as duas escalas previamente apresentadas. O questionário tem como objetivo de caracterizar a sensação de presença e usabilidade do produto.

Os participantes responderam às questões apresentadas na *Slater-Usuh-Steed Questionnaire (SUS)* (Massigli et al., 2011) com a atribuição de uma classificação de 1 a 7, mediante a sensação de

presença experienciada durante a imersão. Contudo, foi necessária a adaptação das descrições das pontuações de modo a clarificar e facilitar o momento de avaliação. Para isso, foram adicionados os termos utilizados na própria questão com o objetivo de enfatizar a correspondência de cada pontuação (Apêndice 1).

Assim, a pontuação de cada participante nesta escala poderia variar num mínimo de 6 e num máximo de 42 pontos, valor a que se atribuiu o nome de *score* total. Estes resultados foram ainda convertidos em percentagem (%) de forma a representar o mínimo (0%, equivalente ao score de 6 pontos) e o máximo (100%, equivalente ao score de 42 pontos) de sensação de presença. Os resultados obtidos são apresentados na Tabela 6.

Tabela 6- Resultados da escala Slater-*Usch*-*Steed Questionnaire (SUS)*(Massigli et al., 2011)

Questões/Participantes	Sensação de presença						Média
	UC1	UC2	UC3	TF1	TF2	TF3	
Q1. Até que ponto se sentiu presente no cenário apresentado?	6	4	7	5	7	6	5,833
Q2. Até que ponto o cenário apresentado se tornou uma realidade e quase se esqueceu do contexto laboratorial em que decorre a experiência?	4	4	7	5	5	5	5
Q3. Até que ponto ao recordar-se do cenário, o vê como uma imagem, ou como um sítio onde esteve presente?	6	5	7	6	6	5	5,833
Q4. Durante o tempo da experiência, foi mais forte o sentido de estar no cenário ou de estar em outro lugar?	6	6	7	7	6	7	6,5
Q5. Em que medida as condições estruturais (cor, objetos) do cenário o recordam de outras situações semelhantes em que já esteve?	7	6	7	7	7	7	6,833
Q6. Durante algum momento da experiência pensou que estivesse realmente no cenário?	6	5	7	4	5	3	5
Score total	35	30	42	34	36	33	35
%	80,56	66,67	100	77,78	83,33	75,00	80,56

Os resultados mostraram uma ampla gama de respostas, o que indica que a sensação de presença é uma experiência individualizada.

Quanto à análise geral dos resultados obtidos, é possível inferir que as pontuações foram globalmente media-altas/altas, não sendo atribuída nenhuma classificação dos valores mínimos 1

e 2. Os scores totais variaram entre 30 e 42, com uma média de 35. Por sua vez, numa representação percentual, os valores variaram entre 77,78% e 100%, com uma média de 80.56%. Numa análise mais detalhada observou-se que o parâmetro “Em que medida as condições estruturais (cor, objetos) do cenário o recordam de outras situações semelhantes em que já esteve?” obteve a maior pontuação, com uma média classificativa de 6,833, seguido do parâmetro “Durante o tempo da experiência, foi mais forte o sentido de estar no cenário ou de estar em outro lugar?” que obteve uma classificação de 6,5.

Por sua vez, os parâmetros “Até que ponto o cenário apresentado se tornou uma realidade e quase se esqueceu do contexto laboratorial em que decorre a experiência?” e “Durante algum momento da experiência pensou que estivesse realmente no cenário?” apresentaram as pontuações mais baixas, contudo com uma média classificativa média-alta de 5 pontos.

Em suma, os valores apresentados sugerem que, em média, os participantes relataram uma sensação de presença elevada durante a experiência imersiva com o ProsoVR.

Por sua vez, a análise da usabilidade do produto é caracterizada com base na eficácia e eficiência da prestação do produto, e na satisfação do utilizador durante a sua aplicação. Seguem-se, por isso, os resultados recolhidos através da aplicação da Escala de Usabilidade do Sistema (Martins et al., 2015).

Para permitir comparabilidade dos itens e, uma vez que existem itens com conotação negativa, inverteu-se a cotação das respetivas afirmações, encontrando-se as mesmas assinaladas com*. Assim, foi possível calcular um *score* total para a usabilidade, estando um *score* mais elevado associado a uma maior usabilidade do produto para o participante.

Os resultados obtidos são apresentados na Tabela 7.

Tabela 7- Resultados da aplicação da Escala de Usabilidade do Sistema (Martins et al., 2015)

Afirmações/Participantes	Usabilidade						Média
	UC1	UC2	UC3	TF1	TF2	TF3	
A1. Acho que gostaria de utilizar este produto com frequência.	5	3	5	5	5	4	4,5
A2. Considerei o produto mais complexo do que necessário.*	4	3	5	5	4	4	4,167
A3. Achei o produto fácil de utilizar.	3	3	5	5	4	5	4,167
A4. Acho que necessitaria de ajuda de um técnico para conseguir utilizar este produto.*	2	1	5	4	4	5	3,5

A5. Considerei que as várias funcionalidades deste produto estavam bem integradas.	4	4	5	5	1	5	4
A6. Achei que este produto tinha muitas inconsistências.*	5	5	5	4	2	5	4,333
A7. Suponho que a maioria das pessoas aprenderia a utilizar rapidamente este produto.	3	3	4	5	4	4	3,833
A8. Considerei o produto muito complicado de utilizar.*	5	4	5	5	5	5	4,833
A9. Senti-me muito confiante a utilizar este produto.	4	5	4	4	4	5	4,333
A10. Tive que aprender muito antes de conseguir lidar com este produto.*	4	5	5	5	5	5	4,833
Score total	39	36	48	47	38	47	42,5
%	72,5	65	95	92,5	70	92,5	81,25

Numa leitura geral, os valores apresentados traduzem-se num *score* médio de usabilidade de 42,5, com um desvio padrão de 5,39. Isso sugere que, em média, os participantes tendem a avaliar positivamente a usabilidade do ProsoVR, com uma média percentual de 81,25% de usabilidade.

