

Índice de Experiência de Usuário em Realidade Virtual imersiva: uma proposta de questionário padronizado

Jean F. P. Cheiran^{1,3}, Denise R. Bandeira², Marcelo S. Pimenta³

¹ Campus Alegrete – Universidade Federal do Pampa (UNIPAMPA)
Av. Tiarajú, 810 – Alegrete/RS – Brasil

²Instituto de Psicologia – Universidade Federal do Rio Grande do Sul (UFRGS)
R. Ramiro Barcelos, 2600 – Porto Alegre/RS – Brasil

³Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Av. Bento Gonçalves, 9500 – Porto Alegre/RS – Brasil

jeancheiran@unipampa.edu.br, bandeira@ufrgs.br, mpimenta@inf.ufrgs.br

Abstract. *Introduction:* Considering particular aspects of the User Experience (UX) in immersive Virtual Reality (VR), psychological constructs such as presence and cybersickness transcend abstract conceptualizations: they have a practical impact on the user-technology interaction. **Objective:** This research presents the development and empirical evaluation of the index of User Experience in immersive Virtual Reality (iUXVR) for Portuguese-speaking contexts. **Methodology or Steps:** The creation process involved proposing items based on a content analysis and selecting the best items through expert evaluations. After a pilot study, an experiment with 77 participants resulted in the collection of 154 questionnaires, which were analyzed using Partial Least Squares Structural Equation Modeling. **Results:** Comprising 25 items distributed across five key components – usability, aesthetics, sense of presence, VR sickness, and emotions – the final version of the iUXVR presents sufficient evidence of validity and a stable structural model, supporting its applicability. **Keywords** User Experience, Virtual Reality, Evaluation, Standardized Questionnaire.

Resumo. *Introdução:* Levando em conta os aspectos particulares da Experiência de Usuário (UX) em Realidade Virtual (RV) imersiva, construtos psicológicos como presença ou cybersickness não são apenas conceitos abstratos: eles têm impacto prático na interação dos usuários com a tecnologia. **Objetivo:** Esta pesquisa apresenta o desenvolvimento e a avaliação empírica do índice de Experiência de Usuário em Realidade Virtual imersiva (iUXRV) para o contexto brasileiro. **Metodologia ou Etapas:** O processo de criação envolveu a proposição de itens com base em análise de conteúdo e a escolha dos melhores itens a partir da avaliação de especialistas. Depois de um estudo piloto, um experimento com 77 participantes garantiu a coleta de 154 questionários completos que foram analisados por meio de Modelagem de Equações Estruturadas por Mínimos Quadrados Parciais. **Resultados:** Contendo 25 itens distribuídos em cinco componentes-chave – usabilidade, estética, presença, VR sickness e emoções – a versão final do iUXRV apresenta evidências de validade e um modelo estrutural estável que viabilizam seu uso.

Palavras-Chave Experiência de Usuário, Realidade Virtual, Avaliação, Questionário Padronizado.

1. Introdução

A Realidade Virtual (RV) tem se popularizado nos últimos anos. Além da comercialização de dispositivos mais baratos, os equipamentos de RV imersiva vêm se tornando significativamente melhores [Kim et al. 2020]. Essa ampliação de consumo por usuários comuns (e não apenas pesquisadores e aficionados por tecnologia) levou a um aumento na disponibilidade de softwares de RV em lojas de aplicativos como Steam e Meta Store [Research e Markets 2025]. Hoje temos aplicações relevantes de RV em áreas como saúde, educação e treinamento, além das conhecidas aplicações de entretenimento (games). O aumento da demanda por esse tipo particular de aplicação, por sua vez, mobilizou *designers* e desenvolvedores a procurarem formas de criar experiências mais envolventes, positivas e intensas para usuários dessas tecnologias.

Considerando que a Experiência de Usuário (UX) é um termo amplo cuja definição depende dos referenciais dos especialistas, pesquisadoras identificaram na comunidade brasileira de Interação Humano-Computador uma concordância maior com o entendimento de UX como “uma consequência do estado interno dos usuários (predisposições, expectativas, necessidades, motivações, humor etc), das características do sistema projetado (por exemplo: complexidade, propósito, usabilidade, funcionalidade etc) e do contexto (ou ambiente) no qual a interação ocorre (por exemplo: ambiente organização/social, significância da atividade, voluntariedade de uso etc)” [Melo e Darin 2019]. Ainda que especialistas sejam céticos sobre uma definição comum do que é Experiência de Usuário [Law et al. 2014], medir essa experiência se refere a uma ampla coleção de métodos, técnicas e ferramentas para avaliar como uma pessoa sente, percebe e responde a um produto antes, durante e depois de interagir com ele [ISO 2019].

Embora pesquisadores e desenvolvedores adotem diferentes formas de avaliação da Experiência de Usuário [Barbosa et al. 2021], questionários continuam sendo uma maneira barata e simples de coletar dados. Considerando que a aplicação dos questionários praticamente não requer habilidades específicas, o esforço de garantir dados úteis e confiáveis recai no seu processo de criação que preferencialmente envolve a avaliação de suas propriedades psicométricas para coletar evidências que reforcem validade (a sustentação da interpretação dos escores do questionário para o uso proposto) e fidedignidade (a consistência dos escores ao longo de replicações de uma avaliação) [AERA; APA; NCME 2014].

Instrumentos psicométricos são fundamentais no contexto da avaliação de experiências subjetivas por serem baseados em princípios científicos para medir aspectos do comportamento, percepções e cognição de indivíduos [AERA; APA; NCME 2014]. Com base em uma vasta coleção de abordagens de modelagem e de métodos estatísticos, esses instrumentos são projetados para uso recorrente e contêm regras bem definidas de interpretação dos seus resultados [Sauro e Lewis 2016]. No contexto da avaliação de UX em Realidade Virtual, a psicometria oferece suporte para o desenvolvimento de instrumentos que são válidos para propósitos de aferição da qualidade da experiência, que resultam em estimativas confiáveis, padronizadas e comparáveis, e que permitem a tomada de decisões melhores sobre o *design* de ambientes imersivos. Mais

especificamente, essas melhorias no processo de tomada de decisão permitem priorizar investimentos em aspectos críticos da interação e da tecnologia, adaptar os cenários para características de grupos de usuários-alvo e sustentar argumentos para políticas de segurança relacionadas a ambientes de RV.

Esses questionários com regras bem definidas de aplicação e interpretação, denominados questionários padronizados [Sauro e Lewis 2016], são amplamente utilizados para aferir experiências em ambientes de RV [Chang et al. 2020, Grassini e Laumann 2020, Kim et al. 2020]. Contudo, pesquisadores e profissionais de RV brasileiros enfrentam um desafio no uso de instrumentos padronizados: não estão disponíveis questionários em português para aferir especificamente a Experiência de Usuário em RV. Essa lacuna frequentemente leva a alternativas que comprometem a qualidade da avaliação como, por exemplo, usar questionários de propósito geral para medir UX (deixando de fora aspectos essenciais da RV), usar múltiplos questionários diferentes para depois compilar um escore global (aumentando a carga cognitiva dos respondentes e causando incertezas sobre a combinação dos escores), usar uma tradução *ad hoc* de um questionário (sem garantias de que a versão traduzida possui as mesmas características psicométricas da original) ou criar um questionário completamente novo sem qualquer garantia de boas propriedades psicométricas.

Considerando a escassez de questionários para avaliar a UX em RV de forma holística e a ausência de instrumentos em português para essa finalidade, este trabalho tem por objetivo descrever o desenvolvimento de um questionário padronizado em português brasileiro para avaliação de Experiência de Usuário em Realidade Virtual imersiva. O questionário foi baseado no *framework* de Componentes de Experiência de Usuário (CUE, do inglês Components of User Experience) [Mahlke 2008] e inclui componentes críticos relacionados à Realidade Virtual: presença e VR sickness. Ainda, o artigo apresenta evidências de validade e estimativas de fidedignidade que sustentam a viabilidade de sua aplicação pela comunidade brasileira de pesquisadores e profissionais de RV.

2. Trabalhos relacionados

Embora existam estudos na área de Interação Humano-Computador que mapeiam questionários padronizados para aferição de Experiência de Usuário global [Vermeeren et al. 2010, Maia e Furtado 2016, Rivero e Conte 2017, Darin et al. 2019, Nur et al. 2021] e questionários padronizados para aplicações em RV [Kim et al. 2020, Marques et al. 2024], os pesquisadores brasileiros ainda são limitados pela disponibilidade dos instrumentos encontrados no idioma português.

