UCLM Universidad Castilla-La Mancha

Modelo arquitectónico y experimental para la implementación de entornos inmersivos en realidad virtual para plataformas de e-commerce

Tesis Doctoral

Rubén Grande Muñoz

Directores:
Javier Alonso Albusac Jiménez
José Jesús Castro Sánchez

Departamento de Tecnologías y Sistemas de la Información

Escuela Superior de Informática

Universidad de Castilla-La Mancha

Octubre, 2026



Autor: Rubén Grande Muñoz
Directores: Dr. Javier Alonso Abusac Jiménez – Dr. José Jesús Castro-Sánchez
LICENCIA: ©Rubén Grande Muñoz. La copia y distribución total o parcial de este documento está permitida con fines no comerciales. Todas las copias, totales o parciales, deben citar expresamente el nombre del autor, la Universidad de Castilla-La Mancha, e incluir esta misma licencia. Si se trata de una copia literal, debe añadirse la mención "copia literal". La modificación y traducción de la obra está autorizada con fines no comerciales, siempre que se cite el nombre del autor, la Universidad de Castilla-La Mancha y se indique que el trabajo se basa en la obra original. El trabajo resultante también debe poder reproducirse, distribuirse, comunicarse al público y transformarse en términos similares a los expresados en esta licencia. Este documento ha sido elaborado con LaTeX.
3 de octubre de 2026

Declaración

Declaro que la presente tesis doctoral, titulada **Modelo arquitectónico y experimental para la implementación de entornos inmersivos en realidad virtual para plataformas de e-commerce**, es un trabajo original y fruto de mi esfuerzo personal. Este trabajo no ha sido previamente presentado ni evaluado para la obtención de ningún otro título en esta ni en ninguna otra universidad. Todas las fuentes utilizadas han sido correctamente citadas en el texto mediante referencias bibliográficas. Asimismo, declaro ser el autor principal de los trabajos utilizados en esta tesis doctoral por compendio de publicaciones. En este sentido, se han considerado las siguientes publicaciones en revistas científicas indexadas:

- Rubén Grande, Javier Albusac, David Vallejo, Carlos Glez-Morcillo, and José Jesús Castro-Schez. Performance Evaluation and Optimization of 3D Models from Low-Cost 3D Scanning Technologies for Virtual Reality and Metaverse E-Commerce. Applied Sciences, 14(14):6037, January 2024. ISSN 2076-3417. Q2. IF: 2.5 [1]. doi: 10.3390/app14146037. URL https://www.mdpi.com/2076-3417/14/14/6037. Publisher: Multidisciplinary Digital Publishing Institute.
- Rubén Grande, Javier A. Albusac, David Vallejo, Carlos González-Morcillo, Santiago Sánchez-Sobrino, and José J. Castro-Schez. Virtual Reality Shopping-Insights: A data-driven framework to assist the design and development of Virtual Reality shopping environments. SoftwareX, 27:101874, September 2024. ISSN 2352-7110. Q2. IF: 2.4 [2]. doi: 10.1016/j.softx.2024.101874. URL https://www.sciencedirect.com/science/article/pii/S2352711024002449.
- Rubén Grande, Javier Albusac, Vanesa Herrera, Dorothy Monekosso, Ana De Los Reyes-Guzman, David Vallejo, and J. J. Castro-Schez. Enhancing Hand Interactions and Accessibility in Virtual Reality Environments for Users With Motor Disabilities: A Practical Case Study on VR-Shopping. IEEE Access, 13:44062–44089, 2025. ISSN 2169-3536. Q2. IF: 3.6 [3]. doi: 10.1109/ACCESS.2025.3549527. URL https://ieeexplore.ieee.org/document/10918638.

- R. Grande, J. Albusac, S. Sanchez Sobrino, D. Vallejo, and J. .J. Castro-Schez. Gamifying virtual reality shopping experiences: a case study on a real-world scenario. Virtual Reality, 30(2):65, February 2026. ISSN 1434-9957. Q1. IF: 5 [4]. doi: 10.1007/s10055-026-01323-4. URL https://doi.org/10.1007/s10055-026-01323-4.

A Elena y a mi familia,

por acompañarme durante este viaje y hacerlo realidad.

Agradecimientos

Quiero comenzar esta sección agradeciendo a mi yo pasado, por tomar la decisión de realizar un doctorado. En aquel momento, por azar de la vida, se abrió ante mi dicha vía con el grupo AIR, al que siempre estaré agradecido por la oportunidad y los años pasados y futuros. Esto me dejó clara una enseñanza: a veces, puedes tener todo planificado, pero en el momento menos esperado, no sabrás que te has levantado en un día que cambiará por completo los próximos años de tu vida.

Si bien realizar una tesis no es siempre un camino de rosas, pues a todos nos surgen dudas, antes o después, y nos vemos influenciados por nuestro entorno por mucho que lo queramos negar, es algo que aporta unas enseñanzas que van más allá de lo medible. En definitiva, uno nunca sabe si tomó la decisión correcta, pero por seguro tengo que recorrer este camino me ha aportado enseñanzas permanentes para la vida, no solo en el ámbito personal.

Quiero comenzar dando las gracias a mis directores, Javi y Pepe, por brindarme la posibilidad allá por el 2023 de comenzar en aquel proyecto nacional del que acabó naciendo esta tesis. Muchas gracias por los consejos, ayudas y todo tipo de consideraciones. He tenido la suerte de tener dos grandes directores de tesis no solo a nivel profesional, si no a nivel humano, que es algo que por desgracia, estoy seguro que no todos los doctorandos pueden afirmar. Muchas gracias por los viajes de trabajo compartidos, los congresos a los que hemos asistido y todos los consejos, tanto vitales como profesionales, y por esa dedicación a que me vaya todo bien en el futuro.

En esta misma línea, quiero agradecer al resto de mis compañeros de AIR, en especial a Vanesa, con la que he compartido preocupaciones de doctorando y buenos momentos. A David, Carlos y Santi, por sus consejos y acompañamiento, y por ofrecerme la posibilidad de trabajar en Furious Koalas, enseñándome un ámbito distinto al académico y que tanto me ha aportado, dándome un aire fresco respecto a todo lo que venía trabajando a lo largo de la tesis.

Cómo no agradecer a mi compañera de vida, Elena, aquella persona que me ha visto en mis buenos y malos tiempos, en mis momentos de fortaleza y duda. Gracias por los casi 10 años vividos juntos, por los días que pasamos juntos e incluso por aquellos días en los que sólo nos podemos ver unas pocas horas. Siempre tienes las palabras exactas y sabes que hacer para hacerme reír en los días malos, gracias por todo.

A mis padres, referentes, que me han hecho entender desde pequeño que quien algo quiere, algo le cuesta. Muchas gracias por moldearme, por hacerme la persona que soy hoy en día, por saber tener mano blanda pero también mano dura cuando lo merecía y quizá no era muy consciente de ello. Gracias por toda la ayuda dada para poder hacer mi vida tal y como ha sido hasta ahora. Aunque no estéis muy familiarizados con la tecnología, esta tesis también es vuestra.

A Jesús, por acompañarme incondicionalmente desde mi niñez, por los momentos divertidos y por toda la vida que nos queda por delante. A Laura, por hacer tan feliz a Jesús y por seguir compartiendo viajes, risas y salidas.

A mi familia, por los que están y los que estuvieron, por ser apoyo incondicional no solo en este viaje, si no durante toda la vida.

A la familia que eliges, gracias por las horas incontables de ocio, de risas en Discord o conversaciones en el gimnasio para desconectar. Aunque dispongamos de menos tiempo que hace unos años, no cabe duda que siempre intentamos juntarnos cuando podemos para pasar un buen rato.

A Paulo y Beatriz, por portarse tan bien conmigo y facilitarme todo lo posible mi estancia en Aveiro.

Y por último, quisiera agradecer a la Universidad de Castilla-La Mancha por la concesión en su día del contrato predoctoral para Formación de Personal Investigador (FPI).

Resumen

El comercio electrónico se ha consolidado como un canal de venta indispensable, pero su crecimiento ha dejado en una posición frágil a las pequeñas y medianas empresas. Estas constituyen la mayor parte del tejido empresarial y son un pilar de las economías locales, pero a menudo carecen de los recursos económicos, técnicos y humanos necesarios para competir con los grandes operadores del comercio digital. La pandemia de COVID-19 evidenció esta brecha de forma abrupta: muchos pequeños comercios sin presencia online se vieron forzados a cerrar, mientras que los gigantes del sector consolidaban su dominio. Resulta urgente, por tanto, ofrecer a estos negocios vías de digitalización que no exijan asumir una complejidad técnica que no pueden afrontar.

En este contexto, las tecnologías inmersivas, y en particular la realidad virtual, emergen como un candidato natural para la próxima evolución del comercio electrónico. Frente a las interfaces web tradicionales, permiten explorar y manipular los productos en un espacio tridimensional, acercando la experiencia de compra online a la de una tienda física y enriqueciendo la información que el consumidor puede obtener antes de decidir. Sin embargo, llevar estas tecnologías al pequeño comercio plantea problemas todavía sin resolver: cómo digitalizar catálogos de productos sin recurrir a equipos profesionales costosos, cómo construir entornos de compra virtual usables sin partir de cero en cada proyecto, cómo ofrecer opciones de accesibilidad, y cómo saber si todo ello influye realmente en el comportamiento de compra y no solo en intenciones declaradas.

Ante estas necesidades surge la presente tesis doctoral, que plantea como hipótesis principal que es posible implementar sistemas de compra inmersiva basados en digitalización de bajo coste e interacción gestual inclusiva de forma técnicamente viable, mejorando la experiencia de compra mediante una usabilidad natural cercana al entorno físico, habilitando la recogida de datos de comportamiento más ricos que en los canales web y generando un impacto cuantificable en la compra real del consumidor, con aplicación directa en el pequeño comercio.

Para dar respuesta a esta hipótesis se ha trabajado de forma integral y progresiva. El punto de partida es un modelo arquitectónico para un mercado virtual inmersivo, pensado específicamente para que varias pequeñas empresas unan fuerzas bajo una misma plataforma que integre de forma nativa la realidad virtual, la inteligencia artificial y mecanismos de fortalecimiento del comercio de proximidad. Sobre esa base se ha demostrado que es posible digitalizar productos con tecnologías de escaneo

accesibles desde un teléfono móvil, obteniendo modelos tridimensionales con calidad visual y rendimiento suficientes para dispositivos de realidad virtual comerciales. Se ha desarrollado además un entorno de software abierto que facilita tanto la creación de tiendas virtuales basadas en interacción manual natural como la recogida y el análisis estandarizado del comportamiento del usuario, cubriendo una carencia habitual en este campo, donde las herramientas rara vez se comparten.

La accesibilidad se ha abordado como un requisito desde el inicio y no como un añadido posterior, con el objetivo de abrir la solución planteada a una amplia diversidad de usuarios sin que las limitaciones puedan suponer un condicionante de uso. En estrecha colaboración con el Hospital Nacional de Parapléjicos de Toledo, centro de referencia en el tratamiento de la lesión medular, se han adaptado los mecanismos de interacción por seguimiento de manos para que personas con limitaciones motoras en los miembros superiores puedan completar las tareas propias de una compra virtual, y estas adaptaciones se han probado y validado con pacientes reales. Finalmente, todo el trabajo se ha llevado a un escenario comercial real, un evento abierto al público en el que los participantes compraban con su propio dinero, lo que ha permitido estudiar el comportamiento de compra efectivo y el papel de la gamificación más allá de los entornos controlados de laboratorio. El resultado es una propuesta coherente y validada por partes que sienta las bases técnicas y metodológicas para que el pequeño comercio pueda incorporar estas tecnologías sin asumir su complejidad.

Abstract

E-commerce has become an essential sales channel, but its growth has left small and medium-sized enterprises in a fragile position. These businesses make up most of the commercial fabric and are a cornerstone of local economies, yet, they often lack the financial, technical, and human resources needed to compete with the large operators of online retail. The COVID-19 pandemic exposed this gap abruptly: many small shops without an online presence were forced to close, while the sector's giants consolidated their dominance. There is, therefore, an urgent need to offer these businesses paths to digitalization that do not require them to take on technical complexity they cannot manage.

In this context, immersive technologies, and virtual reality in particular, emerge as a natural candidate for the next stage in the evolution of e-commerce. Unlike traditional web interfaces, they allow products to be explored and manipulated in a three-dimensional space, bringing the online shopping experience closer to that of a physical store and enriching the information a consumer can gather before deciding. Bringing these technologies to small businesses, however, raises problems that remain unsolved: how to digitize product catalogs without resorting to costly professional equipment, how to build usable virtual shopping environments without starting from scratch in every project, how to avoid excluding people with motor impairments, and how to know whether all of this truly affects purchasing behaviour rather than merely stated intentions.

These needs converged in this doctoral thesis, whose main hypothesis is that immersive shopping systems based on low-cost digitalization and inclusive gesture-based interaction can be implemented in a technically feasible way, improving the shopping experience through a natural usability close to the physical setting, enabling the collection of richer behavioural data than web channels, and producing a measurable impact on real consumer purchasing, with direct application to small retail.

To address this hypothesis, the work has proceeded in an integrated and progressive manner. The starting point is an architectural model for an immersive virtual marketplace, designed specifically so that several small businesses can join forces under a single platform that natively integrates virtual reality, artificial intelligence, and mechanisms to strengthen local commerce. On this foundation, it has been shown that products can be digitized with scanning technologies accessible from a mobile phone, yielding three-dimensional models with visual quality and performance

sufficient for commercial virtual reality headsets. An open software environment has also been developed that supports both the creation of virtual stores based on natural hand interaction and the standardized collection and analysis of user behaviour, addressing a common shortcoming in this field, where tools are rarely shared.

Accessibility has been treated as a requirement from the outset rather than an afterthought, with the aim to make the proposed solution accessible to a wide range of users without any limitations restricting its use. In close collaboration with the National Hospital for Paraplegics in Toledo, a leading center in the treatment of spinal cord injuries, hand-tracking interaction mechanisms have been adapted so that people with upper-limb motor impairments can complete the tasks involved in virtual shopping, and these adaptations have been tested and validated with real patients. Finally, the entire body of work was taken to a real commercial scenario, a public event in which participants shopped with their own money, making it possible to study actual purchasing behaviour and the role of gamification beyond controlled laboratory settings. The outcome is a coherent, module-by-module validated proposal that lays the technical and methodological groundwork for small businesses to adopt these technologies without bearing their underlying complexity.

Índice general

1 Introducción **3**

1.1. Contexto general 3

1.1.1. Pequeñas empresas 4

1.1.2. Efectos de la pandemia del COVID-19 6

1.1.3. Tecnologías inmersivas en el campo del *e-commerce* . . . 8

2 Estado del arte y motivación **15**

2.1. Estado del arte 15

2.1.1. *E-commerce*, transformación digital y *e-marketplaces* 15

2.1.2. Tecnologías inmersivas en *e-commerce* 16

2.1.3. Digitalización de productos en entornos para VR-Shopping 19

2.1.4. Monitorización y análisis de comportamiento de usuario en entornos VR 21

2.1.5. Mecanismos de interacción y accesibilidad en entornos VR 24

2.1.6. Impacto de la gamificación y VR en la decisión de compra 27

2.2. Motivación 29

3 Hipótesis y objetivos **33**

3.1. Preguntas de investigación 33

3.2. Objetivos 34

3.2.1. Objetivos principales 34

3.2.2. Objetivos específicos 35

4 Plan de trabajo **41**

5 Resultados **49**

5.1. Artículos publicados que forman parte del compendio 50

5.2. Otras publicaciones 161
5.2.1. Artículos publicados en revistas científicas indexadas . . 161
5.2.2. Congresos internacionales 162
5.2.3. Congresos nacionales 172
5.2.4. Capítulos de libro 174
5.3. Otras actividades 176
5.4. Premios obtenidos 177

6 Discusión de resultados **179**
6.1. RQ1: marco arquitectónico 181
6.1.1. Estado actual de los *e-marketplaces* 182
6.1.2. Evolución iterativa de la propuesta 182
6.1.3. Formalización del marco arquitectónico 184
6.1.4. Nivel de madurez 187
6.1.5. Posicionamiento frente a la evolución tecnológica reciente 188
6.1.6. Síntesis de la respuesta a RQ1 188
6.2. RQ2: digitalización 3D de bajo coste 190
6.2.1. La digitalización 3D en el entorno PYME 190
6.2.2. Selección de las tecnologías representativas (SO2.1) . . . 191
6.2.3. Calidad visual y error geométrico (SO2.2) 192
6.2.4. Rendimiento en dispositivos VR comerciales (SO2.3) . . 193
6.2.5. Optimización de modelos 3D obtenidos (SO2.4) 195
6.2.6. Posicionamiento frente a la evolución del campo y limitaciones reconocidas 196
6.2.7. Síntesis de la respuesta a RQ2 197
6.3. RQ3: mecanismos de interacción y navegación 198
6.3.1. Interacción natural en dispositivos de VR 198
6.3.2. VRSI como marco de trabajo unificado 199
6.3.3. Interacciones con productos (SO3.1) 200
6.3.4. Navegación (SO3.2) 200
6.3.5. Estructura espacial del entorno (SO3.3) 201
6.3.6. Validación empírica de los mecanismos 202
6.3.7. Limitaciones reconocidas 204
6.3.8. Síntesis de la respuesta a RQ3 205
6.4. RQ4: adaptabilidad de mecanismos de interacción 205

6.4.1. La accesibilidad como problema técnico 206
6.4.2. Análisis de las limitaciones motoras y diseño de los mecanismos adaptados (SO4.1 y SO4.2) 207
6.4.3. Métricas derivadas del hand-tracking: DI y ED (SO4.3) . 209
6.4.4. Caso de estudio en el HNPT (SO4.4) 210
6.4.5. Validación subjetiva: NASA-TLX, SUS, UEQ y MSAQ . . 213
6.4.6. Limitaciones reconocidas 215
6.4.7. Síntesis de la respuesta a RQ4 215
6.5. RQ5: recolección y análisis de datos comportamentales 216
6.5.1. Utilidad de los datos comportamentales 216
6.5.2. Formalización de los eventos de actividad 218
6.5.3. VRSI como marco de trabajo 219
6.5.4. Configuración del entorno de compra (SO5.1) 219
6.5.5. Recolección de datos no invasiva (SO5.2) 221
6.5.6. Estadísticas de comportamiento (SO5.3) 222
6.5.7. Validación de uso por terceros 224
6.5.8. Posicionamiento frente a la literatura y limitaciones . . . 225
6.5.9. Síntesis de la respuesta a RQ5 226
6.6. RQ6: VR y gamificación en compras reales 227
6.6.1. El vacío en la literatura: VR, gamificación y compra real 228
6.6.2. Diseño de los dos escenarios experimentales (SO6.1) . . 228
6.6.3. Caso de estudio (SO6.2) 229
6.6.4. Determinantes actitudinales de la intención de compra . 230
6.6.5. Análisis exploratorio: perfiles comportamentales y compra (SO6.3) 231
6.6.6. Predicción exploratoria de la compra (SO6.4) 233
6.6.7. Hallazgos secundarios y consistencia transversal con otras RQ 235
6.6.8. Limitaciones reconocidas 235
6.6.9. Síntesis de la respuesta a RQ6 236

7 Conclusiones y trabajo futuro **239**
7.1. Respuesta a las preguntas de investigación 239
7.2. Cumplimiento de los objetivos 243
7.3. Líneas de trabajo futuro 244

8 Conclusions and Future Work **246**
8.1. Answering the research questions 246
8.2. Fulfilment of the objectives 250
8.3. Lines of future work 250

Bibliografía **252**

Índice de figuras

1.1. Porcentaje de PYMEs en cada sector económico 6

3.1. Relación entre subobjetivos y preguntas de investigación 39

4.1. Planificación de las tareas de la tesis doctoral 42

5.1. Publicaciones realizadas durante la tesis doctoral 49
5.2. Interfaz web del portal VR-ZOCO 163
5.3. Arquitectura de VR-ZOCO 164
5.4. Composición de fotografías tomadas en reuniones y eventos durante el proyecto VR-ZOCO 176

6.1. Cadena de validación de la tesis: flujo de artefactos entre los artículos del compendio 180
6.2. Arquitectura de VR-ZOCO para dar respuesta a RQ1 184
6.3. Captura de un prototipo VR presentado en [5] 185
6.4. Captura del prototipo web presentado en [6] 187
6.5. Materialización de los módulos de VR-ZOCO a lo largo de la tesis . 189
6.6. Calidad visual por configuración del modelo 193
6.7. Trayectoria de rendimiento por carga poligonal y headset 195
6.8. Catálogo de interacciones implementadas en VRSI 201
6.9. Validación empírica de los mecanismos de VRSI 203
6.10. Marco anatómico de los mecanismos adaptados 208
6.11. Resultados cuantitativos del caso de estudio HNPT 211
6.12. Volúmenes de movimiento adaptado y no adaptado por usuario . . 213
6.13. Arquitectura general del framework VRSI 220
6.14. Ejemplos de mapas de calor volumétricos generados por VRSI . . . 223
6.15. Tasa de compra y perfiles comportamentales por cluster 232

Índice de tablas

5.1. Tabla resumen de los artículos científicos incluidos en el compendio de la tesis doctoral. 50
5.2. Publicaciones en actas de congresos internacionales 165
5.3. Publicaciones en actas de congresos nacionales 172

6.1. Contribuciones a RQ1 181
6.2. Contribuciones a RQ2 190
6.3. Contribuciones a RQ3 198
6.4. Contribuciones a RQ4 206
6.5. Contribuciones a RQ5 217
6.6. Datos capturados por el componente Data Collection de VRSI . . . 221
6.7. Contribuciones a RQ6 227

7.1. Cumplimiento de los objetivos específicos 244

8.1. Fulfilment of the specific objectives 250

Lista de acrónimos

6-DOF 6-Degrees of Freedom 22
AEI Agencia Estatal de Investigación 12
AI Artificial Intelligence 11
AIR Artificial Intelligence and Rendering 11
AOI Area of Interest 202
AR Augmented Reality 8
DVR Desktop Virtual Reality 17
EU Unión Europea 5
FPI Formación de Personal Investigador 12
HCI Human-Computer Interaction 8
HNPT Hospital Nacional de Parapléjicos de Toledo 37
IVR Immersive Virtual Reality 17
MCP Metacarpophalangeal/Metacarpofalángica 209
MR Mixed Reality 8
PYMEs Pequeñas y Medianas Empresas 3
SDK Software Development Kit 27
SVM Support Vector Machine 22
UCLM Universidad de Castilla-La Mancha 11
VR Virtual Reality 8
ZOI Zone of Interest 202

CAPÍTULO 01

01

Introducción

Contexto, problema y propósito de la tesis.

1

Introducción

1.1 | Contexto general

Desde principios de la década de 2010, la evolución tecnológica y la digitalización han cambiado significativamente el perfil de comportamiento del consumidor, desplazando las tendencias tradicionales de compra en tiendas físicas hacia el comercio digital o *e-commerce* [7]. Este cambio fue impulsado fundamentalmente por la adopción generalizada de Internet y los teléfonos inteligentes, lo que ha facilitado el auge del comercio móvil (*m-commerce*) y ha permitido a los usuarios realizar transacciones de manera conveniente en cualquier momento y lugar [8, 9]. Esta evolución tecnológica ha modificado la forma en la que, normalmente, enfocamos el proceso de compra. Primero, tendemos a buscar en un navegador web información sobre lo que queremos y, basándonos en reseñas, las características detalladas del producto y contenido multimedia. Después, muchos deciden realizar la compra a través de la web mientras que otros, cada vez menos, optan por analizar los productos en la tienda física correspondiente.

En este nuevo escenario, la implementación de plataformas de *e-commerce* se ha consolidado como una estrategia determinante para la supervivencia y el crecimiento empresarial, independientemente de la envergadura de la organización [10, 11]. El *e-commerce* puede utilizarse como un mecanismo democratizador que permite, especialmente a las Pequeñas y Medianas Empresas (PYMEs)[1], superar las barreras geográficas tradicionales y acceder a una audiencia global, nivelando así la diferencia competitiva frente a corporaciones

[1] Este término engloba: (1) Microempresa, menos de 10 empleados y menos de 2 millones de euros de facturación anual, (2) empresa pequeña, entre 10 y 50 empleados y menos de 10 millones de euros de facturación anual, y (3) empresa mediana, entre 50 y 250 empleados y menos de 50 millones de euros facturados anualmente.

de mayor tamaño [8, 11]. Además de la expansión del mercado, la adopción de canales de venta *online* proporciona ventajas competitivas tangibles a nivel operativo, tales como la reducción de costes de transacción, particularmente en marketing y distribución, y la optimización de la gestión de inventarios y procesos logísticos [7, 12]. De esta manera, la digitalización de las ventas no solo facilita la diversificación de las fuentes de ingresos, sino que se convierte en un catalizador indispensable para mejorar la eficiencia comercial y responder ágilmente a la demanda en una economía cada vez más interconectada [9, 12].

A diferencia de las tiendas físicas, el *e-commerce* ofrece ventajas competitivas claras, como el ahorro de tiempo y costes de transporte, la capacidad de comparar precios y obtener información detallada de productos de manera eficiente, y la eliminación de barreras geográficas para acceder a un catálogo mayor de productos [12, 13]. Como resultado, el modelo tradicional basado en establecimientos físicos ha sufrido un impacto significativo. A menudo, las plataformas digitales satisfacen las expectativas del cliente de manera más efectiva y fluida [9], obligando a los minoristas tradicionales a adaptar sus estrategias para sobrevivir en un entorno donde los consumidores prefieren la accesibilidad y la eficiencia que ofrece el e-commerce [7, 9, 12] a las que pueden acceder desde la comodidad de sus casas.

1.1.1 | Pequeñas empresas

La evolución de las soluciones de *e-commerce* y su consolidación han supuesto un gran impacto en nuestra sociedad actual. Esta coyuntura beneficiosa para el mercado digital conlleva factores de riesgo inherentes, tales como una mayor probabilidad de desaparición de los pequeños comercios físicos que no logran adaptarse a la presión competitiva y la eficiencia de costos que imponen los gigantes del mundo online como pueden ser Amazon (EE.UU.), Alibaba (China) o Rakuten (Japón) [7, 9]. Este riesgo es notablemente mayor para empresas en sectores que ofrecen productos no diferenciados, como la moda, la tecnología, la literatura o el entretenimiento; por ejemplo, se ha observado cómo la innovación tecnológica y las plataformas de *streaming* han desplazado modelos de negocio tradicionales, como las salas de cine, al ofrecer alternativas más accesibles y económicas, eliminando intermediarios físicos [7].

A pesar de reconocer que la digitalización es crítica para la viabilidad a

largo plazo en la nueva economía, las PYMEs se enfrentan a barreras que a menudo provocan el fracaso en la adopción del comercio electrónico. La limitación más significativa radica en la escasez de recursos internos, tanto financieros como humanos [14], lo que dificulta la inversión en infraestructura tecnológica moderna y la contratación de especialistas cualificados necesarios para gestionar sistemas complejos [15, 16]. Esta carencia se ve agravada por una brecha de conocimiento, donde la falta de comprensión sobre tecnologías emergentes y conceptos de la Industria 4.0, sumada a la insuficiencia de competencias digitales en la fuerza laboral, impide una transición efectiva [14]. Además, factores organizativos como la tendencia a operar con planes de desarrollo a muy corto plazo —a menudo inferiores a un año— y el temor a la pérdida de seguridad de los datos en soluciones en la nube actúan como frenos adicionales [15]. En consecuencia, muchas PYMES fallan en la implementación porque, a diferencia de las grandes corporaciones, carecen de la liquidez y la capacidad dinámica para adaptar rápidamente sus modelos de negocio y procedimientos operativos a los desafíos de un entorno digital competitivo.

Además, debe tenerse en cuenta que el declive del modelo tradicional basado en tiendas puede desencadenar efectos en cadena, donde el cierre de un negocio afecta a otros establecimientos cercanos de sectores complementarios, como el ocio, alterando significativamente las economías locales y la dinámica social al reducir la afluencia de consumidores en espacios físicos [7, 9, 12].

A lo largo de las dos últimas décadas, se han llevado a cabo iniciativas y esfuerzos para promover la digitalización de las PYMEs, tanto por entidades públicas como privadas, entre las que podemos destacar el reciente Kit Digital [2]. Si se amplía el alcance a los 27 países de la Unión Europea (EU), estos esfuerzos están motivados por la importancia de las PYMEs en la economía. De acuerdo con el más reciente informe anual de PYMEs realizado por la Comisión Europea, la EU cuenta con un total aproximado de 26.133.280 PYMEs, entre las cuales 24.514.649 son microempresas y 1.404.631 son pequeñas empresas [17]. Estas constituyen el 99.8 % de todas las empresas de la EU, generando el 53.6 % del valor añadido del sector, aproximadamente 4.1 billones de euros [17]. La Fig. 1.1 muestra el porcentaje del total de PYMEs dedicada a cada sector, entre los que destaca como segundo mayor sector el del comercio minorista con un 21.9 %. Esta cifra pone de manifiesto la importancia de dicho sector para la

[2] `https://www.acelerapyme.gob.es/kit-digital`

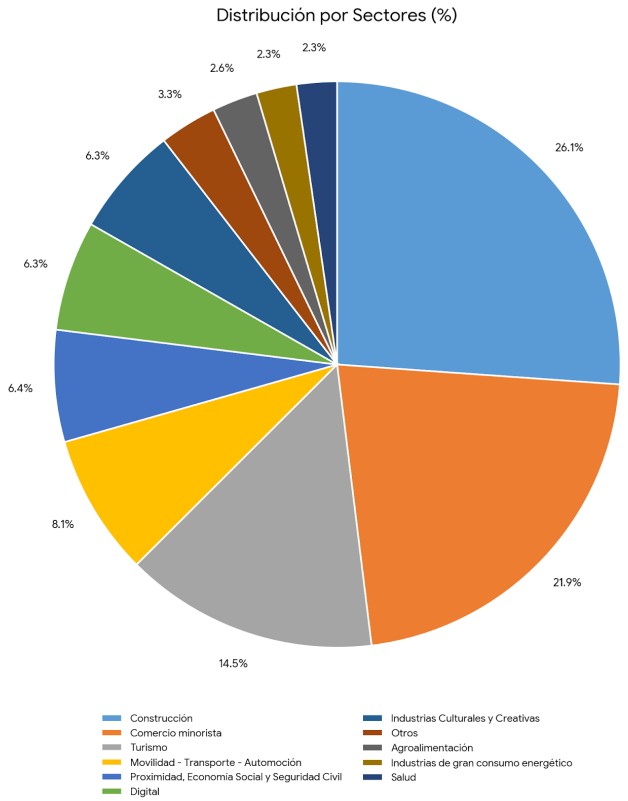


Figura 1.1: Porcentaje del número total de PYMEs destinada a cada sector económico en la EU [17].

economía y la sociedad, no solo a nivel nacional, sino a nivel de la EU.

1.1.2 | Efectos de la pandemia del COVID-19

El año 2020 quedó marcado en la historia por la pandemia del COVID-19. Un hecho sin precedentes que actuó como catalizador en el campo del *e-commerce* a nivel mundial [12, 18, 19]. Las restricciones de movilidad, las medidas de distanciamiento social y los confinamientos obligaron a los consumidores a migrar masivamente hacia canales digitales para satisfacer sus necesidades, convirtiendo al *e-commerce* en un servicio esencial y no solo en una alternativa conveniente [12]. Este cambio forzó a las empresas a reevaluar sus estrategias comerciales, evidenciando que la preparación digital había dejado de ser una ventaja competitiva opcional para convertirse en un requisito fun-

damental para la continuidad del negocio en tiempos de crisis [9, 12]. Para las pequeñas y medianas empresas, este escenario presentó un desafío existencial: a pesar de enfrentar interrupciones en la cadena de suministro y dificultades financieras, se vieron presionadas a superar las barreras de adopción tecnológica rápidamente para evitar la obsolescencia o el cierre [12, 16].

Este hecho se evidenció considerando que, en todo el año 2020, el número de ventas realizadas *online* aumentó un 41 %, mientras que las físicas, en función del país, descendieron entre un 20 % y un 45 % [20]. En el caso de los pequeños comercios, como ya hemos visto anteriormente, el grado de digitalización y las barreras asociadas a su implantación les perjudicó gravemente en este periodo. Por ejemplo, en el caso de Indonesia, un 42 % de las PYMEs sin presencia online se vieron forzadas a parar sus operaciones en la pandemia, mientras que, aquellos que disponían de servicios de *e-commerce*, solo supuso el 24 % [21]. A pesar de la limitación de recursos, diversas investigaciones confirman el interés de las pequeñas empresas por la transformación digital. Un ejemplo notable ocurrió en Bélgica tras la pandemia, donde se registró un incremento del 30 % en el número de pequeñas empresas con presencia en la red [18].

El efecto de la pandemia no fue exclusivo del año 2020, sino que fue el motor del cambio de la forma que consumimos hasta el día de hoy. Los ingresos globales del comercio electrónico se dispararon de 1,4 billones a 2,4 billones de dólares a medida que la pandemia se intensificaba, consolidando un mercado donde aproximadamente 3.500 millones de personas (cerca del 47 % de la población mundial) utilizan plataformas digitales diariamente [12]. El impacto sectorial fue igualmente contundente, registrándose un crecimiento del 65 % en la compra de comestibles en línea en Estados Unidos, mientras que en economías avanzadas como la japonesa, el auge del e-commerce contrastó severamente con una caída del 13.4 % en las ventas minoristas tradicionales y un aumento correlativo del 11.6 % en el volumen de entrega de paquetería [12, 19, 20]. En este entorno de polarización digital, gigantes tecnológicos como Amazon lograron un incremento de ingresos del 20.5 % solo en 2020, evidenciando la inmensa brecha de rendimiento frente a los modelos de negocio que no pudieron digitalizarse a tiempo [22].

Además, la evolución de los hábitos de consumo hacia el comercio electrónico persiste de forma generalizada en todos los rangos generacionales.

Según estudios realizados en Estados Unidos, el volumen de ventas digitales creció cinco veces más rápido que el de las ventas presenciales tras la crisis sanitaria [23]. Y más del 65 % de los consumidores, desde la Generación Z hasta los Baby Boomers, han adoptado el e-commerce como su principal canal de compra. Esta tendencia resulta preocupante para el comercio local con un nivel bajo de digitalización, dado que su base de clientes tradicional —compuesta mayoritariamente por personas de edad avanzada, incluso en zonas rurales— ahora recurre con frecuencia a internet para consultar y adquirir productos [24].

1.1.3 | Tecnologías inmersivas en el campo del *e-commerce*

La consolidación del *e-commerce* como canal de venta preferido por el consumidor promedio tras la pandemia no ha eliminado su limitación estructural más citada desde la perspectiva del comportamiento del consumidor: la reducción de la "experiencia sensorial" disponible durante la evaluación del producto (p. ej., inspección material, escala real, manipulación, prueba previa). En paralelo, la pandemia aceleró la migración de actividades presenciales a entornos digitales y modificó hábitos de compra, reforzando la necesidad de soluciones que preserven eficiencia sin sacrificar la calidad de la experiencia de compra [20]. En este contexto, las tecnologías inmersivas, principalmente Virtual Reality (VR), Augmented Reality (AR) y Mixed Reality (MR) emergen como un candidato natural para evolucionar esta experiencia en el *e-commerce*. Aun así, en el estado actual de estas tecnologías su objetivo no es sustituir el canal web, aportan una serie de beneficios únicos al proceso de compra, ampliándolo. Este hecho se soporta con el estándar WebXR[3] del W3C, formalizando el soporte de VR y AR en navegadores web. Así pues, su uso proporciona un conjunto de interacciones inmersivas (visualización 3D, manipulación y prueba virtual) y con impacto directo en constructos claves usados en experimentaciones en el campo de la Human-Computer Interaction (HCI) como presencia, control percibido, carga cognitiva, confianza y riesgo percibido [25–27].

En esta línea, la aportación diferencial de VR/AR frente al *e-commerce* 2D radica en su capacidad para incrementar la diagnosticidad de la informa-

[3] `https://www.w3.org/TR/webxr/`

ción del producto mediante interacción y espacialidad: el usuario pasa de únicamente leer/ver contenido multimedia 2D a explorar el producto en un espacio egocéntrico, con referencias de tamaño, distancia, oclusión y perspectiva, y con acciones de inspección (rotación, zoom, apertura/cierre, combinación/configuración) que simulan parte del análisis físico previo a la compra. Esta transición se apoya en mecanismos psicológicos ampliamente estudiados en HCI y marketing digital, como la telepresencia, que tienden a mejorar la vividez y la comprensión del producto y a reducir incertidumbre cuando la representación y la interacción son creíbles [26]. Además, la evidencia empírica reciente muestra que entornos de tienda VR y experiencias inmersivas pueden modificar percepciones y comportamientos a través de mediadores experienciales como el disfrute, inmersión, presencia o la carga cognitiva [28, 29]. En AR, los escenarios de *virtual try-on*, especialmente en el campo de la moda y estética, son especialmente relevantes porque aproximan la prueba del producto al contexto corporal del usuario, donde dimensiones de presencia se han relacionado con intención de compra y evaluación del producto [30].

Este potencial se alinea con la lógica defendida por Kim [23] en el marco de la recuperación del sector *retail* tras la COVID-19: la VR puede actuar como mecanismo de continuidad y renovación del *e-commerce* al ofrecer un modelo de experiencia de compra alternativa a la tienda física cuando la movilidad o el contacto son limitantes, y, a la vez, como instrumento de diferenciación competitiva en un escenario post-pandemia dominado por la eficiencia digital. Esto fundamenta la siguiente hipótesis: si el *e-commerce* se ha convertido en estándar, la variable estratégica deja de ser "estar online" y pasa a ser "cómo diseñar experiencias de compra digitales de alta calidad", donde VR/AR aportan capacidades de interacción, exploración y contextualización del producto que el canal 2D no ofrece de forma nativa.

Adicionalmente, existen señales de madurez tecnológica que justifican considerar VR/AR como el próximo gran salto tecnológico en el campo o, al menos, como tecnología con ese potencial. Por ejemplo, la Comisión Europea está impulsando un ecosistema [4] XR europeo con foco explícito en industria y PYMEs, reforzando la plausibilidad de adopción progresiva en el tejido productivo. Como vimos en la Sección 1.1.1, gran parte del comercio minorista europeo se estructura a través de pequeñas empresas con limitaciones estruc-

[4] `https://digital-strategy.ec.europa.eu/en/policies/extended-reality`

turales (capacidad inversora, personal especializado, productividad relativa) que condicionan la adopción de tecnologías avanzadas. En informes recientes sobre desempeño y brechas de productividad, se destaca la persistencia de diferencias entre microempresas y grandes empresas, lo que sugiere que la digitalización avanzada debe diseñarse con criterios de accesibilidad económica y operativa [17]. Así pues, si estas tecnologías van a representar el próximo avance tecnológico en el *e-commerce*, es una gran oportunidad para que las PYMEs alcancen un mayor grado de digitalización con vistas al futuro que les permita, como mínimo, mantenerse competitivas con respecto a las grandes empresas que ya exploran estas opciones. Si estas tecnologías se implementan mediante modelos arquitectónicos adecuados, pueden actuar como tecnología democratizadora: el valor no depende de que cada PYME construya un metaverso propio, sino de habilitar artefactos compartidos, p. ej., tiendas virtuales, catálogos 3D, servicios de recomendación o análisis y recolección de datos que reduzcan barreras de entrada y externalicen complejidad técnica. Esta idea está orientada al pequeño comercio, donde el modelo propuesto libera a las empresas de tareas técnicas difíciles y permite agrupar oferta en un mismo entorno virtual accesible mediante VR, soportando además técnicas de cross-selling/up-selling y recomendaciones asistidas por Inteligencia Artificial (IA).

Desde la perspectiva de la HCI, este enfoque de plataforma compartida tiene implicaciones directas y necesidades para su diseño:

- Análisis de las tendencias actuales de plataformas *e-commerce* y la propuesta de un modelo arquitectónico que aglutine y conecte estas nuevas tecnologías.

- Generación de catálogos de productos 3D mediante tecnologías de digitalización de bajo coste, accesibles y que ofrezcan resultados de alta calidad gráfica.

- Un *framework* que permita implementar, de forma sencilla y abstrayendo dificultades técnicas, entornos virtuales inmersivos enfocados en el shopping (VR-shopping), además de permitir la recogida y análisis de datos adaptado al entorno construido.

- Explorar cómo ofrecer opciones de accesibilidad dentro de entornos virtuales, con el fin de no excluir futuros usuarios y democratizando el uso de estas tecnologías inmersivas.
- Validar de forma prematura mediante sesiones experimentales distintas estrategias, decisiones y técnicas que permitan obtener una percepción positiva de los usuarios a nivel de usabilidad y novedad tecnológica, que pueda impactar en indicadores clave como el ratio de conversión.

Estas consideraciones conectan con factores clásicos de éxito del e-commerce en PYMEs que, trasladados a espacios virtuales inmersivos, se traducen en requisitos de arquitectura y experiencia de usuario, tanto el final como el desarrollador.

Esta tesis doctoral nace en el marco de un proyecto destinado a contribuir a la solución de un problema actual dentro del programa de doctorado en Tecnologías de la Información Avanzadas en la Universidad de Castilla-La Mancha (UCLM) y desarrollada dentro del grupo de investigación Artificial Intelligence and Rendering (AIR). Una de las principales líneas de investigación del grupo es la aplicación de tecnologías innovadoras, como las inmersivas y la Artificial Intelligence (AI) en el ámbito del comercio electrónico.

En este contexto, y con la experiencia previa de miembros del grupo en proyectos de investigación en el campo del comercio electrónico, se plantea la oportunidad de investigar el diseño, desarrollo y validación experimental prematura de un modelo arquitectónico que permita implementar entornos inmersivos de compra (VR-Shopping) de forma efectiva. Una de las principales motivaciones en este ámbito es la propuesta de un modelo que permita a las PYMEs aunar fuerzas bajo una plataforma tecnológica preparada para el próximo salto tecnológico en el campo del comercio, teniendo así la oportunidad de evolucionar mediante una herramienta con la que poder seguir siendo competitivos en el futuro. Las limitaciones a nivel de digitalización de las PYMEs se tienen en cuenta desde el principio, así como la necesidad de estudiar opciones de accesibilidad que permitan que ninguna persona pueda quedarse atrás en el uso de estas nuevas tecnologías.

Esta tesis se desarrolló en el contexto del proyecto **VRZOCO[5]: Diseño y desarrollo de una plataforma basada en VR-Shopping e IA para la digita-**

[5]https://air.esi.uclm.es/vrzoco/

lización y el fortalecimiento de negocios y economías locales (1/12/2022 - 30/04/2025), financiado por el Ministerio de Ciencia, Innovación y Universidades (MCIU) y la Agencia Estatal de Investigación Agencia Estatal de Investigación (AEI). Se trata de un proyecto nacional financiado por los fondos NextGeneration de la EU. Dicho proyecto estuvo financiado con 175.000,00 € a 2 años (01/12/2022 hasta 30/11/2024), extendido con concesión de prórroga hasta el 30/04/2025. Dentro del desarrollo de esta tesis, su autor, ha sido beneficiado inicialmente con un contrato de investigación en el propio proyecto y, posteriormente, con un contrato de Formación de Personal Investigador (FPI) en el Plan Propio de I+D+i de la UCLM, enmarcando esta tesis en la línea estratégica S3 de Turismo, Cultura y Patrimonio Cultural.

La motivación de esta investigación se encuentra justificada tras la detección de los problemas existentes en digitalización en las PYMEs y su impacto económico a nivel global. Como se detalló en la Sección 1.1, ante el escaso grado de digitalización, consideramos conveniente investigar soluciones que permiten abstraer a las PYMEs de las complejidades inherentes de este proceso, pero apostando por las nuevas tecnologías que marcarán el porvenir del campo del *e-commerce*. Así, no hará falta que ocurra otro evento de cisne negro como lo fue el COVID-19 para poner de manifiesto las carencias y el impacto en la competitividad de las PYMEs, pues ya habrá evidencia científica de cómo pueden usar e implantar estas tecnologías de forma eficaz.

Cabe destacar que el pequeño comercio, años previos a la pandemia mencionada, ha realizado grandes esfuerzos (con apoyo económico de iniciativas públicas y privadas) para intentar digitalizarse en distintos niveles. Sin embargo, la evolución tecnológica es constante, por lo que un nuevo avance tecnológico en el campo del e-commerce, como apunta el uso que hacen de la RV/RA las grandes empresas del sector, provocaría una nueva brecha tecnológica respecto a las PYMEs. Así pues, esta tesis propone explorar dichas tecnologías en el sector y conectarlas con el pequeño comercio, de manera que puedan poner a disposición en espacios virtuales inmersivos sus productos en aplicaciones ejecutadas en dispositivos de VR comerciales actuales, puedan tomar decisiones informadas en base a los datos de uso de los usuarios de la aplicación, la cual que debe ser accesible para todo tipo de personas.

CAPÍTULO 02

02

Estado del arte y motivación

El comercio electrónico y la brecha digital de las pymes que motiva esta investigación.

2

Estado del arte y motivación

2.1 | Estado del arte

El uso de tecnologías inmersivas en el ámbito del *e-commerce* ha despertado un interés creciente en los últimos años, evidenciando su potencial para transformar la experiencia de compra *online* y mejorar indicadores clave como la intención de compra, el *engagement* del usuario y la percepción del producto [28, 31–34]. No obstante, aspectos como la digitalización de catálogos de productos, los mecanismos de interacción en entornos virtuales, la accesibilidad para usuarios con diversidad funcional y la validación experimental en contextos comerciales reales presentan desafíos que condicionan la adopción efectiva de estas tecnologías. A continuación, se analizan estos aspectos clave, comenzando por el contexto actual del comercio electrónico y los *e-marketplaces*, para posteriormente profundizar en las tecnologías inmersivas, la digitalización 3D, la monitorización del comportamiento del usuario, los mecanismos de interacción y accesibilidad, y finalmente el impacto de la gamificación en el comportamiento de compra.

2.1.1 | *E-commerce*, transformación digital y *e-marketplaces*

Como se ha expuesto en la Sección 1.1, la consolidación del comercio electrónico y las dificultades de las PYMEs para adoptar tecnologías avanzadas configuran un escenario donde resulta crítico buscar soluciones que faciliten su transformación digital. En este contexto, los mercados electrónicos o *e-marketplaces* se han posicionado como una alternativa estratégica que permite a los pequeños comercios acceder al entorno digital sin necesidad de desarrollar infraestructuras tecnológicas propias [35, 36]. Estas plataformas,

definidas como intermediarios que facilitan transacciones comerciales entre múltiples compradores y vendedores en un espacio virtual compartido [35], ofrecen ventajas significativas: acceso inmediato a una amplia base de clientes, delegación de tareas complejas como el posicionamiento web o el mantenimiento del portal, y acceso a herramientas de análisis del comportamiento del consumidor [36, 37].

No obstante, el modelo actual de *e-marketplaces* presenta limitaciones que cuestionan su sostenibilidad y su capacidad para diferenciarse en un mercado cada vez más competitivo. Por un lado, la revisión sistemática de Cano et al. [38] expone una creciente preocupación por el impacto ambiental derivado del embalaje excesivo y el transporte individualizado de pedidos, así como los retos crecientes en la toma de decisiones debido al alto volumen de datos actual. Por otro lado, la experiencia de compra que ofrecen estas plataformas sigue siendo fundamentalmente bidimensional, limitando la capacidad del consumidor para evaluar sensorialmente los productos antes de la compra. Estas carencias han motivado la búsqueda de nuevos modelos que reconcilien el comercio electrónico con principios de sostenibilidad y proximidad, como el concepto de "ciudad de 15 minutos" acuñado por Moreno et al. [39], y que incorporen tecnologías capaces de enriquecer la experiencia del usuario más allá de las interfaces web tradicionales.

En esta línea, las tecnologías inmersivas han emergido como candidatas naturales para impulsar la siguiente evolución del comercio electrónico [23, 33]. Su capacidad para ofrecer experiencias de visualización tridimensional, manipulación de productos y contextualización espacial abre nuevas posibilidades para superar las limitaciones actuales del canal online. La siguiente subsección examina en profundidad el estado actual de estas tecnologías y su aplicación específica en el ámbito comercial.

2.1.2 | Tecnologías inmersivas en *e-commerce*

La VR permite crear entornos tridimensionales simulados en los que los usuarios pueden interactuar de forma inmersiva mediante dispositivos específicos como cascos de visualización, guantes hápticos o sensores de movimiento [40]. Esta capacidad de recrear experiencias sensoriales que replican o incluso amplían la percepción de la realidad física ha propiciado su aplicación en cam-

pos tan diversos como el entretenimiento [41], la publicidad [42], el turismo [43] o la formación en escenarios complejos [44, 45]. En el ámbito específico del comercio, la revisión sistemática de Xi & Hamari [33] presenta cómo empresas como Amazon, Alibaba, eBay o IKEA han comenzado a integrar soluciones de VR en sus servicios con el objetivo de transformar la experiencia de compra, abordando las limitaciones espaciales y temporales del canal *online* tradicional. El interés académico por el *VR-shopping* ha crecido de forma notable en la última década. Los mismos autores identificaron las principales líneas de investigación en este campo, señalando que la VR ofrece oportunidades únicas para mejorar la presentación de productos, facilitar la toma de decisiones del consumidor y crear experiencias de compra más atractivas y personalizadas [33]. Sin embargo, esta misma revisión advierte que la adopción generalizada de estas tecnologías sigue siendo limitada, existiendo barreras tanto tecnológicas como relacionadas con la aceptación por parte del usuario.

Uno de los aspectos más estudiados es la influencia de la inmersión en la percepción y el comportamiento del consumidor. Diversos trabajos han comparado entornos de Immersive Virtual Reality (IVR), donde el usuario utiliza un casco de visualización, con entornos de Desktop Virtual Reality (DVR), donde la interacción se realiza a través de una pantalla convencional. Un estudio con 60 participantes en el contexto de tiendas virtuales de moda llevado a cabo por Ricci et al. [34] demostró que la IVR genera mayor tiempo de *engagement*, mejores valoraciones en dimensiones hedónicas como el disfrute y la telepresencia, y una percepción más positiva de aspectos utilitarios como la diagnosticidad del producto. Estos hallazgos sugieren que el grado de inmersión influye significativamente en la calidad de la experiencia de compra. No obstante, la relación entre inmersión e intención de compra no es necesariamente lineal. Investigaciones previas han evidenciado que, si bien la inmersión incrementa el disfrute, también puede dificultar la lectura de detalles del producto, generando efectos contrapuestos que se anulan mutuamente en términos de intención de reutilización del sistema [46]. Esta tensión entre valores hedónicos y utilitarios constituye un eje central en la investigación sobre *VR-shopping*, donde se ha demostrado que la decisión de adoptar estas tecnologías puede estar motivada tanto por el disfrute de la experiencia como por su utilidad percibida para evaluar productos [31]. Otros factores contextuales también han sido objeto de estudio. La manipulación de variables como la ubicación de

los productos en las estanterías o expositores usados, la cantidad de unidades expuestas o la escala de los objetos virtuales influye en la intención de compra, observándose que la colocación a la altura de los ojos y un ligero aumento de escala mejoran la percepción del producto [47]. En su estudio comparativo, Xi et al. [47] compararon AR y VR en entornos de *retail* controlados, analizando cómo cada modalidad afecta a las percepciones del consumidor y su intención de revisita, proporcionando evidencia sobre las fortalezas y limitaciones de cada aproximación.

La construcción de prototipos experimentales ha sido una estrategia recurrente para explorar diferentes aspectos del *VR-shopping*. Trabajos previos desarrollaron entornos de compra virtual combinando las ventajas del comercio *online* y *offline*, evaluando cómo los métodos de entrada del usuario, como movimientos de cabeza y voz, y los dispositivos de salida influyen en el rendimiento de las tareas y las preferencias del usuario [48]. Posteriormente, estos mismos autores propusieron la metáfora del apartamento como paradigma de diseño para tiendas virtuales, examinando técnicas de selección y manipulación de productos así como diferentes diseños de carrito de compra, e identificando que la inmersión y la experiencia de usuario constituyen las principales preocupaciones de los participantes [49]. Estos estudios también ofrecieron recomendaciones para reducir el mareo por movimiento y determinaron qué tipos de productos resultan más adecuados para su presentación en entornos VR.

A pesar de los avances, la literatura señala limitaciones metodológicas recurrentes. Gran parte de los estudios se han realizado en entornos controlados de laboratorio donde los participantes expresan intenciones de compra o realizan transacciones con dinero virtual, lo que cuestiona la generalizabilidad de los hallazgos a contextos comerciales reales [46, 50]. Además, la investigación sobre los efectos conjuntos de la VR y elementos de gamificación en tiendas reales ha recibido atención limitada [51]. Estas limitaciones evidencian la necesidad de validar experimentalmente las propuestas en escenarios donde las decisiones de compra impliquen consecuencias económicas reales para los participantes. Para que estas experiencias de compra inmersiva puedan concretarse, resulta imprescindible disponer de representaciones digitales de los productos que se mostrarán en el entorno virtual.

2.1.3 | Digitalización de productos en entornos para VR-Shopping

Una de las variables que afecta a la efectividad de los entornos de *VR-shopping* es la calidad y el realismo de los modelos 3D que representan los productos. Diversos estudios han demostrado que la apariencia realista de los objetos virtuales contribuye significativamente a aumentar la intención de compra, permitiendo al usuario abstraerse y experimentar una mayor sensación de estar en una tienda física [50, 52, 53]. Esta relación entre realismo y comportamiento del consumidor subraya la importancia de contar con métodos de digitalización que produzcan modelos de alta fidelidad visual.

Sin embargo, la generación de modelos 3D de calidad ha requerido tradicionalmente equipos de escaneo profesional de alto coste o el trabajo manual de diseñadores especializados, recursos que suelen estar fuera del alcance de las PYMEs [54]. Esta barrera económica y técnica representa un obstáculo significativo para la democratización del *VR-shopping*. Por tanto, resulta necesario explorar alternativas de bajo coste que permitan obtener resultados de calidad aceptable sin requerir conocimientos técnicos avanzados.

En los últimos años han surgido diversas tecnologías que posibilitan la digitalización 3D mediante dispositivos de consumo como teléfonos móviles o tabletas. La fotogrametría, técnica que reconstruye objetos tridimensionales a partir de múltiples fotografías tomadas desde diferentes ángulos, se ha beneficiado de las mejoras en las cámaras de los dispositivos móviles actuales y del procesamiento en la nube, permitiendo a usuarios sin experiencia técnica generar modelos 3D con relativa facilidad [55]. Los sensores LiDAR, incorporados en los últimos años en dispositivos de Apple, ofrecen una alternativa que permite capturar información de profundidad directamente, aunque estudios de evaluación han mostrado que su precisión es de aproximadamente ±1 cm para objetos pequeños, siendo más adecuados para la captura de espacios amplios que para el escaneo detallado de productos [56, 57].

Una aproximación más reciente es la técnica conocida como Neural Radiance Fields (NeRF), que utiliza redes neuronales profundas para representar escenas u objetos a partir de un conjunto de imágenes bidimensionales y sus posiciones de cámara correspondientes [58]. Esta técnica genera representaciones volumétricas que permiten sintetizar vistas del objeto desde cualquier ángulo

con efectos dependientes de la iluminación, logrando resultados de alta calidad fotorrealista [58]. No obstante, la aplicación de NeRF en entornos de VR presenta desafíos relacionados con la latencia y la pérdida de calidad en escenarios de alta resolución y amplio campo de visión, aspectos críticos para garantizar una experiencia de usuario confortable [59]. Investigaciones recientes como la de Deng et al. [37] han propuesto métodos que optimizan simultáneamente la latencia y la calidad visual alineándolas con la percepción humana, logrando reducciones de hasta el 99 % en tiempo de procesamiento respecto al NeRF estándar mientras mantienen una fidelidad visual perceptualmente idéntica a la resolución completa.

Más recientemente ha surgido una técnica alternativa, el 3D Gaussian Splatting (3DGS), que representa la escena de forma explícita mediante un conjunto de gaussianas tridimensionales anisótropas en lugar de la representación implícita y volumétrica de NeRF [60]. Esta diferencia de representación tiene una consecuencia práctica relevante: el 3DGS permite un renderizado en tiempo real mediante rasterización en GPU, lo que lo hace más eficiente que NeRF para la síntesis de vistas a tasas de fotogramas elevadas. Por este motivo ha despertado un interés creciente para aplicaciones de VR, donde el mantenimiento de un ratio de frames alto es un requisito crítico. No obstante, su aplicación directa en headsets autónomos no está exenta de problemas: aparecen artefactos temporales como el *popping* durante el movimiento de la cabeza, distorsiones de proyección y caídas de rendimiento al renderizar un número elevado de gaussianas. Trabajos recientes orientados específicamente a VR abordan estas limitaciones combinando optimizaciones de rasterización con renderizado foveado, logrando una experiencia inmersiva sin artefactos por encima de los 72 FPS en dispositivos como Meta Quest 3 [61].

Así pues, es fundamental que los modelos digitalizados ofrezcan un rendimiento adecuado cuando se ejecutan en dispositivos VR comerciales. Un exceso de detalle geométrico o texturas de alta resolución puede comprometer el rendimiento del sistema, provocando caídas en la tasa de fotogramas que resultan en mareos y malestar para el usuario [62]. Los cascos de VR actuales requieren mantener tasas de refresco entre 60 y 90 fotogramas por segundo para garantizar una experiencia confortable, pudiendo bajar hasta unos 50 en casos excepcionales. Este hecho impone restricciones sobre la complejidad de los modelos que pueden mostrarse simultáneamente. Este equilibrio entre rea-

lismo y rendimiento resulta especialmente crítico en dispositivos autónomos como Meta Quest 2, Quest Pro o Quest 3, cuya capacidad de procesamiento es inferior a la de sistemas conectados a ordenadores de alto rendimiento.

La evaluación objetiva de la calidad de los modelos 3D generados requiere métricas específicas que capturen tanto la fidelidad geométrica como la calidad visual percibida. Existen métricas que requieren un modelo de referencia, como MSDM2, que evalúa la distorsión visual de mallas 3D mediante un enfoque multiescala [63], y métricas sin referencia como las propuestas por Zhang et al. [64, 65]. Estas métricas son NR-3DQA, que extrae características de color y geometría para predecir puntuaciones de calidad mediante regresión, o MM-PCQA, que combina información geométrica y de textura mediante redes neuronales multimodales. La disponibilidad de estas métricas permite realizar comparaciones sistemáticas entre diferentes tecnologías y configuraciones de escaneo.

Disponer de modelos 3D de calidad constituye el primer paso para construir entornos de *VR-shopping*, pero para comprender cómo los usuarios interactúan con estos espacios virtuales y optimizar su diseño, resulta necesario contar con mecanismos de monitorización y análisis del comportamiento.

2.1.4 | Monitorización y análisis de comportamiento de usuario en entornos VR

La capacidad de los entornos de VR para registrar de forma precisa y continua las acciones del usuario constituye una ventaja significativa frente a los canales de comercio electrónico tradicionales. Mientras que las plataformas web convencionales se limitan a capturar información como *clicks*, tiempo de permanencia en página o productos añadidos al carrito, los sistemas de VR permiten obtener datos mucho más ricos sobre el comportamiento del consumidor, incluyendo movimientos oculares, trayectorias de navegación, cinemática de manos y cabeza, e interacciones con los productos virtuales [66, 67]. Esta información abre posibilidades para comprender en profundidad los procesos de decisión de compra y optimizar el diseño de los espacios comerciales virtuales. El seguimiento ocular o *eye-tracking* ha recibido especial atención como fuente de datos comportamentales en entornos de VR-shopping. A pesar de la ausencia de sensores oculares dedicados en ciertos dispositivos de visualización

comerciales como las Meta Quest 2 y 3, investigaciones recientes demuestran que es posible inferir la dirección de la mirada mediante la integración de la postura de la cabeza y la cinemática de los 6-Degrees of Freedom (6-DOF) [68]. Los patrones de fijación visual proporcionan información valiosa sobre qué elementos captan la atención del usuario, durante cuánto tiempo y en qué secuencia, permitiendo identificar áreas de interés y evaluar la efectividad de la disposición de productos [69].

Para que estos datos sean útiles, resulta necesario disponer de algoritmos capaces de distinguir entre fijaciones, donde la mirada permanece relativamente estable sobre un punto de interés, y movimientos sacádicos, donde el ojo se desplaza rápidamente entre puntos. En este sentido, se han desarrollado y calibrado algoritmos específicos para la identificación de fijaciones en entornos de IVR que tienen en cuenta las particularidades de estos sistemas, como el amplio campo de visión y la necesidad de procesar datos en tiempo real [70]. Más allá del seguimiento ocular, la combinación de múltiples fuentes de datos comportamentales ha demostrado ser especialmente efectiva para caracterizar a los usuarios. El estudio llevado a cabo por Moghaddasi et al. [67], el cual contó con 57 participantes, usó un prototipo de supermercado virtual inmersivo, utilizando un HTC Vive Pro con *eye-tracking* integrado, y capturó cuatro flujos de datos no invasivos: movimientos oculares, cinemática de cabeza y manos, y eventos de interacción. Los investigadores extrajeron características zonales, temporales, espaciales y cinemáticas que, procesadas mediante clasificadores Support Vector Machine (SVM), permitieron distinguir entre usuarios con alta y baja impulsividad con una precisión de hasta el 87 %. De forma similar, utilizando el mismo paradigma experimental, otros trabajos han logrado predecir características demográficas como la edad y el género de los compradores a partir de sus patrones de comportamiento, alcanzando precisiones superiores al 70 % y llegando al 94 % para la predicción de género [29].

Estos hallazgos evidencian que los datos comportamentales en VR contienen información suficiente para inferir características del perfil del usuario que pueden ser relevantes para la personalización de la experiencia de compra. La identificación de rasgos de personalidad a partir del comportamiento en entornos virtuales de *retail* también ha sido objeto de investigación. Mediante el análisis de patrones de navegación, interacción visual y manipulación de

productos en una tienda virtual, se ha demostrado la viabilidad de reconocer dimensiones de personalidad de los consumidores, lo que abre vías para la adaptación dinámica de los espacios comerciales en función del perfil psicológico del usuario [28]. Estos estudios comparten un enfoque metodológico común basado en la definición de zonas de interés dentro del espacio virtual, la segmentación temporal de las sesiones y la extracción de métricas agregadas que caracterizan el comportamiento del usuario en cada zona [69]. A pesar del valor demostrado de estos enfoques, se puede encontrar una carencia en los trabajos de la literatura: los estudios que analizan el comportamiento del consumidor en entornos de *VR-shopping* generalmente no proporcionan las herramientas software utilizadas de forma abierta para la recolección y el procesamiento de los datos [28, 29, 67]. Los investigadores tienden a compartir únicamente los resultados y, en algunos casos, los conjuntos de datos, pero no los sistemas que permitieron obtenerlos. Esta situación dificulta la replicabilidad de los estudios y obliga a desarrollar soluciones propias desde cero. Además, si bien se puede generalizar que los prototipos de *VR-shopping* comparten elementos comunes, como p. ej., categorías de productos, secciones en las que agruparlos, expositores con distintas alturas estandarizadas, no existe una herramienta abierta que aglutine los elementos y conocimientos comunes para construir a partir de una base sólida un prototipo de VR-shopping.

Se han propuesto algunos marcos conceptuales para abordar esta carencia. Un trabajo planteó un framework orientado a la identificación de preferencias del consumidor y la personalización de experiencias de VR-shopping a través de datos de comportamiento [71]. Sin embargo, este marco se centró en proporcionar pasos metodológicos para mejorar el diseño de tiendas virtuales sin formalizar un conjunto de datos estructurado ni abordar el análisis de entornos de *V-commerce* completamente inmersivos, limitándose a *showrooms* virtuales. La ausencia de herramientas integradas que faciliten tanto el desarrollo de entornos de *VR-shopping* como la recolección y el análisis de datos de usuario puede representar, por tanto, una barrera en el avance de la investigación en este campo. Adicionalmente, para que los datos recogidos sean interpretables y comparables entre estudios, es necesario establecer métricas estandarizadas que capturen las dimensiones relevantes del comportamiento del usuario. Entre las métricas más utilizadas en la literatura se encuentran el tiempo de permanencia en zonas específicas, el número y la duración de las fijaciones

visuales por área de interés, la frecuencia y el tipo de interacciones con productos, la distancia recorrida, la velocidad de desplazamiento y los patrones de navegación entre secciones [67, 69, 72]. La estandarización de estas métricas facilitaría la comparación de resultados entre diferentes estudios y plataformas, contribuyendo a la consolidación del conocimiento en el área.

Las posibilidades que ofrece la monitorización del comportamiento en VR se extienden también al ámbito del marketing y la evaluación de estrategias comerciales. La capacidad de realizar pruebas A/B en entornos virtuales controlados permite comparar la efectividad de diferentes configuraciones de tienda, disposiciones de producto o campañas promocionales de forma más económica y rápida que en entornos físicos [73, 74]. Asimismo, la segmentación de mercado basada en patrones de comportamiento observados, en lugar de características demográficas declaradas, ofrece oportunidades para diseñar experiencias de compra más personalizadas y efectivas [74]. No obstante, la utilidad de los datos recogidos depende fundamentalmente de que los usuarios puedan interactuar de forma natural y efectiva con el entorno virtual. Los mecanismos de interacción disponibles y su grado de usabilidad condicionan tanto la experiencia del usuario como la validez de los datos comportamentales registrados.

2.1.5 | Mecanismos de interacción y accesibilidad en entornos VR

La forma en que los usuarios interactúan con los objetos y el espacio en entornos de VR condiciona directamente la calidad de la experiencia y, en el contexto del *VR-shopping*, la capacidad del consumidor para evaluar productos y tomar decisiones de compra informadas. A diferencia de las interfaces web tradicionales basadas en ratón y teclado, los sistemas de VR ofrecen modalidades de interacción más naturales e intuitivas, como el seguimiento de manos o los gestos corporales, que permiten manipular objetos virtuales de forma similar a como se haría en el mundo físico [48, 49]. El seguimiento de manos o *hand-tracking* permite a los usuarios interactuar directamente con sus manos, sin necesidad de sostener controladores, lo que ofrece una experiencia más inmersiva y reduce las barreras de entrada para usuarios no familiarizados con dispositivos de VR [75]. Sin embargo, la comparación entre *hand-tracking* y

controladores tradicionales en términos de rendimiento y usabilidad ha arrojado resultados mixtos. Un estudio experimental evaluó ambas modalidades en entornos de aprendizaje virtual para tareas motoras de alcanzar, coger y colocar objetos, midiendo métricas como la duración de las tareas, la frecuencia de interacciones y el número de objetos manipulados, junto con evaluaciones subjetivas del esfuerzo mental mediante la escala RSME [76]. Los resultados indicaron que, si bien el *hand-tracking* ofrece mayor naturalidad, los controladores proporcionan mayor precisión en tareas que requieren manipulación fina. De forma complementaria, otro estudio con 24 participantes investigó la usabilidad del *hand-tracking* para tareas de agarrar y colocar en IVR, empleando un conjunto detallado de métricas que incluían la precisión medida como distancia euclidiana entre el objeto y su ubicación objetivo, el tiempo de agarre, el tiempo de colocación y las caídas accidentales de objetos [77]. Los participantes también evaluaron su experiencia mediante la escala de usabilidad del sistema SUS, la naturalidad percibida de la interacción y su sentido de agencia sobre las manos virtuales. Estos trabajos evidencian la importancia de seleccionar cuidadosamente los mecanismos de interacción en función de los requisitos específicos de cada aplicación.

La elicitación de gestos definidos por el usuario constituye otra línea de investigación relevante para el diseño de interacciones en *VR-shopping*. Wu et al. [78] realizaron dos experimentos con 32 participantes que debían completar 12 tareas diferentes en escenarios de compra virtual, incluyendo la selección de objetos, el cambio de colores y el redimensionamiento de productos. A partir de los gestos propuestos espontáneamente por los usuarios, los investigadores identificaron patrones comunes que podrían servir de base para el diseño de interfaces más intuitivas. No obstante, los autores no detallaron los métodos utilizados para registrar y procesar la información gestual, limitando la replicabilidad del estudio. Más allá de la elección entre controladores y *hand-tracking*, aspectos ergonómicos como la posición de la mano respecto al cuerpo influyen significativamente en la eficiencia y comodidad de la interacción. La investigación de Lou et al. [79], que contó con 24 participantes, demostró que las interfaces tradicionales centradas en la visión no consideran suficientemente la efectividad y el confort de las interacciones manuales en VR. Sin embargo, cuando se considera la diversidad funcional de los usuarios potenciales, las limitaciones de los sistemas de interacción actuales se hacen más evidentes.

Los mecanismos de *hand-tracking* disponibles en los SDK comerciales están diseñados para reconocer poses y gestos de manos con movilidad completa, lo que excluye de facto a personas con discapacidades motoras en miembros superiores [80]. Condiciones como lesiones medulares cervicales, espasticidad, artritis o secuelas de accidentes cerebro-vasculares afectan la capacidad de los usuarios para realizar los gestos específicos que estos sistemas requieren, como el *pinch* o pellizco entre pulgar e índice para agarrar objetos virtuales. Esta carencia resulta especialmente relevante si se considera la magnitud de la población potencialmente afectada. Según datos de la Organización Mundial de la Salud, más de mil millones de personas en el mundo viven con algún tipo de discapacidad, y se estima que entre el 2 % y el 3 % de la población global, aproximadamente 150-225 millones de personas, presenta limitaciones motoras que afectan específicamente a sus manos. Para que las tecnologías de VR ofrezcan un acceso democrático a experiencias inmersivas, resulta relevante abordar las necesidades de accesibilidad de estos colectivos.

La literatura sobre accesibilidad en VR y AR ha crecido en los últimos años, aunque sigue siendo un campo inmaduro. Una revisión exhaustiva de los esfuerzos para mejorar la accesibilidad en VR, AR y el metaverso identificó las principales barreras y oportunidades en este ámbito [80]. Los autores señalan que, si bien existen avances en adaptaciones para usuarios con discapacidades visuales o auditivas, las soluciones para personas con limitaciones motoras en miembros superiores son escasas y frecuentemente se limitan a recomendaciones de diseño sin implementaciones validadas empíricamente. Algunas contribuciones han abordado aspectos específicos de la accesibilidad en entornos virtuales. El trabajo de Dombrowski et al. [81] propuso directrices para el diseño de experiencias de VR inclusivas, enfatizando la importancia de considerar la diversidad funcional desde las etapas iniciales del desarrollo. En el ámbito de la rehabilitación, donde la VR ha demostrado beneficios para la recuperación motora de pacientes con lesiones neurológicas, se han desarrollado plataformas que adaptan los mecanismos de interacción manual. Herrera et al. [82] propusieron un *framework* para juegos serios orientados a la rehabilitación de miembros superiores en pacientes con lesiones medulares o accidentes cerebrovasculares, el cual incorporó adaptaciones de los mecanismos de *hand-tracking* para la manipulación de objetos virtuales. La adaptación se centró principalmente en modificar los gestos de pinza para individuos

incapaces de realizar el gesto completo de tocar el pulgar con el índice. Esta experiencia en el contexto clínico ofrece conocimientos transferibles al diseño de aplicaciones de *VR-shopping* accesibles. La evaluación del esfuerzo y la dificultad experimentados por los usuarios durante la interacción en VR constituye un aspecto metodológico fundamental para el diseño de sistemas accesibles. Estudios en el ámbito de la rehabilitación virtual han empleado instrumentos estandarizados como el NASA-TLX para evaluar la carga de trabajo, la escala SUS para medir la usabilidad percibida, el cuestionario UEQ para valorar la experiencia de usuario y el MSAQ para cuantificar el mareo inducido por el movimiento [83, 84]. La combinación de métricas objetivas derivadas de los datos de interacción con evaluaciones subjetivas mediante cuestionarios validados permite obtener una visión completa de la experiencia del usuario y detectar áreas de mejora.

A pesar de los avances descritos, persiste una brecha significativa entre las capacidades de interacción que ofrecen los sistemas de VR comerciales y las necesidades de los usuarios con discapacidades motoras. Los Software Development Kit (SDK) proporcionados por fabricantes como Meta no incluyen opciones de accesibilidad específicas para personas con espasticidad, atrofia muscular u otras condiciones que afecten la movilidad de manos y dedos. Esta carencia no solo limita el acceso de estos colectivos a las experiencias de *VR-shopping*, sino que también excluye su participación en estudios de investigación que emplean estas tecnologías, introduciendo sesgos en el conocimiento generado sobre el comportamiento del consumidor en entornos virtuales. Una vez establecidos los mecanismos de interacción y garantizada la accesibilidad del sistema, resulta fundamental evaluar el impacto real de estas tecnologías en el comportamiento de compra.

2.1.6 | Impacto de la gamificación y VR en la decisión de compra

La gamificación, entendida como la integración de mecánicas lúdicas propias de los videojuegos en contextos no recreativos, se ha consolidado como una estrategia para mejorar el engagement y la lealtad del consumidor en plataformas de comercio electrónico [85]. Una revisión sistemática siguiendo las directrices de Kitchenham identificó que los elementos más frecuentemen-

te implementados son puntos, insignias y tablas de clasificación, aunque las mecánicas orientadas a la inmersión emergen como una vía prometedora de investigación futura [85]. La evidencia disponible vincula la gamificación con una mayor intención de compra, mayor frecuencia de compra repetida y una percepción de utilidad más favorable [86]. En el contexto específico de la VR, varios trabajos han explorado cómo la combinación de inmersión y elementos de juego puede potenciar la experiencia de compra. Chen et al. [87] investigaron si la interactividad y la vividez de entornos VR impulsan la compra impulsiva, aplicando el marco Stimulus-Organism-Response en un experimento con 227 participantes equipados con cascos Oculus Gear VR. Sus resultados demostraron que ambas variables incrementan significativamente la telepresencia, la diagnosticidad percibida y la sensación de entretenimiento, factores que a su vez predicen una mayor urgencia de compra impulsiva. De forma complementaria, Wei et al. [88] abordaron el problema de la fatiga del usuario en experiencias VR prolongadas, demostrando que la incorporación de elementos gamificados como desafíos, retroalimentación y narrativas reduce significativamente la fatiga motivacional y cognitiva. Faganel et al. [51] desarrollaron "Fashion Island", un probador virtual gamificado mediante HTC Vive, reportando que más del 90 % de los participantes expresaron disposición a reutilizar el sistema para futuras experiencias de compra.

No obstante, la literatura presenta una limitación metodológica crítica que compromete la validez externa de muchos hallazgos: la mayoría de estudios sobre *VR-shopping* emplean presupuestos virtuales o miden únicamente intenciones de compra declaradas [29, 46, 49, 50]. Esta aproximación, si bien facilita el control experimental, no captura la complejidad psicológica de las decisiones de compra cuando existe un coste económico real para el participante. Además, el efecto conjunto de VR y gamificación en contextos comerciales reales, donde las decisiones de compra están influenciadas por factores emocionales, sociales y situacionales, ha recibido atención limitada [51]. La distinción entre perfiles de consumidor hedónico, motivado por el entretenimiento y los aspectos multisensoriales de la experiencia, y utilitario, orientado a la eficiencia y la consecución de objetivos [89], resulta relevante para comprender cómo diferentes usuarios responden a entornos VR gamificados y cómo estas respuestas se traducen en comportamientos de compra reales.

2.2 | Motivación

La transformación del comercio electrónico hacia experiencias inmersivas representa un cambio de paradigma que, como se ha expuesto en las secciones anteriores, ofrece oportunidades significativas pero también plantea desafíos técnicos, económicos y de diseño que condicionan su adopción efectiva. Si bien las grandes corporaciones tecnológicas ya exploran activamente estas tecnologías, las pequeñas y medianas empresas permanecen en gran medida al margen de esta evolución, perpetuando una brecha digital que amenaza su competitividad a largo plazo. Esta tesis doctoral nace precisamente de la necesidad de abordar esta problemática desde una perspectiva integral, reconociendo que la implementación exitosa de entornos VR-shopping requiere soluciones coordinadas que abarquen desde la arquitectura de plataforma hasta la validación empírica del impacto comercial. A lo largo de las anteriores subsecciones, se han presentado distintos puntos que deben abordarse, fundamentándose en los trabajos previos existentes en la literatura:

- Los *e-marketplaces* actuales, a pesar de su sofisticación, no han sido concebidos para integrar de forma holística las tecnologías inmersivas con las necesidades específicas de los pequeños comercios y las economías locales. Tampoco incorporan criterios de sostenibilidad ni conceptos emergentes como la ciudad de 15 minutos, que podrían reconciliar las ventajas del comercio digital con el fortalecimiento del tejido comercial de proximidad.