Numa leitura mais detalhada observa-se que as afirmações “Acho que necessitaria de ajuda de um técnico para conseguir utilizar este produto” e “Considerei que as várias funcionalidades deste produto estavam bem integradas” suscitaram maior discrepância de valores, uma vez que ambas foram cotadas com as classificações mínima (1) e máxima (5).

Por sua vez, as afirmações “Considerei o produto muito complicado de utilizar”, “Senti-me muito confiante a utilizar este produto” e “Tive que aprender muito antes de conseguir lidar com este produto” obtiveram classificações mais sólidas, com pontuações sem variações entre extremos.

Em suma, os resultados sugerem que o ProsoVR possui uma boa usabilidade, já que os participantes demonstraram sentir-se confiantes durante a utilização do protótipo e consideram que o mesmo não foi complicado de utilizar. No entanto, importa ressaltar que algumas componentes específicas possam beneficiar com ajustes, nomeadamente no apoio técnico e na revisão das funcionalidades do mesmo, de forma a aprimorar ainda mais o protótipo e, assim, aumentar a satisfação dos participantes durante a utilização do mesmo.

Relativamente à entrevista semiestruturada, foram recolhidas informações e *feedback*, de forma individual, sobre a experiência imersiva, bem como foram levantados os pontos fortes e fracos do protótipo.

De forma geral, os participantes mostraram satisfação com a experiência destacando as seguintes categorias: envolvimento e sensação de presença. O resumo das respostas está descrito na Tabela 8.

Tabela 8 – Resumo das respostas sobre a satisfação geral

Categoria	Exemplo de resposta
Envolvimento (2)	«[a experiência] foi muito interessante (...)» (TF1) «O que gostei mais do realismo [da experiência]» (TF2) «(...) estava totalmente envolvida no ambiente (...) foi muito giro, está uma ideia mesmo boa!» (UC1) «[a experiência] foi muito agradável» (UC2)
Sensação de presença (2)	«(...) [gostei] da sensação de estar lá (...)» (UC1) «(...)senti-me mesmo lá (...)» (TF1).
Voz off/orientações (1)	«(...) as introduções estavam boas» (TF1)

Os participantes destacaram como pontos fortes do protótipo os seguintes aspectos: personagens e sensação de presença. O resumo das respostas está descrito na Tabela 9.

Tabela 9- Resumo das respostas sobre os pontos fortes

Categoria	Exemplo de resposta
Sensação de presença (4)	«(...) senti-me mesmo lá» (TF2) «(...) [gostei mais] da sensação de estar lá, ver o todo, nos jogos não acontece isso e sinto que não fico totalmente envolvida no ambiente e aqui senti isso» (UC1) «(...) senti [no contexto 3] que a mulher falou mesmo para mim! » (UC2) «(...) a loja, parecia mesmo que estava na loja! » (UC2)
Personagens (2)	«[no contexto 2] a própria atriz demonstrar tristeza» (UC3) «[no contexto 1] senti mesmo alegria mesmo por causa da atriz» (UC3)

Relativamente aos pontos fracos do protótipo, os participantes destacaram os seguintes aspectos: legendas, frases-chave, sonoridade e efeitos secundários. O resumo das respostas está descrito na Tabela 10.

Tabela 10 – Resumo das respostas sobre os pontos fracos

Categoria	Exemplo de resposta
-----------	---------------------

Legendas (4)	<p>«(...) tiram a espontaneidade(...)» (UC1)</p> <p>«(...) o ter que ler as respostas condiciona (...)» (UC2)</p> <p>«(...) [as legendas] foram mais rápido, precisava de mais tempo(...)» e «(...) [se pudesse] mudava as minhas falas porque numa situação real nunca diria aquilo» (UC3) relativamente ao Cenário 3</p> <p>«...[as legendas] estavam com desfasamento, tinham intervalos grandes(...)» (TF1) relativamente ao Contexto 1</p>
Sonoridade (3)	<p>«(...) [no Contexto 2] ouviu-se um bocadinho baixo» (UC2)</p> <p>«(...) [no Contexto 1 seria importante] adequar os silêncios(...)» (TF3).</p> <p>«(...) [no geral] o som estava baixo» (UC2)</p>
Frases-chave (1)	<p>«...(...) [a frase "uns com tudo e outros com nada"] pode desencadear uma quebra na melodia [no cenário de alegria] porque parece que leva a uma reflexão(...)» (TF3)</p>
Efeitos secundários (1)	<p>«(...) [no Contexto 2] deu-me vertigens (...) devia ter mesmo uma mesa de apoio onde eu pudesse pôr as mãos(...)» (TF1).</p>

No que concerne à indução da emoção, o Contexto 3 (Raiva) teve destaque como o mais eficaz. Por sua vez, o Contexto 1 (Alegria) surge como o menos capaz na indução da emoção proposta. O resumo das respostas está descrito na Tabela 11 por ordem de contextos.