Trabalhos de adaptação dos instrumentos para outro idioma não podem ser encarados como uma simples tradução, mas devem representar um processo sistemático de adaptação cultural para o idioma alvo e de experimentação para avaliação da qualidade psicométrica da nova versão. Considerando que “simplesmente traduzir um teste de um idioma para outro não garante [...] que é comparável em conteúdo e dificuldade à versão original do teste ou que o teste traduzido produz escores que são igualmente confiáveis/precisos e válidos” [AERA; APA; NCME 2014], o uso de traduções *ad hoc* não é uma solução recomendada. A System Usability Scale (SUS), o Simulator Sickness Questionnaire (SSQ) e o Presence Questionnaire (PQ)

estão entre os raros exemplos de instrumentos com traduções para português que costumam ser utilizados em pesquisas que envolvem avaliação de UX em Realidade Virtual [Martins et al. 2015, Lourenço et al. 2022, Carvalho et al. 2011, Silva et al. 2016, Vasconcelos-Raposo et al. 2021]. Contudo, a combinação dos escores obtidos em uma única estimativa de Experiência de Usuário representa um desafio que é agravado pela carga cognitiva provocada pela diferença no formato de resposta dos diferentes questionários e pela grande quantidade de itens de questionários combinados.

O instrumento User Experience Questionnaire (UEQ) divide a experiência global nos componentes de atratividade, transparência, eficiência, controle, estímulo e inovação [Schrepp et al. 2014], possuindo tradução para diversos idiomas. Contudo, a versão em língua portuguesa disponível baseia-se no português europeu para a escolha dos adjetivos da escala de diferencial semântico [Cota et al. 2014], causando dificuldade para participantes brasileiros que encontram palavras incomuns (como obstrutivo, condutor, vanguardista e impraticável) ou com escrita não usual (como cómodo e incômodo). Ainda, a ausência de componentes particulares da experiência em RV causa prejuízo no processo de medição pela omissão de dimensões críticas como presença e *VR sickness*.

Os modelos de UX para Realidade Virtual fundamentam-se na necessidade de capturar dimensões técnicas, sensoriais e psicológicas únicas dos ambientes imersivos. Diferentemente de *frameworks* de propósito geral que priorizam qualidades de UX agnósticas em relação ao domínio e à tecnologia, como o modelo CUE [Mahlke 2008] ou as dimensões pragmática e hedônica do AttrakDiff [Hassenzahl et al. 2003], os modelos para RV geralmente incorporam construtos específicos relacionados à ilusão de se estar em outro lugar, ao realismo de um cenário, aos sintomas de desconforto físico causados pelo uso da tecnologia e ao estado de *flow* na realização de uma atividade.

O modelo proposto por Shin et al. baseia-se na Teoria da Confirmação de Expectativas e no Modelo de Aceitação de Tecnologia para avaliar ambientes virtuais 3D de aprendizagem [Shin et al. 2013]. O questionário desenvolvido inclui componentes de experiência do usuário (UX), organizados em quatro fatores: experiencial, cognitivo, afetivo e facilitador. Apesar da boa qualidade psicométrica dos componentes, o questionário não contempla *VR sickness* (algo justificável dado que o estímulo era um sistema não imersivo acessado por TV 3D) e o componente de estética.

O modelo de Cheng e colegas busca examinar a UX em mercados de RV com base na teoria de *flow* e componentes relacionados [Cheng et al. 2014]. O questionário possui dez componentes divididos em quatro estágios: variáveis do conteúdo midiático, antecedentes do estado de *flow*, estado de *flow* e consequências do *flow*. Destacam-se a interatividade como responsividade do software, o envolvimento relacionado a prazer e interesse, e a vivacidade como realismo sensorial. *VR sickness* é novamente ignorado, partindo da hipótese de que seria um antecedente do fluxo nas sessões experimentais com um simulador imersivo. A usabilidade e a estética também são pouco abordadas. Embora o questionário tenha boas propriedades psicométricas, há respostas com formatos distintos (escala Likert de 7 pontos e diferenciais semânticos), itens com explicações longas (destaque para uma explicação de 8 linhas sobre o conceito de *flow*) e sobreposição das componentes nos itens do questionário.

Finalmente, as pesquisas de Tcha-Tokey e colegas têm como objetivo principal

construir um questionário abrangente para avaliar a UX em ambientes virtuais imersivos, incluindo múltiplos componentes como presença, engajamento, fluxo e adoção tecnológica [Tcha-Tokey et al. 2016, Tcha-Tokey et al. 2018]. Apesar da boa cobertura e de propriedades psicométricas razoáveis, o questionário apresenta problemas como padrões de resposta variados, sobreposição entre itens, versões concorrentes com diferenças relevantes, e ausência da avaliação de validade convergente e discriminante. Ainda assim, esses trabalhos destacam-se positivamente por utilizar óculos de Realidade Virtual como tecnologia imersiva em suas sessões experimentais.

Assim, observa-se que os modelos de UX para VR representam um avanço paradigmático, integrando perspectivas interdisciplinares da psicologia, *design* e ergonomia para orientar o desenvolvimento de sistemas imersivos seguros, éticos e centrados no usuário.

3. Proposta do questionário

A estrutura do iUXRV estende o *framework* de Componentes de Experiência de Usuário (CUE, do inglês Components of User Experience) [Mahlke 2008] e especializa a estrutura da avaliação modular de Componentes-chave de Experiência de Usuário (meCUE, do inglês modular evaluation of key Components of User Experience) [Minge et al. 2016]. Esse modelo foi escolhido por estabelecer uma relação causal clara entre seus componentes e por permitir uma abordagem de componentes críticos que poderia ser estendida para incluir elementos particulares das experiências em RV imersiva.

Tendo como base os três componentes mais amplos do CUE (qualidades instrumentais, qualidades não-instrumentais e reações emocionais), o iUXRV inclui dois outros componentes essenciais da experiência em RV: presença e *VR sickness*.

A sensação de presença pode ser definida como a experiência subjetiva de estar em um lugar embora se esteja fisicamente em outro [Witmer e Singer 1998]. Mesmo que não exista consenso sobre os elementos que compõem a sensação de presença, o iUXRV inclui itens de dois importantes fatores na construção da experiência subjetiva de presença [Slater 2009]: a ilusão de lugar e a ilusão de plausibilidade. Enquanto a ilusão de lugar representa a forte ilusão de estar em um local a despeito do conhecimento inegável de que não se está lá, a ilusão de plausibilidade representa a credibilidade do cenário que está sendo apresentado em comparação com as expectativas.

O *VR sickness* pode ser compreendido como uma coleção de sintomas e efeitos desagradáveis relacionados ao mal-estar visualmente induzido em Realidade Virtual e a outras características contextuais do uso de RV como, por exemplo, conforto do equipamento, predisposição a enjoos e outros. Considerando a extensa investigação da comunidade científica de Realidade Virtual sobre esse fenômeno [Davis et al. 2014, Martirossov e Kopecek 2017, Chang et al. 2020], o iUXRV usa como base os principais sintomas e efeitos mapeados pelas pesquisas na área para os itens de *VR sickness*.

Além de ampliar o *framework* CUE original para incluir aspectos críticos de Realidade Virtual que não se encaixam nos elementos de qualidade propostos, o iUXRV adota apenas um subcomponente para representar cada um dos componentes mais amplos [Minge et al. 2016]: usabilidade para qualidade instrumental, estética para qualidade não-instrumental e emoções para reações emocionais.

A usabilidade está relacionada a como usuários interagem com um produto para atingir objetivos em um contexto de uso [ISO 2018], sendo um aspecto frequentemente avaliado em pesquisas de Realidade Virtual imersiva. Considerando as qualidades mais comuns descritas na literatura especializada [Nielsen 1993, Preece et al. 2015, ISO 2018, Barbosa et al. 2021], os itens do iUXRV contemplam três fatores de usabilidade: facilidade de aprendizagem (o esforço para usuários se tornarem especialistas no uso do produto), eficiência (o desempenho e a quantidade de recursos necessários para atingir um resultado) e eficácia (a precisão e a completude da realização de uma tarefa).

A estética reflete um fenômeno complexo que envolve estímulo e apreciação de algo. Essa pesquisa adota uma perspectiva de dois fatores críticos na experiência estética [Lavie e Tractinsky 2004]: a estética clássica que envolve as noções clássicas do *design* estético e a estética expressiva que reflete a riqueza de criatividade e originalidade do *designer*. Mais especificamente, os itens do iUXRV abordam especificamente a estética visual relacionada à estética clássica e o estímulo relacionado à estética expressiva.