- Disponer de una arquitectura adecuada resulta insuficiente si los comerciantes no pueden poblar estos entornos virtuales con representaciones digitales de sus productos. El análisis de la literatura reveló que, a pesar del potencial demostrado del realismo 3D para incrementar la intención de compra, no existían directrices claras sobre cómo las PYMEs pueden digitalizar sus catálogos de forma asequible y con resultados de calidad suficiente para dispositivos VR comerciales. Las tecnologías de escaneo profesional exceden claramente las capacidades económicas y técnicas de estos negocios, mientras que las alternativas de bajo coste basadas en dispositivos móviles no habían sido evaluadas de forma sistemática

considerando simultáneamente la calidad visual y el rendimiento en tiempo real

- La ausencia de herramientas software que facilitasen tanto el desarrollo de aplicaciones *VR-shopping* como la recolección y análisis de datos del comportamiento del usuario. Los trabajos revisados en la literatura raramente comparten las herramientas empleadas, lo que dificulta la replicabilidad de los estudios y obliga a cada equipo de investigación a desarrollar soluciones propias desde cero. Esta fragmentación no solo ralentiza el avance del conocimiento, sino que también eleva las barreras de entrada para investigadores de disciplinas como el marketing o la psicología del consumidor, que podrían aportar perspectivas valiosas pero carecen de la formación técnica necesaria para implementar entornos virtuales.

- Las tecnologías de interacción incluidas en las gafas de VR comerciales actuales, basadas en *hand-tracking*, cada vez más prevalentes en dispositivos VR comerciales, han sido diseñadas asumiendo usuarios con movilidad completa en sus manos. Esta asunción excluye de facto a un porcentaje significativo de la población mundial que presenta limitaciones motoras derivadas de lesiones medulares, espasticidad, artritis u otras condiciones. Si las tecnologías inmersivas van a representar el futuro del comercio electrónico, resulta imperativo que este futuro sea accesible para todos.

- Finalmente, la revisión de la literatura evidenció una limitación metodológica transversal a gran parte de los estudios sobre VR-shopping: el uso predominante de presupuestos virtuales e intenciones de compra declaradas como variables dependientes. Si bien estas aproximaciones facilitan el control experimental, no capturan la complejidad psicológica de las decisiones de compra cuando existe un coste económico real

CAPÍTULO 03

03

Hipótesis y objetivos

La conjetura que vertebra la tesis y las metas que la hacen verificable.

3

Hipótesis y objetivos

En este capítulo se presentarán tanto la hipótesis que motivó esta tesis doctoral como los objetivos que permitirán dar respuestas a las preguntas de investigación formuladas. Dichas preguntas de investigación se verán respondidas en los artículos que forman el compendio.

La hipótesis sobre la que se cimienta esta tesis es que **la implementación de sistemas de VR basados en la digitalización de bajo coste y en la interacción gestual inclusiva es técnicamente viable y mejora significativamente la experiencia de compra mediante una usabilidad natural cercana al entorno físico, generando un impacto positivo y cuantificable en la intención de compra en entornos de venta directa, además de habilitar la recolección de datos comportamentales más ricos que en entornos web, lo que permite validar un modelo de venta inmersiva accesible, con potencial aplicación en el pequeño comercio**.

3.1 | Preguntas de investigación

Esta tesis busca responder un conjunto de preguntas sobre la viabilidad y potenciales beneficios del uso de tecnologías inmersivas en el campo del *e-commerce*, con un enfoque particular en las PYMEs.

- **Pregunta de investigación 1 (RQ1)**: ¿Cómo deben evolucionar los *e-marketplaces* actuales para integrar de forma holística tecnologías inmersivas y de IA, atendiendo a las necesidades específicas de las PYMEs y las economías locales?
- **Pregunta de investigación 2 (RQ2)**: ¿Pueden producirse modelos 3D con calidad visual y rendimiento suficientes para su integración en entornos

de *VR-shopping* ejecutados en dispositivos comerciales a bajo coste?

- **Pregunta de investigación 3 (RQ3)**: ¿Cómo debe ser la interacción y navegación en entornos VR-shopping haciendo uso de *hand-tracking* para proporcionar una experiencia cercana a la compra en entornos físicos?

- **Pregunta de investigación 4 (RQ4)**: ¿Pueden este tipo de entornos inmersivos, con interacciones basadas en *hand-tracking*, ser accesibles a todo tipo de personas, especialmente a aquellas con limitaciones motoras en miembros superiores?

- **Pregunta de investigación 5 (RQ5)**: ¿Puede el comportamiento del usuario en este tipo de entornos comerciales ser recopilado de manera no invasiva y analizado para optimizar la experiencia de compra, incrementar las ventas, personalizar la oferta y tomar decisiones estratégicas basadas en datos reales?

- **Pregunta de investigación 6 (RQ6)**: ¿Influye positivamente la VR, combinada con estrategias de gamificación, en el comportamiento de compra real del consumidor?

3.2 | Objetivos

3.2.1 | Objetivos principales

El objetivo principal de esta investigación consiste en diseñar, desarrollar y validar experimentalmente un modelo arquitectónico y un conjunto de herramientas que permitan implementar entornos inmersivos de compra (VR-Shopping) de forma efectiva, accesible y de bajo coste. Para ello, se propone una aproximación integral que abarca desde la definición de la arquitectura del sistema y la digitalización del catálogo de productos, hasta la creación de herramientas software para el desarrollo y análisis de estos entornos, la validación experimental del impacto de la VR y la gamificación en el comportamiento de compra real del consumidor, y la incorporación de mecanismos de accesibilidad que garanticen la inclusión de personas con diversidad funcional motora en miembros superiores.

3.2.2 | Objetivos específicos

A partir de las preguntas de investigación formuladas en la Sección 3.1 y del objetivo general planteado, se establecen seis objetivos parciales que guían el desarrollo de esta tesis. Estos objetivos abarcan desde el análisis del estado del arte y la propuesta de un modelo arquitectónico, hasta la validación experimental en escenarios comerciales reales. Cada objetivo parcial se descompone en sub-objetivos verificables cuya resolución contribuye a dar respuesta a la pregunta de investigación correspondiente.

- **Objetivo 1 (O1)**. Analizar el estado del arte de los *e-marketplaces* y proponer un modelo arquitectónico para *e-marketplaces* inmersivos orientados a PYMEs y economías locales.

    1. **Subobjetivo 1 (SO1.1)**. Analizar y sistematizar las limitaciones funcionales de los *e-marketplaces* actuales en cuanto a experiencia de usuario, integración de tecnologías inmersivas, atención a las necesidades de las PYMEs e impacto sobre las economías locales.
    2. **Subobjetivo 2 (SO1.2)**. Diseñar un modelo arquitectónico multicapa para un *e-marketplace* inmersivo (VR-ZOCO).

- **Objetivo 2 (O2)**. Estudiar y evaluar tecnologías actuales de escaneo 3D de bajo coste para su uso en entornos de VR-shopping ejecutados en dispositivos comerciales.

    1. **Subobjetivo 1 (SO2.1)**. Seleccionar y caracterizar las tecnologías de escaneo 3D de bajo coste disponibles en dispositivos móviles comerciales, incluyendo fotogrametría, sensores LiDAR y Neural Radiance Fields, analizando sus características técnicas, requisitos de hardware y accesibilidad para usuarios sin formación técnica.
    2. **Subobjetivo 2 (SO2.2)**. Evaluar la calidad visual de los modelos 3D obtenidos con cada tecnología sobre un conjunto de productos con diferentes materiales y geometrías, empleando métricas objetivas con y sin referencia, así como midiendo el error geométrico inherente de cada método de escaneo mediante la comparación con modelos de referencia modelados a partir de las medidas reales.

3. **Subobjetivo 3 (SO2.3)**. Analizar el rendimiento de los modelos 3D en dispositivos VR autónomos comerciales (Meta Quest 2, Quest Pro, Quest 3) en *modo standalone*, usando métricas para determinar la escalabilidad de cada tipo de modelo en escenas con volúmenes crecientes de productos.
4. **Subobjetivo 4 (SO2.4)**. Definir y validar un proceso de optimización que permita reducir el número de polígonos y el tamaño de las texturas de los modelos de alta resolución, manteniendo una calidad visual aceptable y siendo viables para su despliegue masivo en entornos de VR-shopping.

- **Objetivo 3 (O3)**. Diseñar e implementar los mecanismos de interacción y navegación basados en *hand-tracking* para entornos de VR-shopping.

  1. **Subobjetivo 1 (SO3.1)**. Diseñar e implementar mecanismos de interacción con productos mediante *hand-tracking* que permitan al usuario agarrar, examinar, escalar, rotar y soltar productos virtuales de forma natural, así como añadirlos a un carrito de compra, proporcionando retroalimentación visual que guíe al usuario durante cada acción.
  2. **Subobjetivo 2 (SO3.2)**. Implementar mecanismos de navegación adaptados a espacios comerciales virtuales cerrados, incluyendo teletransportación mediante gesto de mano y rotación del cuerpo virtual, ofreciendo alternativas al desplazamiento continuo.
  3. **Subobjetivo 3 (SO3.3)**. Diseñar la estructura espacial del entorno de compra virtual, incluyendo la disposición de estanterías, secciones de productos, zonas de interacción y áreas de interés, de forma que sea configurable y reutilizable para diferentes configuraciones.

- **Objetivo 4 (O4)**. Adaptar las interacciones basadas en *hand-tracking* para personas con limitaciones motoras en miembros superiores y validar su efectividad con personas que sufren dichas limitaciones.

  1. **Subobjetivo 1 (SO4.1)**. Analizar, en colaboración con el personal clínico, las limitaciones motoras específicas de los pacientes con lesión medular que impiden el uso de las interacciones estándar de

hand-tracking proporcionadas por los SDK de facto, identificando qué gestos y posturas resultan inviables para este colectivo.

2. **Subobjetivo 2 (SO4.2)**. Diseñar e implementar un conjunto de interacciones adaptadas que incluya tanto interacciones realizadas con las manos como gestos o posturas con las mismas que requieren motricidad fina para ser realizadas. Estas deben tener mecanismos de retroalimentación visual y ser lo más genéricas posibles para cubrir un amplio espectro de grados de lesión.
3. **Subobjetivo 3 (SO4.3)**. Definir y calcular métricas objetivas derivadas del *hand-tracking*, que permitan establecer cuantitativamente la diferencia percibida en dificultad y esfuerzo realizado entre interacciones adaptadas y no adaptadas.
4. **Subobjetivo 4 (SO4.4)**. Ejecutar un caso de estudio con pacientes reales con lesión medular y sujetos de control sanos en el Hospital Nacional de Parapléjicos de Toledo (HNPT), donde los participantes completen un conjunto de tareas representativas de un entorno de VR-shopping utilizando primero interacciones no adaptadas y después interacciones adaptadas.

- **Objetivo 5 (O5)**. Desarrollar un *framework* abierto que facilite la creación de entornos de VR-shopping y la recolección y análisis estandarizado de datos de comportamiento del usuario.

  1. **Subobjetivo 1 (SO5.1)**. Diseñar e implementar un componente de configuración del entorno de compra virtual que proporcione una capa de abstracción sobre las tecnologías VR subyacentes, ofreciendo elementos reutilizables propios de entornos de compra y facilitando la configuración de *hand-tracking*, *eye-tracking* y mecanismos de teletransportación.
  2. **Subobjetivo 2 (SO5.2)**. Desarrollar un componente de recolección de datos que registre de forma no invasiva la actividad del usuario durante la experiencia de compra, incluyendo datos de eye-tracking, cinemática de cabeza y manos, datos posicionales del usuario en el espacio 3D, eventos de interacción con productos y eventos de navegación.

3. **Subobjetivo 3 (SO5.3)**. Implementar un componente de estadísticas de comportamiento que, a partir de los datos en crudo recolectados, genere métricas espacio-temporales y cinemáticas estandarizadas.

- **Objetivo 6 (O6)**. Validar experimentalmente el impacto de la VR y la gamificación en el comportamiento de compra real del consumidor mediante un caso de estudio en un entorno comercial real.

  1. **Subobjetivo 1 (SO6.1)**. Diseñar y desarrollar una aplicación de VR-shopping con dos escenarios experimentales diferenciados: (i) un entorno de tienda virtual con productos reales digitalizados, donde el usuario puede explorar y manipular productos; y (ii) un entorno gamificado basado en la narrativa de una novela gráfica.
  2. **Subobjetivo 2 (SO6.2)**. Ejecutar un caso de estudio con participantes reales en el contexto de un evento comercial abierto al público, donde los participantes puedan realizar compras con su propio dinero, recogiendo datos comportamentales objetivos durante la sesión y datos actitudinales subjetivos mediante cuestionarios post-sesión.
  3. **Subobjetivo 3 (SO6.3)**. Analizar los datos comportamentales recolectados para identificar perfiles de usuario diferenciados, analizar sus diferencias en métricas clave de interacción, navegación y eficiencia, y determinar si existe una asociación estadísticamente significativa entre los perfiles comportamentales y la decisión de compra real.
  4. **Subobjetivo 4 (SO6.4)**. Realizar un análisis predictivo exploratorio para evaluar la viabilidad de predecir la decisión de compra real a partir de los datos comportamentales recogidos en el entorno inmersivo.

La Figura 3.1 muestra de forma esquemática la relación existente entre los distintos subobjetivos planteados y las preguntas de investigación, y cómo estas conforman la presente tesis. Como se puede observar, RQ1 sirve como los cimientos sobre los que construir las bases que suponen RQ2, RQ3 y RQ4, que permiten el desarrollo de los entornos inmersivos de VR-shopping accesibles y

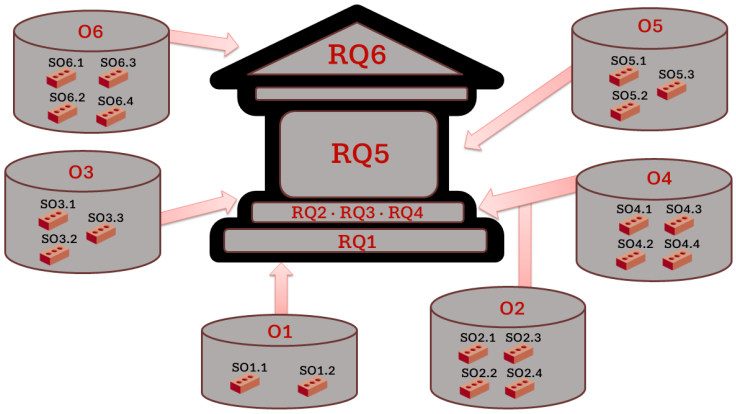


Figura 3.1: Relación entre subobjetivos y preguntas de investigación de la presente tesis.

ofrecen todo el soporte tecnológico para poder validar la hipótesis establecida. Sobre esta base se puede construir la respuesta que se dará a RQ5, que es clave para poder responder RQ6, en el techo de todo, como caso de estudio que permita explorar la influencia de la VR en la decisión de compra.

CAPÍTULO 04

04

Plan de trabajo

Metodología, fases y recursos para llevar la hipótesis a la práctica.

4

Plan de trabajo

En esta sección se va a presentar la división en tareas realizadas de forma cronológica para poder completar los distintos objetivos establecidos en la Sección 3.2. Como ya se mencionó en la Sección 1.1, esta tesis se desarrolló en el contexto del proyecto nacional **VR-ZOCO: Diseño y desarrollo de una plataforma basada en VR-Shopping e IA para la digitalización y el fortalecimiento de negocios y economías locales**, financiado por el MICIU y la AEI, que abarcó desde el 1/12/2022 hasta el 30/04/2025. Sin embargo, la investigación de esta tesis comenzó en febrero del 2023, desarrollándose durante la vigencia del proyecto nacional y tras su finalización, terminando en junio del 2026.

El desarrollo de esta tesis doctoral se ha estructurado en torno al objetivo general formulado en la Sección 3.2.1, que se descompone en seis objetivos parciales (O1–O6) orientados a dar respuesta a las preguntas de investigación y a la hipótesis planteadas en el Capítulo 3. Este desglose se traduce en un conjunto de tareas prácticas, organizadas en cuatro fases, que han permitido alcanzar los hitos científicos y técnicos establecidos.

La Figura 4.1 muestra la planificación temporal de las tareas realizadas. Estas se organizan en cuatro fases diferenciadas: elaboración de la propuesta, desarrollo y validación experimental, plan de movilidad internacional y diseminación de resultados. A continuación, se presenta el desglose de cada tarea, incluyendo su relación con los objetivos parciales y las publicaciones del compendio.

- **T0. Planificación y organización:** fase inicial de definición del marco de trabajo, que incluyó la formulación del problema de investigación, el estudio preliminar de las tecnologías disponibles (dispositivos VR, frameworks de desarrollo, tecnologías de escaneo 3D), la identificación

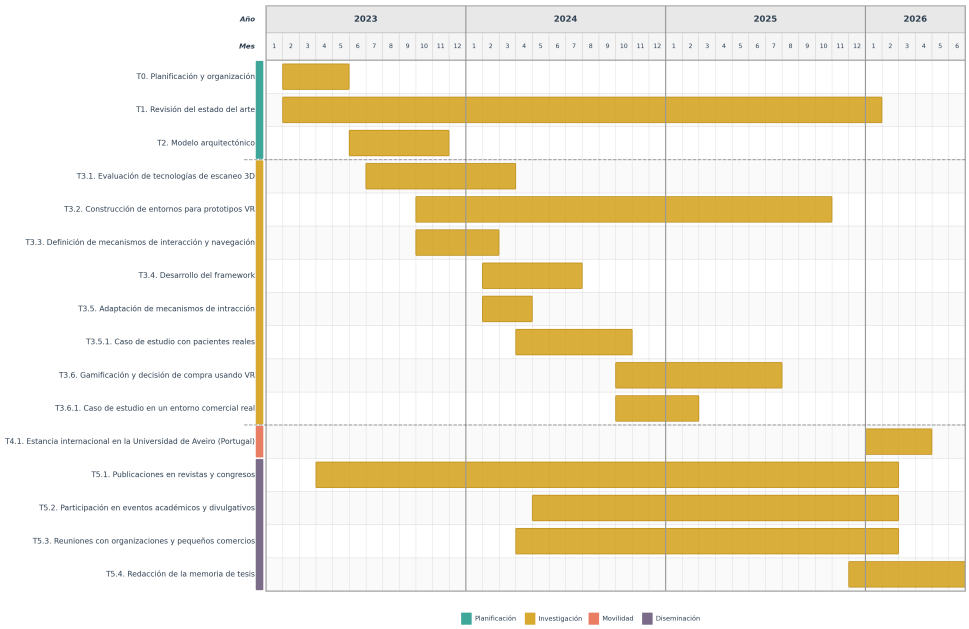


Figura 4.1: Diagrama de Gantt con la planificación temporal de las tareas realizadas durante el desarrollo de la tesis doctoral.

de los actores clave (PYMEs, consumidores, personas con diversidad funcional) y la planificación temporal del proyecto (febrero 2023 – mayo 2023).

- **T1. Revisión del estado del arte:** análisis continuo y sistemático de la literatura científica en los dominios que vertebran esta tesis: evolución de los *e-marketplaces* en el contexto de las PYMEs, tecnologías de escaneo 3D de bajo coste, entornos de *VR-shopping*, gamificación en comercio electrónico, comportamiento del consumidor en entornos virtuales, y accesibilidad en VR para personas con limitaciones motoras. Dada la evolución de estas áreas, esta tarea se ha mantenido activa durante prácticamente toda la duración de la tesis, incorporando avances como la aparición de Meta Quest 3, las mejoras en *hand-tracking*, los grandes modelos de lenguaje o los avances en Neural Radiance Fields (NeRF) y Gaussian Splatting (febrero 2023 – enero 2026). Relacionada con SO1.1, si bien alimenta transversalmente a todos los objetivos.

- **T2. Modelo arquitectónico:** diseño de la propuesta arquitectónica multicapa para VR-ZOCO, el *e-marketplace* inmersivo. Incluyó la definición de la capa de interacción del usuario, la capa de razonamiento basada en IA, la capa de persistencia y el módulo de digitalización de productos, así como los mecanismos de sostenibilidad y economía de proximidad. El trabajo se concretó en publicaciones fuera del compendio (junio 2023 – noviembre 2023) que se presentan en la Sección 5.2. Relacionada con O1 (SO1.1, SO1.2).

- **T3.1. Evaluación de tecnologías de escaneo 3D:** selección, caracterización y evaluación comparativa de tecnologías de escaneo 3D de bajo coste disponibles en dispositivos móviles comerciales: fotogrametría y LiDAR mediante Polycam, y NeRF mediante Luma AI. La evaluación abarcó la calidad visual de los modelos generados métricas con y sin referencia, su rendimiento en dispositivos VR autónomos (Meta Quest 2, Quest Pro, Quest 3), y la definición de un proceso de optimización para despliegue masivo. Los resultados constituyen el **Artículo I** [1] del compendio (julio 2023 – marzo 2024). Relacionada con O2 (SO2.1, SO2.2, SO2.3, SO2.4).

- **T3.2. Construcción de entornos para prototipos VR:** tarea transversal de desarrollo continuo que engloba la construcción progresiva de los entornos virtuales de compra utilizados como soporte experimental. Incluye el diseño de los espacios de *retail* virtual (disposición de estanterías, zonas de interacción, áreas de interés), la integración de los modelos 3D en Unity, y la iteración sobre el diseño en función de los resultados de cada fase experimental (octubre 2023 – octubre 2025). Relacionada con O3 (SO3.3) y, de forma instrumental, con O4, O5 y O6.

- **T3.3. Definición de mecanismos de interacción y navegación:** diseño e implementación de los mecanismos de interacción con productos y navegación en el entorno de *VR-shopping* mediante *hand-tracking*: agarre, examen y rotación de productos, teletransportación, rotación del cuerpo virtual y retroalimentación visual. Estos mecanismos constituyen la base interactiva sobre la que se construyeron las adaptaciones de accesibilidad (T3.6) y la aplicación de gamificación (T3.7) (octubre 2023 – febrero 2024). Relacionada con O3 (SO3.1, SO3.2).

- **T3.4. Desarrollo del *framework* VRSI:** diseño e implementación de Virtual Reality Shopping-Insights (VRSI), un *framework* de código abierto publicado bajo licencia MIT que proporciona: (i) un componente de configuración del entorno de *retail* virtual; (ii) un componente de monitorización y recolección no invasiva de datos de *eye-tracking*, cinemática, posición y eventos de interacción; y (iii) un componente de estadísticas de comportamiento que genera métricas estandarizadas, mapas de calor volumétricos y visualizaciones. Como resultado se obtuvo el **Artículo II** [2] del compendio (febrero 2024 – julio 2024). Relacionada con O5 (SO5.1, SO5.2, SO5.3).

- **T3.5. Adaptación de mecanismos de interacción:** diseño e implementación de interacciones adaptadas para personas con limitaciones motoras en miembros superiores, en colaboración con el personal clínico del HNPT. Incluyó el análisis de las limitaciones motoras específicas de los pacientes con lesión medular (espasticidad, atrofia muscular, mano en garra), las adaptaciones de los gestos de pinza, teletransportación y rotación, la retroalimentación visual compensatoria y la definición de las métricas DI (*Difficulty Index*) y ED (*Effort Degree*) (febrero 2024 – abril 2024). Relacionada con O4 (SO4.1, SO4.2, SO4.3).

- **T3.5.1. Caso de estudio con pacientes reales:** ejecución del caso de estudio en el HNPT con pacientes con lesión medular y sujetos de control sanos, donde los participantes completaron tareas de *VR-shopping* con interacciones no adaptadas y después adaptadas. Se evaluaron tasas de completitud, tiempos de ejecución, métricas DI y ED, y cuestionarios estandarizados (NASA-TLX, SUS, UEQ, MSAQ). La ventana temporal estuvo condicionada por la disponibilidad de pacientes y la coordinación con el equipo clínico. Los resultados constituyen el **Artículo III** [3] del compendio (abril 2024 – octubre 2024). Relacionada con O4 (SO4.4).

- **T3.6. Gamificación y decisión de compra usando VR:** diseño y desarrollo de una aplicación de *VR-shopping* con dos escenarios experimentales: un entorno de tienda virtual con productos digitalizados mediante NeRF, y un entorno gamificado basado en una narrativa de novela gráfica con mecánicas de puzzle. Abarcó el diseño de los escenarios, la integración

de productos digitalizados, la implementación de las mecánicas de gamificación y la preparación de los instrumentos de evaluación (octubre 2024 – julio 2025). Relacionada con O6 (SO6.1).

- **T3.6.1. Caso de estudio en un entorno comercial real:** ejecución del caso de estudio con 48 participantes durante un evento comercial público, donde los usuarios podían realizar compras con su propio dinero. Se recogieron datos comportamentales objetivos mediante VRSI y datos actitudinales subjetivos mediante cuestionarios. El análisis incluyó técnicas de *clustering* no supervisado para identificar perfiles comportamentales y un análisis predictivo exploratorio mediante aprendizaje supervisado. Los resultados constituyen el **Artículo IV** [4] del compendio (octubre 2024 – febrero 2025). Relacionada con O6 (SO6.2, SO6.3, SO6.4).

- **T4.1. Estancia internacional en la Universidad de Aveiro (Portugal):** estancia de investigación predoctoral en la Universidade de Aveiro, dentro del grupo de investigación VARLab, que cuenta con expertos en tecnologías inmersivas e Interacción Persona-Ordenador. Esta estancia permite la colaboración con investigadores internacionales y el contraste de resultados y metodologías, contribuyendo a la obtención de la mención internacional en el título de Doctor (enero 2026 – abril 2026).

- **T5.1. Publicaciones en revistas y congresos:** redacción, envío, revisión y publicación de los artículos científicos que conforman el compendio de publicaciones de esta tesis, así como de contribuciones adicionales a congresos internacionales. Esta tarea acompaña de forma paralela a las fases de desarrollo y validación (abril 2023 – febrero 2026).

- **T5.2. Participación en eventos académicos y divulgativos:** asistencia y presentación de resultados en conferencias nacionales e internacionales, jornadas de investigación del programa de doctorado, seminarios y talleres de divulgación científica (mayo 2024 – febrero 2026).

- **T5.3. Reuniones con organizaciones y pequeños comercios:** colaboración con comercios locales y organizaciones empresariales para contrastar las necesidades reales de las PYMEs con las soluciones diseñadas, identificar productos representativos para la digitalización, y facilitar la

organización del caso de estudio comercial (T3.7.1) (abril 2024 – febrero 2026). Relacionada con O1 y O6.

- **T5.4. Redacción de la memoria de tesis:** elaboración del documento final de la tesis doctoral, incluyendo los capítulos de introducción, estado del arte, hipótesis y objetivos, plan de trabajo, resumen de las contribuciones del compendio, discusión general, conclusiones y líneas de trabajo futuro (diciembre 2025 – junio 2026).

CAPÍTULO 05

05

Resultados

Sobre digitalización, mecanismos de interacción, usabilidad, registro y medición de comportamientos, e impacto real en la compra.

5

Resultados

Esta tesis doctoral se presenta como compendio de publicaciones científicas. Así pues, en el presente capítulo se exponen los artículos científicos que forman el compendio, demostrando su validez investigadora al estar indexados en la base de datos Journal Citation Reports (JCR). El compendio está formado por cuatro artículos. De forma adicional, se incluyen otras investigaciones realizadas durante el desarrollo de la tesis, entre las que se encuentran artículos publicados en revistas científicas que no forman parte del compendio [AR], Congresos Nacionales [CN], Congresos Internacionales [CI] y capítulos de libro [CL]. Para facilitar la visualización del volumen de trabajos publicados y su ubicación en la línea temporal, se ha elaborado un gráfico (ver Fig. 5.1).

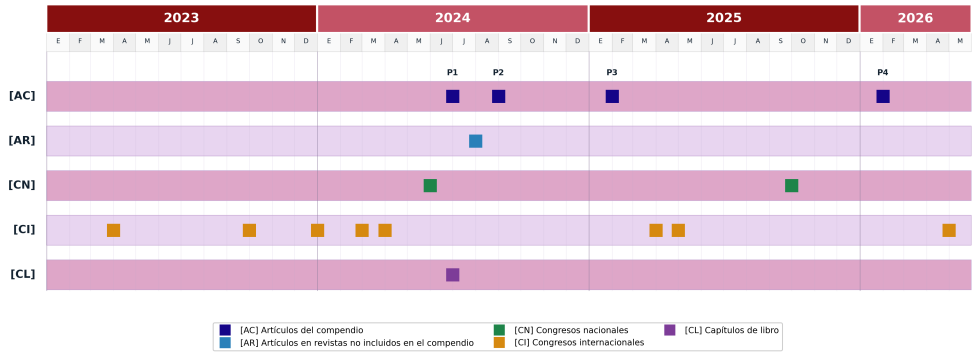


Figura 5.1: Publicaciones realizadas durante la realización de esta tesis doctoral, clasificadas en: [AC] artículos del compendio, [AR] artículos en revistas científicas no incluidos en el compendio, [CN] congresos nacionales, [CI] congresos internacionales y [CL] capítulos de libro.

5.1 | Artículos publicados que forman parte del compendio

Los cuatro artículos científicos que forman el compendio de esta tesis doctoral están publicados en revistas indexadas en el Journal Citation Reports (JCR) según la base de datos de Clarivate, abarcando el primer cuartil (Q1) y el segundo cuartil (Q2). A continuación, en la Tabla 5.1 se muestran los datos más relevantes de las publicaciones, incluyendo la referencia bibliográfica del artículo, revista, editorial, año de publicación, Índice de Impacto (IF) y cuartil. Cada una de las publicaciones incorpora un repositorio público asociado, con el objetivo de facilitar la reutilización y validación de los resultados presentados.

Tabla 5.1: Tabla resumen de los artículos científicos incluidos en el compendio de la tesis doctoral.

Paper	Revista	Editorial	Año	IF	Cuartil	Referencia
I	Applied Sciences	MDPI	2024	2.5	Q2	[1]
II	SoftwareX	Elsevier	2024	2.4	Q2	[2]
III	IEEE Access	IEEE	2025	3.6	Q2	[3]
IV	Virtual Reality	Springer	2026	5	Q1	[4]

Artículo I: Performance Evaluation and Optimization of 3D Models from Low-Cost 3D Scanning Technologies for Virtual Reality and Metaverse E-Commerce

- **Autores:** R. Grande, J. Albusac, D. Vallejo, C. Glez-Morcillo, J.J. Castro-Schez.
- **Tipo:** artículo de revista.
- **Revista:** Applied Sciences.
- **Editorial:** MDPI.
- **ISSN:** 2076-3417.
- **Año:** 2024.

- **DOI:** `https://doi.org/10.3390/app14146037`
- **Categoría:** Engineering, Multidisciplinary.
- **Factor de Impacto (2024):** 2.5.
- **Factor de Impacto a 5 años (2024):** 2.7.
- **Ranking JCR:** Q2.
- **Número de citas (junio 2026):** 34.
- **Objetivos relacionados:** O2 (SO2.1, SO2.2, SO2.3, SO2.4).
- **Repositorio:** `https://github.com/AIR-Research-Group-UCLM/PDIVR-ZOCO`

applied sciences

MDPI

Article

Performance Evaluation and Optimization of 3D Models from Low-Cost 3D Scanning Technologies for Virtual Reality and Metaverse E-Commerce

Rubén Grande, Javier Albusac *, David Vallejo, Carlos Glez-Morcillo and José Jesús Castro-Schez

Department of Information Technologies and Systems, University of Castilla-La Mancha, Paseo de la Universidad 4, 13071 Ciudad Real, Spain; ruben.grande@uclm.es (R.G.); david.vallejo@uclm.es (D.V.); carlos.gonzalez@uclm.es (C.G.-M.); josejesus.castro@uclm.es (J.J.C.-S.)
* Correspondence: javieralonso.albusac@uclm.es

Abstract: Virtual Reality (VR) is and will be a key driver in the evolution of e-commerce, providing an immersive and gamified shopping experience. However, for VR shopping spaces to become a reality, retailers' product catalogues must first be digitised into 3D models. While this may be a simple task for retail giants, it can be a major obstacle for small retailers, whose human and financial resources are often more limited, making them less competitive. Therefore, this paper presents an analysis of low-cost scanning technologies for small business owners to digitise their products and make them available on VR shopping platforms, with the aim of helping improve the competitiveness of small businesses through VR and Artificial Intelligence (AI). The technologies to be considered are photogrammetry, LiDAR sensors and NeRF.In addition to investigating which technology provides the best visual quality of 3D models based on metrics and quantitative results, these models must also offer good performance in commercial VR headsets. In this way, we also analyse the performance of such models when running on Meta Quest 2, Quest Pro and Quest 3 headsets (Reality Labs, Reality Labs, CA, USA) to determine their feasibility and provide use cases for each type of model from a scalability point of view. Finally, our work describes a model optimisation process that reduce the polygon count and texture size of high-poly models, converting them into more performance-friendly versions without significantly compromising visual quality.

Keywords: virtual reality; e-commerce; VR shopping; product digitisation; virtual reality performance; digital retail transformation

check for updates

Citation: Grande, R.; Albusac, J.; Vallejo, D.; Glez-Morcillo, C.; Castro-Schez, J.J. Performance Evaluation and Optimization of 3D Models from Low-Cost 3D Scanning Technologies for Virtual Reality and Metaverse E-Commerce. *Appl. Sci.* **2024**, *14*, 6037. https://doi.org/10.3390/app14146037

Academic Editor: João M. F. Rodrigues

Received: 4 June 2024
Revised: 5 July 2024
Accepted: 8 July 2024
Published: 10 July 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Emerging technologies, such as Virtual Reality (VR) and Artificial Intelligence (AI), are contributing to the further evolution of e-commerce [1]. They offer considerable advantages, such as the ability to provide a shopping experience that resembles face-to-face shopping and the opportunity to present products that are highly personalised to the customer's profile [2–4]. However, the adoption of these technologies requires considerable investment in skilled staff and state-of-the-art technological resources [5]. While large international companies have these resources at their disposal, marking the next evolutionary steps, small businesses are generally at a disadvantage in terms of competitiveness. This dynamic is leading to a growing gap between the e-commerce giants and small businesses, further widening existing inequalities.

Faced with these challenges, there is a need to look for solutions that reduce this gap, improve the competitiveness of small businesses and strengthen local economies. To improve this competitiveness, it is necessary to help small businesses go digital with technologies that will shape the commercial situation in the years to come. Due to the changes in consumption patterns in recent years, it is critical to provide these small businesses with technological solutions that abstract them from the difficulties of implementation and

management, while enabling interaction with potential buyers. All of this with the aim of not being continually in a notable technological gap.

Therefore, a first step must be overcome: the digitisation of the products to be displayed in the virtual environment. Studies show that the realistic appearance of these products contributes significantly to a higher purchase intention [3,4]. Unfortunately, product scanning requires high-cost devices for high-quality results, such as advanced 3D scanners, which are often beyond the economic reach of retailers and require technical know-how to operate. Without the possibility of acquiring these scanners, software is needed to facilitate this task and integrate the products into e-commerce platforms.

Therefore, it is necessary to consider low-cost and user-friendly digitisation methods that enable realistic digitisation of products. In recent times, state-of-the-art hardware and software technologies have emerged which, combined with AI techniques, can improve the results [6]. To the best of our knowledge, these technologies have not yet been analysed and evaluated in depth within the context of VR shopping.

The aim of this work is to analyse and experiment with technologies that enable small retailers to digitise their product catalogues at low cost. The results of this work will be used to establish guidelines for the low-cost scanning technologies that we are investigating in this work and that will be available to such retailers.

In addition to realism, it is essential to ensure that digitised products are integrated into a virtual environment without negatively affecting their performance. Too much detail, or the simultaneous display of a large number of products, can compromise the performance of the system, leading to dizziness [7], uncomfortable usage of VR glasses and, in general, a poor User Experience (UX). Therefore, this work includes a study that seeks a balance between realism and performance to ensure customer well-being and future effectiveness in product sales. The results provide guidelines for the design of optimal VR experiences for e-commerce.

Our approach is therefore based on two fundamental aspects: (i) to quantitatively study the visual quality of 3D models generated by different low-cost scanning tools and (ii) to analyse the performance offered by those models that ensure higher visual quality in order to determine the workload that current VR headsets on the market can support while providing a pleasant user experience. For the first part, we will select a set of objects whose characteristics do not differ from the best practices of the selected scanning tools and then scan and analyse them according to a set of metrics. For the second part, we will create scenes in the Unity game engine with the models of the highest visual quality, until we determine a theoretical maximum where the performance is poor from the user's point of view. These scenes will be run standalone, and the data collected through a profiling tool will be further analysed.

The rest of the paper is structured as follows. In Section 2, a brief description of VR-ZOCO's architecture is provided. In Section 3, related works regarding the scanning technologies that will be addressed and their usage for VR applications are presented, as well as the current state of VR in e-commerce field. Section 4 describes the process of selecting scanning technologies, the hardware used for scanning and the scanning methodology. Moreover, it details the metrics used for 3D model quality evaluation, as well as the methodology used for scanning and the evaluation of the 3D models. Section 5 provides details of the performance of these models in VR environments tested on the Meta Quest 2, Quest Pro and Quest 3 headsets. Section 6 presents and optimisation process for 3D models, motivated by the results shown in Section 5, and it shows a comparison that addresses visual quality and performance with the optimised models. Finally, Section 6 presents the conclusions drawn and outlines future research work.

2. Related Work

In this section, we review papers related to current scanning technologies and their applications in object scanning. We will focus on technologies such as LiDAR sensors and

NeRF. Subsequently, we discuss works that explore the integration of scanned objects into VR shopping environments. Lastly, we address the influence of realistic 3D models in e-commerce.

2.1. NeRF (Neurance Radiance Field)

NeRF is a state-of-the-art deep-learning approach for creating high-quality 3D models from 2D images. NeRF uses a fully connected neural network to represent a scene or object [8], which is trained using a set of images and a set of camera poses. Based on the viewing direction and camera position, the network generates the volume density and emitted radiance at that location in 3D space. The volumetric representation is then used to render novel views of the scene from any viewpoint in real-time [9].

The work discussed in [10] addressed the challenges in VR applications arising from the demand for high-quality, low-latency rendering in environments with wide field-of-view and stereoscopic viewing. NeRF has shown potential in rendering photorealistic 3D scenes with view-dependent effects. However, their application in VR faces hurdles due to significant latency and quality drop in high-resolution and wide-view scenarios, which can negatively affect the user experience and even induce sickness. To overcome these issues, the authors presented the first gaze-contingent 3D neural representation and view synthesis method for VR. This method incorporated human visual and stereo acuity into an egocentric neural representation of 3D scenery. By optimising both latency and visual quality and aligning them with human perception, their approach achieved a significant reduction in latency (up to 99% time reduction compared to standard NeRF) while maintaining high-fidelity rendering. The results are perceptually identical to full-resolution ground truth, marking a crucial step towards advanced VR systems that can capture and visualise remote environments in real-time. The effectiveness of this method has been validated through both objective analysis and subjective studies.

The authors of [11] presented an innovative end-to-end pipeline for creating dynamic 3D video content from monocular video footage. It can efficiently process video on consumer-grade hardware, reducing the processing time from days to mere minutes per second of footage. Key aspects of this pipeline include the estimation of camera parameters, depth maps and the 3D reconstruction of both dynamic foreground and static background elements. The pipeline also encompasses the rendering of these 3D videos on either a computer or a VR headset. A notable feature is the use of a state-of-the-art visual transformer model for estimating depth maps, which enhances the capability for RGB-D fusion with estimated depth data. The paper reports preliminary experiments where the output was rendered in a VR headset and visually compared against ground-truth datasets and other leading NeRF-based methods.

Li et al. presented and evaluated a NeRF-based framework that can render scenes in immersive VR, allowing users to freely explore complex real-world scenes [12]. The framework's performance was benchmarked across three different NeRF scenes, demonstrating its capability to achieve frame rates of 30 frames per second at a resolution of 1280×720 pixels per eye, using super-resolution techniques. The authors noted that their implementation of instant neural graphics primitives for immersive NeRF in VR environments shows promising results, particularly in rendering small-sized scenes with frame rates of up to 38 fps. This framework stands out for its ability to rapidly digitise real-world 3D environments from unconstrained 2D images and to compress scene data efficiently. The Unity-based implementation allows for adaptability to various Head-Mounted Displays (HMDs) and sensors, opening up possibilities for features like foveated rendering Feld of View (FoV) adjustments.

A more recent work of some of the previous authors [13] focused on an implementation of real-time Variable Rate Shading (VRS) in VR, building upon the advancements of NeRF. The key innovation lies in merging multiple pixels into one, thereby reducing overall rendering time and allowing for the real-time visualisation of small to medium NeRF scenes. This system is also applicable in 2D settings, showcasing its versatility. It enabled

Appl. Sci. **2024**, *14*, 6037

rendering small to medium NeRF scenes on moderate and gaming-grade hardware, and on commercially available HMDs like Quest 2.

2.2. Photogrammetry

Photogrammetry is one of the oldest techniques for digitising objects in 3D graphics. The process of data handling involves using software to reconstruct the 3D geometric shape of an object from various images. Recent advancements in automated reconstruction techniques now allow for quicker creation of 3D models from a collection of photos.

Obradović et al. developed a VR application based on photogrammetric models, specifically applied to the iconostasis of the Serbian Orthodox Cathedral Church of Saint Nicholas [14]. The research outlined a comprehensive workflow, including the generation of an accurate 3D reconstruction of the iconostasis using photogrammetry, model optimisation, retopology, control and analysis, followed by the creation of the VR experience using a game engine. The result is an interactive walkthrough of the church. This method showed the significant role of photogrammetry in creating realistic 3D models for VR applications, particularly in cultural heritage preservation.

The work of [15] also provided insights into cultural heritage promotion. They proposed a VR solution using photogrammetry techniques to enhance heritage promotion in a tourist centre in Trujillo. The research identified deficiencies in heritage promotion, particularly in its diffusion and perception by the population. The authors tested with 30 participants and confirmed that the proposed VR solution significantly improves the promotion process. In addition, they underscored the role of realistic 3D models created through photogrammetry in delivering a more engaging and authentic experience to tourists.

In another application domain, Tadeja et al. researched the use of photogrammetry to generate realistic 3D models of physical structures for use in VR, particularly in engineering surveys [16]. The study explores how digital twins (digital replicas of physical objects) created through photogrammetry can serve as effective substitutes for real-world engineering inspections. The authors explored the effectiveness of bimanual gestural input, augmented with gaze-tracking, in interacting with these 3D models in VR. A key outcome of the study was the demonstration that such an interaction mechanism is an effective modality for domain experts in VR, facilitating efficient manipulation and accurate measurement of elements within the photogrammetric 3D models. The qualitative feedback from the six domain experts involved in the case study underscores the potential of this method in professional settings.

2.3. LiDAR Sensors

LiDAR (Light Detection and Ranging) sensors are another technology used for 3D scanning. LiDAR works by emitting laser beams and measuring the time it takes for the beams to reflect off objects in the environment and return to the sensor [17]. By measuring the time of flight of the laser beams, LiDAR sensors can obtain precise depth information about objects in the environment, which can be used to create 3D models.

LiDAR sensors have been used in works from several application domains. Mikita et al. explored the use of various laser scanning technologies to assess damage to forest road surfaces [18]. The study focused on comparing the accuracy and applicability of three different laser scanning methods: Terrestrial Laser Scanning (TLS), hand-held Personal Laser Scanning (PLShh) with GeoSLAM ZEB Horizon and hand-held personal laser scanning with iPhone 13 Pro using 3D Scanner and Polycam apps. The research aimed to determine the effectiveness of these technologies in accurately mapping forest road surfaces and identifying damage compared to traditional tacheometric surveying methods. Specifically, the iPhone 13 Pro with 3D Scanner app demonstrated a positional height accuracy with a Root Mean Square Error (RMSE) of 0.185 m (X,Y) and 0.021 m (Z), while with the Polycam app, the RMSE was 0.31 m (X,Y) and 0.045 m (Z). The GeoSLAM ZEB Horizon achieved a (X,Y) RMSE of 0.108 m and a (Z) RMSE of 0.025 m. Comparatively, Terrestrial Laser Scanning (TLS) achieved the best results with a (XY) RMSE of 0.049 m and a (Z) RMSE of

0.0077 m. These results indicate that while TLS provides the most precise measurements, both the Apple's iPhone 13 Pro and the Geonorte's GeoSLAM ZEB Horizon scanners produce sufficiently accurate results for practical applications. The iPhone 13 Pro, in particular, offers a cost-effective and rapid alternative for assessing road conditions.

The work discussed in [19] studied the viability of using Apple's iPad Pro's LiDAR and True-Depth camera systems as 3D scanning solutions compared to an industrial 3D scanner. The authors used different coloured Lego bricks to assess the accuracy of these technologies, focusing on shape and position tolerances. The data from the study revealed that the iPad Pro's LiDAR and True-Depth systems, while not as accurate as the industrial 3D scanner, still showed considerable potential. The industrial scanner consistently outperformed the iPad in terms of accuracy, but the iPad's technologies were deemed sufficient for specific applications where extreme precision is not a critical requirement.

Ferreira et al. presented a methodology for using a Digital Surface Model (DSM) within a VR environment focusing on its application in military operations in the Amazon [20]. The study details the use of an Unmanned Aircraft System (UAS) equipped with LiDAR sensors to gather data, which is then modelled using software like Blender and Sketchup. This model is used to create a VR environment for immersive terrain analysis, making it a simple and low-cost way to explore a DSM in VR. The paper highlights the potential of integrating modern technology like VR and UAS in strategic and tactical planning, particularly in difficult terrains, enhancing situational awareness and operational efficiency. Finally, the authors verified that a Ground Sample Distance (GSD) greater than 1.93 cm/pixel begins to present a significant loss of the terrain characteristics considered important.

2.4. VR in E-Commerce Field

In the literature, several works on the prototypes of virtual environments for commerce can be found. Ricci et al. [21] carried out a comparative study of Immersive Virtual Reality (IVR) and Desktop Virtual Reality (DVR) in virtual fashion stores, noting a significant enhancement of the shopping experience through IVR. The research, involving 60 participants, focused on evaluating various dimensions of the shopping experience: duration, hedonic and utilitarian values, cognitive load and overall user experience. The findings revealed that IVR led to a longer engagement time in shopping activities, with an average duration notably higher in IVR than in DVR. This extension in engagement time was attributed to the more immersive and interactive nature of IVR. Moreover, both the hedonic aspects (such as enjoyment and telepresence) and utilitarian aspects (like product diagnosticity and usefulness) were significantly enhanced in IVR, indicating a more satisfying and effective shopping experience. Interestingly, the cognitive load was comparable between the two modes, suggesting that the mental effort required by the users in IVR was not excessively higher than in DVR. The study also employed the User Experience Questionnaire (UEQ) to assess various user experience dimensions, which mostly favoured IVR, though some inconsistencies were observed in dependability scores. These findings underscore the potential of IVR in enriching the shopping experience in fashion retail.

The authors of [22] researched how immersive VR shopping experiences influence users' propensity to adopt the technology for future applications. For this purpose, the researchers conducted an experimental study using two different settings: a highly immersive environment enabled by the HTC Vive and a less immersive setup on a standard desktop display. The more immersive scenario involved monitoring and recording the participants' hand and head movements, including interactions with products such as picking up, dropping or passing items from one hand to the other. Eye-tracking data was also collected. The authors' findings indicate that the immersive nature of the VR environment may affect a user's decision to continue using the system through two distinct pathways: hedonic, motivated by enjoyment, and utilitarian, motivated by practicality.

Wu et al. conducted two experiments within the realm of user-defined gestures for interaction in immersive VR shopping applications [23]. Consequently, they suggested a more pragmatic approach for generating more dependable gestures compared to con-

ventional gesture elicitation studies. One of these experiments involved VR shopping scenarios using the HTC Vive headset, where 32 participants were required to engage in 12 different tasks, including selecting objects, changing colours and resizing, among others. However, the authors did not detail the methods used for recording and processing the gesture information.

The work discussed in [24] detailed the development and evaluation of a prototype for a VR online shopping environment, focusing on how user inputs such as head movements and speech, and outputs through desktop and HMD, influence task performance and user preferences and behaviour. The prototype had two versions, employing different input devices for the desktop and HMD versions, which changed how users interacted with the prototype. Following a survey where participants identified the advantages and disadvantages of online shopping, a prototype was constructed for both desktop and mobile VR environments, restricting user input to voice and head gestures. After analysing a case study, they proposed design principles for VR shopping environments. Subsequently, the same authors developed a new VR shopping prototype using the 'Apartment' metaphor [25]. This was aimed at examining various product selection and handling techniques (such as grab and beam) and exploring different shopping cart designs (like basket and sphere). In their study using this new prototype, they found that immersion and user experience were key concerns for the participants. They also offered recommendations to reduce motion sickness and identified which types of products are most suitable for VR stores.

Shravani et al. introduced VR Supermarket, an online VR shopping platform implemented using Unity [26]. This VR system was designed for use with the HTC Vive headset. Additionally, it incorporates a dynamic recommendation system that relies on user purchase history and is hosted on a web server. The VR application not only manages user interactions via controllers and the headset but also interfaces with the web server, which is backed by an SQL database for generating recommendations. Simultaneously, the VR application requires a NoSQL database to handle data storage and retrieval tasks related to products, supplementary information and user data, among other functionalities. The authors presented a brief case study with well-known datasets such as Amazon Groceries to showcase the recommendations made by the system.

2.5. Influence of 3D Models in E-Commerce Activities

It is noteworthy that high quality 3D models of products add realism to the virtual environment, allowing the user to abstract and increase their feeling of being in a physical shop [27]. In addition, studies have investigated the impact of product graphics quality on characteristics such as informativeness or playfulness, finding that those with higher visual quality graphics have a greater positive impact [3]. The aspects of the user's shopping experience that are favoured have the potential to increase the number of retail sales or the users' preferred use when choosing between a physical or VR-based shop to shop in [4].

In addition, a 3D model with higher-quality graphics can better capture the details of products that may contain minor flaws, which can enable small shops to display second-hand products and allow shoppers to feel more confident about the condition of the product they are buying. In addition, products that identify the towns in which small shops are located can benefit if the models can capture a greater number of details. Given that customers respond more positively to promotions in VR shops than in physical stores [27], it is worthwhile to obtain realistic 3D models so that small businesses can benefit even more from the platform. Therefore, we seek to analyse and compare scanning tools that are low-cost so that they do not pose a financial problem for retailers, easy to use so that they do not require skilled staff to use them, and that produce high-quality results to attract more attention from customers.

3. Low-Cost Digitisation for VR E-Commerce

Having established the importance of realistic 3D product display and its influence on immersive e-commerce, as well as the resource constraints faced by smaller businesses,

we proceed to the study of various scanning methods that try to find a balance between quality and cost, as well as good performance during real-time graphics processing using VR Headsets.

Therefore, in the following subsections we will present in detail the software and hardware used in the study for the digitisation of physical products, characteristics and quality metrics for the comparison of the digitisation techniques and, finally, the experimentation carried out with some products considered as samples.

3.1. Software Tools Selection Based on Scanning Techniques

In Section 2, state-of-the-art 3D scanning techniques were introduced, including photogrammetry, LiDAR mapping and NeRF. In the current section, we will focus on choosing low-cost tools and technologies that support these techniques and the needs of small businesses. We would like to point out that all the selected software tools have a user interface very similar to that of the camera application integrated with the mobile phone. In addition, they usually include tooltips and visual guides to perform scans in video mode. This makes the learning curve very easy for people who are not experts in the use of scanning technologies [28], which is an incentive to suggest the use of these applications as opposed to a 3D scanner.

LiDAR sensors were first integrated in smartphones only in iPhone and iPad devices from the 4th generation (2020). Several authors have conducted studies with these sensors that show good results in terms of accuracy of the generated models. For example, ref. [29] stated that the accuracy is of ± 1 cm for small objects with side length greater than 10 cm, with the limit of detection for objects being around 5 cm. This suggests that LiDAR sensors might be more accurate and preferable to use for large objects. Even though these devices are considerably cheaper compared with professional 3D scanners, they might not be affordable for every small business owner. Therefore, it is necessary to explore scanning tools available in Android-based devices, as this operating system for mobile devices has 71.44% market share globally (https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/ accessed on 7 July 2024) and is the operating system installed in the majority of mobile devices. Therefore, we searched for scanning applications available in either Android or iOS to select a set of tools that allow us to study the three technologies mentioned above. Even though some applications found had both Android and iOS versions, the features offered differed in some cases. Therefore, we proceeded to explore scanning applications from various sources, including the literature, to identify the most used ones by researchers.

Regarding the current state of photogrammetry applications, the image acquisition phase of such a process is easily done by the cameras of current mobile devices [30]. These cameras have superior specifications to those on previous years' mobile phones, especially when they are high-end devices, allowing one to obtain high-quality images from which to generate the 3D models. In addition, the rest of the phases (image processing, point cloud generation, mesh generation and post processing) are done by more powerful processors in current mobile phone devices or even by a cloud-based computing process. Therefore, we explored free or low-cost applications available for Android devices. Moreover, we considered user's ratings and reviews as well as the published results made with such applications in order to decide whether to consider them or not. Since all of them are based on the photogrammetry technique, we also considered additional features made public by the developers of such applications, such as editing tools for the obtained model inside the application, estimated time for generating the model, subscription prices, exporting formats or the guides offered for scanning.

With respect to NeRF, we believe that this novel technique can offer high-quality results, even when compared with professional scanners. For instance, the authors of [31] showed encouraging results, where, compared with Agisoft Metashape software 2.1.0, the accuracy comparisons performed over different datasets gave differences of less than 1 cm, averaging around 0.5 cm. Due to the novelty of this technique, we have only chosen Luma AI since

it is free and accessible to small retailers, aligning with our goal of identifying low-cost solutions, and it has been more used in the literature, demonstrating its effectiveness and reliability in generating high-quality 3D models using the NeRF technique. Table 1 shows the relevant features of the found scanning technologies regarding the economic resources of small businesses.

Table 1. Features of found scanning applications for Android and iOS devices . P = Photogrammetry, L = LiDAR, N = NeRF.

Feature	Polycam 1.3.10	Widar 4.1.3	Kiri Engine 4.4.0	Luma AI 1.3.8	MagiScan 1.9.7	Scaniverse 3.1.2 (iOS Only)
Technique Supported	P, L	P, L	P	N	P, L	P, L
Subscription Price	USD 14.99/m	USD 9.99/m	USD 15.99/m	Free	USD 9.99/m	Free
Free Export Formats	.gltf	No free export	.obj	.fbx, .obj, .glb, .gltf, .usdz, .stl, .ply, .xyz, .dae	.fbx, .obj, .glb, .gltf, .usdz, .stl, .ply	.fbx, .obj, .glb, .gltf, .usdz, .stl, .ply
Free Number of Export	Unlimited	No free export	3/week	Unlimited	3/account	Unlimited
Edition Tools Included	Yes	Yes	Yes	Yes	No	Yes
Object Masking (Photogrammetry)	Yes	No	Yes	-	No	No
User Rate	4.3/5 (16,227 reviews)	3.7/5 (2268 reviews)	4.1/5 (1231 reviews)	4.7/5 (1575 reviews)	4.0/5 (4936 reviews)	4.8/5 (7585 reviews)

The feature for exporting the models is essential in order to both evaluate its characteristics and to use it in further VR applications. Note that the pricing for these applications is very similar, hence, the pricing is not a differential factor. However, after conducting a test scan under the same conditions on one of the objects selected with these applications, we decided to use Polycam as the application for evaluating photogrammetry and LiDAR techniques. Polycam was selected for its user-friendly interface, extensive usage in the literature, and comprehensive feature set. It supports both photogrammetry and LiDAR scanning technologies, which are crucial for our comparative analysis. Additionally, Polycam offers multiple export formats and editing tools, making it a versatile option for small businesses aiming to digitize their product catalogs for VR shopping environments

Therefore, we selected Polycam and Luma AI as the scanning applications to be used in this study, the second one for being the only one to use NeRF technology by the time we carried out the selection process. Once we finished the selection process described previously, we began the scanning process as well as the definition of characteristics to be measured.

3.2. Hardware Used

The deviceswe employed to scan the objects are as follows:

1. Apple iPad Pro 2022 (Ciudad Real, Spain). Dual camera of 12 megapixels (MP) and 10 MP ultra-wide, plus 3D LiDAR sensors incorporated. It is the device on which we used Luma AI and Polycam's LiDAR feature.
2. Xiaomi Redmi Note 8 (Ciudad Real, Spain). Quad camera of 48 MP (main), 8 MP (ultra-wide), 2 MP (macro) and 2 MP (depth).

The reasons for selecting these two devices are as follows. First, we required an iOS device to be able to install the Luma AI app, as well as to have access to LiDAR sensors

on a device that is cheaper and easier for the public to use than a professional 3D scanner. Second, we chose an Android device to use the Polycam app with photogrammetry. We consider that such a device has a camera with sufficient MP and number of lenses to provide good quality photographs. Taking into account the economic possibilities of small shops, the chosen Android device presents a good reference, and the results of photographs with cameras of higher-end devices could improve the exposed results.

3.3. Features and Quality Metrics Used

We aimed to retrieve characteristics that could help us identify its quality and detail and gain an insight into the performance of a VR application that uses a considerable number of such models. We aimed to select state-of-the-art 3D graphics quality metrics that worked with 3D model representations, either meshes or point clouds. Hence, to select and use such metrics, we required that (a) the process followed by the metric takes as input a 3D mesh or 3D point cloud, easily obtained from the scanned model and (b) the source code of the metric was public in order to be able to use it with our data, for which papers with code (https://paperswithcode.com accessed on 7 July 2024) web page was useful.

The features of the 3D models we collected were:

1. Size in MB of the .obj file.
2. The number of polygons of the model (polygon count).
3. The texture's size of the model in MB.
4. NR-3DQA 1.0 (https://github.com/zzc-1998/NR-3DQA accessed on 7 July 2024) [32]. This metric, which does not need a reference model, takes as input a point cloud or 3D mesh, extracting colour and geometric features, such as curvature, sphericity or linearity. After estimating four statistical parameters, a vector of 64 features is constructed and fed to a Support Vector Regression (SVR) neural network to provide a quality score in the form of a Mean Opinion Score (MOS) obtained by regression. For training, the public database Waterloo Point Cloud (WPC) [33] is used, which comprises 760 point clouds including 20 original ones and others distorted through different methods. These clouds are evaluated by obtaining an MOS, which represents the visual quality with which the point cloud is perceived by the human visual system.
5. MM-PCQA 1.0 (https://github.com/zzc-1998/MM-PCQA accessed on 7 July 2024) was proposed by Zhang et al. [34]. This metric, which does not need a reference model, takes as input a point cloud in .PLY format. The proposed metric works by first splitting the point cloud into sub-models to capture local geometric distortions such as point offsets and down-sampling. It then converts the point clouds into 2D image projections to extract texture features. Both the sub-models and the projected images are encoded using point-based neural networks, such as PointNet++ and DGCNN (Dynamic Graph CNN), and image-based neural networks. The main innovation of this metric is the symmetric cross-modal attention mechanism, which fuses quality-aware multi-modal information to predict the quality score. The fused features are then used to predict the quality score of the point cloud. The loss function used to train the model includes components to minimise the difference between the predicted and actual quality scores and to maintain the quality ranking between the different point clouds. In this case, the WPC database was also used to train the neural networks.
6. CMDM (Color Mesh Distortion Measure) 0.15.1 (https://github.com/MEPP-team/MEPP2 accessed on 7 July 2024) [35]. This metric, which uses a reference mesh, evaluates the quality of a 3D mesh with colour, combining geometric and colour features to determine the perception of distortion. It is based on the extraction of features at a local level and their subsequent analysis at a global multi-scale scale. Geometric features include mean curvature and curvature structure comparison, while colour features are extracted by transforming the colours of vertices to the LAB200HL colour space, considering luminance, chroma and hue. These features are compared between the reference and the distorted or compared mesh, and an optimised linear combination

of these features, determined through logistic regression, results in a measure of perceived quality.

7. Accuracy of the 3D model. This can be obtained by performing a comparison between triangular meshes of a reference model and the model you want to evaluate for accuracy. For this purpose, we will use the Root Mean Square (RMS) of the distances computed with the software CloudCompare 2.6.3 (https://www.danielgm.net/cc/ accessed on 7 July 2024), which is widely used in the literature for comparing meshes and point clouds. Thus, we will follow the recommendations provided by the tool for performing the comparisons. Since we do not have a ground truth model of the product being scanned, we are taking the 3D models provided by the most accurate scanning technology based on the results of Section 3.4.

To obtain the values of the features numbered from 1 to 3 for each 3D model, we used the MeshLab 2023.12 (https://www.meshlab.net accessed on 7 July 2024) software tool for inspecting them. Moreover, it was used for exporting the .obj 3D models into .ply files for usage in NR-3DQA, MM-PCQA and CMDM. With regard to quality metrics numbered from 4 to 6, the code from their public repositories was installed and configured for use. For the metrics that required a cloud point as input, we sampled a cloud point of 1 million points based on the 3D mesh obtained, since the scanning applications provided a point cloud of the whole scene. Even though we attempted a cleaning process, the number of points in the resulting clouds was low when compared with those in the WPC database, used for training the neural networks used in two of the three metrics. Therefore, we decided to sample the point clouds with the same number of points for each model to ensure that all the meshes' details were retained. Meanwhile, for evaluating the accuracy of 3D models, we used CloudCompare for comparing meshes (see Section 3.3). For metrics that need a reference model, such as CMDM and accuracy, we will use the model that provides the most detail based on the number of polygons. We lack a ground truth model, as we do not have a professional 3D scanner to provide us with a highly accurate and reliable model. Therefore, we will see in the next section how we can deal with the bias of selecting the most detailed model provided by the selected software applications.

3.4. Scanning of Basic 3D Primitives

Because of the above, we decided to build, scan and model the objects shown in Figure 1. The cube, pyramid and sphere are basic primitives used in computer graphics software and are easy to model in a program like Blender 3.5 (https://www.blender.org/ accessed on 7 July 2024) with little experience. So are the letters 'A', 'I' and 'R', in this case in Times New Roman font. After obtaining the necessary dimensions for the model of the expanded polystyrene objects and modelling them in Blender, we obtained a model that was close to the real one, with a significantly reduced amount of polygons, which allowed us to calculate the error of the technology we used as a reference in the comparisons. Thus, despite having the most detailed model of one of the tools we use as a reference, we will be aware of the error it has in advance.

Following the scanning process defined in Section 3.5, these figures were scanned with Luma AI to export the high-poly models as a result, as well as with both devices. Then, we used the 3D models designed with Blender as a reference model to obtain the results of the MSDM2 (Mesh Structural Distortion Measure 2) metric [36] and the accuracy of the 3D model (see Section 3.3) obtained as a result of the scanning. In this way, we quantified the inherent errors these scanning applications have. The philosophy behind MSDM2 is very similar to that of CMDM (see Section 3.3), which aims to provide a score of the subjective visual quality, as perceived by human vision, of a 3D model with respect to a reference model. However, unlike CMDM, this metric considers features related to colour, and we decided not to use it since the models created with Blender did not have textures. This decision was made because it could compromise the accuracy when obtaining the score.

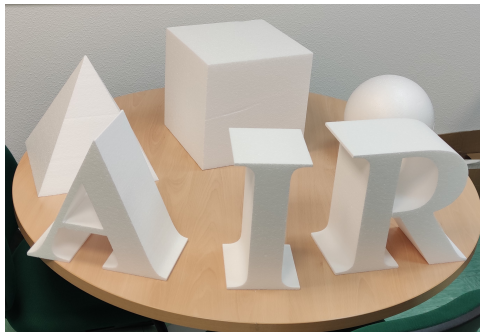


Figure 1. Set of 6 expanded polystyrene Objects used.

Table 2 shows the results obtained for the mentioned metrics. The RMS values obtained with the comparisons obtained from CloudCompare 2.9.3 have resulted in 0.00425 ± 0.00253 m. That is, on average we can count on an error of approximately 0.42 cm between the distance from one point to another of the real object with respect to the generated 3D model. Depending on the dimensions of the object, 0.42 cm could be critical, therefore, considering the dimensions of real expanded polystyrene objects (around 30 cm), which can be consulted in the following repository: https://github.com/AIR-Research-Group-UCLM/PDIVR-ZOCO accessed on 7 July 2024. Taking these dimensions into account, the average error is 1.4 %. With respect to the MSDM2 values obtained, the mean is 0.112375. A value of 0, as in CMDM, indicates that the meshes are identical. Furthermore, in the work of the authors of the metric, graphical examples are shown in which a value of 0.14 is barely noticeable to the human eye [36]. With these results, we can anticipate the error made in the scans when using this type of model as a reference for comparisons.

Table 2. Metrics obtained from the expanded polystyrene figures comparisons.

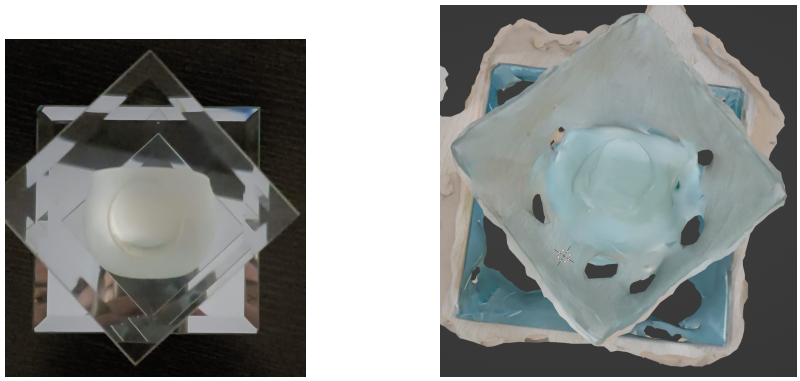
Metric	Cube	Sphere	Pyramid	A	I	R
NeRF (Luma AI High-Poly Models)						
Root Mean Square (RMS)	0.00144	0.00751	0.00125	0.00750	0.00386	0.00394
MSDM2	0.10303	0.14509	0.08394	0.12386	0.10182	0.11651
Photogrammetry (Polycam Android 1.3.10)						
Root Mean Square (RMS)	0.00483	0.00945	0.00432	0.00221	0.00391	0.00417
MSDM2	0.13715	0.18423	0.10411	0.11910	0.13607	0.11754
LiDAR sensors (Polycam iOS 1.3.10)						
Root Mean Square (RMS)	0.00701	0.01179	0.00471	0.01045	0.01350	0.0055
MSDM2	0.12862	0.24176	0.14615	0.16083	0.34365	0.17143

3.5. Scanning of Objects Selected for the Study

We have followed best practice guidelines provided by Polycam 1.3.10 (https://learn.poly.cam/about accessed on 7 July 2024) and Luma 1.3.8 (https://docs.lumalabs.ai/MCrGAEukR4orR9 accessed on 7 July 2024), including good lighting and moving slowly while capturing or avoiding objects with transparent materials or complex reflections. There are a set of well-known materials for which it can be hard to obtain good-quality results from a scan. Those are transparent materials (i.e., glass or plastic), reflective materials (i.e., mirror-like or metallic surfaces), smooth and even surfaces (i.e., a white wall) and furry

Appl. Sci. **2024**, *14*, 6037

materials (i.e., a carpet). It can be seen in Figure 2b that the mirror-like material of the object is badly represented, with holes in the 3D model and the bottom part not being filled. Therefore, we tried to avoid objects for this work with such materials in order to obtain high visual quality results for later VR experimentation, in which a high level of realism is important. However, we selected two objects that have one of the materials listed above to research the impressions caused on users in further studies.



(**a**) (**b**)

Figure 2. Object based on a mirror material. (**a**) Decorative candle with a mirror; (**b**) 3D model of decorative candle with a mirror in Blender.

Although the guidance interface in Luma is more comprehensive than that of Polycam, it required, on average, 4 min more to complete the rings. We used a round table with a radius of 50 cm to rotate around the object placed on it to capture the different points of view. We did not take into account the processing time for generating the model since both applications are based on cloud computing to process it. However, the usage of LiDAR sensors in Polycam allows for local processing, generating the model in a range of 1 min 30 s and 3 min, depending on the product scanned.

We note that the scanning time with Polycam was around 4 min, while with Luma it was around 8 min. However, the model generation time depends on the workload of the cloud servers at that time. Bearing in mind that these times can vary, models generated with Polycam took around 3–5 min to generate, while those with Luma AI could take up to 10 min. However, these times were obtained at the time of the study and may change in the near future.

It is important to consider the various formats that these applications offer for generating 3D models. Ordered by the lowest to the highest detailed model given, Luma AI offers low-poly, medium-poly and high-poly, while Polycam provides optimized, medium, full and raw. Regarding Polycam, the last two options are better for VFX effects and professional workflows use cases, whereas the two first are designed for game engines [37]. Nonetheless, the optimised option is more convenient for fast loading and real-time rendering environments, which is our target by allowing potential customers to interact with the 3D models of products in the VR environment. Although it could seem that, among the options provided by Luma AI, the best for our use case is low-poly, we have explored every format since there is no existing research specifically focused on this technology for our use case.

We selected five objects (see Figure 3) made with different materials and from different sizes that could be potentially showcased in a VR e-commerce platform. In addition, we considered objects with a size large enough for LiDAR sensors to detect them effectively, since they commonly provide worse results with small objects, such as LEGO blocks [19].

Appl. Sci. **2024**, *14*, 6037

Therefore, we selected objects sized between 20 and 100 cm in height or width. It was also important to select objects with different colours, as LiDAR sensors seem to return more 3D data points when lighter coloured objects are scanned [6,19]. Figure 4 shows the 3D models in .obj extension exported in low-poly format from Luma AI.

The selected objects for this scanning study represent a spectrum of surface types and their interactions with light, which are essential for assessing scanning effectiveness and detail accuracy. We include organic materials and textiles, which predominantly absorb and scatter light, serving as examples of diffuse surfaces. Ceramics are chosen for their semi-specular properties, providing a balanced mix of light absorption and moderate reflection. Metallic objects, known for their highly reflective, specular surfaces, are included to test the scanner's ability to handle intense reflections and mirroring effects.



Figure 3. Collage of the 5 objects scanned in real life.



Figure 4. Collage of the 3D models in low-poly from Luma AI.

Appl. Sci. **2024**, *14*, 6037

3.6. 3D Model Quality Evaluation

Table 3 shows the measurements taken from each 3D model obtained from the scanning process for each of the features described above. On the one hand, it is noteworthy that the polygon count from Luma AI's high-poly models are between 9 and 10 times higher compared to those of medium-poly models, and approximately 50 times higher than those of low-poly models. This seems to indicate that high-poly models will not be the most feasible to use in VR environments in terms of the impact they will have on VR headset performance. However, this hypothesis will be tested in Section 4.

Table 3. Data extracted of defined features from each 3D model. HP = High Poly, MP = Medium Poly, LP = Low Poly, OP = Optimised.

Object	Type	Scanning Application	Device Used	Size (MB)	Polygon Count	Texture's Size (MB)	NR-3DQA	MM-PCQA
Munich shoes	HP	LUMA AI	iPad Pro	98.94	999,846	18.50	0.7272	**86.06**
	MP	LUMA AI	iPad Pro	13.67	149,915	13.10	0.8101	83.52
	LP	LUMA AI	iPad Pro	2.16	24,893	10.20	0.7617	84.85
	OP	Polycam	iPad Pro	4.04	39,627	0.42	0.7363	80.67
	OP	Polycam	Xiaomi Note 8	1.94	19,424	2.30	**0.8111**	82.18
Octopus teddy	HP	LUMA AI	iPad Pro	99.47	999,869	8.60	0.6619	**78.51**
	MP	LUMA AI	iPad Pro	13.50	149,916	5.19	**0.6914**	70.92
	LP	LUMA AI	iPad Pro	2.10	24,920	3.67	0.6537	66.25
	OP	Polycam	iPad Pro	0.25	2689	0.45	0.4288	20.51
	OP	Polycam	Xiaomi Note 8	1.84	18,623	1.99	0.5539	65.98
Burner	HP	LUMA AI	iPad Pro	53.54	528,442	4.64	0.6190	70.43
	MP	LUMA AI	iPad Pro	14.30	149,972	3.71	**0.6622**	69.73
	LP	LUMA AI	iPad Pro	2.30	24,981	2.77	0.6487	**70.63**
	OP	Polycam	iPad Pro	0.64	6693	0.29	0.6381	67.50
	OP	Polycam	Xiaomi Note 8	1.93	19,678	1.54	0.6238	69.97
Ceramic mushroom	HP	LUMA AI	iPad Pro	98.94	999,731	2.97	0.7065	**78.59**
	MP	LUMA AI	iPad Pro	13.67	149,875	1.16	**0.7948**	61.10
	LP	LUMA AI	iPad Pro	2.16	24,814	0.87	0.7226	72.60
	OP	Polycam	iPad Pro	0.96	10,161	0.30	0.6748	42.66
	OP	Polycam	Xiaomi Note 8	1.05	10,917	2.26	0.6542	63.071
Trophy	HP	LUMA AI	iPad Pro	39.2	329,208	6.43	0.7606	**76.17**
	MP	LUMA AI	iPad Pro	13.9	149,923	4.93	**0.8001**	73.03
	LP	LUMA AI	iPad Pro	2.2	24,964	3.97	0.7974	69.61
	OP	Polycam	iPad Pro	1.49	13,416	1.83	0.4474	69.85
	OP	Polycam	Xiaomi Note 8	3.39	36,105	12.2	0.4460	72.90

As can be seen, the sizes of the Luma AI high-poly files are very large compared to the other types, as are the sizes of the textures, since they have a higher resolution. In the case of textures, the scanning applications do not offer control over the desired resolution when exporting them. In the case of Polycam, a single texture file was generated, 4096 × 4096 in the case of the Android version (photogrammetry), and in the case of the iOS version (LiDAR sensors) the file obtained is 4096 pixels wide and a variable height depending on the scanned object. On the other hand, Luma AI generates several texture files in the three formats shown in a much more flexible way. For example, for the sneakers in high-poly format, 78 files were generated in which the highest resolution was 4096 × 4096, while for the octopus 56 files were generated in which the highest resolution was 2048 × 2048. The same applies for the rest of the data, generating 3–4 files for low-poly models and between 10–30 for medium-poly ones. This may indicate that the algorithms adapt the textures and their resolution more intelligently than in Polycam depending on the scanned object, since the process and the conditions under which the scanning was performed were identical. However, a downside for these models is

Appl. Sci. **2024**, *14*, 6037

the need to apply more than one material to the model when rendered by a graphics engine, which may worsen performance.

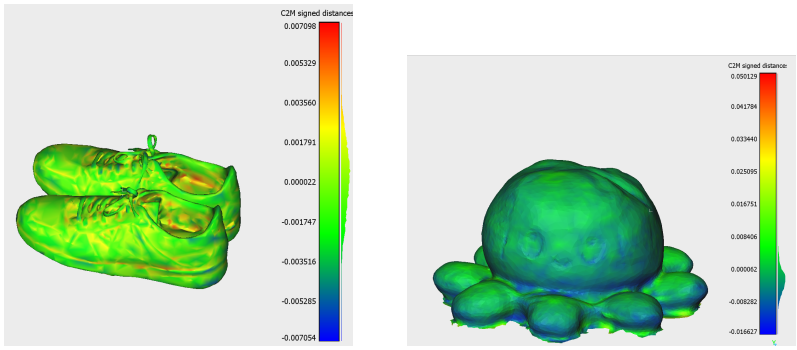
Regarding the quality metrics NR-3DQA and MM-PCQA, a low result does not necessarily indicate low visual quality. Because the neural networks are regression-based, they are trained with the features extracted from the corresponding models and the MOS obtained from the volunteer study present in the WPC database [33]. Since not only original models are used, but also distortions based on them, it is also necessary to take into account the results obtained considering the group of models of the same object.

Analysing the results, we observed discrepancies between the metrics, due to the characteristics of the models used in their algorithms. In the case of NR-3DQA, we noted that the models from Luma performing best were the medium-poly ones, except for the sport shoes, for which by a few decimals it was the Polycam model in its Android version. It can also be seen that, apart from the sport shoes, Luma's models outperform Polycam's in both versions. When we calculate the percentage difference between the top-performing models and those with inferior results, we find the following. The greatest difference is seen in the trophy, where Polycam's models are outperformed by 45% compared to Luma's medium-poly model, followed by the octopus (39% on the iOS version and 20% on the Android version) and the ceramic mushroom (Polycam's Android model) with a 17% difference. The least difference is found in the case of the burner, with just a 6% gap between the best Luma model and Polycam's, and it is the object where the quality of all models is most similar.

The situation with the MM-PCQA results is slightly different, as we see that the high-poly models yield the best results, except for the burner, as it is the low-poly model. These outcomes continue to demonstrate good visual quality for both the Luma and the Polycam models, with the difference narrowing compared to the NR-3DQA results. For Polycam's Android version models, the results (from top to bottom in Table 3) are 4%, 16%, 1%, 19% and 5% lower compared to the highest scoring model. Once again, the burner is the object whose models have been most similarly generated by both applications. For the iOS version models, we can see that the percentage differences are even greater in some cases, with two models standing out where the visual quality is low: the plush octopus (74%) and the ceramic mushroom (46% less).