Tabela 11- Resumo das respostas sobre o contexto que melhor e pior induziram a emoção

Contexto	Exemplo de resposta
Contexto 1 (Alegria)	<p>«(...) forçado (...) [a interação] seria mais natural sem tanto entusiasmo [da personagem]» (TF3)</p> <p>«(...) como [o Cenário 1] foi o primeiro ainda estava muito consciente que não era real» (TF2)</p>
Contexto 2 (Tristeza)	<p>«(...) senti mesmo tristeza porque pensei num amigo meu» (TF2)</p>
Contexto 3 (Raiva)	<p>«(...) [fazia sentir] que a mulher falou diretamente para mim(...) e apesar de ser fictício acaba por ser real porque já aconteceu, a situação foi parecida com algo que já aconteceu comigo» (UC2)</p> <p>«(...) foi o que me pareceu estar mais no sítio» (UC1)</p>

5 Discussão

O desenvolvimento do presente estudo surge como necessidade de construção de um protótipo que permitisse criar um ambiente de recolha de amostras de fala ecológicamente válidas para se avaliar a prosódia emocional dessas mesmas amostras na vertente da expressão.

O estudo propôs-se a estudar/desenvolver três tópicos gerais – construir o protótipo do ProsoVR e estudar a validade ecológica e usabilidade inerente ao mesmo – com vista a identificar pontos de melhoria da versão alfa e avançar para a fase seguinte do processo de desenvolvimento deste instrumento.

Relativamente ao protótipo, os ambientes virtuais foram gravados e editados e encontram-se passíveis de ser utilizados em qualquer espaço físico, desde que os óculos de RV possuam ligação à internet. Note-se que a experiência é possível com equipamentos de outras marcas e modelos. A experiência virtual com o ProsoVR dura 13:08 minutos, o que vai ao encontro do que a bibliografia sugere como espaço de tempo recomendado para as sessões de RV, ser inferior a 60 minutos (Kourtesis et al., 2020).

O ProsoVR ambiciona incrementar o realismo das avaliações em TF, de modo a complementar as avaliações tradicionais já validadas para PE. Por isso a validade ecológica do protótipo foi estudada no sentido de perceber em que medida os ambientes virtuais proporcionam uma interação semelhante a uma situação conversacional humana (Massigli et al., 2011). Para isso, foi importante analisar as respostas dos participantes e medir a sensação de presença sentida pelos mesmos durante a imersão virtual.

A sensação de presença é uma variável contínua, oscilando durante uma mesma experiência o que leva a atingir valores diferentes aquando da imersão no ambiente virtual (Slater *et al.*, 2009 citado por (A. Ribeiro & Monteiro, 2015) mas a sua classificação deve feita de forma geral. Isto é, a experiência deve ser avaliada como um todo, independentemente dos picos de sensação de presença que possa apresentar.

O facto de o mesmo ter sido construído com a colaboração de atores reais em locais reais, promove uma interação com o utilizador mais próxima daquilo que seria uma situação quotidiana (Manuel & Pinto, 2015), o que por sua vez aumenta os níveis validade ecológica.

O Contexto 3 surge nos resultados como o mais eficaz na indução da emoção.

O que diferencia este dos restantes contextos foi o facto de este ter sido gravado numa loja local, em que todos os adereços eram efetivamente reais e materiais da loja, ao passo que os restantes cenários foram gravados em ambientes simulados. Apesar de ter sido o primeiro a gravar, o que trouxe alguns desafios técnicos, as condições gerais das gravações repercutem-se na falta de

homogeneidade dos contextos, uma vez que não foram construídos de igual forma. A título exemplificativo, no Contexto 1 que pretende simular um café, o facto de não existir o ruído de fundo típico (p.e. máquinas em funcionamento, conversas, música, etc.) ou movimento de outros intervenientes, faz ampliar a artificialidade do contexto. Compreende-se que a introdução de estímulos desta natureza no ambiente virtual, poderá vir a potenciar a sensação de presença que a participante experiencia durante a imersão.

Os resultados quantitativos demonstraram elevado nível percentual de sensação de presença: 80,56%. Sendo este um dos indicadores para a determinação da validade ecológica, pode-se aferir que o ProsoVR é um produto que permite ao utilizador experienciar uma imersão ecologicamente válida.

Importa, no entanto, ressaltar que um dos participantes atribuiu a pontuação máxima em todos os parâmetros da avaliação da sensação de presença. Atendendo ao número participantes do estudo, a equipa achou importante estudar quais seriam os resultados se não fosse contabilizada a pontuação deste mesmo participante, para averiguar se o nível de sensação de presença seria igualmente elevado. O nível percentual dos resultados da sensação de presença sem os dados acima referidos é de 75,00%, o que é um indicador igualmente favorável.

Assim, apesar de estar documentada a existência de alguma resistência ao uso das novas tecnologias (Wanderley et al., 2018), a experiência imersiva com o ProsoVR foi vista com agrado por todos os participantes, de acordo com os resultados obtidos na entrevista. Importa, por isso, ressaltar o facto de o grupo de participantes ser constituído por utilizadores comuns e terapeutas da fala. Estarão, por isso, os terapeutas da fala recetivos à implementação de novas tecnologias nas suas práticas clínicas?

Para perceber e testar essa recetividade à aplicação das novas tecnologias, nomeadamente a RV e o ProsoVR nas suas práticas clínicas, será necessário testar o protótipo com um número de participantes mais significativo, já que poderá haver viés de seleção, pela amostragem de conveniência, nesta que é uma versão Alfa do protótipo. Num estudo posterior, a equipa de investigação necessitará de desenvolver a versão Beta do ProsoVR, em que poderá ser usada num número maior de utilizadores e poderá haver comparação de grupos e análise de comportamentos ou atitudes específicas.

Os resultados para análise da usabilidade sugerem que, no geral, o grupo apresenta uma média percentual alta (Fonsêca RO, *et al*, 2015 & World Health Organization, 2011 citado por Scanferla, 2022). Segundo um estudo realizado em Portugal sobre as iniciativas *mhealth* no país, elevados níveis de usabilidade estão, por norma, associados a sistemas/produtos que se encontram em "... fases estabelecidas" (Duque, 2017) o que vai de encontro aos valores previamente apresentados e, por isso, assume-se como um indicador positivo da presente investigação.