Finalmente, as emoções são consideradas “a experiência consciente do afeto complementada pela atribuição de sua causa e identificação de seu objeto” [Norman 2004], podendo ser mapeadas em uma pequena coleção de emoções básicas [Izard 1977, Izard 2007]. Diferente da perspectiva de emoções positivas e negativas do meCUE [Minge et al. 2016], essa pesquisa assume uma abordagem não polarizada que pode contribuir com desenvolvedores de aplicações em RV para ajustarem os procedimentos de pontuação para refletir a intencionalidade da experiência projetada. Por exemplo, uma emoção de “medo” deveria contribuir positivamente para a UX em um jogo de horror, mesmo que tradicionalmente o medo seja considerado uma emoção negativa.

Dessa forma, o modelo estrutural da Figura 1 é considerado para o índice de Experiência de Usuário em Realidade Virtual imersiva.

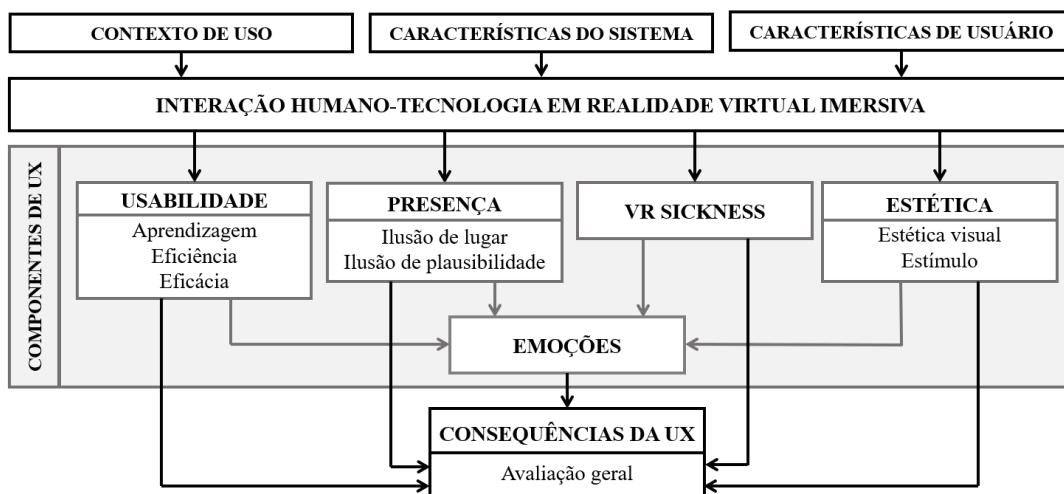


Figura 1. Modelo de Experiência de Usuário em Realidade Virtual imersiva.

3.1. Métodos

A avaliação das propriedades psicométricas do índice de Experiência de Usuários em Realidade Virtual imersiva foi baseada na Modelagem de Equações Estruturadas por

Mínimos Quadrados Parciais (PLS-SEM, do inglês Partial Least Square Structured Equation Modeling). Esse método assume que os componentes do modelo são aproximações válidas dos fenômenos reais, sendo formados por combinações lineares dos indicadores. As aproximações (comumente denominadas ‘compósitos’) são estimadas a partir dos dados amostrais por meio de um processo iterativo de minimização de erros que se baseia em regressões lineares múltiplas. Essa abordagem foi escolhida por não requerer dados em distribuição normal, além de ser viável para amostras reduzidas, eficaz para o contexto de modelos complexos e adequada para situações que incluem subcomponentes ligados a um componente formativo de mais alta ordem [Hair e Alamer 2022].

A estimativa de tamanho amostral foi realizada por meio do software G*Power (v. 3.1.9.7). Considerando um tamanho de efeito médio ($f^2 = .15$), a quantidade de preditores em nosso modelo e probabilidades de erro padrão ($\alpha = .05$, $\beta = .80$), uma amostra com, ao menos, 114 observações é necessária. Considerando que cada pessoa da pesquisa contribui com dois questionários preenchidos, a quantidade mínima de participantes foi estimada em 57.

Todas as estatísticas e estimativas dos modelos foram realizadas por meio da linguagem de programação R (v. 4.4.2) e do pacote SEMinR (v. 2.3.4). Ainda, a significância estatística e os intervalos de confiança foram calculados a partir de *bootstrapping* com 1000 subamostras.

3.2. Procedimentos

O processo de criação e avaliação do questionário proposto seguiu as etapas representadas na Figura 2.

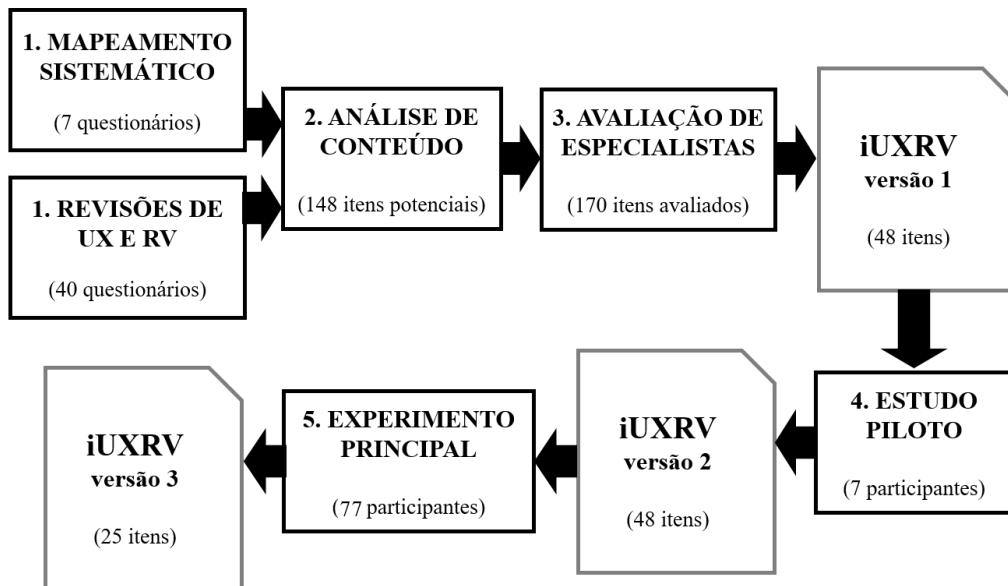


Figura 2. Desenho da pesquisa.

A identificação de questionários de avaliação de Experiência de Usuário em ambientes de Realidade Virtual foi realizada por meio de um **mapeamento sistemático** [Kitchenham et al. 2011] que identificou sete instrumentos específicos.

A partir dos instrumentos mapeados por essa pesquisa e dos instrumentos identificados em outras revisões da área, um pesquisador conduziu uma **análise de conteúdo** [Moraes 1999] contendo 47 instrumentos relacionados à Experiência de Usuário para Realidade Virtual, Experiência de Usuário geral, usabilidade, presença, estética, emoções, *VR sickness* e outros componentes.

Essa análise dividiu os instrumentos em 693 unidades de avaliação e, posteriormente, agrupou esses itens em 155 categorias com propriedades específicas da técnica. Esse processo de agrupamento é inherentemente subjetivo e considerou a similaridade textual das unidades originais e suas intenções dentro dos instrumentos de origem. O tema de cada categoria emergiu a partir da investigação de seu conteúdo, dando origem aos 148 itens iniciais do questionário após a filtragem de categorias com propósito redundante. Esses itens foram então agrupados nos respectivos componentes do modelo teórico dessa pesquisa.

Em seguida, a **avaliação por especialistas** contou com quatro pesquisadoras experientes em Realidade Virtual ou Experiência do Usuário para avaliarem a coleção de itens iniciais, atribuindo um escore para sua relevância na avaliação de UX e um escore para a qualidade do texto. Esses escores foram usados na escolha dos itens para a primeira versão do questionário.

Além disso, as especialistas também sugeriram a troca do componente de alguns itens, mudanças nos textos de itens existentes e itens completamente novos que não foram identificados na análise de conteúdo. Essa etapa contribuiu tanto para a escolha de itens relevantes a partir dos escores das especialistas quanto para a melhoria da compreensão dos itens a partir dos ajustes nos textos.