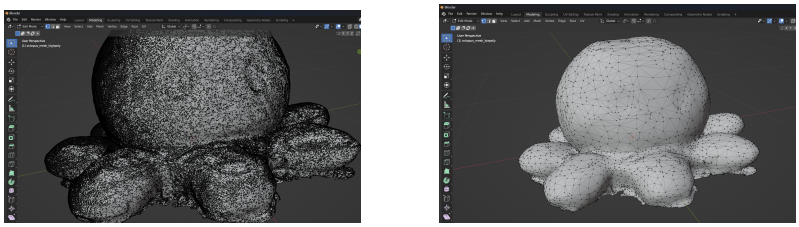
Based on these results, we can determine that Luma AI's models offer the highest visual quality, with percentage differences ranging from 1% (MM-PCQA of the burner) to 74% (MM-PCQA of the octopus teddy). By calculating the mean and median of the percentage differences between Luma AI's top-scoring model and Polycam's models, we can more generally visualise Luma AI's superior visual quality of models. Thus, for both metrics, we obtain the following results: (i) NR-3DQA. Mean: 17.78%, median: 17.72% better results compared to Polycam's photogrammetric models (Android), mean: 22.73%, median: 15.18% better results compared to the models obtained with Polycam's LiDAR sensors (iOS); (ii) M-PCQA. Mean: 9.43%, median: 5.2% better results compared to Polycam's photogrammetric models, mean: 28.19%, median: 9.21% better results compared to the models obtained with Polycam's LiDAR sensors.

In Figure 5, we can observe a comparison of models where a colour scale visually indicates the distances between the points of the point clouds of both models. In those parts of the compared model where colours show more red or blue hues, it indicates that the differences in distances are greater, which translates into less precision in the compared model. Figure 6a,b present a comparison of the meshes from high-poly and low-poly 3D models. It can be noted that the number of vertices and edges in the high-poly model is significantly greater than in the low-poly one, but this implies much more computation time and effort by the game engine that would render such a model. Moreover, Figure 7 shows a comparison in details with the different technologies used and the real object scanned.



(**a**) (**b**)

Figure 5. Comparison of 3D models performed using CloudCompare 2.9.3. (**a**) High-poly shoes as reference and low-poly shoes as compared; (**b**) High-poly teddy as reference and Polycam in Android's version teddy as compared.



(**a**) Octopus teddy's high-poly mesh. (**b**) Octopus teddy's low-poly mesh.

Figure 6. High-poly and low-poly meshes generated by Luma AI from one of the objects.

Furthermore, the accuracy of the generated 3D model is presented in the remainder of this subsection. As stated before, given the absence of a ground truth model, high-poly models of each object have been employed as a reference model, as the data obtained confirm their ability to capture the most intricate details. The process implemented in the CloudCompare 2.9.3 tool is comprehensively outlined in its corresponding documentation. In addition, this tool provides a statistical distribution of the mean and standard deviation of the distances calculated during the comparison. Table 4 shows the RMS given by CloudCompare 2.9.3 from the comparisons performed, as well as the results of using the MEEP2 0.15.1 platform (https://github.com/MEPP-team/MEPP2 accessed on 7 July 2024) to obtain CMDM metric values for each comparison. In the case of a high difference of polygons, the application's manual recommends sampling points in the compared model. Therefore, the points were sampled to match the number of vertices in the reference models.

Analysing the RMS results, it is expected that the medium and low-poly models of Luma would be the most precise compared to the reference model, maintaining accuracy in millimetres relative to the reference model. For the rest of the models, it can be observed that, except on one occasion, the results of the models obtained with Polycam using photogrammetry indicate greater precision compared to the models obtained with LiDAR sensors. Regarding the models obtained by photogrammetry, the average of the RMS results indicates a 2.36 cm difference in the measured distances.



(**a**) (**b**)

(**c**) (**d**)

Figure 7. Comparison of details between Luma AI low-poly model, Polycam from Xiaomi, Polycam from iPad and the real object. (**a**) Shoe model generated with Xiaomi phone (Polycam); (**b**) Shoe low-poly model generated with Luma AI; (**c**) Shoe model generated with iPad (Polycam); (**d**) Shoes in real life.

By inspecting the CMDM results obtained, we can see that medium-poly models achieved the best results overall, with an average perceived distortion of 13.22% considering all comparisons. Only in the case of the mushroom did another comparison achieve better results, which was the one between Polycam models, the reference model being that from the Android version. However, this was the model that obtained the least realistic results due to the material it is made of. Again, the object that achieved the best results in both applications was the burner. It is interesting to note that the difference in perceived distortion between medium-poly and low-poly results is on average 3.4%, which indicates really good quality for low-poly models as well. This metric also shows the overall low capability of LiDAR sensors to provide good visual quality results when scanning moderate-sized objects.

Having analysed the obtained data, we can conclude that the use of LiDAR sensors with the Polycam application should not be the preferred option for obtaining realistic 3D models that can be showcased in VR shopping environments. The various metrics employed have yielded the worst visual quality level results for these models, both in comparisons using a reference model and those without. Therefore, for obtaining a realistic 3D model, the preferred option would be NeRF, specifically in the form of a mobile application like Luma AI, when available on a device. The data from the metrics used, both with and without a reference model, show that the models from Luma AI have the highest visual quality, except in NR-3DQA in the case of sports shoes. Taking into account high-poly models, values between 5 and 20% higher were obtained compared to the photogrammetry application used. Considering low-poly models, this range is reduced to between 2 and 15%. Given the visual results of the photogrammetry application, and considering that the difference is not very large in the results, this option is also viable for obtaining high visual quality 3D models that can be used in VR.

Table 4. 3D mesh comparisons performed with CloudCompare and CMDM. Regarding the labelling of types compared: HP, LP, MP and OP stands for the same as the previous table, Poly<device> = Polycam using such <device>, Polycam = both models obtained from both devices are compared.

Object	Type	Types Compared	Scanning Application	RMS	CMDM
Munich shoes	MP	HP, MP	LUMA	0.00041	**0.132868**
	LP	HP, LP	LUMA	0.00171	0.190245
	OP	Polycam	Polycam	0.01269	0.152718
	OP	HP, PolyXiaomi	Polycam	0.03012	0.223199
	OP	HP, PolyiPad	Polycam	0.03767	0.224947
Octopus teddy	MP	HP, MP	LUMA	0.00023	**0.151651**
	LP	HP, LP	LUMA	0.00098	0.17674
	OP	Polycam	Polycam	0.03360	0.204733
	OP	HP, PolyXiaomi	Polycam	0.00522	0.222374
	OP	HP, PolyiPad	Polycam	0.05976	0.318699
Burner	MP	HP, MP	LUMA	0.00009	**0.0480725**
	LP	HP, LP	LUMA	0.00040	0.0637373
	OP	Polycam	Polycam	0.01451	0.12820
	OP	HP, PolyXiaomi	Polycam	0.01568	0.121982
	OP	HP, PolyiPad	Polycam	0.01423	0.151532
Ceramic mushroom	MP	HP, MP	LUMA	0.00018	0.239902
	LP	HP, LP	LUMA	0.00072	0.290423
	OP	Polycam	Polycam	0.04873	**0.181069**
	OP	HP, PolyXiaomi	Polycam	0.03710	0.34902
	OP	HP, PolyiPad	Polycam	0.02446	0.350014
Trophy	MP	HP, MP	LUMA	0.00027	**0.0902698**
	LP	HP, LP	LUMA	0.00141	0.1052886
	OP	Polycam	Polycam	0.08235	0.239079
	OP	HP, PolyXiaomi	Polycam	0.03008	0.29231
	OP	HP, PolyiPad	Polycam	0.23793	0.390321

Discussion

Our evaluation revealed that the difference in visual quality between Luma AI models and those obtained through photogrammetry was generally minor, ranging from 10% to 20%. Specifically, the Root Mean Square (RMS) values for the medium-poly models of Luma AI were as low as 0.00041, while Polycam's photogrammetry models showed slightly higher RMS values of approximately 0.01269 for the optimised models. However, the difference with models obtained using LiDAR sensors was significantly greater, with some cases showing RMS differences up to 70% higher, such as the RMS value of 0.05976 for low-poly models.

This substantial disparity leads us to consider the models produced by Luma AI as having the highest visual quality among the technologies we tested. As such, we recommend Luma AI for product digitisation when the highest visual fidelity is required. Additionally, Luma AI provides comprehensive support for NeRF technology, which is critical for achieving high-quality 3D models.

It is crucial to remember that we are focusing on technologies that can be easily used by individuals without advanced technical knowledge, such as small business owners, who may not have access to expensive hardware and software resources. Therefore, it is important to find solutions that offer a good balance between quality and accessibility. Polycam stands out in this regard due to its user-friendly interface and extensive usage in the literature. It supports both photogrammetry and LiDAR scanning technologies, which are crucial for our comparative analysis. Polycam's average NR-3DQA score for optimised models was 0.8111, indicating good quality while maintaining accessibility for non-experts.

Conversely, our results indicate that models obtained using LiDAR sensors are less suitable for 3D object scanning in the context of product digitisation. The lower visual quality and higher error rates observed with LiDAR make it less ideal for creating detailed product models. For example, the CMDM values for LiDAR models reached up to 0.318699, significantly higher than those of photogrammetry models. However, LiDAR can be more

appropriate for scanning larger environments, such as entire stores or shopping spaces, where the focus is on capturing the overall layout rather than fine details.

Further limitations of these scanning technologies were detected. Products made with transparent material or specular surfaces require a specific scanning environment where light sources do not produce poor results. Moreover, objects with very small dimensions are quite difficult to scan with the technologies proposed. Instead, it would require a professional camera where the optical zoom allows for quality photos, as well as the use of photogrammetry software to transfer the photos to.

In Section 4, we will analyze the performance of these models on various current market VR headsets, including the Meta Quest 2, Quest Pro and Quest 3. This analysis aims to determine the practical feasibility of using these models in VR environments and to provide guidelines for their optimal use in different VR shopping scenarios.

4. Performance Analysis of the Generated Models in Virtual Environments

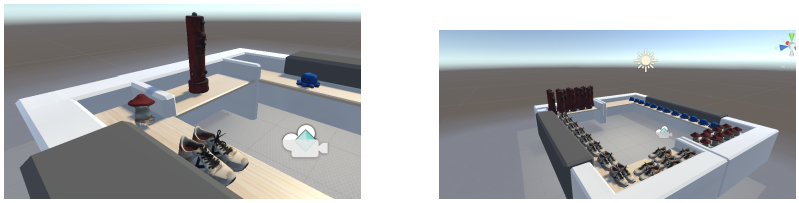
After comparing and evaluating the 3D models generated by the various scanning tools based on different characteristics, we proceeded to measure their performance on Meta's VR headsets: Quest 2, Quest Pro, Quest 3. To this end, we imported the .obj files of the models into a Unity project, where we configured scenes that were subsequently built and deployed on the headset for native execution, aiming to collect performance metrics without interference from the hardware of the connected PC. The chipset, GPU and RAM memory of the Meta Quest devices used are as follows:

- Meta Quest 2. Chipset: Qualcomm Snapdragon XR2, 8 cores (1 × 2.84 GHz, 3 × 2.42 GHz, 4 × 1.8 GHz), integrated GPU: Qualcomm Adreno 650 @ 587 MHz, 1.267 GFLOPS. RAM: 6 GB.
- Meta Quest Pro. Chipset: Qualcomm Snapdragon XR2+, 8 cores (1 × 2.84 GHz, 3 × 2.42 GHz, 4 × 1.8 GHz), integrated GPU: Qualcomm Adreno 650 @ 587 MHz, 1.267 GFLOPS. RAM: 12 GB.
- Meta Quest 3. Chipset: Qualcomm Snapdragon XR2 Gen 2, 6 cores (1 × 3.19 GHz, 2 × 2.8 GHz, 3 × 2.0 GHz) GHz, integrated GPU: Qualcomm Adreno 740 @ 492 MHz, 3.1 GFLOPS. RAM: 8 GB.

While it would have been preferable to include a more comprehensive range of VR headset models in the study, financial constraints within the research project necessitated a selective approach. The three Meta models were selected for their potential to facilitate an examination of the progression among models that share a direct lineage. Furthermore, the Meta Pro model is equipped with internal cameras, which will facilitate the integration of eye tracking in future research. The Meta Quest 3 model supports switching between virtual reality and augmented reality, which is also promising for subsequent investigations. Two further reasons influenced the decision to procure and utilise the Meta models in the study. Firstly, they are the best-selling models currently, as confirmed by data from sources such as Counterpoint (https://www.counterpointresearch.com/insights/global-xr-ar-vr-headsets-market-share/), Statista (https://www.statista.com/topics/10052/vr-headsets/) and Amazon (accessed on 7 July 2024). Secondly, they exhibit an excellent balance between quality/features (as outlined above) and cost, making them accessible to future VR consumers across various economic backgrounds.

On the other hand, it is worth noting that the design of the scenes is very minimalist, with a display stand for positioning the products. We chose this design to avoid overloading the hardware of the headsets by rendering objects in the scene that are not part of the study, as this could introduce noise in the measurements that would not faithfully represent the performance offered by the models. An example of a scene can be seen in Figure 8. We are aware of the limitations of not considering dynamic optimisation techniques in the analysis, but our objective is to see the performance provided by the models that small businesses will provide to the platform in the future, without considering future performance improvements that may be included in the development of the application. Moreover, even though we will not analyse the effects of lightning sources or materials

in the rendering process, we have shown in Section 3.3 the features of the textures of the 3D models used in the study. Next, the process of collecting performance data and the decisions made regarding the objectives of this work will be described.



(**a**) (**b**)

Figure 8. Two of the scenes developed for analysing the performance of models. (**a**) Simple scene (1 object of each model) of low-poly Luma AI models; (**b**) 12x scene (12 objects of each model) of low-poly Luma AI models.

4.1. Data Collection Method

Following the construction and deployment of each scene, usage metrics were collected using the native profiling tool OVR Metrics Tool 1.6.5 (https://developer.oculus.com/documentation/native/android/ts-ovrstats/ accessed on 7 July 2024). As the primary objective of analysing the performance offered by the models is their integration into a VR shopping environment, we made the following decisions when developing the scenes:

- Employ the official Meta SDK for Unity to incorporate into each model the necessary components to interact with the models. Thus, these can be picked up with virtual hands and moved freely to observe how a typical interaction with objects affects performance.
- Replicate the number of models in different scenes to estimate the limit of products that could be included in a virtual environment for each type of model generated by Luma AI. This allows estimating a maximum number of objects that can be displayed for the application to run smoothly and for the user experience to be satisfactory. This was performed with the low-poly models of Luma AI and Polycam for Android as they have been the most realistic models and the ones that have provided the best data, as seen in Section 3.3. In total, 14 scenes were developed to achieve this objective.
- For the collection of performance data, the corresponding application for each scene was run, and during an application execution lasting approximately 60 s, there was interaction with the objects in the environment, as well as free movement around the scene. This was done with the aim of simulating a user's behaviour in a VR shopping environment.

OVR Metrics exports data from each application execution to a .csv file, recording the value of up to 88 metrics every second regarding the CPU, GPU, memory, temperature, etc. We have selected for analysis the metrics we are interested in with respect to performance, as metrics such as voltage, wattage, or colour intensity in each eye of the headset are also collected but are not of particular interest to us.

In the next section, we provide a performance analysis of the models from the execution of the scenes developed. The main constraint to further replicating objects of one type was the smoothness of the application, which is a key factor in the user experience. With the recommended rate of frames per second being between 60 and 72 for the Meta Quest devices used, a rate below half that (in this case around 30 frames per second) starts to severely affect the user experience, causing lag that, in the case of VR, can cause motion sickness. We started with 4 objects to represent a minimum number of objects that could appear simultaneously in a real shopping environment, one of each type that was scanned. The replications, as shown below, allowed us to reach a maximum of 96 low-poly, 24 medium-poly and 12 high-poly objects, figures at which the user experience was affected quite negatively.

4.2. Performance Analysis of Models

The analysis of the data collected in the different .csv files has been conducted in a Python notebook that can be visited and executed in a Google Colab environment which can be accessed in the public repository of the paper. In the rest of this section, we will present the analysis of the metrics that we have considered relevant to understanding the performance of a VR shopping application using the obtained models. In Table 5, we can observe some values collected from metrics that have helped us to carry out the analysis, and that we will examine in more detail below. Please note that, during the analysis, we will refer to quantities of objects as a reference. To avoid repetition, bear in mind that this expression refers to objects with graphic characteristics (number of polygons, vertices, size, quantity of textures, etc.) similar to those used in our study. For quick reference, Table 5 includes a column to represent the polygonal load of each scene, while the size of the textures can be found in Table 1.

Table 5. Metrics measured with the tool OVR Metrics. Scene names mean the following: (a) Simple: 4 objects, one of each model, (b) <Number>x: <Number>*4 objects, <Number> of each model, (c) LowPoly/MediumPoly/HighPoly: the type of the models used in the scene. In the second column, LP = Low Poly, MP = Medium Poly, HP = High Poly.

Scene	N° of Objects (N° Polygons)	Average FPS	Stale Frames %	App GPU Time %	GPU U $\geq$ 99%	Average Prediction %
			Meta Quest 2			
SimpleLowPoly	4 LP (99,806)	71.37	6.35	0.00	0.00	0.00
6xLowPoly	24 LP (597,648)	71.16	10.20	0.00	0.00	0.00
12xLowPoly	48 LP (1,195,296)	71.74	15.15	0.00	6.06	0.00
18xLowPoly	72 LP (1,792,944)	59.47	62.90	30.71	32.26	3.22
24xLowPoly	96 LP (2,390,592)	56.21	64.52	48.38	32.26	11.29
SimpleMediumPoly	4 MP (599,600)	69.22	16.00	0.00	0.00	0.00
3xMediumPoly	12 MP (1,798,800)	48.81	80.00	57.33	17.33	29.33
6xMediumPoly	24 MP (3,597,600)	29.61	45.00	61.67	56.66	8.33
SimpleHighPoly	4 HP (3,527,888)	9.30	10.0	30.00	93.33	78.33
3xHighPoly	12 HP (10,583,644)	8.75	1.66	33.33	93.33	56.66
			Meta Quest Pro			
SimpleLowPoly	4 LP (99,806)	71.63	11.66	0.00	0.00	0.00
6xLowPoly	24 LP (597,648)	71.56	16.66	0.00	0.00	0.00
12xLowPoly	48 LP (1,195,296)	69.28	43.33	5.00	11.66	1.67
18xLowPoly	72 LP (1,792,944)	60.16	65.00	35.00	13.33	5.00
24xLowPoly	96 LP (2,390,592)	44.56	78.33	71.67	30.00	6.67
SimpleMediumPoly	4 MP (599,600)	71.48	20.00	0.00	3.33	0.00
3xMediumPoly	12 MP (1,798,800)	56.94	88.33	63.33	20.00	30.00
6xMediumPoly	24 MP (3,597,600)	30.62	35.00	33.33	61.67	26.66
SimpleHighPoly	4 HP (3,527,888)	12.40	5.00	18.33	93.33	48.33
3xHighPoly	12 HP (10,583,644)	8.28	80.00	53.33	91.66	90.00
			Meta Quest 3			
SimpleLowPoly	4 LP (99,806)	71.52	4.91	0.00	0.00	0.00
6xLowPoly	24 LP (597,648)	71.95	6.55	0.00	0.00	0.00
12xLowPoly	48 LP (1,195,296)	71.96	13.11	3.28	0.00	0.00
18xLowPoly	72 LP (1,792,944)	62.60	45.90	21.31	0.00	11.47
24xLowPoly	96 LP (2,390,592)	55.52	50.82	31.15	4.92	21.31
SimpleMediumPol y	4 MP (599,600)	71.85	16.66	0.00	0.00	0.00
3xMediumPoly	12 MP (1,798,800)	52.43	84.51	67.60	30.03	19.71
6xMediumPoly	24 MP (3,597,600)	48.08	27.87	45.90	34.43	14.75
SimpleHighPoly	4 HP (3,527,888)	15.97	13.11	83.60	81.97	40.98
3xHighPoly	12 HP (10,583,644)	10.95	4.92	91.80	91.80	57.37

4.2.1. Frame Rate

This metric, commonly measured in Frames Per Second (FPS), critically affects the user experience in an application, as a low frame rate results in the application feeling slower and less fluid, showing much greater latency in response to user actions. Additionally, a low frame rate can also indicate that the workload is not suitable for the hardware on which it is running. In our case, the most critical components are the CPU and GPU of the headset responsible for rendering the 3D models.

It is important to note that the frame rate should be close to the refresh rate (60–72 Hz) for the VR application experience to be pleasant and smooth. In Figure 9, we can observe the frame

Appl. Sci. **2024**, *14*, 6037

rate offered by different types of objects in the scenes executed for each of the Meta VR headsets used. As can be seen, all devices offer performance close to the 72 Hz refresh rate in the form of FPS for LP objects up to a quantity of 48 objects. Beyond this number, the frame rate starts to fall to 60 FPS, which is still smooth enough to offer a pleasant user experience. However, in the last scene with 96 LP objects, there is a severe drop in the frame rate, especially in the Quest 2 and Quest Pro, due to complexity spikes when one or two objects were manipulated while a large number of them were being rendered.

Regarding the MP models, all devices struggle to varying degrees when handling 12 objects, with complexity spikes affecting the average FPS, as can be observed. This is more evident in the next scene with 24 MP objects, although the Quest 3 handles these types of spikes better due to its more powerful CPU and GPU hardware.

Looking at the data obtained for the scenes with HP objects, neither offers a reasonable frame rate to provide a good user experience. A considerable detriment in frame rate can be observed for a similar quantity of polygons compared to the scene with 24 MP objects, whose quantity of polygons is even higher. Focusing on the scenes with LP objects, specifically the one with 72 objects, executed on the Quest 3, we obtain a frame rate 3.91 times higher despite the scene with 4 HP objects having only almost twice the number of polygons (1.97 times more). These data highlight the effect that other factors such as the number and resolution of the textures generated for HP and MP models have on performance. As previously mentioned, the number of textures generated by Luma AI for these types of model is much higher compared to LP.

In terms of fluidity, we can observe that LP models would be the most suitable for inclusion in a VR shopping space that aims to house a considerable number of models in the same scene, trying not to exceed 72 objects with characteristics similar to those used in our study to avoid complexity spikes that negatively impact the frame rate. On the other hand, it is not advisable to use HP models from Luma AI due to the low frame rate they offer, which could only be resolved if executed with VR headsets connected to a PC with powerful hardware to help render the scene, thereby limiting the user's shopping experience. Finally, the use of MP objects could be an alternative in scenes that do not require a large number of simultaneous objects, such as simple shopping spaces with a single virtual counter. However, this would mean losing flexibility in setting up the scene, as computing power would be used in a greater number of polygons and textures which, according to the results of Section 3.3, offer a visual quality improvement of less than 10% compared to LP.

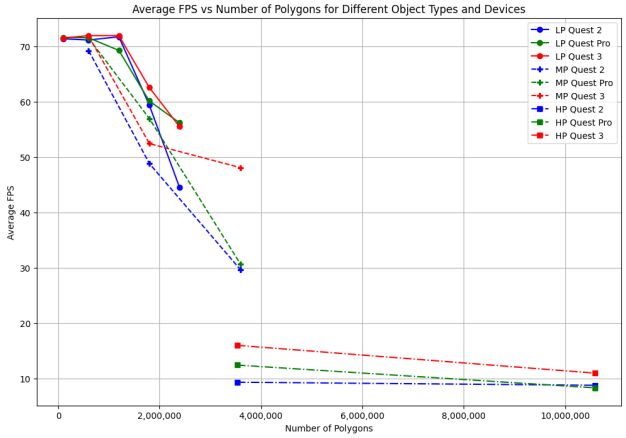


Figure 9. Average FPS obtained for each device and type of object. The number of polygons is the total of those for the objects loaded in each scene. LP = Low-Poly objects, MP = Medium-Poly objects, HP = High-Poly objects.

4.2.2. Stale Frames

Stale frames indicate the number of times a frame could not be delivered on time due to the parallel processing of the GPU and CPU, resulting in the reuse of the previous frame. Stale frames can cause problems when the value recorded is greater than 0 but less than the refresh rate, i.e., 72. The issues arising from this may include increased latency between user actions and their visual representation in the virtual environment or even motion sickness due to the perceived lag between these actions and their visual feedback. Table 5 shows the percentage of records where values less than 72 but greater than 0 were collected.

Although it can be observed that even in scenes with a frame rate close to the refresh rate there are percentages close to or above 50% of stale frames, this does not necessarily signify a problem for the user experience. According to Meta's own documentation, extra latency mode is enabled by default in applications built with Unity. This mode establishes communication with the Asynchronous TimeWarp (ATW) system (a technology that helps reproject frames based on the most recent head tracking information) to always wait for an extra frame. This could explain why the fluidity of the application is not affected. On the other hand, the low percentage of stale frames in scenes with HP objects indicates better synchronisation between generating a new frame and the refresh cycle. However, the low rate of generation of these new frames (frame rate) makes it more difficult for them to become stale.

4.2.3. GPU U (GPU Utilization)

A GPU utilisation percentage equal to or greater than 99% can be one of the reasons indicating that the application is GPU-limited, with performance issues potentially starting to appear from 90% utilisation. Table 5 shows the percentage of application's execution time where the GPU U percentage exceeded or equalled 99%.

In scenes with LP objects, up to 24 objects, the application does not reach 99% utilisation at any point, just like scenes with 4 MP objects, except in the Quest Pro in this last case. From 48 LP objects onwards, the percentage of time varies between the Quest 2 and Pro, as the Quest 3 does not maximise GPU usage until the scene with 96 LP objects, for which it only does so 5% of the total execution time. While in the Quest 2, nearly a third of the execution time of the two scenes with the most LP objects maximised GPU usage, in the Quest Pro, it only happened in the scene with the highest number of these objects (30% of the time).

For scenes with 12 or more MP objects and those with HP objects, there is an upward trend. While in scenes with 12 MP objects, the percentage remains between 20 and 30% in all headsets; again, the difference in power of the Quest 3 is noticeable in the scene with 24 MP objects, where the Quest 2 and Pro reached nearly 60% of the time with the GPU at maximum use, against 34.43% in the Quest 3. However, all headsets remained near 90% of the time with the GPU at maximum use in all HP scenes, demonstrating how demanding these models are in terms of power when running applications as standalone VR.

With these results, we can observe once again that it is advisable to use LP objects up to a reasonable amount based on the number of polygons or a few MP objects. Thus, the application will avoid maximising GPU usage for an extended period. This can lead to overheating in the headset, causing discomfort or loss of stability and fluidity in the application when run for a longer period of time.

4.2.4. Application GPU Time (App T)

This metric helps determine if an application is potentially GPU-bound, as it collects the amount of time the application takes to render a single frame. If the time measured in the metric is longer than the length of a frame (13.88 ms for 72 frames per second), the application can be considered GPU-bound. Meta's documentation provides a simple rule to indicate when an application is limited by the CPU or GPU. If App T > 13.88 ms and FPS < 72, then there is a 100% bottleneck caused by the GPU (it could also be happening in the CPU but is masked by the GPU), if App T < 13.88 and FPS < 72 then there is a bottleneck

Appl. Sci. **2024**, *14*, 6037

caused by the CPU. This rule was used to develop the charts shown in Figure 10, showing the frames where bottlenecks occur or not, and of what type (CPU or GPU) if they occur, for the data obtained in the Meta Quest 3.

As can be seen in Table 5, for scenes with LP models, in the Quest 2 and Pro, they begin to need more time than necessary to render a frame smoothly, starting from scenes with 48 objects, albeit with low percentages. However, in the scene with 96 objects, they obtained percentages of 48 and 71% in the Quest 2 and Pro, respectively, while in the Quest 3, it was 31%. These are moderately high percentages, since, taking the best performance obtained in the Quest 3 as a case, nearly a third of the total execution time, the application could have been limited by the GPU.

Considering scenes with MP objects, the polygonal load begins to affect more severely, reducing the GPU's working capacity. In any of the three devices, at least one of the scenes reached a percentage close to 60%, representing a high percentage of the time in which the GPU was a bottleneck due to the work needed to render. Finally, we can see how, in scenes with HP objects, the Quest 3 presents an extremely high percentage, while the other two headsets present much lower percentages, even lower than in the scenes with MP models. This could be due to the main cause of poor performance in these headsets being a bottleneck in the CPU due to the low percentages of GPU usage time by the application.

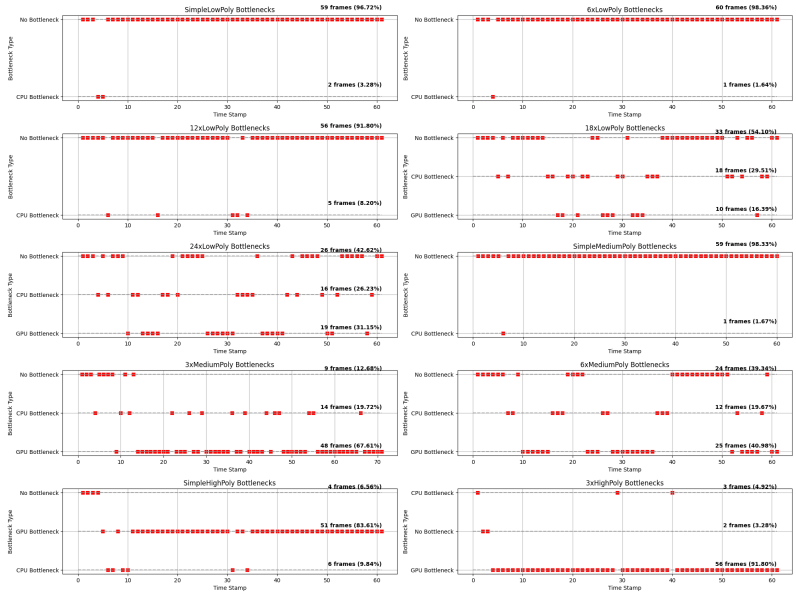


Figure 10. Meta Quest 3 bottleneck identification based on frame rate and App T.

4.2.5. CPU Utilization Percentage

Similarly, the tool records the CPU utilization percentage of the headset, represented by the worst-performing core. Figure 11 shows the average CPU usage percentage, taking into account that it captures the core with the worst performance in each measurement.

Appl. Sci. **2024**, *14*, 6037

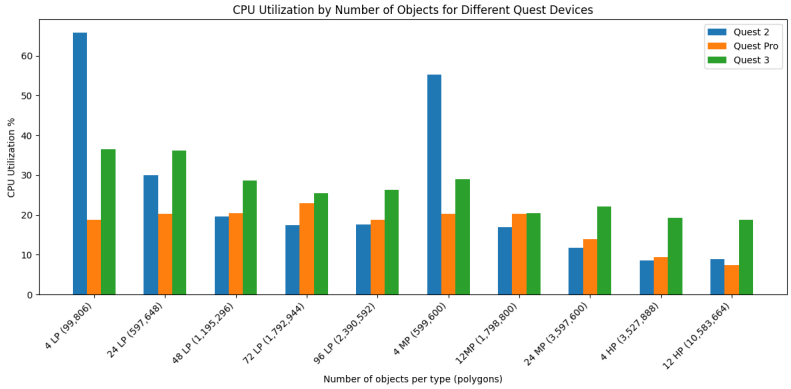


Figure 11. CPU utilization % obtained of each scene run on each VR Headset. This metric represents the worst performing core. Therefore, for multi-threaded applications, the main thread of the app may not be represented in this metric.

In the case of Quest 2 and 3, there is a clear trend of a higher percentage of CPU usage in scenes with low polygonal load, which decreases as the polygonal load of the scenes increases. In the case of the Quest Pro, the percentages remain close, between 20 and 30% usage, except in the three scenes with the highest polygonal load of all. The reason for the downward trend in CPU usage in devices may be due to the bottlenecks generated by the GPU in more demanding scenes during rendering, so the CPU waits for the GPU to finish rendering before starting to process the next frame. Conversely, in the case of scenes with lower polygonal load, the CPU is more used as it needs to prepare and send a greater number of frames to the GPU, meeting the frame rate obtained in these scenes as shown in Section 4.2.1. Finally, we highlight the difference in CPU usage in the Quest 2 compared to the other two more recent Meta devices. While in scenes with 4 LP and MP objects its usage is considerably higher, it is not concerning since these scenes feel fluid. However, it can be observed, especially with the data from the Quest 3, that the processor of the latter is used approximately 10% more for the rest of the scenes, as the power of its CPU allows it to try to reach a higher frame rate.

4.2.6. GPU and CPU Levels

The levels of the GPU and CPU clocks refer to the frequency in MHz. The clock level can be from 1 to 5 and, although they can be set manually, it is recommended to use dynamic clock throttling by default. Thus, these levels automatically increase if the application does not run at an adequate frame rate at the current levels. Figure 12 shows graphs of the evolution of GPU and CPU levels throughout the application execution in each scene.

At the device level, we observe how the Quest 3 stands out for reaching level 4 or even 5 more frequently in scenes with a higher graphic workload. These graphs are interesting as they show how, for example, in scenes with HP objects, it is demonstrated that the Quest 3 was limited by the GPU (see Section 4.2.4), as the GPU level was 5 while the CPU level remained at 4. The opposite occurs with the Quest 2, where the levels remained the same. In general, until the scene with 48 LP objects, no device reached level 5, with it being in this scene where the Quest 3 increased the GPU level to this number during specific periods of time. We also highlight in general that the Quest 3 and Pro increase the GPU level more frequently. This indicates that the default dynamic adjustment increases the clock frequency to try to process the rendering more quickly, especially in scenes with higher polygonal load, where we can observe complexity spikes in the graphs. Meanwhile, the Quest 2 has maintained the same level for both CPU and GPU in most application executions.

Appl. Sci. **2024**, *14*, 6037

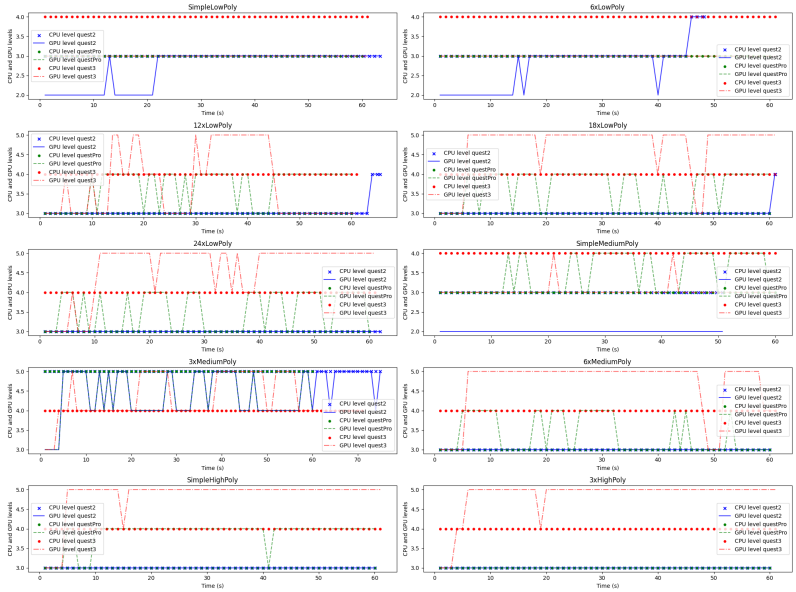


Figure 12. CPU and GPU levels in each scene run in each VR headset.

In general, we can observe how the Quest 3 and Pro adapted the use of the GPU dynamically better. The reason for this may be due to the difference in hardware power between these headsets and the Quest 2. In any case, it is shown once again that, from a certain number of polygons, as a guide, the different devices may start to experience problems due to the continuous increase in the clock frequency, either of the CPU or the GPU.

4.2.7. Average Prediction

This metric collects "the absolute time between when an app queries the pose before rendering and the time the frame is displayed on the HMD screen". That is, this metric measures the delay between the hand tracking data collected by the headset and its rendering in the corresponding frame. This value should be a fixed number between 45 and 50 ms; however, if the value is much higher than this, the application may suffer latency problems when rendering. Since Meta's documentation does not provide a reference value that quantifies this "much higher", we have opted to check those records of the .csv files with a value greater than 60 ms (20% higher than the upper limit of the documented range). Table 5 shows the percentage of gaming time with this condition for this metric.

In all three devices, the latency recorded by this metric remains below 12% up to the scenes with 72 LP objects, while in the Quest 3 this percentage increases up to 21% in the scene with 96 LP objects. This higher value may indicate that, in that scene, there was more interaction and more frequent modification of the hand poses than in the same scene executed on the other two devices, as the application appeared more fluid and encouraged more interaction, as can be seen in the obtained frame rate (see Section 4.2.1). Thus, with more interaction these delays may occur more frequently in a scene that does not run with complete fluidity. In scenes with a higher number of MP and HP models, we can observe disparate percentages depending on the devices, from 8% to 90%. Although it is true that in HP scenes the percentages are intrinsically high because, what little the headset could track and render, it did so with a very high delay.

4.3. Discussion

Analysing the data with the OVR Metrics Tool, we have determined that, for 3D models with characteristics similar to ours (polygons, texture size, etc.), an acceptable limit could be 72 objects (1,792,944 polygons) of low-poly models from Luma AI. Beyond this limit, the frame rate falls below 60 FPS and presents worrying complexity spikes, which severely impact the fluidity and performance of all the VR headsets used. For example, in scenes with 72 low-poly objects, the frame rate dropped to an average of 56.21 FPS on the Meta Quest 2, 60.16 FPS on the Meta Quest Pro and 62.60 FPS on the Meta Quest 3, indicating that performance starts to degrade significantly beyond this point.

In the case of medium-poly models, reasonable results were obtained with up to 4 objects (599,600 polygons). When the number of medium-poly objects increased to 12, the total polygon count was similar to that of 72 low-poly objects, but the performance suffered more severely. Specifically, the average FPS dropped to 29.61 on the Meta Quest 2, 30.62 on the Meta Quest Pro and 48.08 on the Meta Quest 3, making these configurations less feasible for a smooth VR experience.

Lastly, the performance offered by the high-poly models was very low, so their use is discouraged. For instance, high-poly scenes resulted in frame rates as low as 9.30 FPS on the Meta Quest 2, 12.40 FPS on the Meta Quest Pro and 15.97 FPS on the Meta Quest 3. These results clearly indicate that high-poly models are not suitable for standalone VR shopping environments due to their detrimental impact on performance.

Conversely, low-poly models are deemed appropriate depending on the hardware currently available in VR headsets, as they allow greater scalability by adding more objects without severely impacting performance up to a limit of 72 objects. These could be suitable for large VR shopping spaces that simulate environments such as supermarkets or shopping centres. While the use of medium-poly models is very limited, they could be used in simpler spaces with fewer objects. Despite the visual quality of medium-poly models being slightly higher (approximately 5% improvement), the performance trade-off makes them less viable for extensive use.

Finally, we discuss some limitations of these 3D models for their use in VR shopping environments. The tests performed completely isolated the 3D models; therefore, other aspects such as the mesh complexity of the virtual environment are not taken into account. It would be interesting to further research the effects of other optimisation techniques such as Levels of Detail (LoD) considering both the environment and the products. In order to obtain a better performance, an optimisation of the 3D model can be made in 3D software tools like Blender. Even though it could be beneficial, it would require people with technical expertise in the field in order to obtain improved results, which might not be possible in the circumstances of small businesses.

5. 3D Model Optimisation of High-Poly Models for VR

The main focus of our work has been to analyse various low-cost scanning technologies and mobile applications that implement them, enabling small business owners to use them. This analysis aimed to examine whether the obtained 3D models offered good visual quality and whether their performance was suitable for various VR shopping use cases. Nevertheless, the low-poly models offered by both Luma and Polycam, which are around 20,000 polygons, may seem excessive for a mesh's complexity to be labelled low poly. Additionally, for medium- and high-poly models, the number of polygons increases exponentially. Aware of the existence of performance optimisation techniques at both the model and application levels, which can be included in a future production version of the VR-ZOCO platform, we plan to explore the impact of optimising high-poly models in terms of visual and performance levels.

This process requires a series of technological knowledge not commonly possessed by small business owners. Therefore, we must adopt an approach that allows this process to be practically automated, taking the 3D model generated by a small merchant from one of the analysed applications as input and returning a simplified model. Following this

Appl. Sci. **2024**, *14*, 6037

approach, we chose Blender because it offers an API with which to create and execute Python scripts, allowing us to automate actions performed in the software's GUI.

5.1. Optimisation Process

Next, the step-by-step optimisation process carried out with the tools that *Blender* offers is described:

1. Apply a 3D mesh modifier with the "Decimate" tool [38], a technique that reduces the number of polygons by a specified percentage, preserving details while reducing the mesh geometry.
2. Unwrap the 3D model imported to Blender 3.5 in order to create a 2D UV map that fits the mesh of the model with the Smart UV project tool, with an island margin of 0.02 to avoid overlapping.
3. The "Shrinkwrap" tool was used to adjust the mesh of the optimised low-polygon model more closely to the shape of the high-poly model.
4. Bake both a normal map and diffuse textures. While the first allows for the transfer of the surface details of the high-poly model without having to add more polygons, the second bake allows for the transfer of the colour details from the different texture files of the model (in the case of high-poly models, they had more than 50 files) into a single file.

Following this process, we have managed to optimise, not only in terms of polygons but also at the texture level for high-poly models, where we now have only 2 files: a normal map and a texture file. Figure 13 shows the result of the optimisation process both in terms of mesh and when rendering the model.



Figure 13. High-poly (**left**) and optimised (**right**) models from the octopus teddy, showing the differences in both meshes and rendering.

5.2. Quality and Performance Evaluation of Optimised Models

Briefly, we will analyse both the graphic quality and performance of the models obtained through the optimisation process. Figure 14 presents five bar charts comparing the visual quality of the optimised models to the original high-poly models from which they were derived. In the first subgraph, starting at the top left, we can see that the percentage difference in the number of polygons is at least 99%, meaning that the number of polygons is almost completely reduced compared to the high-poly models. In the second subgraph, we observe differences in percentages, as in some cases the textures of the high-poly models were smaller in total size than the two files generated for the optimised model. This is largely due to the size of the normal maps. Thus, we see that in models with more complex geometry, such as the sport shoes, the texture size can be

Appl. Sci. **2024**, *14*, 6037

considerably reduced in the optimised model, the opposite being case for objects with simpler meshes. In both the third and fourth subgraphs, which present the NR-3DQA and MM-PCQA metrics used in Section 3.3, we observe that, except in one case, there is some loss of visual quality in the optimised models. However, the losses in NR-3DQA do not exceed 15.5%, while in MM-PCQA two of them exceeded 25%. The case of the octopus teddy is surprising, which obtained a 30% loss in the metric value, but visually there is no significant difference (see Figure 13). Finally, the fifth subgraph compares the results of the CMDM metric, comparing the low-poly and optimised models with the corresponding high-poly models as the reference model. We recall that a lower CMDM value indicates less perceived distortion in the compared mesh relative to the reference. Thus, we can observe that in three of the five objects the low-poly models obtained better values in the comparison than the optimised ones, while in the other two objects (octopus teddy and ceramic mushroom) the comparison was more favourable to the optimised models with values 7.26% and 14.22% lower, respectively.

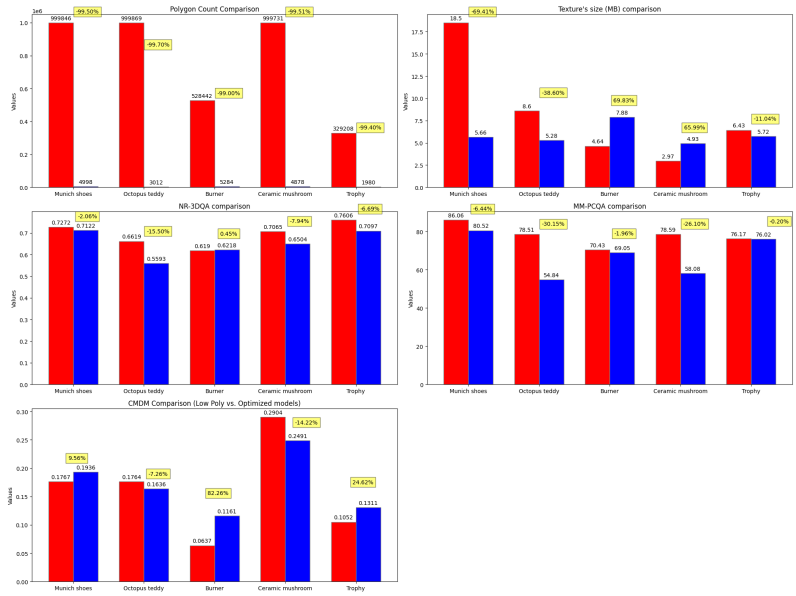


Figure 14. Visual quality comparison between high-poly models (red bars) and optimised models from high poly (blue bars). In the fifth subgraph, the red bar represents data obtained from low-poly models. The yellow text box contains the percentage difference between high-poly and optimised models.

On the other hand, Table 6 shows the performance data obtained when running a scene with 96 optimised objects in the three VR headsets used in the study. We have only used this scene to offer a perspective on the performance offered by these models with low polygonal load compared to the models with the lowest number of polygons in the data shown in Section 4.2. As can be seen, on any of the devices the frame rate is very close to the refresh rate, the GPU and CPU do not generate bottlenecks due to excessive use at any time and the percentage of stale frames and average prediction are not a cause for concern. Thus, it is evident how these optimised models can allow the inclusion of an even greater number of objects due to their low polygonal and texture load.

Table 6. Performance data obtained from running a 96 (483,648 polygons) optimised 3D models scene for 1 min.

Device	Average Frame Rate	Stale Frames %	App GPU Time %	GPU U >= 99%	CPU U %	Average Prediction %
Meta Quest 2	71.36	15.00	0.00	0.00	19.6	0.00
Meta Quest Pro	71.33	15.00	0.00	0.00	21.70	0.00
Meta Quest 3	71.65	11.66	0.00	0.00	26.15	0.00

6. Conclusions and Future Work

In this work, we have selected a series of scanning techniques (photogrammetry, LiDAR sensors and NeRF) to analyse, both in terms of visual quality and performance in current VR headsets on the market, the results of scanning a series of products using software applications on mobile devices (Polycam and Luma AI). This analysis was conducted with the goal of investigating and exploring low-cost 3D scanning tools, enabling small business owners to afford to digitise their product catalogues and include them in any VR shopping platform. In this way, we aim for small businesses with limited technological and financial resources to participate in the evolution of e-commerce in the coming years. At the same time, this would simplify the management of VR shopping platforms on the developer's side, as it would be the small businesses themselves who digitise their products and keep their catalogues updated, without having to depend on graphic designers to model the products.

For this reason, it is essential that the results offered by the tools recommended to small merchants provide good visual quality, encouraging VR shopping users to explore immersive shopping experiences. In this first part of the analysis, we first measured the inherent error of the models obtained with the selected applications for the study, as we did not have a professional 3D scanner to provide us with a ground truth model. For this, we modelled six pieces of expanded polystyrene in Blender, with their real measurements, and used them as a reference to compare the 3D models obtained after scanning these pieces with the applications. The results showed that the high-poly models introduced the least error. After this, we evaluated the visual quality of five products with different materials and geometries, using metrics that do not require a reference model such as NR-3DQA [32] or MM-PCQA [34] and metrics that do require one such as CMDM [35]. The results obtained positioned the models from Luma AI (low poly, medium poly, and high poly) as those with the best visual quality, followed by Polycam using photogrammetry and Polycam using LiDAR sensors.

Following this first part, we began the second analysis, focused on the performance provided by the models offering the highest graphic quality (see Section 4), as we must ensure that, despite having good visual quality, they must provide a pleasant user experience when executed on VR headsets in standalone mode. Thus, we built scenes in Unity with the least possible amount of assets, logic, and other engine elements in order to isolate the obtained data as much as possible to the rendering of the Luma AI models in the Meta Quest 2, Pro and 3. According to the perceived fluidity, the number of objects of the corresponding type was multiplied to try to reach a theoretical maximum from which the application did not run fluidly. The data showed that the high-poly models from Luma AI should not be used for standalone VR shopping environments, despite being the most detailed and visually quality-rich. Instead, medium-poly models should be used if the shopping environment will host very few objects, or low-poly models with the possibility that the virtual space can accommodate a much larger number of objects (up to 18 times more for acceptable performance). As we saw in the data shown in Section 3.3, the difference in visual quality of the high-poly models compared to the low-poly ones is approximately 5%, so much better performance is provided at the cost of losing very little visual quality.

Finally, we explored the use of optimising high-poly models to convert them into ultra low-poly models, with a lower number of polygons than those offered by the low-poly models from Luma AI. The results, despite confirming that in terms of visual quality they

can lose up to 30% depending on the mesh, are very good in terms of performance. Thus, these types of models could be used to implement level of detail techniques, loading them when the user is far from them and, as they approach, loading the low-poly models from Luma AI, replacing these optimised models. This would allow even more objects to be added to the scene.

We are currently planning contacts with small businesses in our city that will allow us to carry out further research along the lines of the VR-ZOCO project. Some of them include a cost study when the platform starts operating, as shown in the timeline of one of our previous works [39]. Additionally, we highlight the exploration of a new technique for the generation of 3D models, Gaussian splattering, analysing its results similarly to what was done in this work. Additionally, we plan to organise experimental sessions in collaboration with local chambers of commerce. These sessions will be coordinated with real businesses and shops to more accurately assess the impact of object realism on consumer purchase intentions and preferences. This approach will not only yield more relevant and applicable data but will also foster a direct relationship with the commercial sector, allowing us to gain practical and firsthand feedback from market participants.

In addition, we are currently scanning small businesses in our city in order to further research the benefits and challenges of these technologies. It will also allow us to study the adoption of these technologies by these businesses, giving us insights on how easy it is for people without expertise to digitize their shops. Moreover, it will serve us well to gather feedback about the scanning technologies studied from real users, as well as how the digitization process affects the operation of the small businesses. Considering that it would require an as-empty-as-possible shop to protect the privacy of customers and employees, it could affect the normal businesses operations.

Author Contributions: R.G.: conceptualizacion, methodology, software, validation, formal analysis, investigation, visualization, writing; J.A.: methodology, validation, formal analysis, investigation, writing, supervision, funding acquisition, project administration; D.V.: methodology, validation, formal analysis; C.G.-M.: methodology, validation, formal analysis; J.J.C.-S.: methodology, validation, formal analysis, investigation, supervision, funding acquisition, project administration. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been funded by the Spanish Ministry of Science and Innovation MICIN/AEI/10.13039/501100000033, and the European Union (NextGenerationEU/PRTR), under the Research Project: Design and development of a platform based on VR shopping and AI for the digitalization and strengthening of local businesses and economies, TED2021-131082B-I00.

Data Availability Statement: All data generated during the study are publicly available in the GitHub repository: https://github.com/AIR-Research-Group-UCLM/PDIVR-ZOCO accessed on 7 July 2024. It contains links for downloading the 3D models obtained and used in the study, the data obtained for carrying out the performance analysis and more. The repository's structure is fully detailed in its README.md.

Acknowledgments: This work has been funded by the Spanish Ministry of Science and Innovation MICIN/AEI/10.13039/501100000033, and the European Union (NextGenerationEU/PRTR), under the Research Project: Design and development of a platform based on VR shopping and AI for the digitalization and strengthening of local businesses and economies, TED2021-131082B-I00. During the preparation of this work, the author(s) used ChatGPT for assistance in coding content in LaTeX and as a support tool for translation. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Chodak, G.; Ropuszyńska-Surma, E. Virtual Reality and Artificial Intelligence in e-Commerce. In *Advanced Computing*; Garg, D., Narayana, V.A., Suganthan, P.N., Anguera, J., Koppula, V.K., Gupta, S.K., Eds.; Springer: Cham, Switzerland, 2023; pp. 328–340. [CrossRef]
2. Fedorko, R.; Kráľ, Š.; Bačík, R. Artificial Intelligence in E-commerce: A Literature Review. In *Congress on Intelligent Systems*; Saraswat, M., Sharma, H., Balachandran, K., Kim, J.H., Bansal, J.C., Eds.; Springer: Singapore, 2022; pp. 677–689. [CrossRef]

3. Kang, H.J.; Shin, J.H.; Ponto, K. How 3D Virtual Reality Stores Can Shape Consumer Purchase Decisions: The Roles of Informativeness and Playfulness. *J. Interact. Mark.* **2020**, *49*, 70–85. [CrossRef]
4. Pizzi, G.; Scarpi, D.; Pichierri, M.; Vannucci, V. Virtual reality, real reactions?: Comparing consumers' perceptions and shopping orientation across physical and virtual-reality retail stores. *Comput. Hum. Behav.* **2019**, *96*, 1–12. [CrossRef]
5. Cortinas, M.; Berne, C.; Chocarro, R.; Nilssen, F.; Rubio, N. Editorial: The Impact of AI-Enabled Technologies in E-commerce and Omnichannel Retailing. *Front. Psychol.* **2021**, *12*, 718885. [CrossRef]
6. Miller, S.H.; Hashemian, A.; Gillihan, R.; Helms, E. A Comparison of Mobile Phone LiDAR Capture and Established Ground based 3D Scanning Methodologies. In Proceedings of the WCX SAE World Congress Experience, Detroit, MI, USA, 5–7 April 2022. [CrossRef]
7. Skrupskaya, Y.; Skibina, V.; Taratukhin, V.; Kozlova, E. The Use of Virtual Reality to Drive Innovations. VRE-IP Experiment. In *Information Systems and Design*; Taratukhin, V., Matveev, M., Becker, J., Kupriyanov, Y., Eds.; Springer: Cham, Switzerland, 2022; pp. 336–345.
8. Mildenhall, B.; Srinivasan, P.P.; Tancik, M.; Barron, J.T.; Ramamoorthi, R.; Ng, R. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *arXiv* **2020**, arXiv:2003.08934. [CrossRef]
9. Gao, K.; Gao, Y.; He, H.; Lu, D.; Xu, L.; Li, J. NeRF: Neural Radiance Field in 3D Vision, A Comprehensive Review. *arXiv* **2023**, arXiv:2210.00379. [CrossRef]
10. Deng, N.; He, Z.; Ye, J.; Duinkharjav, B.; Chakravarthula, P.; Yang, X.; Sun, Q. FoV-NeRF: Foveated Neural Radiance Fields for Virtual Reality. *IEEE Trans. Vis. Comput. Graph.* **2021**, *28*, 3854–3864. [CrossRef] [PubMed]
11. Dickson, A.; Shanks, J.; Ventura, J.; Knott, A.; Zollmann, S. *VRVideos: A Flexible Pipeline for Virtual Reality Video Creation*; University of Otago: Dunedin, New Zealand, 2022; pp. 199–202. [CrossRef]
12. Li, K.; Rolff, T.; Schmidt, S.; Bacher, R.; Frintrop, S.; Leemans, W.; Steinicke, F. Immersive Neural Graphics Primitives. *arXiv* **2022**, arXiv:2211.13494. [CrossRef]
13. Rolff, T.; Li, K.; Hertel, J.; Schmidt, S.; Frintrop, S.; Steinicke, F. Interactive VRS-NeRF: Lightning fast Neural Radiance Field Rendering for Virtual Reality. In Proceedings of the SUI '23: ACM Symposium on Spatial User Interaction, Sydney, Australia, 13–15 October 2023; pp. 1–3. [CrossRef]
14. Obradović, M.; Vasiljević, I.; Durić, I.; Kićanović, J.; Stojaković, V.; Obradović, R. Virtual Reality Models Based on Photogrammetric Surveys—A Case Study of the Iconostasis of the Serbian Orthodox Cathedral Church of Saint Nicholas in Sremski Karlovci (Serbia). *Appl. Sci.* **2020**, *10*, 2743. [CrossRef]
15. Andree León Tejada, R.; Alexander Jimenez Azabache, J.; Javier Berrú Beltrán, R. Proposal of virtual reality solution using Photogrammetry techniques to enhance the heritage promotion in a tourist center of Trujillo. In Proceedings of the 2022 IEEE Engineering International Research Conference (EIRCON), Lima, Peru, 26–28 October 2022; pp. 1–4. [CrossRef]
16. Tadeja, S.; Lu, Y.; Rydlewicz, M.; Rydlewicz, W.; Bubas, T.; Kristensson, P. Exploring gestural input for engineering surveys of real-life structures in virtual reality using photogrammetric 3D models. *Multimed. Tools Appl.* **2021**, *80*, 31039–31058. [CrossRef]
17. Raj, T.; Hashim, F.H.; Huddin, A.B.; Ibrahim, M.F.; Hussain, A. A Survey on LiDAR Scanning Mechanisms. *Electronics* **2020**, *9*, 741. [CrossRef]
18. Mikita, T.; Krausková, D.; Hrůza, P.; Cibulka, M.; Patočka, Z. Forest Road Wearing Course Damage Assessment Possibilities with Different Types of Laser Scanning Methods including New iPhone LiDAR Scanning Apps. *Forests* **2022**, *13*, 1763. [CrossRef]
19. Vogt, M.; Rips, A.; Emmelmann, C. Comparison of iPad Pro®'s LiDAR and TrueDepth Capabilities with an Industrial 3D Scanning Solution. *Technologies* **2021**, *9*, 25. [CrossRef]
20. Ferreira, V.S.; Martins, S.G.; Figueira, N.M.; Pochmann, P.G.C. The Use of a Digital Surface Model with Virtual Reality in the Amazonian Context. In Proceedings of the 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), Cape Town, South Africa, 9–10 December 2021; pp. 1–6. [CrossRef]
21. Ricci, M.; Evangelista, A.; Di Roma, A.; Fiorentino, M. Immersive and desktop virtual reality in virtual fashion stores: A comparison between shopping experiences. *Virtual Real.* **2023**, *27*, 2281–2296. [CrossRef] [PubMed]
22. Peukert, C.; Pfeiffer, J.; Meißner, M.; Pfeiffer, T.; Weinhardt, C. Shopping in Virtual Reality Stores: The Influence of Immersion on System Adoption. *J. Manag. Inf. Syst.* **2019**, *36*, 755–788. [CrossRef]
23. Wu, H.; Wang, Y.; Qiu, J.; Liu, J.; Zhang, X.L. User-defined gesture interaction for immersive VR shopping applications. *Behav. Inf. Technol.* **2019**, *38*, 726–741. [CrossRef]
24. Speicher, M.; Cucerca, S.; Krüger, A. VRShop: A Mobile Interactive Virtual Reality Shopping Environment Combining the Benefits of On- and Offline Shopping. *Proc. Acm Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *1*, 102. [CrossRef]
25. Speicher, M.; Hell, P.; Daiber, F.; Simeone, A.; Krüger, A. A virtual reality shopping experience using the apartment metaphor. In Proceedings of the AVI'18 2018 International Conference on Advanced Visual Interfaces, New York, NY, USA, 29 May–1 June 2018; pp. 1–9. [CrossRef]
26. Shravani, D.; R, P.Y.; Atreyas, P.V.; G, S. VR Supermarket: A Virtual Reality Online Shopping Platform with a Dynamic Recommendation System. In Proceedings of the 2021 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), Taichung, Taiwan, 15–17 November 2021; pp. 119–123. [CrossRef]
27. van Herpen, E.; van den Broek, E.; van Trijp, H.C.; Yu, T. Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. *Appetite* **2016**, *107*, 196–207. [CrossRef] [PubMed]

28. Nightingale, R.; Ross, M.T.; Cruz, R.; Allenby, M.; Powell, S.K.; Woodruff, M. Frugal 3D scanning using smartphones provides an accessible framework for capturing the external ear. *J. Plast. Reconstr. Aesthetic Surg. JPRAS* **2021**, *74*, 3066–3072. [CrossRef] [PubMed]
29. Luetzenburg, G.; Kroon, A.; Bjørk, A.A. Evaluation of the Apple iPhone 12 Pro LiDAR for an Application in Geosciences. *Sci. Rep.* **2021**, *11*, 22221. [CrossRef]
30. Fawzy, H.E.D. The Accuracy of Mobile Phone Camera Instead of High Resolution Camera in Digital Close Range Photogrammetry. *Int. J. Civ. Eng. Technol. (IJCIET)* **2015**, *6*, 76–85.
31. Croce, V.; Billi, D.; Caroti, G.; Piemonte, A.; De Luca, L.; Véron, P. Comparative Assessment of Neural Radiance Fields and Photogrammetry in Digital Heritage: Impact of Varying Image Conditions on 3D Reconstruction. *Remote Sens.* **2024**, *16*, 301. [CrossRef]
32. Zhang, Z.; Sun, W.; Min, X.; Wang, T.; Lu, W.; Zhai, G. No-Reference Quality Assessment for 3D Colored Point Cloud and Mesh Models. *arXiv* **2021**, arXiv:2107.02041. [CrossRef]
33. Liu, Q.; Su, H.; Duanmu, Z.; Liu, W.; Wang, Z. Perceptual Quality Assessment of Colored 3D Point Clouds. *IEEE Trans. Vis. Comput. Graph.* **2022**, *29*, 3642–3655. [CrossRef] [PubMed]
34. Zhang, Z.; Sun, W.; Min, X.; Zhou, Q.; He, J.; Wang, Q.; Zhai, G. MM-PCQA: Multi-Modal Learning for No-reference Point Cloud Quality Assessment. *arXiv* **2023**, arXiv:2209.00244. [CrossRef]
35. Nehmé, Y.; Dupont, F.; Farrugia, J.P.; Le Callet, P.; Lavoué, G. Visual Quality of 3D Meshes With Diffuse Colors in Virtual Reality: Subjective and Objective Evaluation. *IEEE Trans. Vis. Comput. Graph.* **2021**, *27*, 2202–2219. [CrossRef] [PubMed]
36. Lavoué, G. A Multiscale Metric for 3D Mesh Visual Quality Assessment. *Comput. Graph. Forum* **2011**, *30*, 1427–1437. [CrossRef]
37. Webster, N.L. High poly to low poly workflows for real-time rendering. *J. Vis. Commun. Med.* **2017**, *40*, 40–47. [CrossRef]
38. Schroeder, W.J.; Zarge, J.A.; Lorensen, W.E. Decimation of triangle meshes. *SIGGRAPH Comput. Graph.* **1992**, *26*, 65–70. [CrossRef]
39. Grande, R.; Albusac, J.; Castro-Schez, J.; Vallejo, D.; Sánchez-Sobrino, S. A Virtual Reality Shopping platform for enhancing e-commerce activities of small businesses and local economies. In Proceedings of the 2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Sydney, Australia, 16–20 October 2023; pp. 793–796. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Artículo II: Virtual Reality Shopping-Insights: A data-driven framework to assist the design and development of Virtual Reality shopping environments

- **Autores:** R. Grande, J.A. Albusac, D. Vallejo, C. González-Morcillo, S. Sánchez-Sobrino, J.J. Castro-Schez.
- **Tipo:** artículo de revista.
- **Revista:** SoftwareX.
- **Editorial:** Elsevier BV.
- **ISSN:** 2352-7110.
- **Año:** 2024.
- **DOI:** `https://doi.org/10.1016/j.softx.2024.101874`
- **Categoría:** Computer Science, Software Engineering.
- **Factor de Impacto (2024):** 2.4.
- **Factor de Impacto a 5 años (2024):** 3.2.
- **Ranking JCR:** Q2.
- **Número de citas (junio 2026):** 4.
- **Objetivos relacionados:** O3 (SO3.1, SO3.2, SO3.3), O5 (SO5.1, SO5.2, SO5.3).
- **Repositorio:** `https://github.com/ElsevierSoftwareX/SOFTX-D-24-00351`

SoftwareX 27 (2024) 101874

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original software publication

Virtual Reality Shopping-Insights: A data-driven framework to assist the design and development of Virtual Reality shopping environments

Rubén Grande *, Javier A. Albusac, David Vallejo, Carlos González-Morcillo, Santiago Sánchez-Sobrino, José J. Castro-Schez

Department of Information Technologies and Systems University of Castilla-La Mancha Paseo de la Universidad 4, 13071 Ciudad Real, Spain

ARTICLE INFO

Keywords:
Virtual reality
VR shopping
Behaviour analysis
Human–computer interaction
Software development
Marketing

ABSTRACT

In this paper, Virtual Reality Shopping Insights (VRSI) is presented, a framework that aids the development of Virtual Reality (VR) shopping applications as well as collecting data and analysing it. VRSI aims to help software developers and researchers by abstracting the layers needed to setup VR technology on applications developed with Unity. Moreover, it provides data registration tools that monitors user activity from non invasive data sources. This data can be useful for marketing analysts to understand user behaviour inside V-commerce applications, helping them to improve the setup and layout of such environments with data-driven decisions. Furthermore, we present an example of VR shopping application developed with VRSI.

Code metadata

Current code version	1.0
Permanent link to code/repository used for this code version	https://github.com/ElsevierSoftwareX/SOFTX-D-24-00351
Permanent link to Reproducible Capsule	https://github.com/AIR-Research-Group-UCLM/Virtual-Reality-Shopping-Insights/tree/master/apk
Legal Code License	MIT License
Code versioning system used	git 2.33.0
Software code languages, tools, and services used	C#, OpenXR VERSION, Meta XR All-in-One SDK 59.0, TextMeshPro 3.0.6, Unity 2022.3.15f1, Python 3.9, pandas, matplotlib, numpy
Compilation requirements, operating environments & dependencies	Oculus Meta Quest 2/Pro/3, Unity 2022.3.15f1, Android Support for Unity Editor, OVRBuild APK (optional)
If available Link to developer documentation/manual	https://github.com/AIR-Research-Group-UCLM/Virtual-Reality-Shopping-Insights/tree/master/doc
Support email for questions	ruben.grande@uclm.es

1. Motivation and significance

In recent years, Virtual Reality (VR) has been reported as a key technological tool in the digitisation process [1]. VR creates simulated environments in 3D spaces, allowing users to interact immersively through specific devices such as headsets, haptic gloves, and motion sensors. These virtual environments are designed to recreate sensory experiences that replicate or enhance the perception of physical reality [2].

VR is used in various fields, such as entertainment [3], advertising [4], tourism [5], and commerce [6,7]. In these fields, the use of VR promotes high levels of immersion, such as in video games or virtual replicas of dilapidated historic buildings. Moreover, it has been proven to provide positive effects for tourism promotion [5], for emulating of complex scenarios for training [8–10] and for facilitating work during the conceptualisation and prototyping stages of product design field [11].

In the specific domain of commerce and retail, companies like Amazon, Alibaba, eBay, and IKEA are integrating VR into their services to transform the future of shopping. VR addresses space and time limitations, enhancing the shopping experience by offering a more immersive

* Corresponding author.
E-mail address: Ruben.Grande@uclm.es (Rubén Grande).

https://doi.org/10.1016/j.softx.2024.101874
Received 1 July 2024; Received in revised form 20 August 2024; Accepted 29 August 2024
Available online 6 September 2024
2352-7110/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

and accessible experience at any time [6]. Although its adoption has accelerated after the COVID-19 pandemic [12], its widespread use remains limited. However, the interaction mechanisms that VR offers between customers and companies are already significantly influencing the creation of new business ecosystems, shaping both their value chains and decision-making processes [13].

This work focuses on e-commerce in immersive environments and particularly on the possibilities offered by analysing user interactions when using VR. Our contribution is an open-source framework called Virtual Reality Shopping Insights (VRSI), which enables the development and deployment of such systems, built on a modular and scalable architecture that provides functional modules for configuring VR-based shopping spaces, recording user interaction data, and providing an initial analysis of this data towards understanding the user behaviour.

In literature, papers that rely in standardised questionnaires to gather data for their V-commerce application's case study can be found [14]. In addition, there are studies that use data gathered from users' behaviour during experiments that is further used in their discussions [15–17]. However, these studies do not provide the software tools used to obtain these data, bur rather they focus on only sharing the data. The work of Elboudali et al. [18] proposed a conceptual framework to help with the identification of consumer preferences and early stages insights to personalise VR shopping experiences. This framework focused on giving steps to polish VR shop designs through behaviour data. Nevertheless, the study lacks of a formalised dataset to achieve such goals, as well as an approach to analyse the data of a fully immersive V-commerce, since they focused their analysis on virtual showrooms.

On one hand, VRSI enables the configuration and customisation of immersive shopping environments on a functional core that offers all together interaction mechanisms for visualising and manipulating objects in VR spaces. The customisation process involves both the visual aspects, related to components such as the scenario, shopping objects, or user interface, among others, and the functional aspects, through the integration of new functional modules. This capability of VRSI offers small businesses the potential opportunity to increase their operational efficiency and reduce the productivity gap with large companies that have digital e-commerce solutions [19].

On the other hand, VRSI eases data analysis tasks and experiment evaluation for researchers from various domains, such as e-commerce, marketing, psychology, accessibility, HCI, or AI. This feature of VRSI, aimed at understanding how users interact in VR-based immersive shopping environments, aligns with the previously mentioned idea of creating new business ecosystems supported by VR. In this sense, VRSI can be used by researchers without requiring technical knowledge of VRSI's internal design and development. Instead, they can use it as an integrated solution to focus on analysing the data generated by the framework.

This open-source framework provides developers with an additional abstraction for developing VR applications, saving them effort and time in configuring the layers required for proper functioning. Compared to other software tools that can aid developers, VRSI allows for effortless data collection configuration and aims to help professionals from other fields by providing a dataset for analysing user behaviour or identifying improvements in the immersive environment.

2. Software description

The VRSI framework is comprised of three distinct components: (i) *retail environment setup component*, (ii) *data collection component*, and (iii) *behaviour statistics component*. Fig. 1 shows, from a general perspective, the architecture that supports the VRSI framework and manifests the relationships between these components.

The *retail environment setup component* allows developers and researchers to configure the virtual space. Its main responsibility is to enable data collection and provide users interaction capabilities with the environment and products. Besides that, the *data collection component* is responsible for recording user activity data during the shopping experience. Finally, the *behaviour statistics component* provides a tool for obtaining statistics and data derived from the data collected by the *data collection component*. The objective of this last layer is to offer a set of statistics, based on metrics proposed in previous works in the literature [15–17,20], to help understand user behaviour in each session. Among these, we can mention those based on eye-tracking, user positional data and hand data, navigation in the immersive environment, and user interactions with it.

Overall, the proposed software not only provides the recording of raw data so that any researcher has the freedom to process it as they see fit, but it also offers a general understanding of it. Through Python *scripts* and the help of libraries like Pandas and Numpy, data processing is performed, allowing it to be structured and analysed in the form of graphs and numerical data based on different features: observation time, number of interactions, hands used, distance travelled, or volumetric heat-maps, among other data listed in Section 2.3.

2.1. Retail environment setup component

This component establishes the foundation for building a VR Shopping application, enabling the configuration of the shopping environment and providing the user with the ability to interact with the whole immersive space.

Regarding environment configuration, this component offers a series of *prefabs*[1] that allow design the 3D scene, starting with the shelves or product exhibitors of the virtual shopping environment. In the literature, the effects of product placement and the arrangement of shelves on potential consumers have been studied [21,22]. This influences, for example, product placement based on price at a certain height, or enhancing product visibility to boost the number of sales.

The type of shelves chosen will limit both the distribution of displayed products and the areas of interest (AOIs) [20]. By linking a *script* to the desired shelves, three configurable *colliders*[2] for three heights: *Top*, *Middle*, and*Bottom*, will be added. This allows knowing which specific area of the shelf has been interacted with.

Finally, it is highly recommended to setup teleportation navigation method using the *retail environment setup component*. This interaction mechanism allows building larger commercial environments and navigating comfortably without experiencing motion sickness [23]. Additionally, the component provides *prefabs* that facilitate the inclusion of teleportation zones or points.

Zones of Interest (ZOIs) allow obtaining structured information about the user's shopping experience when in a specific zone of the environment with regards the shelves that contain products [17]. The information obtained from the data collected through these features has been used in studies to recognise age [16] and gender [16,20], personality traits [17], or consumer's compulsivity [15]. ZOIs are significantly useful for understanding user behaviour and the analysis of proposed shopping tasks for case studies.

Moreover, the chosen product distribution is a major aspect for enhancing the user shopping experience [22]. Due to the usefulness of these features for analysing consumer behaviour, VRSI provides the possibility to easily incorporate them into any V-commerce application. Additionally, the sizes of the ZOIs and AOIs can be adjusted to match the dimensions of the shelves and the environment floor used. Regarding AOIs, the *script* adds the three aforementioned height levels, but it can be modified if the shelf cannot be divided in these heights. VRSI

[1] Components created within the Unity editor composed of a series of elements. These can be instantiated without recreating the group anew and are configurable, allowing them to act differently based on certain parameters.

[2] Components of the Unity physics engine that allow detecting events and collisions

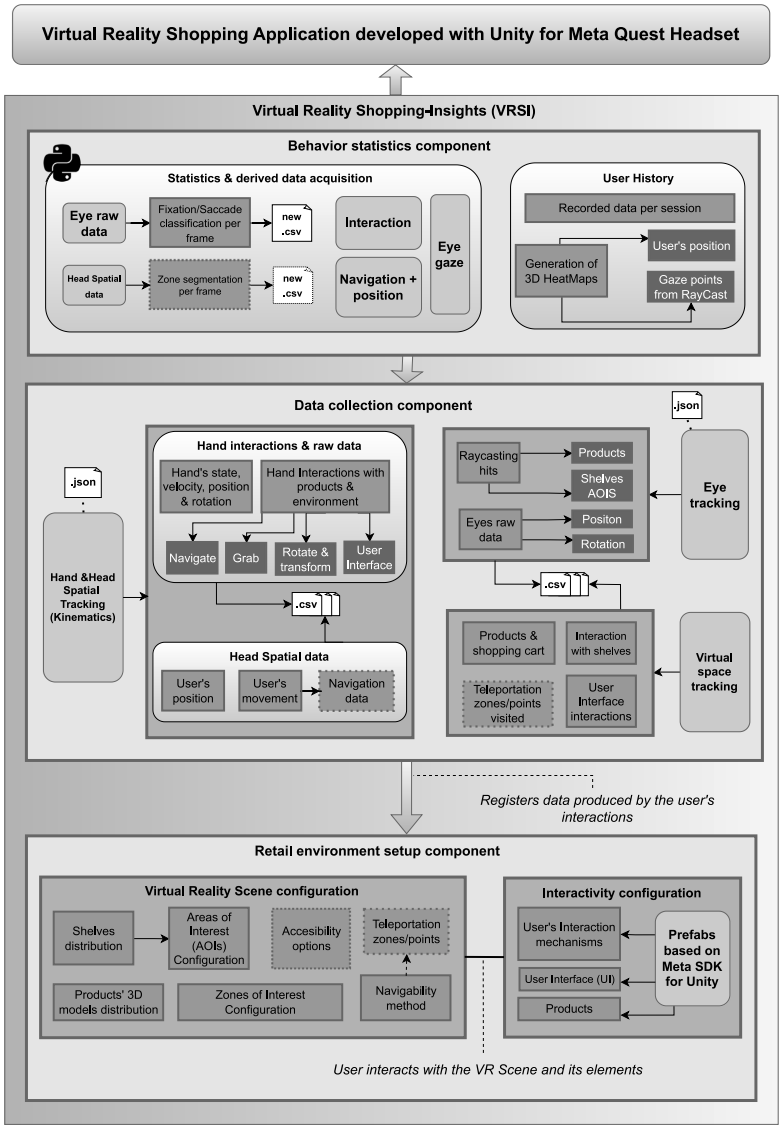


Fig. 1. Overview of the VRSI framework's architecture.

offers the option to segment data by ZOIs based on the user's position in each frame (see Section 2.3). Based on previous literature [15–17,20], the segmentation of the environment floor into ZOIs is based on four zones: (i) the *Shelf* itself, (ii) *Adjacent* to the shelf, (iii) *Near* the shelf, and (iv) *Far* from the shelf.

2.2. Data collection component

This component is responsible for collecting and structuring data generated during a V-commerce application session for behaviour analysis. The recorded data comes from non-invasive data sources to ensure

Table 1
Data items obtained by the Data collection component of VRSI.