No entanto, seria benéfico realizar um estudo que posicionasse os terapeutas como utilizador do ProsoVR na perspectiva de técnicos para obtenção de dados mais concretos para o âmbito do uso do ProsoVR a nível clínico.

Neste sentido, foi possível gravar sinais de fala com qualidade para posterior análise áudio-percetiva e acústica. Contudo, estas análises não foram efetivamente realizadas neste estudo, por não estar previsto nos seus objetivos. Assim, será importante, no futuro, não apenas estudar a usabilidade da experiência em realidade virtual, mas também o produto final obtido, ou seja, a qualidade das amostras de fala obtidas, a sua capacidade de predição das emoções retratadas e a sua capacidade de discriminar/caracterizar pessoas com e sem perturbações da comunicação. Para tal, poderá recorrer-se a métodos mais comuns, como a avaliação áudio-percetiva ou a análise acústica, ou a métodos mais avançados, nomeadamente o uso de ferramentas de inteligência artificial.

Durante o desenvolvimento do presente estudo surgiram algumas dificuldades de carácter organizacional e logístico, que resultaram no atraso da produção do protótipo o que, por sua vez, prejudicou o cronograma inicialmente planeado. Nesta medida, a fase de teste foi condicionada. Assim, o facto de o teste ter sido aplicado a um número reduzido de pessoas é uma limitação dos resultados, ainda que se assuma tratar-se de uma versão Alfa. Como tal, de futuro será importante testar o protótipo junto de um número maior de participantes bem como deve ser feita a comparação da perceção do utilizador comum *vs* terapeuta da fala.

Outra das limitações deste estudo passa pela avaliação da validade ecológica, que se baseou apenas no estudo do nível de realismo e análise eclética (Massigli et al., 2011), pelo que seria importante efetuar a análise da união (Massigli et al., 2011) de modo a obter uma avaliação de todos os critérios da validade ecológica. A análise da união ajuda a manter a fiabilidade e a validade interna dos resultados, através da manipulação da variável em estudo (no caso, a prosódia emocional na vertente expressiva) e comparação com o grupo de controle.

O protótipo também pode beneficiar de ajustes. Uma das sugestões que os participantes apresentaram durante a entrevista de feedback passaria pela alteração da interação com os personagens representados no ProsoVR, de forma que estes oferecessem respostas diferentes mediante a interação do utilizador. Tal seria possível através da incorporação de dispositivos de entrada como microfones ou acesso aos comandos remotos, o que permitiria experienciar sensações de realismo mais ajustadas e proporcionar ao utilizador uma participação mais ativa. Mecanismos de controlo por parte do avaliador (terapeuta da Fala) ou de inteligência artificial, poderiam também contribuir para solucionar a questão de os tempos de resposta se encontrarem desalinados, uma vez que se forem introduzidos dispositivos de entrada a resposta do personagem já irá surgir apenas após a interação do utilizador. Contudo, estes elementos

umentam substancialmente a complexidade na execução do protótipo e a programação subjacente.

Outra das alterações que o protótipo pode beneficiar, esta mais simples, passa pelos ajustes do volume, de forma a uniformizar a sonoridade entre todos os cenários. No mesmo sentido, na versão poderá ser pertinente rever o guião e a produção do contexto 1 (alegria), já que foi o que encontrou menor sensação de realismo e indução da emoção por parte dos participantes.

Estas alterações serão cruciais para o sucesso da construção da versão melhorada deste primeiro protótipo, a versão Beta, a qual deverá ser submetida a novos testes. No mesmo sentido, e considerando que estes ambientes de realidade virtual foram desenvolvidos para induzir traços emocionais na fala, será importante estudar se as amostras de fala representam e discriminam de forma eficiente as emoções pretendidas. Para tal, será também importante explorar diferentes métodos de avaliação da prosódia que possam ser utilizados na prática clínica dos terapeutas da fala.

6 Conclusão

O presente estudo procurou demonstrar como a realidade virtual pode ser uma alternativa válida na criação de soluções terapêuticas inovadoras em TF, neste caso, aplicada à área da prosódia. Os resultados obtidos sugerem que o protótipo (versão Alpha) do ProsoVR possui elevados níveis de validade ecológica e usabilidade e, por isso, tem-se que os objetivos gerais para este estudo foram cumpridos. Não obstante, identificaram-se pontos-chave que necessitam de atenção para o desenvolvimento da versão Beta.

De futuro, será importante realizar estudos que contemplem uma amostra maior de participantes, onde seja destacada a opinião dos terapeutas da fala, tal como efetuar o estudo sobre qual a melhor forma de enquadrar o ProsoVR numa prática clínica comum.

Em suma, o protótipo do ProsoVR e o estudo envolvente demonstram-se promissores no que concerne à inovação das ferramentas de avaliação da prosódia a nível nacional. A atitude globalmente positiva demonstrada pelos participantes sugere uma possível aceitação na implementação de novas tecnologias na prática clínica na área da TF, que podem ser aplicadas ao nível da avaliação e exploradas as possibilidades para a aplicação também em intervenção.