O questionário foi estruturado na forma de afirmações em formato Likert com uma faixa de resposta de sete pontos, numerada de 1 a 7 e rotulada com uma escala de concordância crescente: discordo totalmente, discordo, discordo um pouco, nem concordo nem discordo, concordo um pouco, concordo e concordo totalmente.

Um **estudo piloto** foi conduzido com 7 participantes nas mesmas condições experimentais do experimento principal dessa pesquisa. O recrutamento desses participantes partiu de convites pontuais a pessoas conhecidas com perfil desejado para essa etapa. Esses participantes variaram em idade (entre 23 e 57 anos), escolaridade (de ensino médio completo a doutorado) e experiência em Realidade Virtual imersiva (de nunca ter experimentado até uso frequente) de forma a contribuir com a representatividade. Procedimentos relacionados à ética em pesquisa estão na Seção 3.3.

Os participantes responderam a diferentes leiautes do questionário impresso depois de usarem três aplicações em Realidade Virtual e, ao final da sessão experimental, foram questionados sobre as instruções e sobre os itens do iUXRV conforme o protocolo de uma entrevista semi-estruturada cujas respostas foram registradas por meio de anotações [Barbosa et al. 2021]. Os resultados dessa entrevista e a observação dos participantes respondendo ao questionário permitiram fazer novos refinamentos de texto e escolher o leiaute mais claro e com menor taxa de erros.

O **experimento principal** contou com 77 participantes (51 se identificam com gênero masculino, 24 se identificam com gênero feminino e 2 preferiram não informar sua identificação de gênero) com idades entre 18 e 56 anos ($M = 21.21$, $SD = 5.30$, $Mdn =$

20). A maior parte dos participantes é estudante de graduação ($n = 50$, ou 65%), pessoas com ensino médio completo ($n = 18$, ou 24%) e pessoas com graduação completa ($n = 6$, ou 8%). Ainda, 31 participantes (40%) informaram uma experiência prévia com dispositivos de Realidade Virtual imersiva: 24 indicaram ter usado uma vez, 5 relataram uso mensal e 2 responderam uso semanal. Essas pessoas foram recrutadas a partir de redes sociais, listas de e-mail, cartazes e convites pontuais. Os procedimentos relacionados à ética em pesquisa são os mesmos descritos na Seção 3.3.

Os experimentos ocorreram nas instalações da Universidade Federal do Rio Grande do Sul (UFRGS) e da Universidade Federal do Pampa (UNIPAMPA) em salas silenciosas, climatizadas e com dimensões mínimas de 3 x 3 metros. Os participantes usaram um equipamento Meta Quest 2 com proteção de silicone lavável, sem fones de ouvido e com os controles Quest 2 VR. As aplicações de Realidade Virtual escolhidas foram Oculus First Contact¹, Painting VR² e Beat Saber³. Enquanto a aplicação Oculus First Contact foi usada apenas para familiarização dos participantes, o Painting VR foi escolhido por representar uma interação em modo de objetivo [Hassenzahl 2003] com meta de longo prazo de criar uma pintura em um ambiente tranquilo. O Beat Saber, por sua vez, representa uma interação em modo de ação [Hassenzahl 2003] por sua dinâmica intensa e pelos mini-objetivos de cortar blocos no ritmo certo e desviar de obstáculos.

Os participantes usaram cada aplicação por cerca de 15 minutos antes de responderem versões impressas dos questionários. Além do questionário iUXRV, os instrumentos SUS, SSQ e PQ foram usados em suas versões em português [Martins et al. 2015, Carvalho et al. 2011, Silva et al. 2016] de forma intercalada: depois do uso de uma aplicação, apenas o iUXRV e mais um questionário eram respondidos. Essa abordagem permitiu reduzir a fadiga dos participantes e a aplicação de todos os questionários em diferentes ordens foi garantida por meio de randomização em *latin square*. A Figura 3 apresenta a estrutura geral de uma sessão experimental.

3.3. Cuidados éticos

Todos os participantes tiveram seu consentimento registrado em papel por meio de um termo contendo dados da pesquisa, benefícios e riscos, seguindo os padrões do Comitê de Ética em Pesquisa local. Nenhum participante foi remunerado pela participação, e todos os dados resultantes da pesquisa são reportados de forma anônima.

Além do Termo de Consentimento, um questionário de triagem foi aplicado antes do início da sessão experimental para identificar pessoas que não deveriam usar equipamentos de Realidade Virtual imersiva de acordo com os manuais do dispositivo usado. Condições de triagem incluíam pessoas grávidas, pessoas com transtornos psiquiátricos severos, pessoas com problemas cardíacos graves e outros. De forma a preservar a privacidade em relação a cada critério, as pessoas respondiam ao final do questionário apenas se cumpriam todos os requisitos de participação ou não.

Essa pesquisa foi aprovada pelo Comitê de Ética em Pesquisa da UFRGS (CAAE 53245721.1.0000.5347) e seguiu todas as recomendações e normas para pesquisas com seres humanos.

¹<https://www.oculus.com/experiences/quest/2188021891257542>

²<https://www.oculus.com/experiences/quest/3106117596158066>

³<https://www.oculus.com/experiences/quest/2448060205267927>

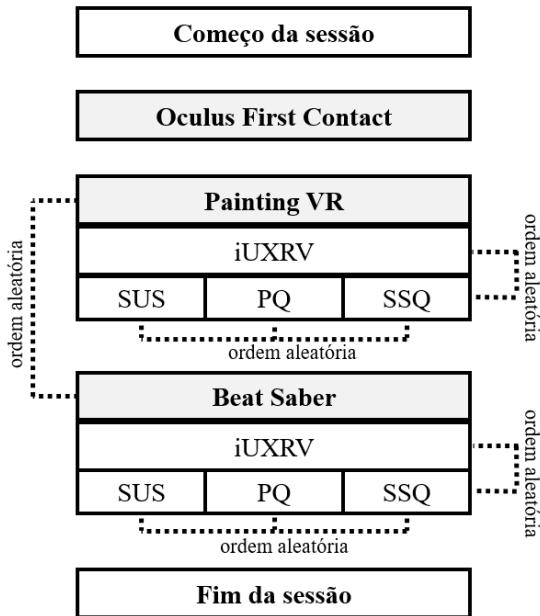


Figura 3. Estrutura de uma sessão experimental com ordem de aplicações e aplicação de questionários.

4. Avaliação de qualidade psicométrica

Essa pesquisa adota a perspectiva de evidências de validade a partir de três fontes [AERA; APA; NCME 2014]: conteúdo do teste (a análise das relações entre o componente medido e o conteúdo do questionário), estrutura interna (o grau de relacionamento entre os itens, componentes e escores) e relação com outras variáveis (a comparação com outros instrumentos que medem o mesmo componente).

Destacam-se como **evidências baseadas no conteúdo do teste**:

1. A cuidadosa análise de conteúdo que deu origem aos 148 itens originais;
2. A avaliação por especialistas que ampliou a quantidade de itens para 170 e resultou na classificação de relevância e qualidade;
3. O processo de criação da primeira versão do questionário que escolheu 48 itens baseados em sua relevância e qualidade;
4. A definição da ordem dos itens no questionário que não permitiu que itens de um mesmo componente fossem contíguos, exceto pelos itens de *VR sickness* que foram agrupados no início do questionário por medirem sintomas transientes.
5. A aplicação da primeira versão do questionário em um estudo piloto que permitiu escolher o layout mais adequado e refinar a escrita de itens para estarem comprehensíveis para pessoas com diferentes idades e escolaridades.

As **evidências de validade baseadas na estrutura interna** incluem a estimativa das cargas dos itens do questionário, coeficientes de fidedignidade, coeficientes de caminho e outras métricas.

O modelo de medição do questionário foi estimado usando a abordagem de indicadores repetidos [Becker et al. 2012] e incluiu um indicador adicional no final do questionário para medir a experiência geral: **Escolha uma nota de 0 (zero) a 10 (dez)**

para sua experiência no mundo virtual, sendo 0 uma experiência péssima e 10 uma experiência ótima. A análise das cargas dos indicadores do modelo estimado garantiu a remoção de itens com carga inferior a .40, sendo esse considerado o valor mínimo para pesquisas exploratórias usando PLS-SEM [Hair et al. 2021]. Adicionalmente, os itens com baixa contribuição para a variância explicada foram removidos, assim como os itens cuja exclusão aumentaria significativamente os coeficientes de fidedignidade. Esse procedimento reduziu a quantidade de itens de 48 para 25. Os itens resultantes, suas cargas, as estatísticas T e os intervalos de confiança são apresentados na Tabela 1.