Data item	Type	Range of values	Description
Frame number	Integer	[1, n]	The frame number from the start of the session
Timestamp	Float	[0, n]	The time measured in seconds from the start of the session
Hand position	Vector	(x, y, z) coordinates	Position of both left and right hands in 3D space
Hand rotation	Vector	(x, y, z) euler angles	Euler angles in each of the 3D space axes for both hands
Hand velocity	Vector	(x, y, z) coordinates	Velocity of both hands in each 3D space axis
Head position	Vector	(x, y, z) coordinates	HMD's position in 3D space
Head rotation	Vector	(x, y, z) euler angles	Euler angles in each of the 3D space axes for the HMD
Eye position	Vector	(x, y, z) coordinates	3D space position of both left and right eyes
Eye rotation	Vector	(x, y, z) euler angles	Euler angles in each of the 3D space axes for both eyes
Raycast hit	Vector	(x, y, z) coordinates	Averaged collision point obtained from the collision of both eyes' rays
AOI	String	Top, Middle or Bottom	Level of the shelf being monitored
Product	String	3D scene dependent	The name of the object in the scene representing the product in Unity
Product's state	String	Disabled, Normal, Hover or Select	State of the interactable product in the Interactor-Interactable lifecycle of Meta SDK
Section	String	3D scene dependent	Themed product section within the immersive shopping area
Shelf	String	3D scene dependent	The name of the shelf given in the Unity editor
Right hand state	String	Disabled, Normal, Hover or Select	State of a right hand's Interactor in the Interactor-Interactable lifecycle of Meta SDK
Left hand state	String	Disabled, Normal, Hover or Select	State of a left hand's Interactor in the Interactor-Interactable lifecycle of Meta SDK
Teleportation zone (TPHostspot)	String	3D scene dependent	The given in the Unity editor to a teleportation zone
Teleport duration	Float	[0,n]	The time the user spends targeting the teleport zone
Successful TP (WasTP)	Boolean	True or False	Indicates whether the targeting performed resulted in a successful teleport
ZOI	String	Start, Adjacent, Near or Far	If enabled, registers the zone of interest for each teleportation made by the customer
Distance	Float	[0, n]	Either the distance covered in one teleport or the distance from the user to the close's shelf
Shopping cart action	String	Add or Remove	Indicates if a given Product was added or removed from the shopping cart
UI item	String	3D scene dependent	Name of the UI element interacted
UI action	Boolean	True or false	True represents click or open while False represents close

the user's shopping experience is unaffected. This data is collected passively through the integrated technologies in current headsets and SDKs. Specifically, these data sources are (i) *hand tracking*, (ii) *head tracking*, (iii) *eye tracking*, and (iv) *virtual space tracking*.

The 3D coordinates obtained from the user's position via *head tracking* are stored in a JSON file, used by the *behaviour statistics component* to generate a 3D heatmap. Also stored in a JSON file for generating this heatmap are the collision points representing what a consumer has observed, thanks to Unity's physics engine ray-casting and the VR headset's eye tracking. These heatmaps help understand the areas where a user has spent the most time or the products or shelves that have caught their attention, with warmer colours indicating higher interest.

In addition to these two JSON files, this component structures the data into nine CSV files. The data recorded can be categorised as follows: (i) *Eye tracking* data regarding what the users saw and 3D spatial data, (ii) kinematics data from head and *hand tracking*, (iii) interactions with products and the environment, (iv) navigation data including teleportation and (v) shopping activities within the virtual environment by interacting with a shopping cart and user Interface (UI) elements. Further details and description about the structure of each of the files can be found in the developer documentation of VRSI's public repository: https://github.com/AIR-Research-Group-UCLM/Virtual-Reality-Shopping-Insights/tree/master/doc. Moreover, Table 1 describes each of the fields stored in such CSV files.

2.3. Behaviour statistics component

Additionally, our framework offers a Python-based command-line tool for obtaining reports to analyse user behaviour patterns. It also allows for maintaining a user history, with all data collected per session, and generating 3D heatmaps to graphically understand user interests (see Fig. 2). Furthermore, 3D heatmaps can be viewed and navigated by loading the session's JSON files in the Unity editor.

Based on the CSV files obtained with raw *eye-tracking* data, the component will generate two new CSV files identifying eye fixations and saccades during the session, using the dispersion-threshold identification (I-DT) algorithm for immersive virtual environments proposed by Llanes-Jurado et al. [24]. Additionally, the tool allows segmenting the CSV file related to the user's spatial data, dividing the floor plane into ZOIs (see Section 2.1). It allows the user to define its own distance parameters from the shelf, as the distances may vary depending on the virtual environment's configuration. Moreover, the script detects the frames in which the user is moving or stationary based on a configurable threshold.

Finally, the behaviour statistics component provides an initial analysis by generating three PDF reports that include graphs and statistics, mainly based on data features used in the literature [16] to identify user characteristics based on their behaviour in V-commerce environments. These reports are divided into (i) eye gaze, (ii) interaction, and (iii) navigation + positional reports. The features contained in these reports are shown in Table 2.

3. Illustrative example

In order to demonstrate the capabilities that VRSI offers, this section will illustrate and describe the use of VRSI for the development of a VR Shopping environment that we will use for future research. The Unity project, as well as the analysis scripts from which the

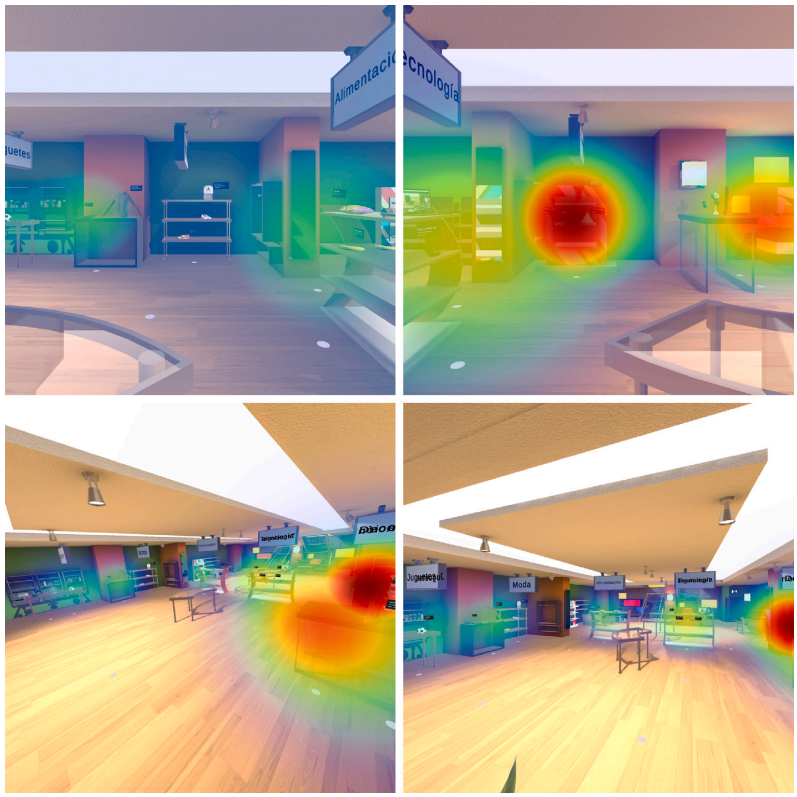


Fig. 2. A set of screenshots from a volumetric heatmap generated with VRSI. It also allows to change the gradient colour gamut.

figures have been obtained, can be found in the following repository: https://github.com/AIR-Research-Group-UCLM/Virtual-Reality-Shopping-Insights.git. This repository also contains an example analysis with generated reports, as well as the CSV files used, extracted from a sample execution. Additionally, a video showing the execution of a demo implementing VRSI is available here: https://youtu.be/fQeqDPCmEqI.

Fig. 3.a shows an overview of the virtual environment from its entry. Furthermore, Fig. 3.b and Fig. 3.c display elements that a person intending to use VRSI would utilise for implementation. The first figure shows part of the object hierarchy of a *prefab* representing the user within the VR environment, while the second illustrates the configurable parts of a script from the Data Collection component. The first figure abstracts configuration and lower layers for VR application implementation, whereas the second requires a simple *drag-and-drop* configuration for data collection. Fig. 3.d shows a group of volunteers participating in a study that used the prototype shown in Fig. 3.a.

In the context of interactions, Figs. 4.a and 4.b show the grasping of the same object. The first shows one-handed grasping, while the second shows grasping and rescaling which is only possible with both hands. Fig. 4.c shows the visual feedback indicating when a hand is grasping an object (green outline on the right hand) and when it is possible to grasp it (blue outline on the left hand). Moreover, Fig. 4.d illustrates a gesture that invokes a UI element to confirm the purchase when pressed, while Fig. 4.e shows adding an object to a shopping cart that follows the user's movement.

On the one hand, Fig. 5.a depicts the gesture that invokes the arc allowing targeting of available teleportation points, and Fig. 5.b shows the activation of this teleportation by joining the fingertips of the index finger and thumb. On the other hand, Figs. 5.c and 5.d depict the interaction adapted for users with upper limb motor issues, such as atrophy or spasticity issues, that prevent them from performing these gestures. Thus, it is enough to aim to the points with the palm facing the ceiling, regardless of its shape. This removes the difficulty of making a specific pose with the hands and fingers.

Finally, Fig. 6 shows three graphs and part of a table of numerical values related to statistics reported by the Behaviour Statistics component, listed in Table 2. Fig. 6.a shows a stacked bar graph to see the number of visits (a visit is counted as a consecutive block of lines in the .csv file depending on the variables being measured) with the eyes to each Shelf and within these, to each AOI, allowing to see which one stands out among those of the same Shelf. Furthermore, Fig. 6.b shows the user's presence in each section over the session time, indicating teleportation events with vertical bars. NIAS stands for "Not In Any

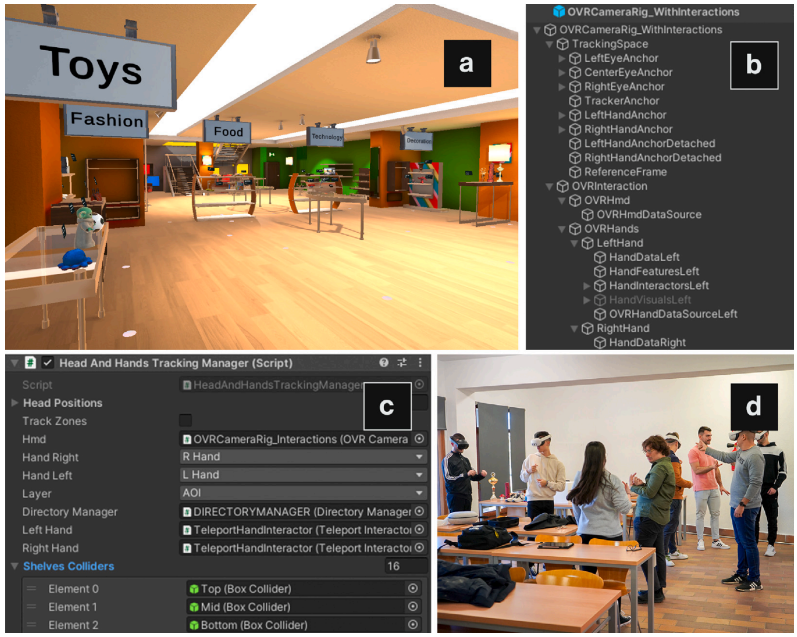


Fig. 3. Screenshots of the general view of the V-commerce environment and VRSI's elements inside Unity Editor.

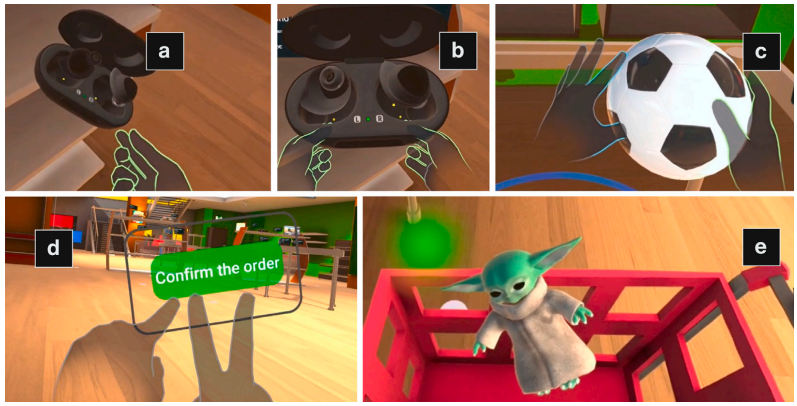


Fig. 4. Interactions with objects and UI elements along with visual feedback provided.

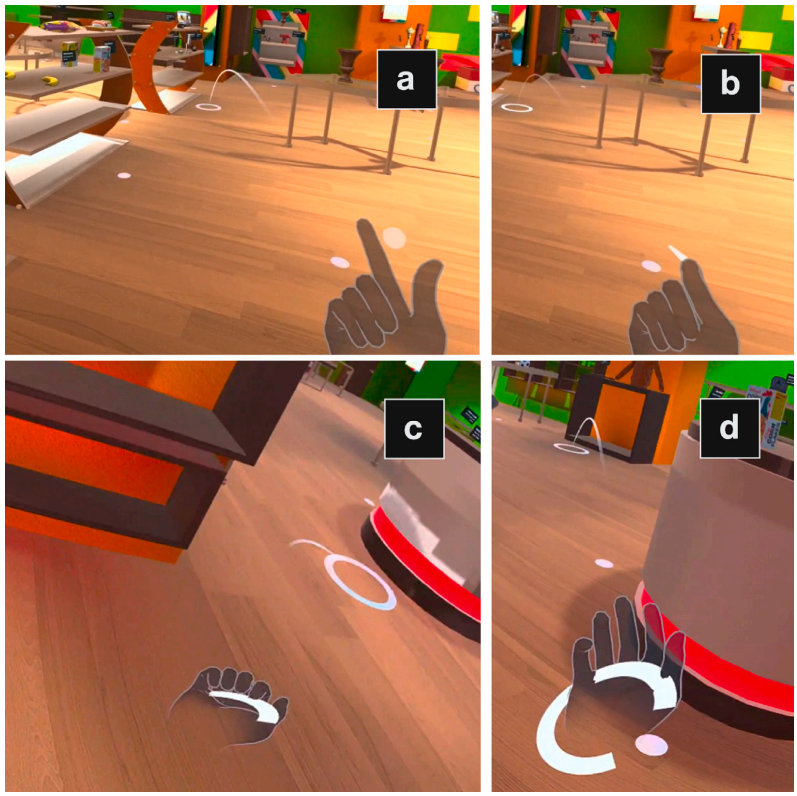


Fig. 5. Original and adapted teleportation interactions provided by VRSI.

Section", i.e., when the user is in the *Far* zone or at the start of the application's execution. Fig. 6.c shows part of a table of statistics about the number of times a gaze fixation has been made on each Shelf and each AOI, as well as the total duration of these fixations. At last, Fig. 6.d illustrates two graphs: on the left, the total time the user has spent in each zone and, on the right, the average time per visit to each zone in seconds, as well as the percentage of the total application time that each zone represents.

4. Impact

The proposed framework for recording and analysing user interactions in immersive shopping environments could have significant implications for various fields of research, including e-commerce, marketing, psychology, accessible design, human–computer interaction, and artificial intelligence. As an open-source tool, this framework allows researchers with programming expertise to adapt and customise it to their specific needs. Conversely, for researchers whose expertise does not lie in software design & development, the framework offers an out-of-the-box solution that facilitates data collection and experimentation. This dual accessibility enables a wide range of scholars to conduct in-depth studies and leverage the technological components without needing to develop them independently.

From an e-commerce perspective, the framework supports developers of immersive shopping environments by offering a extensible base platform. This allows for the comparison of the effectiveness of immersive commerce against traditional e-commerce [7]. By examining user interactions in a controlled yet flexible environment, researchers can gain insights into how immersive experiences can enhance customer engagement and sales [25]. We would like to note the value VRSI could provide for product evaluation in terms of both the V-commerce application and design and the products displayed. Relying on the data gathered, researchers can obtain insights whether products are properly displayed or need to be changed. In addition, it can help them to re-design the VR shop layout by identifying design flaws in shelves distribution or the shop itself, based on interaction and navigational data.

In the Marketing sector, the framework enables the evaluation of marketing strategies and techniques within immersive environments [26]. Researchers can conduct A/B testing to determine which versions of a product or marketing campaign perform better [27]. Additionally, the tool allows for market segmentation based on user behaviour, facilitating the design of more effective and targeted marketing campaigns [28]. This could lead to new strategies that are specifically tailored to immersive shopping experiences.

The field of Psychology benefits from this framework by providing a platform to study consumer behaviour in immersive shopping environments [22]. Researchers can explore how different environmental factors influence purchasing decisions and user engagement [29]. This could lead to new insights into consumer psychology and contribute to the development of more effective marketing and retail strategies.

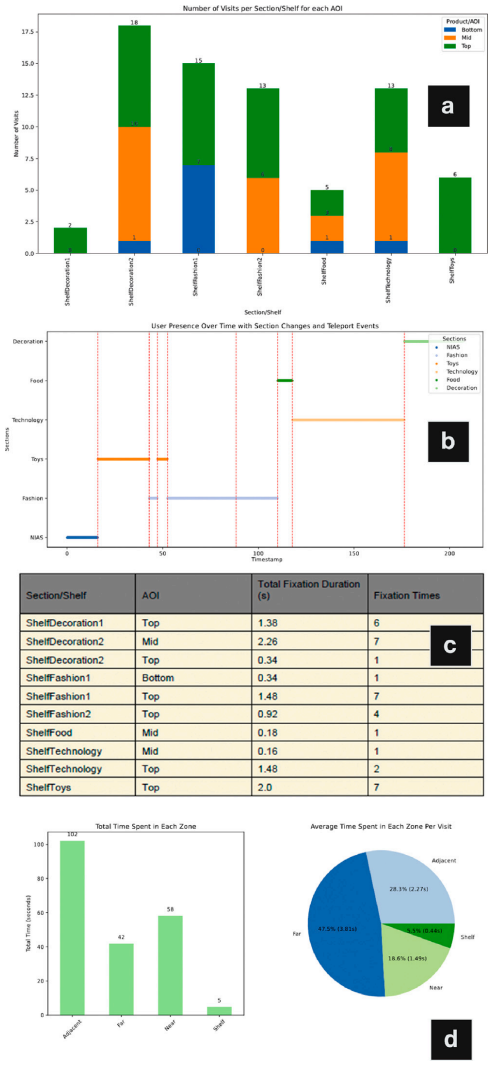


Fig. 6. Example of graphs and statistics of the initial analysis that VRSI provides.

Human–Computer Interaction (HCI) researchers can use the framework to enhance user experience in immersive shopping environments [30]. Unlike the extensively explored mobile apps and web interfaces, immersive shopping environments remain relatively unexplored in HCI research. This framework allows for the study and improvement of user interfaces and interactions within these novel environments, potentially leading to more intuitive and engaging user experiences.

Finally, in the field of Artificial Intelligence (AI), the framework provides a platform for creating dynamically adaptive shopping experiences based on user behaviour [31]. The framework enables the collection of rich datasets, which can be subsequently processed using predictive models for sales forecasting and pattern detection in shopping behaviour [15]. Furthermore, by integrating virtual avatars that act as sales assistants, researchers can study their interactions with real shoppers, thereby gaining valuable insights into human–avatar interaction dynamics and enhancing the overall shopping experience [32].

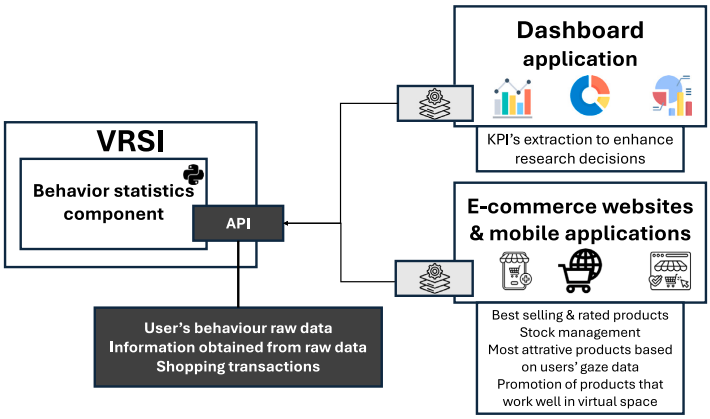


Fig. 7. Illustration of integration of VRSI with other software tools.

Regarding current limitations of VRSI, developers using it might design an interaction system based only on hand gestures. This would require them to modify the data gathering of the framework to be able to acquire data from the hand gestures used by the user, as well as to obtain insights of how useful are them based on error or trials made to make them. Along with this, VRSI currently lacks of integration with any other software tool or service than could consume the raw or analysed data for any further decision making. In this respect, Fig. 7 shows a diagram that includes the component that exposes an API that allows other services to obtain the information described in the image. This API is very beneficial to allow the synthesis of information gathered from the V-commerce application that implements VRSI.

5. Conclusions

In this work, we have presented VRSI, a framework intended to aid the development of VR Shopping applications and the analysis of user behaviour within these applications. As stated at the beginning, the field of e-commerce seems to be evolving toward the future adoption of VR to create V-commerce solutions [6,13]. To further research the effects of this new technology in retail and e-commerce, it is essential to have technologies that allow for agile development of VR Shopping applications and, additionally, to record and analyse data generated by users in these virtual environments. In this way, VRSI can assist various roles as shown in Section 4.

In terms of future work, we are researching the inclusion of accessibility elements for people with motor difficulties. The utilities offered in current SDKs for VR application development do not easily consider such options, such as hand tracking adapted for people with spasticity, arthritis, or other conditions, as well as navigation for those in wheelchairs. For further development directions, an improvement of the analysis tool that allows the analysis of a user's history, and not only a single session, is being designed.

Moreover, the diagram showed in Section 4 illustrates a future work lines that allow for the integration of the framework with other software tools. Regarding current and future trends in the e-commerce field, we will make efforts to research how to include social aspects of V-commerce into VRSI. Hence, it would provide researchers with the means to analyse the impact of sharing reviews or opinions through social networks as well as the effect of shopping with partners in the same virtual space.

CRediT authorship contribution statement

Rubén Grande: Writing – original draft, Software, Investigation, Conceptualization. **Javier A. Albusac:** Writing – review & editing, Software, Funding acquisition, Conceptualization. **David Vallejo:** Writing – review & editing, Supervision. **Carlos González-Morcillo:** Writing – review & editing, Methodology, Funding acquisition. **Santiago Sánchez-Sobrino:** Writing – review & editing, Resources, Funding acquisition. **José J. Castro-Schez:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Javier Alonso Albusac Jimenez reports financial support was provided by Spain Ministry of Science and Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The public repository with all the data and source code are fully open through the link provided in Code metadata table.

Acknowledgements

This work has been founded by the Spanish Ministry of Science, Innovation and Universities MICIU/AEI/10.13039/501100000033, and the European Union (NextGenerationEU/PRTR), under the Research Project: Design and development of a platform based on VR-Shopping and AI for the digitisation and strengthening of local businesses and economies, TED2021-131082B-I00.

Table 2
Statistics of the reports given by Statistics behaviour component.

Metric	Type	Dataset report
Number of visits per AOI	Graph	Eye Gaze
Number of visits per Shelf	Graph	
Number of visits per Shelf for each AOI	Graph	
Total time spent per AOI	Graph	
Total time spent per Shelf	Graph	
Total time and mean time spent per Shelf for each AOI	Graph	
Mean eye velocity per AOI for each Shelf	Graph	
Fixations and saccades over time	Graph	
Total and average fixation and saccade times	Numerical	
Number of fixations and duration per shelf	Numerical	
Same metrics are extracted for products being watched with eye tracking	Graph	Eye Gaze
User presence in each Section over time with teleportations noted	Graph	Navigation + Positional
User presence in Section over time for each zones	Graph	
Hand velocity (m/s) over time	Graph	
Stop and movement counts	Graph	
Total time spent in each zone	Graph	
Average time spent in each zone per visit		
Number of visits to each teleport point	Graph	
Average duration for successful teleport	Numerical	
Average velocity in each zone	Numerical	
Average + standard deviation velocity and acceleration for each hand	Numerical	
Number of visits per zone	Numerical	
Total distance travelled (m) without teleport	Numerical	
Number of interactions per product	Graph	Interaction
Average interaction time per product	Graph	
Percentage of interactions by Hand	Graph	
Percentage of interactions by Section	Graph	
Number of interactions, additions and removals from shopping cart per product	Graph	
Conversion rate (total additions to cart/total interactions with products added)	Numerical	
Average time difference between addition and removal from shopping cart	Numerical	

References

[1] Xiong J, Hsiang E-L, He Z, Zhan T, Wu S-T. Augmented reality and virtual reality displays: Emerging technologies and future perspectives. Light: Sci Appl 2021;10(1):1–30.

[2] LaValle SM. Virtual reality. Cambridge University Press; 2023.

[3] Jang Y, Park E. An adoption model for virtual reality games: The roles of presence and enjoyment. Telemat Inform 2019;42:101239.

[4] de Regt A, Plangger K, Barnes SJ. Virtual reality marketing and customer advocacy: Transforming experiences from story-telling to story-doing. J Bus Res 2021;136:513–22.

[5] Lo WH, Cheng KLB. Does virtual reality attract visitors? The mediating effect of presence on consumer response in virtual reality tourism advertising. Inf Technol Tourism 2020;22(4):537–62.

[6] Xi N, Hamari J. Shopping in virtual reality: A literature review and future agenda. J Bus Res 2021;134:37–58.

[7] Ricci M, Evangelista A, Di Roma A, Fiorentino M. Immersive and desktop virtual reality in virtual fashion stores: A comparison between shopping experiences. Virtual Real 2023;27(3):2281–96.

[8] Bracq M-S, Michinov E, Jannin P. Virtual reality simulation in nontechnical skills training for healthcare professionals: A systematic review. Simul Healthc 2019;14(3):188–94.

[9] Bracq M-S, Michinov E, Arnaldi B, Caillaud B, Gibaud B, Gouranton V, et al. Learning procedural skills with a virtual reality simulator: An acceptability study. Nurse Educ Today 2019;79:153–60.

[10] Lerner D, Mohr S, Schild J, Göring M, Luiz T, et al. An immersive multi-user virtual reality for emergency simulation training: Usability study. JMIR Serious Games 2020;8(3):e18822.

[11] Berni A, Borgianni Y. Applications of virtual reality in engineering and product design: Why, what, how, when and where. Electronics 2020;9(7):1064.

[12] Billewar SR, Jadhav K, Sriram V, Arun DA, Mohd Abdul S, Gulati K, et al. The rise of 3D E-commerce: The online shopping gets real with virtual reality and augmented reality during COVID-19. World J Eng 2022;19(2):244–53.

[13] De Regt A, Barnes SJ, Plangger K. The virtual reality value chain. Bus Horiz 2020;63(6):737–48.

[14] Speicher M, Hell P, Daiber F, Simeone A, Krüger A. A virtual reality shopping experience using the apartment metaphor. In: Proceedings of the international conference on advanced visual interfaces. 2018, p. 1–9. http://dx.doi.org/10.1145/3206505.3206518.

[15] Moghaddasi M, Marín-Morales J, Khatri J, Guixeres J, Chicchi Giglioli IA, Alcañiz M. Recognition of customers' impulsivity from behavioral patterns in virtual reality. Appl Sci 2021;11(10):4399. http://dx.doi.org/10.3390/app11104399, URL https://www.mdpi.com/2076-3417/11/10/4399. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.

[16] Gil-López C, Guixeres J, Moghaddasi M, Khatri J, Marín-Morales J, Alcañiz Raya M. Recognizing shopper demographics from behavioral responses in a virtual reality store. Virtual Real 2023;27. http://dx.doi.org/10.1007/s10055-023-00767-2.

[17] Khatri J, Marín-Morales J, Moghaddasi M, Guixeres J, Giglioli IAC, Alcañiz M. Recognizing personality traits using consumer behavior patterns in a virtual retail store. Front Psychol 2022;13. http://dx.doi.org/10.3389/fpsyg.2022.752073, URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.752073/full. Publisher: Frontiers.

[18] Elboudali A, Aoussat A, Mantelet F, Bethomier J, Leray F. A customised virtual reality shopping experience framework based on consumer behaviour: 3DR3CO. Int J Interact Des Manuf (IJIDeM) 2020;14. http://dx.doi.org/10.1007/s12008-020-00645-0.

[19] Pavlova V, Murovana T, Sarai N, Velychko V, Illyashenko K, Hrysshyna H. Crisis phenomena as an incentive to intensify e-commerce of the enterprise. J Theor Appl Inf Technol 2021;99(19):4646–57.

[20] Moghaddasi M, Khatri J, Llanes-Jurado J, Spinella L, Marín-Morales J, Guixeres J, et al. Segmentation of areas of interest inside a virtual reality store. In: Stephanidis C, Antona M, editors. HCI international 2020 - posters. Cham: Springer International Publishing; 2020, p. 92–8. http://dx.doi.org/10.1007/978-3-030-50729-9_13.

[21] Chittaro L, Ranon R. New directions for the design of virtual reality interfaces to e-commerce sites. In: Proceedings of the working conference on advanced visual interfaces. New York, NY, USA: Association for Computing Machinery; 2002, p. 308–15. http://dx.doi.org/10.1145/1556262.1556311.

[22] Lombart C, Millan E, Normand J-M, Verhulst A, Labbé-Pinlon B, Moreau G. Effects of physical, non-immersive virtual, and immersive virtual store environments on consumers' perceptions and purchase behavior. Comput Hum Behav 2020;110:106374.

[23] Plouzeau J, Chardonnet J-R, Merienne F. Using cybersickness indicators to adapt navigation in virtual reality: A pre-study. In: 2018 IEEE conference on virtual reality and 3D user interfaces. 2018, p. 661–2. http://dx.doi.org/10.1109/VR.2018.8446192.

[24] Llanes-Jurado J, Marín-Morales J, Guixeres J, Alcañiz M. Development and calibration of an eye-tracking fixation identification algorithm for immersive virtual reality. Sensors 2020;20(17). http://dx.doi.org/10.3390/s20174956, URL https://www.mdpi.com/1424-8220/20/17/4956.

[25] Eisingerich AB, Marchand A, Fritze MP, Dong L. Hook vs. hope: How to enhance customer engagement through gamification. Int J Res Market 2019;36(2):200–15.

[26] Raji MA, Olodo HB, Oke TT, Addy WA, Ofodile OC, Oyewole AT. Business strategies in virtual reality: A review of market opportunities and consumer experience. Int J Manag Entrepreneurship Res 2024;6(3):722–36.

[27] Kohavi R, Tang D, Xu Y. Trustworthy online controlled experiments: A practical guide to a/b testing. Cambridge University Press; 2020.

[28] Wedel M, Bigné E, Zhang J. Virtual and augmented reality: Advancing research in consumer marketing. Int J Res Market 2020;37(3):443–65.

[29] Sun Y, Zhong Y, Zhang Z, Wang Y, Zhu M. How technical features of virtual live shopping platforms affect purchase intention: Based on the theory of interactive media effects. Decis Support Syst 2024;114189.

[30] Su K-W, Chen S-C, Lin P-H, Hsieh C-I. Evaluating the user interface and experience of VR in the electronic commerce environment: A hybrid approach. Virtual Real 2020;24(2):241–54.

[31] Raji MA, Olodo HB, Oke TT, Addy WA, Ofodile OC, Oyewole AT. E-commerce and consumer behavior: A review of AI-powered personalization and market trends. GSC Adv Res Rev 2024;18(3):066–77.

[32] Zhu S, Hu W, Li W, Dong Y. Virtual agents in immersive virtual reality environments: Impact of humanoid avatars and output modalities on shopping experience. Int J Hum–Comput Interact 2023;1–23.

Artículo III: Enhancing Hand Interactions and Accessibility in Virtual Reality Shopping Environments for Users with Upper Limb Motor Impairments

- **Autores:** R. Grande, J. Albusac, D. Vallejo, S. Sánchez-Sobrino, J.J. Castro-Schez.
- **Tipo:** artículo de revista.
- **Revista:** IEEE Access.
- **Editorial:** IEEE.
- **ISSN:** 2169-3536.
- **Año:** 2025.
- **Número de citas (junio 2026):** 5.
- **DOI:** `https://doi.org/10.1109/ACCESS.2025.3549527`
- **Categoría:** Computer Science, Information Systems.
- **Factor de Impacto (2024):** 3.6.
- **Factor de Impacto a 5 años (2024):** 3.9.
- **Ranking JCR:** Q2.
- **Objetivos relacionados:** O4 (SO4.1, SO4.2, SO4.3, SO4.4).
- **Repositorio:** `https://github.com/AIR-Research-Group-UCLM/VR-Hands-Interactions-Adapted`

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 11 February 2025, accepted 4 March 2025, date of publication 10 March 2025, date of current version 17 March 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3549527

RESEARCH ARTICLE

Enhancing Hand Interactions and Accessibility in Virtual Reality Environments for Users With Motor Disabilities: A Practical Case Study on VR-Shopping

RUBÉN GRANDE[1], JAVIER ALBUSAC[1], VANESA HERRERA[1], DOROTHY MONEKOSSO[2], ANA DE LOS REYES-GUZMAN[3], DAVID VALLEJO[1], AND J. J. CASTRO-SCHEZ[1]

[1]Department of Information Technologies and Systems, University of Castilla-La Mancha, 13071 Ciudad Real, Spain
[2]School of Engineering and Physical Sciences, Aston University, B4 7ET Birmingham, U.K.
[3]Biomechanics and Technical Aids Unit, National Hospital for Paraplegics, 45004 Toledo, Spain

Corresponding author: Rubén Grande (ruben.grande@uclm.es)

This work was supported in part by the Research Project, entitled Mixed Reality-Based Platform on 5G Infrastructure for Remote Support and Diagnosis Among Nursing, General Practitioners, and Medical Specialists (MRP-5G) under Grant ID SBPLY/23/180225/000182; in part by the European Union through European Regional Development Fund (FEDER); in part by Junta de Comunidades de Castilla-La Mancha through Castilla-La Mancha Agency for Research and Innovation (INNOCAM); in part by the Project Design and Development of a Platform based on VR Shopping and AI for the Digitalization and Strengthening of Local Businesses and Economies, which was funded by the Spanish Ministry of Science and Innovation MICIN/AEI/10.13039/501100000033 and the European Union (NextGenerationEU/PRTR) under Grant TED2021-131082B-I00; and in part by the Project Platform for Upper Extremity Rehabilitation based on Immersive Virtual Reality (Rehab-Immersive), funded by the Spanish Ministry of Science and Innovation MCIN/AEI/10.13039/501100011033 under Grant PID2020-117361RB-C21 and Grant PID2020-117361RB-C22. The work of Rubén Grande was supported by the Pre-doctoral FPI (Research Staff Training) contract at the University of Castilla-La Mancha, Spain, from the European Union through the European Social Fund Plus (FSE+).

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Health Care Ethics Committee of Hospital Nacional de Parapléjicos de Toledo.

ABSTRACT Over the past decade, Virtual Reality (VR) has achieved significant advancements in both quality and accessibility of its devices, particularly with VR headsets that offer an enhanced immersive experience at a reduced cost. This improvement is not only in graphical fidelity but also in interactivity within virtual environments, highlighted by advancements in tracking systems that allow direct manipulation with hands-free of external devices. However, these technologies are not fully adapted for use by individuals with motor disabilities in their arms and hands. This study addresses the ethical and moral obligation to make VR accessible to all users by proposing specific adaptations to the manual interaction mechanisms. Focusing on VR e-commerce as a use case, which is anticipated to revolutionize the shopping experience in the coming years, this study explores a scenario where users need to navigate and manipulate virtual products to examine and make purchase decisions. The experimental phase was conducted in a real-world setting at the Hospital Nacional de Parapléjicos in Toledo (HNPT), involving patients with spinal cord injuries and motor limitations. The results demonstrate that the adapted interactions not only enable users to perform tasks that are impossible with conventional mechanisms but also reduce the time and effort required to complete these tasks. Specifically, completion times achieved by users increased up to 91.6%. Also, the number of tasks completed highly increased in comparison with unadapted interactions. We elaborate an Effort Degree (ED) formula with various data items based on hand movements, that demonstrated that adapted interactions

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Wei.

© 2025 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

required around 40% less effort for individuals with mobility issues. These findings underscore the potential of adapted VR technologies to significantly enhance the inclusivity and utility of emerging e-commerce platforms, providing equitable access to new forms of digital engagement.

INDEX TERMS Virtual reality, adapted hand interaction, motor limitations, accessibility, inclusivity, e-commerce.

I. INTRODUCTION

Virtual Reality (VR) has experienced significant growth over the past five years due to various factors. A key driver has been the strong commitment from major technology companies such as META, Apple, Microsoft, HTC, Sony, and Samsung (among others). These companies have contributed to the evolution of VR by developing higher quality VR headset models (improving resolution, refresh rate, ergonomics, etc.) at a cost that is becoming affordable to a broader segment of society. This affordability could lead to VR and Extended Reality (XR) headsets becoming the next widely used devices, potentially reaching the level of mobile phone usage.

This momentum has contributed to a growing interest in this technology from numerous companies [32] and the developer community, thereby expanding the range and catalogue of products related to immersive spaces and motivating the emergence of the Metaverse [37]. Simultaneously, there is an increased interest within the scientific community, with a rise in research focused on areas such as Education [5], Health [38], Simulation [1], Entertainment [25], and E-commerce [2], among others. Notably, in the field of health rehabilitation, the use of VR is increasingly employed to provide users with controlled and gamified environments that enhance patient motivation and complement traditional therapies [10], resulting in positive effects on mobility recovery. Additionally, in e-commerce, VR is expected to drive a new evolution in the way we shop, offering experiences similar to those in a physical store but from the comfort of home [3].

In addition to technological advancements in hardware and the availability of software proposals, the growing interest is also driven by improvements in the interaction between users and virtual elements within the virtual space, resulting in a significant enhancement of the sense of immersion. This is mainly due to improved hand interaction without the use of external devices such as joysticks, or any other kind of controllers. Devices like the META Quest models integrate a set of external cameras in the headset itself and tracking algorithms that enable precise hand tracking and virtual representation [29]. The increased computational capacity of these devices, integrated into a smaller form factor, allows the execution of these demanding algorithms. Additionally, the API associated with the headset facilitates developers in implementing functionality related to hand and virtual component interactions. This is complemented by the physics engine included in the development environment (such as Unity), which endows objects with realistic behaviour.

Despite all these improvements, current hand interaction in immersive spaces requires good health, normal mobility in hands and fingers, and the ability to perform recognizable gestures [4]. Otherwise, most systems face significant difficulties in detecting user intentions within the virtual space [13]. Therefore, this excludes individuals with hemiplegia and motor limitations, who have difficulty moving their hands and fingers or suffer from spasticity. For example, those who have suffered a stroke or a spinal cord injury affecting their upper limbs, and do not have total disability in their hands. Hence, there is a need to adapt some of the basic hand interactions for individuals with such disabilities to create more accessible virtual environments for everyone.

As previously mentioned, VR has seen a significant boost in recent years, entering a phase of maturation. When a technology reaches this stage, broader concerns such as usability and, importantly, accessibility must be considered. Just as there is a global effort to improve accessibility in the real world, virtual environments must not be an exception. In fact, this is a fundamental requirement for VR/AR devices to be used widely on a global scale. Moreover, it is a moral imperative to make technological advances accessible to all.

In this article, we present a set of basic hand interactions in virtual spaces and their potential adaptations to make them more accessible for individuals with upper limb motor limitations. As a case study, we propose their use in VR-Shopping, a virtual shopping space where real products, previously digitized, are displayed. In this environment, users can explore the environment, navigate, and interact with products using their hands.

For the experimentation, we collaborated with the Hospital Nacional de Parapléjicos (HNPT) in Toledo, Spain, specifically with clinical and technical staff from the hospital's Biomechanics Unit and patients with Spinal Cord Injury (SCI) who suffer from motor limitations. All participants in the experiment were assigned a series of tasks within the immersive shopping space, which they had to complete using both non-adapted and subsequently adapted hand interactions. The results demonstrate that the use of adapted interactions allows people with SCI to fully complete not only the tasks designed for the case study but also to complete them in much less time. In addition, the effort made was also substantially reduced thanks to the adaptation made to the interactions, so muscular fatigue is alleviated.

This proposal has the potential to benefit a significant number of people globally. According to the World Health Organization (WHO),[1] it is estimated that over 1 billion people globally live with some form of disability, with a significant portion experiencing motor impairments.

[1] https://www.who.int/publications/i/item/9789241564182

Specifically, it is reported that approximately 2-3% of the global population, or roughly 150-225 million people, have motor limitations affecting their hands. These individuals could greatly benefit from the proposed adaptations in virtual hand interactions, enhancing their ability to engage with immersive virtual environments and participate more fully in activities such as VR-Shopping.

The remainder of the article is organised as follows. Section II presents related work on the need to adapt hand interactions for virtual reality applications and the commercial context by different VR headset manufacturers. Section III details the VR interactions that were adapted, as well as how they were adapted for people with SCI. Next, Section IV describes the case study carried out, including the methodology followed and the research questions. Following it, V shows both quantitative and qualitative analysis of the case study described previously. Finally, Section VI presents the conclusion of our work, summarising it and briefly describing our future work.

II. RELATED WORK

In this section, we discuss some significant contributions in the scientific-academic and commercial sectors related to interaction mechanisms and accessibility improvement in semi-immersive and immersive environments for individuals with disabilities. Specifically, in the scientific-academic context, we categorise the noteworthy research into four groups: i) studies that express and justify the need to enhance accessibility in VR (Virtual Reality) and AR (Augmented Reality) environments. Once this need is established, ii) other authors focus their efforts on assessing the current state of accessibility to raise awareness. iii) Subsequently, based on the measured accessibility levels, proposals aimed at enhancing it. iv) Finally, some studies involve hand interaction in VR environments among individuals with disabilities.

A. SCIENTIFIC-ACADEMIC CONTEXT

Recent studies have analysed the current shortcomings in terms of accessibility in systems based on VR and AR, highlighting the need to propose solutions for creating more inclusive virtual spaces. An example is the work conducted by [13], where the need to improve accessibility in both VR and AR systems for various types of disabilities is identified and justified. This article meticulously examines proposals to date for this purpose, in both the academic and commercial spheres. Among the proposals related to hand usage, there are none that suggest adapting interactions; instead, alternative methods such as the use of gloves or joysticks, which are unsuitable for people with hand deformities, or the use of voice commands, or issuing orders through pupil or head movement are discussed.

Following the same line of work, namely the detection of adaptation needs in interactions for accessibility improvement, [4] addresses the need to enhance accessibility in VR and AR environments for people with diverse abilities. It proposes using user model-based customization to tailor these technologies to individual needs, highlighting the limited research and application outside of gaming and navigation contexts. Through a workshop that brings together experts and developers, the paper explores new AR/VR applications and services, in addition to promoting the development of standards and guidelines to expand access and inclusion of these technologies in various social and educational settings.

[31] presents an exploration of the accessibility of VR systems for people with limited mobility. Through semi-structured interviews with 16 participants, the authors identified seven barriers related to the physical accessibility of VR devices, ranging from the initial setup of the system to keeping the VR controllers within the field of view of the cameras integrated into the VR headsets. Additionally, possible improvements were proposed to address some accessibility concerns, emphasizing the importance of considering the abilities of people with limited mobility when designing VR systems, as the capabilities of many participants did not match the design assumptions of current systems. For individuals with motor limitations in arms and hands it suggests interactions through voice commands and gaze control. In addition to the three previous studies, other authors such as [9], [12], and [30] complement what was mentioned above, following the same line of work.

Having highlighted the importance of constructing more immersive virtual environments, it makes sense to have the ability to measure the current state of accessibility in any system. Thus, there are studies aimed at measuring and evaluating the user experience in a VR environment, which subsequently allows for improvements to that experience. Such is the case with [28], which presents a model to assess and enhance the user experience in VR environments, focusing on hand movements and considering laterality (dominant and non-dominant hand). The study includes a practical case with two groups of users performing specific tasks, evaluating aspects such as engagement, enjoyment, and usability. The results indicate that customizing and adapting virtual environments according to laterality and individual responses significantly improves the user experience, suggesting the need for further research to optimize interaction and maximize usability and enjoyment in virtual reality contexts.

A second study that conducts an accessibility analysis in VR environments is [41]. This work aims to evaluate the effectiveness of accessibility features in a VR environment. An open-source solution called Gear VRF Accessibility was tested, which provides a framework for developers to adapt functions such as Zoom, Inverted Colours, Auto Reading (Screen Reader), and Subtitles in a VR environment. These features were implemented in an application and evaluated in a controlled environment with 12 participants with disabilities. The study concluded that while users appreciated the accessibility of the solution, there are still challenges in using VR without tactile contact, and further

studies are needed to improve accessibility in various contexts and for different types of disabilities.

Given the importance of creating inclusive spaces, and the availability of mechanisms to measure the degree of accessibility, it makes sense to propose solutions to improve accessibility. [8] addresses the optimization of the placement of interactable elements in VR environments. This approach focuses on modifying the layout of interface elements to adapt them to the specific conditions of each user using manual interaction techniques, such as direct touches and air gestures. However, the study does not specifically address the needs of users with motor limitations, which implies that the proposed adaptations might not be fully accessible to people with such limitations.

Other authors, like [17], rely on VR-based systems to complement rehabilitation methods for patients with motor limitations, particularly patients with spinal cord injuries who require upper limb rehabilitation. They propose an advanced calibration method that individually detects limitations on the left and right sides of the body. Based on this information, the location of objects with which the user must interact is automatically adjusted to make them accessible and to ensure they do not require excessive effort or compensations that could further endanger the patient.

Moreover, we want to show use cases where accessibility has been improved, particularly through the use of hands as the primary mechanism of interaction. Reference [35] propose an intelligent wheelchair control system that operates solely through hand use instead of joysticks for people suffering from dystonia. The method is primarily based on a recurrent neural network. It includes a camera integrated into the wheelchair and vision algorithms that recognize different hand poses with motor limitations and translate them into commands. It is designed for patients who cannot use joysticks. This is not implemented in an immersive environment, but rather with a 3D simulator displayed on a 2D screen. It starts from pre-recorded poses, but each patient is a different case. They look for similarities with reference poses. Tested with three patients. Other similar articles, in which gesture recognition is used for wheelchair control, include [11] and [19].

Reference [26] discusses how traditional user interfaces in virtual reality (VR), which are vision-centred, do not sufficiently consider the effectiveness and comfort of manual interactions. In response, an adaptive hand movement user interface technique is proposed to improve the performance of manual interactions in VR. Through an experiment with 24 participants, it was demonstrated that the choice of hand and position of the interaction significantly influence the efficiency and accuracy of target selection. The results indicated that adaptive hand interaction results in greater interaction efficiency and less physical effort and perceived difficulty of the task compared to the traditional vision-centred interface. Guidelines are provided for designing more efficient and user-friendly user interfaces in VR applications. All is based on the position of the hand relative to the chest, with no gestures involved.

Herrera et al. [18] proposed a serious games platform based on VR for the rehabilitation of upper limbs in patients with motor limitations who have suffered a stroke or spinal cord injury. The platform features a core common to all serious games, with basic functionality. Part of this functionality is focused on hand tracking and the adaptation of hand interaction mechanisms to manipulate virtual objects. The adaptation primarily centres on modifying pinch grips for individuals who are unable to perform the complete gesture of touching the thumb and index finger together.

We conclude this section by focusing on how the difficulty faced and effort made by participants in case studies involving the use of VR applications has been measured. In this sense, we will not limit ourselves to case studies whose volunteers have limited upper limb motor skills in order to cover a broader spectrum of participants. Reference [15] focuses on evaluating the interaction modalities of hand-tracking and handheld controllers in virtual reality learning environments (VRLEs) for performing motor tasks such as reach-pick-place activities. Quantitatively, in-game analytics were utilised to measure performance metrics such as task duration, frequency of clicks (right and left), and the number of objects picked. These metrics were systematically recorded across four experimental conditions (hand-tracking vs. controller $\times$ photorealistic vs. non-photorealistic environments) to capture detailed behavioural data during task execution. Qualitatively, the study implemented the Rated Scale Mental Effort (RSME) questionnaire to evaluate the perceived mental workload of participants. After completing each task, participants provided subjective feedback using this scale, which allowed researchers to gauge their cognitive and physical effort.

Jamalian et al. [21] evaluated user effort and difficulty during a memory puzzle task relied on a detailed collection of performance data. To assess the accuracy of task execution, the researchers measured the percentage of trials where participants correctly replicated both the order and sequence of the patterns (referred to as ''Correct Order and Pattern''). A related metric, ''Correct Pattern Only'', captured instances where participants reproduced the correct pattern without maintaining the proper sequence, providing an additional layer of insight into user performance. Timing metrics were also central to the study. The ''First Selection Time'' recorded the time elapsed before participants selected the first button in a trial, serving as an indicator of their response immediacy. Additionally, the ''Trial Completion Time'' measured the total duration required to complete each trial, offering a quantitative view of task efficiency. To evaluate motor behaviour and interaction dynamics, the study tracked the usage of participants' dominant and non-dominant hands using hand-tracking data. By calculating the percentage of interactions performed by each hand, the researchers investigated whether the naturalistic interaction

of hand-tracking encouraged more balanced hand usage compared to controller-based systems.

The work of [27] researched the usability of hand-tracking for grab-and-place tasks in immersive virtual reality employed a detailed set of metrics to evaluate the effort and difficulty associated with this interaction method. Performance data from hand-tracking interactions were central to the analysis, focusing on accuracy, task timing, and error rates. Accuracy was measured as the Euclidean distance between the virtual object's placement and the target location, providing a precise indicator of interaction precision. Timing metrics included the duration from when the object appeared to when it was successfully grabbed, referred to as grab time, and the time from grabbing to placing the object, termed release time. These measurements highlighted the efficiency and responsiveness of the hand-tracking interface. Total task time, which encompassed the entire interaction from the object's appearance to its placement, further contextualized the interaction's overall efficiency. Accidental drops, defined as unintentional releases of objects before placement, were recorded to identify errors specific to the hand-tracking modality and assess its reliability under experimental conditions. Subjective measures complemented these performance metrics, capturing user perceptions of the interaction. Participants rated their experiences using the System Usability Scale (SUS), evaluated the naturalness and precision of hand-tracking, and reported their sense of agency—how much they felt in control of their virtual hands. These qualitative responses added depth to the analysis by reflecting the cognitive and physical demands experienced during hand-tracking use.

The authors of [20] assessed fine motor skills, specifically focusing on reaching accuracy under various sensory and cognitive conditions. Hand-tracking data were central to evaluating effort and difficulty in the tasks performed by healthy participants and individuals with cerebellar stroke. Metrics included the accuracy of reaching movements, measured as the angular difference between the target location and the participant's finger trajectory. This provided a precise indicator of task performance. Additionally, the study recorded reaching precision, calculated as the standard deviation of endpoint accuracy across trials. Reaching time, defined as the duration from the appearance of the target to the completion of the movement, was also measured to provide insights into the speed and control of hand movements. These metrics allowed for a detailed analysis of the interplay between sensory modalities, such as vision and proprioception, and their impact on motor task execution. Hand-tracking data specifically enabled the evaluation of performance under conditions where visual feedback of the hand was manipulated or absent, isolating the reliance on proprioceptive input.

B. COMMERCIAL ENVIRONMENT

In this section, we will explore accessibility features included in current VR headsets of the market per manufacturer. META - accessibility features https://www.meta.com/es-es/help/quest/articles/in-vr-experiences/oculus-features/accessibility-features/, allow for height adjustment, audio balance, colour correction, contrast, subtitles, mono audio, joystick control remapping, font size adjustment, and thumb calibration for joystick handling. There is no adaptation for hand interaction.

HTC https://www.htc.com/us/accessibility/ - HTC virtual reality headsets offer various accessibility settings to enhance the experience for users with disabilities. These settings include options to improve visibility such as the TalkBack screen reader, which allows users to navigate and receive auditory feedback without needing to look at the screen. Font sizes can also be adjusted and magnification gestures can be enabled to enlarge parts of the screen. For those with visual impairments, options such as color inversion and color correction for different types of color visual impairment are available. In terms of hearing, HTC devices offer compatibility with hearing aids and special modes that reduce interference, as well as options for closed captions and visual or vibration notifications. The devices are also equipped for voice input and voice commands to facilitate interaction without the need to touch the device, and functions can be accessed through the use of external switches or device buttons. Additionally, features like motion gestures and touch delay adjustments help improve interaction for people with manual dexterity difficulties.

Sony https://www.playstation.com/en-us/ps-vr/ - Concerning the accessibility configurations for the PlayStation VR2, it is important to note that all of the PlayStation 5's system-level accessibility settings, with the exception of Zoom, remain available when operating in virtual reality mode. This encompasses support at the system level for functionalities such as Colour Correction, Invert Colour, and Custom Button Assignment.

HP https://support.hp.com/us-en/product/setup-user-guides/hp-reverb-g2-virtual -reality-headset/33835976 - The main configuration options that primarily affect accessibility are found in Audio & Visual and, on the other hand, Display and Graphics.

Finally, Pico 4 https://www.picoxr.com/global - No official page was found that includes features or options to improve accessibility.

To conclude the section, it is worth mentioning that most of the cited works do not address the adaptation of hand interactions for people who do not have a total disability and can move them with limitations. As an alternative, they usually propose interaction with other parts of the body, such as the use of voice, head movements, or eye tracking. In fact, in [29], a review that includes 80 studies, it is found that voice interfaces are the most studied, followed by eye and head tracking. Other novel interfaces include brain control and facial expressions. Therefore, there is a need to improve and adapt hand interactions without the necessity of resorting to other parts of the body.

III. ADAPTED HAND INTERACTIONS FOR VR ENVIRONMENTS

In this section, we will present our proposal of adapted interaction mechanisms used in this case study. The prototype we based the following interactions on was developed for Meta Quest headsets. Consequently, we utilised the API and features provided by Meta's SDK version 59.9 to adapt such interactions. This choice was driven by Meta's position as a leading and successful company in the field of VR, alongside the extensive development and continuous updates of its API. Additionally, Meta devices feature hand tracking, enhancing the shopping experience by allowing users to use their hands, which, as demonstrated in Section II, is easier for individuals with upper limb motor difficulties compared to using controllers. Moreover, people without precise motor coordination cannot grab a controller and use the buttons, joystick or crosshead of it. In any case, the proposed adaptations are not dependent on a particular technology and are implementable on other VR headset models through their development API.

Furthermore, we would like to note that the adaptations were made with the aim of being as generic as possible. That is, whichever injury degree or upper limb part affected by such injury is, the adapted interaction should work for the user. In the following subsections, we will describe the interaction mechanisms with and without adaptions, highlighting the features included for its adaption. We have categorised the interactions mechanisms in the following categories: i) interaction with products, involving grasping and releasing products as well as adding or removing them to a shopping cart, ii) navigation, involving adapted teleportation and iii) virtual environment inspection, involving adapted rotation of the user's view.

A. VIRTUAL OBJECT GRASPING

In the context of VR applications, grasping or grabbing objects is a fundamental interaction mechanism that enables users to thoroughly explore them by rotating, scaling, or performing more complex actions, such as moving or detaching sub-parts. Like any other interaction implemented by Meta's API, grasping relies on the *Interactor-Interactable* framework. *Interactors* are added to input devices, in this case, the virtual hands—and *Interactables*, which are added to objects in the environment intended for interaction with the corresponding *Interactors*. This framework embedded in the Meta SDK allows these components to switch between four states: *Disabled*, *Normal*, *Hover* and *Select*. These states indicate whether a certain interaction defined by the Meta SDK can occur on a certain object configured as *interactable* for certain interaction.[2] Based on this, the interaction requires the hand to be near the object with the *Interactable component*



FIGURE 1. A set of screenshots showing the grab interaction: (A) approaching the hand near to a grab Interactable object (Hover state), (B) the virtual hand grabbing an object, along with the hand bones mapping, (C) performing a two hand grab and (D) scaling an object by using the two hand grab.

and other necessary elements, such as a *Collider*,[3] and to perform a *pinch* gesture grip or a palmar grip. The latter involves bringing the fingertips towards the palm. Grasping an object can be achieved using either the *pinch* grip, the palmar grip, or both. While the *pinch* is detected through an algorithm of the API, which takes into account grip strength and the fingers involved based on *hand-tracking* data and 3D coordinates, the palmar grip and its strength are based on the curl values of the fingers to determine if there is a grip.

To release a grasped object, certain distance thresholds values must be met for the API to recognise that the grip is no longer being performed. The distance values are calculated taking into account the 3D coordinates of the *Collider* bounds of the object and the 3D coordinates of the fingers used to grab the object. Generally, it is recommended to fully extend the fingers to facilitate this detection and release the grasped object.

1) BASE INTERACTION

Having described how to grasp and release an object, we will now illustrate an unadapted grasp. Figure 1.A shows a virtual hand of the user approaching an object to grasp it. Upon contacting the boundaries of its Collider, the user's hand lights up, providing visual feedback to indicate that performing a *pinch* or *palmar* gesture will grasp the object.

Furthermore, Figure 1.B illustrates the mapping of the hand bones onto the virtual hand grasping an object. A visual component rendering the hand's skeleton has been added to show that, even when performing a *pinch* gesture to grasp,

[2] More information in the official documentation: https://developer.oculus.com/documentation/unity/unity-isdk-interactor-interactable-lifecycle/.

[3] A component from the Unity's physics engine that defines a surrounding shape over the object to allow physics interactions. A component, in the context of Unity, can be comprised of configurable elements provided by the game engine, such as Colliders, Cameras or Materials, or C# scripts.

the API allows modifying the visual form of the hand to fit a specific shape, enhancing immersion This visual form must be recorded for the object.

Additionally, Figure 1.C and Figure 1.D demonstrate a two-handed grasp, which allows for scaling an object. While Figure 1.C shows that when grasping the ball, the grasp is indicated by the other hand lighting up in blue, Figure 1.D shows wireless earbuds being scaled using this grasp. Following this, we will show the adapted grab interactions.

2) ADAPTED INTERACTION

Grasping by pinching, which involves bringing the index finger and thumb together and exerting force, assumes neuromuscular integrity and fine coordination that may be severely compromised in individuals with SCI or those who have suffered a stroke. In patients with SCI, depending on the level and severity of the injury, there can be dysfunction of both upper and lower motor neurons, resulting in varying degrees of paralysis and spasticity. This affects the ability to perform precise movements, such as the pinching gesture, due to a loss of muscle strength, fine motor control, and coordination. Additionally, patients who have experienced a stroke often face additional challenges such as hemiparesis or hemiplegia, which results in weakness or paralysis on one side of the body, further complicating the ability to manipulate small objects and perform precise hand gestures. Moreover, some users might not be able to keep the object grabbed as the force exerted might not be the needed to maintain the grasp.

In order to address these limitations, object grasping has been modified using two different approaches: i) by adjusting Meta's API script and ii) through a C# script developed by the authors. The modification involves extending an interface to detect if a given finger is grabbing as well as scoring the strength of such pinch (see Algorithm 1). The function to get the score of the pinch grab can be further used with other finger joints to adapt other grabs such as palmar grab, or define a new one. This adaptation can assist individuals with limited finger mobility in performing grips through the movements they are capable of making, by not requiring them a full pinch. The float values assigned to the given parameters can potentially be identical across multiple parameters, streamlining the configuration process while maintaining functionality. These adaptations coexist, so that if a user with SCI can perform a grasp to some degree, i) will apply. However, if he/she cannot, due to proximity and time, the self-grasp ii) will be performed. In the absence of haptic gloves or any other tool that allows us to measure the force with sensors, Algorithm 1 allows to modify grab detection by relying on the inverse of the Euclidean distance between joint positions. A grab is detected when the computed strength exceeds the grab threshold G but remains below the maximum allowed threshold T. This ensures both proximity and validity for detecting the intention of grabbing an object. Those threshold values were set according the range of values managed by the Meta API and the help of HNPT therapists by looking how the values were mapped to joint positions when a pinch was being performed.

Algorithm 1 Detect Pinch Grab and Score Strength

Data: Vector3 position of Joint 1 *pos1*, Vector3 position of Joint 2 *pos2*, strength threshold T, grab threshold G
Result: Boolean indicating if a pinch grab is detected, strength of the pinch grab

function ScorePinchStrength(*pos1*, *pos2*)
    $distance \leftarrow \text{EuclideanDistance}(pos1, pos2)$
    Return $\frac{1}{distance}$
end function

function DetectPinchGrab(*pos1*, *pos2*, T, G)
    $strength \leftarrow \text{ScorePinchstrength}(pos1, pos2)$
    if $strength \geq G$ & $strength \leq T$ **then**
        $isPinchGrab \leftarrow$ true
    else
        $isPinchGrab \leftarrow$ false
    end if
    Return (*isPinchGrab*, *strength*)
end function

However, these modifications may not be feasible for individuals with significantly reduced hand mobility. Therefore, a developed script (see Algorithm 2) detects when the hand is over the surface of the object, and after one second, an automatic grip is performed. Although this may seem straightforward, the *Interactor-Interactable* lifecycle must be considered, requiring the states of the *Interactors* and *Interactables* to be checked to execute the automatic grip.

Automatic grasping with both hands does not happen independently in code; instead, it occurs simultaneously when both hands grasp the object automatically, following the *Interactor-Interactable* lifecycle. The algorithm makes use of specific functionalities of the Meta SDK, e.g. checking the state of the hand grab Interactor attached to hand H, forcing the grip or checking if the hand is hovering over an object. On the other hand, Update is a native function of the Unity engine that runs once per frame of the application, which allows us to execute a functionality throughout the execution of the application. Following this lifecycle ensures that functionalities like adapting the visual form of the virtual hand during a grip are retained, as shown in Fig. 2.B.

Figure 2.A displays the state of the hand before initiating a grip, highlighted by a blue outline, which is the same visual feedback in the unadapted interaction. Once the grip is performed, the outline turns green as shown in Figure 2.B. Meanwhile, Figure 2.C illustrates the distance grip implemented, based on *raycasting*. This ray is rendered with the LineRenderer component, which belongs to a Unity library. In addition, Unity's physics engine allows to detect

Algorithm 2 Auto-Grab of Grabbable Objects

```
Data: Required time to grab r, timers t1 and t2 that determine whether or not perform an automatic grab, object references that control the progressive rendering of the feedback circle P1 and P2, Interactors H1 and H2 (representing left and right hand grab interactors), boolean that allows auto-grab a
Result: The grabbable object O is grabbed by one or both of the user's hand
function AutoGrabHandler(H,t,P,a,r)
    if a is true then
        if H is not grabbing then
            a ← false
        end if
    else
        if HandIsHovering(H) then
            Set P.pointing ← false
            t ← t + Time.deltaTime
            if t ≥ r then
                H.ForceSelect(H.Candidate, true)
                t ← 0
                a ← false
            end if
        else
            t ← 0
        end if
    end if
end function
while Update is called do
    AutoGrabHandler(H1, t1, P1, a1, r)
    AutoGrabHandler(H2, t2, P2, a2, r)
end while
```

the collision of the ray with the objects marked as attractable to start the grabbing process at a distance. The ray originates from the carpometacarpal joint using the forward direction of the object, taking into account its current orientation in the world space. After pointing at an object for a second and a half at a certain minimum distance, beyond the reach of a stretched arm, the object will move towards the user's hand and perform a grip, as shown on the right side of Figure 2.D. The ray originates from the metacarpophalangeal (MCP) joint of the index finger.

Finally, we address the release of a grasped object. Our proposed solution involves pointing to an object within the virtual environment, which follows the user and maintains a fixed distance and orientation for easy location. In this case, we use the recycling bin metaphor to release the hand from the selected object. This solution allows users to release grasped objects more precisely, acknowledging that not all users can exert sufficient force with their fingers for the hand tracking system to recognise a released grip. The left image of Fig. 2.D shows the ray that, after pointing at the bin for one second, will release the object and it will automatically return to its original position.



FIGURE 2. A set of screenshots showing the adapted grab interaction: (A) auto-grab by approaching the hand near to a grab Interactable object (Hover state), (B) an object being grabbed (Select state) after one second has passed, (C) an user aiming with its right hand to perform a distant grab and (D) automatic release of objects by pointing to a bin.

B. VIRTUAL ENVIRONMENT NAVIGATION

Locomotion in VR environments has been and is of interest in the field of VR research. One of the main limitations to proposing navigation methods that do not involve physical movement of the user, i.e. physically moving around, is motion sickness. In large virtual spaces, it is difficult to have a physical space of the same dimensions, so the need arises to offer the user another form of locomotion. This can be through the movement of the joysticks of the controllers, which causes more motion sickness, or through teleportation, which causes hardly any motion sickness and is more enjoyable for the user than other techniques. These facts were demonstrated by Puritat et al. [34] in their case study of a large virtual space in a museum.

Therefore, teleportation is one of the most popular and frequently employed navigation methods in VR, offering a fast and convenient way to move within the environment while minimising motion sickness [36]. Despite this, teleportation can be very difficult to perform by potential users, who we intend to help with adapted hand interactions. Firstly, because the use of controllers may be too difficult for them, as they require a certain amount of upper limb mobility and hand shape in order to be able to use buttons or joysticks included. Secondly, developers of VR SDKs often define a default hand shape to invoke this interaction so as not to conflict with other interactions or gestures. This is why we have modified, in the case of the Meta SDK, the shape and the way of activating this interaction with the hands so that it can be accessible.

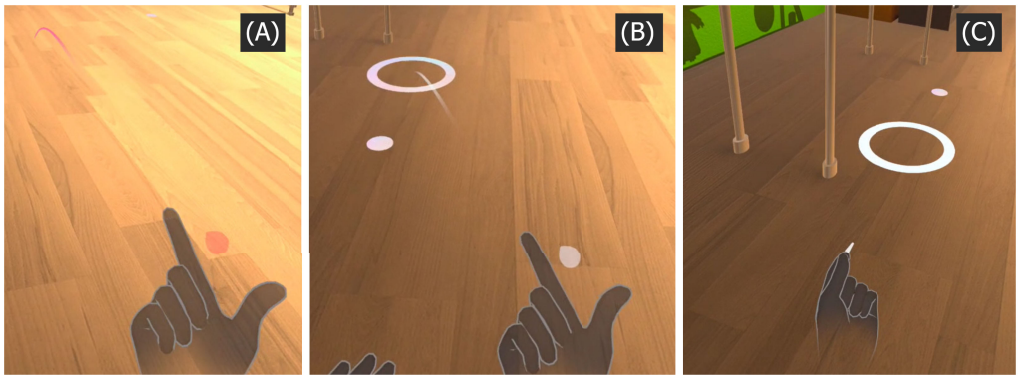


FIGURE 3. A set of screenshots showing the base teleportation interaction: (A) pointing to a place where teleportation is not allowed (Normal state), (B) pointing to a teleportation point, where teleportation is allowed (Hover state) and (C) performing the pinch gesture to activate the teleportation (Select state).

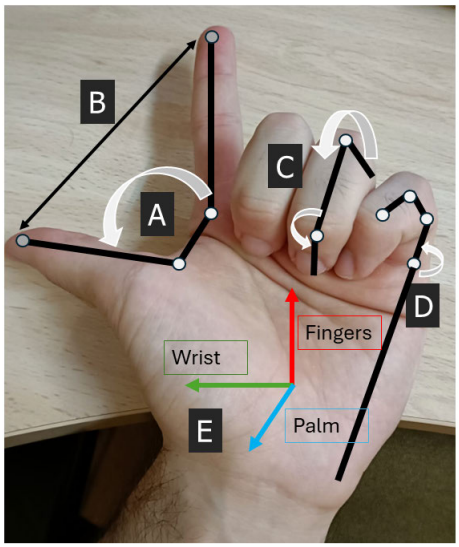


FIGURE 4. The L-shape with a left hand that enables the teleportation arc. The finger features of Meta's SDK are as following: (A) *Abduction*, (B) *Opposition*, (C) *Curl* and (D) *Flexion*. (E) represents the world coordinate axes that allow transform recognition of the hand's position in 3D space.

1) BASE INTERACTION

To be able to perform a teleportation to some point in the virtual space, Meta's API requires the configuration of the relevant objects in the Unity scene with appropriate scripts, as well as the identification of a *hand pose*. The concept of *hand pose* is essential for adapting these component interactions:

- **Shape recognition**. This is based on a set of boolean conditions regarding the position of one or more fingers. When the data collected by *hand tracking* satisfies these conditions, the shape is "activated". The boolean conditions are derived from comparing the data obtained from *hand tracking* with defined *finger features*,[4] which specify finger positions to define a shape. These features include *curl*, *flexion*, *abduction*, and *opposition*. Thresholds are established to determine the transition from one feature to another.
- **Transform recognition**. The hand transform represents the orientation and position of the hand in a 3D space. This recognition is based on boolean conditions that must be met concerning the fingers, wrist, and/or palm. For instance, it can be verified that the palm is orientated upward or downward, or that the fingers or wrist are oriented upward, among other possibilities.

With regards **shape recognition**, Figure 4 shows the different finger features of the L-shape that are needed to invoke the teleportation arc. We provide a description of the used finger features for each finger and the values taken to detect this L-shape with the hand. For further information on other possible values and their meaning, please refer to the official documentation.

Figure 4.A shows *abduction*, which is the angle between two adjacent fingers, measured at the base of those two. In this case, it measures the angle between the thumb and its adjacent that is closer to the pinky, i.e., the index finger. In this case, the value taken is *Open* as the two fingers are spread apart. Figure 4.B illustrates *opposition*, described as how close a given fingertip is to the thumb tip. In this case, the value taken is *Near* which means that the fingertip joints are between 1.5 cm and 15 cm apart. Figure 4.C shows the curl performed

[4]https://developer.oculus.com/documentation/unity/unity-isdk-hand-pose-detection/#finger-features

with the ring finger, which quantifies how bent the top two joints, represented with white circles, are. In this case, the value taken is *closed*, which means that the fingers are tightly curled inwards with the tips almost touching the palm or doing it. Finally, Figure 4.D represents the *flexion*, which is the extent that the Proximal joint is bent relative to the palm. In this case, the value taken is *Closed* as the joint is fully bent, almost perpendicular to the palm.

Figure 3 illustrates the hand poses required to initiate a teleportation or rotation interaction. Figure 3.A shows the a hand making a shape similar to an L facing upwards, which invokes the arc that allows the user to aim where the user wants to teleport. However, the arc is shown in red as teleportation is not allowed where the user is aiming. Nevertheless, Figure 3.B shows that the colour changes to white, as well as the teleportation point aimed at is enlarged when it is pointed. Finally, Figure 3.C the action that activates teleportation when pointed to a place that allows teleportation, which involves making the *pinch* gesture.

2) ADAPTED INTERACTION

Meta's SDK offers an API that allowed us to use a single component of *transform recognition* (see Section III-B1) to activate the arc to point where the user wants to teleport.

The basic interaction defined for teleportation in a VR space using the L-shaped hand gesture, with the index and thumb fully extended, is unrealistic for patients with claw hand deformities. Therefore, it is necessary to find alternative solutions. In this case, we have decided that the user only needs to point to the desired teleportation site for one second and a half, as illustrated in Figure 5. In this case, the shape of the user's fingers is not considered; they only need to point the palm of the hand upward to invoke the arc. This allows for generic adaptation to a wide spectrum of injuries, as we removed the requirement of performing a concrete shape with the hand.

To avoid errors when pointing, the environment has been marked with small circles indicating teleport points. These points will cause the arc to snap to the indicated location, allowing for some aiming error rather than only allowing that single circle. This is achieved by providing the teleportation point a wider *Collider* area. We also added visual feedback in the form of a loading circle, which progresses until completion as the pointing time advances, preventing users from stopping too early and having to restart. When the circle is 100% complete, which means that the user has kept their hand in the same position for a certain amount of time, the user is then teleported to the chosen destination point. The following Google Drive link contains a video shows the adaptation interaction to teleportation as we have described it https://github.com/AIR-Research-Group-UCLM/VR-Hands-Interactions-Adapted/blob/master/video_link.txt.

The algorithm 3 shows the pseudocode of the Update function that controls whether to teleport or not a user to a given destination point based on pointing time over a certain teleportation point. The *Select* and *Unselect* events are manually sent in the code to not alter the *Interactor-Interactable* lifecycle, since the adaptation is based on time rather than on user inputs through hand-tracking detection.

Algorithm 3 Auto-Teleportation

```
Data: Boolean h: Indicates if the teleportation arc is hitting a teleportation point. Teleport Interactor tpi: Manages teleportation interactions. Timer t: Tracks hover time over a teleportation point. Time required for teleportation r: Hover time required for teleportation. Feedback manager ipm: Controls the visual feedback progression (circle). Boolean wp: Tracks if the teleportation arc was hitting a teleportation point in the previous frame.
Result: The user being teleported to a certain tp point that was aimed to over r seconds
while Update is called do
    if tpi.State is Active then
        if tpi.State is not Select and h = true then
            t ← t + Time.DeltaTime
            if notwp then
                ipm.SetPointing(true)
                wp ← true
            end if
            if t ≥ r then
                ipm.SetPointing(false)
                wp ← false
                 Send event of Select occurred
                t ← 0
            end if
        else
            ipm.SetPointing(false)
             Send event of Unselect occurred
            if wp then
                wp ← false
                ipm.SetPointing(false)
            end if
        end if
    end if
end while
```

C. VIRTUAL ENVIRONMENT INSPECTION

Another functionality we are adapting is body rotation in virtual spaces. This feature allows the user's view to be rotated by a specific angle relative to the vertical world axis of the scene, thereby simulating a turn without requiring physical rotation. Although this option may appear less essential for an person/people without disabilities, it is crucial for users who have experienced a stroke or other incidents that require the use of a wheelchair, as it facilitates the exploration of a virtual space 360 degrees.

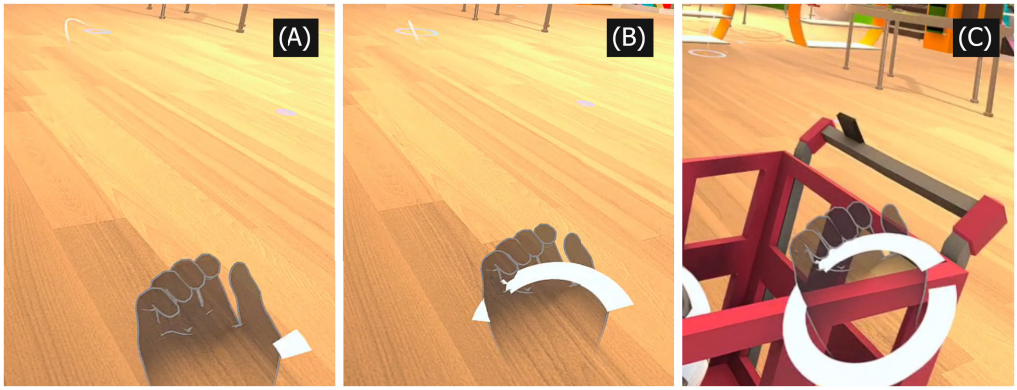


FIGURE 5. Screenshots that show the adapted teleportation interaction: (A) pose pointing to a place where teleportation is allowed (Normal state), (B) the adapted teleportation shows a circle that progressively fills as the user keeps pointing to the teleportation point and (C) shows the circle almost filled completely before teleporting the user.

1) BASE INTERACTION

The Meta SDK combines the invocation of two interactions, the teleportation we have shown in the previous section, and the rotation of the virtual body, depending on a certain condition. This combination is implemented with a series of scripts that are ultimately based on the concept of gates. Thus, when a condition is met on a given variable, the teleportation arc is invoked. If it is desired to invoke the virtual body rotation interaction, the user will modify these values by moving his hands, and this data is collected by the hand-tracking technology. These variables are a minimum and maximum degrees of wrist rotation angle, which determines which interaction should be invoked in order to be performed, as long as the hand shape that the hand-tracking checks these angles, i.e. the L-shape, has been detected.

In the case of the base interaction, the wrist must be rotated in order to make the palm of the hand point towards the ground. Once a certain angle of rotation is reached, arrows will be invoked that can be selected by the user to rotate 45 degrees with respect to the vertical axis in the direction marked by these arrows.

Therefore, Figure 6.A shows the arrows being invoked. The selection of rotation towards one direction or the other consists of performing the *pinch* gesture on one of the arrows.

2) ADAPTED INTERACTION

As we have already shown in Section III-B2, the restriction imposed by Meta's SDK of combining the wrist turning angle to determine whether to activate the teleportation arc or the rotation arrows, we have chosen to recognise the hand position with the palm facing upwards, avoiding the need to flex or extend the fingers. In the case of virtual body rotation, it still requires turning the wrist a certain amount of degrees to invoke either the left arrow (see Figure 6.B) or the right arrow (see Figure 6.C). However, we are aware that the minimum and maximum wrist rotation angles may vary depending on the user and his or her limitations, as well as the difficulty of performing the *pinch* to select to perform the interaction.

Therefore, we have implemented two features to avoid these problems. First, we have modified how to select the arrows. Instead of having to perform a *pinch* over them, the action required is moving the palm's face towards the direction of the arrow. This was achieved by i) implementing a **transform recognition** component that states the orientation of the wrist and the fingers as shown in Figure 4.E, without considering the shape, and ii) using a component of Meta's SDK that allows activating a state based on the rotation performed with a joint over a span of time. The second step used a joint named "Hand Start" that represents the wrist, and the minimum time to perform the rotation was set to 0.1 seconds. This allows for slow and faster movements.