Referências Bibliográficas

- Alexandre, N. M. C., & Coluci, M. Z. O. (2011). Content validity in the development and adaptation processes of measurement instruments. *Ciência & Saúde Coletiva*, 16(7), 3061–3068.
- Behlau, M. (2004). *Voz: O livro do Especialista I*. Revinter.
- Bodolay, A. N., & Forte, G. N. (2023). *Não confusa prosódia com pronúncia: fundamentos de prosódia para PLE* (Editora CRV Ltda, Ed.).
- Boutsen, F. (2003). Prosody: The Music of Language and Speech. *The ASHA Leader*, 8(4), 6–8. <https://doi.org/10.1044/leader.FTR1.08042003.6>
- Bowers D. Florida Affect Battery: University of Florida; 1991.
- Bryant, L., Brunner, M., & Hemsley, B. (2020). A review of virtual reality technologies in the field of communication disability: implications for practice and research. Em *Disability and rehabilitation. Assistive technology* (Vol. 15, Número 4, pp. 365–372). NLM (Medline). <https://doi.org/10.1080/17483107.2018.1549276>
- Burdea, G. C., & Coiffet, P. (2003). *Virtual Reality Technology*. Wiley-Interscience.
- Carlos, D. de A. O., Magalhães, T. de O., Filho, J. E. V., da Silva, R. M., & Brasil, C. C. P. (2016). Concepção e Avaliação de Tecnologia mHealth para promoção da saúde vocal. *RISTI – Revista Iberica de Sistemas e Tecnologias de Informacao*, 19, 46–60. <https://doi.org/10.17013/risti.19.46-60>
- Carvalho, A. (2013). *Patologia da Comunicação Humana e Intervenção Terapêutica: Algumas Reflexões Teóricas e Metodológicas*. Centro Hospitalar de Lisboa Central, EPE.
- Carvalheda, G. (2021, Novembro). *Speech Immersion Hub*. Centro de Tratamento de Gaguez. <https://www.centrotratamentogaguez.pt/post/speech-immersion-hub>
- Catecati, T., Gomes Faust, F., Roepke, G. A. L., Araujo, F. S., Albertazzi, D., Ramirez, A. R. G., & Ferreira, M. G. G. (2018). *Métodos Para a Avaliação da Usabilidade no Design de Produtos*. <https://doi.org/10.5965/1808312906082011564>
- Chen, Y. P. P., Johnson, C., Lalbakhsh, P., Caelli, T., Deng, G., Tay, D., Erickson, S., Broadbridge, P., el Refaie, A., Doube, W., & Morris, M. E. (2016). Systematic review of virtual speech therapists for speech disorders. Em *Computer Speech and Language* (Vol. 37, pp. 98–128). Academic Press. <https://doi.org/10.1016/j.csl.2015.08.005>
- da Silva, M. F. G. (2011). *Avaliação da Comunicação Não-Verbal em Pacientes com Doença de Parkinson: Reconhecimento da Emoção de Faces, Gestos e Prosódia* [Programa de Pós-Graduação em Medicina: Ciências Médicas]. Faculdade de Medicina.
- da Silva, M. I. R. (2021). *Contributo para o desenvolvimento do ProsoVR: estudo preliminar de validade de protocolos de avaliação audio-percetiva e acústica da prosódia emocional* [Projeto de Investigação para a obtenção de Grau de Licenciatura]. Politécnico do Porto.

- Delft, M. van. (2018). *Prosody in speech before and after a total laryngectomy* [MSc Linguistics Speech and Language pathology]. Radboud Universiteit Nijmegen.
- Dias, C. O. (2019). *Cuidados Paliativos: O Papel do Terapeuta da Fala na Promoção da Funcionalidade Comunicativa de Utentes Adultos*. Universidade do Porto.
- Dias, M. C. (2022). *Contributo para o desenvolvimento do ProsoVR: pré-teste em casos de Voz Alaríngea*. Politécnico do Porto.
- Donga, J., Gomes, P. V., Marques, A., Pereira, J., & Azevedo, J. (2020). *Application of Adaptive Virtual Environments Through Biofeedback for the Treatment of Phobias*. 42. <https://doi.org/10.3390/proceedings2020054042>
- Dores, A. R., Barbosa, F., Marques, A., Carvalho, I. P., De Sousa, L., & Castro-Caldas, A. (2012). Realidade Virtual na Reabilitação: Por Que Sim e Por Que Não? Uma Revisão Sistemática. *Acta Médica Portuguesa*, 25(6), 414–421.
- Duque, C. M. S. (2017). *Iniciativas de mHealth em Portugal* [Dissertação para a obtenção do grau de Mestre]. Universidade Aberta | Instituto Superior Técnico.
- Enderby PM, Palmer R. FDA-2: Frenchay Dysarthria Assessment–Second Edition. 2nd ed. Austin, TX: Pro-Ed; 2008.
- Fernandes, F. D. (2017). *Tratamento de Fobias com Ambientes Virtuais*.
- Ferreira, T. (2022). *Contributo para o desenvolvimento do ProsoVR: pré-teste em casos de Parkinsonismo*. Politécnico do Porto.
- Filipe, & Vicente. (2011). *Teste de Competências Prosódicas para Falantes do Português Europeu*. 447–463.
- Franco, M. R. (2022). *Contributo para a criação de um Protocolo de Avaliação da Prosódia no contexto de Terapia da Fala*. Politécnico do Porto.
- Freitag, R. M. Ko., & Lucente, L. (2017). *Prosódia da fala : pesquisa e ensino*. Edgard Blucher Ltda.
- Freitas, B. (2020). *Avaliação da Prosódia pelos Terapeutas da Fala em Portugal*.
- Galdino, J. C., & Oliveira JR, M. (2023). Prosódia e síntese da fala: uma revisão integrativa da literatura. *Revista Abralín*, 1–15. <https://doi.org/10.25189/rabralin.v.22i1.2130>
- Gramuglia, A. C. J. (2013). *Parâmetros vocais perceptivo-auditivos e acústicos em crianças com nódulos vocais*.
- Guimarães, Inês (2023). Contributo para o desenvolvimento do ProsoVR em normofalantes e profissionais de voz de elite: avaliação acústica da prosódia emocional.
avaliação acústica da prosódia emocional
- Hawthorne, K., & Fischer, S. (2020a). Speech–language pathologists and prosody: Clinical practices and barriers. *Journal of Communication Disorders*, 87. <https://doi.org/10.1016/j.jcomdis.2020.106024>