Tabela 1. Cargas dos indicadores para a solução de 5 componentes usando bootstrapping (25 itens). Todas as cargas são estatisticamente significativas e relevantes, embora Q4, Q5 e Q9 tenham cargas levemente abaixo do desejável.

	Componente	Carga	Estatística T	IC 2.5%	IC 97.5%
Q1	VR sickness	.919	5.822	.763	.951
Q2	VR sickness	.816	4.416	.489	.932
Q3	VR sickness	.890	5.332	.659	.941
Q4	VR sickness	.538	2.472	-0.079	.745
Q5	VR sickness	.648	3.500	0.224	.857
Q6	Usabilidade	.799	14.802	.677	.888
Q7	Estética	.788	10.812	.612	.880
Q8	Presença	.856	31.036	.792	.902
Q9	Emoções	.699	5.102	.380	.885
Q10	Usabilidade	.807	12.336	.626	.885
Q11	Estética	.910	17.371	.743	.951
Q12	Presença	.871	37.700	.819	.909
Q13	Emoções	.883	23.653	.780	.929
Q14	Usabilidade	.846	26.843	.780	.900
Q15	Estética	.784	13.054	.638	.877
Q16	Presença	.776	18.993	.683	.842
Q17	Emoções	.820	11.311	.626	.905
Q18	Usabilidade	.823	16.090	.686	.891
Q19	Estética	.921	23.236	.799	.954
Q20	Presença	.822	24.164	.746	.880
Q21	Emoções	.720	8.438	.490	.835
Q22	Usabilidade	.730	8.181	.534	.860
Q23	Estética	.919	19.443	.776	.957
Q24	Presença	.855	35.713	.807	.900
Q25	Emoções	.833	10.618	.631	.920

Os coeficientes de fidedignidade representam o grau de associação entre os indicadores de um mesmo componente [Russo e Stol 2021], variando de zero (nenhuma associação) até 1 (homogeneidade perfeita e sem erros de medição). Embora não exista consenso sobre as medidas mais adequadas para esses coeficientes, valores abaixo de .60 costumam representar uma significativa falta de associação ou um elevado erro de medição e valores acima de .95 indicam redundância excessiva de medida ou padrões

de resposta indesejados [Hair et al. 2021]. Dentre os vários coeficientes existentes, o coeficiente alfa de Cronbach é um estimador popular tau-equivalente derivado da correlação entre os itens e o coeficiente rô de Henseler é uma medida de fidedignidade congenérica que é mais precisa quando os pressupostos do coeficiente alfa são violados [Hair et al. 2021].

Ainda, a Variância Média Explicada (VME) é outra métrica baseada nos indicadores para medir a força da relação entre os itens e a medida em que são intercambiáveis [Russo e Stol 2021]. O limite de $VME \geq .50$ é comumente adotado por sugerir que o componente explica 50% ou mais da variância entre seus itens [Hair et al. 2021]. A Tabela 2 apresenta os coeficientes de fidedignidade e a VME dos componentes do iUXRV depois da exclusão dos itens que diminuíam a estabilidade do instrumento. Destaca-se ainda que essas estimativas contribuem como **evidências de validade convergente** para o instrumento.

Tabela 2. Estimativas de fidedignidade para a solução de 5 componentes (25 itens). Todas as estimativas estão acima dos valores mínimos desejáveis.

Componente	coeficiente alfa	coeficiente rô _C	VME
Usabilidade	.863	.900	.643
Estética	.916	.938	.752
Emoções	.851	.894	.630
Presença	.892	.921	.700
VR sickness	.803	.880	.603

Em outra perspectiva, a razão de correlações Heterotraço-Monotraço (HTMT) é geralmente utilizada para verificar se componentes diferentes estão medindo conceitos diferentes [Russo e Stol 2021], contribuindo como **evidência de validade discriminante**. Considerando que valores inferiores a .90 são considerados aceitáveis para componentes similares, observa-se na Tabela 3 que apenas o par Estética–Emoções viola esse critério. As implicações práticas desse efeito são discutidas na Seção 5.

Tabela 3. Razão Heterotraço-Monotraço para a solução de 5 componentes usando bootstrapping (25 itens). Todas as razões estão dentro dos limites aceitáveis, exceto aquela entre Estética e Emoções.

	Usabilidade	Estética	Emoções	Presença	VR sickness
Usabilidade	-	-	-	-	-
Estética	.604	-	-	-	-
Emoções	.661	.964	-	-	-
Presença	.229	.329	.349	-	-
VR sickness	.119	.090	.302	.140	-

As **evidências de validade baseadas na relação com outras variáveis** foram estabelecidas nessa pesquisa por meio de investigação da correlação entre os escores de componentes (usabilidade, VR sickness e presença) e os escores de instrumentos desenvolvidos por outros pesquisadores para medirem os mesmos conceitos (SUS, SSQ e PQ). O coeficiente de correlação de Pearson (r) foi calculado para medir a força desse

relacionamento, sendo que intensidades típicas são fraca ($r < .30$), moderada ($.30 \leq r \leq .70$) e forte ($r > .70$) [Reynolds et al. 2021].

Tendo em vista que os instrumentos SUS, SSQ e PQ não foram aplicados para todos os participantes devido ao *design* experimental dessa pesquisa, o número de observações usadas para o cálculo das correlações varia.

A correlação entre o componente de usabilidade e a SUS é considerada forte em nossa amostra ($n = 53$, $r = .761$, $p\text{-value} < .001$), a correlação entre o componente *VR sickness* e o SSQ é considerada moderada ($n = 50$, $r = .644$, $p\text{-value} < .001$) e a correlação entre o componente de presença e o PQ é considerada fraca ($n = 51$, $r = .278$, $p\text{-value} < .05$). Uma hipótese para essa baixa correlação é apresentada na Seção 5.

Finalmente, a **avaliação do modelo estrutural** contribui para estabelecer o iUXRV como um instrumento fundamentado em relações significativas e relevantes entre seus componentes.

Considerando a Figura 1, observa-se a extensão do *framework* CUE pela adição dos componentes de presença e *VR sickness*. Ainda, é importante destacar o papel do componente emoções como mediador parcial entre os componentes de usabilidade, presença, estética e *VR sickness* e o componente de Experiência de Usuário.

A significância de coeficientes de caminho é estimada por meio de *bootstrapping*, indicando que um determinado efeito envolvendo componentes tem relevância e pode ser assumido naquela população [Hair et al. 2021]. Ainda, a relevância de coeficientes de caminho comumente varia entre -1 (uma relação fortemente negativa) e +1 (uma relação fortemente positiva), representando a influência de um componente sobre outro [Hair et al. 2021]. Tendo em vista que não há regras estabelecidas sobre medidas de relevância, sua análise depende do contexto da pesquisa e dos pressupostos dos pesquisadores. A Figura 4 indica que todos os efeitos previstos são significativos, exceto o caminho entre *VR sickness* e UX.

O poder de explicação de um modelo é apresentado como um coeficiente de determinação (R^2), indicando quanto da variação dos dados pode ser explicada pelo modelo. O coeficiente varia entre 0 e 1 e padrões típicos descrevem o poder de explicação como fraco ($.25 \leq R^2 < .5$), moderado ($.5 \leq R^2 < .75$) ou forte ($R^2 \geq .75$) [Hair et al. 2021]. Um coeficiente muito baixo sugere que o modelo não explica adequadamente a variabilidade dos dados devido a uma especificação incorreta, ausência de variáveis importantes ou outros motivos. Por outro lado, um coeficiente extremamente elevado indica que o modelo pode ser especializado demais para a amostra usada (não conseguindo generalizar para outros dados) ou que há multicolinearidade entre os preditores. É essencial destacar que o coeficiente $R^2 = 1$ na componente de UX é um efeito colateral previsto na abordagem de indicadores repetidos [Becker et al. 2012].