While our first goal was to adapt the way the interaction was used, the second one focused on the invocation of the virtual elements that allow to perform teleportation and virtual head turning. Based on each user wrist mobility limitation, the invocation of the arrows and the teleport arc were adapted. Therefore, we develop a *script* that collects the minimum and maximum wrist rotation angle between 0 (palm face towards ceiling) and 180 (palm face towards floor) degrees. Algorithm 4 describes the logic behind this feature. This pseudocode heavily relies in the Vector3 class provided by the Unity engine, which represents a point in 3D world coordinates of the application. In order to provide further explanation on the variables used, a brief description of them is provided. While *minAngle* and *maxAngle* represent the range of angles that can be detected from wrist rotation to invoke the arrows or teleportation arc, *wrist.angle* represents the current rotation's angle of the wrist,

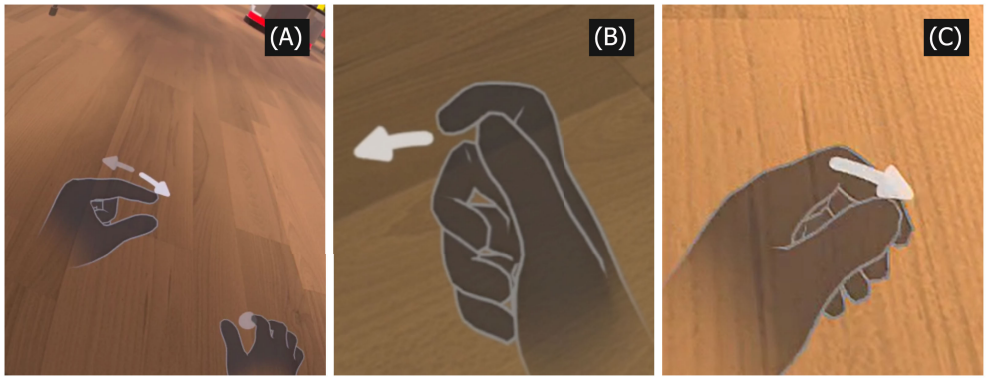


FIGURE 6. Screenshots that shows the invocation of the arrows that allow virtual body rotation. (A) shows the base interaction, where a *pinch* over the arrow is required, (B) illustrates the adapted interaction with the left direction (right hand) and (C) shows the adapted interaction with right direction (left hand).

Algorithm 4 Adaptation of the Invocation of Teleportation Arc and Turning Arrows

Data: boolean that checks if the handedness of the hand is right *rh*, a Vector3 that represents the up direction (0,1,0) in world coordinates *up*, a hand Pose object that contains the rotation and position of the hand obtained from hand-tracking data *HP*, the shoulder of the same body's half that the hand *sh*, boolean that controls if wrist rotation data is being collected *ic*, a public variable that stores the wrist rotation angle *angle*, an object that holds the script with the function HandleHandUpdated *wrist*, two floats that hold the minimum and maximum angles of wrist rotation *minAngle* and *maxAngle*.
Result: The rotation angle of the wrist adjusted to the handedness in every frame
function HandleHandUpdated
    Vector3 shoulderToHand ← (*HP*.position - *sh*.position).normalized
    Vector3 trackingRight ← Vector3.Cross(*up*, shoulderToHand).normalized
    trackingRight ← *rh* ? trackingRight: -trackingRight
    Vector3 wristDir ← *rh* ? -*HP*.forward: *HP*.forward
    wirstDir ← Vector3.ProjectOnPlane(wirstDir, houlderToHand).normalized
    wrist.angle ← Vector3.SignedAngle(wristDir, trackingRight, shoulderToHand)
    wrist.angle ← *rh* ? -angle: angle;
end function
while Update is called **do**
    if ic **then**
        float currentAngle ← *wrist.angle*
        currentAngle ← Mathf.Clamp(currentAngle, 0, 180)
        minAngle ← Mathf.Min(minAngle,currentAngle)
        maxAngle ← Mathf.Min(maxAngle,currentAngle)
        update the framecount to check if it has already passed 7 seconds
    end if
end while

which is further used in the Meta SDK. *ShoulderToHand* holds a vector with the direction of the shoulder to the hand, while *trackingRight* holds a vector that points to the right side relative to the direction of *shoudlerToHand*. *WristDir* holds the direction vector of the wrist, which is later reassigned to project the wrist direction onto a plane perpendicular to the direction pointed by *shoulderToHand*. An image that might help to understand the relationship between these vectors can be found in the following link https://github.com/AIR-Research-Group-UCLM/VR-Hands-Interactions-Adapted/blob/master/vectors.png. The function *HandleHandUpdated* is a callback function from one of the Meta SDK's library that we modified as shown in Algorithm 4. We would like to note that this pseudocode involves adjusting the sign of the variables, as well as the final angle, according to the hand where the script is attached.

IV. CASE STUDY: ADAPTED INTERACTIONS IN VR SHOPPING APPLICATIONS

To evaluate the suitability of the proposed adapted hand interactions for individuals with upper limb motor impairments, a case study was designed and executed, which will be presented in this section. For this purpose, a VR Shopping application prototype was developed using Unity 2022.3.15f1, as it allows us to evaluate the various proposed adaptations, as well as advance research on accessibility options in this type of virtual environment.

The two experimental sessions, conducted at the HNPT, involved a total of twelve participants (see Table 1), comprising ten patients and two healthy individuals (HNPT staff). All the patients had a chronic cervical SCI (more than one year after the injury). The sessions were conducted in two different days, spanning a duration of five hours each. The inclusion of two healthy participants was not meant for statistical comparison with the group of ten users with disabilities, but rather to ensure usability, validate feasibility, and obtain expert opinions on accessibility adaptations.

The primary objective of the experimental sessions was to conduct a functional analysis. This analysis enabled an

early assessment of the tool's safety and efficacy, ensuring that potential risks were identified and mitigated before expanding the patient pool. It also provided an opportunity to identify and resolve any technical or operational challenges in a controlled setting, which is crucial for successful project scaling. Moreover, this approach facilitated the validation of research methodologies, including data analysis protocols, ensuring their robustness and reliability for a large-scale study.

In addition to the functional analysis, other objectives of the experimental session included assessing the well-being of users in the VR environment and gathering their personal feedback on its application in rehabilitation. This involved understanding their comfort levels and subjective opinions about the tool's role in their recovery process. In addition, our goal was to track any improvements in exercise performance during the course of the session, focusing on how quickly and effectively users could adapt to the exercises in the virtual setting.

A. RESEARCH QUESTIONS

The case study aims to answer the following research questions:

1) (RQ1) Do adapted interactions enable people with motor impairments to perform tasks in immersive spaces?
2) (RQ2) Do adapted interactions improve the performance of a person with motor impairment when performing tasks in immersive spaces?
3) (RQ3) Do users positively evaluate adapted interactions in terms of usability and user experience?

To address these questions, data was collected during the execution of two experimental sessions with 10 patients from HNPT, a biomechanical engineer and a therapist. These data were obtained from non-invasive sources for the participants and can be classified as: i) interaction data, ii) head and hand tracking data, and iii) positional and kinematic data. In addition, a questionnaire was prepared for the participants to evaluate their experience based on various standardised questionnaires. We compiled questions from NASA-TLX [16] to assess workload, the System Usability Scale (SUS) [6] to determine usability of adaptations, the User Experience Questionnaire (UEQ) [39] to collect feedback on user experience, and the Motion Sickness Assessment Questionnaire (MSAQ) [14] to measure the degree of motion sickness caused.

The next section describes the methodology of the study carried out.

B. DESCRIPTION OF THE VOLUNTEERS

For the execution of the case study, we collaborated with HNPT patients and staff. All participants gave us their consent to use the data obtained during the experimentation, as well as to participate in the case study. Table 2 shows the mean and standard deviation (SD) of various parameters to characterise the degree of neurological injury of the study participants. The use of average data to characterise patients is the standard in clinical studies. Patients had a SCI between the metameric levels C2 (cervical injury) and D8 (dorsal injury), and neurological classification according AIS (ASIA Impairment Scale) [24] between A (complete lesion in motor and sensitive aspects) and D (incomplete injury in the motor and sensitive aspects). The right arm was dominant in 90% of the patients, and was the treated arm in 70% of them. The UEMS (Upper Extremity Motor Score) for the treated arm was 14.14 (SD = 8.62), reflecting the impairment in upper extremity function, as assessed through the evaluation of five specific upper extremity muscles. The maximum score, 25 points, is reached for neurologically healthy subjects. By means of the SCIM scale (Spinal Cord Independence

TABLE 1. Volunteers' age and relationship with HNPT.

User	Age	Status within HNPT
U1	14	Patient
U2	28	Patient
U3	34	Patient
U4	21	Patient
U5	42	Patient
U6	48	Patient
U7	19	Patient
U8	54	Patient
U9	22	Patient
U10	48	Patient
U11	44	HNPT's Biomechanical Engineer
U12	38	HNPT's therapist

TABLE 2. Characterisation of the degree of neurological injury of the case study participants.

	SCI Patients (n=10)	Healthy subjects (n=2)
	Mean (SD)	Mean (SD)
Male	10.00 (100.00)	1.00 (50.00)
Age	32.00 (13.22)	41.00 (4.24)
Injury level	C2-D8	-
AIS scale	A-D	-
Dominant arm (Right)	8.00 (80.00)	2.00 (100.00)
Treated arm (Right)	7.00 (70.00)	2.00 (100.00)
UEMS treated arm	14.14 (8.62)	25.00 (0.00)
SCIM (0-100)	48.00 (25.25)	100.00 (0.00)



FIGURE 7. Photographs taken during experimental sessions at the National Hospital for Paraplegics with patients suffering from spinal cord injuries and having motor limitations in legs, arms, and hands.

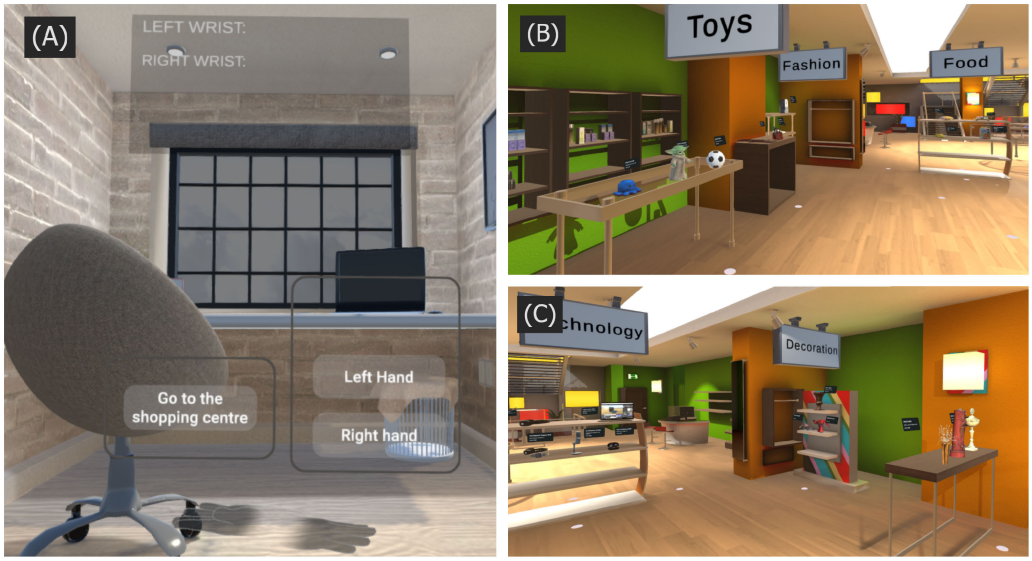


FIGURE 8. Two screenshots of the environments used in the case study. (A) shows the environment where wrist rotation angles were obtained, B and C shows two points of view of the environment used for the case study.

Measure) [7] was used to evaluate the independence level of the patients, scored between 0 and 100. In this scale a higher score corresponds to a higher degree of independence. In combination with this characterisation, the knowledge and experience of the HNPT staff was of great help in designing the tailored interactions, as it gave us a more complete picture of the patients' limitations. Thus, we always had in mind to develop the adaptations as generically as possible so that any patient could use them regardless of whether they had more or fewer limitations. In the case of the patients, we asked which of their upper limbs was most affected by their motor impairments (treated arm in Table 2). This information was crucial, as they were later asked to use their most affected limb during VR environment sessions with adapted interactions. In the case of participant U4, the patient expressed that both limbs were similarly impaired.

C. METHODOLOGY

To determine whether the adapted interaction mechanisms help individuals with motor issues complete the proposed tasks and whether there is an improvement in their performance, the following experimental session flow was designed:

1) Each volunteer was shown the environment in which he or she would be immersed during the session in a video, with a particular emphasis on the products and sections relevant to the tasks.
2) Next, the users were introduced to the different ways of interacting in the non-adapted environment, and they were given between 1 to 2 minutes of prior training in an environment different from the one used in the experiment. It was emphasized that the non-adapted interactions could be difficult to perform, and they were encouraged not to become overly frustrated.
3) After this, they began performing the tasks listed in Table 3. One of the authors guided them, indicating the objective of the next task after completing the previous one.
4) Once this first execution was completed, the adapted interactions and how to perform them were introduced. In this case, no prior training was provided, to later assess how simple and user-friendly the adapted interactions were, considering they were already familiar with the environment.
5) Before the execution of the adapted environment, Algorithm 4 was applied in an additional scene created with Unity (see Fig. 8.A) to utilize the wrist rotation angles obtained in the case study.
6) During the execution in the adapted environment, the same author guided the participants through the same tasks in the same order.
7) After both executions, the participants were asked to complete the questionnaire prepared for the study.

To avoid causing excessive fatigue for the volunteers, the following measures were taken. First, the total experimental time for each participant was limited, with an ideal duration between 25 and 30 minutes to explain, conduct the training in non-adapted interactions, experiment with both environments, and complete the questionnaire. Based on our previous experience and the knowledge of the hospital staff,

experimental sessions of longer duration cause considerable mental and physical fatigue in patients, causing them to lose concentration and provoking a desire to finish as soon as possible. Second, in the event of a participant abandoning a task in the non-adapted environment, such as navigating through the environment, one of the authors would intervene to position them for the next task. In these cases, the task was marked as failed or incomplete. Volunteers were encouraged to prioritize their health and avoid overexertion or compensating with unaffected parts of their body.

After formulating the research questions and determining approximate time limits for the tasks based on previous data, the tasks described in Table 3 were proposed for the case study participants. The fourth link of this. txt file from tout Github repository shows one of the authors conducting each of the tasks described in the table: https://github.com/AIR-Research-Group-UCLM/VR-Hands-Interactions-Adapted/blob/master/video_link.txt. All tasks, except the last one, were guided, meaning that the section and product were predefined for the participant. This decision was made to allow for more precise and objective comparisons of results, given that the same situation would be presented to all volunteers across both environments, as will be discussed below. Even though it may seem that the number of tasks is not balanced according to the number of adapted interactions shown, note that grabbing tasks inherently require movement within the environment. Therefore, teleportation and virtual body rotation are implicit actions needed to complete such tasks.

We did not consider adding the removal of a product from the shopping cart since the user's position and movements are practically the same as when adding a product. Additionally, we must take into account the physical and mental load that long experimental sessions impose on patients with spinal cord injuries. The fatigue experienced when performing an activity is greater than that of a healthy person, especially in muscles such a the elbow flexors [40].

D. MATERIALS AND SOFTWARE USED

The materials used for the study were as follows:

1) MSI Raider GE68 HX 13V laptop with the following hardware: Intel i7-13700HX @ 2.1GHz, 32 GB RAM Memory, NVIDIA RTX 4060 8GB, Windows 11 64-bit.
2) Meta Quest 3 headset.
3) Optic fibre Link Cable for Meta Quest.

On the other hand, the software employed was as follows:

1) Unity 2022.3.15f1 for the design and development of the applications used in the experimental sessions.
2) Meta XR All-in-One SDK v59.0 to integrate the Unity application with the Meta headset.
3) Meta Quest Link application for PC to run the application during the experimental sessions.
4) Microsoft Forms for creating the questionnaire and collecting responses.

E. CASE STUDY LIMITATIONS

The sample of volunteers with motor difficulties may seem small. However, we must consider the challenges associated with recruiting individuals with these types of injuries for experimental sessions. In addition to travelling to HNPT, we must consider the patients' tight schedules at the hospital, as rehabilitation and physiotherapy activities often consume much of their time. Furthermore, we must account for the dependency of these individuals, which means that neither them nor their caregivers are always available. We also decided to include two biomechanical engineers in the sessions, who work regularly with these patients, to provide feedback on the adaptations. However, the results presented in Section V show a clear trend among patients with SCI, and thus we consider the sample sufficient to answer the research questions posed.

Another limitation to consider is that this case study does not conduct long-term follow-up. Nevertheless, in the context of this work, as shown in Section V, it can be seen that the adapted interactions helped the volunteers perform tasks they could not previously accomplish or improve their performance in executing them. This case is not the same as one that monitors the long-term influence on a patient's recovery or rehabilitation.

The VR application was developed using Unity and the Meta SDK for the Meta Quest headset line. This decision was driven by three main factors: (i) the authors' experience in developing VR applications with this SDK, (ii) Meta's leading position in the VR device market,[5] with a substantial advantage over competitors, enabling it to reach a larger potential user base, and (iii) the Meta SDK provides direct integration with hardware-specific features of the Meta Quest headsets that are particularly relevant for the case study. These features include advanced hand tracking, allowing precise and high-frequency tracking, as patients may suffer trembling during the usage of the application, causing undesired outcomes if the tracking is not precise enough. Additionally, the abstraction layers and the ability to modify the interaction mechanisms provided by the SDK led to the decision to use the Meta Quest 3 for the case study. While this may initially seem like a significant limitation, the Meta SDK is compatible with the OpenXR standard, facilitating the potential future porting of the application to this standard.

Lastly, one could consider a limitation the fact that the same virtual environment was used for both interactions. However, this allows us to objectively and directly compare the actions taken to complete a given task, whether it was completed or not, as well as the time spent on it.

V. RESULTS

In this section, we will present the results obtained from the analysis of the data extracted from the experimentation, as well as the responses to the questionnaire.

[5]https://techwireasia.com/2023/07/meta-headset-sales-outperform-other-brands-in-2023/

TABLE 3. Tasks to be performed by volunteers of the case study. In the case of Task #6, 4 more seconds are added to compensate the user movement of reaching an object that is placed high in the shelf.

Task ID	Description	Minimum number of interactions required	Consideration for the adapted interaction
#1	Move to the "Toys" section	It requires 1 teleportation	-
#2	Grab the Baby Yoda stuffed toy and release it	Grab + moving to the shopping cart	Release made by pointing to the bin (see Fig. 2.D)
#3	Turn your body using the arrows until you see the table where the Burner is	Volunteer can turn using either left or right hand directions. 3 turns are required	Volunteer can turn using either left or right hand directions (see Fig. 6.B and C)
#4	Move to the "Decoration" section where the Burner is	It can be performed within 1 teleportation	Volunteer might need to turn itself in the virtual environment
#5	Grab the Burner and add it to the shopping cart	It requires grab and addition to shopping cart shopping cart	Volunteer might add it through distance or manually (see Section III-A2)
#6	Move to the "Fashion" section	It requires 2 turns and 2 teleportation	-
#7	Grab the Backpack and add it to the shopping cart	Grab in a high shelf + move to the shopping cart	Volunteer might not be able to grab it in the non-adapted environment
#8	Move again to the "Toys" section	It requires 1 turn and 2 teleportations	Volunteer might need to turn itself in the virtual environment
#9	Grab the Ball and add it to the shopping cart	Grab with both hands is required.	-
#10	Add to the shopping cart the product of your desire	Depends on what each user wants to add	Free exploration task. Volunteer cannot add a product from "Toys" section

A. QUANTITATIVE ANALYSIS

The quantitative analysis regarding the degree of improvement provided by the adapted interactions focuses on a study of effectiveness and an estimation of effort during task completion in the immersive space. Both studies help provide answers to Research Questions $RQ1$ and $RQ2$ at the end of the section.

1) EFFECTIVENESS

Table 4 shows the time, in seconds, that each user took to complete each task in both environments, unadapted and adapted (tasks described in Table 3). Cells highlighted with a gray background indicate that the user abandoned the task without completing it. The time reflected in these cells corresponds to the duration the user spent trying to perform the necessary interactions before abandoning it due to fatigue or their inability to complete the task because of motor limitations. The last row indicates the percentage of tasks completed in each VR environment by each user. Among the volunteers, it is worth noting that U2, U5, U6, U7, U8, and U10 were the ones who were least able to complete tasks in the unadapted environment due to their injuries and limited mobility.

It is important to point out that, in the case of the first three tasks in the unadapted environment, we asked users to attempt the tasks using their most impaired hand so that they could later assess the impact of the adapted interactions. Notably, tasks #4, #6, #8, and #10 took the most time in both environments, as they involved the most movement around the environment, either via teleportation or virtual body rotation. However, it is evident that users with injuries to both hands had a high abandonment rate, reflected in the tasks within the unadapted environment. The tasks that took the least time for users were those related to object grabbing and adding items to the shopping cart. In terms of effectiveness, we observe that all users were able to complete all tasks in the adapted environment, demonstrating that these interactions significantly improved the effectiveness of users with injuries in interacting with the environment.

Fig. 9 shows, for each user, the percentage of performance gain when performing the same task in the environment with adapted interactions compared to the same task in the unadapted environment. We can observe a general trend of performance improvement when using adapted interactions. Specifically, tasks involving movement through the virtual environment show increases of up to 91.6%. Tasks with lower increases, or even in some cases negative increases, are those related to grasping, particularly two-handed grasping. This may be due to the waiting time established for performing auto-grasp when the virtual hand comes into contact with the object. It is also important to consider that, for instance, in tasks #5 and #7, if the user opts for distance grasping, they must also wait a certain amount of time while pointing at the object, plus the time it takes for the object to move towards the hand to perform the auto-grasp. Finally, it is noteworthy that healthy users, U11 and U12, show worse performance in some tasks with the adapted interactions. Despite this, one possible reason could be the need for them to simulate having no hand mobility, in addition to the required waiting times for grasping and teleportation.

Taking these waiting times into account, we can perform a paired t-test taking into account the data collected from U1 to U10 (HNPT's patients) with the following hypotheses:

1) H_0: There is no significant difference in task completion times between using unadapted and adapted interactions.
2) H_1: There is a significant difference in task completion times between using unadapted and adapted interactions.

Thus, we obtain that $t - statistic = 5.75$ and $p - value = 7.06 \times 10^{-8}$. The former indicates a significant difference in

TABLE 4. Time in seconds spent by each user on every task in both Unadapted (U) and Adapted (A) environments. Cells with grey background mean that the user could not complete the task by itself. The time showed is the time they spent trying to carry out the task before abandoning it. The thicker vertical line is used to separate injured users (U1-U10) between healthy users (U11 and U12). C.R. % stands for Completion rate.

User	U1		U2		U3		U4		U5		U6		U7		U8		U9		U10		U11		U12	
Task	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A
#1	30.50	8.40	26.20	5.57	20.05	5.97	22.36	6.46	51.24	7.68	59.65	6.48	70.09	5.86	43.10	9.97	22.36	5.47	34.16	8.43	8.89	6.73	11.09	5.01
#2	13.50	6.50	8.99	7.46	13.46	7.97	12.07	7.01	9.40	9.12	10.75	7.82	12.90	9.97	10.13	6.90	12.97	7.39	12.33	7.26	4.29	7.37	6.39	8.15
#3	33.90	13.38	32.04	16.01	36.88	7.57	38.51	13.70	62.34	18.76	43.82	10.60	76.47	13.90	96.13	18.77	29.18	12.68	69.03	18.10	9.59	11.76	11.99	6.08
#4	22.93	10.97	30.39	6.19	43.16	8.33	16.36	12.69	73.04	22.48	22.64	15.05	28.37	16.39	74.80	13.03	23.47	8.68	64.53	19.42	5.17	5.71	7.24	6.43
#5	19.67	10.46	12.19	10.15	16.65	8.94	15.00	7.83	18.04	12.40	25.18	8.46	14.14	12.05	24.32	8.86	18.78	9.94	9.31	10.26	5.82	6.15	7.36	6.95
#6	36.96	13.51	42.89	10.87	36.03	23.81	27.37	18.90	46.19	21.30	58.87	22.80	26.44	16.29	105.51	22.01	22.18	20.50	71.67	28.47	10.33	10.04	14.42	12.03
#7	16.46	8.01	13.64	8.40	12.50	10.10	36.78	7.17	36.78	8.96	21.32	7.67	52.54	11.99	35.23	16.99	13.35	9.44	19.32	6.87	5.85	7.17	10.90	8.15
#8	24.87	18.61	43.37	21.40	45.08	30.35	30.00	16.47	55.52	27.37	46.93	20.54	25.37	18.10	46.23	33.32	43.22	20.73	26.56	23.36	10.72	15.43	17.65	18.43
#9	18.94	14.22	13.61	16.00	13.68	10.53	14.04	11.56	17.04	17.17	15.35	11.03	11.83	12.49	27.43	8.05	15.64	10.68	16.93	14.46	8.29	10.18	7.85	12.89
#10	34.14	28.56	48.88	18.96	28.00	20.89	36.94	23.29	49.04	35.58	75.45	30.43	49.79	42.40	64.38	37.19	44.99	28.40	57.93	44.47	20.00	14.90	21.01	24.50
C.R.	80%	100%	30%	100%	50%	100%	50%	100%	0%	100%	10%	100%	30%	100%	0%	100%	60%	100%	30%	100%	100%	100%	100%	100%



FIGURE 9. Bar graph per user showing the increased performance in the Adapted environment against the Unadapted environment.

terms of standard deviation between the times. Additionally, the latter is a value far below the common significance threshold of $\alpha = 0.05$. Given that the p-value is much smaller than the significance threshold $\alpha = 0.05$, we reject the null hypothesis H_0 and accept the alternative hypothesis H_1. This means that there is statistically significant evidence of a difference in task completion times when using adapted interactions compared to unadapted ones. The t-statistic also supports this conclusion, as such a high value indicates a significant difference between the means of the two groups.

2) EFFORT ESTIMATION

Having demonstrated that the adapted interactions helped users complete tasks that they could not perform before, and in less time, we will now estimate the effort they had to exert with both types of interaction. To do this, we will inspect the

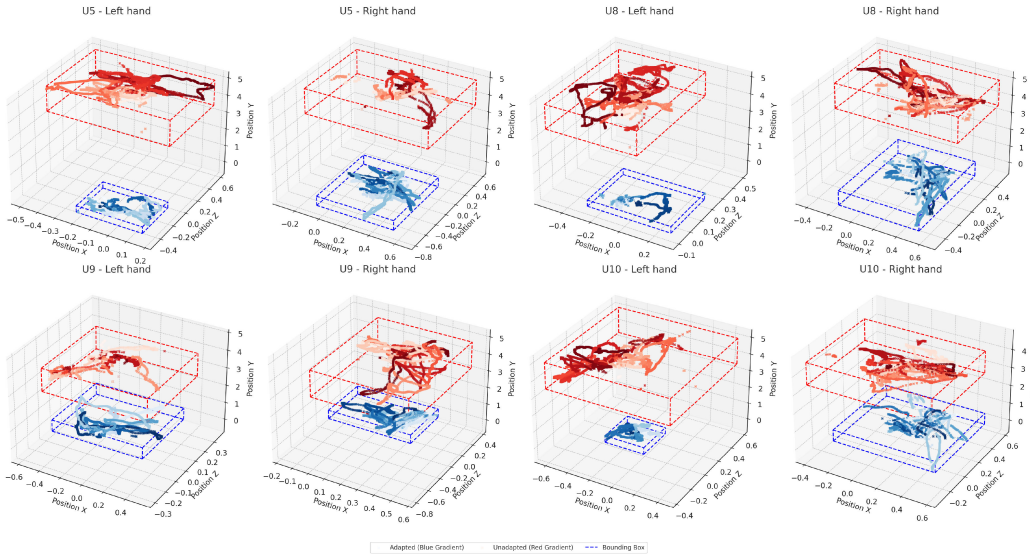


FIGURE 10. Hand movements made with the hands by some volunteers during the case study. These subgraphs include all the movements made for every task. These movements are translated in the Y axis to ease the readability of the movements and avoid superposition. The gradient in the colours indicates the passage of time: the colour starts very warm at the beginning of the user's session and progressively increases towards cooler colours. The bounding boxes encapsulate the range of movement of the hands by each of the users. The legend of the 3D charts indicate: Red colour - Unadapted, Blue colour - Adapted.

spatial movement data recorded during the experimentation phase.

Figure 10 shows the hand trajectories of some of the injured volunteers during the case study. A Z-score filtering has been used to remove outliers likely caused by tracking errors. As can be observed, each volunteer with injuries performs movements differently, depending on their motor capabilities and height. However, one general characteristic is visually apparent: the graphs representing the adapted interactions appear more compact, while those representing the unadapted interactions are broader and less condensed. This implies movements of greater amplitude and more erratic patterns, which could indicate higher effort and discomfort for users with spinal cord injuries. These movements may be due to the users attempting to perform the required interaction for a task by adjusting their limbs according to their degree of injury.

In an effort to quantify this apparent trend in the degree of effort exerted, we will attempt to estimate the difficulty faced by each user based on the spatial tracking data of the users' hands, the number of attempts made to complete each task, and the time taken to either complete or abandon each task, as previously presented in Section V-A1. Having described in Section II the works in literature to evaluate physical effort and the difficulty encountered in VR case studies, a significant reliance on qualitative data from questionnaires for such assessments was identified. However, the quantitative data used largely depended on the context of the task, such as the completion time. Moreover, other data to estimate physical effort is used for rehabilitation purposes and depends on other hardware tools such as haptic gloves to determine pinch strength precisely [33]. Therefore, based on these studies and the limited use of motion data obtained through VR headset hand tracking systems, we propose two quantitative metrics to estimate the difficulty encountered and the physical effort made by volunteers in our case study.

The Difficulty Index (DI) is a metric we developed, taking into account various characteristics extracted from the hand tracking data of volunteers during the case study. The index returns a value in the range [0,1], quantifying the difficulty encountered by the user based on their recorded movements. Below, we present the formula and a detailed description of its components.

Before describing the components of the index, we must first contextualize the data collection process and its implications. The application running on the Meta Quest 3 operated at an average frame rate of 72 frames per second (fps), recording spatial tracking data at each frame with four significant decimal digits of precision. This implies that the difference in the recorded coordinates between one row of the generated CSV files and the next is generally minimal, as 72 rows of data are stored per second. This provides detailed data records, but the noise generated by small, involuntary movements must be filtered out

before meaningful information, such as actual movements or distance travelled, can be obtained. To achieve this, a Kalman filter [43] was applied to eliminate noise, making the changes in the data smoother and more representative of the actual movement performed.

The Kalman filter used is a simplified version that does not require complex state transition models, as the goal is simply to smooth the collected data. Thus, the filter consists of the following steps:

1) Prediction.

$$x_k = x_{k-1} \tag{1}$$

where x_k is the prediction of the state at time k (smoothed position), and x_{k-1} is the corrected state at the previous time step. Here, we assume that the previous state is the best prediction for the next.

2) Covariance of the state and its prediction. The covariance P_k predicts the uncertainty of the estimation.

$$P_k = P_{k-1} + Q \tag{2}$$

where Q is the process noise. Q is parametrizable, and in our context, $Q = 0.001$ since the uncertainty in the position change between two consecutive frames is minimal.

3) Correction with the measurement (Kalman Gain)

$$K_k = \frac{P_{k|k-1}}{P_{k|k-1} + R} \tag{3}$$

where R is the covariance of the measurement noise and $P_{k|k-1}$ is the covariance estimated in time k based on the information available until time $k-1$. K_k quantifies how much to trust the prediction compared to the new measurement. In our context, $R = 0.01$ represents a low value because our measurements contain insignificant noise, with only small changes due to the high measurement frequency (72 Hz). Therefore, the measured data is highly reliable.

4) Correction of the state with the measurement.

$$x_{k|k} = x_{k|k-1} + K_k \times (z_k - x_{k|k-1}) \tag{4}$$

where $x_{k|k-1}$ represents the prediction of the state at time k based on the available information at time $k-1$, before knowing the new measurement at time k. Thus, $x_{k|k}$ is the corrected or smoothed state estimate at time k, incorporating the new measurement z_k and the Kalman gain K_k.

5) Covariance correction.

$$P_{k|k} = (1 - K_k) \times P_{k|k-1} \tag{5}$$

where the uncertainty of the state is updated for the next iteration.

The result of applying the filter to the dataset returns the corrected hand positions, $x_{k|k}$, resulting in a smoothing of minor fluctuations, allowing significant movements to be detected as changes in position while ignoring small changes caused by the high sampling frequency.

Taking this smoothing of the data into account, DI is represented by the following equation:

$$DI = \frac{NV_{BX} + N_{TD}N_{ME} + N_{MPS}}{4} \tag{6}$$

where NV_{BX} represents the normalized volume of the *bounding box* that encapsulates the range of motion for each hand, N_{TD} is the Euclidean distance between two points (i,j) covered by the hands, normalized, N_{ME} quantifies how efficiently the user has moved over a given number of frames, and, finally, N_{MPS} represents the number of movements per second (MPS). We wanted to include range of motion as it is one of the features used in the literature study regarding AR and VR applications in hand rehabilitation [33]. The parameters chosen for DI —movement volume, total distance traveled, movement efficiency, and number of movements per second— were selected based on their direct relationship with motor effort and precision, which are critical factors in assessing the difficulty experienced by users with mobility impairments. These parameters have been widely used in prior research to evaluate interaction complexity in VR and AR environments, particularly in rehabilitation contexts [22], [23]. The inclusion of these specific metrics ensures that DI captures different aspects of task execution: movement volume reflects the extent of hand displacement, total distance accounts for the overall effort required, movement efficiency quantifies precision, and movements per second indicate the control and stability of interactions. The letter N at the beginning of each component indicates that the data is normalized, as they have different scales, in order to assign equal weight in the calculation of the index. Therefore, the normalization process followed for each metric was the max absolute scaling. That is, divide each metric item by the maximum absolute value obtained for such metric from all volunteer's data obtained. This ensures us that the data items are within range [0,1]. The following paragraphs detail each of these characteristics:

1) NV_{BX}. The total volume of the *bounding box* in cubic meters (m^3) is calculated as follows:

$$\begin{aligned} V_{BX} = {} & V((\max(R_X) - \min(R_X)) \times (\max(R_Y) \\ & - \min(R_Y)) \times (\max(R_Z) - \min(R_Z))) \\ & + V((\max(L_X) - \min(L_X)) \times (\max(L_Y) \\ & - \min(L_Y)) \times (\max(L_Z) - \min(L_Z))) \end{aligned} \tag{7}$$

That is, the volumes of the *bounding boxes* for the left hand L and right hand R are summed, obtained by calculating the maximum and minimum values along the three-dimensional axes.

2) N_{TD}. The distance is computed for each of the used hands L and R in meters (m). This distance is calculated as following:

$$TD_h = \sum_{i=1,j=1}^{i=F,j,F} \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2 + (Z_i - Z_j)^2} \tag{8}$$

where h represents the hand and F the number of records in the data file once the Kalman filter is applied. In order to consolidate the calculated distance into a single value, the arithmetic mean is computed: $\frac{TD_L+TD_R}{2}$

3) N_{ME}.We start with the assumption that the most efficient movement between two points in three-dimensional space is the straight-line distance that connects them:

$$D = \sqrt{(X_N - X_1)^2 + (Y_N - Y_1)^2 + (Z_N - Z_1)^2} \quad (9)$$

where N represents the last frame of the analysed sequence. On the other hand, the total distance covered between these points is:

$$D_T = \sum_{i=1}^{N-1} \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2 + (Z_{i+1} - Z_i)^2} \quad (10)$$

A larger total distance covered may indicate more erratic and imprecise movements, which can be further compounded by the angle between consecutive frames to measure trajectory change by summing all the angles, such that:

$$\theta_T = \sum_{i=1}^{N-1} \theta_i = \sum_{i=1}^{N-1} arccos\left(\frac{(v_i \times v_{i+1})}{|v_i| \times |v_{i+1|}}\right) \quad (11)$$

v_i represents the displacement vector between points i and $i+1$. Lastly, redundant movements, where the user returns to a previous or very close position, are counted, indicating imprecision and a failure to stabilize the hand, for example, when trying to teleport. This allows us to observe hand stabilization and compare how the non-adapted interaction requires maintaining the hand shape while performing a *pinch*, whereas the adapted interaction does not require maintaining or forming a specific hand shape. To compute redundant movements, a threshold $\epsilon = 0.01$ centimeters is applied concerning the distance between consecutive points:

$$RM = \sum_{i=1}^{N-1} 1(distance(i, i+1) < \epsilon) \quad (12)$$

where $1(distance(i, i+1) < \epsilon)$ is a function that adds 1 if the distance is higher than the threshold ϵ. Thus:

$$ME = \frac{1}{\frac{D}{D_T} \times \frac{1}{1+\theta_T+RM}} \quad (13)$$

This final inversion is done so that low efficiency values have a greater impact on increasing the value of the DI, giving us a dimensionless score. Taking into account both hands, the final value is $\frac{ME_L+ME_R}{2}$

4) N_{MPS}. Number of movements per second (MPS), normalized. A higher number of movements per unit of time contributes to user fatigue, indicating that they are facing greater difficulty and require more movements to try to accomplish the task. Thanks to the Kalman filter, we can compute each movement as the distance between two consecutive points i and $i+1$, establishing a threshold $\epsilon = 0.01$, which, as before, refers to 1 cm. Taking into account both hands, the final value is $\frac{MPS_L+MPS_R}{2}$.

Our decision to normalize over the population was driven by the need to ensure fair comparisons among users with different levels of mobility. This was also suggested by expert staff of the hospital we collaborated with. Since the parameters used to calculate DI—such as movement volume, total distance traveled, movement efficiency, and movements per second—vary significantly between individuals, raw values would be inherently biased by each user's motor capabilities rather than reflecting the relative challenge of the interaction. Normalization mitigates this issue by adjusting values within a common scale, allowing DI to capture differences in difficulty across participants rather than being skewed by individual limitations. Table 5 shows the values obtained for each component of the DI based on the collected data, as well as the DI value itself. It can be observed that, generally, the volume of the bounding box encapsulating the hand movements is smaller when using adapted interactions, which is consistent with what is shown in Fig. 10. A smaller movement range results in a smaller volume, indicating that the user was able to perform the movements more comfortably and in a more controlled manner, without having to compensate with other parts of the body, which would have resulted in a larger range of motion. Moreover, in general, the DI components related to movement show lower values in the adapted environments. This reduced number of movements, executed more efficiently and covering less distance, helps reduce user fatigue and indicates that the movements were performed more precisely and with greater control for the same purposes as with the unadapted interactions. We now calculate the percentage difference of the DI for each user as

$$PD = \frac{DI_A - DI_U}{DI_U} \times 100 \quad (14)$$

where DI_A is the score obtained for DI in the adapted environment and DI_U is the score obtained for DI in the unadapted environment. Therefore, we find that the overall mean is -37.06%, the median is -40.10%, and the standard deviation is 14.67%. In other words, users experienced an average reduction of nearly 40% in the required hand movements, as well as a smaller movement range and greater precision, resulting in less fatigue and perceived difficulty in completing the proposed tasks using the adapted interactions.

On the other hand, Table 6 shows the number of gestures or attempts made by a user in each task. Each attempt refers to the intention or success in making a hand shape or gesture required to perform an interaction. The effective execution of the interaction—for example, making the L-shape or attempting it with the hand—ecounts as 1 attempt, as does the pinch gesture or palmar grasp. In this case, the CSV

TABLE 5. Values of the components that make up the Difficulty Index (DI) and the final value of the DI for each user and each type of interaction used. In bold, adapted interactions' values are highlighted.

User	Interaction type	Bounding box volume	Total distance	Movement Efficiency	Movements per second	Normalized Volume (NV_BX)	Normalized Total distance (N_TD)	Normalized Movement Efficiency (N_ME)	Normalized Movements per second (N_MPS)	Difficulty Index (DI)
1	Unadapted	2.59	141.07	122882.88	6.29	1.0	0.56	0.2	0.67	0.49
1	**Adapted**	**1.16**	**71.38**	**105583.26**	**4.23**	**0.45**	**0.29**	**0.17**	**0.45**	**0.27**
2	Unadapted	1.54	114.05	190554.25	4.2	0.59	0.46	0.31	0.45	0.36
2	**Adapted**	**0.9**	**80.52**	**63884.47**	**7.4**	**0.35**	**0.32**	**0.1**	**0.79**	**0.31**
3	Unadapted	2.11	141.26	128044.12	9.35	0.81	**0.56**	0.21	1.0	0.52
3	**Adapted**	**1.97**	**80.4**	**135048.0**	**4.19**	**0.76**	**0.32**	**0.22**	**0.45**	**0.35**
4	Unadapted	1.74	90.13	160468.54	4.76	0.67	0.36	0.26	0.51	0.36
4	**Adapted**	**0.88**	**35.58**	**21700.54**	**2.03**	**0.34**	**0.14**	**0.04**	**0.22**	**0.15**
5	Unadapted	2.08	158.63	236371.68	4.51	0.8	0.63	0.39	0.48	0.46
5	**Adapted**	**0.84**	**89.65**	**110759.79**	**4.76**	**0.32**	**0.36**	**0.18**	**0.51**	**0.27**
6	Unadapted	1.62	221.77	359340.09	6.48	0.63	0.89	0.59	0.69	0.56
6	**Adapted**	**1.08**	**112.46**	**118344.92**	**7.59**	**0.42**	**0.45**	**0.19**	**0.81**	**0.37**
7	Unadapted	1.85	250.06	361796.36	7.21	0.71	1.0	0.59	0.77	0.61
7	**Adapted**	**1.18**	**91.11**	**67787.3**	**5.67**	**0.46**	**0.36**	**0.11**	**0.61**	**0.31**
8	Unadapted	1.55	170.54	609840.37	3.67	0.6	0.68	1.0	0.39	0.53
8	**Adapted**	**1.16**	**59.77**	**42681.67**	**1.96**	**0.45**	**0.24**	**0.07**	**0.21**	**0.19**
9	Unadapted	1.53	121.69	129441.35	6.56	0.59	0.49	0.21	0.7	0.4
9	**Adapted**	**0.77**	**67.15**	**60204.39**	**4.92**	**0.3**	**0.27**	**0.1**	**0.53**	**0.24**
10	Unadapted	2.45	215.86	583853.42	5.94	0.95	0.86	0.96	0.64	0.68
10	**Adapted**	**1.15**	**64.21**	**79041.2**	**2.67**	**0.44**	**0.26**	**0.13**	**0.29**	**0.22**
11	Unadapted	1.55	49.14	29673.7	8.97	0.6	0.2	0.05	0.96	0.36
11	**Adapted**	**0.84**	**71.31**	**41180.57**	**7.86**	**0.32**	**0.29**	**0.07**	**0.84**	**0.3**
12	Unadapted	1.23	73.04	65908.17	7.5	0.47	0.29	0.11	0.8	0.33
12	**Adapted**	**0.84**	**45.92**	**19028.54**	**6.05**	**0.32**	**0.18**	**0.03**	**0.65**	**0.24**

TABLE 6. Comparison of user interactions in adapted and unadapted environments. The table presents the number of attempts made by each participant for various tasks. The values represent the number of gestures or interaction attempts recorded during the study.

User	Environment	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
1	**Adapted**	**1**	**1**	**4**	**2**	**1**	**5**	**1**	**4**	**2**	**7**
1	Unadapted	17	1	4	8	3	12	2	9	3	9
2	**Adapted**	**2**	**1**	**8**	**2**	**3**	**5**	**1**	**6**	**2**	**4**
2	Unadapted	4	2	4	10	1	13	2	11	2	12
3	**Adapted**	**1**	**1**	**3**	**1**	**1**	**6**	**1**	**6**	**2**	**7**
3	Unadapted	8	1	4	15	1	17	1	10	4	11
4	**Adapted**	**2**	**1**	**6**	**1**	**1**	**4**	**2**	**5**	**3**	**5**
4	Unadapted	26	1	4	3	1	10	3	15	4	10
5	**Adapted**	**3**	**2**	**6**	**3**	**1**	**3**	**1**	**4**	**2**	**5**
5	Unadapted	19	2	9	17	2	-	3	-	4	2
6	**Adapted**	**2**	**1**	**3**	**1**	**3**	**6**	**1**	**4**	**2**	**7**
6	Unadapted	10	1	5	5	3	17	2	9	4	10
7	**Adapted**	**1**	**1**	**4**	**1**	**1**	**3**	**2**	**6**	**2**	**7**
7	Unadapted	9	1	5	6	7	-	4	-	5	5
8	**Adapted**	**2**	**1**	**5**	**3**	**1**	**8**	**3**	**6**	**1**	**9**
8	Unadapted	7	1	-	18	3	14	5	-	2	3
9	**Adapted**	**1**	**1**	**4**	**3**	**1**	**4**	**1**	**5**	**2**	**7**
9	Unadapted	9	1	4	8	1	9	2	9	4	9
10	**Adapted**	**1**	**1**	**5**	**1**	**3**	**6**	**1**	**4**	**2**	**10**
10	Unadapted	14	1	6	16	2	13	3	4	4	7
11	**Adapted**	**3**	**2**	**7**	**1**	**1**	**3**	**1**	**4**	**2**	**4**
11	Unadapted	3	1	3	4	1	7	3	8	2	10
12	**Adapted**	**2**	**1**	**6**	**2**	**1**	**6**	**1**	**7**	**2**	**7**
12	Unadapted	6	1	4	5	1	7	3	10	2	9

files recorded from each user allows us to determine when an interaction occurred, looking at he states of the relevant interactors for each interaction. In the case of adapted interactions, we recall that these gestures or hand shapes are simplified, counting as 1 attempt the fact of making such gesture. With both unadapted and adapted interactions, users could make attempts without carrying out the intended interaction. Delving into the Meta SDK code, certain classes allowed us to compute attempts. For instance, if a given threshold determine whether or not a pinch occurs, values lower than such threshold in a determined range help us determine that an attempt was made. Th API references for such classes we used are the following:

1) *FingerPinchGrabAPI* (https://developers.meta.com/horizon/reference/interaction /v71/class_oculus_interaction_grab_a_p_i_finger_pinch_grab_a_p_i). Helped us determine pinch/grab attempts as well as unadapted teleportations and turnings.
2) *FingerFeatureStateProvider* (https://developers.meta.com/horizon/reference/interaction/v71/class_oculus_interaction_pose_detection_finger_feature_state_provider). Helped us determine attempts related with gestures that trigger interactions (invoking teleportation arc, the arrows for turning or distance grab). This class uses recorded data from the Hand with the *Hand* class as well has thresholds for determining the state of fingers, *FingerStateThresholds*.

For volunteers U5, U8, and U7 in the unadapted environment, there are certain tasks in which no attempts were recorded. This is because the attempts made by the user did not exceed the required thresholds for the headset to detect the spatial position or required hand shape, meaning that the confidence levels provided by the Meta API were too low to detect these attempts. Additionally, this is combined with the fact that they abandoned the task before completing it, resulting in even fewer attempts.

Before addressing the research questions, we will consolidate into a single value: i) the time each user spent on each task ($Time_{Ti}$) and whether they completed it or not, ii) the number of attempts made ($Attempts_T$), and iii) the difficulty experienced by user U in the environment env, $DI_{U,env}$. Moreover, $Attempts_{max}$ and $Time_{max}$ represent the maximum values obtained for attempts and time in a single task T, respectively, over the user population, in order to normalize both terms. We will call this value the Effort Degree (ED), which is calculated for each task T and user U using the type

TABLE 7. Effort degree values taking into account number of attempts, time spent and difficulty index.

Task	U1		U2		U3		U4		U5		U6		U7		U8		U9		U10		U11		U12	
	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U	A	U
#1	**0.55**	3.72	**0.67**	2.15	**0.52**	1.24	**0.7**	2.15	**0.91**	3.48	**0.68**	3.15	**0.37**	4.09	**0.58**	2.22	**0.42**	2.1	**0.35**	3.27	**1.12**	1.01	**0.61**	1.5
#2	**0.47**	0.63	**0.62**	0.46	**0.62**	0.48	**0.54**	0.48	**0.75**	1.34	**0.55**	1.31	**0.5**	0.45	**0.35**	1.23	**0.5**	0.56	**0.32**	0.39	**0.99**	0.43	**0.59**	0.54
#3	**1.32**	2.72	**2.29**	2.31	**1.03**	2.42	**1.83**	2.57	**1.94**	2.94	**1.06**	2.36	**1.18**	3.5	**1.26**	2.39	**1.26**	2.53	**1.11**	3.25	**2.29**	1.06	**1.37**	1.29
#4	**0.85**	1.63	**0.71**	1.89	**0.63**	3.48	**0.82**	0.73	**1.44**	3.77	**0.87**	1.93	**0.69**	2.67	**0.81**	3.61	**0.91**	1.97	**0.65**	4.19	**0.67**	0.9	**0.68**	1.12
#5	**0.65**	1.06	**1.11**	1.44	**0.66**	0.58	**0.58**	0.58	**0.65**	1.51	**0.97**	1.8	**0.56**	1.55	**0.42**	1.61	**0.61**	0.74	**0.65**	0.43	**0.7**	0.53	**0.53**	0.6
#6	**1.51**	2.54	**1.48**	3.55	**2.5**	3.43	**1.7**	2.49	**1.4**	1.92	**2.2**	3.78	**1.06**	1.56	**1.76**	3.72	**1.6**	3.09	**1.51**	4.04	**1.4**	1.65	**1.67**	1.84
#7	**0.54**	0.84	**0.68**	0.59	**0.72**	0.45	**0.74**	1.46	**0.53**	1.97	**0.54**	1.62	**0.74**	2.82	**0.94**	1.94	**0.59**	0.73	**0.31**	0.77	**0.79**	0.81	**0.59**	1.09
#8	**1.55**	1.79	**2.29**	3.36	**2.82**	3.14	**1.77**	2.84	**1.82**	2.11	**1.71**	2.8	**1.68**	1.53	**1.86**	1.67	**1.79**	3.74	**1.13**	2.04	**2.04**	1.82	**2.17**	2.45
#9	**1.0**	1.03	**1.31**	1.59	**0.96**	0.73	**1.34**	0.65	**1.04**	1.65	**0.89**	0.68	**0.76**	1.14	**0.39**	1.57	**0.82**	1.11	**0.64**	1.82	**1.23**	0.84	**1.01**	0.76
#10	**2.54**	2.17	**1.82**	3.62	**2.57**	2.71	**2.11**	2.83	**2.33**	2.13	**2.74**	3.48	**2.62**	2.94	**2.38**	2.19	**2.48**	3.8	**2.44**	3.09	**2.0**	2.72	**2.48**	2.53

of interactions i:

$$ED = \left(\frac{Attempts_T}{Attempts_{max}} \times \frac{Time_{Ti}}{Time_{max}}\right) \times (1 + DI_{U,env}) + 1(NotCompleted(T, U, i)) \quad (15)$$

As can be seen, the terms for attempts and time are normalized so that neither predominates. On the other hand, the term $1 + DI_{Ui}$ ensures that there is always an impact of time and attempts on the effort, even if the DI is very low. Moreover, it ensures that the ED is amplified based on the DI obtained from the movements made by the user with their upper limbs. Finally, $1(NotCompleted(T, U, i))$ represents a function that adds 1 if the task T was not completed by U using the interactions i. Table 7 shows the results obtained, with the columns related to the adapted interactions highlighted in bold.

Observing Table 7, we can see the significant difference in effort considering the previously mentioned components. Those tasks where the values are not as significant, or where the resulting value is even lower in the unadapted interactions, are those exclusively related to object grasping: #2, #5, and #9. This may be due to the waiting times required for auto-grasp and distant grasping, as well as the injuries of each user, as those with both hands affected show an opposite trend to those who are healthy or have only one hand injured. This can also be seen in task #10, where, being the last task, users with more severe injuries abandoned the task earlier, resulting in fewer attempts compared to the same task with the adapted interactions. This also impacts time and DI, as the autonomy and ease provided by the adapted interactions allowed them to complete the task, making more movements and spending more time.

3) ANSWERING RESEARCH QUESTIONS

Having conducted the previous quantitative analysis, we will now address the following research questions.

a: RQ1: DO ADAPTED INTERACTIONS ENABLE PEOPLE WITH MOTOR IMPAIRMENTS TO PERFORM TASKS IN IMMERSIVE SPACES?

As demonstrated in Section V-A1, the adapted interactions allowed the injured users to complete 100% of the tasks in the shopping environment, enabling them to complete tasks they had to abandon when using unadapted interactions. This effect is particularly noticeable in patients with both hands injured, where the task completion rate is, at most, 30% of the proposed tasks. These interactions not only allowed them to complete the tasks without abandoning them due to fatigue but also significantly improved the time taken to complete them, as shown by the paired t-test conducted.

b: RQ2: DO ADAPTED INTERACTIONS IMPROVE THE PERFORMANCE OF A PERSON WITH MOTOR IMPAIRMENT WHEN PERFORMING TASKS IN IMMERSIVE SPACES?

As previously highlighted, Section V-A1 demonstrated the performance improvement achieved with the adapted interactions. However, this has been proven not only in terms of time and task completion but also in terms of the effort exerted and the difficulty encountered, as shown in Section V-A2. The reduced effort was due to a decrease in difficulty, facilitated by adapting the interactions to the users' upper limb impairments. This was demonstrated throughout Section V-A1. Additionally, we can perform a paired t-test concerning the ED:

1) H_0: There is no significant difference in the effort required to complete the tasks in the case study using the adapted interactions.
2) H_1: There is a significant difference in the effort required to complete the tasks in the case study using the adapted interactions.

We will compare all the values obtained for each task with the adapted interactions and those obtained with the unadapted interactions, shown in Table 7. Thus, we obtain $t - statistic = -6.79$, indicating that the mean values with the adapted interactions are lower. Furthermore, $p - value = 8.71 \times 10^{-11}$, which is much smaller than $\alpha = 0.05$, so we can affirm that there is a highly statistically significant difference in the effort required to complete the case study tasks using the adapted interactions, implying a lower effort.

To evaluate the relationship between the perceived difficulty reported by users and their performance in task completion, a Spearman correlation analysis was conducted between the DI and two key performance metrics: the task Completion Rate (CR) and the task Completion Time (CT), see Table 4. This analysis was performed separately for both the unadapted and adapted environments, allowing for a comparison of how these relationships vary depending on the presence of interaction adaptations.

The Spearman correlation coefficient (ρ) was chosen due to its suitability for non-parametric data, as it does not assume

TABLE 8. Questionnaire's questions along with the mean and standard deviation of answers per questions.

Question ID	Question	Scoring system	Mean		Standard deviation	
			A	U	A	U
Q1	How much experience do you think you have with the use of Virtual Reality (VR) glasses?	1 - None, 6 - A lot	**2.42**	2.42	**1.38**	1.38
Q2	Were the tasks very mentally demanding?	1 - None demanding, 6 - Really demanding	**1.50**	1.50	**0.80**	0.80
Q3	Were the tasks very physically demanding?	1 - None demanding, 6 - Really demanding	**1.91**	4.33	**0.79**	1.15
Q4	Did you have to make a great effort to carry out the tasks?	1 - None, 6 - A lot of effort	**1.75**	4.33	**0.97**	1.70
Q5	How frustrated, stressed or irritated did you feel?	1 - None, 6 - A lot	**1.50**	3.50	**0.67**	1.40
Q6	Did the interactions seem easy to use?	1 - Really easy, 6 - Really difficult	**1.67**	4.16	**0.89**	1.40
Q7	The interactions demotivated me from using the application	1 - Not at all, 6 - A lot	**1.33**	2.66	**0.65**	0.77
Q8	The interactions are understandable and easy to learn to use	1 - Strongly understandable and easy to learn, 6 - Very little	**1.58**	4.33	**0.51**	1.57
Q9	The interactions seemed friendly and appropriate for my limitations	1 - Very friendly and appropriate, 6 - Very little	**1.92**	4.50	**0.67**	1.56
Q10	I found it difficult to grab, release, move around the environment, and add products	1 - Very little, 6 - Very difficult	**1.50**	4.41	**0.67**	1.57
Q11	I would like to use these interactions frequently if I had to use the application	1 - Strongly agree, 6 - Strongly disagree	**1.67**	4.75	**0.65**	1.42
Q12	I think I would need help from a person with technical knowledge to use the system	1 - Strongly disagree, 6 - Strongly agree	**2.08**	3.50	**0.67**	0.90
Q13	I found the system very cumbersome to use	1 - Strongly disagree, 3- Neither agree nor disagree, 6 - Strongly agree	**1.50**	4.42	**0.90**	1.56
Q14	I felt very safe using the system	1 - Strongly agree, 6 - Strongly disagree	**1.33**	3.67	**0.49**	1.15
Q15	I felt very immersed in the virtual environment	1 - Strongly agree, 6 - Strongly disagree	**1.50**	1.50	**0.67**	0.67
Q16	I felt general discomfort or dizziness	1 - Strongly disagree, 6 - Strongly agree	**1.00**	1.00	**0.00**	0.00
Q17	I ended up with a headache	1 - Strongly disagree, 6 - Strongly agree	**1.17**	1.17	**0.39**	0.39
Q18	I felt blurred vision or eye fatigue	1 - Strongly disagree, 6 - Strongly agree	**1.25**	1.25	**0.45**	0.45
Q19	I felt tired	1 - Strongly disagree, 6 - Strongly agree	**1.17**	3.67	**0.39**	1.44

a normal distribution and effectively measures the strength and direction of monotonic relationships. The correlation was calculated as follows:

In the adapted environment, all users achieved a 100% completion rate, while in the unadapted environment, the completion rates varied depending on the users' motor abilities. For the CT, the total time taken to complete all tasks was calculated for each user in both environments. This allowed for an assessment of whether the DI was related to the time required to complete the tasks.

In the unadapted environment, the analysis revealed a moderate negative correlation between the DI and the task completion rate ($\rho = -0.578$, $p = 0.049$), indicating that users with higher DI scores completed fewer tasks. This suggests that the perceived difficulty had a direct impact on users' ability to complete tasks when no adaptations were in place. Additionally, there was a strong positive correlation between the DI and the task completion time ($\rho = 0.711$, $p = 0.009$), implying that users with higher DI scores required more time to complete the tasks. These results confirm that the DI reflects the real-world challenges faced by users, as greater perceived difficulty was associated with both lower success rates and longer completion times in the unadapted environment.

In contrast, in the adapted environment, the correlation between the DI and the task completion rate could not be computed because all users achieved a 100% completion rate. This lack of variability in performance outcomes indicates the positive impact of the adaptations, as they enabled all users to successfully complete the tasks regardless of their perceived difficulty levels. Furthermore, the correlation between the DI and the task completion time was not statistically significant ($\rho = -0.084$, $p = 0.794$), suggesting that users completed tasks in similar amounts of time regardless of their DI scores.

These findings highlight a clear distinction between the two environments. In the unadapted environment, the DI is strongly correlated with performance metrics, confirming that it accurately reflects the difficulties encountered by users

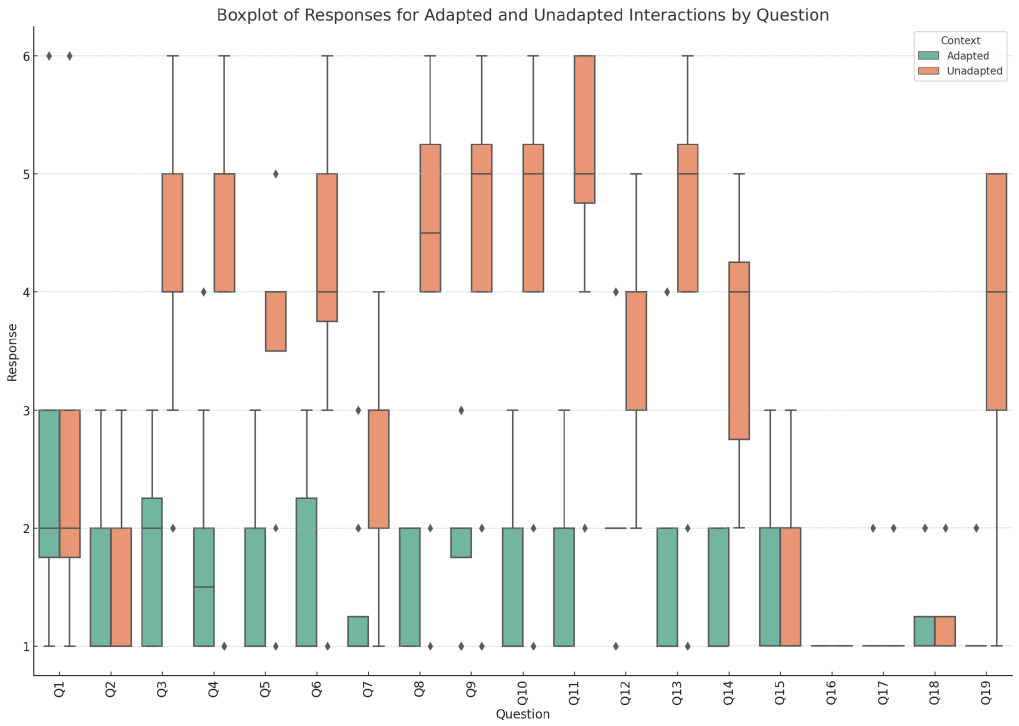


FIGURE 11. Boxplot of the questionnaire responses, containing two boxes per question.

without interaction adaptations. On the other hand, in the adapted environment, the absence of significant correlations indicates that the adaptations effectively mitiagated the impact of perceived difficulty on task performance. Even users who reported higher levels of perceived difficulty were able to complete tasks successfully and in comparable times to those with lower DI scores.

Importantly, the lack of correlation in the adapted environment should not be interpreted as a limitation of the DI. Rather, it serves as evidence of the success of the implemented adaptations. By reducing the influence of users' motor limitations, the adaptations created a more inclusive environment where performance disparities based on individual capabilities were minimized. This outcome is analogous to universal design principles, where accessibility improvements, such as ramps or assistive technologies, eliminate barriers without creating performance gaps between different user groups.

B. QUALITATIVE ANALYSIS

Next, we will analyse the responses to the questionnaire that each user completed after using both the adapted and the unadapted interactions. The volunteers answered the same questions considering their experience and feelings with both types of interactions. In cases where the questions were not intrinsically related to the types of interaction, the responses were identical.

Table 8 presents the questions from the questionnaire, as well as the mean and standard deviation of the 12 responses to each question within the two contexts: with the adapted interactions highlighted in bold and the unadapted interactions not highlighted. Additionally, to statistically analyse these responses, Fig. 11 shows a boxplot containing two boxes for each question. These plots summarize the distribution of a dataset and visualize its key features, with the box representing the interquartile range containing 50% of the data, the central line corresponding to the median, and the whiskers extending from the box with outliers shown outside them. The orange boxes represent the responses regarding the unadapted interactions, while the green boxes represent those for the adapted interactions. On one hand, we can clearly observe an opposite trend between both contexts. Additionally, the outliers in the responses to the unadapted interactions correspond to healthy HNPT staff, who perceived

TABLE 9. Wilcoxon Signed-rank test p-values obtained from the questionnaire's responses.

Question ID	Wilcoxon p-value
Q3	0.0005
Q4	0.0045
Q5	0.0066
Q6	0.0048
Q7	0.0050
Q8	0.0046
Q9	0.0041
Q10	0.0045
Q11	0.0046
Q12	**0.0069**
Q13	**0.0069**
Q14	0.0041
Q19	0.0046

these interactions as more typical from their perspective. In the case of the unadapted interactions, we also observe elevated medians in questions such as Q9, Q10, Q11, and Q13.

We will analyse these results by grouping the questionnaire questions into categories:

1) **Physical and mental demand**. In Q2 (mental demand), the responses are consistent in both contexts, with a low median and minimal dispersion. This suggests that users perceived a low mental load in both the adapted and unadapted contexts. However, in Q3 (physical demand), the differences are much more pronounced. The unadapted interactions show a higher median, with considerably greater dispersion compared to the adapted ones. This indicates that, while some users did not perceive a significant increase in physical demand, others experienced much higher levels of physical effort in the unadapted context. Q4 reinforces this trend, showing greater variability and a considerably higher median. This confirms that physical effort was perceived as more demanding in the unadapted context, whereas in the adapted context, the responses are more clustered toward the lower end of the scale.
2) **Usability and comprehension**. The responses related to ease of use (Q6, Q13) show a clear difference between contexts. With the adapted interactions, responses are concentrated at lower values, with limited dispersion, indicating that most users found the interactions easy to use. In contrast, the unadapted interactions show greater dispersion and a significantly higher median, suggesting more variability in user experience and a general perception of greater difficulty. Similarly, Q8 (understandability of the interactions) and Q9 (friendly and appropriate interactions) present a pattern where the adapted interactions are perceived as more accessible and easier to understand, while the unadapted interactions are seen as more complex, with greater variability in the responses. Finally, Q12 shows very concentrated responses with the adapted interactions, unlike the unadapted ones, where there is a greater perceived need for technical assistance.
3) **Frustration and motivation**. Questions Q5 (frustration) and Q7 (demotivation) reveal a similar pattern. In both cases, the median and dispersion with the unadapted interactions are notably higher, indicating that, for many users, the unadapted context generated more frustration and demotivated them from using the application. In contrast, the adapted context shows responses concentrated at lower levels.
4) **Accessibility**. In Q10 (difficulty in manipulating the environment), the responses in the unadapted context are significantly more dispersed, with a high median. This indicates that interacting with the environment was perceived as considerably more challenging in this context, whereas the adapted context was perceived as more manageable, with responses clustered at the lower end of the scale. Q11 (willingness to use the interactions) reflects a similar behaviour, suggesting that users would prefer to avoid these interactions in the unadapted context, compared to the adapted one.
5) **Fatigue and safety**. In Q14 (feeling of safety), the unadapted interactions show a higher median, indicating that users felt less safe in this context. The dispersion is also higher, suggesting that some users felt significantly less safe than others, particularly the more severely injured patients. Q19 (feeling of fatigue) shows a similar trend. This reinforces the perception that the unadapted context resulted in a greater physical and mental load for the users.
6) The remaining questions show no variability (Q1, Q16, Q17, Q18).

Additionally, we will check if there are significant differences between the responses to the questions that show variability using a Wilcoxon Signed-rank test [42]. This non-parametric test will be used instead of the t-test due to the non-continuous nature of the data obtained from the questionnaire, which is based on Likert-scale questions.

Except for questions Q12 and Q13, the p-value is significant as $p < \alpha, \alpha = 0.005$, indicating a statistically significant difference between the responses after using the adapted and unadapted interactions. Given the relatively small sample size and the need to ensure robustness in our statistical findings, we adopted a conservative significance threshold of 0.005 for the Wilcoxon signed-rank test. This choice aligns with best practices in studies with multiple comparisons, reducing the likelihood of false positives while maintaining confidence in the results. Observing the nature

of the responses, we can affirm that users rated the adapted interactions significantly more positively.

Finally, we answer RQ3, as indicated in Section IV-A: *Do users positively evaluate the adapted interactions in terms of usability and user experience?* The statistical analysis of the questionnaire responses demonstrates that users, particularly the patients from HNPT, as can be seen from the individual responses available in the public repository, rated the adapted interactions positively. Thanks to this case study, we were also able to gain the acceptance of these patients and therapists regarding the use of these adapted interactions.

VI. CONCLUSION

In this paper, we have addressed the need to adapt hand interactions in immersive spaces based on Virtual Reality, aiming to improve accessibility for individuals with motor limitations. This proposal contributes to enhancing inclusivity in Virtual Reality, which is beginning to reach a maturity phase.

As a case study, VR Shopping was chosen. The number of people making their purchases online through a mobile device is increasing in an ever more digitalised society. VR Shopping is poised to be the next revolution in e-commerce, allowing users to have shopping experiences similar to physical ones, but from a distance.

Collaboration with the Hospital Nacional de Parapléjicos in Toledo has allowed us to evaluate the proposal in a real environment with individuals who suffer from real limitations, in this case, patients with spinal cord injuries and motor limitations in legs, arms, and hands. As observed in the results obtained, the adaptation of interaction mechanisms has enabled users to successfully complete tasks they could not finish with traditional mechanisms. This increase in effectiveness is accompanied by a reduction in effort around 40% and completion rates and time. While completion rates with adapted interactions are 100% for all users, completion times increase between a 3% and 91.6% depending on the tasks and interactions required. The average increase is 38.63% with a standard deviation of 32.84%, since the completion times significantly improves in tasks where hand shapes and specific positions are required and those more simple such as grabbing do not present such a high increase. Moreover, we hope that our metrics DI and ED can be used extended with other quantitative data for other purposes, i.e., fatigue muscle detection using signals from electrodes.

Besides quantitative evaluation, the proposal was also assessed qualitatively. We conducted a survey with questions from standard questionnaires including, NASA-TLX, System Usability Scale (SUS), User Experience Questionnaire (UEQ) and Motion Sickness Assesment Questionnaire (MSAQ), among others. This allowed us to assess how users perceived adapted interactions against unadapted ones. The responses to the questionnaire, based on Likert-scale questions, were analysed by using Wilcoxon Signed-rank test and boxplots. Wilcoxon p-values demonstrated significant differences among the responses given about adapted and unadapted interactions, being the adapted ones positively evaluated in terms of user experience, fatigue and usability, while unadapted ones where negatively evaluated. Moreover, among the volunteers that were patients the boxplots showed remarkable statistical differences among the responses given. The percentage difference among responses between adapted and unadapted interactions was 40.15%, with a standard deviation of 8.60%, considering all respondents.

Lastly, our intention is to continue to evolve the proposed solution. In future work, we will explore how to adapt more complex interaction mechanisms with hands in VR for individuals with mobility issues. In this way, we will explore how gesture recognition could work according to the limitations of each individual. Moreover, interactions with user interface for people with upper limb limitations were not researched. These interactions are crucial for interacting with VR applications. Therefore, we will make efforts to propose and evaluate solutions to adapt. Finally, in the context of VR Shopping as case study, we will research how VR shop layouts can be improved to make them more accessible, adapting not only interactions, but also the distribution of items, sections and shelves around a VR shop.

ACKNOWLEDGMENT

GPT-4o has been used for grammar enhancement.

DATA AVAILABILITY

All data collected during the experiment are publicly available in the GitHub repository: https://github.com/AIR-Research-Group-UCLM/VR-Hands-Interactions-Adapted. The repository is organized into two folders. One of them contains twelve folders, one for each of the users, containing all the data. The other, contains some Python code used to extract the information shown in the results section. The following Google Drive link contains a video showing the adapted interactions used in our case study https://github.com/AIR-Research-Group-UCLM/VR-Hands-Interactions-Adapted/blob/master/video_link.txt.

REFERENCES

[1] S. Arlati, V. Colombo, G. Ferrigno, R. Sacchetti, and M. Sacco, "Virtual reality-based wheelchair simulators: A scoping review," *Assistive Technol.*, vol. 32, no. 6, pp. 294–305, Nov. 2020.

[2] S. Baltierra, "Virtual reality and augmented reality applied to e-commerce: A literature review," in *Proc. HCI-COLLAB, Iberoamerican Workshop Hum.-Comput. Interact.*, Havana, Cuba. Cham, Swizerland: Springer, Oct. 2022, p. 201.

[3] S. R. Billewar, K. Jadhav, V. P. Sriram, D. A. Arun, S. M. Abdul, K. Gulati, and D. N. K. K. Bhasin, "The rise of 3D e-commerce: The online shopping gets real with virtual reality and augmented reality during COVID-19," *World J. Eng.*, vol. 19, no. 2, pp. 244–253, Mar. 2022.

[4] P. Biswas, P. Orero, M. Swaminathan, K. Krishnaswamy, and P. Robinson, "Adaptive accessible AR/VR systems," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, May 2021, pp. 1–7.

[5] B. Boyles, "Virtual reality and augmented reality in education," United States Mil. Acad., Center For Teaching Excellence, West Point, NY, USA, Tech. Rep. 1, 2017.

[6] J. Brooke, "Sus: A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, no. 194, pp. 4–7, 1995.

[7] A. Catz, M. Itzkovich, E. Agranov, H. Ring, and A. Tamir, "The spinal cord independence measure (SCIM): Sensitivity to functional changes in subgroups of spinal cord lesion patients," *Spinal Cord*, vol. 39, no. 2, pp. 97–100, Feb. 2001.

[8] Y. F. Cheng, C. Gebhardt, and C. Holz, "InteractionAdapt: Interaction-driven workspace adaptation for situated virtual reality environments," in *Proc. 36th Annu. ACM Symp. User Interface Softw. Technol.*, Oct. 2023, pp. 1–14.

[9] H. T. Chong, C. K. Lim, M. F. Ahmed, K. L. Tan, and M. B. Mokhtar, "Virtual reality usability and accessibility for cultural heritage practices: Challenges mapping and recommendations," *Electronics*, vol. 10, no. 12, p. 1430, Jun. 2021.

[10] A. Cuesta-Gómez, P. Sánchez-Herrera-Baeza, E. D. Oña-Simbaña, A. Martínez-Medina, C. Ortiz-Comino, C. Balaguer-Bernaldo-De-Quirós, A. Jardón-Huete, and R. Cano-De-La-Cuerda, "Effects of virtual reality associated with serious games for upper limb rehabilitation in patients with multiple sclerosis: Randomized controlled trial," *J. NeuroEngineering Rehabil.*, vol. 17, no. 1, pp. 1–10, Dec. 2020.

[11] P. Darshini, B. S. Kumar, S. J. Shree, U. K. Saravanan, and V. Bharath, "Design and development of smart wheelchair system using hand gesture control," in *Proc. 6th Int. Conf. Electron., Commun. Aerosp. Technol.*, Dec. 2022, pp. 312–315.

[12] M. Dombrowski, P. A. Smith, A. Manero, and J. Sparkman, "Designing inclusive virtual reality experiences," in *Virtual, Augmented and Mixed Reality. Multimodal Interaction*, J. Y. Chen and G. Fragomeni, Eds., Cham, Swizerland: Springer, 2019, pp. 33–43.

[13] J. Dudley, L. Yin, V. Garaj, and P. O. Kristensson, "Inclusive immersion: A review of efforts to improve accessibility in virtual reality, augmented reality and the metaverse," *Virtual Reality*, vol. 27, no. 4, pp. 2989–3020, Dec. 2023.

[14] P. Gianaros, E. Muth, J. Mordkoff, M. Levine, and R. Stern, "A questionnaire for the assessment of the multiple dimensions of motion sickness," *Aviation, Space, Environ. Med.*, vol. 72, p.115, Feb. 2001.

[15] A. Hameed, A. Perkis, and S. Möller, "Evaluating hand-tracking interaction for performing motor-tasks in VR learning environments," in *Proc. 13th Int. Conf. Quality Multimedia Exper. (QoMEX)*, Jun. 2021, pp. 219–224.

[16] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," in *Human Mental Workload* (Advances in Psychology), vol. 52, P. A. Hancock and N. Meshkati, Eds., Amsterdam, The Netherlands: North Holland, 1988, pp. 139–183.

[17] V. Herrera, J. Albusac, S. Schez-Sobrino, A. Reyes-Guzmán, and C. Glez-Morcillo, "Dynamic adjustment of interactive objects in virtual environments for upper limb rehabilitation: A patient-centred solution," in *Proc. Spanish Comput. Graph. Conf. (CEIG)*, J. Marco and G. Patow, Eds., 2024, pp. 1–11.

[18] V. Herrera, D. Vallejo, J. J. Castro-Schez, D. N. Monekosso, A. de los Reyes, C. Glez-Morcillo, and J. Albusac, "Rehab-immersive: A framework to support the development of virtual reality applications in upper limb rehabilitation," *SoftwareX*, vol. 23, Jul. 2023, Art. no. 101412.

[19] M. R. Huda, Md. L. Ali, and M. S. Sadi, "Real-time hand-gesture recognition for the control of wheelchair," in *Proc. 12th Int. Conf. Electr. Comput. Eng. (ICECE)*, Dec. 2022, pp. 384–387.

[20] E. L. Isenstein, T. Waz, A. LoPrete, Y. Hernandez, E. J. Knight, A. Busza, and D. Tadin, "Rapid assessment of hand reaching using virtual reality and application in cerebellar stroke," *PLoS ONE*, vol. 17, no. 9, Sep. 2022, Art. no. e0275220.

[21] N. Jamalian, M. Gillies, F. F. Leymarie, and X. Pan, "The effects of hand tracking on user performance: An experimental study of an object selection based memory game," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Oct. 2022, pp. 768–776.

[22] R. Khokale, G. S. Mathew, S. Ahmed, S. Maheen, M. Fawad, P. Bandaru, A. Zerin, Z. Nazir, I. Khawaja, I. Sharif, Z. U. Abdin, and A. Akbar, "Virtual and augmented reality in post-stroke rehabilitation: A narrative review," *Cureus*, vol. 15, Apr. 2023, Art. no. e37559.