- Hawthorne, K., & Fischer, S. (2020b). Speech-language pathologists and prosody: Clinical practices and barriers. *Journal of Communication Disorders*, *87*.
<https://doi.org/10.1016/j.jcomdis.2020.106024>
- Hoekert, M., Vingerhoets, G., & Aleman, A. (2010). Results of a pilot study on the involvement of bilateral inferior frontal gyri in emotional prosody perception: an rTMS study. *BMC Neuroscience*, *11*(93).
- Holmes, D. P. W. (2018). *Avaliação da Atenção com recurso ao Teste da Galeria Virtual*. Universidade Lusófona de Humanidades e Tecnologias.
- Jorge, A. C. A. (2020). Prosódia afetiva na esquizofrenia. *Estudos Linguísticos (São Paulo. 1978)*, *49*(3), 1393–1412. <https://doi.org/10.21165/el.v49i3.2567>
- Kerr, M. de S. (2012). *Adaptação da Bateria Montreal de Avaliação da Comunicação (Versão Brasileira; MAC-BR) para o Português Europeu*.
- Kourtesis, P., Korre, D., Collina, S., Doumas, L. A. A., & MacPherson, S. E. (2020). Guidelines for the Development of Immersive Virtual Reality Software for Cognitive Neuroscience and Neuropsychology: The Development of Virtual Reality Everyday Assessment Lab (VR-EAL), a Neuropsychological Test Battery in Immersive Virtual Reality. *Frontiers in Computer Science*, *1*. <https://doi.org/10.3389/fcomp.2019.00012>
- Ladd, D. R., Couto, T. de L. R., & Seara, I. C. (2019). O que é prosódia? *Working Papers em Linguística*, *20*(1), 8–46. <https://doi.org/10.5007/1984-8420.2019v20n1p8>
- Lazzari, M., & Petry, I. (2020). Disfluência e narratividade: a importância do olhar prosódico sobre a aquisição da linguagem. *Revista Abralín*, *19*(2), 1–5.
<https://doi.org/10.25189/rabralin.v19i2.1638>
- Lee, S. A. S. (2019). Virtual Speech-Language Therapy for Individuals with Communication Disorders: Current Evidence, Limitations, and Benefits. Em *Current Developmental Disorders Reports* (Vol. 6, Número 3, pp. 119–125). Springer. <https://doi.org/10.1007/s40474-019-00169-7>
- Lopes, L. W., & Lima, I. L. B. (2014). Prosódia e transtornos da linguagem levantamento das publicações em periódicos indexados entre 1979 e 2009. *Revista CEFAC*, *651–659*.
- Manuel, C., & Pinto, N. (2015). *Estratificação da interação com o utilizador em ambientes de Realidade Virtual*. Politécnico do Porto.
- Martins, A. I., Rosa, A. F., Queirós, A., Silva, A., & Rocha, N. P. (2015). European Portuguese Validation of the System Usability Scale (SUS). *Procedia Computer Science*, *67*, 293–300.
<https://doi.org/10.1016/j.procs.2015.09.273>

- Massigli, M., Nunes, M. E. de S., Freudenheim, A. M., & Corrêa, U. C. (2011). Estrutura de prática e validade ecológica no processo adaptativo de aprendizagem motora. *Escola de Educação Física e Esporte, Universidade de São Paulo*.
- Mertens, P. (2004). *The Prosogram: Semi-Automatic Transcription of Prosody Based on a Tonal Perception Model*. <http://www.isca-speech.org/archive>
- Miguel Teixeira de Jesus, L., Helena Borges Aguiar Vilarinho Machado Castro Ricardo Jorge Ferreira dos Santos, M., Helena Borges Aguiar Vilarinho Machado Castro, M., & Jorge Ferreira dos Santos, R. (2015). *Protocolo de Avaliação da Qualidade Vocal da Universidade de Aveiro (PAQVUA) 2.3.6*(Número 1). ACSA. <http://acsa.web.ua.pt/Quoteas>:<http://acsa.web.ua.pt/>
- Miranda, P. H. de S., Ferreira, J. V. R., Ferreira, B. A., Dos Santos, N. Í. M., Ribeiro, B. de F., Lopes, C. C., Ataíde, A. G. de O., Da Silva, I. C., Da Silva, M. C. M., & Da Costa, A. C. F. (2021). O uso da realidade virtual na reabilitação do paciente após acidente vascular encefálico / The use of virtual reality in patient rehabilitation after stroke. *Brazilian Journal of Development*, *7*(10), 95350–95365. <https://doi.org/10.34117/bjdv7n10-39>
- Nunes, A. M. B. (2009). *Voz e Emoção em Português Europeu*. Universidade de Aveiro.
- Obana, F. Y., & Tori, R. (2021). *Presença: Conceitos, Definições e Avaliações*. <https://www.researchgate.net/publication/357748855>
- Oliveira Vassoler, A. M. de, & Raposo de Medeiros, B. (2013). Frequência fundamental e emoções: um estudo a partir da fala atuada em português brasileiro. *Signum: Estudos da Linguagem*, *16*(2), 327. <https://doi.org/10.5433/2237-4876.2013v16n2p327>
- Pasquali, L. (2017). Validade dos Testes. *Revista Examen*, *1*(1), 14–48.
- Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport*, *19*(2), 209–213.
- Ribeiro, A. C. O. da S. (2013). *A indução afetiva em cenários de realidade virtual: a importância da tridimensionalidade na avaliação da sensação de presença* [Dissertação apresentada no Instituto Superior de Ciências da Saúde - Norte, para obtenção do grau de Mestre em Neuropsicologia Clínica.]. CESPU - Instituto Superior de Ciências da Saúde Norte.
- Ribeiro, A., & Monteiro, L. (2015). A indução afetiva em cenários de realidade virtual: avaliação da sensação de presença. *Psic. Clin.*, *27*(1), 139–160.
- Scanferla, W. H. (2022). *Avaliação de uma tecnologia mHealth para a promoção da saúde vocal do professor* [Dissertação de mestrado]. Universidade Federal de Minas Gerais.
- Serra, D. (2021). *O papel da prosódia como portadora de sentido na desambiguação de enunciados em língua francesa* [Doutorado em Linguística Aplicada e Estudos da Linguagem]. Pontifícia Universidade Católica de São Paulo.