5. Discussão

Uma característica importante no processo de refinamento do iUXRV foi a necessidade de agregar os fatores das componentes usabilidade, presença e estética. De acordo com a fundamentação teórica dos componentes apresentada na Seção 3, a usabilidade foi baseada nos fatores de facilidade de aprendizagem, eficiência e eficácia, sendo que cada fator possuía itens próprios no questionário. O mesmo ocorreu para o componente

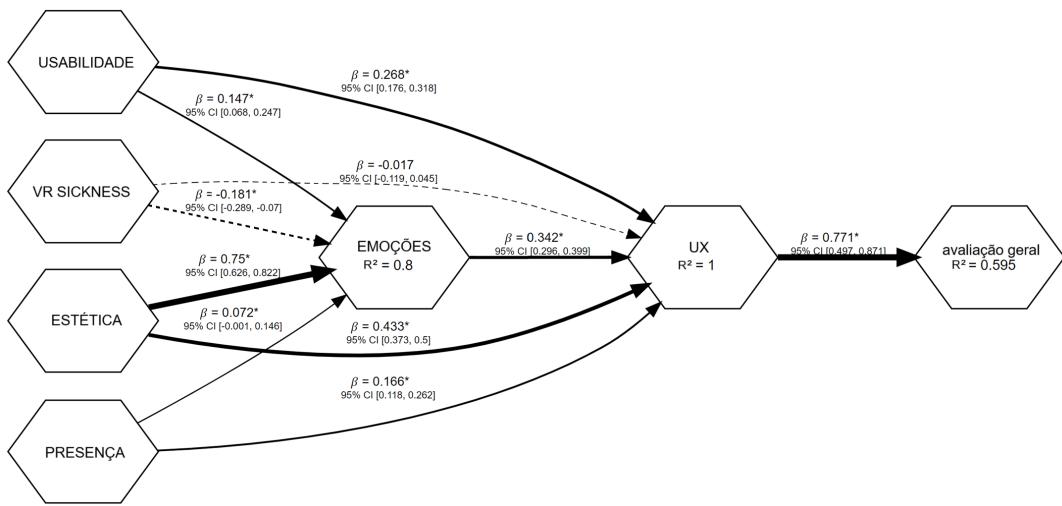


Figura 4. Modelo estrutural do iUXRV com coeficientes de caminho e poder de explicação estimados usando bootstrapping. Coeficientes com asterisco indicam valores estatisticamente significativos, setas mais escuras indicam um efeito maior e setas tracejadas indicam um efeito inverso.

de presença, que é composto por ilusão de lugar e ilusão de plausibilidade, e para o componente de estética, que envolve estética visual clássica e estímulo.

As análises preliminares do modelo envolvendo nove fatores separados (facilidade de aprendizagem, eficiência, eficácia, ilusão de lugar, ilusão de plausibilidade, estética visual clássica, estímulo, VR sickness e emoções) apresentaram baixa VME em vários fatores e elevadas razões HTMT entre os fatores de um mesmo componente. Considerando que esses são indícios substanciais de um agrupamento inadequado dos itens [Russo e Stol 2021, Hair et al. 2021], um modelo contendo cinco componentes mais amplos mostrou-se mais adequado no contexto de medição da Experiência do Usuário. Essa tendência também pode ser observada em outros modelos de UX para RV [Cheng et al. 2014, Shin et al. 2013, Tcha-Tokey et al. 2016, Tcha-Tokey et al. 2018].

Destaca-se que a avaliação qualitativa dos itens a partir dos resultados de PLS-SEM também foi essencial para a manutenção de itens importantes e para o entendimento de algumas remoções. Por exemplo, a despeito de possuir uma carga relativamente baixa (.538), o item Q4 “Eu estou sentindo meus olhos cansados” é crítico na avaliação de sintomas oculomotores, constando em todos os principais questionários para avaliar VR sickness [Sevinc e Ilker 2020]. Por outro lado, o item “Eu estou suando demais”, removido por apresentar carga de indicador baixa, permitiu supor que o suor pode nem sempre representar um sintoma de VR sickness mas, sim, um reflexo do esforço físico ou de fatores ergonômicos (como a proteção de silicone colocada nos óculos de RV).

Outro aspecto importante a ser discutido é a elevada razão HTMT entre estética e emoções apresentada na Tabela 3. Enquanto essa medida sozinha sugere que os dois componentes podem estar medindo a mesma coisa, a análise do texto dos itens e a verificação de colinearidade entre os componentes indicam que eles representam construtos distintos, embora estejam bastante relacionados. Primeiramente, é possível observar que não há sobreposição aparente nos itens de estética (que avaliam se o mundo

virtual é elegante, fascinante, bonito, divertido e empolgante) e de emoções (que avaliam se a pessoa se sentiu bem, contente, irritada, frustrada e feliz). A própria perspectiva de avaliação diferente contribui para afastar a sobreposição dos itens: itens de estética têm perspectiva de avaliação do ambiente e itens de emoções possuem perspectiva de avaliação da própria pessoa e daquilo que ela sentiu. Adicionalmente, a colinearidade (que indica que dois ou mais componentes estão tão correlacionados que podem causar inflação nos coeficientes de caminho) verificada por meio do Fator de Inflação de Variância (FIV) apresentou valor abaixo do limite de 5 para colinearidade crítica ($FIV = 1.542$, FIV limite = 5) [Hair et al. 2021]. Finalmente, coloca-se como hipótese que o “efeito UAU” frequentemente experienciado no primeiro contato com a Realidade Virtual imersiva pode gerar valores elevados simultaneamente nos itens de estética e emoções, causando pouca variabilidade nas respostas e aumentando artificialmente as correlações entre os componentes diferentes.

Embora as correlações entre o componente de usabilidade e a SUS e entre o componente de *VR sickness* e o SSQ tenham sido significativas e elevadas, o componente de presença e o Presence Questionnaire apresentaram uma correlação abaixo do esperado. Uma hipótese para isso é a diferença entre as bases teóricas do iUXRV e o PQ [Witmer et al. 2005]: enquanto o questionário proposto nessa pesquisa usa os conceitos de ilusão de lugar e plausibilidade como únicos fatores para a sensação de presença, o trabalho de Witmer e colegas se estrutura sobre vários componentes (envolvimento, adaptação/imersão, fidelidade sensorial e qualidade de interface). Ainda, versões concorrentes do questionário traduzido para o português apresentam estruturas distintas após avaliação psicométrica do instrumento [Silva et al. 2016, Vasconcelos-Raposo et al. 2021], sugerindo instabilidade estrutural no PQ que já havia sido identificada pelos autores [Witmer et al. 2005].

Um resultado significativo dessa pesquisa é a identificação da importância do componente estético na composição da Experiência de Usuário, sendo esse um componente comumente negligenciado em pesquisas sobre Realidade Virtual imersiva. Na Figura 4, nota-se um coeficiente de caminho significativo e elevado de estética para UX (.433), indicando que o componente influenciado de UX aumenta em .433 unidades de desvio padrão quando o componente preditor de estética aumenta em uma unidade de desvio padrão [Hair et al. 2021]. Dessa forma, um instrumento que não considere aspectos de estética visual e estímulo falha em capturar um componente que contribui significativamente para a experiência geral em Realidade Virtual.

Ainda, observa-se que o componente de *VR sickness* não possui caminho significativo para UX, embora tenha influência negativa sobre as emoções de usuários como previsto em outros modelos de UX para RV [Tcha-Tokey et al. 2018]. Essa configuração indica que o efeito dos sintomas de *VR sickness* é mediado pelo componente de emoções quando esse último está presente no modelo, sugerindo que os escores dos itens de *VR sickness* não contribuem significativamente para a experiência geral. Isso não quer dizer que o componente *VR sickness* não seja importante ou que seus itens devam ser removidos do questionário, mas aponta uma tendência do escore de *VR sickness* ser calculado de forma separada do escore de UX a partir do iUXRV. A aplicação do procedimento de Haberman [Reise et al. 2013] para verificar o potencial dos escores dos componentes individuais como preditores mais precisos de UX (em vez do escore do

questionário como um todo) revelou que a redução proporcional do erro médio quadrático do componente *VR sickness* com base nos escores totais é 10 vezes menor que os demais componentes, ou seja, os itens de *VR sickness* são muito mais precisos quando reportados como um escore do componente isolado do que quando são agrupados com os demais componentes e reportados como escore global de UX. Naturalmente, essa abordagem só é recomendada quando o componente de emoções está incluído no modelo, tendo em vista que é o componente mediador para UX.

Finalmente, os procedimentos para cálculo do escore de Experiência de Usuário devem considerar a intencionalidade dos desenvolvedores. Por exemplo, se a experiência desejada inclui um ambiente opressivo e caótico para um jogo de horror, os itens “Q7. Eu achei que o mundo virtual é elegante” e “Q15. Eu achei que o mundo virtual é bonito” poderiam contribuir negativamente para a pontuação da experiência. Por outro lado, se a experiência desejada inclui um ambiente relaxante e calmo para uma aplicação de meditação, então o item “Q23. Eu achei que o mundo virtual é empolgante” poderia contar negativamente para o escore final.