[23] S. Kiani, I. Rezaei, S. Abasi, S. Zakerabasali, and A. Yazdani, "Technical aspects of virtual augmented reality-based rehabilitation systems for musculoskeletal disorders of the lower limbs: A systematic review," *BMC Musculoskeletal Disorders*, vol. 24, no. 1, p. 4, Jan. 2023.

[24] S. Kirshblum and W. Waring, "Updates for the international standards for neurological classification of spinal cord injury," *Phys. Med. Rehabil. Clinics North Amer.*, vol. 25, no. 3, pp. 505–517, Aug. 2014.

[25] M. Kour and K. Rani, "Challenges and opportunities to the media and entertainment industry in metaverse," in *Applications of Neuromarketing in the Metaverse*. Hershey, PA, USA: IGI Global, 2023, pp. 88–102.

[26] X. Lou, X. A. Li, P. Hansen, and P. Du, "Hand-adaptive user interface: Improved gestural interaction in virtual reality," *Virtual Reality*, vol. 25, no. 2, pp. 367–382, Jun. 2021.

[27] A. Masurovsky, P. Chojecki, D. Runde, M. Lafci, D. Przewozny, and M. Gaebler, "Controller-free hand tracking for grab-and-place tasks in immersive virtual reality: Design elements and their empirical study," *Multimodal Technol. Interact.*, vol. 4, no. 4, p. 91, Dec. 2020.

[28] A. Mitre-Ortiz, J. Muñoz-Arteaga, and H. Cardona-Reyes, "Developing a model to evaluate and improve user experience with hand motions in virtual reality environments," *Universal Access Inf. Soc.*, vol. 22, no. 3, pp. 825–839, Aug. 2023.

[29] P. Monteiro, G. Gonçalves, H. Coelho, M. Melo, and M. Bessa, "Hands-free interaction in immersive virtual reality: A systematic review," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 5, pp. 2702–2713, May 2021.

[30] M. Mott, E. Cutrell, M. G. Franco, C. Holz, E. Ofek, R. Stoakley, and M. R. Morris, "Accessible by design: An opportunity for virtual reality," in *Proc. IEEE Int. Symp. Mixed Augmented Reality Adjunct (ISMAR-Adjunct)*, 2019, pp. 451–454.

[31] M. Mott, J. Tang, S. Kane, E. Cutrell, and M. R. Morris, "'I just went into it assuming that I wouldn't be able to have the full experience' understanding the accessibility of virtual reality for people with limited mobility," in *Proc. 22nd Int. ACM SIGACCESS Conf. Comput. Accessibility*, Oct. 2020, pp. 1–13.

[32] J. M'utterlein and T. Hess, "Exploring the impacts of virtual reality on business models: The case of the media industry," in *Proc. 25th Eur. Conf. Inf. Syst. (ECIS)*, 2017, pp. 3213–3222.

[33] M. F. Pereira, C. Prahm, J. Kolbenschlag, E. Oliveira, and N. Rodrigues, "Application of AR and VR in hand rehabilitation: A systematic review," *J. Biomed. Inform.*, vol. 111, Nov. 2020, Art. no. 103584.

[34] K. Puritat, P. Thongthip, W. Nadee, O. Sirasakamol, and S. Sangamuang, "Virtual reality locomotion in place for virtual museum exhibitions of culture heritage: Comparing joystick controller, point & teleport and arm swinging," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 12, no. 8, pp. 91–99, Aug. 2022.

[35] Y. Rabhi, M. Mrabet, and F. Fnaiech, "Intelligent control wheelchair using a new visual joystick," *J. Healthcare Eng.*, vol. 2018, no. 1, pp. 1–20, 2018.

[36] M. Ricci, A. Evangelista, A. Di Roma, and M. Fiorentino, "Immersive and desktop virtual reality in virtual fashion stores: A comparison between shopping experiences," *Virtual Reality*, vol. 27, no. 3, pp. 2281–2296, May 2023.

[37] G. D. Ritterbusch and M. R. Teichmann, "Defining the metaverse: A systematic literature review," *IEEE Access*, vol. 11, pp. 12368–12377, 2023.

[38] T. Rose, C. S. Nam, and K. B. Chen, "Immersion of virtual reality for rehabilitation–review," *Appl. Ergonom.*, vol. 69, pp. 153–161, May 2018.

[39] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Design and evaluation of a short version of the user experience questionnaire (UEQ-S)," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 4, p. 103, Jan. 2017.

[40] D. Severijns, F. Van Geel, and P. Feys, "Motor fatigability in persons with multiple sclerosis: Relation between different upper limb muscles, and with fatigue and the perceived use of the arm in daily life," *Multiple Sclerosis Rel. Disorders*, vol. 19, pp. 90–95, Jan. 2018.

[41] M. Teófilo, V. F. Lucena, J. Nascimento, T. Miyagawa, and F. Maciel, "Evaluating accessibility features designed for virtual reality context," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2018, pp. 1–6.

[42] R. F. Woolson, "Wilcoxon signed-rank test," in *Encyclopedia Biostatistics*, vol. 8, Jul. 2005.

[43] L. Yang, M. Wang, and T. Li, "Hand tracking based on the Kinect and Kalman filter," in *Proc. IEEE 3rd Int. Conf. Inf. Syst. Comput. Aided Educ. (ICISCAE)*, Dalian, China. IEEE Press, Sep. 2020, pp. 708–714.



RUBÉN GRANDE received the bachelor's and master's degrees in computer engineering, in 2022 and 2023, respectively, where he was honored with the award for the best academic record. He is currently pursuing the Ph.D. degree with the University of Castilla-La Mancha, focusing on the application of virtual reality and artificial intelligence in e-commerce, supported by a research training grant. He is an Active Contributor to the scientific community, authoring various papers for international conferences and scientific journals. His research is dedicated to evolving e-commerce by integrating virtual reality and artificial intelligence technologies to enhance inclusivity and accessibility, particularly for small-scale commerce, and individuals with disabilities.



JAVIER ALBUSAC received the degree in computer engineering and the European Ph.D. degree from the University of Castilla-La Mancha (UCLM), Ciudad Real, in 2005 and 2010, respectively. He is currently an Associate Professor and the Vice Director of the School of Computer Science, UCLM. He is the author of more than 30 articles in high-impact journals. His current research interests include the application of virtual reality and artificial intelligence to rehabilitation for individuals with motor disabilities



VANESA HERRERA received the degree in computer engineering and the master's degree in advanced information technologies from the University of Castilla-La Mancha (UCLM), where she is currently pursuing the degree. With over a decade of experience in the private software development sector, she has participated in various projects, contributing to her extensive industry expertise. Since 2012, she has been a member of the Artificial Intelligence Rendering (AIR) Research Group, UCLM. In collaboration with the National Hospital for Paraplegics in Toledo, she works on a national research project developing virtual reality applications aimed at facilitating upper limb recovery in spinal cord injury patients. In this context, she is currently working on her Ph.D. thesis.



DOROTHY MONEKOSSO received the bachelor's degree in electronic engineering, the master's degree in satellite engineering from the Surrey Space Centre, University of Surrey, and the Ph.D. degree in spacecraft engineering, in 2000. She is currently a Professor of computer science and Deputy Dean Research and Enterprise with the College of Engineering and Physical Sciences, Aston University. She began her career in space sector, developing onboard computers and control systems for spacecraft and satellites. She became interested in artificial intelligence during her Ph.D. applying machine learning methods and techniques to spacecraft autonomy. Her current research applies the same techniques to develop innovative healthcare technologies. Based on this work, she was awarded a Royal Academy of Engineering, Engineering Foresight Award. In 2020, she was awarded an Honorary Fellowship of the British Computer Society for work in assistive and rehabilitation technologies.



ANA DE LOS REYES-GUZMAN received the Ph.D. degree in biomedical engineering from the Polytechnic University of Madrid, in 2015. She is currently a Biomedical Engineer with the Biomechanics and Technical Aids Department, National Hospital for Paraplegics, Toledo, Spain. Her research has focused on the rehabilitation of movement disorders, particularly through the development and clinical application of virtual reality and low-cost technologies for upper limb rehabilitation in spinal cord injury patients. Her extensive work includes leading innovative projects that integrate advanced biomechanical models and rehabilitation devices to improve the functional assessment and recovery of motor abilities in patients with neurological disorders.



DAVID VALLEJO received the Ph.D. degree in computer science from the University of Castilla-La Mancha (UCLM), Spain, in 2010. He is currently an Associate Professor of computer science with UCLM, where he was the Academic Director, from 2016 to 2020. His research interests include eHealth, multi-agent systems, distributed programming, augmented reality, and gamification.



J. J. CASTRO-SCHEZ received the M.Sc. and Ph.D. degrees in computer science from the University of Granada, Spain, in 1995 and 2001, respectively. He is currently a Full Professor with the Department of Information Technologies and Systems, University of Castilla-La Mancha, Spain. He is a member of the Artificial Intelligence and Rendering Research Group. His research interests include artificial intelligence applied to real-world problems, more specifically, the study and development of formalisms of knowledge representation, knowledge acquisition, learning systems, decision systems, and knowledge-based systems.

Artículo IV: Gamifying Virtual Reality shopping experiences: a case study on a real-world scenario

- **Autores:** R. Grande, J. Albusac, S. Sánchez-Sobrino, D. Vallejo, J.J. Castro-Schez.
- **Tipo:** artículo de revista.
- **Revista:** Virtual Reality.
- **Editorial:** Springer.
- **ISSN:** 1359-4338.
- **Año:** 2025.
- **DOI:** `https://doi.org/10.1007/s10055-026-01323-4`
- **Categoría:** Computer Science, Software Engineering; Imaging Science & Photographic Technology.
- **Factor de Impacto (2024):** 5.
- **Factor de Impacto a 5 años (2024):** 6.6.
- **Ranking JCR:** Q1.
- **Objetivos relacionados:** O6 (SO6.1, SO6.2, SO6.3, SO6.4).
- **Repositorio:** `https://github.com/AIR-Research-Group-UCLM/Mancha comic-case-study`

Virtual Reality (2026) 30:65
https://doi.org/10.1007/s10055-026-01323-4

ORIGINAL ARTICLE

Gamifying virtual reality shopping experiences: a case study on a real-world scenario

R. Grande[1] · J. Albusac[1] · S. Sanchez Sobrino[1] · D. Vallejo[1] · J. J. Castro-Schez[1]

Received: 29 July 2025 / Accepted: 15 January 2026
© The Author(s) 2026

Abstract
Virtual Reality (VR) is transforming how consumers interact with products in e-commerce, enabling immersive, engaging, and emotionally connected experiences that resemble those of physical retail stores. However, the extent to which these experiences influence actual purchase intention and behaviour remains largely unexplored. This paper presents an exploratory study examining how immersive and gamified shopping experiences affect consumer behaviour. A custom VR application was developed with two scenarios: one featuring real products digitized using Neural Radiance Fields (NeRF), and another offering a gamified experience based on a graphic novel, requiring users to solve a puzzle to follow the storyline. The study involved 48 participants during a real commercial event, where they could use their own money to purchase the products explored in the virtual store or the graphic novel from the gamified scene, facilitating authentic purchase decisions. A mixed-methods approach was employed, combining qualitative feedback with quantitative analysis of in-app behaviour. The study extracted a range of space–time and kinematic metrics, such as interaction frequency, navigation efficiency, and locomotion patterns, to characterize user behaviour. Additionally, unsupervised clustering techniques were applied to identify distinct behavioural profiles, revealing significant differences between buyers and non-buyers. The results, including a conversion rate of 35.42%, show that specific user behaviours and attitudes are closely linked to final purchasing decisions. This study provides empirical evidence of the transformative role of VR and gamification in the evolution of e-commerce, highlighting their potential to enhance user engagement and drive effective purchase behaviour in real-world commercial contexts.

Keywords Virtual reality · E-commerce · Gamification · Consumer behaviour · User engagement

1 Introduction

In recent years, electronic commerce has gained notable interest research, motivated by the widespread adoption of digital technologies and evolving consumer habits. To remain competitive, retailers and brands are experimenting with innovative modes of customer interaction that aim to replicate and improve traditional in-store experiences (Castro-Schez et al. 2024). Within this landscape, emerging technologies such as Virtual Reality (VR) and Augmented Reality (AR) have become crucial, providing the means to craft immersive and highly personalised shopping environments (Xi and Hamari 2021). Interest in these tools now stretches well beyond the context of e-commerce, drawing attention from a broad range of industries (Mütterlein and Hess 2017).

VR, in particular, has proven effective at simulating and extending physical experiences by letting users handle

✉ R. Grande
ruben.grande@uclm.es

J. Albusac
JavierAlonso.Albusac@uclm.es

S. Sanchez Sobrino
Santiago.Sanchez@uclm.es

D. Vallejo
David.Vallejo@uclm.es

J. .J. Castro-Schez
JoseJesus.Castro@uclm.es

[1] Department of Information Technologies and Systems, University of Castilla-La Mancha, Paseo de la Universidad 4, Ciudad Real 13071, Spain

Published online: 17 February 2026

products and explore virtual spaces in previously impossible ways. These immersive interactions not only broaden the horizons of online retail but also provide a substantial influence on consumer behaviour by enhancing product perception and strengthening purchase intent (Khatri et al. 2022). One of the most promising applications is gamification, which integrates playful, game-like mechanics into the shopping process and thereby boosts user engagement and motivation (Azmi et al. 2021a).

Although VR and gamification applications are gaining popularity, significant knowledge gaps remain. Nevertheless, various studies examine theoretically or qualitatively (Peukert et al. 2019; Xue et al. 2020; Mkedder et al. 2024; Afwa et al. 2024) the impact of VR on purchase intention, most are conducted in controlled experimental settings where participants express their purchase intentions or make purchases using simulated money or virtual currencies (van Herpen et al. 2016; Speicher et al. 2017; Gil-López et al. 2023; Moghaddasi et al. 2021), i.e., a virtual quantity of money with no real value associated. These conditions limit how well earlier findings translate to real shopping situations. In addition, the joint effect of VR and game features in real stores, where buying choices are shaped by feelings, social cues, and the setting, has received limited scholarly attention (Donatiello et al. 2018). Moreover, works that address the effects of game elements on online retail have gained research interest recently (Azmi et al. 2021b; Rahmadhan et al. 2023). Current evidence links gamification with higher purchase and repeat-purchase intention, higher buying frequency, and stronger perceived usefulness (Aparicio et al. 2021).

This paper has the following goals: (i) conduct a case study in which participants' own money is involved instead of allocating virtual currencies to be used, (ii) measure, both quantitatively and qualitatively, how VR and gamification influence the shopping experience and the purchase intention of volunteers and (iii) to evaluate the relationship between user behaviour during the experimentation with their purchase behaviour. We address these gaps through a field experiment at a three-day annual fair that convenes authors to present their works, while also providing a marketplace for local businesses to offer and sell their products. Visitors who voluntarily decided to try out the VR application could view and interact with digitised products, which they could then choose to buy them later by visiting the physical stand of the business with their own money, as they would do if they had the VR application in their home. This design enhances the external validity of our findings and offers actionable guidance for small retailers who want to adopt immersive technology as a sales tool (Peukert et al. 2019). It is important to note that the decision to purchase remained entirely in the hands of the participants after trying the VR application. As a result, the psychological implications differ from scenarios in which users are provided with virtual currency.

To build a clear theoretical foundation for this study, we adopt the Stimulus-Organism-Response (S-O-R) framework (Jacoby 2002). This model proposes that environmental features act as stimuli (S) that influence an individual's internal cognitive and affective states (O), which subsequently drive behavioral responses (R). In our context, the gamified VR application serves as the primary stimulus designed to enhance the user's internal organismic state, such as immersion. This internal state is expected to manifest in a specific behavioral response, encompassing both in-app interaction patterns and the final, real-money purchase decision.

Based on this framework, we test the following hypotheses:

- **H1:** A user's positive internal state (Organism), measured through self-reported attitudes like gamified engagement and usability or immersion, is positively associated with their purchase intention and subsequent purchase behavior (Response).
- **H2:** A specific behavioral profile (Response), characterized by high interaction intensity and active exploration, is positively associated with a higher likelihood of making a real-money purchase.
- **H3:** The gamified VR experience (Stimulus) is an effective mechanism for driving purchase conversions in a real-world commercial setting.

Working with a Publishing House that also owns retail physical stores, we constructed a virtual replica of one part of its physical store. Techniques such as Neurance Radiance Fields (NeRF) were used to capture a set of items, providing high-quality 3D models to include in the VR application that was developed for our study. We then turned the path to one specific product, a €20 graphic novel, into a short gamified experience. The volunteers followed the plot, solved small puzzles tied to the storyline, and could unlock the ending only by purchasing the book. It is important to mention that we did not contact the volunteers previously, and none of them knew the graphic novel or the products exhibited in the VR scene. A subjective Likert-like questionnaire was answered by each of the 48 volunteers. Finally, we analysed the data collected and structured it into behavioural features to research if behaviour has any impact or relationship with purchase decision and the attitude showed in the questionnaire answers. The approach followed was to perform unsupervised learning (to check if there are differences in behaviour that lead to purchase wihtout knowing who purchased) as well as supervised learning solutions (once we

know who bought, what are the key factors that made this decission?).

The remaining paper is structured as follows. Section 2 provides related work about case studies of VR in the context of e-commerce, as well as works about consumer behaviour, including purchase intention. Moreover, papers that researched the effects of gamification in e-commerce and VR are described. Section 2.3 details the background of this paper, describing the research project in which it is framed. Section 3 shows the case study methodology employed and Sect. 3.2 details technical aspects regarding the implementation of the application used in the case study. Finally, Sect. 4 provides the qualitative and quantitative results of the study, while Sect. 5 discuss the results of this paper and the future work is presented in Sect. 6.

2 Related work

In this section, we first present those studies that have focused on the use of VR in the context of e-commerce, emphasizing psychological aspects of customer behaviour. Then, we review recent work on the application of gamification in VR and e-commerce applications.

2.1 VR shopping experiences and consumer behaviour

As outlined in the previous section, VR promises to be one of the technologies that will drive the evolution of e-commerce. In this regard, Peukert et al. (2019) conducted a study to understand why - and under what conditions - consumers would adopt fully immersive virtual-reality shopping environments. Building on earlier work in e-commerce that highlights both utilitarian (task-oriented) and hedonic (experience-oriented) motivations, they compared two versions of a virtual supermarket shelf. They tested two display modes, a VR headset versus a desktop screen, and proposed that immersion would boost both utilitarian outcomes (by improving perceived product diagnosticity) and hedonic outcomes (by enhancing enjoyment through telepresence). Their results showed that while immersion significantly increased enjoyment, it actually made product details harder to read, undermining diagnosticity; these opposing effects nullified one another on overall reuse intention. To encourage realistic decision-making, participants began with a "virtual" €14 budget and completed ten product-choice tasks, selecting from various muesli products. Wölfel and Reinhardt (2019) extended this line of inquiry by manipulating shelf placement, product quantity, scale, and context in a VR setting with 70 participants, revealing that eye-level placement, multiple units, and slight up-scaling (+10%) all enhanced purchase intentions, without any monetary endowment or real-purchase tasks.

In a similar approach to the previous study, Xi et al. (2024) aimed to determine how AR and VR shape consumers' perceptions and revisit intentions in a controlled retail setting. In a 2 × 2 between-subjects design, 157 university students spent ten minutes each in one of four shopping conditions (AR on/off × VR on/off) before completing validated scales for perceived ease of use, usefulness, enjoyment, attitude, and intention. To ensure authentic decision-making, participants shopped in a purpose-built second-hand record store stocked with 600 LPs and received a €10 gift card as an experimental budget, which they could spend and then keep the records they "purchased". Results showed that AR reduced perceived ease of use without harming overall attitude, while VR increased enjoyment but not usefulness; notably, enjoyment and perceived usefulness both drove positive attitudes and revisit intentions despite any usability challenges.

Now, we will focus on papers addressing consumer behaviour and psychological aspects of using VR for shopping experiences. The work of Mkedder et al. (2024) researched how various aspects of VR stores interact to shape consumers' purchase intentions by applying a multi-level, expert-driven modelling approach. After a literature review, they found twenty-one potential determinants down to ten key factors, ranging from e-service quality and perceived ease of use to presence, enjoyment, escapism, trust, and satisfaction. Then, they recruited eight academics and industry practitioners to structure their interrelationships using Interpretive Structural Modeling (ISM). Through pairwise comparisons, a Structural Self-Interaction Matrix, reachability matrices, and level partitions, the authors built a five-layered hierarchical framework; subsequent MICMAC analysis then classified e-service quality as the dominant independent driver, identified perceived ease of use, perceived usefulness, presence, enjoyment, and escapism as linkage factors, and placed shopping experience, trust, attitude toward VR stores, and satisfaction at the dependent top level.

The research effort of Wu (2023) focused on developing and testing a theoretical framework grounded in Self-Determination Theory (SDT) for understanding how virtual reality (VR) and augmented reality (AR) service technologies influence consumers' purchase intentions in the fashion-brand context. After conducting in-depth, semi-structured interviews to identify core antecedent variables (perceived value, enjoyment, informativeness, presence, and consumer experience), the author fielded an online survey of 675 fashion-brand consumers in Taiwan (70% female; 44% aged 16–25) between April and June 2022. Respondents completed a 32-item questionnaire on seven-point scales,

and the data were analysed using structural equation modelling (SEM) to test hypothesized paths, alongside fuzzy-set qualitative comparative analysis (fsQCA) to explore necessary and sufficient conditions. The results demonstrate that perceived value, enjoyment, informativeness, presence, and consumer experience each exert significant positive effects on purchase intention, with perceived value, enjoyment, and informativeness also enhancing the overall consumer experience, which in turn further enhances purchase intention.

Many brands exploit techniques and features to promote impulsive buying behaviour. The work of Chen et al. (2022) researched whether VR's interactivity and vividness drive impulsive buying in immersive virtual stores by applying the Stimulus–Organism–Response framework. In a lab experiment with 227 participants, each used an Oculus Gear VR headset and a custom shopping app to explore one of eight store scenarios (varying interactivity, vividness, and product type) for 5–15 min, then completed measures of telepresence, perceived diagnosticity, playfulness, and impulse-buying urge. Their results show that both interactivity and vividness significantly enhance telepresence, diagnosticity, and playfulness, each of which in turn predicts a stronger urge to buy impulsively, with trait impulsiveness further amplifying this effect. With the same goal, the work of Moghaddasi et al. (2021) wanted to determine whether shoppers' trait impulsivity can be inferred from their behaviour in a fully immersive VR supermarket. In their study, 57 adults (27 women, 30 men; mean age 25.1 ± 5.1 years) navigated a 6 × 6 m virtual store, using an HTC Vive Pro with integrated eye-tracking and hand controllers, through three tasks: a free 5-minute exploration, a goal-directed snack purchase with a €5 budget, and a sneaker-buying task with €180 to spend. The authors captured four non-invasive data streams - eye movements, head/hand kinematics (navigation and posture), and interaction events - and distilled these into zonal, temporal, spatial, and kinematic features before training support-vector-machine classifiers via stratified 10-fold cross-validation (50 repeats). They found that combined human-behaviour tracking (HBT) signals yielded up to 87% accuracy in distinguishing high vs. low impulsivity, with posture and interaction features during free exploration proving most informative.

The study of Gil-López et al. (2023) aimed to identify shoppers' age and gender that can be inferred from their behaviour in a fully immersive VR supermarket. In their study, 57 adults (27 women, 30 men; ages 18–36) used an HTC Vive Pro with eye-tracking and hand controllers to complete three tasks: free exploration, a €5 snack-buying task (up to three items), and a €180 shoe-buying task. Participants kept whatever they purchased, creating real budget constraints to elicit authentic choices. By extracting features from eye movements, head and hand motions, navigation paths, and interaction events, the authors trained SVM classifiers that achieved over 70% accuracy for both demographics (peaking at 94% for gender and 84% for age) demonstrating that VR behavioural data can reliably predict shopper profiles when real monetary incentives are used.

2.2 Gamification in VR and e-commerce

Gamification has become a popular strategy for e-commerce platforms seeking to improve customer engagement, loyalty, and purchase behaviours. Rahmadhan et al. (2023) conducted a systematic literature review following Kitchenham's guidelines to identify trends and impacts of gamification in e-commerce. From an initial pool of 550 articles, they reduced this quantity through their criteria to 25 high-quality studies published between 2016 and 2021. Their analysis revealed that points, badges, and leaderboards are the most frequently deployed elements, while immersion-focused mechanics are emerging as a promising avenue for future research. Across these studies, gamification was demonstrated to enhance consumer loyalty, engagement, and various user behaviours (e.g., impulse buying, gifting).

The works of Faganel et al. (2024) and Donatiello et al. (2018) both investigate how game-like features can enrich online and virtual fashion experiences, yet they approach the topic from complementary angles. Faganel et al. surveyed 118 Slovenian online shoppers aged 16–65 via a 12-item questionnaire to assess how points, badges, and "top reviewer" endorsements affect each stage of the purchase funnel. They found that gamification most strongly influences the alternatives-evaluation phase and reviewer trust, though only a third of respondents were willing to spend extra for rewards. In contrast, Donatiello et al. invited 13 master's-level fashion students (ages 20–35) to play "Fashion Island", a Unity-based VR fitting-room game in which users dress an avatar on a tropical island using HTC Vive controls. Measuring cybersickness and usability through standard questionnaires, they reported minimal discomfort and high creativity support and satisfaction, with over 90% of participants eager to reuse the system for shopping or entertainment, again without any virtual money incentives. Together, these studies suggest that immersive VR gamification delivers strong user engagement and usability.

Building on prior research that highlighted how gamification elements can address consumer decision processes in e-commerce and enhance user engagement and satisfaction in immersive VR fashion applications, Wei et al. (2023) addressed a key challenge in VR tourism: user fatigue. To explore this, Wei et al. conducted two interlinked studies with a total of 233 participants. In Study 1, 60 adults (15 men, 45 women; aged 18–27) were randomly assigned to a 50–60 min lab-based VR tour of four "Journey to the

West" scenes, either with or without integrated gamification elements (quests, points, feedback, storyline). Using validated scales for four fatigue dimensions and satisfaction, they found that gamified participants reported significantly lower physical, motivational, affective, and cognitive fatigue alongside higher satisfaction, and that only motivational fatigue fully mediated the link between gamification and satisfaction. Study 2 then isolated the four most common elements (challenge, feedback, reward, leaderboard) in an online 360° Palace Museum tour with 173 new users, revealing that each element reduced motivational fatigue and, in turn, boosted satisfaction, with feedback and rewards proving most effective.

The papers addressed showed that gamification in the context of VR e-commerce warrants further exploration, as well as engaging participants of case studies in contexts where real money is used to realistically evaluate the impact of the VR experiences proposed in consumer behaviour.

2.3 VRZOCO research project

In this section, we briefly describe the research project within which this work is framed, aiming to provide a holistic overview of both the project and the present study, to help contextualize the importance of our research. The project, entitled VRZOCO,[1] has as its main objective the design and development of an e-commerce platform based on VR and Artificial Intelligence (AI) that helps strengthen the connection between consumers and local businesses, whose operations have been negatively impacted since the COVID-19 pandemic (Kim 2022).

One of the project's innovative features is the implementation of a VR version of the portal, leveraging the capabilities of this immersive technology to allow the display and manipulation of virtual products (Castro-Schez et al. 2024). Acknowledging that many individuals still cannot access or afford VR headsets, a web version of the platform is also provided.

One method the platform employs to foster the enhancement of local economies is establishing synergies among small local businesses. This can be achieved through direct cooperation on the platform, pooling efforts to maintain a single point of contact, either via the web or the VR application. Thus, technological efforts are focused on a unified portal rather than providing one per business. Additionally, the platform offers indirect synergy: using fuzzy-logic and collaborative filtering algorithms, VRZOCO employs data collected from user activity on both the web and VR platforms (Arcos et al. 2024) to construct a user profile for personalized recommendations. After each session, the system recalculates preferences based on any interactions to update recommendations for that user. To generate synergies without requiring business intervention, the system computes and offers leisure or cultural plans around the city with each purchase. These plans are based, among other factors, on weather forecasts, time of day, temporary city events, and, most importantly, the user's preferences and feedback on previous plans (Garrido et al. 2024). These recommendations are provided regardless of whether the user chooses to pick up the product in-store or receive it by delivery. In the former case, a route is calculated to allow the user to enjoy local events or businesses. In doing so, a strong bond is established between the focal business and the rest of the city's enterprises, with the intention of boosting the local economy.

Within the context of the present work, we intend to start research line focused on studying the synergies and influences of methods and techniques that can improve user engagement with a VR-based e-commerce platform. Specifically, this study will address how gamification can enhance this aspect, as well as help understand natural user behaviour in immersive shopping environments, which can be an invaluable source of data for the platform's recommendation system. Finally, Fig. 1 illustrates the architecture with the modules most relevant to this study. While the Product Digitisation Module enables us to obtain 3D models of the products to be displayed in the virtual environment, the VR Information Collector gathers user activity data during the conducted experiments for subsequent analysis. Lastly, the Output Layer provides the visualization of the developed virtual environments for the case study, utilizing the development technologies described in Sect. 3.

3 Case study

In this section, we will detail the context, methodology, and technological implementation necessary for the case study carried out to answer the research questions presented in Sect. 1.

3.1 Context

The case study was conducted within the framework of a major event held in the city where our university and research group are located. This annual fair, called "*Manchacómic!*"[2] (hereafter referred to as the *fair/event*), features over 50 independent businesses and artists from the city and, more broadly, the surrounding region (Castilla La-Mancha), with partial funding from the Ministry of Culture and Sport and

[1] https://airproy.esi.uclm.es/vrzoco/vr-project

[2] https://manchacomic.es

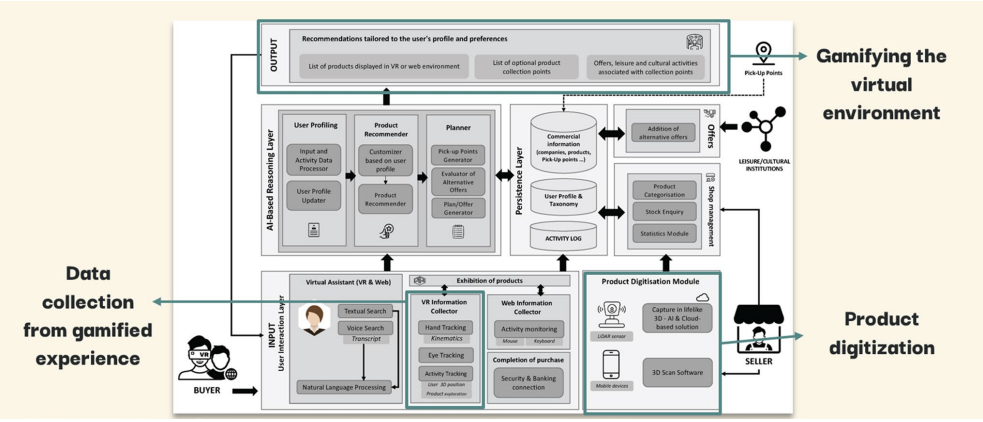


Fig. 1 Key modules of VRZOCO's architecture for this case study

the Spanish Cultural Association. Each exhibitor offers its products at an individual stand within the event venue. Held over three days each year, the fair-event hosts various creators showcasing their works or latest releases, including graphic novels, indie video games, and other artistic creations. Additionally, a range of educational activities is offered, such as workshops, author talks, and panel discussions. Our case study took place on one of the three days of the event.

Among the participating businesses, in the months preceding the event we collaborated with one in particular to conduct the case study at *Manchacómic!*. This partner was *Serendipia*,[3] a prominent regional enterprise with its own book publishing arm, a science club, and brick-and-mortar stores selling authorial works and other products. The significance of this business in the sector is underscored by its recent receipt of a national award granted by the Federation of Publishers' Guilds of Spain. Although they operate physical stores year-round, during the fair-event they also set up an individual stand to offer products to interested attendees. Initial contact was established through collaboration with the city's Chamber of Commerce, the body responsible for representing, promoting, and defending general commercial interests, as well as providing services and support to businesses to enhance competitiveness and facilitate development both nationally and internationally. Our collaboration with *Serendipia* enabled us to:

1. Digitize one of their physical stores using a 360° camera to recreate it virtually in the VR application used for the case study. This approach was intended to simulate for fair-event attendees the experience of visiting the actual store, even though they were in the event venue. Although the digitized result can be viewed on the web portal (https://airproy.esi.uclm.es/vrzoco/tiendas/streetview/4), in Sect. 3.2 we discuss the rationale for opting for a slightly less faithful virtual replica rather than the full scan.
2. Create high-quality 3D models of a heterogeneous set of products available in the store by employing a low-cost NeRF-based tool (Grande et al. 2024), so that volunteers could purchase items of interest following the experimental sessions.
3. Establish a collaboration with the creators of the graphic novel *Pedro Girón* from the series *Calatrava*, set in the medieval and cultural history of our region. This close collaboration was essential to obtain their consent and to faithfully respect the artists' work when virtually representing both the environments and the protagonist, Pedro Girón. Fig. 2 displays the cover and selected panels from the graphic novel.
4. Coordinate with the fair-event organizers, facilitated by *Serendipia*, who allocated us a space within the venue and included our experiment in the event website's program, equivalent to any other scheduled activity, without disclosing study details.

After having provided the context, the theme, and how we secured participation in *Manchacómic!* through our collaboration with *Serendipia*, we now describe the design and implementation of the VR application used during the case study, as well as the steps taken to conduct it. The scenes in the application are as follows:

[3] https://serendipiaeditorial.com



Fig. 2 **A** the cover of the €20 graphic novel named "Calatrava: Pedro Girón". **B** extract from the graphic novel, showing the castle where the gamified experience is set

- The first scene contained the virtual replica of *Serendipia*'s physical store, and was the first that volunteers experienced upon launching the VR application. In this scene, they could freely interact with and inspect digitized products (see Fig. 4) that were on sale at *Serendipia*'s stand. Most of the digitized and displayed items were anime collectible figures and cartoon art, since these products offer more information when viewed in VR than, for example, a book. When a volunteer wished to proceed to the next scene, they would interact with a door at the back of the store, which then displayed a "loading screen" (see the top image in Fig. 7) presenting the title of the graphic novel featured in the second scene: the gamified experience of *Calatrava: Pedro Girón*.
- The second scene, as mentioned above, comprised the gamified experience of the graphic novel. To exploit VR's immersive capabilities, we modeled an interior courtyard of the "Calatrava la Nueva Castle" in Blender,[4] shown in Fig. 2.B. In this virtual environment, the novel's protagonist, *Pedro Girón*, appears before the user and begins narrating the story's opening from a first-person perspective. Once introduced, volunteers had to explore the environment to find pieces that formed a puzzle of the novel's cover. For each piece found, *Pedro Girón* continued the narration. Upon completing the puzzle, the protagonist delivered a final segment of the story, leaving an open ending and encouraging the user to discover the full narrative in the graphic novel. We chose the puzzle mechanic to present a challenge that would keep participants motivated (Werbach and Hunter 2012), as well as to encourage environment exploration, thereby fostering greater immersion in the novel's experience and storyline.

The following subsections will briefly detail, at a technical level, the implementation of the two VR scenes that comprise the application used in the case study, along with the limitations and challenges encountered. These subsections are not intended to present technical novelties, but rather to document the work required in the VR domain to execute the case study. Technology selection was based on the authors' prior expertise, favoring open-source tools such as Blender and free-to-use widely used platforms like Unity, to ensure the developed software can serve as a foundation for replication by the scientific community using accessible, open technologies.

3.2 First scene: *Serendipia*'s virtual store

As mentioned previously, we visited *Serendipia*'s physical store on two separate days during approximately two-hour sessions, with the following objectives: (i) to scan the store for a faithful virtual reconstruction, and (ii) to scan products intended for display at their *Manchacómic!* stand for inclusion in the virtual environment. Figure 3A and B show part

[4] https://en.wikipedia.org/wiki/Calatrava_la_Nueva



Fig. 3 Inside *Serendipia*'s store. **A** and **B** show photographs taken in it, while **C** and **D** are parts of the 360 degree scan made. **E** one of the owners of *Serendipia* trying out a prototype of the VR application used in the case study



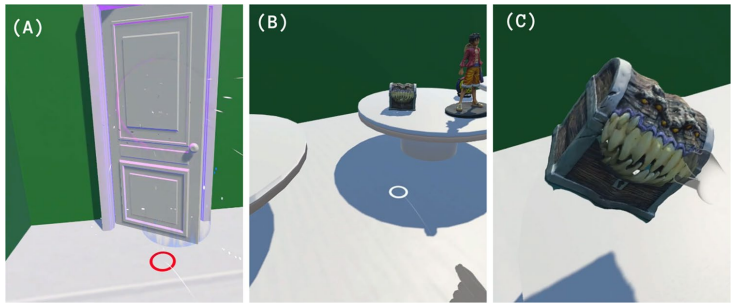
Fig. 4 A set of screenshots of the virtual replica of *Serendipia*. All 3D models were obtained using Luma AI, and they were made interactable with Meta SDK for Unity

of the store's interior, while Fig. 3C and D depict portions of the 360° scan, available at https://airproy.esi.uclm.es/vrzoco/tiendas/streetview/4.

Despite having discarded that approach for the development of the VR application used in the case study, the 360° scan has proven useful for its inclusion in the VRZOCO project's web portal.To ensure a satisfactory user experience, a simplified virtual environment was created from scratch to include the 3D models of *Serendipia*'s products in it. This environment can be seen in Fig. 4, while Fig. 5 shows the teleportation and grabbing interactions that could be performed, including how to trigger the beginning of the next scene in Fig. 5A.

We built a simplified, performant virtual replica of *Serendipia*'s store to ensure stable standalone VR rendering. Products were digitised using a low-cost NeRF workflow and made interactable with hand tracking via the Meta SDK. Further capture and optimisation details are provided in Appendix B.1 and the public repository.

Fig. 5 Interactions that can be performed in the first scene. While **A** and **B** shows teleportation to be able to move around the environment, **C** shows the grabbing of a product, which can be made larger or smaller to some extent as well as moving and rotating it



3.3 Second scene: gamified graphic novel *Calatrava: Pedro Girón*

This scene, which the volunteer accesses by interacting with the door at the back of the first scene's environment, contains the virtual setting for the gamified experience of the €20 graphic novel *Calatrava: Pedro Girón*. For the development of this scene, we modelled a castle courtyard (see Fig. 6A) and an NPC (see Fig. 6B) that delivers short narrative cues while users solve a cover-puzzle mechanic designed to foster exploration and engagement. The scene supports hand-based interaction and teleport locomotion. Implementation specifics (asset creation, rigging, and animation) are detailed in Appendix B.2. Figure 7 shows a set of screnshots of this scene while the application is running.

Regarding the efforts in developing this second scene (see Fig. 7), we maintained use of the Meta SDK for Unity, enabling users to use their hands to interact with objects and the environment. As in the first scene, the entire environment is navigable via teleportation, since the experiment was conducted in a physical space not large enough for room-scale navigation. Having described Scenes 1 and 2, Table 1 shows the data items collected and to be processed later in Sect. 4.

Before beginning the next section, which presents the quantitative results, a video of the experiment, simulated by one of the authors, can be viewed via the following link to observe both scenes in action: https://github.com/AIR-Research-Group-UCLM/Manchacomic-case-study/blob/master/EXAMPLE_VIDEO.mp4.

3.4 Methodology

Of the three days of the event, the experimental session was conducted on the intermediate day. Table 2 summarizes the design and the procedure carried out to conduct the case study. To reduce expectation effects, participants had no prior information about the activity's content or goals; at most, those who checked the event website could have inferred that VR devices were used..

It is worth noting that volunteers were not given a strict time limit, although they were advised not to exceed 5 min in the first scene so that the total experimental duration per participant would be around 20 min. We had three Meta Quest 3 headsets available during the experimental session and a finite amount of time allocated by the event organizers; therefore, to obtain a moderately large sample of approximately 50 participants, it was necessary to monitor the duration of each participant's session.

Figure 8 shows some of the study participants. Before beginning each participant's session, we instructed them on how to interact within the scenes. By using the Meta Quest3 "hand-tracking", volunteers interacted with the environment using their hands, both to grab and inspect products and to teleport within the scenes. We also explained how to transition autonomously from the first to the second scene, including an overview of the graphic novel on which it was based and the objective (finding puzzle pieces and completing the puzzle).

Upon completing the VR experience, each participant filled out a 15-question survey on topics such as VR as a shopping tool, usability, and user experience. The average time to complete the survey was 5 min and 32 s. Some questions were directly related to purchase intention, allowing us to qualitatively analyze VR's impact on that intention and the potential of gamification to increase it. After finishing the survey, we provided each volunteer with two numbered tickets, which they could present if they chose to make a purchase from *Serendipia*, enabling us to anonymously

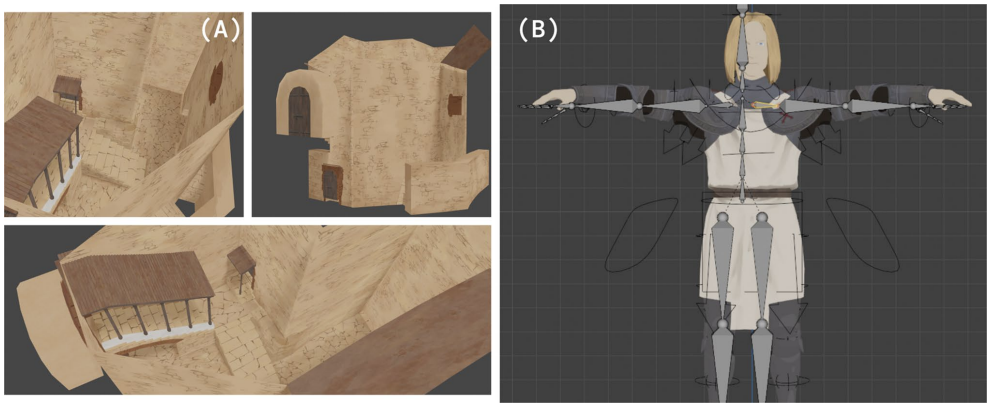


Fig. 6 **A** Calatrava's castle 3D space modelled in Blender. **B** Pedro Girón, graphic novel main character, manually rigged

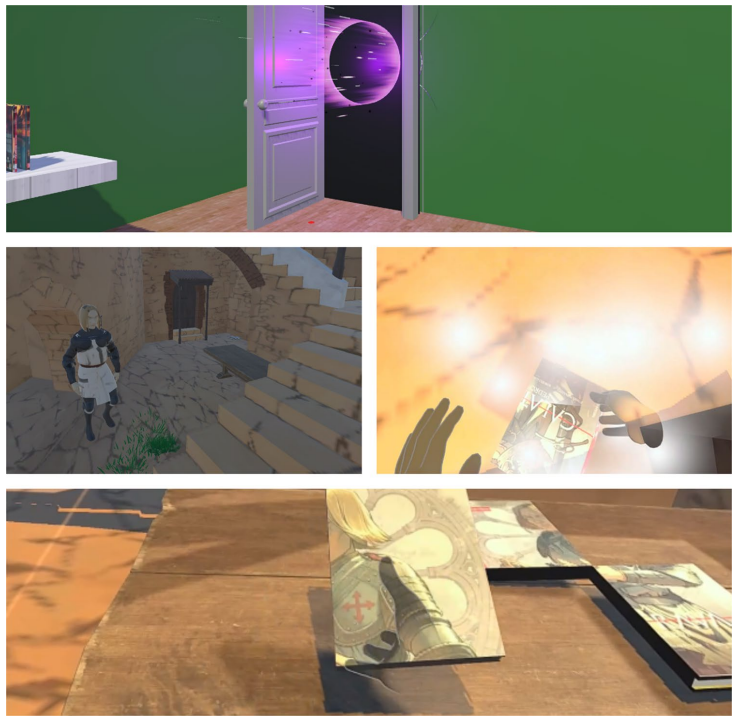


Fig. 7 A set of screenshots taken from the gamified experience scene running. The middle screenshot contains the door that serves as transitions between both game scenes is shown

track which survey ID corresponded to each sale. One ticket identified the sale of at least one product displayed in the virtual replica, and the other identified the sale of at least one copy of the graphic novel *Calatrava: Pedro Girón*. Table 3 briefly details the demographic data collected in the questionnaire for study participants.

4 Results

In this section, we present the results obtained during the conducted experiment. Unless otherwise noted, we report central tendency with mean and variability with standard deviation (SD). For comparative charts we display 95% confidence intervals (CIs). For proportions (e.g., purchase rates) we report Wilson 95% CIs. Heatmaps and dimensionality-reduction visualizations (e.g., Spearman matrix, UMAP) do not include error bars because they are not estimates of a sampling mean.

4.1 Questionnaire results

The survey consisted of 14 Likert-scale questions ranging from 1 to 5, where 1 represents the most negative value and 5 the most positive. These are reported in Table 4, which also includes the mean, standard deviation, mode, and median of the responses. Additionally, Q15 was an optional open-ended question included to gather more free-form feedback and impressions. We obtained a total of 28 responses (58.3% of participants), with respondents averaging 29.3 ± 12.7 years of age, representing a slightly older sample than the mean age of all 48 participants shown in Table 3.

Conducting a brief thematic analysis of the items in Table 4, we find:

- Participants reported little prior experience with VR devices (Q1) yet strong intrinsic interest (Q2) and an even higher sense of novelty (Q3). The combination between low experience and high enthusiasm might suggests that the VR application was encouraging and appealing even to novices. Mean preference for browsing products in VR versus Web sat near the positive side of neutral, indicating that most respondents were either ambivalent

Table 1 Raw data items collected from both scenes

Data item	Type	Range of values	Description
Frame number	Integer	[1,n]	The frame number from the start of the session
Timestamp	Float	[0,n]	The time measured in seconds from the start of the session
Head position	Vector	(x,y,z) coordinates	HMD's position in 3D space
Head rotation	Vector	(x,y,z) euler angles	Euler angles in each of the 3D space axes for the HMD
Product	String	Set of products used Piece{0-3} for second scene	The name of the object in the scene representing the product in Unity
Product's state	String	Disabled, Normal, Hover or Select	State of the interactable product in the Interactor-Interactable lifecycle of Meta SDK
Interactor's state	String	Desabled, Normal, Hover or Select	State of a interactor in the Interactor-Interactable lifecycle of Meta SDK
Right hand state	String	Disabled, Normal, Hover or Select	State of a right hand's Interactor (Grab/Teleport/Turn) in the Interactor-Interactable lifecycle of Meta SDK
Left hand state	String	Disabled, Normal, Hover or Select	State of a left hand's Interactor (Grab/Teleport/Turn) in the Interactor-Interactable lifecycle of Meta SDK
Teleport duration	Float	[0,n]	The time the user spends targeting the teleport destination
Successful teleport	Boolean	True or False	Indicates whether the targeting performed resulted in a successful teleport
Distance	Float	[0,n]	Distance covered in one teleport event
Session_Timestamp	String	Date & time	Datetime identifier of collected data files associated with one user ID
Piece pick-up Piece drop	Float	[0,n]	Timestamp associated with a pick-up and drop of a given puzzle piece
Final drop	Boolean	3D scene dependent	Indicates if that piece was the last one to be dropped (puzzle completed)
Error	Boolean	True or false	True if the piece drop was not done in the table. Might capture false positives such as hand being lost from hand tracking

or already leaning toward VR after the experience, but with appreciable variance that future work could target by clarifying comparative benefits

- Scores on natural interactions (Q5) and immersion (Q6) were high, while reported motion sickness remained very low (Q7). Together, these figures show that the prototype succeeded in creating presence without compromising comfort. The overall experience item (Q8) mirrors these positives, reinforcing that interaction smoothness and presence translate directly into satisfaction.
- The potential shift of purchase intention toward VR (Q9) is slightly above the scale midpoint, whereas willingness to reuse the VR store (Q10) is very high. This discrepancy may suggests that while users are eager to return, conversion to actual purchasing may still depend on complementary factors (e.g., pricing, trust, gamification elements to make it more appealing).
- Motivation to explore products via gamification elements (Q11), perceived entertainment (Q12) and interest in continuing the story (Q13) form a coherent block of high scores that culminates in an encouraging yet slightly lower intention to buy the graphic novel (Q14). However, we will later provide purchase data and analyze it accordingly.
- For most items the median equals or exceeds the mean, indicating slight left-skew (a minority of low scores pulling means downward). This pattern reinforces that favourable judgments dominate the distribution rather than being artefacts of a few extreme ratings

4.1.1 Spearman's correlation matrix

To perform a more in-depth analysis of the Likert-scale responses before applying factorial techniques and constructing reliable composites, we calculated a Spearman correlation matrix. Since these items represent ordinal rather than interval scales, this approach preserves the original coherence and avoids scale-bias. Moreover, as Table 4

Table 2 Summary of the experimental design and procedure

Parameter	Description
Participants	48 volunteers recruited on-site at the "Manchacómic!" event. Participants were event attendees with no prior knowledge of what activity they were going to take part in
Experimental Design	A single-sequence, within-subjects design. All participants experienced both scenes in a fixed order
Session Flow & Duration	The total experience per participant was approximately 20 min and followed these steps 1. Briefing: Instructions on VR interactions (hand-tracking for grabbing and teleporting). 2. Scene 1 - Virtual Store: Free exploration of digitized products. Advised duration was approximately 5 min. 3. Scene 2 - Gamified Experience: Narrative-driven puzzle-solving task based on the graphic novel. 4. Post-Session Survey: Completion of a 15-item questionnaire. The average completion time was 5 min and 32 s. 5. Ticket Issuance: Participants received numbered tickets for the optional purchase phase
Randomization	None. The sequence Scene 1 → Scene 2 was fixed for all participants to maintain a logical narrative flow, transitioning from a general store to a gamified experience for a specific product found within it
Purchase Tracking & Anonymization	The procedure was designed to be anonymous and ecologically valid • After the VR session, participants could use their own money to purchase items at the physical store stand. • To track sales anonymously, each participant received two numbered tickets. • One ticket corresponded to products from the virtual store (Scene 1) and the other to the graphic novel (Scene 2), allowing us to link sales back to the anonymized user ID without collecting personal data

shows a positive skew toward high values, Spearman's ρ does not require normality and is less sensitive to such skew. Finally, this exploratory tool allows us to detect monotonic relationships and identify thematic clusters that could later be confirmed via Exploratory Factor Analysis (EFA).

Figure 9 shows two primary clusters. First, Q9, Q10, and Q11 cluster together: Q10–Q11 correlate at $\rho = 0.66$, and both link to Q9 at $\rho = 0.58$. This indicates that gamification elements in the graphic novel, along with VR technology, not only improve enjoyment but also enhance intentions to reuse similar applications.

A second strong link appears between Q13 and Q14 with $\rho = 0.63$, suggesting that curiosity about the storyline directly converts into willingness to purchase. These findings reinforce the commercial value of gamifying such products in retail.

Moderate positive relationships spread across usability-focused items. Q3 and Q5 correlate at $\rho = 0.53$, while both connect moderately ($\rho \approx 0.55$–0.58) to Q2. Motion-related discomfort (Q7) remains weakly associated ($|\rho| < 0.15$) with nearly all outcomes, indicating that in this sample nausea did not systematically harm user experience or purchase inclination. Q1 exhibits only weak ties ($\rho \approx 0.28$–0.34) to engagement and purchase metrics.

In the next section, we will analyze these correlations in depth by segmenting the sample based on participants' purchase intention.

4.1.2 Segment analysis by purchase intention

In the questionnaire shown in Table 4, we included an explicit purchase-intention question, Q14, by which we will segment the results to observe, based on responses to the other questionnaire items, which factors carry the greatest weight for individuals who scored high versus low on their interest in purchasing after the experience.

1. We conducted an EFA following the procedure recommended by Watkins (2018) to validate whether the Q1–Q13 Likert items shown in Table 4 cluster around the four theorized themes. Q1 was excluded as the segmenting variable from this computation. The full workflow is detailed in Appendix A. This analysis revealed four multi-item factors (VR Appeal; Usability/Immersion; Purchase Intention & Reuse; Gamification's Engagement) plus three single-item indicators (Channel Preference; Discomfort; VR Experience). Treating these last three individually prevents construct dilution.
2. Once the multi-item factors capturing distinct psychological dimensions were obtained, we calculated Cronbach's alpha (α) for each. This statistic measures the internal consistency of composite scales. Generally, $0.6 \leq \alpha \leq 0.8$ is considered to demonstrate acceptable internal consistency (Raharjanti et al. 2022). Table 5 displays the α values for each factor from the EFA, confirming the reliability of the constructs used in the segment analysis.
3. Finally, we applied Welch's t-tests to compare low-intent (Likert 1–3) and high-intent (Likert 4–5) users on each composite. This variant is robust to unequal variances and sample sizes. Effect sizes are reported as Cohen's *d*, allowing practical interpretation alongside statistical significance (Lakens 2013). When the thirteen neutral respondents (score 3) were analyzed as a separate "mid-intent" group, a Welch-ANOVA confirmed a monotonic increase in Gamification's Engagement ($Low < Mid < High$, $F(2, 45) = 7.1$, $p = 0.002$), although only the Low–High comparison



Fig. 8 Photos taken during the experimental session. **A** showing a few volunteers how to interact in the VR application, **B** and **C** show volunteers of different ages taking part of the case study

Table 3 Demographic characteristics of survey participants (n = 48). IQR = Interquartile range

Variable	Category/Statistic	n (%)/Value
Gender	Male	33 (68.8%)
	Female	14 (29.2%)
	No response	1 (2.1%)
Age (years)	Mean (Std. dev.)	23.7 (13.6)
	Median (IQR)	21 (12–29.5)
	Range	8–55

reached post-hoc significance after Bonferroni correction. This pattern reinforces that narrative gamification scales proportionally with purchase intention, even at intermediate levels.

Table 6 presents the segment-analysis results. Although both groups rate the system near the top of the 1–5 scale, participants with high purchase intention are systematically more positive on the motivational and ergonomic composites.

The most pronounced disparity is observed for *Gamification's engagement* ($\Delta M = 0.53$; $p < 0.001$; $d = 1.00$), indicating that story-driven and gamification elements are critical for converting curiosity into buying intent. *Usability/Immersion* also differs meaningfully ($\Delta M = 0.45$; $p = 0.02$; $d = 0.67$), as does *Purchase intention & reuse* ($\Delta M = 0.43$; $p = 0.03$; $d = 0.63$), confirming that smooth interaction and expectations of future use accompany willingness to purchase.

By contrast, *VR Appeal* varies negligibly between groups ($\Delta M = 0.03$; $p = 0.87$; $d = 0.05$), suggesting that novelty and intrinsic interest are already at ceiling even for not experienced users. Likewise, *Channel preference* (Q4) shows only a non-significant medium shift toward VR ($\Delta M = 0.29$; $p = 0.42$; $d = 0.23$), and *Discomfort* (Q7) remains uniformly low ($p = 0.96$; $d = 0.02$). Finally, prior VR experience (Q1) does not differ between segments ($\Delta M = 0.10$; $p = 0.75$; $d = 0.07$), ruling out VR expertise as a direct driver of purchase intention.

Overall, the results suggest that conversion from curiosity to purchase depends primarily on gamified engagement and on users' perception of a seamless, immersive interaction, rather than on the technology's novelty or on motion-sickness concerns.

Before presenting the analysis of the recorded interaction data, we will present the sales data collected during the case study.

4.2 Purchases data

After completing the questionnaire, participants were free to make purchases at *Serendipia*'s stand, either of a product displayed in the virtual replica of Scene 1 or of the €20 graphic novel *Calatrava: Pedro Girón*, which is priced higher than typical graphic novels (approximately €5–€10). The products shown in the virtual store ranged in price from plush toys to collectible figures and board games. Figure 10.A presents the purchase data collected during the case study. Of the 48 participants, 8 (16.67%) bought the graphic novel (none of whom had prior knowledge of it) while 9 (18.75%) purchased a product examined in the first scene, resulting in a total VR application conversion rate of 35.42%.

Separately, *Serendipia* provided us with overall sales data for the featured graphic novel, shown in Fig. 10.B. During the case study period, 8 copies of the novel were sold, matching the total sold on Friday and nearly matching Sunday's total (10). Out of 47 copies sold over the entire

Table 4 All questionnaire's questions along with descriptive statistics from the 48 participants. Please, do note that Q15 is an open question and will be analyzed later on. Moreover, note that from Q9 to Q14 the Likert scale is the same as Q3

ID	Question	Mean	Std. Dev.	Mode	Median
VR expertise and predisposition					
1	What experience do you have with virtual reality devices? (1- None, 2- Little, 3- Occasional, 4- Frequent, 5- A lot)	2.58	1.47	1.0	2.0
2	How interested are you in Virtual Reality and being able to shop using this technology? (1- None, 2- Little, 3- Neutral, 4- Quite a bit, 5- A lot)	4.17	0.86	5.0	4.0
3	How novel do you find the option of shopping on an e-commerce portal with virtual reality? (1- Not novel, 2- Not very novel, 3- Neither novel nor not novel, 4- Novel, 5- Very novel)	4.44	0.68	5.0	5.0
4	Do you prefer exploring store products in VR or via their web portal? (1- Much more on the web, 3- I do not prefer one over the other, 5- Much more in VR)	3.75	1.25	5.0	4.0
Usability and immersion					
5	Did you find the interactions natural and comfortable? (including teleportation) (1- Very little, 2- Little, 3- Neutral, 4- Quite a bit, 5- A lot)	4.21	0.80	4.0	4.0
6	Did you feel immersed in the VR application? (1- Very little, 2- Little, 3- Neutral, 4- Quite a bit, 5- A lot)	4.29	0.85	5.0	5.0
7	Feeling of dizziness or general discomfort (1- None, 2- Little, 3- Neutral, 4- Quite a bit, 5- A lot)	1.54	0.71	1.0	1.0
Purchase and reuse intention					
8	How positive was your experience using both VR scenes? (1- Very negative, 2- Negative, 3- Neutral, 4- Positive, 5- Very positive)	4.29	0.72	5.0	5.0
9	Would your purchase intention change when seeing products in VR instead of on a web page? (1- None, 5- A lot) (Same as Q3)	3.73	1.13	5.0	4.0
10	Would you use such applications again to explore products you're interested in buying? (1- None, 5- A lot)	4.54	0.70	5.0	5.0
Gamification's impact and engagement					
11	Does exploring products through play as seen in the graphic novel motivate you? (1- None, 5- A lot)	4.40	0.80	5.0	5.0
12	Did you find the graphic novel's VR experience entertaining? (1- None, 5- A lot)	4.23	0.83	5.0	5.0
13	Would you be interested in learning more about the comic's story? (1- None, 5- A lot)	4.33	0.83	5.0	5.0
14	Do you intend to buy the graphic novel or any product showed in the VR Store? (1- None, 5- A lot)	3.90	1.12	5.0	4.0
15	Any feedback or improvements you would like to share?	–	–	–	–

event, these figures include sales to attendees who did not participate in our activity and may have been driven by related events, such as the author's graphic novel signing on Friday or an open author talk on Saturday afternoon.

Having presented sales data resulting from case study participants, we now analyze in depth whether participant behaviour within the VR application influences the decision to purchase either the displayed products or the graphic novel.

4.3 Registered data analysis

All CSV files recorded during the case study and used for result generation, as well as the Python scripts employed, are available in the public repository: https://github.com/AIR-Research-Group-UCLM/Manchacomic-case-study. The structure of each CSV file is detailed in that repository.

To provide an overview of the tools and techniques used for analyzing the recorded data, a conceptual schema has been created, shown in Fig. 11.

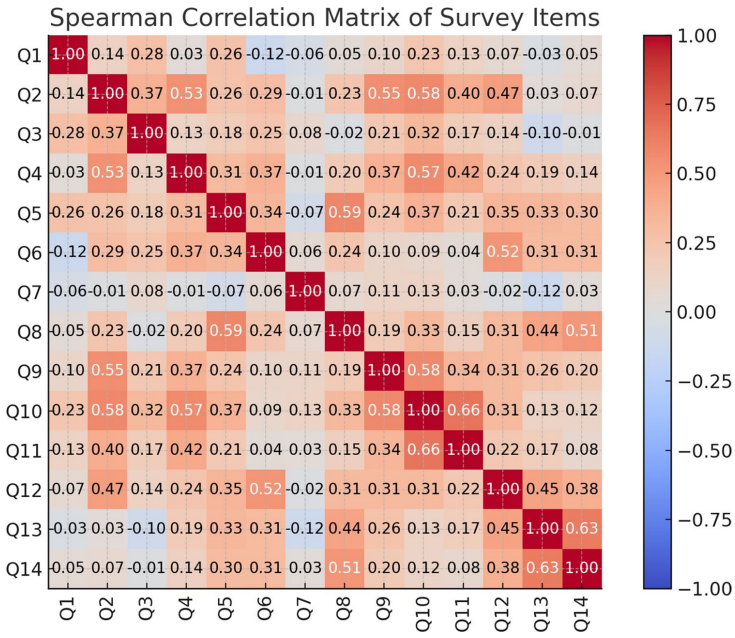


Fig. 9 Spearman Correlation Matrix, including Spearman's ρ values for each cell. Values ≥ 0.60 denote strong positive association, 0.40 to 0.59 moderate, 0.20 to 0.39 weak, and < 0.20 negligible

Table 5 Cronbach's alpha values for each composited formed as a result of the EFA

Composite	Cronbach α
VR Appeal (Q2–Q3)	0.71
Usability/Immersion (Q5–Q6)	0.65
Purchase intention & reuse (Q8–Q10)	0.76
Gamification's engagement (Q11–Q13)	0.77

4.3.1 Clustering process

While the aggregate count of post-experimentation purchases (17 out of 48 participants) provides a relevant high-level outcome, the decision to precede predictive modeling with an unsupervised clustering analysis was driven by the objective of uncovering hidden, data-driven user segments based exclusively on participants' in-VR behavioural patterns. Unlike the singular binary outcome of actual purchase, clustering on these multi-dimensional behavioural features allows for the identification of natural groupings based purely on how they interacted with the virtual environments. We selected a set of behavioural metrics present in VR behavioural analysis literature, which we could compute based on the data we were able to collect.

This method aims to move beyond a simple buyer/non-buyer classification by revealing underlying heterogeneity in user engagement. Such behaviourally-defined segmentation can provide a richer understanding of diverse user types, potentially informing personalized intervention strategies or tailored VR experiences. The subsequent characterization of these emergent clusters then provides a deeper, explanatory context for the predictive performance of the models.

Table 6 Low vs High prior VR experience on composites and single items (Welch's *t*), showing Mean (M), Starndad Deviation (SD), $p - value$ and *d* effect size

	Low (n = 27)		High (n = 21)		p	d
Composite factors	M	SD	M	SD		
VR Appeal	4.29	0.60	4.32	0.66	0.87	0.05
Usability/Immersion	4.12	0.78	4.57	0.52	**0.02**	0.67
Purchase intention & reuse	4.01	0.70	4.44	0.65	**0.03**	0.63
Gamification's engagement	4.21	0.66	4.74	0.34	**0.0001**	1.00
Channel preference (Q4)	3.62	1.24	3.91	1.27	0.42	0.23
Discomfort (Q7)	1.57	1.05	1.55	0.87	0.96	0.02
Experience (Q1)	2.54	1.30	2.64	1.68	0.75	0.07

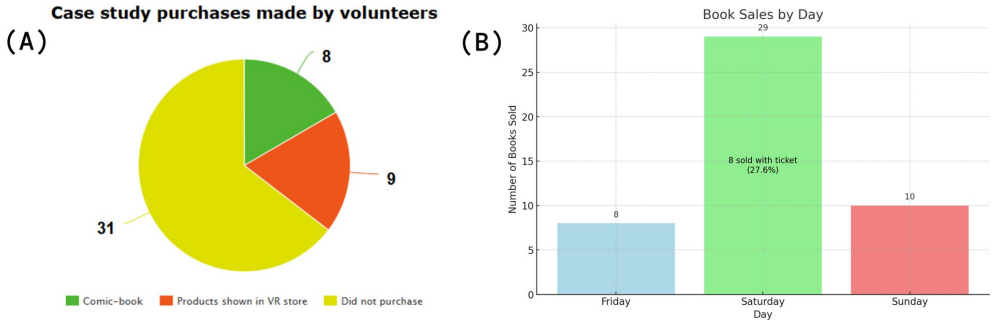


Fig. 10 **A** Purchases made by case study's participants. **B** Sales of "Calatrava: Pedro Girón" graphic novel over the three days of the event

Table 7 presents the behavioural metrics employed for clustering and for the other methods used in the quantitative analysis. We selected a diverse set of variables to establish a comprehensive profile of users based on their behaviour, because each metric captures a distinct aspect of how consumers engage with the two VR scenes, and together they form a compact yet multi-faceted description of user behaviour. Exploration metrics quantify where and how participants move; manipulation metrics provide insights into which items they choose to inspect; fixation duration gauges visual focus; teleport counts and distances summarize navigation strategies; and error count plus puzzle-completion time measure task efficacy. As shown in the last column of Table 7, we prioritized metrics already used in other studies rather than ad-hoc metrics for our VR application (such as `PuzzleCompletionTime` or `ErrorCountS2`), in order for the methods described in this section to be applicable to similar contexts or case studies and to promote generalisation of results.

Although 21 metrics may seem numerous for a dataset of $N = 48$ samples, we applied appropriate contingency measures in clustering and other analyses. We opted to perform a single clustering rather than separate clusterings per scene, which implies splitting some metrics by scene. Scene 1 is a store environment centred on pick-and-inspect products, whereas Scene 2 is gamified and emphasizes locomotion, way-finding, and problem solving. Therefore, we represent each metric, where possible, per scene, hypothesizing that behaviour will differ according to the nature and goals of each scene. This approach allows us, later, to analyse how participants adapt their behaviour by scene or whether their patterns remain stable. Consequently, the resulting clusters capture a broader spectrum of behaviour.

With regards to the metrics selected, using raw event counts (e.g., total number of grabs) is methodologically flawed as it is heavily mitigated by the total time a user spends in the virtual environment. Therefore, a longer session would naturally produce higher counts, irrespective of the user's actual interaction patterns. To mitigate this, we employed for a set of the features computed two standard feature engineering techniques. First, we computed temporal metrics (e.g., GrabsPerMin, PathLenPerMin) by normalizing event counts by the duration of the user's session. This transforms the features from a measure of total volume to a measure of rate or intensity, providing a much more accurate representation of a user's characteristic pace and activity level. Second, we used aggregate metrics (e.g., MeanTeleportDist, MeanContactTime) to capture the quality and nature of interactions, rather than just their frequency. For instance, MeanTeleportDist distinguishes between a user who makes many short, cautious movements and one who makes fewer, more confident leaps across the environment. By using these normalized and aggregated metrics, we ensure that our subsequent clustering analysis identifies genuine differences in behavioural patterns, not artifacts of session length.

We employed a two-stage approach combining Uniform Manifold Approximation and Projection (UMAP) with Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN). This pipeline was chosen to effectively address the "curse of dimensionality", a common challenge in high-dimensional spaces with small sample sizes, where distances between points can become less meaningful.

First, UMAP was used to perform dimensionality reduction. As a manifold learning technique, UMAP excels at finding the underlying topological structure of the data, creating a lower-dimensional embedding that preserves significant relational patterns while reducing noise. Second, HDBSCAN was applied to this UMAP embedding. We selected HDBSCAN over alternatives like k-means or Gaussian Mixture Models due to its key advantages for behavioral data: it does not require pre-specifying the number of clusters, it can identify participants as noise (outliers),

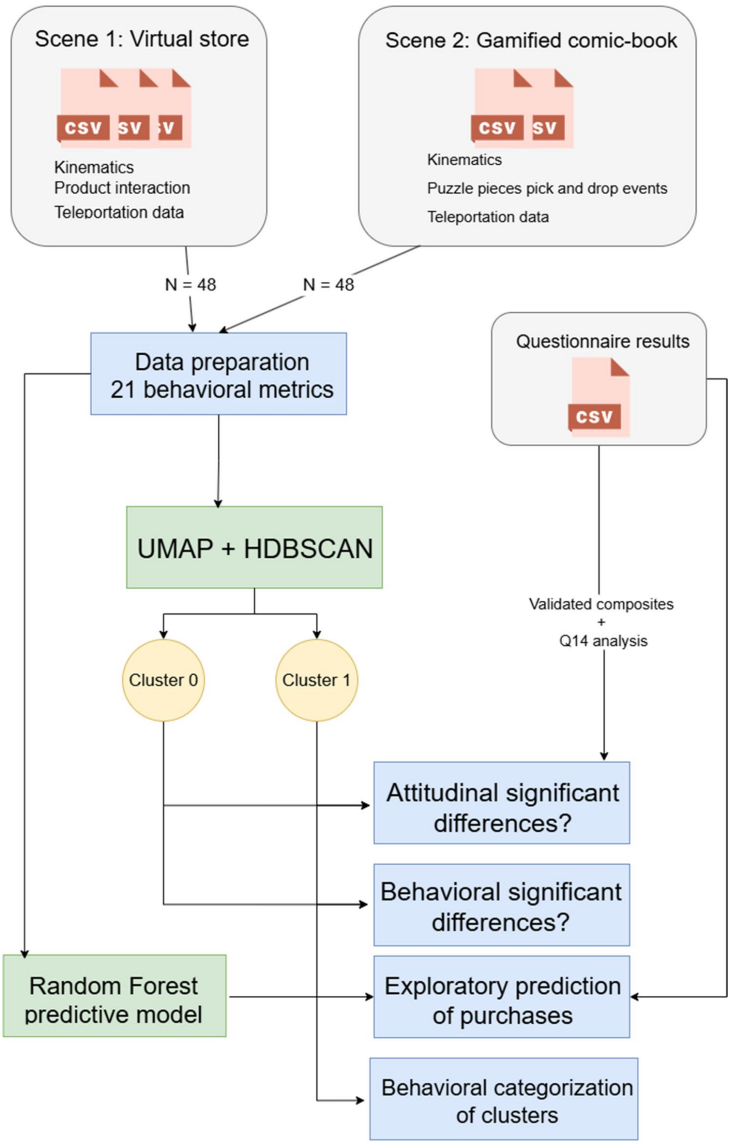


Fig. 11 Conceptual schema of the quantitative analysis presented in this subsection

and it can discover clusters of arbitrary shape and density, which is ideal for modeling complex, non-linear human behaviors.

To find the optimal parameters for clustering process, we performed a systematic grid search over a combined hyperparameter space for both UMAP (e.g., `n_neighbors`, `n_components`) and HDBSCAN (e.g., `min_cluster_size`, `min_samples`). More details can be found in the code of the public repository, The best-performing combination was selected by maximizing the Density-Based Clustering Validation (DBCV) index (Moulavi et al. 2014). The DBCV score ranges from -1 to 1, with values closer to 1

Table 7 Behavioural metrics employed in the quantitative data analysis. S1 refers to the first scene, while S2 points to the second scene

#	Metric (per user)	Definition	Psychological usefulness	Literature example
Spatial exploration				
1–2	*Entropy_S1 / S2*	Shannon entropy of 0.4m × 0.4m floor tiles where p_i = dwell-time fraction in tile i	Captures breadth of space sampled. Higher entropy ⇒ more curiosity / presence and sistematically spread visual attetion	Entropy reliably distinguished "thorough" vs. "shortcut" explorers in VR (Ward et al. 2025). Free-roam social VR scenes (Sánchez Pacheco et al. 2025a)
3–4	*FloorCoverage _S1 / S2*	Unique visited tiles ÷ total reachable tiles (scene-level denominator)	Simple % coverage that complements entropy. Easier to interpret for retail layout analysis. Higher % ⇒ more spatial curiosity	Visual-entropy studies often report both entropy and coverage to separate breadth from evenness (Sánchez Pacheco et al. 2025)
5– 6	*PathLenPer Min_S1 / S2*	$\sum \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \div \text{minutes played}$	Overall locomotor activity. Captures engagement for exploring	Path length per minute correlated with perceived exertion in locomotion studies (Kim et al. 2023)
7–8	*Tortuosity_S1 / S2*	Total path length ÷ straight-line displacement	High tortuosity = wandering; low = goal-directed. Aids to detect "aimless browsing" attitude	Tortuosity separates efficient vs. exploratory routes in VR wayfinding tasks (Yaremych et al. 2019)
Interaction intensity				
9–10	*GrabsPerMin _S1 / S2*	Completed object releases ÷ session minutes	Interaction rate. Higher rates implies higher engagement and may raise purchase intent	Object-grab frequency predicted task load in XR grabbing studies (Brzica et al. 2025; Shishir and Zhao 2025)
11–12	*MeanContact _S1 / S2*	Mean grab duration	Higher time suggest a higher cognitive processing, detailed evaluation and higher interest	Contact time used as inspection-time indicator in XR grasp feasibility research (Alabood et al. 2025)
13	*UniqueObjects _S1*	Count of distinct store products grabbed	Curiosity in the objects presented in the environment	Variety metrics employed in VR choice experiments (Meißner et al. 2020)
Visual attention				
14–15	*AvgFixDur_S1 / S2 (s)*	Mean fixation run length (angular speed $< 3°$ /s for ≥ 0.10s)	Longer fixations = more attetion; tied to cognitive load	Used threshold is standard in VR I-VT fixation detection (Uramune et al. 2024)
Locomotion activity				
16–17	*TeleportCount Per Min_S1 / S2*	Teleport activations ÷ minutes	High rates can lower presence but reduce sickness. Higher count ⇒ better usability of the interaction mechanism & spatial curiosity	Teleport frequency analysed in locomotion-technique trials (Franzluebbers and Johnsen 2023; Ribeiro et al. 2024)
18–19	*MeanTeleport Dist _S1 / S2*	Mean Euclidean step length per teleport	Differentiates small-hops from large-jumps. showing greater exploration confidence in long jumps	Mixed-method teleport studies log both count and distance to tune comfort vs. speed (Franzluebbers and Johnsen 2023)
Task performance (Scene 2)				
20	*ErrorCount_S2*	Number of mis-drops logged	Direct usability / performance metric and perceived dificulty	Error tallies used to benchmark grab-and-release spell interfaces (Alabood et al. 2025)
21	*Puzzle Completion Time*	Time from scene start to final piece drop	Overall task efficiency. Shorter times might map to higher VR skill	Completion time used as motor-learning performance metric (Özdel et al. 2025)

indicating denser, better-separated clusters. Note that scores above 0.5 are generally considered to indicate good clustering structure.

Prior to clustering, all 21 behavioural metrics were standardized using the z-score transformation. Our features existed on vastly different scales; for example, *PathLenPerMin* is measured in meters per minute, while a metric like *Entropy* is a unitless value. Without standardization, features with larger scales would disproportionately dominate the distance calculations that underpin the density estimation. The model would effectively overweight the importance of locomotion and be blind to subtle but

potentially crucial interaction patterns. By applying z-scoring, we transformed each feature to have a mean of 0 and a standard deviation of 1. This places all metrics on a common scale, ensuring that each behavioural feature contributes equally to the density estimation.

The raw results, as well as descriptive statistics for each cluster by metric, are available in the public data repository; here, we focus on synthesising and explaining the key findings. The optimal hyperparameter combination, reducing dimensionality from 21 to 11 components, produced a DBCV score of 0.653, considered moderately good based on the interpretation of the value range of this index. This clustering yielded two clusters: cluster 0 (C0) with 20 participants and cluster 1 (C1) with 28 participants, and no outliers. Using UMAP to reduce to two components, we visualise the clusters in Fig. 12. At first glance, the algorithm may have formed two clusters based on participant behaviour, such that C0 includes the 17 users who made at least one purchase, while C1 includes none, indicating a direct and strong relationship between in-experiment behaviour and final purchase.

To structure the cluster analysis and test this premise, we pose the following questions and outline our approach according to the conceptual map in Fig. 11:

1. Do the behaviourally-defined user groups exhibit statistically significant differences in their demographics?
2. What specific behavioural metrics most significantly differentiate the clusters, thereby defining their fundamental interaction styles?
3. Do the behaviourally-defined user groups exhibit statistically significant differences in their self-reported attitudinal profiles derived from questionnaire data?
4. What is the user behavioural profile in each cluster based in the data collected?

Before proceeding with the cluster analysis, Fig. 13 shows the distribution of purchasers across clusters, refuting the premise of a single "buyer" cluster. Our analysis immediately revealed a strong link between a user's behavioral profile and their likelihood to make a purchase. Specifically, participants belonging to Cluster 0 were more than twice as likely to buy a product, showing a purchase rate of 55.0% compared to just 21.4% for participants in Cluster 1. This significant association was confirmed statistically with a Chi-sqaure test $\chi^2(1, N = 48) = 4.37, p = 0.036$, with a medium effect size (Cramér's $V = 0.30$) and Fisher's exact test, $p = 0.031$, indicating that the behavioral patterns identified by our clustering are directly relevant to real-world commercial outcomes.

Specifically, participants in C0 demonstrated a much higher purchase behaviour, with 55.0% (11 out of 20) making a purchase. In contrast, only 21.4% of participants in C1 (6 out of 28) made a purchase. This might indicate that the behavioural groups identified by the clustering algorithm are strongly linked to real-world purchase outcomes.