- Slater, M., Usoh, M., & Steed, A. (1994). Depth of Presence in Virtual Environments. Em *Presence: Teleoperators and Virtual Environments*, (Vol. 3, Número (2)). University of London.
- Souza, R. L. de, & Cardoso, M. C. de A. F. (2013a). Fluência e prosódia: Aspectos diferenciais frente aos distúrbios. Em *Revista Neurociências* (Vol. 21, Número 3, pp. 468–473). <https://doi.org/10.4181/RNC.2013.21.766.6p>
- Souza, R. L. de, & Cardoso, M. C. de A. F. (2013b). Fluência e prosódia: Aspectos diferenciais frente aos distúrbios. Em *Revista Neurociências* (Vol. 21, Número 3, pp. 468–473). <https://doi.org/10.4181/RNC.2013.21.766.6p>
- Teixeira, D. (2021). *Contributo para o desenvolvimento do ProsoVR: proposta guião para protótipo de avaliação da prosódia emocional em ambiente de realidade virtual*. Politécnico do Porto.
- Teixeira, R. M. D. B. T. (2014). *Realidade Virtual no Tratamento de Ansiedade Social (medo de falar em público)*. Universidade de Lisboa.
- Vasconcelos, C., Ferreira Valenzuela de Oliveira, B., Alves Carvalho, S., & Reis, C. (2011). *A expressão da atitude de certeza em indivíduos com perda auditiva bilateral: análise prosódica*.
- Vianello, L. (2006). *O Uso da Voz em Sala de Aula: o Caso Dos Professores Readaptados Por Disfonia*.
- VIRTUS. (2023). *VIRTUS*. <https://www.project-virtus.eu>. <https://www.project-virtus.eu>
- Walkom, G. (2016). Virtual reality exposure therapy: To benefit those who stutter and treat social anxiety. *Proceedings – 2016 International Conference on Interactive Technologies and Games: EduRob in Conjunction with iTAG 2016, iTAG 2016*, 36–41. <https://doi.org/10.1109/iTAG.2016.13>
- Wanderley, T. P. S. P., Batista, M. H. de J., Dutra Júnior, L. D. S., & Silva, V. C. (2018). Docência em saúde: tempo de novas tecnologias da informação e comunicação. *Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*, 12(4). <https://doi.org/10.29397/reciis.v12i4.1522>
- Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, 31(4), 749–778. <https://doi.org/10.1017/S030500090400652X>
- Wilson, K., Bell, C., Wilson, L., & Witteman, H. (2018). Agile research to complement agile development: a proposal for an mHealth research lifecycle. *npj Digital Medicine*, 1(1). <https://doi.org/10.1038/s41746-018-0053-1>
- Zuanetti, P. A., Silva, K. da, Pontes-Fernandes, Â. C., Dornelas, R., & Fukuda, M. T. H. (2018). Characteristics of the emissive prosody of children with autism spectrum disorder. *Revista CEFAC*, 20(5), 565–572. <https://doi.org/10.1590/1982-021620182051718>

Apêndice

Apêndice 1 – Questionário

O presente questionário tem como objetivo recolher o feedback relativo à experiência com o ProsoVR. Deve ser aplicado no final da experiência e o utilizador deve responder de forma autónoma e anónima: No final, deve ser entregue à pessoa responsável.

Este questionário apresenta três páginas. Na primeira encontrará questões relacionadas com a sensação de presença durante a imersão virtual, que deve cotar de 1 a 7. Na segunda página encontrará afirmações relativas à usabilidade do ProsoVR, que deve classificar numa escala de 1 a 5, consoante se concorda ou discorda da afirmação.

Qualquer questão/dúvida que surja durante o preenchimento do questionário, deverá apresentá-la à pessoa responsável.

No final de responder ao questionário, a pessoa responsável colocar-lhe-á algumas questões adicionais, que deve responder oralmente.

1. Sensação de presença

Este questionário apresenta seis questões, em que cada uma das respostas abrange uma escala de 1 a 7, variando entre a pontuação mínima (1 = baixa presença), até à pontuação máxima (7 = elevada presença). Desta forma, solicitamos que responda a cada uma das seis questões registrando um número de 1 a 7 no campo “resposta”.