Propõe-se nesse trabalho a soma simples dos itens individuais (exceto pelos itens 1, 2, 3, 4, 5, 17 e 21 que devem ser invertidos) como estratégia preliminar de cálculo dos escores do iUXRV. Ainda, sugere-se que o valor 25 seja subtraído do resultado da soma anterior para que a faixa de valores possíveis fique exatamente entre zero e 150, levando a uma interpretação mais natural dos escores.

6. Conclusões

Esse artigo apresenta os principais aspectos do desenvolvimento e da avaliação psicométrica de um instrumento em português brasileiro para aferir a Experiência de Usuário em aplicações de RV imersiva: o índice de Experiência de Usuário em Realidade Virtual imersiva (iUXVR). O questionário comprehende 25 itens apresentados na Tabela 4 envolvendo cinco dimensões-chave: usabilidade, estética, presença, *VR sickness* e emoções.

Além da coleção de evidências de validade relacionadas ao conteúdo, à estrutura interna e à relação com outros questionários sustentar o entendimento de que os escores obtidos são representações adequadas da Experiência dos Usuários, as estimativas de fidedignidade identificadas apontam precisão na capacidade de medição do instrumento em um estudo controlado.

A partir dessas características favoráveis de qualidade psicométrica, destaca-se que o iUXRV é uma ferramenta útil no cenário brasileiro, tendo em vista os problemas enfrentados por pesquisadores e desenvolvedores para avaliar experiências no contexto de Realidade Virtual como, por exemplo, a escassez geral de instrumentos em língua portuguesa e a completa ausência de um instrumento em português específico para RV imersiva.

A quantidade reduzida de itens no iUXRV torna mais factível sua aplicação em sessões de avaliação contendo múltiplos cenários quando comparado ao questionário UXIVE que possui 68 itens [Tcha-Tokey et al. 2016]. Ainda, diferente de alternativas existentes [Cheng et al. 2014, Tcha-Tokey et al. 2018], o uso padronizado de um modelo de respostas em escala do tipo Likert em todo o iUXRV reduz a carga cognitiva dos respondentes ao mesmo tempo que evita equívocos nas respostas.

Tabela 4. Versão final do índice de Experiência de Usuário para Realidade Virtual imersiva (25 itens). Itens com asterisco (1, 2, 3, 4, 5, 17 e 21) devem ter o escore invertido.

1. Eu estou me sentindo enjoado.*
2. Eu estou me sentindo tonto.*
3. Eu estou sentindo um mal-estar geral.*
4. Eu estou sentindo meus olhos cansados.*
5. Eu estou com dor de cabeça.*
6. Eu achei fácil fazer as coisas no mundo virtual.
7. Eu achei que o mundo virtual é elegante.
8. Eu cheguei a esquecer que o mundo era apenas virtual.
9. Eu senti bem no mundo virtual.
10. Eu fiz as coisas no mundo virtual com confiança.
11. Eu achei que o mundo virtual é fascinante.
12. Eu cheguei a esquecer que estava usando equipamento de Realidade Virtual.
13. Eu senti contente no mundo virtual.
14. Eu consegui fazer as coisas que queria em poucos passos.
15. Eu achei que o mundo virtual é bonito.
16. Eu acreditei que as coisas no mundo virtual estavam acontecendo de verdade.
17. Eu senti irritado no mundo virtual.*
18. Eu tive controle sobre as coisas que eu fazia no mundo virtual.
19. Eu achei que o mundo virtual é divertido.
20. Eu esqueci totalmente do mundo real enquanto estava no mundo virtual.
21. Eu senti frustrado no mundo virtual.*
22. Eu acho que aprendi rápido a fazer as coisas no mundo virtual.
23. Eu achei que o mundo virtual é empolgante.
24. Eu senti que o mundo virtual poderia ser real.
25. Eu me senti feliz no mundo virtual.

Adicionalmente, a inclusão de componentes-chave de RV garante que os aspectos de presença e *VR sickness* não sejam negligenciados na avaliação da experiência como ocorreria no uso de um questionário de propósito geral como o UEQ [Schrepp et al. 2014, Cota et al. 2014].

Finalmente, nossa proposta estende o modelo CUE [Mahlke 2008], incluindo os componentes de presença e *VR sickness* que são particulares e críticos de RV. A estimativa do modelo estrutural ainda permitiu identificar aspectos essenciais do *framework* aplicado ao contexto de RV imersiva (Figura 4): o forte efeito do componente estético sobre emoções do usuário e sobre a UX como um todo, o baixo impacto do componente de presença sobre as emoções dos usuários, a ausência de uma contribuição significativa e direta de *VR sickness* na UX, o forte poder de explicação de emoções por seus antecedentes (usabilidade, *VR sickness*, estética e presença), e o moderado poder de explicação da avaliação geral (dada por uma nota de zero a dez) a partir da Experiência de Usuário calculada com base nos itens do questionário. Destaca-se que os coeficientes de determinação encontrados são superiores àqueles presentes em modelos de UX para

RV concorrentes [Shin et al. 2013].

Através dessa extensão do modelo CUE, o iUXVR reforça a postura da aplicação de dimensões significativas para contextos especializados (RV imersiva, Realidade Aumentada, jogos, *chatbots* etc.) [Schrepp e Thomaschewski 2019] ao mesmo tempo que preserva a perspectiva de uma coleção mínima de componentes para estimar a experiência.

As principais ameaças à validade dessa pesquisa incluem a adoção de aplicações de RV relacionadas exclusivamente ao lazer, o perfil homogêneo dos participantes e o uso de apenas um tipo de equipamento de RV imersiva. Consequências possíveis dessas limitações são a alteração das estimativas dos modelos, alterando a relevância e a significância das cargas e dos coeficientes para domínios de aplicação sérios, grupos de usuários experientes e tecnologias de simuladores e CAVEs (do inglês Cave Automatic Virtual Environment).

Por fim, trabalhos futuros sobre o iUXRV visam ampliar as evidências de validade sobre o instrumento com a avaliação de outras aplicações de RV imersiva fora do contexto de lazer como, por exemplo, ambientes de treinamento técnico, ferramentas de produtividade, software de visualização de dados, salas de reunião virtuais e outros. A inclusão de uma amostra mais heterogênea de participantes (especialmente participantes que usam frequentemente Realidade Virtual imersiva e pessoas mais velhas) também garantirá uma generalização mais ampla do modelo proposto e permitirá a investigação de diferentes perfis de usuário em uma perspectiva de análise de invariância da medida [Liengaard 2024].

As investigações futuras também incluem a variação do equipamento utilizado, uma vez que a adoção de um único modelo de óculos de Realidade Virtual nesse artigo (assim como a limitação de três aplicações em RV diferentes) teve o objetivo de garantir condições consistentes ao longo dos experimentos e de aumentar o controle sobre variáveis externas como conforto e qualidade geral do hardware.

7. Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Tecnologias de Inteligência Artificial geradora foram usadas exclusivamente para revisão de textos e para apoiar a criação dos textos alternativos para as figuras neste artigo.

Referências

- AERA; APA; NCME (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association, Washington, DC, EUA.
- Barbosa, S. D. J., Silva, B. S. d., Silveira, M. S., Gasparini, I., Darin, T., e Barbosa, G. D. J. (2021). *Interação Humano-Computador e Experiência do Usuário*. autopublicação.
- Becker, J.-M., Klein, K., e Wetzels, M. (2012). Hierarchical latent variable models in pls-sem: Guidelines for using reflective-formative type models. *Long Range Planning*, 45(5):359–394.