Figure 14 illustrates the aggregated navigational patterns for C0 and C1 across the two experimental scenes. To facilitate a fair comparison of spatial preference independent of total session duration, each heatmap was normalized relative to its own maximum value (min-max scaling).

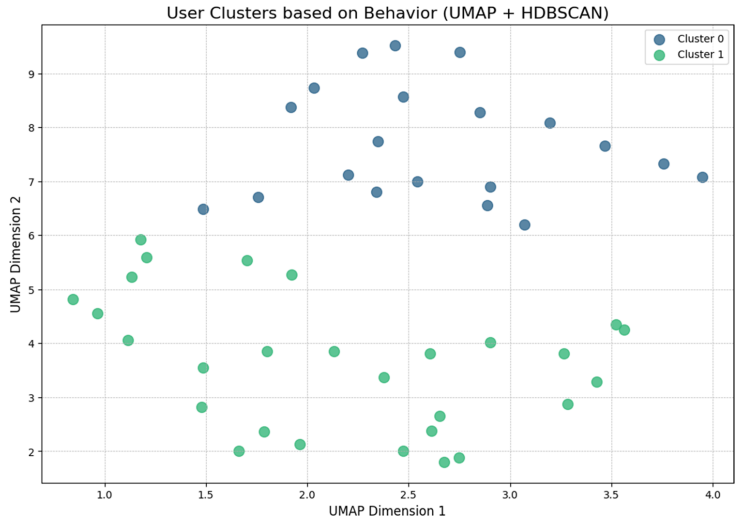


Fig. 12 A visual representation of the clusters formed. This visualization is possible thanks to reducing the dimensionality of the dataset to 2 dimensions

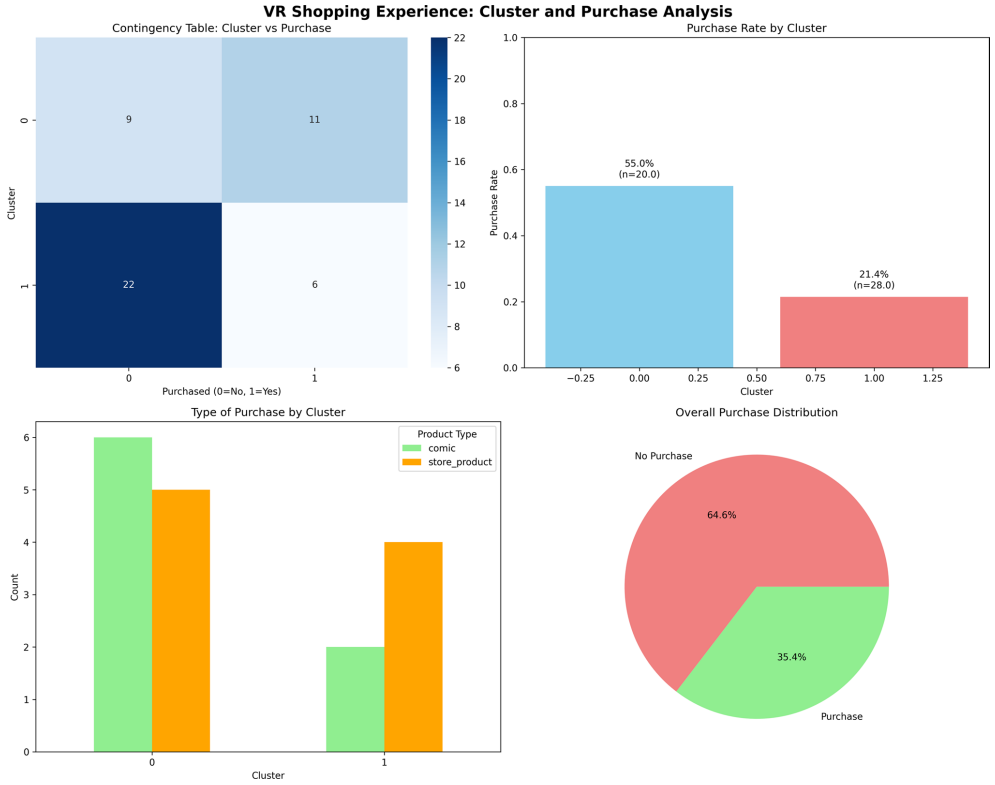


Fig. 13 Total number and distribution of buyers per cluster, along with what they actually purchased. The "Purchase Rate by Cluster" subplot shows sample proportions; Wilson 95% CIs are provided in Table 8

Table 8 Uncertainty summary for purchase rate by cluster

Purchase rate by cluster (Wilson 95% CI)				
Cluster	*n*	Purchases	Purchase rate	95% CI
0	20	11	55.0%	[34.2%, 74.2%]
1	28	6	21.4%	[10.2%, 39.5%]

In the scene 1 environment, both clusters demonstrated nearly indistinguishable behavioural patterns. As shown in the top image of Fig. 14, user activity for both groups was primarily concentrated along the central walkway and around the interactive displays. This suggests that the simple layout and clear points of interest prompted a standardized navigational approach.

In the gamified environment, the navigational patterns remain very similar, since the objective and puzzle pieces distribution were the same. However, C1 shows a broader spatial exploration as there are more activity points but rather cold than hot. This increasing activity might reveal that they spent more time completing the puzzle, since those users grouped in C0 present less spreadness but more hot areas in key tiles of the environment (where puzzle pieces are or in the table, where they deliver the pieces and stand for hearing the final discourse of the protagonist).

Demographic differences

After showing the statistical differences between clusters with regards to purchases, we will start to perform an external validation of the discovered clusters. First, we are ensuring they did not merely reflect pre-existing demographic groupings. Therefore, we performed a comparative analysis of participant age, gender, and prior VR experience. As shown in Fig. 15, no statistically significant differences were found across these variables.

The age distribution for Cluster 0 (M=25.8, SD=14.7) and Cluster 1 (M=22.2, SD=12.9) was compared. The violin plot (Fig. 15, top-left) visualizes these distributions, where the width of each "violin" indicates the density of

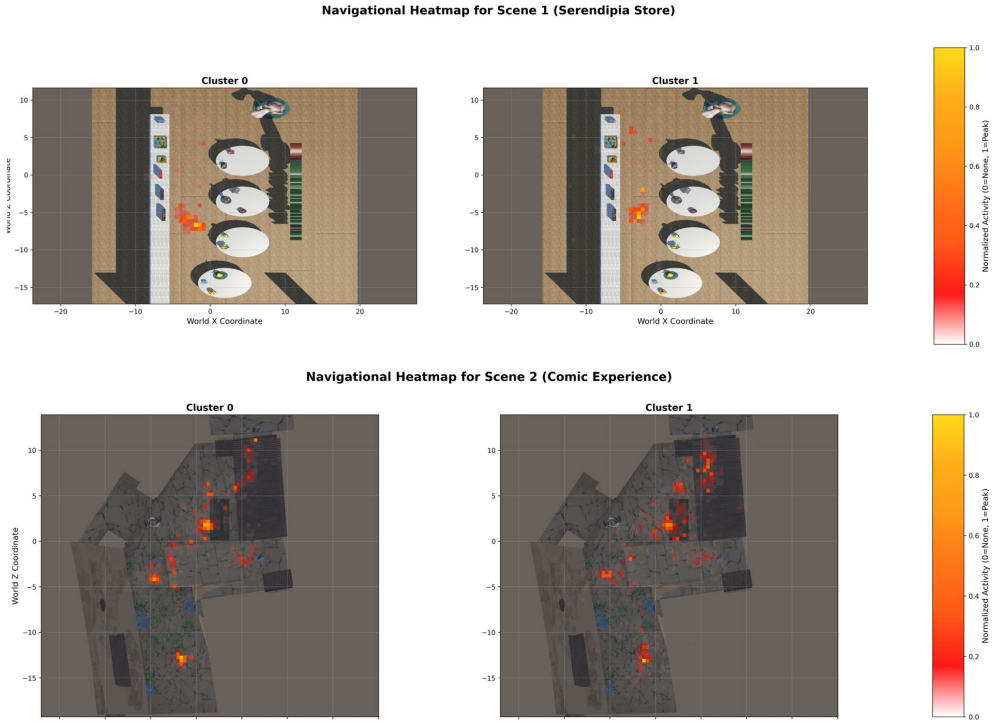


Fig. 14 Averaged heatmaps for both clusters in both scenes

participants at a given age. The solid red line within each violin marks the median age, providing a robust measure of central tendency. Despite slight visual differences in the distributions, a Mann–Whitney U test confirmed that there was no significant difference in age between the two clusters ($U = 238.0, p = 0.457$).

Similarly, the distribution of self-reported gender was analyzed. A Chi-square test of independence found no significant association between gender and cluster membership ($\chi^2(2) = 2.33, p = 0.311$, Cramér's V = 0.22). Finally, prior VR experience, rated on a 5-point Likert scale, was also found to be statistically indistinguishable between Cluster 0 (Median = 2.0) and Cluster 1 (Median = 2.0), as confirmed by a Mann–Whitney U test ($p = 0.369$).

Behavioural differences

Beyond simply predicting purchases, the two clusters revealed fundamentally different ways of interacting with the virtual world. After performing a comparative analysis (controlling for multiple comparisons), we found seven behavioral metrics that showed statistically significant differences, all with large effect sizes. These differences paint a clear picture of two distinct user behavioural profiles. All significant differences were characterized by large effect sizes (Rank-biserial $r > 0.58$ or Cohen's $d > 0.99$), indicating substantial practical importance.

Figure 16 shows the distributions of these seven key differentiating features. We used the same plot type (violin plot) as we used in the top left plot of Fig. 15.

The analysis of these key features reveals two contrasting interaction styles:

- Interaction Intensity: C0 demonstrated a significantly higher level of interaction, particularly within the virtual store. They grabbed objects far more frequently in both scenes, and interacted with a greater number of unique objects in the store. Even though for the gamified scene it could have revelaed a significant difference in number of errors, this is not the case, and the additional number of grabs may have been made for repositioning the graphic novel for further inspection in a more comfortable way. This points to a more hands-on, explorative interaction style for C0.

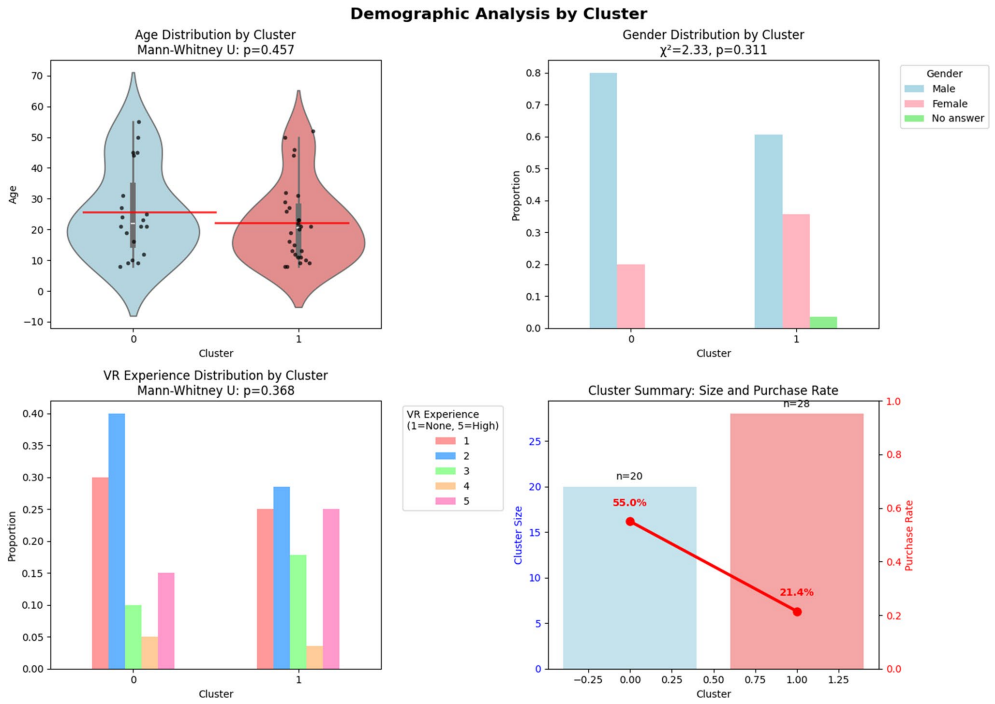


Fig. 15 Total number and distribution of buyers per cluster, along with what they actually purchased

- Locomotion and Task Performance: The clusters also exhibited different approaches to navigation and task completion. In the gamified scene, C0 covered more distance per minute and teleported more frequently, suggesting more active movement. Nonetheless, in the store, C1 utilized significantly longer, more deliberate teleports. This navigational attitude in C1 is mirrored in their task performance, as they took significantly longer to finish the puzzle in the second scene.

Additionally, we aimed to categorize into qualitative profiles the individuals inside each cluster based on the dimensions in which the behavioural features are classified. To ground our categorization in established theory, we mapped our data-driven dimensions to the holistic VR experience taxonomy proposed by Hameed et al. (2024). Our *Interaction Intensity* dimension serves as a direct, objective measure of Interactivity. Our *Spatial Exploration* and *Locomotion Activity* dimensions together provide a quantitative profile of *Explorability* and *Locomotion*. Furthermore, our inclusion of *Visual Attention* and *Task Performance* aligns with the framework of Croucher et al. (2024), who identify these as critical components of user behaviour in VR. It is important to note that our analysis, being focused on objective behavioural metrics, does not capture the purely subjective dimensions of the Hameed et al. taxonomy, such as Plausibility or Believability.

By calculating the average z-score for each cluster within these dimensions, we derived two distinct data-driven personas. The profiles of these personas are visualized in the radar charts in Fig. 17, which map the relative behavioural signatures of each cluster.

On the one hand, C0 individuals' behavioural signature is dominated by significantly higher levels of *Interaction Intensity* ($\Delta z = 0.26$) and *Locomotion Activity* ($\Delta z = 0.16$) compared to the other cluster. These users are actively moving through the environment and exhibit a hands-on approach, demonstrated by their significantly higher rates of object grabs. Interestingly, this high level of active engagement comes at the cost of task efficiency, as this profile scored lower on *Task Performance* ($\Delta z = -0.40$), taking their time to complete the puzzle.

On the other hand, the C1 profile is defined by efficiency and methodical evaluation. They scored significantly higher

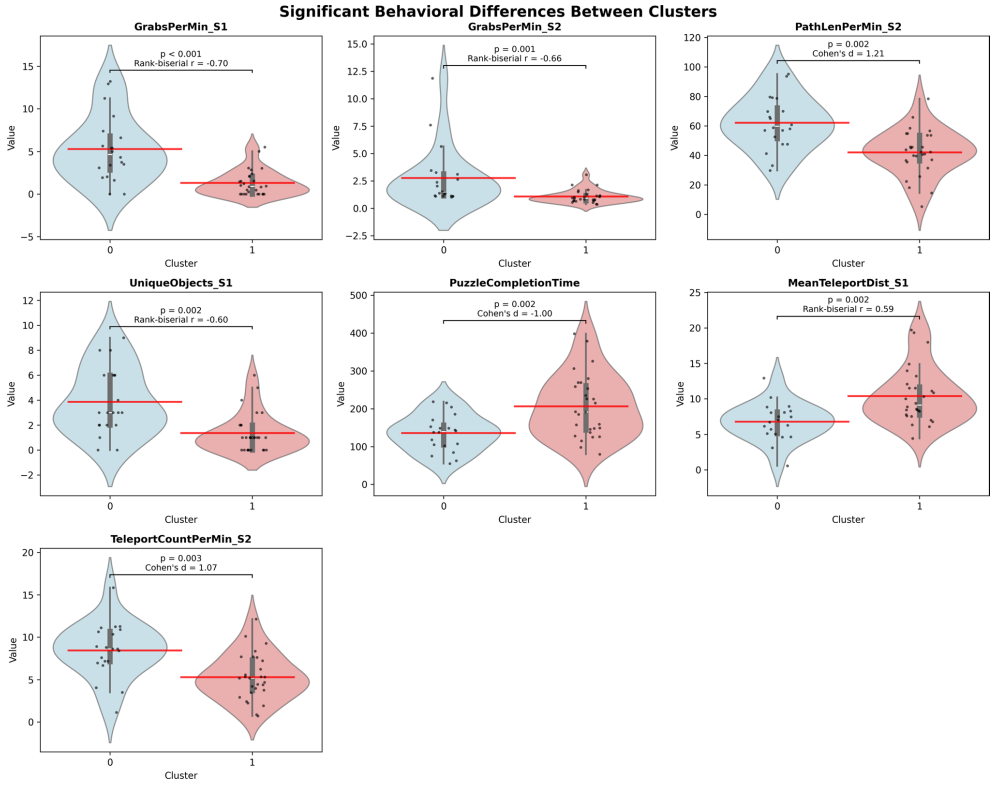


Fig. 16 Violin plots for each of the statistically significant different behavioural metrics between C0 and C1

on *Task Performance* ($\Delta z = 0.29$) and *Visual Attention* ($\Delta z = 0.12$), indicating a more goal-oriented and efficient approach to completing tasks. This is coupled with lower *Interaction Intensity* ($\Delta z = -0.19$) and Locomotion Activity ($\Delta z = -0.11$), defining a more stationary and observational user.

The primary trade-off between the two clusters lies in the inverse relationship between Interaction Intensity and Task Performance. While C0 users prioritize interaction over efficiency, which leads to a significantly higher conversion rate, C1 users prioritize efficiency over interaction, resulting in fewer purchases. This finding suggests that fostering high levels of direct, hands-on interaction with products appears to be a more effective driver of commercial outcomes than optimizing for streamlined, efficient task completion. The behavioural signature, not the task efficiency, is a stronger predictor of conversion.

Attitudinal differences

The next step in our external validation is to research whether the distinct behavioural individuals also corresponded to different self-reported attitudes through their answers to the questionnaire. We compared the two clusters across four validated attitudinal composites derived from the post-session questionnaire (see Sect. 4.1.2) and four single item questions that did not form into any composite (Q1, Q4, Q7 and Q14). Given that the data for these composite scales frequently violated normality assumptions (as confirmed by Shapiro-Wilk tests Shapiro and Wilk 1965), we employed non-parametric Mann–Whitney U tests for all comparisons.

The results of this analysis are summarized in Fig. 18. Despite interacting with the VR world in different ways, statistical analysis confirmed that there were no significant differences between Cluster 0 and Cluster 1 on any of our eight attitudinal measures, including key composites like VR Appeal ($p = 0.344$) and Usability & Immersion ($p = 0.785$). In our case study, it might suggest that

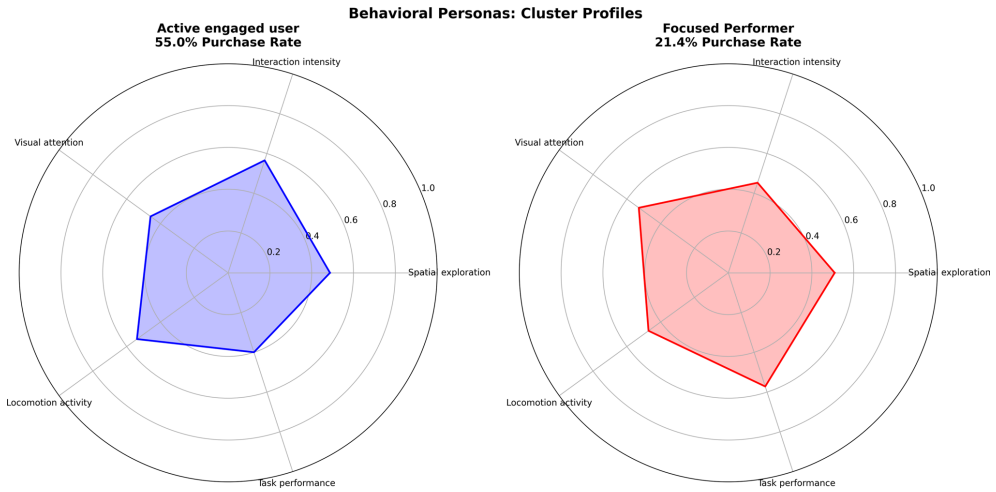


Fig. 17 Radar plot that maps a qualitative profile for each cluster member based on the dimensions that categorized our 21 behavioural features

observed behavior is a more potent predictor of commercial outcomes than self-reported attitudes.

With these results, we finish the analysis of the clusters formed. In the next section, we will change the focus of the analysis, basing it on a supervised learning problem.

4.3.2 Exploratory predictive analysis

While the previous sections identified and characterized distinct behavioural individuals, a critical next step is to assess the utility of our data in predicting purchases. To this end, we transitioned from unsupervised discovery to a supervised learning approach, with the goal of predicting a user's purchase decision. The main motivation for this analysis is **illustrative** rather than inferential. Given the limited statistical power afforded by our sample size, our goal is not to build a definitive, generalizable classifier. Instead, we conduct this exploratory predictive analysis to achieve two objectives: first, to demonstrate the feasibility of predicting real-money purchase decisions from in-VR data, and second, to identify the relative importance of different feature sets (i.e., objective behavioral metrics versus subjective attitudinal responses) to guide future, larger-scale research.

The objective is dual: first, to determine the feasibility of predicting purchase decisions using behaviourally collected data, and second, to identify the relative importance of different feature sets (i.e., objective behavioural metrics versus subjective questionnaire responses) in driving this prediction. By framing this as an exploratory study, we aim to provide a consistent methodological sketch and a reliable proof-of-concept that can inform and be replicated by future research with larger user samples.

To achieve this, we systematically developed and compared a series of Random Forest (RF) models. This choice was motivated by key advantages for a dataset with a small sample size. To specifically address the high risk of overfitting, we applied the following steps. First, we leveraged the inherent functioning of Random Forest, which averages the predictions of multiple decorrelated decision trees to reduce overall variance. Second, we conducted a comprehensive hyperparameter grid search to tune parameters that directly control model complexity, such as `max_depth`, `min_samples_split`, and `min_samples_leaf`. By optimizing these, we aim to prevent any single tree from becoming overly complex and fitting noise in the training data. Finally, the best model was selected based on its average performance across our cross-validation framework, ensuring a focus on generalization to unseen data.

Our analysis was conducted using a rigorous cross-validation strategy to ensure our performance estimates are as stable and unbiased as possible. Specifically, we employed a **Repeated Stratified K-Fold** procedure with **5 splits and 3 repeats**, resulting in 15 distinct model evaluations for each set of hyperparameters. This approach is particularly robust for the size of our dataset.

With regards to the data to be used, we will (i) evaluate a model built exclusively on behavioural data, then (ii) demonstrate the impact of feature selection, and (iii) finally assess a holistic model that integrates both behavioural and attitudinal data.

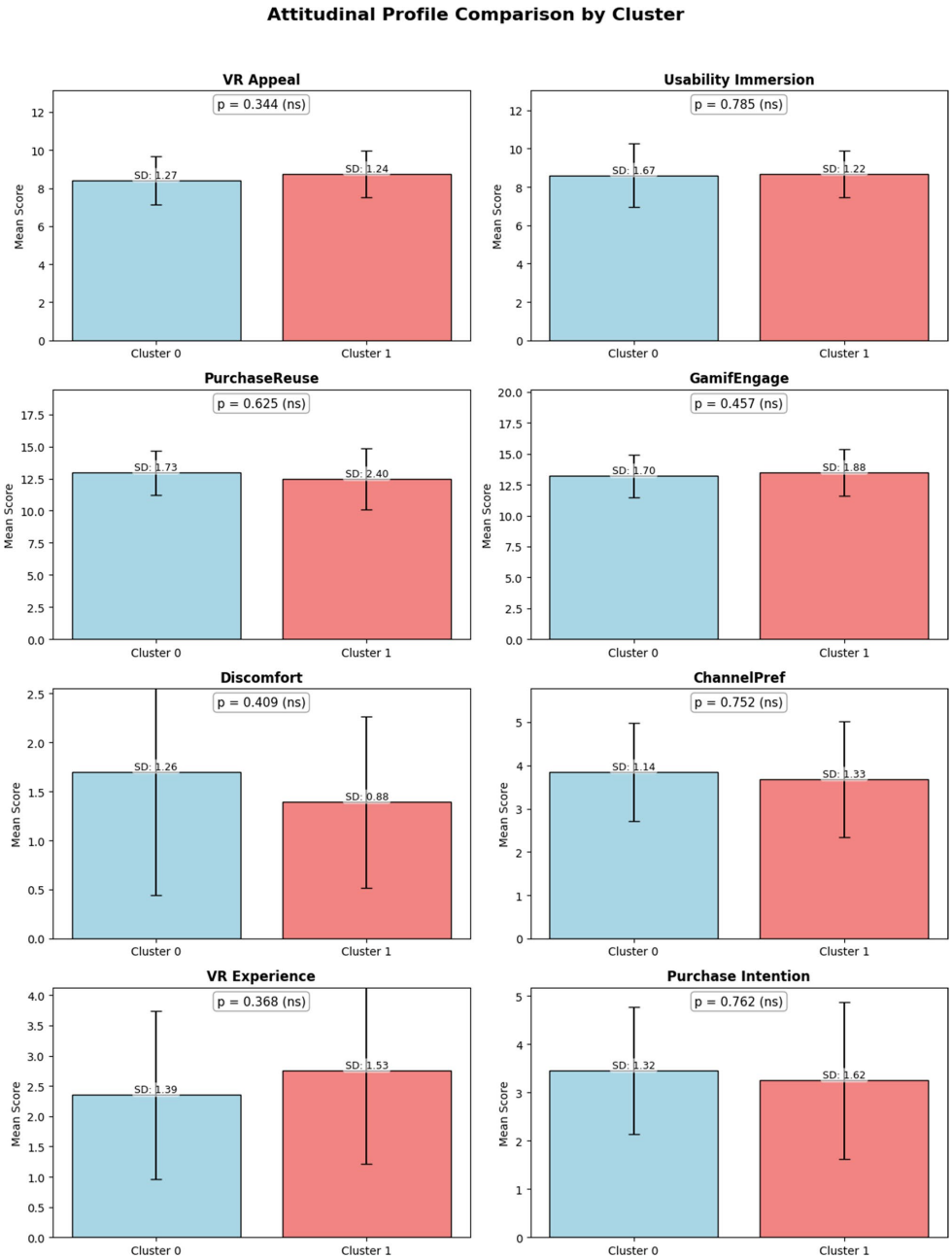


Fig. 18 Total number and distribution of buyers per cluster, along with what they actually purchased. Vertical whiskers on bars depict sample standard deviation; 95% CIs for the mean are provided in Table 9

Table 9 Uncertainty summary for attitudinal composites by cluster

Composite	Cluster	Mean ± SD	95% CI	*n*
(B) Attitudinal composites by cluster (Mean ± SD; 95% CI)				
VR Appeal	0	8.00 ± 1.27	[7.44, 8.56]	20
	1	8.90 ± 1.24	[8.44, 9.36]	28
Usability Immersion	0	8.60 ± 1.67	[7.87, 9.33]	20
	1	8.40 ± 1.22	[7.95, 8.85]	28
PurchaseReuse	0	13.20 ± 3.73	[11.57, 14.83]	20
	1	12.70 ± 2.40	[11.81, 13.59]	28
Gamification Engagement	0	13.10 ± 1.70	[12.35, 13.85]	20
	1	13.40 ± 1.88	[12.70, 14.10]	28
Discomfort	0	1.70 ± 1.26	[1.15, 2.25]	20
	1	1.40 ± 0.88	[1.07, 1.73]	28
Channel Preference	0	3.70 ± 1.14	[3.20, 4.20]	20
	1	3.80 ± 1.33	[3.31, 4.29]	28
VR Experience	0	2.40 ± 1.39	[1.79, 3.01]	20
	1	2.80 ± 1.53	[2.23, 3.37]	28
PurchaseIntention	0	3.40 ± 1.32	[2.82, 3.98]	20
	1	3.20 ± 1.62	[2.60, 3.80]	28

To rigorously evaluate the performance of our predictive models, we used the standard metrics, with a particular focus on those robust to the class imbalance present in our dataset (17 purchasers vs. 31 non-purchasers):

- AUC (Area Under the Receiver Operating Characteristic Curve). This is our main metric for overall model performance. It AUC represents the model's ability to correctly distinguish between a random positive (purchase) and a random negative (no purchase) case. An AUC of 1.0 signifies a perfect classifier, while an AUC of 0.5 represents performance equivalent to random chance. AUC was the score that we looked at when the hyperparameter tuning process was carried, selecting those values that provided the best AUC score.
- Precision: This metric answers the question: "Of all the users the model predicted would make a purchase, what percentage actually did?". High precision indicates a low false-positive rate, meaning the model's predictions of purchase behaviour are trustworthy.
- Recall. This metric answers the question: "Of all the users who actually made a purchase, what percentage did our model correctly identify?". High recall indicates a low false-negative rate, meaning the model is effective at finding the majority of the true purchasers.
- F1-Score. It is the harmonic mean of Precision and Recall, providing a single, balanced measure of a model's performance on the positive class. It is particularly useful in imbalanced datasets where there is a trade-off between identifying all positive cases and being correct when doing so.

Having presented the tools we used, we are presenting now the different models and the results obtained.

Full behavioural-only model

The purpose of this model was to determine the raw predictive power inherent in the behavioural data, prior to any feature selection or inclusion of attitudinal data. As mentioned previously, a RF classifier was trained on all 21 features, with the optimal hyperparameters based on AUC value.

The performance of this baseline model, evaluated through rigorous cross-validation, is summarized in Fig. 19. The model achieved a n average AUC of 0.655, indicating a predictive capability modestly better than random chance. However, other performance metrics were notably lower, with a mean F1-Score of 0.315 and Recall of 0.256. The high variance in these scores across the cross-validation folds suggests a degree of model instability.

The feature importance analysis, also shown in Fig. 19, reveals that no single metric is predominantly responsible for the model's predictions. Importance is distributed across numerous features, with *AvgFixDurS2* (Average Fixation Duration in Scene 2) emerging as the most influential, yet still only accounting for 7.6% of the model's decision-making weight. This distribution suggests that purchase behaviour is not predicted by a single action but rather by a complex, multivariate pattern of interactions.

Top 11 important features model

Building on the baseline model, we hypothesized that its performance was constrained by the inclusion of noisy, less-informative features, a common issue in datasets with a high feature-sample ratio. To address this, we employed a post-hoc feature selection strategy to create a more robust model.

This involved retraining a RF classifier using only the top k=11 most predictive behavioural features. The selection of k was decided to be aligned with the number of components that UMAP provided for the best DBCV score in the clustering process. This reduced the feature set by nearly 50%, focusing the model on the metrics with the strongest signal.

The results, shown in Fig. 20, demonstrate an improvement in predictive performance. The refined model achieved a mean AUC of 0.702, an increase of 0.046 over the full-feature model. The F1-Score also saw a notable improvement to 0.412. The most important features in this refined model, such as *AvgFixDurS1*, *GrabsPerMinS1*, and *TeleportCountPerMinS1*, reinforce the importance of visual attention, interaction, and locomotion patterns as key predictors. Of

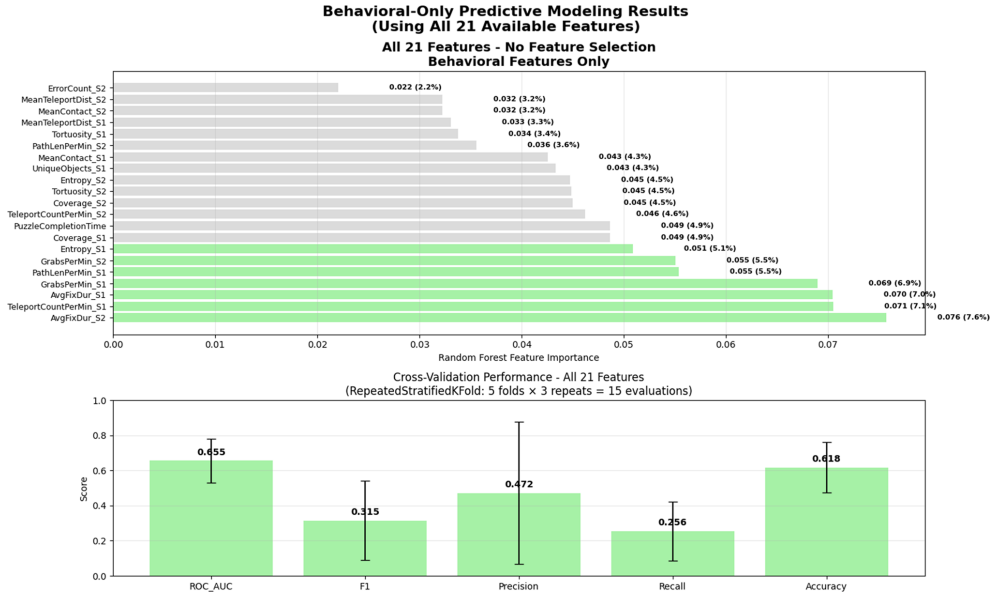


Fig. 19 Top subplot: bar chart with the importance given by the model to each of the features. Bottom subplot: bar chart with the average and Standard Deviation of the metrics

course. That's an excellent and critical point to address. It is important to note that the feature importance ranking in the refined 11-feature model (Fig. 20) differs from the ranking in the initial 21-feature model. This change is an expected outcome of the feature selection process. In a RF model, the calculated importance of any single feature is relative to the other features present. By removing the 10 least informative features, we reduced noise. This allows the model to more consistently leverage the stronger predictive patterns within the remaining data, resulting in a clearer and more stable hierarchy of feature importance.

Behavioural and attitudinal model

Having established that a refined set of behavioural metrics offers moderate predictive power, our final analysis sought to determine if integrating subjective attitudinal data could yield a superior holistic model. For this, we constructed a Random Forest classifier using all 29 available features: the 21 objective behavioural metrics, the 4 z-scored questionnaire composites and the 4 z-scored questionnaire single items.

The results demonstrate a dramatic and significant improvement in predictive performance. As shown in Fig. 21, this model achieved a mean cross-validated AUC of 0.849, a substantial increase over the behavioural-only models. This finding strongly supports the hypothesis that combining objective user actions with their subjective self-reported attitudes provides the most powerful framework for predicting purchase decisions.

An examination of the feature importance ranking provides critical insights into why this model performs so well. The two most influential predictors by a significant margin are the *PurchaseReuse* and *PurchaseIntention* composites from the questionnaire. In this model, it indicates that a user's stated intention is a potent predictor of their eventual action. However, the model does not rely solely on attitudes. Key behavioural metrics, including *TeleportCountPerMinS1*, *GrabsPerMinS2*, and *AvgFixDurS1*, remain highly ranked, confirming their value.

In conclusion, this exploratory analysis provides strong evidence that while behavioural data is a valuable predictor, a holistic approach that integrates both behavioural and attitudinal data is superior. For researchers and practitioners aiming to understand and predict user actions in VR, our findings suggest that capturing both implicit behaviours and explicit attitudes is essential for building robust and accurate predictive models.

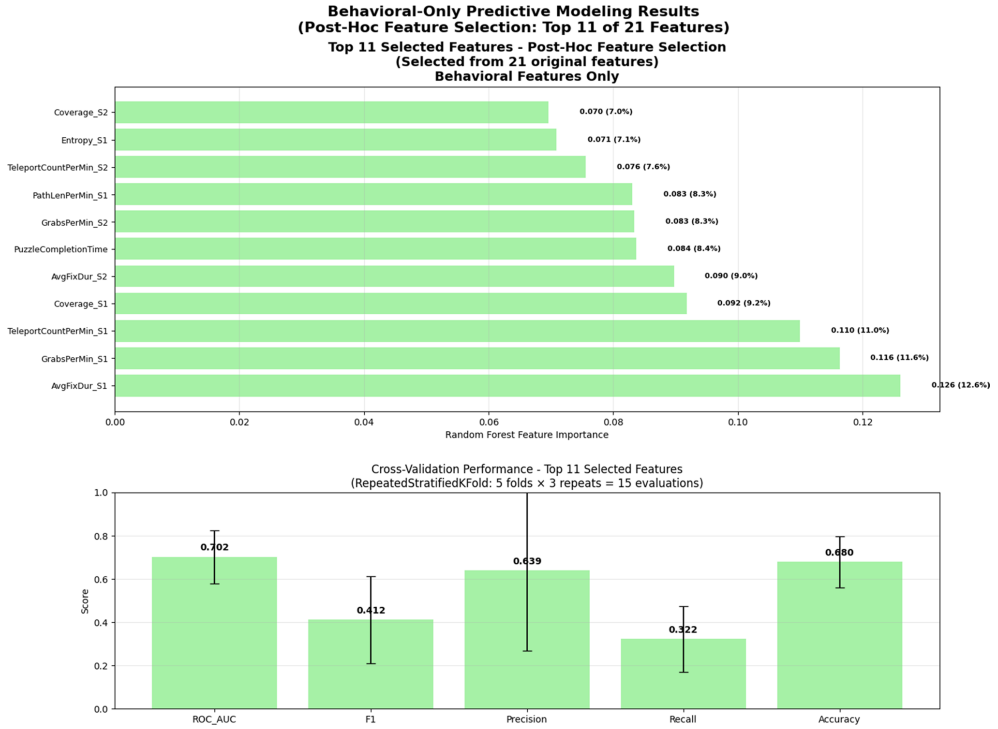


Fig. 20 Top subplot: bar chart with the importance given by the model to each of the features. Bottom subplot: bar chart with the average and Standard Deviation of the metrics

5 Discussion

In this work, we have addressed and discussed the potential of VR to provide immersive shopping experiences. Beyond conducting a survey-based analysis, we combined it with the recording and analysis of user behavioural data to classify participants and determine which behavioural characteristics may translate into higher purchase intention. Furthermore, we conducted a pilot/exploratory validation with a sample of 48 volunteers featuring diverse demographic characteristics (see Table 3) and, in general, very little prior VR headset experience, to assess the prospective impact of methods such as gamification on business metrics like conversion rate. Additionally, we exploratorily confirmed the potential of low-cost scanning technologies for representing real products in immersive spaces. Participants responded positively to these 3D products, with 9 of the 48 participants choosing to purchase one of the items they inspected in the virtual replica scene of *Serendipia*. Regarding the gamified graphic-novel scene, 8 of the 48 participants ended up purchasing a copy of *Calatrava: Pedro Girón*.

The empirical findings from the case study corroborate the proposed hypothesis, yielding a substantial overall conversion rate of 35.42%. This figure can be disaggregated into a conversion rate of 18.75% directly attributable to the sale of the €20 graphic novel and 16.67% stemming from virtual products presented within the virtual store. The intensity of consumer response was also significant; for instance, the sales volume from a single morning (Saturday) equaled that of the entire preceding day. This level of performance suggests that the strategy under examination offers a considerable advantage over more conventional marketing modalities in e-commerce.

Our analysis of the qualitative questionnaire (see Sect. 4.1 data revealed that the primary attitudinal drivers of purchase intention were not related to the novelty of VR itself. While

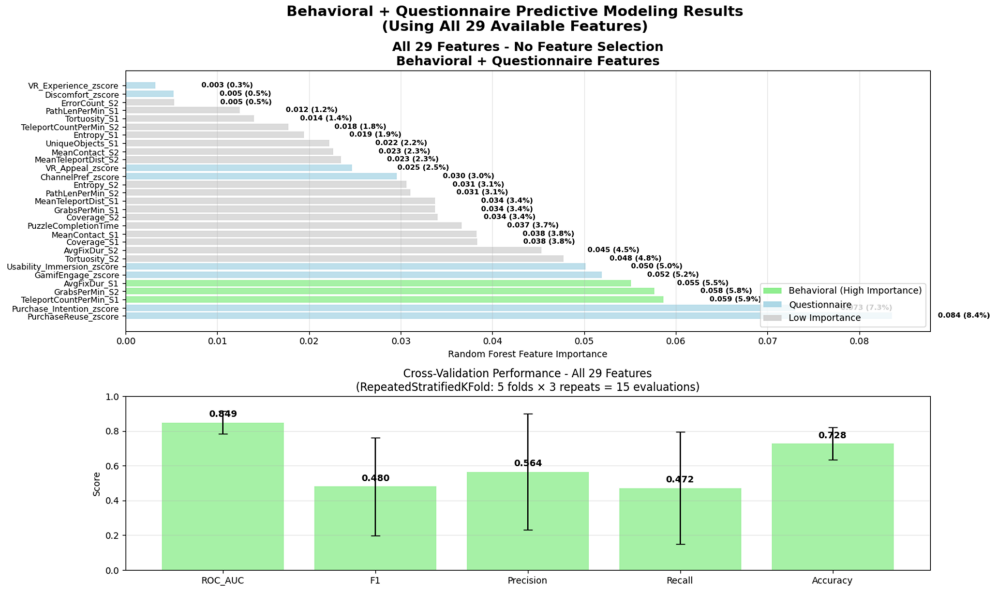


Fig. 21 Top subplot: bar chart with the importance given by the model to each of the features. Bottom subplot: bar chart with the average and Standard Deviation of the metrics

participants reported high general *VR Appeal*, this sentiment was consistent across both high and low intent purchasers and thus did not differentiate them. Instead, the key factors were *Gamification's Engagement* and *Usability/Immersion*. Users who reported higher motivation from the story-driven, gamified elements and perceived the interactions as easy, natural and immersive were significantly more likely to express an intention to buy. This initial finding suggests that for retail VR, the quality of the experience, specifically its engaging and usable design, is more critical for driving conversion than the novelty of the technology in the field.

Building on this, our quantitative analysis provided a deeper behavioural layer of understanding. Unsupervised clustering of 21 behavioural metrics (see Sect. 4.3.1) successfully identified two distinct user profiles: the "Active Engaged User" and the "Focused Performer", which were not defined by demographics but by their fundamental interaction styles. While the former was characterized by significantly higher *Interaction Intensity* and *Locomotion Activity*, engaging in more object grabs and covering more distance, the latter demonstrated higher *Task Performance* and *Visual Attention*, completing tasks more efficiently and showing a more observational and stationary approach. These behaviourally-defined groups showed a statistically significant difference in a key commercial outcome: the "Active Engaged User" persona had a much higher purchase rate (55.0%) than the "Focused Performer" (21.4%). This objective result directly complements our qualitative findings, suggesting that the high engagement reported by purchasers manifests as a specific, measurable pattern of active, hands-on behaviour.

5.1 Relation to prior work

Our behavioural segmentation shows that "Active Engaged Users" purchase substantially more often than "Focused Performers" (55.0% vs. 21.4%). This is consistent with prior evidence that interactivity and immersion/presence foster enjoyment, attention, and purchase propensity in VR shopping experiences (Peukert et al. 2019; Xi et al. 2024). In our data, concrete indicators of active engagement act as measurable proxies of hands-on exploration and visual attention, mechanisms linked with diagnosticity/usefulness and choice confidence during product inspection (Chen et al. 2022; Xi et al. 2024). At the same time, the near-equivalence of attitudinal composites across clusters alongside large behavioural differences suggests a slight

attitude–behaviour gap: self-reports remain uniformly positive, whereas revealed behaviour differentiates who actually converts. A possible reason is our real-money context (as opposed to virtual budgets typical of laboratory studies), which can accentuate discrepancies between stated intentions and actual purchases (Wölfel and Reinhardt 2019; Gil-López et al. 2023). Finally, our graphic-novel gamification likely supported motivational engagement and narrative transportation, encouraging exploration and inspection—the very behaviours that co-vary with purchase in our data(Rahmadhan et al. 2023; Wei et al. 2023).

Furthermore, our data-driven behavioral user profiles show a remarkable alignment with the foundational archetypes of hedonic and utilitarian shopping motivations in consumer psychology literature (Babin et al. 1994).

On the one hand, the "Active Engaged User" profile, characterized by high interaction intensity, broad exploration, and a significantly higher purchase rate, mirrors the behavior of a hedonic shopper. Hedonic consumers are primarily motivated by the entertainment, multisensory, and emotional aspects of the shopping experience itself (Alzayat and Lee 2021). Our findings suggest that a gamified, hands-on VR environment is highly effective at capturing this user type, converting their high engagement into tangible sales.

On the other hand, the "Focused Performer" profile, with its emphasis on task efficiency, aligns with the utilitarian shopper. Utilitarian consumers are typically rational, task-oriented, and seek to complete their shopping goal efficiently (Babin et al. 1994). The lower conversion rate for this group suggests that while they may successfully complete tasks within the VR environment, the experiential and gamified elements are less persuasive for them (Alzayat and Lee 2021). This aligns with research showing that utilitarian value is often linked to outcomes like convenience and efficiency, which were not the primary focus of our gamified design.

5.2 Case study limitations

Our sample size (N=48) necessarily limits statistical power and increases the risk of Type II errors, so inferences were made with caution. Participants were volunteers recruited on-site at a themed fair. Although they had no prior knowledge of the activity's content or goals, which minimizes expectation-driven bias, a minimal self-selection of interests typical of such events may still be present. For these reasons, the supervised models are positioned as exploratory, and future studies should replicate these analyses with larger and more diverse samples to test the robustness and generalizability of our findings. While our rigorous cross-validation provides confidence in our findings, the results should be considered a strong proof-of-concept rather than a definitive, generalizable model. Future work should aim to replicate this methodology with a larger and more diverse user population to validate these personas and refine the predictive model. Further research could also explore these behavioural patterns across different types of VR content and over longer periods to understand how user styles evolve with experience.

Finally, we aimed to construct an exploratory predictive model, detailing the methodological steps undertaken to combine both behavioural and attitudinal data. This final model was the most effective at predicting purchase decisions, achieving an average cross-validated AUC of 0.849 using a Repeated Stratified K-Fold strategy. The most influential predictors were the specific, action-oriented attitudes, followed closely by key behavioural metrics such as grabs per minute or number of teleport actions per minute. This highlights a strong synergy between both data sets, indicating that the most comprehensive understanding and the highest predictive accuracy are achieved by combining both data types. Furthermore, as discussed, the predictive models presented are intended to be **illustrative** of the potential of in-VR data, not inferential tools for predicting consumer behavior at large. Future work must aim to replicate this methodology with a larger and more diverse user population to validate these behavioral personas, confirm the observed effect sizes with greater statistical power, and refine the predictive models for general use.

6 Future work

The findings from this exploratory analysis have several implications for both designers and researchers of e-commerce VR experiences. For designers, it highlights that focusing solely on high-level satisfaction scores may mask underlying behavioural patterns that are more closely linked to commercial success. This, combined with the qualitative finding that gamified engagement was a key driver, suggests that fostering hands-on, playful interaction and exploration may be more critical for driving sales than optimizing for pure task efficiency.

Future work will focus on replicating this study with the aim of recruiting more participants to acquire a larger sample size to robustly validate our statements and methods.

Moreover, we are exploring Large Language Models powered solutions that can help to shape users from low-purchase intent to high-purchase intent based on their historical behaviour inside a VR shopping application. We aim to to detect early on any difficulties they may have in interacting or any lack of interest (using behavioural data collected) and try to offer natural assistance to improve both, so that they can be brought closer to the cluster behaviour that leads to better conversion rate.

Appendix A: EFA process detailed

The workflow was carried out as follows:

1. Every item is z-scored to that all questiones contributed equally regardless of their raw variance
2. We extracted factors using Principal-Components analysis (PCA) due to the moderate samle (N = 48). Following Kaiser criterion, we retained six components with eigenvalues ≥ 1.00, accounting for 78% of total variance, balancing parsimony and information retention.
3. The unrotated solution often mixes high loadings across factors. We therefore applied orthogonal varimax rotation, which redistributes variance so that each item loads strongly on a single factor and weakly elsewhere. A conventional salience threshold of $|\lambda| \geq 0.40$ is boldfaced in Table 10. Loadings below this mark are treated as negligible.
4. Items that clustered on one rotated axis were interpreted conceptually and grouped into a composite scale. The mapping is shown in Table 11

Appendix B: Technical implementation details

This appendix documents some technical details underpinning the two scenes used in the case study. These are provided to support replication and complement the open repository.

B.1 First scene

Although the original plan was to use the physical store's scan and photogrammetry techniques to construct the virtual environment replicating *Serendipia*, we encountered the following technical difficulties that led us to abandon this option:

- Lack of volumetric information. Despite using structure-from-motion photogrammetry tools, the quality of the resulting 3D reconstruction was insufficient for the case study's requirements.
- Lighting and reconstruction inconsistencies. User movement (whether walking or teleporting with the Meta Quest 3 headset) caused discrepancies in illumination and reconstruction based on the various 360° panoramic images, resulting in visible visual artifacts when rendering in Unity.
- Performance limitations. Running the standalone VR headset did not provide a satisfactory experience due to the high polygon count required to reconstruct *Serendipia*'s numerous products and shelves. The Meta Quest 3 struggled to maintain a stable frame rate (around 72 FPS). Moreover, imported photogrammetry meshes contained unnecessary polygons in areas such as floors

Table 10 Rotated factor loadings (varimax). Loadings $|\lambda| \geq 0.40$ in **bold**

Item	F1	F2	F3	F4	F5	F6
Q1	**0.73**	0.10	0.05	0.09	0.05	0.02
Q2	−0.12	−0.05	**0.68**	0.07	0.01	0.08
Q3	0.03	−0.04	**0.70**	0.22	0.02	0.05
Q4	**0.63**	0.11	−0.06	0.14	0.09	0.07
Q5	0.05	−0.07	0.09	**0.71**	0.11	−0.10
Q6	0.02	0.04	0.18	**0.78**	−0.03	0.01
Q7	0.06	0.02	0.03	0.06	**0.83**	0.05
Q8	−0.07	0.08	0.09	−0.09	0.06	**0.66**
Q9	−0.12	0.12	0.15	0.18	0.18	**0.60**
Q10	−0.11	0.03	0.11	0.22	0.04	**0.71**
Q11	**0.79**	-−0.04	0.05	0.06	0.03	0.09
Q12	**0.70**	0.15	0.08	0.07	0.02	0.01
Q13	**0.65**	0.19	0.02	0.05	0.04	0.10

Table 11 Rotated main loadings resulted from EFA

Main loading (rotated)	Composite label	Items
F1	Gamification's engagement	Q11-Q13
F2	Channel preference	Q4
F3	Usability/Immersion	Q6-Q5
F4	VR Appeal	Q3-Q2
F5	Discomfort	Q7
F6	Purchase intention & reuse	Q8-Q10
F7	VR Experience	Q1

and walls, which could be represented by simple planes, further degrading performance.The main digitization effort was devoted to *Serendipia*'s products. For this, we used the cameras of an iPhone 15 Pro. We employed Meshroom,[5] which uses photogrammetry to reconstruct 3D models from photographs of simple products such as board games, while for complex, highly detailed products like collectible figures, we used Luma AI.[6] The latter is free software that uses Neural Radiance Fields (NeRF) to provide high-quality 3D models, guiding the user through a video-based capture process from multiple viewpoints.

For interaction with the products and the environment, we used the Meta SDK. This allowed, through hand tracking, interaction with the user's bare hands, dispensing with controllers. Examples can be seen in Fig. 5, which show teleportation for moving through the environment and switching scenes by interacting with the door at the back, as well as product manipulation. In this scene, in addition to recording the user's spatiotemporal position over time, we also recorded interaction-level metrics - such as grab events and their durations and which specific objects were grabbed - thanks to the VRSI framework (Grande et al. 2024).

B.2 Scene

For the development of this scene, we first modeled a part of Calatrava Castle in 3D using Blender. Figure 6.A shows different viewpoints of the modeled environment, where we placed the graphic novel's protagonist (also modeled) as well as the puzzle pieces forming the novel's cover. Once the interior space was modeled, we added the floor and stairs as navigable areas so that, upon export to Unity, they could be marked as navigable with the Meta SDK, allowing the user to teleport through the scenario.

The 3D modeling effort also encompassed the creation and animation of the graphic novel's protagonist, Pedro Girón. This character is included in the experience as a Non-Playable Character (NPC), introducing the user to the story at the scene's start and continuing the narration each time a puzzle piece is collected. In Fig. 6.B, the protagonist's 3D model in Blender is shown; it was manually rigged (without stock animations) for more realistic movement by creating its 3D skeleton within the software. The anthropomorphic skeleton features joints to simulate articulation and nodes to simulate mouth movements during speech.

Regarding the efforts in developing this second scene (see Fig. 7), we maintained use of the Meta SDK for Unity, enabling users to use their hands to interact with objects and the environment. As in the first scene, the entire environment is navigable via teleportation, since the experiment was conducted in a physical space not large enough for room-scale navigation. Having described Scenes 1 and 2, Table 1 shows the data items collected and to be processed later in Sect. 4.

Acknowledgements This work has been funded by the Spanish Ministry of Science, Innovation and Universities MICIU/AEI/10.13039/501100000033, and the European Union (NextGenerationEU/PRTR), under the Research Project: Design and development of a platform based on VR-Shopping and AI for the digitisation and strengthening of local businesses and economies, TED2021-131082B-I00.
In addition, the first author has a pre-doctoral FPI (Research Staff Training) contract at the University of Castilla-La Mancha, Spain, for which funding has been received from the European Union through the European Social Fund Plus (FSE+).

Author contributions R.G wrote the main manuscript supported by J.A, S.S, J.A and D.V reviewed it and suggested imrpovements in the manuscript, while J.J C.S helped with the preparation and conduction of the experimental session

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data availability The anonymized data generated during the experimentation sessions of this work is available through the following public repository: https://github.com/AIR-Research-Group-UCLM/Manchacomic-case-study

Declarations

Conflict of interest The authors declare no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

[5] https://alicevision.org/

[6] https://lumalabs.ai/interactive-scenes

References

Azmi LF, Ahmad N, Iahad NA (2021) Gamification elements in e-commerce—a review. In: 2021 international congress of advanced technology and engineering (ICOTEN), pp 1–5. https://doi.org/10.1109/ICOTEN52080.2021.9493475

Azmi LF, Ahmad N, Iahad NA (2021) Gamification elements in e-commerce—a review. In: 2021 international congress of advanced technology and engineering (ICOTEN), pp 1–5. https://doi.org/10.1109/ICOTEN52080.2021.9493475

Aparicio M, Costa CJ, Moises R (2021) Gamification and reputation: key determinants of e-commerce usage and repurchase intention. Heliyon 7(3):06383. https://doi.org/10.1016/j.heliyon.2021.e06383

Arcos MG, Grande R, Martínez-Cid S, Morcillo CG, Jesús Castro Sánchez J, Albusac J (2024) User activity monitoring and personalized recommendations for enhancing the vr shopping experience. In: 2024 IEEE international conference on artificial intelligence and extended and virtual reality (AIxVR), pp 176–184. https://doi.org/10.1109/AIxVR59861.2024.00030

Alzayat A, Lee SHM (2021) Virtual products as an extension of my body: exploring hedonic and utilitarian shopping value in a virtual reality retail environment. J Bus Res 130:348–363. https://doi.org/10.1016/j.jbusres.2021.03.017

Alabood L, Nazari A, Dow T, Alabood S, Jaswal VK, Krishnamurthy D (2025) Grab-and-release spelling in xr: a feasibility study for nonspeaking autistic people using video-passthrough devices. In: Proceedings of the 2025 ACM designing interactive systems conference. DIS '25, pp 81–102. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3715336.3735719

Afwa IL, Rubiyanti N, Widodo A, Silvianita A (2024) The influence of virtual reality technology on purchase intention through user experience as an intervening variable in the marketing strategy of samsung products in the virtual world: a conceptual paper. East Asian J Multidiscip Res 3(6):1997–2008. https://doi.org/10.55927/eajmr.v3i6.9350

Babin BJ, Darden WR, Griffin M (1994) Work and/or fun: measuring hedonic and utilitarian shopping value. J Consum Res 20(4):644–656. https://doi.org/10.1086/209376 (https://academic.oup.com/jcr/article-pdf/20/4/644/5090984/20-4-644.pdf)

Brzica L, Matanović F, Vlahović S, Bernardić NP, Skorin-Kapov L (2025) Analysis of user experience and task performance in a multi-user cross-reality virtual object manipulation task. MMVE '25, pp 22–28. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3712677.3720461

Chen J, Ha QA, Vu M (2022) The influences of virtual reality shopping characteristics on consumers' impulse buying behavior. Int J Hum-Comput Interact 39(17):3473–3491. https://doi.org/10.1080/10447318.2022.2098566

Croucher C, Powell W, Stevens B, Miller-Dicks M, Powell V, Wiltshire TJ, Spronck P (2024) Locomote—a framework for classification of natural locomotion in VR by task, technique and modality. IEEE Trans Vis Comput Graphics 30(8):5765–5781. https://doi.org/10.1109/TVCG.2023.3313439

Castro-Schez JJ, Grande R, Herrera V, Schez-Sobrino S, Vallejo D, Albusac J (2024) E-marketplace state of the art and trends: VR-ZOCO–an architectural proposal for the future. Appl Syst Innov 7(5):76. https://doi.org/10.3390/asi7050076

Donatiello L, Morotti E, Marfia G, Di Vaio S (2018) Exploiting immersive virtual reality for fashion gamification. In: 2018 IEEE 29th annual international symposium on personal, indoor and mobile radio communications (PIMRC), pp 17–21. https://doi.org/10.1109/PIMRC.2018.8581036

Franzluebbers A, Johnsen K (2023) Versatile mixed-method locomotion under free-hand and controller-based virtual reality interfaces. In: Proceedings of the 29th ACM symposium on virtual reality software and technology. VRST '23. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3611659.3615701

Faganel A, Pačarić F, Rižnar I (2024) The impact of gamification on Slovenian consumers' online shopping. Adm Sci 14(5):86. https://doi.org/10.3390/admsci14050086

Grande R, Albusac JA, Vallejo D, González-Morcillo C, Sánchez-Sobrino S, Castro-Schez JJ (2024) Virtual reality shopping-insights: a data-driven framework to assist the design and development of virtual reality shopping environments. SoftwareX 27:101874. https://doi.org/10.1016/j.softx.2024.101874

Grande R, Albusac J, Vallejo D, Glez-Morcillo C, Castro-Schez JJ (2024) Performance evaluation and optimization of 3D models from low-cost 3d scanning technologies for virtual reality and metaverse e-commerce. Appl Sci 14(14):6037. https://doi.org/10.3390/app14146037

Garrido M, Jesús Castro-Schez J, Grande R, Schez-Sobrino S, Vallejo D (2024) Enhancement of the online presence of small and medium sized enterprises with minimum impact on traditional business activities in towns and cities. In: Proceedings of the 26th international conference on enterprise information systems - Volume 1: ICEIS, pp 989–996. SciTePress. https://doi.org/10.5220/0012630600003690. INSTICC

Gil-López C, Guixeres J, Moghaddasi M, Khatri J, Marín-Morales J, Alcañiz Raya M (2023) Recognizing shopper demographics from behavioral responses in a virtual reality store. Virtual Real 27(3):1937–1966. https://doi.org/10.1007/s10055-023-00767-2

Hameed A, Möller S, Perkis A (2024) A holistic quality taxonomy for virtual reality experiences. Front Virtual Real 5:1434016. https://doi.org/10.3389/frvir.2024.1434016

Jacoby J (2002) Stimulus-organism-response reconsidered: an evolutionary step in modeling (consumer) behavior. J Consum Psychol 12(1):51–57

Kim RY (2022) Retail after COVID-19: use virtual reality to enhance ecommerce, pp 118–123. https://doi.org/10.1109/TEMSCONEUROPE54743.2022.9801972

Kim A, Lee J-E, Lee K-M (2023) Exploring the relative effects of body position and locomotion method on presence and cybersickness when navigating a virtual environment. ACM Trans Appl Percept 21(1):1–25. https://doi.org/10.1145/3627706

Khatri J, Marín-Morales J, Moghaddasi M, Guixeres J, Chicchi Giglioli I, Alcañiz Raya M (2022) Recognizing personality traits using consumer behavior patterns in a virtual retail store. Front Psychol 13:752073. https://doi.org/10.3389/fpsyg.2022.752073

Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and anovas. Front Psychol 4:863. https://doi.org/10.3389/fpsyg.2013.00863

Mütterlein J, Hess T (2017) Exploring the impacts of virtual reality on business models: the case of the media industry, pp 3213–3222

Moulavi D, Jaskowiak PA, Campello RJ, Zimek A, Sander J (2014) Density-based clustering validation. In: Proceedings of the 2014 SIAM international conference on data mining, pp 839–847. SIAM

Mkedder N, Jain V, Salunke P (2024) Determinants of virtual reality stores influencing purchase intention: An interpretive structural modeling approach. J Retail Consum Serv 78:103757. https://doi.org/10.1016/j.jretconser.2024.103757

Moghaddasi M, Marín-Morales J, Khatri J, Guixeres J, Chicchi Giglioli I, Alcañiz Raya M (2021) Recognition of customers' impulsivity from behavioral patterns in virtual reality. Appl Sci 11:4399. https://doi.org/10.3390/app11104399

Meißner M, Pfeiffer J, Peukert C, Dietrich H, Pfeiffer T (2020) How virtual reality affects consumer choice. J Bus Res 117:219–231. https://doi.org/10.1016/j.jbusres.2020.06.004

Özdel S, Buldu KB, Kasneci E, Bozkir E (2025) Exploring context-aware and LLM-driven locomotion for immersive virtual reality. arXiv:2504.17331

Peukert C, Pfeiffer J, Meißner M, Pfeiffer T, Weinhardt C (2019) Shopping in virtual reality stores: the influence of immersion on system adoption. J Manag Inf Syst 36(3):755–788. https://doi.org/10.1080/07421222.2019.1628889

Ribeiro RA, Gonçalves I, Piçarra M, Seixas Pereira L, Duarte C, Rodrigues A, Guerreiro Ja (2024) Investigating virtual reality locomotion techniques with blind people. In: Proceedings of the 2024 CHI conference on human factors in computing systems. CHI '24. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3613904.3642088

Raharjanti N, Wiguna T, Purwadianto A, Soemantri D, Indriatmi W, Poerwandari E, Mahajudin M, Nugrahadi N, Roekman A, Saroso O, Ramadianto A, Levania M (2022) Translation, validity and reliability of decision style scale in forensic psychiatric setting in Indonesia. Heliyon 8:5. https://doi.org/10.1016/j.heliyon.2022.e09810

Rahmadhan P, Wana MA, Sensuse DI, Suryono RR et al (2023) Trends and applications of gamification in e-commerce: a systematic literature review. J Inf Syst Eng Bus Intell 9(1):90

Speicher M, Cucerca S, Krüger A (2017) Vrshop: a mobile interactive virtual reality shopping environment combining the benefits of on- and offline shopping. Proc ACM Interact Mob Wearable Ubiquitous Technol 1(3):1–31. https://doi.org/10.1145/3130967

Sánchez Pacheco T, Nolte D, König SU, Pipa G, König P (2025) Beyond the first glance: How human presence enhances visual entropy and promotes spatial learning. bioRxiv. https://doi.org/10.1101/2025.05.29.656764. https://www.biorxiv.org/content/early/2025/05/30/2025.05.29.656764.full.pdf

Sánchez Pacheco T, Nolte D, König SU, Pipa G, König P (2025) Beyond the first glance: How human presence enhances visual entropy and promotes spatial learning. bioRxiv. https://doi.org/10.1101/2025.05.29.656764. https://www.biorxiv.org/content/early/2025/05/30/2025.05.29.656764.full.pdf

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples)†. Biometrika 52(3–4):591–611. https://doi.org/10.1093/biomet/52.3-4.591 (https://academic.oup.com/biomet/article-pdf/52/3-4/591/962907/52-3-4-591.pdf)

Shishir NS, Zhao R (2025) Is the jedi force pull method effective? evaluating eye-tracked object grabbing in vr. FDG '25. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3723498.3723695

Uramune R, Ikeda S, Ishizuka H, Oshiro O (2024) Fixation-based Self-calibration for eye tracking in VR headsets. arXiv:2311.00391

van Herpen E, van den Broek E, van Trijp HCM, Yu T (2016) Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. Appetite 107:196–207. https://doi.org/10.1016/j.appet.2016.07.033

Watkins MW (2018) Exploratory factor analysis: a guide to best practice. J Black Psychol 44(3):219–246. https://doi.org/10.1177/0095798418771807

Ward EM, Carlson JM, Chrastil ER (2025) Divide (evenly) and conquer (quickly): spatial exploration behaviors predict navigational learning and differ by sex. Cognition 261:106144. https://doi.org/10.1016/j.cognition.2025.106144

Werbach K, Hunter D (2012) For the win: how game thinking can revolutionize your business. Wharton Digital Press, Philadelphia, p 144

Wölfel M, Reinhardt A (2019) Immersive shopping—presentation of goods in virtual reality, pp 119–130

Wu C-W (2023) The impact of virtual reality and augmented reality service technologies on consumer purchase intention for fashion brands. Preprints.org. https://doi.org/10.20944/preprints202307.1408.v1

Wei Z, Zhang J, Huang X, Qiu H (2023) Can gamification improve the virtual reality tourism experience? Analyzing the mediating role of tourism fatigue. Tour Manag 96:104715. https://doi.org/10.1016/j.tourman.2022.104715

Xi N, Chen J, Gama F, Korkeila H, Hamari J (2024) Acceptance of the metaverse: a laboratory experiment on augmented and virtual reality shopping. Internet Res 34:82–117. https://doi.org/10.1108/INTR-05-2022-0334

Xi N, Hamari J (2021) Shopping in virtual reality: a literature review and future agenda. J Bus Res 134:37–58. https://doi.org/10.1016/j.jbusres.2021.04.075

Xue L, Parker C, Hart C (2020) How to engage fashion retail with virtual reality: a consumer perspective, pp 23–35. https://doi.org/10.1007/978-3-030-37869-1_3

Yaremych HE, Kistler WD, Trivedi N, Persky S (2019) Path tortuosity in virtual reality: a novel approach for quantifying behavioral process in a food choice context. Cyberpsychol Behav Soc Netw 22(7):486–493. https://doi.org/10.1089/cyber.2018.0644. (**PMID: 31241349**)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

5.2 | Otras publicaciones

En esta sección se incluyen los artículos publicados en revistas científicas indexadas, que no forman parte del compendio pero que se encuentran estrechamente relacionados con la tesis y el proyecto en el que se enmarcó. Además, se incluyen publicaciones realizadas en congresos nacionales e internacionales, así como un capítulo de libro. Quedan fuera del ámbito de esta sección aquellas publicaciones en revistas indexadas que no están relacionadas con el tema de la tesis.

5.2.1 | Artículos publicados en revistas científicas indexadas

Además de los artículos reseñados en la sección anterior, el trabajo realizado durante la investigación llevada a cabo en esta tesis doctoral, ha producido un artículo en una revista indexada JCR.

E-Marketplace State of the Art and Trends: VR-ZOCO—An Architectural Proposal for the Future

- **Autores:** J.J. Castro-Schez, R. Grande, V. Herrera, S. Schez-Sobrino, D. Vallejo, J. Albusac.
- **Tipo:** artículo de revista.
- **Revista:** Applied System Innovation.
- **Editorial:** MDPI.
- **ISSN:** 2571-5577.
- **Volumen/Número:** 7(5), 76.
- **Año:** 2024.
- **DOI:** `https://doi.org/10.3390/asi7050076`
- **Categoría:** Engineering, Electrical & Electronic.
- **Factor de Impacto (2024):** 3.7.
- **Factor de Impacto a 5 años (2024):** 4.0.

- **Ranking JCR:** Q2.
- **Figuras relacionadas:** 5.2, 5.3.
- **Resumen:** E-commerce has become uniquely relevant to small- and medium-sized enterprises (SMEs) as an essential catalyst for their growth and sustainability. SMEs see e-commerce portals as a strategic way to engage in digital business activities without having to implement costly proprietary e-commerce solutions. However, e-commerce portals or e-marketplaces are not without their challenges. Today, they face increasing pressure to reduce their environmental impact and to empower local commercial businesses, as well as local businesses in the entertainment and culture industry. To address these challenges, there is a pressing need to propose new types of e-marketplaces that support the concept of the 15-minute city and in which virtual and augmented reality play a key role. These marketplaces would not only boost environmental sustainability but also strengthen the connection between local businesses and the community, creating a stronger and more collaborative network that benefits both businesses and consumers.
- **Relación con la tesis:** este artículo constituye la base conceptual y arquitectónica de VR-ZOCO, el *e-marketplace* inmersivo propuesto en esta tesis. En concreto, el trabajo aborda el análisis y sistematización de las limitaciones funcionales de los *e-marketplaces* actuales (SO1.1) y presenta el diseño de la arquitectura multicapa del sistema, incluyendo las capas de interacción del usuario, razonamiento basado en IA y persistencia (SO1.2). Este artículo se relaciona directamente con el objetivo O1 de la tesis. Aunque no se incluye en el compendio, sus contribuciones se citan y discuten a lo largo de los capítulos de introducción y estado del arte, y proporcionan el marco donde se insertan las restantes contribuciones del compendio.

5.2.2 | Congresos internacionales

Además de los artículos reseñados en la sección anterior, el trabajo realizado durante la investigación llevada a cabo en esta tesis doctoral, ha dado como fruto un conjunto de publicaciones y ponencias en congresos internacionales.

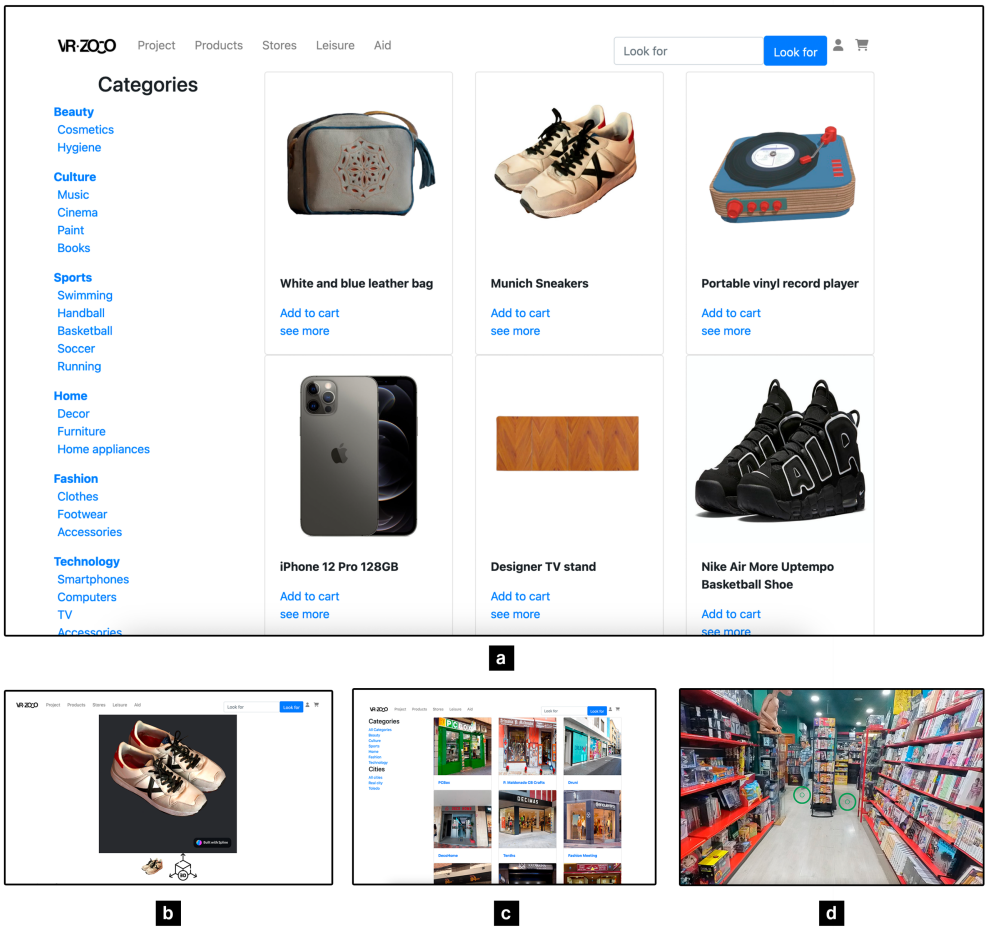


Figura 5.2: (a) Vista de la versión web del portal VR-ZOCO, donde los productos se categorizan y muestran al usuario. (b) Posibilidad de interactuar con cada producto en vista 3D; los productos fueron previamente escaneados y digitalizados. (c) Visualización de los comercios de la ciudad localizados en torno a los puntos de recogida de productos.

Supporting Small Businesses and Local Economies Through Virtual Reality Shopping and Artificial Intelligence: A Position Paper

- **Autores:** R. Grande, S. Sánchez-Sobrino, D. Vallejo, J.J. Castro-Schez, J. Albusac.

- **Tipo:** artículo de congreso (*position paper*).

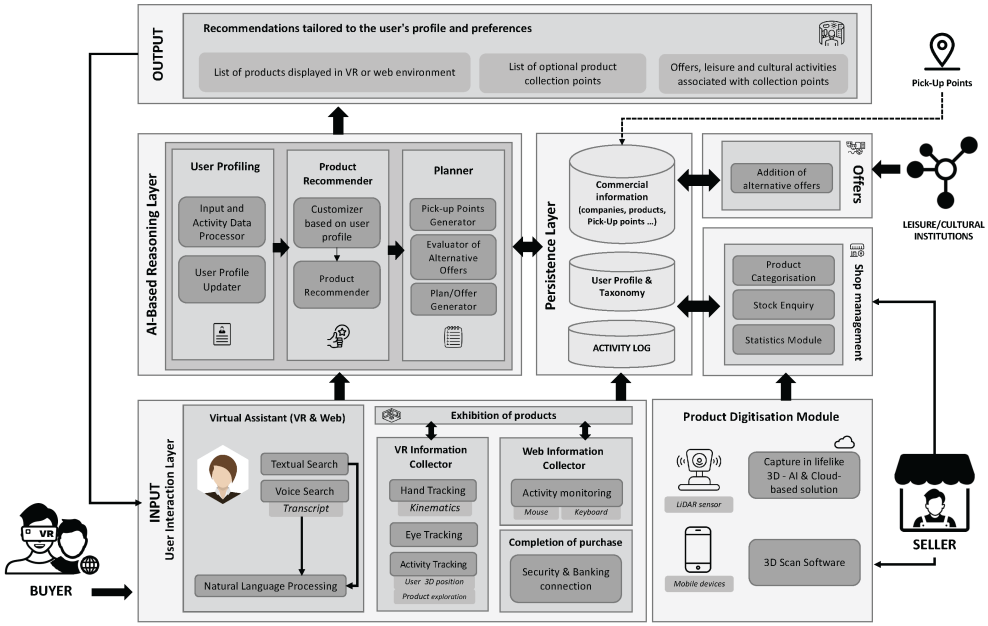


Figura 5.3: Arquitectura de VR-ZOCO: componentes principales, usuarios e interacciones. La capa inferior (*User Interaction Layer*) gestiona la interacción del usuario mediante búsquedas textuales y por voz, *hand-tracking*, *eye-tracking* y procesamiento de lenguaje natural. La capa intermedia (*AI-Based Reasoning Layer*) se encarga del perfilado de usuario, la recomendación de productos y la generación de planes de ocio asociados a puntos de recogida. La capa de persistencia almacena catálogos, perfiles y registros de actividad. El módulo de digitalización de productos permite a los vendedores integrar sus catálogos mediante tecnologías de escaneo 3D de bajo coste.

- **Congreso:** 25th International Conference on Enterprise Information Systems (ICEIS 2023).
- **Editorial:** SciTePress.
- **Páginas:** 312–319.
- **Año:** 2023.
- **Resumen:** The rise of e-commerce and online sales has had a detrimental effect on small businesses that lacked an online presence in recent years, with negative consequences for local services and economies. Despite attempts to digitize businesses, large corporations continue to hold a

Tabla 5.2: Publicaciones en actas de congresos internacionales. Se indica referencia bibliográfica, título, conferencia, año y categoría.

Cita	Título	Conferencia	Año	Categoría
[90]	Supporting Small Businesses and Local Economies Through VR Shopping and AI	ICEIS	2023	C[1] / B1[2]
[91]	A Virtual Reality Shopping Platform for Enhancing E-Commerce Activities of Small businesses and local economies	ISMAR	2023	A*[1] / A2[2]
[6]	User Activity Monitoring and Personalized Recommendations for Enhancing the VR Shopping Experience	IEEE AIxVR	2024	Scopus[3]
[92]	Exploring Interaction Mechanisms and Perceived Realism in Different VR Shopping Setups	ICEIS	2024	C[1] / B1[2]
[93]	Enhancement of the Online Presence of SMEs with Minimum Impact on Traditional Business Activities	ICEIS	2024	C[1] / B1[2]
[94]	VR-ADAPT: An Immersive Learning and Training Environment for Wheelchair Users with Recent SCI	ICEIS	2025	C[1] / B1[2]
[95]	Pilot Study on Gamification and Purchase Behaviour in VR Shopping Environments	ICEIS	2025	C[1] / B1[2]
[96]	Occlusion-Resilient Mixed Reality: A Smart Object Architecture for Upper Limb Rehabilitation	XR Salento	2026	Springer[4]

[1] CORE/ERA Conference Ranking: ICEIS = C, ISMAR = A*. Fuente: `http://www.conferenceranks.com` (Última consulta, marzo de 2026).
[2] Qualis CAPES: ICEIS = B1, ISMAR = A2. Fuente: `http://www.conferenceranks.com` (Última consulta, marzo de 2026).
[3] Indexado en IEEE Xplore y Scopus. No clasificado en CORE ni en Qualis.
[4] Publicado en *Lecture Notes in Computer Science* editado por Springer.

privileged position that allows them to capture the majority of sales. Small businesses may regain a competitive edge against large platforms by anticipating and adapting to the next phases of commercial evolution, which are likely to be heavily reliant on Virtual Reality shopping and Artificial Intelligence. In this article, we propose a platform that combines these two technologies and enables local businesses to join forces and overcome physical barriers, thereby providing a virtual world of unrestricted retail spaces. The system also proposes retail collection points and develops attractive leisure plans around these points through the

outcomes of a recommender system, thereby promoting city activity and bolstering local economies.