	RESPOSTA
1. Até que ponto se sentiu presente no cenário apresentado? 1 a 7, em que 1 corresponde a baixa sensação de presença, e 7 elevada sensação de presença.	
2. Até que ponto o cenário apresentado se tornou uma realidade e quase se esqueceu do contexto laboratorial em que decorre a experiência? 1 a 7, em que 1 corresponde nunca se tornou numa realidade, e 7 foi sempre uma realidade.	
3. Até que ponto ao recordar-se do cenário, o vê como uma imagem, ou como um sítio onde esteve presente? 1 a 7, em que 1 apenas como imagem, e 7 como um sítio em que esteve presente.	
4. Durante o tempo da experiência, foi mais forte o sentido de estar no cenário ou de estar em outro lugar? 1 a 7, em que 1 corresponde a estar em outro lugar, e 7 estar no cenário.	
5. Em que medida as condições estruturais (cor, objetos) do cenário o recordam de outras situações semelhantes em que já esteve? 1 a 7, em que 1 recorda pouco, e 7 recorda muito.	
6. Durante algum momento da experiência pensou que estivesse realmente no cenário? 1 a 7, em que 1 equivale a pensei pouco, e 7 pensei muito.	

2. Usabilidade

Escala de usabilidade do sistema – SUS * validado para PE

Martins, A. I., Rosa, A. F., Queirós, A., Silva, A., & Rocha, N. P. (2015). European Portuguese Validation of the System Usability Scale (SUS). *Procedia Computer Science*, 67, 293–300. <https://doi.org/10.1016/j.procs.2015.09.273>

1. Acho que gostaria de utilizar este produto com frequência.

Discordo totalmente				Concordo totalmente
1	2	3	4	5

2. Considerei o produto mais complexo do que necessário.

Discordo totalmente				Concordo totalmente
1	2	3	4	5

3. Achei o produto fácil de utilizar.

Discordo totalmente				Concordo totalmente
1	2	3	4	5

4. Acho que necessitaria de ajuda de um técnico para conseguir utilizar este produto.

Discordo totalmente				Concordo totalmente
1	2	3	4	5

5. Considerei que as várias funcionalidades deste produto estavam bem integradas.

Discordo totalmente				Concordo totalmente
1	2	3	4	5

6. Achei que este produto tinha muitas inconsistências.

Discordo totalmente				Concordo totalmente
1	2	3	4	5

7. Suponho que a maioria das pessoas aprenderia a utilizar rapidamente este produto.

Discordo totalmente				Concordo totalmente
1	2	3	4	5

8. Considerei o produto muito complicado de utilizar.

Discordo totalmente				Concordo totalmente
1	2	3	4	5

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

Discordo totalmente				Concordo totalmente
1	2	3	4	5

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

Discordo totalmente				Concordo totalmente
1	2	3	4	5

Apêndice 2 – Entrevista

Protótipo do ProsoVR – Feedback da experiência imersiva

Objetivos gerais:

- Recolher informação relevante para alterar ou complementar o protótipo
- Perceber qual é a perspetiva dos utilizadores face à imersão
- Perceber se o protótipo apresenta contextos capazes de simular situações reais

Questões:

- I. Qual a sua opinião sobre esta experiência de imersão virtual?
 - Gostei mais de:
 - Gostei menos de:

- II. Em que medida o cenário da alegria (café) induziu a emoção pretendida?
 - Efeitos sonoros, interação com personagens, tempo...

- III. Em que medida o cenário da tristeza (gabinete médico) induziu a emoção pretendida?
 - Efeitos sonoros, interação com personagens, tempo...

- IV. Em que medida o cenário da raiva (loja) induziu a emoção pretendida?
 - Efeitos sonoros, interação com personagens, tempo...

- V. Há alguma coisa que mudaria na experiência? Se sim, o quê?

- VI. Há algum aspeto que gostasse de acrescentar, que não tenha sido dito anteriormente?

Muito obrigada pela colaboração!

PARECER DA COMISSÃO DE ÉTICA

Número de Registo da Comissão de Ética: CE0020D
Data receção do Documento: 02/01/2024, processo de ADENDA
Existência de entradas anteriores: Sim

TÍTULO DO TRABALHO: ProsoVR – desenvolvimento do protótipo de uma ferramenta de avaliação da prosódia mediada por realidade virtual

INVESTIGADOR RESPONSÁVEL: Diana Catarina FreitasTeixeira

DATA PREVISTA PARA A REALIZAÇÃO DO TRABALHO: Início | Fim

RESUMO DO ESTUDO

OBJETIVOS: sem alteração.

AMOSTRA: Na 1ª fase refere que é uma amostra de conveniência, a partir dos contactos pessoais da equipa de investigação.

Tem como critérios de inclusão - ter mais de 18 anos, sem alterações autorrelatadas ao nível da comunicação, linguagem, fala e/ou audição-, e são critérios de exclusão - ser terapeutas da fala, atores ou indivíduos experientes em representação, ter problemas de visão (usar óculos, ser cego) e limitações da fluência da leitura e pessoas analfabetas expressos.

FORMULÁRIO DE DADOS A RECOLHER: foram acrescentadas 3 escalas a Slater- Usoh- Steed Questionnaire (SUS), a Usabilidade do sistema, e modelo de entrevista semiestruturada, que cumprem estarem publicadas e/ou questões ajustadas aos objetivos de estudo e não serem sensíveis.

MATERIAL: Nada a referir.

MÉTODOS: Descreve métodos usados referindo sem alterações

RISCOS: Não

CONSENTIMENTO INFORMADO: sem alterações.

AUTORIZAÇÃO PELOS RESPONSÁVEIS LOCAIS: sem alterações.

APRECIAÇÃO DA COMISSÃO DE ÉTICA: Da análise realizada por esta Comissão o processo encontra-se bem instruído.

PARECER FINAL DA COMISSÃO DE ÉTICA: De acordo com os dados analisados, o parecer é favorável desde que cumpridas todas as diretrizes submetidas a esta Comissão, recomendando-se que a decisão seja suspensa caso haja algum incumprimento grave.

29/01/2024

Assinado por: **Pedro Manuel Ribeiro da Rocha Monteiro**
Num. de Identificação: 09132856
Data: 2024.02.07 18:38:37 +0000