- Carvalho, M. R. d., Costa, R. T. d., e Nardi, A. E. (2011). Simulator Sickness Questionnaire: tradução e adaptação transcultural. *Jornal Brasileiro de Psiquiatria*, 60(4):0–5.
- Chang, E., Kim, H. T., e Yoo, B. (2020). Virtual Reality Sickness: A Review of Causes and Measurements. *International Journal of Human-Computer Interaction*, 36(17):1658–1682.
- Cheng, L. K., Chieng, M. H., e Chieng, W. H. (2014). Measuring virtual experience in a three-dimensional virtual reality interactive simulator environment: A structural equation modeling approach. *Virtual Reality*, 18(3):173–188.
- Cota, M. P., Thomaschewski, J., Schrepp, M., e Gonçalves, R. (2014). Efficient Measurement of the User Experience . A Portuguese Version. *Procedia - Procedia Computer Science*, 27:491–498.
- Darin, T., Coelho, B., e Borges, B. (2019). Which Instrument Should I Use? Supporting Decision-Making About the Evaluation of User Experience. In *Design, User Experience, and Usability. Practice and Case Studies*, pages 49–67, Cham, Switzerland. Springer International Publishing.
- Davis, S., Nesbitt, K., e Nalivaiko, E. (2014). A Systematic Review of Cybersickness. In *Proceedings of the Conference on Interactive Entertainment*, pages 1–9, New York, NY, USA. Association for Computing Machinery.
- Grassini, S. e Laumann, K. (2020). Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology*, 11:349.
- Hair, J. e Alamer, A. (2022). Partial least squares structural equation modeling (pls-sem) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics*, 1(3).
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., e Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer, Cham, Switzerland.
- Hassenzahl, M. (2003). The thing and i: Understanding the relationship between user and product. In Blythe, M. A., Overbeeke, K., Monk, A. F., e Wright, P. C., editors, *Funology: From Usability to Enjoyment*, pages 31–42. Springer Netherlands, Dordrecht, Netherlands.
- Hassenzahl, M., Burmester, M., e Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Szwilus, G. e Ziegler, J., editors, *Mensch & Computer 2003: Interaktion in Bewegung*, pages 187–196. Vieweg Teubner Verlag, Wiesbaden.
- ISO (2018). ISO 9241-11:2018 - Ergonomics of human-system interaction - part 11: Usability: Definitions and concepts. Standard, ISO, Geneva, Switzerland.
- ISO (2019). ISO 9241-210:2019 - Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems. Standard, ISO, Geneva, Switzerland.
- Izard, C. E. (1977). *Theories of Emotion and Emotion-Behavior Relationships*, pages 19–42. Springer US, Boston, MA.

- Izard, C. E. (2007). Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. *Perspectives on Psychological Science*, 2(3):260–280.
- Kim, Y. M., Rhiu, I., e Yun, M. H. (2020). A Systematic Review of a Virtual Reality System from the Perspective of User Experience. *International Journal of Human-Computer Interaction*, 36(10):893–910.
- Kitchenham, B. A., Budgen, D., e Brereton, O. P. (2011). Using mapping studies as the basis for further research – A participant-observer case study. *Information and Software Technology*, 53(6):638–651.
- Lavie, T. e Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human Computer Studies*, 60(3):269–298.
- Law, E. L. C., Van Schaik, P., e Roto, V. (2014). Attitudes towards user experience (UX) measurement. *International Journal of Human Computer Studies*, 72(6):526–541.
- Liengaard, B. D. (2024). Measurement invariance testing in partial least squares structural equation modeling. *Journal of Business Research*, 177:114581.
- Lourenço, D. F., Carmona, E. V., e de Moraes Lopes, M. H. B. (2022). Translation and cross-cultural adaptation of the System Usability Scale to Brazilian Portuguese. *Aquichan*, 22(2).
- Mahlke, S. (2008). *User Experience of Interaction with Technical Systems*. Doktor der ingenieurwissenschaften, Technische Universität Berlin, Fakultät V - Verkehrs- und Maschinensysteme, Alemanha.
- Maia, C. L. B. e Furtado, E. S. (2016). A systematic review about user experience evaluation. In Marcus, A., editor, *Design, User Experience, and Usability: Design Thinking and Methods*, pages 445–455, Cham, Switzerland. Springer International Publishing.
- Marques, L., Barcellos, M. P., Gadelha, B., e Conte, T. (2024). Characterizing ux assessment in the context of immersive experiences: A systematic mapping study. *International Journal of Human-Computer Interaction*, 0(0):1–17.
- Martins, A. I., Rosa, A. F., Queirós, A., Silva, A., e Rocha, N. P. (2015). European Portuguese validation of the System Usability Scale (SUS). *Procedia - Procedia Computer Science*, 67:293–300.
- Martirosov, S. e Kopecek, P. (2017). Cyber sickness in virtual reality - Literature review. In Katalinic, B., editor, *Proceedings of the DAAAM International Symposium*, pages 718–726, Vienna, Austria. DAAAM International.
- Melo, B. e Darin, T. (2019). Scope and definition of user experience in brazil: a survey to explore community's perspectives. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*, pages 1–11, New York, NY, USA. ACM.
- Minge, M., Thüring, M., Wagner, I., e Kuhr, C. V. (2016). The meCUE questionnaire: A modular tool for measuring user experience. In Soares, M., Falcão, C., e Ahram, T. Z., editors, *Advances in Ergonomics Modeling, Usability and Special Populations*, pages 115–128, Cham, Switzerland. Springer International Publishing.
- Moraes, R. (1999). Análise de conteúdo. *Revista Educação*, 22(27).

- Nielsen, J. (1993). *Usability Engineering*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- Norman, D. A. (2004). *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books, New York, NY, USA.
- Nur, A. I., Santoso, H. B., e Putra, P. O. H. (2021). The method and metric of user experience evaluation: A systematic literature review. In *Proceedings of the International Conference on Software and Computer Applications*, pages 307–317, New York, NY, EUA. ACM.
- Preece, J., Rogers, Y., e Sharp, H. (2015). *Interaction Design: Beyond Human-Computer Interaction*. Wiley, Hoboken, NJ, 4 edition.
- Reise, S. P., Bonifay, W. E., e Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95(2):129–140.
- Research e Markets (2025). Virtual Reality Market Report by Component, Type, Application, Technology, Regions and Company Analysis 2025-2033. Report, Research and Markets, s.l.
- Reynolds, C. R., Altmann, R. A., e Allen, D. N. (2021). *Mastering Modern Psychological Testing: theory and methods*. Springer, Cham, Switzerland, 2 edition.
- Rivero, L. e Conte, T. (2017). A Systematic Mapping Study on Research Contributions on UX Evaluation Technologies. In *Proceedings of the Brazilian Symposium on Human Factors in Computing Systems*, pages 1–10, New York, NY, USA. ACM.
- Russo, D. e Stol, K. J. (2021). PLS-SEM for software engineering research: An introduction and survey. *ACM Computing Surveys*, 54(4).
- Sauro, J. e Lewis, J. (2016). *Quantifying the user experience: practical statistics for user research*. Morgan Kaufmann, Massachusetts, USA.
- Schrepp, M., Hinderks, A., e Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. In *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*, pages 383–392, Cham. Springer International Publishing.
- Schrepp, M. e Thomaschewski, J. (2019). Design and validation of a framework for the creation of user experience questionnaires. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(7):88–95.
- Sevinc, V. e Ilker, B. M. (2020). Psychometric evaluation of simulator sickness questionnaire and its variants as a measure of cybersickness in consumer virtual environments. *Applied Ergonomics*, 82:1–12.
- Shin, D. H., Biocca, F., e Choo, H. (2013). Exploring the user experience of three-dimensional virtual learning environments. *Behaviour and Information Technology*, 32(2):203–214.
- Silva, G. R., Donat, J. C., Rigoli, M. M., de Oliveira, F. R., e Kristensen, C. H. (2016). A questionnaire for measuring presence in virtual environments: factor analysis of the presence questionnaire and adaptation into Brazilian Portuguese. *Virtual Reality*, 20(4):237–242.

- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B - Biological Sciences*, 364(1535):3549–3557.
- Tcha-Tokey, K., Christmann, O., Loup-Escande, E., e Richir, S. (2016). Proposition and Validation of a Questionnaire to Measure the User Experience in Immersive Virtual Environments. *International Journal of Virtual Reality*, 16(1):33–48.
- Tcha-Tokey, K., Loup-Escande, E., Christmann, O., Canac, G., Farin, F., e Richir, S. (2018). Towards a user experience model in immersive virtual environments. *Advances in Human-Computer Interaction*, 2018:1–10.
- Vasconcelos-Raposo, J., Melo, M., Barbosa, L., Teixeira, C., Cabral, L., e Bessa, M. (2021). Assessing presence in virtual environments: adaptation of the psychometric properties of the Presence Questionnaire to the Portuguese populations. *Behaviour and Information Technology*, 40(13):1417–1427.
- Vermeeren, A., Law, E., Roto, V., Obrist, M., Hoonhout, J., e Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: Current state and development needs. In *Proceedings of the Nordic Conference on Human-Computer Interaction*, pages 521–530, New York, NY, USA. ACM.
- Witmer, B. G., Jerome, C. J., e Singer, M. J. (2005). The Factor Structure of the Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 14(3):298–312.
- Witmer, B. G. e Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240.