- **Relación con la tesis:** Relacionado con O1.

A Virtual Reality Shopping Platform for Enhancing E-Commerce Activities of Small Businesses and Local Economies

- **Autores:** R. Grande, J. Albusac, J.J. Castro-Schez, D. Vallejo, S. Sánchez-Sobrino.
- **Tipo:** artículo de congreso
- **Congreso:** IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct 2023).
- **Editorial:** IEEE.
- **Año:** 2023.
- **DOI:** `https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00175`
- **Resumen:** This doctoral work-in-progress presents VR-ZOCO, a VR-powered shopping platform enhanced with AI that aims to support small businesses and local economies. The platform enables local businesses to offer their products in an immersive virtual environment, incorporating a recommendation system and intelligent agents. The poster discusses the overall architecture, interaction mechanisms, and the research plan for evaluating low-cost 3D scanning technologies for product digitization.
- **Relación con la tesis:** Relacionado con O1 y O2.

User Activity Monitoring and Personalized Recommendations for Enhancing the VR Shopping Experience

- **Autores:** M.G. Arcos, R. Grande, S. Martínez-Cid, C. González-Morcillo, J.J. Castro-Schez, J. Albusac.
- **Tipo:** artículo de congreso.

- **Congreso:** IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR 2024).
- **Editorial:** IEEE.
- **Páginas:** 176–184.
- **Año:** 2024.
- **DOI:** `https://doi.org/10.1109/AIxVR59861.2024.00030`
- **Resumen:** Artificial Intelligence (AI) has had a significant impact on the e-commerce sector since its inception, enabling major retailers to gain a competitive edge. In the forthcoming years, Virtual Reality (VR) is anticipated to emerge as a transformative advancement in human-computer interaction within the e-commerce domain. Therefore, it is crucial to incorporate AI within VR-developed environments to monitor user interactions therein to construct accurate user profiles. In this paper, we propose the definition and collection of four key actions that users can perform in VR Shopping environments. These are further used to build user profiles and provide recommendations tailored to each user. Moreover, the algorithm built for product recommendation is detailed, as well as an evaluation of such recommendation system performed with a custom dataset of categorized products.
- **Relación con la tesis:** Relacionado con O1 y O5.

Exploring Interaction Mechanisms and Perceived Realism in Different Virtual Reality Shopping Setups

- **Autores:** R. Grande, J. Albusac, S. Sánchez-Sobrino, D. Vallejo, J.J. Castro-Schez, C. González-Morcillo.
- **Tipo:** artículo de congreso.
- **Congreso:** 26th International Conference on Enterprise Information Systems (ICEIS 2024), Volume 2.
- **Editorial:** SciTePress.

- **Páginas:** 505–512.
- **ISBN:** 978-989-758-692-7.
- **Año:** 2024.
- **Resumen:** Within the e-commerce field, disruptive technologies such as Virtual Reality (VR) are beginning to be used more frequently to explore new forms of human-computer interaction and enhance the shopping experience for users. This study presents an experiment with a set of participants that addresses: (1) users' evaluation of a set of pre-formalised interaction mechanisms, (2) preference for a large-scale or small-scale shopping environment and the degree of usability while navigating, and (3) the usefulness of monitoring user activity to infer user preferences. The results show that interaction mechanisms made with users' hands are fluid and natural, high usability in both small and large shopping spaces with the latter being preferred by users, and the recorded interactions can be employed for user profiling that improves future shopping experience.
- **Relación con la tesis:** contribución donde se evalúan experimentalmente los mecanismos de interacción y navegación basados en *hand-tracking*, así como la percepción de realismo en diferentes configuraciones de entornos de *VR-shopping*. Relacionado con O3 y O5.

Enhancement of the Online Presence of Small and Medium Sized Enterprises with Minimum Impact on Traditional Business Activities in Towns and Cities

- **Autores:** M. Garrido, J.J. Castro-Schez, R. Grande, S. Schez-Sobrino, D. Vallejo.
- **Tipo:** artículo de congreso.
- **Congreso:** 26th International Conference on Enterprise Information Systems (ICEIS 2024).
- **Editorial:** SciTePress.
- **Páginas:** 989–996.

- **Año:** 2024.
- **DOI:** `https://doi.org/10.5220/0012630600003690`
- **Resumen:** This paper introduces an innovative strategy for an e-commerce portal designed to support small and medium-sized enterprises (SMEs), integrating local businesses not directly related to product sales, referred to as "satellite businesses", such as bars, restaurants, cinemas or sports facilities. This proposal modifies the existing VR-ZOCO e-commerce portal structure to strengthen local economies, facilitating a symbiosis between online shopping and physical leisure activities. Following a purchase on the portal, users are offered the option to collect their products at specific local points. Linked to this collection act, personalized "leisure plans" are generated, based on "leisure activities" from satellite businesses.
- **Relación con la tesis:** Relacionado con O1.

VR-ADAPT: An Immersive Learning and Training Environment for Wheelchair Users with Recent Spinal Cord Injuries

- **Autores:** J. Albusac, D. Cordero, M. Jiménez, R. Grande, V. Herrera, R. Perales, M.E. San Félix, A. de los Reyes.
- **Tipo:** artículo de congreso.
- **Congreso:** 27th International Conference on Enterprise Information Systems (ICEIS 2025).
- **Editorial:** SciTePress.
- **Año:** 2025.
- **Resumen:** Each year, thousands of people worldwide suffer injuries that limit their mobility, affecting not only their ability to walk but also, in many cases, the functionality of their upper limbs. To address this challenge, we present VR-ADAPT, an innovative virtual reality-based platform designed to facilitate the transition to a more autonomous life. VR-ADAPT integrates an advanced simulator for learning to operate electric wheelchairs and digitized environments based on domestic and

workplace settings. These environments are gamified through serious games, allowing users to practice and develop essential skills for confidently navigating their daily lives. Additionally, the platform includes a kinematic recording and analysis module that collects detailed data during exercises, providing clinical teams with a valuable tool for objectively evaluating patients' progress.

- **Relación con la tesis:** Relacionado con el contexto de O4.

Pilot Study on the Effects of Gamification and Virtual Reality on the Shopping Experience

- **Autores:** R. Grande, J. Albusac, S. Sánchez-Sobrino, D. Vallejo, J.J. Castro-Schez.
- **Tipo:** artículo de congreso.
- **Congreso:** 27th International Conference on Enterprise Information Systems (ICEIS 2025).
- **Editorial:** SciTePress.
- **Año:** 2025.
- **Páginas:** 482–493.
- **DOI:** `https://doi.org/10.5220/0013423500003929`
- **Resumen:** E-commerce has embraced emerging technologies such as Virtual Reality (VR) and Augmented Reality (AR), which are transforming consumer interaction. In particular, VR enables the creation of immersive environments that simulate and enhance physical experiences, offering advantages such as gamification of the shopping process, thereby boosting user engagement and interaction. This pilot study addresses two main questions: how does the exploration of digitalized products in VR influence purchase intention? And, does a gamified experience with a specific product have an additional impact? In collaboration with a local business, an experiment involving 48 participants was designed. A portion of the store, its surroundings, and several products were digitalized

using advanced techniques such as Neural Radiance Fields (NeRF) and Gaussian Splatting, achieving realistic models integrated into a virtual store accessible through Meta Quest 3 headsets. One specific product, a comic book, was gamified, allowing users to interact with its narrative, solve challenges, and be incentivized to purchase the product to discover the ending. Preliminary results, including a conversion rate of 41.9 %, suggest that VR, especially when it incorporates gamification, can increase purchase intention and interest in local products, highlighting its potential in digital commerce

- **Relación con la tesis:** Relacionado con O6.

Occlusion-Resilient Mixed Reality: A Smart Object Architecture for Upper Limb Rehabilitation

- **Autores:** R. Echevarría, R. Grande, V. Herrera,A. De los Reyes, J. Albusac.
- **Tipo:** artículo de congreso.
- **Congreso:** XR Salento 2026.
- **Editorial:** Springer.
- **Año:** 2026.
- **Páginas:** No disponibles a junio de 2026.
- **DOI:** No disponible a junio de 2026.
- **Resumen:** Mixed Reality (MR) has emerged as a powerful tool for upper limb rehabilitation, offering immersive serious games that encourage motor recovery. However, current optical tracking systems in commercial headsets face a critical limitation: the loss of tracking during hand-object interaction due to visual occlusion. This disruption breaks the immersive experience and hinders the use of tangible objects with real weight which are essential for restoring haptic feedback. This paper proposes an Occlusion-Resilient Smart Object Architecture designed to ensure continuous interaction in rehabilitation scenarios. The system integrates a low-power embedded system based on the Seeed Studio XIAO nRF52840 Sense, featuring a 6-axis Inertial Measurement Unit (IMU) and pressure

sensors directly into the physical object. We present a hybrid tracking framework where inertial and tactile data transmitted via Bluetooth Low Energy (BLE) are combined with hand tracking and visual object detection to support task-oriented rehabilitation exercises even when the object is obscured by the patient's hand. Technical validation experiments assessed latency, visual freshness, drift, and tracking continuity under severe occlusion conditions. Results demonstrate that our architecture maintains real-time spatial coherence when visual detection fails, providing a robust, scalable, and cost-effective solution for tangible interaction in e-Health systems.

- **Relación con la tesis:** el trabajo se fundamenta en la investigación realizada en la tesis sobre interacción en aplicaciones VR mediante *hand-tracking* y la recogida de datos sobre estas interacciones.

5.2.3 | Congresos nacionales

Además de los artículos reseñados en la sección anterior, el trabajo realizado durante la investigación llevada a cabo en esta tesis doctoral, ha resultado en dos publicaciones y ponencias en congresos nacionales.

Tabla 5.3: Publicaciones en actas de congresos nacionales. Se indica referencia bibliográfica, título, conferencia y año.

Cita	Título	Conferencia	Año
[97]	Adaptation of interaction mechanisms in Virtual Reality shopping environments for people with upper limb motor difficulties	CEIG 2024 (CEDI)	2024
[98]	VR-ADAPT: A Virtual Reality Environment for Skill Development in New Wheelchair Users with SCI	XoveTIC	2025

Adaptation of interaction mechanisms in Virtual Reality shopping environments for people with upper limb motor difficulties

- **Autores:** R. Grande, V. Herrera, C. Glez-Morcillo, A. Reyes-Guzmán, J.J. Castro-Schez, J. Albusac.

- **Tipo:** artículo de congreso.

- **Congreso:** Spanish Computer Graphics Conference (CEIG 2024), dentro del Congreso Español de Informática (CEDI 2024).
- **Editorial:** The Eurographics Association.
- **Año:** 2024.
- **Resumen:** In recent years, there has been research and exploration into the development of new shopping experiences within the field of electronic commerce (e-commerce). One of the technologies that can offer a more immersive shopping experience is Virtual Reality (VR). Retail giants such as Amazon and Alibaba Group have begun to use it. The technological advancement of VR, motivated by its use in various domains like e-commerce, has driven the development of software tools like APIs which allow developers to easily develop applications for these devices. One of the latest technologies included in recent VR headsets is hand tracking, which allows users to use their own hands as an input method to interact with the virtual environment. However, software tools for the development of VR applications are not fully adapted to include accessibility options for people with motor difficulties in their bodies, making it very difficult for these people to use this technology with both controllers and hand tracking. To promote accessibility options in the use of VR shopping environments, this study will present the adaptation of a set of interaction mechanisms, among which we highlight: automatic object grabbing, release of grabbed objects, navigation through the environment, attraction of distant objects, and interaction with the shopping cart. These adaptations will be made using Meta's API for Meta Quest devices as a base. The adapted environment has been tested by healthy students from the faculty and one of them with reduced mobility in the left half of his body after suffering a stroke. In this paper, we present the feedback provided by the volunteers, as well as the verification that these interaction mechanisms meet our expectations. This is an essential previous step to carry out a planned experimental session with patients with spinal cord injuries and therapist at the National Hospital for Paraplegics in Toledo (HNPT).
- **Relación con la tesis:** Relacionado con el contexto de O4.

VR-ADAPT: A Virtual Reality Environment for Skill Development in New Wheelchair Users with Spinal Cord Injuries

- **Autores:** D. Cordero, R. Grande, V. Herrera, M. Jiménez, J. Albusac.
- **Tipo:** artículo de congreso.
- **Congreso:** XoveTIC 2025.
- **Año:** 2025.
- **Resumen:** Annually, hundreds of thousands of individuals worldwide experience mobility limiting injuries that compromise not only their walking ability but frequently the functionality of their upper extremities as well. Such injuries constitute a profound life transformation, necessitating an intensive period of learning and adjustment to new physical realities. We introduce VR-ADAPT, a novel virtual reality platform conceived to support this critical transition toward greater independence. The system comprises three integrated components: an advanced electric wheelchair operation simulator, immersive digital recreations of home and professional environments, and a suite of therapeutic serious games embedded within these spaces. Through gamification strategies, users can safely acquire and refine the competencies needed to navigate their everyday environments with confidence. A comprehensive kinematic capture and analysis module continuously records detailed performance metrics throughout training sessions, furnishing rehabilitation professionals with objective, quantifiable data to track patient advancement. This evidence-based approach enables more precise therapy customization and improved clinical outcomes.
- **Relación con la tesis:** Uso de las interacciones adaptadas en el apartamento adaptado escaneado para su inclusión en la aplicación usada. Relacionado con O4.

5.2.4 | Capítulos de libro

Narrowing the Technological Gap by Promoting Small Commerce Through VR and AI for a Lifelike E-Commerce Experience: Needs and Solutions

- **Autores:** R. Grande, S. Sánchez-Sobrino, D. Vallejo, J.J. Castro-Schez, J.A. Albusac.
- **Tipo:** capítulo de libro.
- **Libro:** Enterprise Information Systems (ICEIS 2023).
- **Serie:** Lecture Notes in Business Information Processing (LNBIP), vol. 519.
- **Editorial:** Springer.
- **Páginas:** 92–112.
- **ISBN:** 978-3-031-64754-3 (impreso), 978-3-031-64755-0 (online).
- **Año:** 2024.
- **DOI:** `https://doi.org/10.1007/978-3-031-64755-0_5`
- **Resumen:** In recent years, the increasing popularity and demand of e-commerce services and online retail has adversely impacted small businesses without an online presence, leading to unfortunate consequences for local services and economies. Despite efforts to digitise small commerce, large retail businesses continue to maintain a superior position in technological capability. In this chapter, we introduce VR-ZOCO, a platform that integrates Virtual Reality and Artificial Intelligence, enabling local businesses to collaborate and overcome physical limitations, consequently offering a limitless virtual retail space. The system also incorporates retail collection points and constructs appealing leisure plans around these points via the results of a recommendation system, thus stimulating urban activity and strengthening local economies. Moreover, we explore the needs of small businesses to obtain 3D models of their products and incorporate them into VR-ZOCO, as well as proposed solutions and how we will research them.
- **Relación con la tesis:** versión extendida del *paper* presentado en ICEIS 2023, donde se detalla la propuesta arquitectónica de VR-ZOCO, las necesidades de digitalización de productos para PYMEs y las soluciones propuestas basadas en VR e IA. Relacionado con O1 y O2.

5.3 | Otras actividades



Figura 5.4: Fotografías tomadas durante reuniones con Cámara de Comercio de Ciudad Real (arriba izquierda) y Toledo (arriba derecha) y el evento Feria Transforma CLM (resto de fotografías)

Durante el desarrollo de esta tesis, se ha tratado de ser proactivo a la hora de divulgar y dar a conocer la iniciativa en la que se basa (ver 5.4). Algunas de estas actividades fueron la asistencia a la **feria tecnológica Transforma CLM**, participación en el **III Foro de Digitalización** y reuniones formales con las Cámaras de Comercio de Ciudad Real y Toledo. En cierta medida, esto ha favorecido el desarrollo de sesiones experimentales en colaboración con distintas organizaciones y pequeños comercios, como **Librería Serendipia**, permitiendo obtener conocimiento empírico para validar las preguntas de investigación e hipótesis de este trabajo.

5.4 | Premios obtenidos

Durante el desarrollo de esta tesis se han obtenido los siguientes premios y reconocimientos, que ponen de manifiesto el impacto y la calidad de las contribuciones realizadas:

- **Premio a mejor expediente de Máster Oficial en Ingeniería Informática (UCLM)**. Premio obtenido tras la realización del máster habilitante para matricularme en el programa del presente doctorado, siendo el TFM de dicho máster el punto de partida de la presente tesis.

- **Best Student Paper Award – ICEIS 2025.** Premio al mejor artículo científico de estudiantes de doctorado en la 27ª edición de la *International Conference on Enterprise Information Systems* (ICEIS 2025), celebrada en Oporto (Portugal) del 4 al 6 de abril de 2025, por el trabajo *"Pilot study on the effects of gamification and virtual reality on the shopping experience"*. Este premio fue otorgado por el comité científico y organizador de la conferencia, reconociendo la relevancia y calidad del estudio piloto sobre el efecto de la gamificación y la Realidad Virtual en el comportamiento de compra real del consumidor.

- **NTT Data Chair Award – VIII Congreso XoveTIC 2025.** Premio de la Cátedra NTT Data otorgado a miembros del grupo de investigación AIR en el VIII Congreso XoveTIC 2025, celebrado en octubre de 2025, por la contribución *"VR-ADAPT: A Virtual Reality Environment for Skill Development in New Wheelchair Users with Spinal Cord Injuries"*. En este trabajo se integraron los mecanismos de interacción adaptados planteados en el Artículo III del compendio.

CAPÍTULO 06

06

Discusión de resultados

Lectura crítica de los hallazgos y su alcance frente al estado del arte.

6

Discusión de resultados

Este capítulo analiza e interpreta de manera integrada los resultados obtenidos a lo largo de la presente tesis doctoral con el propósito de dar respuesta explícita a las preguntas de investigación formuladas en la Sección 3.1 y de evaluar el grado en que la hipótesis enunciada en el Capítulo 3 queda validada por la evidencia acumulada. A diferencia del Capítulo 5, donde cada artículo del compendio se presenta de forma autocontenida, la discusión que se desarrolla a continuación adopta una perspectiva transversal: las contribuciones no se exponen una por una, sino que se presentan de forma conjunta, de modo que se haga visible la cadena causal y las implicaciones conjuntas que de ella se derivan.

La hipótesis de esta tesis sostiene que la implementación de sistemas de *VR-shopping* basados en la digitalización de bajo coste e interacción gestual inclusiva es técnicamente viable y mejora significativamente la experiencia de compra, generando un impacto cuantificable en el comportamiento de compra real del consumidor y habilitando la recolección de datos comportamentales más ricos que los disponibles en entornos web convencionales. Contrastar esta afirmación requiere ir más allá del análisis individual de cada publicación: exige demostrar que los resultados obtenidos en la digitalización de productos, el desarrollo de herramientas de software, la adaptación de las interacciones para llegar a una solución plural y global que no excluya usuarios por sus limitaciones y la validación empírica en un escenario comercial real conforman un conjunto de trabajos coherente que se sostienen mutuamente.

La estructura de este capítulo se organiza en torno a las seis preguntas de investigación (RQ1-RQ6). Las RQs no son independientes entre sí: como se discutirá, RQ1 actúa como marco arquitectónico sobre el que se basan las restantes; RQ2 cimenta técnicamente el resto al resolver la digitalización de

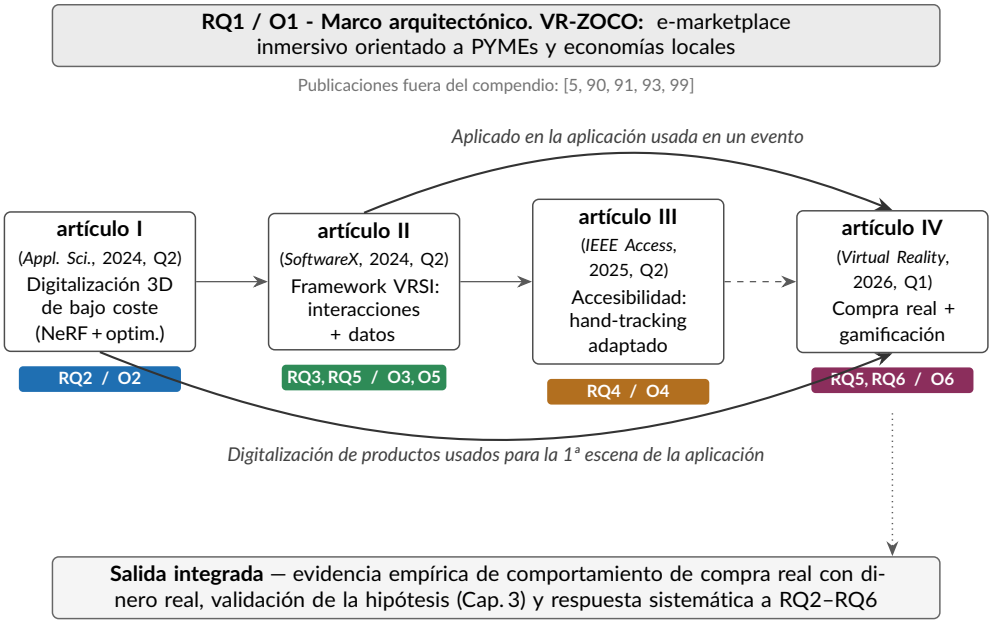


Figura 6.1: **Cadena de validación de la tesis.** Los cuatro artículos del compendio se ordenan según las dependencias técnicas y narrativas entre sus contribuciones, no según su cronología estricta. Las flechas continuas señalan transferencia de artefactos materializados; la flecha discontinua indica transferencia de principios de diseño. Las dos flechas en arco más gruesas evidencian que artículo IV integra simultáneamente productos digitalizados de artículo I y la infraestructura de captura y análisis de artículo II, cerrando el bucle empírico en un evento comercial real.

objetos con resultados de buena calidad a bajo coste; RQ3 y RQ4 conforman el plano interactivo, con RQ4 ampliando la base poblacional alcanzable por RQ3; y RQ5 y RQ6 completan el ciclo aportando, respectivamente, la infraestructura analítica y la validación empírica del impacto.

La Fig.6.1 sintetiza visualmente esta relación. Los cuatro artículos del compendio se representan como nodos de un grafo dirigido en el que las aristas materializan transferencias concretas de artefactos entre publicaciones: modelos 3D digitalizados que se utilizan en la fase experimental del último trabajo del compendio, una infraestructura software que se despliega en una feria comercial, un conjunto de interacciones gestuales que se adapta a personas con limitaciones motoras.

Cada sección dedicada a una RQ comienza con una tabla compacta que

enumera las publicaciones que contribuyen a esa pregunta y las relaciona con los objetivos y subobjetivos parciales definidos en la Sección 3.2. Estas tablas integran tanto los artículos del compendio (Sección 5.1, destacados en negrita) como las publicaciones realizadas durante la tesis fuera del compendio (Sección 5.2), siempre que aporten evidencia adicional o contexto relevante.

Las Secciones 6.1 a 6.6 abordan secuencialmente las seis preguntas de investigación, siguiendo la lógica de dependencias descrita anteriormente.

6.1 | RQ1: marco arquitectónico

> ***Pregunta de investigación 1.*** *«¿Cómo deben evolucionar los e-marketplaces actuales para integrar de forma holística tecnologías inmersivas y de IA, atendiendo a las necesidades específicas de las PYMEs y las economías locales?»*

La tabla 6.1 resume las contribuciones por publicación para responder esta pregunta de investigación.

Tabla 6.1: **Mapeo de contribuciones a RQ1.** Ninguna de las publicaciones que aborda RQ1 pertenece al compendio: la respuesta arquitectónica se construyó en seis publicaciones realizadas durante la tesis.

Publicación (abreviatura)	Venue / año	Tipo	O1	
			SO1.1	SO1.2
VR-ZOCO architectural proposal [5]	Appl. Syst. Innov., 2024	AR	●	●
Narrowing the Technological Gap	Springer LNBIP, 2024	CL	●	●
Supporting Small Businesses (position artículo)	ICEIS, 2023	CI	●	•
VR Shopping Platform	ISMAR Adj., 2023	CI	●	●
Enhancement of SMEs' Online Presence	ICEIS, 2024	CI	●	●
User Activity Monitoring & Recommendation	IEEE AIxVR, 2024	CI	○	•

Contribución: ● principal • parcial ○ no contribuye **Tipo:** AC compendio AR revista CL cap. libro CI congr. int. CN congr. nac.

6.1.1 | Estado actual de los *e-marketplaces*

La revisión bibliométrica de Cano et al. [38] identifica ocho líneas de investigación activas en el desarrollo de *e-marketplaces*. Entre ellas destacan la adopción por parte de PYMEs, el diseño de la experiencia de usuario y la sostenibilidad, pero ninguna contempla la VR como canal nativo de interacción. Las plataformas comerciales dominantes y los trabajos de la literatura analizados en el estado del arte (Sección 2.1) coinciden en este punto. Cuando incorporan tecnologías inmersivas, lo hacen mediante extensiones puntuales (visualización 3D, AR sobre catálogos web, salas de muestra virtuales) que se añaden a una arquitectura concebida para el navegador.

Los trabajos sobre VR-shopping de Speicher [48, 49], Pizzi et al. [52] y Peukert [46] junto con los estudios más recientes sobre tiendas de moda inmersivas de Ricci et al. [34], demuestran que la sinergia entre el uso de VR y el campo del e-commerce es positiva. Sin embargo, todos ellos abordan tiendas aisladas, no e-marketplaces. Las revisiones de Xi y Hamari [33] y de Wedel et al. [66] confirman la madurez tecnológica y su efecto positivo sobre la presencia, pero señalan que llevar la VR al e-commerce a escala de plataforma, y sobre todo al ámbito de la PYME, sigue siendo una asignatura pendiente.

A esta limitación arquitectónica se suma una limitación de mercado. La aceleración de la digitalización tras la pandemia de COVID-19 [23] ha afectado a la posición competitiva de las PYMEs, que disponen de menos recursos para incorporar tecnologías punteras. Cualquier propuesta que pretenda responder a RQ1 debe cumplir entonces con tres condiciones simultáneas: incorporar la VR como canal de primer nivel, integrar capacidades de IA que aprovechen la riqueza informativa del entorno inmersivo, e incluir mecanismos explícitos para que el e-marketplace no actúe como sustituto del comercio físico local sino como su aliado.

6.1.2 | Evolución iterativa de la propuesta

El capítulo de libro publicado en la serie Springer LNBIP [99] dio nombre al problema central: la brecha tecnológica entre PYMEs y grandes operadores como objeto explícito de cierre. Este trabajo aporta evidencia empírica sobre el interés de las PYMEs por digitalizarse (incremento del 30 % de presencia online en Bélgica tras la pandemia, 88 % de PYMEs en Indonesia interesadas en adop-

tar e-commerce) y profundiza por primera vez en los módulos individuales de la arquitectura. Anticipa además aspectos que el artículo I del compendio desarrollaría después en el plano técnico de digitalización 3D.

El artículo de ICEIS 2023 [90] presentó la idea inicial: una plataforma de VR-shopping con tres capas conceptuales (entrada VR/web, capa de razonamiento basada en IA, salida personalizada) y un mecanismo de asociación entre pequeños comercios que distribuyera de forma equitativa las ventas generadas.

El trabajo presentado en el ISMAR-Adjunct de 2023 [91] llevó esta idea al ámbito específico de las tecnologías inmersivas y refinó los componentes de captura de datos. Más allá del valor de la validación externa que supone su aceptación en un congreso internacional de valoración A*, este trabajo aporta la primera publicación del diagrama de arquitectura completo que más adelante se formalizaría, y una caracterización analítica del problema desde tres tipos de actores: el vendedor, el comprador y la ciudad o pueblo.

El capítulo de libro publicado en la serie Springer LNBIP [99] extendió la propuesta y enfatizó cómo la combinación de VR e IA puede reducir la brecha tecnológica que separa a las PYMEs de los grandes comerciantes. Además, profundiza por primera vez en los módulos individuales de la arquitectura. Anticipa además aspectos que el artículo I del compendio desarrollaría después en el plano técnico de digitalización 3D.

La formalización completa de la propuesta llegó con Castro-Schez et al. [5], donde el e-marketplace se define como la quíntupla $VR-ZOCO = \{E, C, M, U, B\}$: E es el conjunto de entornos en los que la plataforma está disponible (dispositivos VR, smartphones y navegadores web), C es el catálogo, M es el conjunto de módulos funcionales, U son los usuarios registrados y B son los negocios participantes. Cada componente se descompone a su vez de manera explícita.

El catálogo C se divide en cuatro secciones: C_1 son productos físicos o digitales (libros, electrónica, moda, alimentación, etc.), C_2 son espectáculos públicos celebrados en la ciudad, C_3 son actividades recreativas y C_4 son tipos de alojamiento turístico. El conjunto de módulos funcionales de la plataforma son: $M = \{m_{UI}, m_{PC}, m_{PM}, m_{UM}, m_{OP}, m_{SMI}, m_{VRE}, m_{AI}, m_{ABI}, m_{CC}, m_{BM}, m_{SC}\}$, los cuales establecen doce subsistemas con responsabilidades disjuntas, desde la gestión de la interfaz de usuario hasta el cumplimiento normativo. Por último,

los usuarios U se modelan desde cinco perspectivas complementarias: datos demográficos (dd), datos psicográficos (pd), comportamiento en la plataforma digital ($dcpw$), comportamiento en el mundo virtual ($dcmv$) y comportamiento en el mundo real ($dcmr$, que incluye visitas a tiendas físicas y puntos de recogida).

6.1.3 | Formalización del marco arquitectónico

La formalización anterior se desarrolla en el trabajo [5] fuera del compendio de artículos. Dicha formalización contempla de forma holística todas las características que un *e-marketplace* aplicando la arquitectura de VR-ZOCO debería poseer. Si se desea consultar de nuevo la arquitectura de VR-ZOCO propuesta en este trabajo, vea Fig. 6.2. En esta subsección nos limitaremos a los aspectos relevantes a la presente tesis. Por ejemplo, no se van a detallar el catálogo extendido a actividades de ocio o turismo o los mecanismos de fortalecimiento del tejido comercial local.

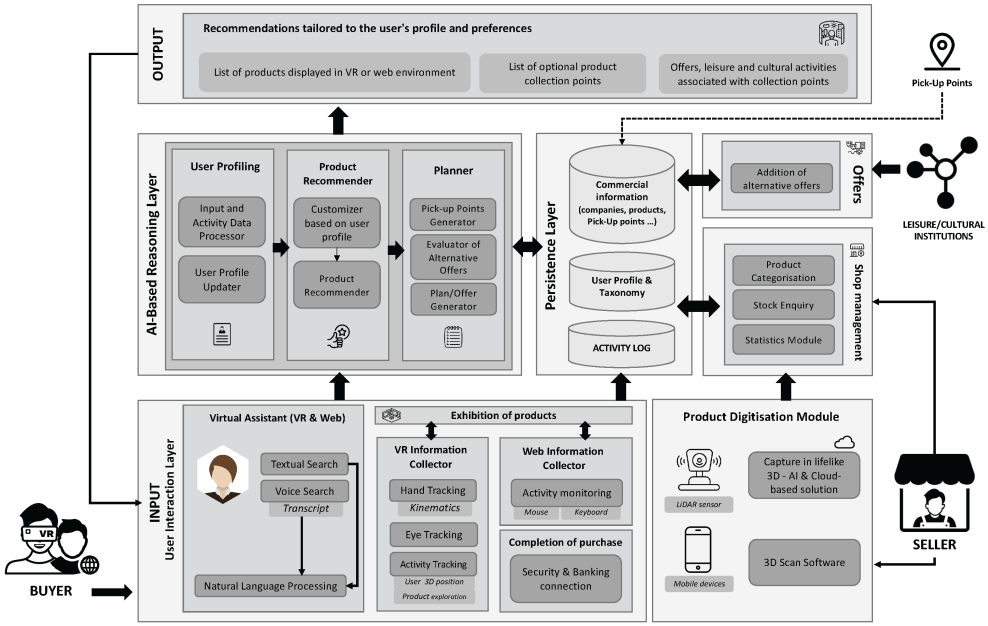


Figura 6.2: Arquitectura de VR-ZOCO. Por comodidad del lector, se vuelve a incluir en la sección de discusión.

Multientornos como propiedad nativa. El conjunto $E = \{VR, mobile, web\}$ se trata de manera simétrica en la arquitectura. La funcionalidad FBU8 (Multichannel User Data Integration) y FBA12 (Unified Content Management) descritas en [5] garantizan que la información se sincroniza automáticamente entre canales sin necesidad de carga múltiple por parte del comerciante. Esto contrasta con la situación habitual en plataformas como Amazon o Etsy, donde la incorporación de visualización inmersiva (cuando existe) opera como capa superpuesta sobre un catálogo concebido para el navegador. En la Fig. 6.3 se muestra un prototipo, desarrollado en la presente tesis, que incluye el mapa de una ciudad (Toledo) digitalizado y puntos de interés. En VR-ZOCO, los datos de comportamiento *dcmv* recogidos en el entorno virtual se tratan con el mismo nivel de importancia que los datos *dcpw* recogidos en la web.

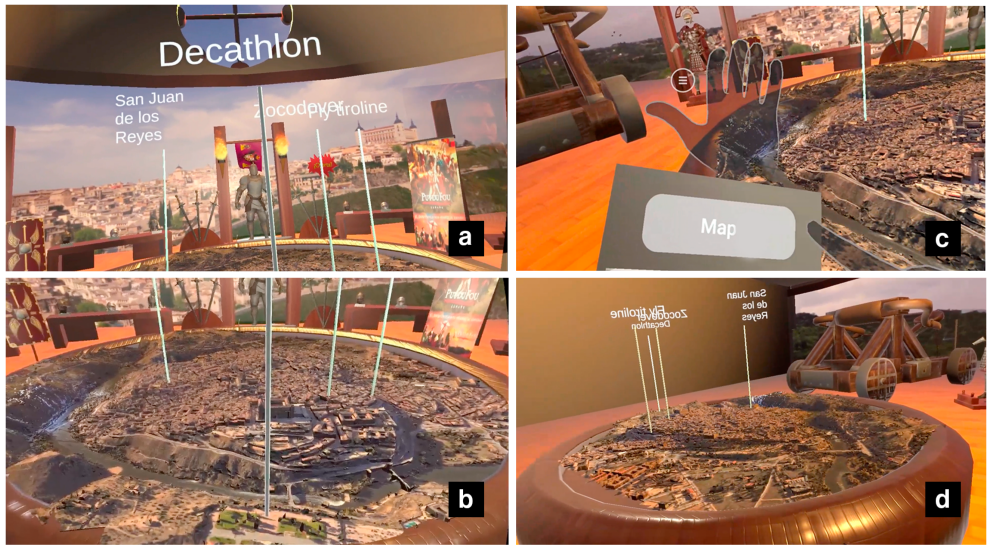


Figura 6.3: (a) El entorno está inspirado en la ciudad de Toledo y su cultura. Al fondo se pueden ver algunas réplicas de armas medievales, junto con una vista real de 360 grados de la ciudad de Toledo. (b) Al bajar ligeramente la mirada, se puede ver una réplica real de un mapa en 3D de la ciudad de Toledo, en el que varios haces de luz resaltan los puntos de interés que forman parte del itinerario generado asociado a la compra. (c) El usuario tiene un panel virtual acoplada a su mano izquierda, que puede utilizar para ocultar o mostrar el mapa de la ciudad en cualquier momento. (d) Vista del mapa desde otra perspectiva en el espacio virtual.

Digitalización de productos integrada como módulo de la plataforma. La funcionalidad FBS11 (Support for 3D Model Digitization) y el módulo de digitalización descrito en [5] reconocen que la barrera de entrada para la PYME no es solo de tipo financiero sino también de proceso. VR-ZOCO no se limita a aceptar modelos 3D ya generados, sino que proporciona herramientas de escaneo accesibles desde smartphone (cámara, LiDAR si está disponible) y un pipeline de procesado pensado para usuarios sin formación técnica especializada. La validación experimental de esta funcionalidad es el contenido del artículo I del compendio [1] y se discute en detalle en la Sección 6.2.

Recomendación basada en señales del entorno inmersivo. La capa m_{AI} incorpora un sistema de recomendación cuyo prototipo (ver Fig. 6.4) se desarrolló y publicó en [6]. A diferencia de los recomendadores tradicionales que operan sobre clics, vistas de página y compras históricas, este sistema se alimenta de las señales más ricas que el módulo de captura m_{UI} obtiene en el entorno virtual: trayectorias de navegación, fijaciones de la mirada, tiempos de manipulación de cada producto y patrones de interacción con el carrito. La hipótesis subyacente, según la cual estas señales son indicadores más fiables del interés y la predisposición a comprar que los datos disponibles en web, se valida posteriormente en RQ5 (Sección 6.5) y RQ6 (Sección 6.6) sobre datos empíricos reales.

Accesibilidad como funcionalidad nativa. La funcionalidad FE10 (Virtual Reality Environment Adaptability) descrita en [5] incluye explícitamente ajustes de navegación e interacción para usuarios en silla de ruedas, opciones de control para personas con movilidad reducida y configuraciones visuales y auditivas específicas. Esta declaración a nivel arquitectónico encuentra su validación experimental en el artículo III del compendio [3], en el que se prueban adaptaciones concretas con pacientes con lesión medular del Hospital Nacional de Parapléjicos de Toledo. La accesibilidad no aparece como un parche posterior a la implementación, sino como un requisito declarado en la propia especificación del módulo m_{UI}.

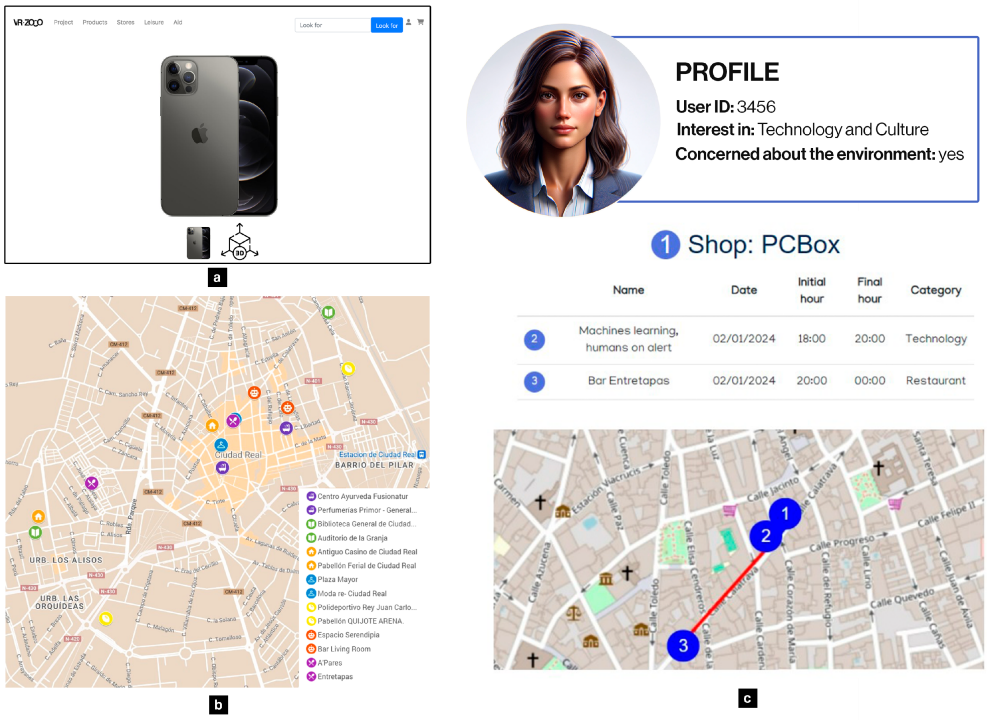


Figura 6.4: (a) Búsqueda y compra de un producto realizadas a través del prototipo por el usuario con el ID 3456. Este decide adquirir un producto tecnológico. (b) El usuario interactúa con el sistema desde Ciudad Real (España). (c) El usuario, según su perfil, está interesado en productos tecnológicos y puede considerarse una persona con estudios, ya que en el pasado ha comprado libros y asistido a conferencias recomendadas por el portal. Además, siempre ha preferido comprar en tiendas locales siempre que ha sido posible. El sistema recomienda comprar el producto en una tienda de Ciudad Real, PCBox, que tiene el producto en stock, y luego asistir a la conferencia *Machine Learning, Humans on Alert* en Espacio Serendipia, una librería cercana al punto de recogida. Por último, el sistema sugiere tomar un aperitivo en un bar cercano.

6.1.4 | Nivel de madurez

Conviene ser riguroso respecto al alcance real de la contribución. VR-ZOCO no se ha desplegado como plataforma comercial, ni esta tesis lo exhibe de tal forma. La contribución es el *modelo arquitectónico* con su formalización completa como referencia de diseño para futuros *e-marketplace*s inmersivos. Dicho esto, los componentes que conforman la arquitectura sí se han implementado y validado experimentalmente por módulos, a lo largo de la tesis. El módulo de

digitalización de productos se valida en el artículo I del compendio [1]. Los componentes de la capa m_{UI} correspondientes a hand-tracking, eye-tracking y captura de comportamiento *dcmv* se implementan y validan en el artículo II [2]. Las adaptaciones de FE10 para usuarios con limitaciones motoras se prueban en el artículo III [3]. El despliegue integrado en un escenario comercial real, que ejercita simultáneamente el módulo m_{VRE}, el m_{PM} y el m_{UI}, se reporta en el artículo IV [4]. Por su parte, los componentes del módulo m_{AI} correspondientes a recomendación y planificación se prototipan en [6] y [93] respectivamente.

La Fig. 6.5 sintetiza este estado de madurez por módulo. De los trece elementos identificados (los doce módulos de M más el módulo transversal de digitalización), siete se han validado en publicaciones del compendio JCR, tres en publicaciones de soporte fuera del compendio y tres permanecen como propuestas a nivel de diseño (búsqueda textual con NLP, procesado transaccional y onboarding completo de PYMEs).

6.1.5 | Posicionamiento frente a la evolución tecnológica reciente

Los avances en síntesis 3D mediante Gaussian Splatting y modelos generativos de difusión hacen aún más accesible la digitalización de productos por parte de actores con recursos limitados, lo que refuerza la viabilidad del módulo de digitalización en el futuro inmediato. La aparición de grandes modelos de lenguaje multimodales abre la puerta a que partes del módulo m_{AI} se sustituyan por agentes conversacionales con conciencia espacial del entorno virtual, lo que es coherente con la funcionalidad FE11 (AI-powered Virtual Assistants) ya declarada en la especificación de VR-ZOCO. La decisión de separar la capa de razonamiento de las capas de interacción y persistencia permite incorporar estos avances mediante la sustitución de componentes, sin necesidad de un rediseño estructural.

6.1.6 | Síntesis de la respuesta a RQ1

La respuesta a RQ1 puede formularse en dos niveles complementarios. En el nivel descriptivo, los *e-marketplaces* actuales presentan tres carencias funcionales: tratan la VR como una extensión ocasional en lugar de un canal

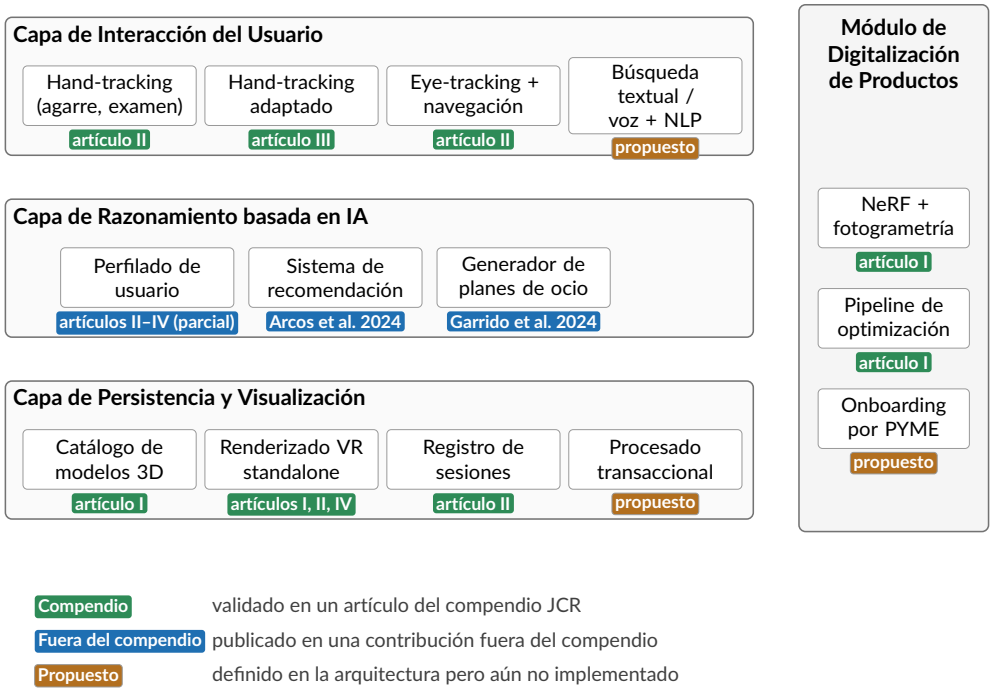


Figura 6.5: **Estado de materialización de los módulos de la arquitectura VR-ZOCO.** De forma simplificada, la arquitectura propuesta se descompone en tres capas funcionales y un módulo transversal de digitalización de productos. De los trece módulos identificados, siete se han implementado y validado experimentalmente en artículos del compendio JCR, tres en publicaciones fuera del compendio y tres permanecen como elementos propuestos a nivel de diseño.

principal, no proporcionan mecanismos accesibles para que las PYMEs digitalicen sus catálogos y operan al margen de las dinámicas comerciales locales. Esta tesis propone una respuesta concreta y formalizada: la evolución de los e-marketplaces requiere un modelo arquitectónico que (i) trate los entornos VR, móvil y web de manera simétrica como conjunto E, (ii) extienda el catálogo a entidades no estrictamente comerciales para permitir la vinculación con la economía de proximidad, (iii) modele al usuario desde cinco perspectivas comportamentales para alimentar un sistema de IA que opere como planificador y no como recomendador clásico, e (iv) incluya como módulos de alta importancia la digitalización de productos por parte del comerciante y la accesibilidad para usuarios con diversidad funcional. VR-ZOCO instrumenta

esta evolución y sus componentes se han validado de manera modular a lo largo del compendio y de las publicaciones de soporte.

6.2 | RQ2: digitalización 3D de bajo coste

> ***Pregunta de investigación 2.*** *«¿Pueden producirse modelos 3D con calidad visual y rendimiento suficientes para su integración en entornos de VR-shopping ejecutados en dispositivos comerciales a bajo coste?»*

La tabla 6.2 resume las contribuciones por publicación para responder esta pregunta de investigación.

Tabla 6.2: **Mapeo de contribuciones a RQ2.** El artículo I del compendio (resaltado en negrita) es la única publicación que aborda los cuatro subobjetivos en el plano experimental. El capítulo de libro [99] y el artículo presentado en el congreso ISMAR [91] contextualizan la necesidad de digitalización por parte de las PYMEs (SO2.1) pero no aportan evaluación técnica.

Publicación	Venue / año	Tipo	O2			
			SO2.1	SO2.2	SO2.3	SO2.4
Performance Evaluation of low-cost 3D scanning	Appl. Sci., 2024	AC	●	●	●	●
Narrowing the Technological Gap	Springer LNBIP, 2024	CL	●	○	○	●
VR Shopping Platform	ISMAR Adj., 2023	CI	●	○	○	○

Contribución: ● principal ● parcial ○ no contribuye **Tipo:** AC compendio AR revista CL cap. libro CI congr. int. CN congr. nac.

6.2.1 | La digitalización 3D en el entorno PYME

La digitalización 3D de productos físicos, históricamente, no ha sido un aspecto de preocupación de las PYMEs. Los escáneres profesionales ofrecen calidad excelente pero su precio oscila entre miles y decenas de miles de euros, una barrera económica inaccesible para la mayor parte del tejido comercial local. El trabajo de [99] aporta evidencia cuantitativa sobre esta brecha: el interés de las PYMEs por digitalizarse es alto, pero la mayoría carece de los recursos económicos y humanos para hacerlo con herramientas profesionales. La arquitectura VR-ZOCO definida en [5] reconoce este problema pero no responde

la pregunta técnica: ¿qué tecnologías concretas, accesibles desde dispositivos comerciales, producen modelos con buena calidad gráfica y rendimiento suficientes para uso real en VR? Esa es exactamente la pregunta que el artículo I del compendio [1] aborda de manera sistemática, y es por ello la fuente de toda la evidencia que se discute en lo que resta de sección.

6.2.2 | Selección de las tecnologías representativas (SO2.1)

El primer subobjetivo exige identificar qué tecnologías de escaneo cumplen el criterio de bajo coste y son operables sin formación técnica especializada. El artículo I parte de un análisis comparativo de seis aplicaciones móviles que cubren las tecnologías representativas del mercado en ese momento: Polycam, Widar, Kiri Engine, Luma AI, MagiScan y Scaniverse. La comparación contempla la técnica subyacente (fotogrametría, LiDAR, NeRF), la suscripción (la mayoría en torno a USD 10-15 al mes; Luma AI y Scaniverse gratuitas), los formatos de exportación soportados, las limitaciones del plan gratuito y la valoración de usuarios. De este análisis salen dos herramientas seleccionadas: **Polycam**, por ser la única que soporta tanto fotogrametría como LiDAR en un mismo flujo, ofrecer múltiples formatos de exportación, contar con herramientas de edición y disponer de documentación extensa; y **Luma AI**, por ser la única implementación móvil comercial de NeRF en el momento del estudio, gratuita, y respaldada por trabajos previos que demuestran su fiabilidad [58, 59].

El hardware utilizado refleja también la decisión de mantener el coste bajo. Para fotogrametría se eligió un *Xiaomi Redmi Note 8*, un dispositivo Android de gama media-baja con cámara cuádruple de 48 MP. Para LiDAR y NeRF se empleó un *iPad Pro 2022*, equipado con cámara dual de 12 MP, lente ultra-angular de 10 MP y sensor LiDAR integrado. La elección del iPad para NeRF no responde a una restricción técnica de Luma AI sino a la compatibilidad de la aplicación, pues en ese momento solo estaba disponible en dispositivos basados en iOS. Esta selección satisface SO2.1 con tres aportaciones distinguibles: (i) una caracterización sistemática del mercado de aplicaciones de escaneo móvil, (ii) un criterio explícito de selección basado en accesibilidad económica, versatilidad y reproducibilidad, y (iii) la identificación de las dos herramientas que, en conjunto, cubren las tres familias técnicas relevantes

(reconstrucción multivista clásica, escaneo asistido por sensor activo, síntesis neuronal de escena).

6.2.3 | Calidad visual y error geométrico (SO2.2)

La evaluación de calidad combina cuatro métricas complementarias para evitar los sesgos de cualquiera de ellas por separado. Dos métricas no-referenciales sobre nubes de puntos: NR-3DQA, basada en una red neuronal SVR entrenada sobre la base de datos Waterloo Point Cloud, y MM-PCQA, basada en un mecanismo de atención multimodal entre características geométricas y proyecciones 2D. Una métrica con referencia: CMDM, que mide la distorsión perceptual entre una malla y su referencia combinando curvatura geométrica y características de color en el espacio LAB200HL. Y finalmente, el error geométrico cuantitativo medido como RMS (mediante CloudCompare) y como MSDM2, calibrados sobre seis figuras de poliestireno expandido modeladas previamente en Blender con medidas físicas exactas (un cubo, una esfera, una pirámide y las letras A, I y R) que actúan como *ground truth* del experimento.

Los resultados sobre el conjunto de figuras de control son contundentes. El error geométrico medio del NeRF en su configuración de alta resolución es de 0,42 cm sobre figuras de aproximadamente 30 cm, lo que supone un 1,4 % de error relativo. La fotogrametría sobre el mismo conjunto produce un error medio del orden de 2,36 cm, y el LiDAR queda por detrás de ambas. El valor MSDM2 medio del NeRF es de 0,112, dato comparable al umbral de 0,14 que los autores de la métrica identifican como apenas perceptible para el ojo humano. La precisión geométrica resultante es muy aceptable para uso comercial de visualización, aunque no para mediciones críticas (un repuesto industrial con tolerancia inferior al milímetro, por ejemplo).

Sobre el conjunto de cinco productos reales de distintos materiales (tela, cuero, madera, metal y cerámica) y tamaños pequeños y medianos (menos de 1 metro de longitud/anchura). Los resultados de calidad perceptual confirman el mismo patrón, los cuales pueden verse resumidos y a grandes rasgos en la Fig. 6.6. NeRF mediante Luma AI gana en las cuatro métricas. Frente a la fotogrametría, las mejoras medias son del 17,78 % en NR-3DQA y del 9,43 % en MM-PCQA; frente al LiDAR, las mejoras crecen al 22,73 % y al 28,19 %

respectivamente. La diferencia varía con el producto: en el quemador, una pieza con materiales mate y geometría simple, los tres métodos producen modelos casi indistinguibles (diferencia de solo el 6 % entre el mejor y el peor). En el extremo opuesto, en el trofeo metálico y en el peluche, las diferencias llegan al 45 % y al 74 % respectivamente, lo que pone de manifiesto que la dificultad del LiDAR para tratar superficies reflectantes es alta. La conclusión que esto aporta a SO2.2 es más fuerte que un ranking entre tecnologías. Una PYME que digitalice su catálogo con NeRF mediante un smartphone gratuito obtiene modelos que un usuario en VR percibe como realistas y que matemáticamente difieren menos del 2 % de las medidas físicas reales del producto.

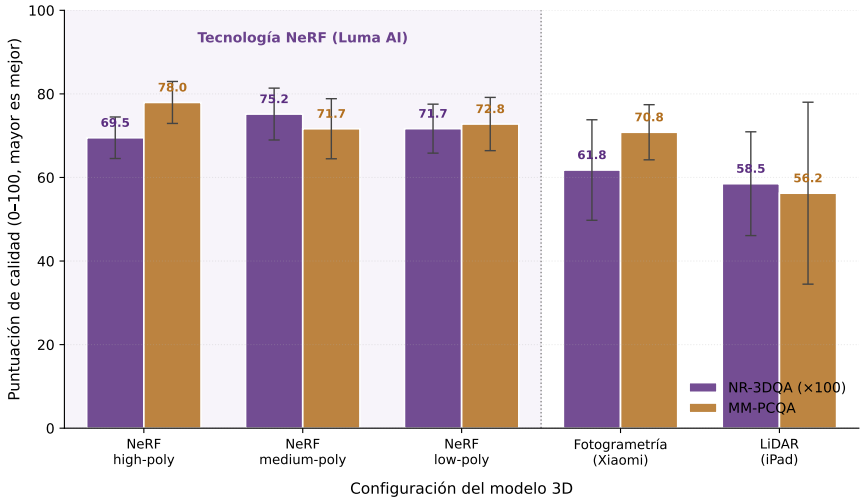


Figura 6.6: **Calidad visual por configuración del modelo.** Medias de NR-3DQA (reescalada a 0–100 para representación conjunta) y MM-PCQA sobre los cinco productos reales evaluados en el artículo I del compendio. Las barras de error reflejan la desviación estándar entre productos. El LiDAR es la peor opción en ambas métricas y muestra una variabilidad muy alta en MM-PCQA, reflejo de su fragilidad ante materiales complejos.

6.2.4 | Rendimiento en dispositivos VR comerciales (SO2.3)

La calidad visual es solo la mitad del problema. La otra mitad es si los modelos resultantes se ejecutan con fluidez en los dispositivos VR autónomos que una PYME puede esperar que sus clientes posean. El artículo I evalúa

tres de los headsets más extendidos del mercado: *Meta Quest 2*, *Meta Quest Pro* y *Meta Quest 3*. La instrumentación se realiza mediante OVR Metrics Tool, que registra hasta 88 métricas por segundo (FPS sostenidos, utilización de GPU y CPU, *stale frames*, *average prediction*, temperatura, etc.) durante sesiones de aproximadamente 60 segundos en las que el usuario interactúa con los productos (agarrarlos, examinarlos, moverlos) y se desplaza libremente por la escena. Esta carga de interacción simula el uso real en VR-shopping, no un renderizado pasivo. La métrica de referencia es el FPS sostenido, con el umbral de 60–72 FPS recomendado por Meta para evitar cybersickness.

Para evaluar la escalabilidad, el artículo I construyó 14 escenas con un número creciente de productos, alcanzando un máximo de 96 objetos en configuración low-poly, 24 en medium-poly y 12 en high-poly. Los resultados clarifican la naturaleza del compromiso entre calidad y rendimiento. Los modelos NeRF de alta resolución, con entre 1,2 y 1,8 millones de polígonos por objeto y texturas individuales de 4096×4096, son sencillamente inviables en modo standalone: 9,30 FPS en Quest 2, 12,40 en Quest Pro y 15,97 en Quest 3, incluso en escenas con pocos objetos. La causa es directa: el ancho de banda de la memoria de las GPU móviles integradas en los headsets autónomos no soporta el volumen poligonal y la resolución de textura de un modelo high-poly.

La transición al low-poly que la propia Luma AI genera por defecto (decimación automática a unos 25 000 polígonos por objeto) cambia el panorama. La misma escena con 72 objetos, que suma aproximadamente 1,79 millones de polígonos totales, sostiene 56,21 FPS en Quest 2, 60,16 en Quest Pro y 62,60 en Quest 3. La pérdida de calidad visual perceptual entre high-poly y low-poly es solo del 5 % según NR-3DQA y de un 3,4 % según CMDM, datos que el artículo I documenta de manera explícita. La conclusión práctica que responde a SO2.3 es directa: *low-poly es la configuración por defecto sensata para VR-shopping en dispositivos comerciales*. La calidad visual sigue siendo alta, los tiempos de carga se reducen entre 9 y 50 veces y el rendimiento entra en la zona de confort visual en los dos headsets más modernos. La Figura 6.7 (que se discute en detalle en la siguiente subsección al introducir el pipeline de optimización) muestra el comportamiento completo de los tres tipos de modelo frente a la carga poligonal y permite situar visualmente la frontera de viabilidad.

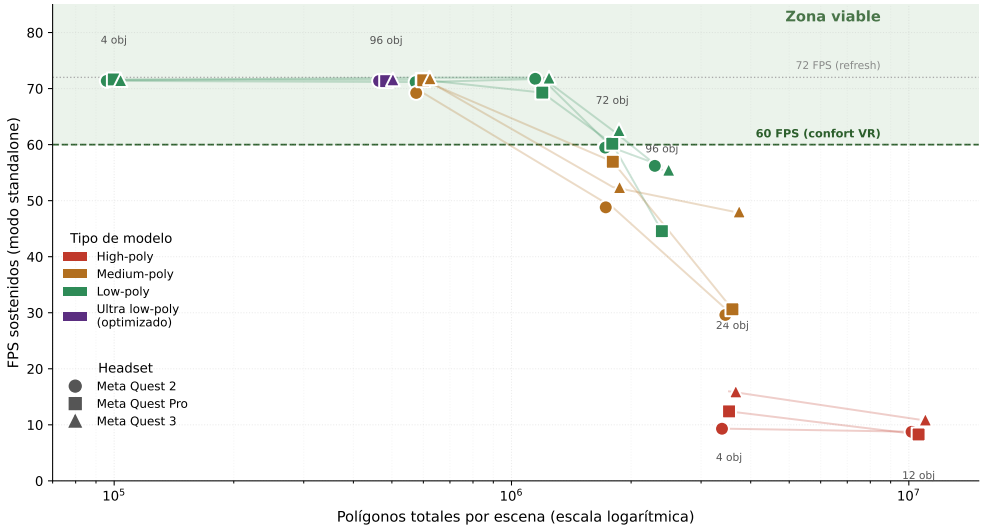


Figura 6.7: **Trayectoria de rendimiento por carga poligonal y headset.** FPS sostenidos en modo standalone sobre los tres dispositivos comerciales de Meta evaluados, agrupados por tipo de modelo (high-poly, medium-poly, low-poly y ultra low-poly post-optimización) y representados frente a la carga poligonal total de la escena en escala logarítmica. Los puntos rojos (high-poly) permanecen siempre por debajo de los 20 FPS, lo que los descarta para uso en VR-shopping standalone. Los puntos naranjas (medium-poly) sostienen los 60 FPS solo en escenas mínimas y caen rápidamente al añadir productos. Los puntos verdes (low-poly) se mantienen en la zona viable hasta unos 48 objetos en los tres headsets y empiezan a degradarse a partir de los 72. El punto morado (ultra low-poly tras optimización en Blender) sitúa una escena de 96 objetos en 484 K polígonos totales con los tres headsets superando los 71 FPS, lo que demuestra que el pipeline de optimización abre un espacio de viabilidad antes inalcanzable.

6.2.5 | Optimización de modelos 3D obtenidos (SO2.4)

Para escenas comerciales realistas con un centenar de productos o más, incluso el low-poly por defecto empieza a tensionar a las GPU móviles, sobre todo en Quest 2, que es el dispositivo con menor capacidad gráfica. El artículo I propone y valida un pipeline adicional de optimización con dos componentes: decimación dirigida que preserva polígonos en zonas de alto detalle visual y los reduce en zonas planas o de fondo, y compresión de texturas con reducción de resolución de las áreas no críticas. El resultado es lo que el artículo denomina modelos *ultra low-poly*.

Sobre escenas de 96 objetos con esta configuración (un 33 % más grandes que las que ya se sostenían con el low-poly por defecto), los tres headsets superan los 71 FPS sin caídas de GPU, CPU ni *stale frames*. Esta holgura es suficiente para soportar la interacción gestual del usuario sin comprometer la experiencia. Este pipeline responde a SO2.4 y cierra el bucle: la digitalización de bajo coste no es solo viable a nivel de modelo aislado, sino que escala a tamaños de tienda virtual realistas en hardware comercial.

6.2.6 | Posicionamiento frente a la evolución del campo y limitaciones reconocidas

Gaussian Splatting (3DGS) ha emergido como alternativa a NeRF con tiempos de entrenamiento órdenes de magnitud menores y rendimiento de renderizado superior, especialmente en hardware sin núcleos tensoriales dedicados. La uso de 3DGS en el módulo de digitalización de VR-ZOCO puede ser considerado a futuro, dado que el resto del sistema consume el modelo final y no la representación intermedia. Más recientemente, los modelos generativos 3D por difusión plantean un escenario futuro en el que la digitalización deje de ser estrictamente necesaria: una descripción textual del producto y un conjunto reducido de fotografías podrían bastar para generar un modelo plausible. La separación modular del artículo I permite incorporar estos avances cuando alcancen madurez suficiente, sustituyendo el módulo de digitalización sin tocar el resto del sistema.

Conviene reconocer abiertamente las limitaciones del estudio. El artículo I documenta familias de productos problemáticas para todas las tecnologías evaluadas, incluida la NeRF, como lo son los objetos con materiales transparentes o altamente reflectantes. Para estas categorías, los modelos resultantes presentan agujeros, distorsiones o falta de detalle. Es importante señalar que estas dificultades no son específicas del enfoque de bajo coste: persisten incluso con escáneres profesionales y constituyen un problema abierto en la disciplina del escaneo 3D. Una segunda limitación a reconocer es que el proceso de optimización a modelos ultra low-poly requiere todavía conocimiento técnico específico (decimación dirigida, compresión de texturas) que no es realista exigir directamente a una PYME del sector *retail*. Una tercera limitación es que el estudio no contempla la complejidad del entorno completo (iluminación

dinámica, sombras, materiales con *shaders* avanzados), lo que ofrecería un panorama más realista del rendimiento en una tienda virtual con escenografía cuidada.

6.2.7 | Síntesis de la respuesta a RQ2

La respuesta a RQ2 es afirmativa con tres matizaciones precisas. Primero, las tres tecnologías evaluadas no son equivalentes: NeRF mediante Luma AI supera consistentemente a la fotogrametría y al LiDAR tanto en calidad visual perceptual (mejoras del 17,78 % y 22,73 % en NR-3DQA respectivamente) como en error geométrico (1,4 % de error relativo frente al *ground truth*), y lo hace además con la herramienta más accesible de las tres (gratuita, basada en cloud, ejecutable desde un smartphone). Segundo, la configuración de alta resolución de NeRF es inviable en modo standalone debido a las restricciones de hardware de los headsets autónomos, por lo que la digitalización efectiva pasa por la versión low-poly que la propia herramienta genera por defecto, una elección que cuesta solo el 5 % de calidad perceptual y reduce el tamaño de los modelos en más de un orden de magnitud. Tercero, para escenas comerciales de tamaño realista (96 productos o más) un pipeline de optimización adicional consigue sostener tasas por encima de los 71 FPS en los tres headsets comerciales evaluados, lo que cubre con holgura el umbral de confort visual en VR. La digitalización de productos físicos por parte de una PYME desde un smartphone comercial es, por tanto, técnicamente viable, satisface los requisitos de rendimiento de los dispositivos VR autónomos del mercado actual y produce experiencias percibidas como realistas por el usuario. Esta conclusión es la que después habilita el despliegue del artículo IV del compendio en un escenario comercial real, donde los productos digitalizados con NeRF se emplean en la escena de tienda virtual sin que la calidad o el rendimiento se hayan reportado como problemas durante el estudio de campo.

6.3 | RQ3: mecanismos de interacción y navegación

> ***Pregunta de investigación 3.*** *«¿Cómo debe ser la interacción y navegación en entornos VR-shopping haciendo uso de hand-tracking para proporcionar una experiencia cercana a la compra en entornos físicos?»*

La tabla 6.3 resume las contribuciones por publicación para responder esta pregunta de investigación.

Tabla 6.3: **Mapeo de contribuciones a RQ3.** El artículo II del compendio (VRSI) es la pieza principal e implementa los tres subobjetivos. La publicación de congreso *Exploring Interaction Mechanisms* aporta la validación empírica con usuarios reales sobre interacciones (SO3.1) y estructura espacial (SO3.3). El artículo IV del compendio despliega las interacciones en condiciones de campo. La arquitectura VR-ZOCO declara la capa a nivel funcional, y el artículo III valida indirectamente la modularidad de SO3.1 al extender las interacciones para accesibilidad.

Publicación	Venue / año	Tipo	O3		
			SO3.1	SO3.2	SO3.3
Virtual Reality Shopping-Insights (VRSI)	SoftwareX, 2024	AC	●	●	●
Exploring Interaction Mechanisms	ICEIS, 2024	CI	●	●	●
Gamifying VR-shopping (estudio de campo)	Virtual Reality, 2026	AC	●	●	○
VR-ZOCO architectural proposal	Appl. Syst. Innov., 2024	AR	●	●	●
Enhancing accessibility	IEEE Access, 2025	AC	●	○	○

Contribución: ● principal ● parcial ○ no contribuye **Tipo:** AC compendio AR revista CL cap. libro CI congr. int. CN congr. nac.

6.3.1 | Interacción natural en dispositivos de VR

Hasta la maduración del hand-tracking en los dispositivos VR comerciales (Meta Quest 2, Pro y 3), la interacción con productos virtuales se realizaba principalmente mediante controladores físicos. Esta limitación rompía la analogía con la compra en una tienda real, donde el comprador agarra, modifica, gira, acerca y deja productos con sus propias manos. Los trabajos previos sobre

VR-shopping [46, 48, 49] ya advertían que la naturalidad de la interacción es un factor determinante de la experiencia de compra inmersiva. A esto se suma una preocupación adicional sobre el desplazamiento continuo en entornos amplios, que la literatura sobre *cybersickness* identifica como fuente recurrente de malestar para sesiones prolongadas [44]. Para poder ofrecer una flexibildiad máxima, la situación idónea es no depender de grandes superficies físicas para poder navegar cómodamente por el entorno virtual creado.

La pregunta que se desprende de este contexto es doble: qué conjunto de mecanismos de interacción y navegación son adecuados para un uso cómodo y continuado en VR-shopping, y cómo encapsularlos en un marco de trabajo reutilizable. La capa de interacción del usuario de la arquitectura VR-ZOCO [5] declara este conjunto a nivel funcional, siendo algunos de estos FE1: navegación por categorías; FE2: vista 3D del producto con manipulación; FE10: adaptabilidad para diversidad funcional, pero la declaración arquitectónica no garantiza nada sobre la efectividad de los gestos concretos. La operacionalización experimental de esa capa, junto con su validación con usuarios, es lo que abordan las publicaciones que se discuten a continuación.

6.3.2 | VRSI como marco de trabajo unificado

El artículo II del compendio [2] introduce VRSI (Virtual Reality Shopping-Insights), un framework de código abierto publicado en SoftwareX bajo licencia MIT. SoftwareX es una revista JCR Q2 cuyo proceso editorial se basa en *Original Software Publication*: lo que se revisa por pares no es solo el artículo descriptivo sino el propio código y su capacidad de reproducción por parte de terceros. La validación, por tanto, reside en su utilidad demostrable como marco de trabajo, no en una hipótesis con sujetos.

VRSI se organiza en tres componentes con responsabilidades disjuntas. El primero, Retail Environment Setup, proporciona prefabs reutilizables para construir tiendas virtuales y los mecanismos de interacción y navegación. El segundo, Data Collection, captura de manera no invasiva las señales del usuario (hand-tracking, eye-tracking, head-tracking, posición espacial) y las exporta en formato estructurado JSON y CSV. El tercero, Behaviour Statistics, aplica análisis post hoc sobre esos datos y produce reportes en PDF con métricas de comportamiento. Para RQ3, el primer componente es el que da la respuesta

directa, mientras que el segundo y el tercero alimentan la discusión de RQ5 (Sección 6.5).

6.3.3 | Interacciones con productos (SO3.1)

VRSI implementa cinco interacciones principales con los productos virtuales. La de agarre puede ser unimanual o bimanual, esta última útil para objetos grandes donde la experiencia física natural requiere las dos manos, o para reescalar el objeto dinámicamente. El examen libre permite al usuario rotar el producto en cualquier eje una vez agarrado, replicando el gesto de mirar un objeto desde distintos ángulos antes de comprarlo. La acción de soltar libera el agarre al abrir la mano. La inclusión en el carrito de compra se realiza mediante un gesto de aproximación al carrito flotante o mediante una pulsación sobre su botón virtual, ofreciendo así dos vías alternativas que el usuario puede usar según su preferencia. Un ejemplo de las interacciones puede verse en la Fig. 6.8.

Cada interacción incorpora retroalimentación visual mediante outlines del producto: un contorno luminoso que cambia de color según el estado (idle, hovered, grabbed). Este feedback compensa la ausencia de retroalimentación háptica en los headsets autónomos comerciales y es coherente con los principios de naturalidad y reversibilidad que la literatura sobre interacción gestual recomienda. La elección de gestos sigue además una lógica de simetría: cada acción puede deshacerse con un movimiento equivalente al original (agarrar/soltar, escalar más/escalar menos), lo que reduce la carga cognitiva del aprendizaje y permite al usuario explorar la interfaz sin miedo a errores irreversibles.

6.3.4 | Navegación (SO3.2)

El desplazamiento dentro de la tienda virtual combina dos mecanismos complementarios. La caminata natural cuando el espacio físico del usuario lo permite, y la teletransportación por gesto cuando no. El gesto de teletransportación se activa al apuntar con la palma de la mano hacia el suelo en la posición destino, lo que dibuja un arco con un punto de aterrizaje visible. Al cerrar la mano o tras un breve mantenimiento del gesto, el usuario aparece en el punto seleccionado. Esta elección de gesto se alinea con la convención dominante en los SDK comerciales (Meta SDK, OpenXR) y evita el cybersickness asociado al

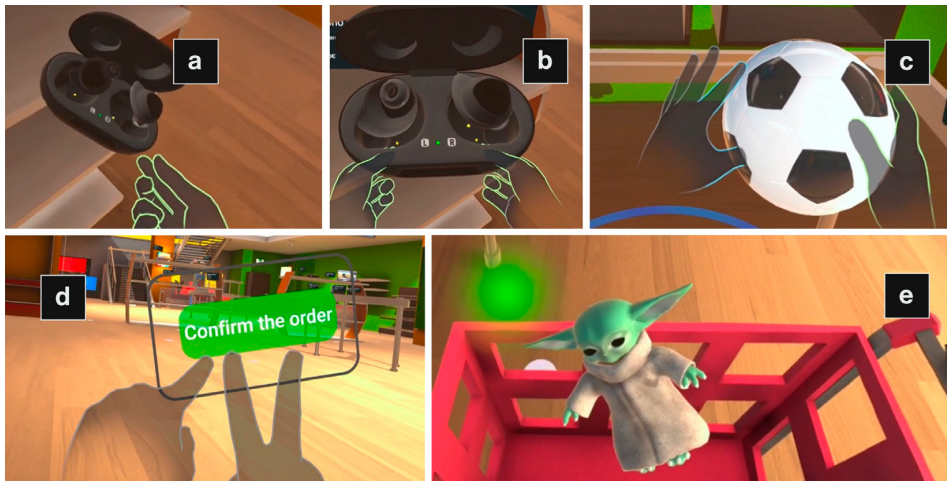


Figura 6.8: **Catálogo de interacciones implementadas en VRSI.** (a) Inspección unimanual de un producto pequeño (auriculares) tras un agarre con pinza. (b) Inspección bimanual del mismo producto, gesto que permite también el rescalado al separar las dos manos. (c) Agarre bimanual de un producto grande (balón), donde el modelo de dos manos refleja el gesto natural del comprador físico. (d) Interacción con un elemento de UI gestual (botón *Confirm the order*) que cierra el flujo de compra. (e) Adición de un producto al carrito mediante aproximación, con retroalimentación visual en forma de partículas verdes que confirman la acción. Capturas extraídas del artículo II del compendio [2].

desplazamiento continuo, una preocupación bien documentada en la literatura [44].

A la teletransportación se añade la rotación del cuerpo virtual mediante un gesto lateral con la mano abierta, útil cuando el usuario necesita girarse en un espacio físico pequeño. El conjunto de ambos mecanismos permite que un usuario equipado con un Quest standalone recorra una tienda virtual de cualquier tamaño desde una sala doméstica de pocos metros cuadrados, requisito habitual del público objetivo que VR-ZOCO contempla en sus escenarios de uso. La validez de esta decisión queda confirmada en los resultados empíricos que se discuten en la subsección de validación.

6.3.5 | Estructura espacial del entorno (SO3.3)

El tercer componente arquitectónico que VRSI proporciona es un sistema configurable para definir la estructura espacial del entorno de compra. Las es-

tanterías están organizadas en tres alturas de exposición (Top, Middle, Bottom) que reflejan la convención del retail físico, donde la altura del producto en el lineal tiene impacto demostrado sobre la atención del comprador. Sobre estas estanterías se superpone un sistema de áreas de interés Area of Interest (AOI) que permite asociar cada producto y cada sección a una región del espacio.

A este nivel se añade un sistema de zonas de interés Zone of Interest (ZOI) que segmenta el plano del suelo en cuatro anillos concéntricos respecto al usuario en cada momento (Start, Adjacent, Near, Far). La combinación de AOI y ZOI permite calcular en post hoc métricas como tiempo de presencia por sección, transiciones entre zonas y mapas de calor volumétricos de atención y trayectoria, datos que alimentan las contribuciones a RQ5. El sistema completo es configurable mediante archivos JSON, lo que satisface el requisito de SO3.3 de ser reutilizable para diferentes layouts de tienda sin necesidad de modificar el código fuente. Este punto es relevante para una PYME que quiera adoptar el framework: cambia el layout, no la implementación.

6.3.6 | Validación empírica de los mecanismos

Esta validación se aborda en [92], que reporta un experimento con 10 participantes (8 hombres y 2 mujeres, edades entre 22 y 28 años, más 2 voluntarios de 58) realizado sobre un Meta Quest Pro conectado por cable Oculus Link de 5 m a un MSI Vector GP66 con i7-12700H, 32 GB de RAM y GeForce RTX 3060. El estudio se ejecutó, iniciando con un cuestionario inicial, sobre dos escenas Unity sucesivas (entorno pequeño con 5 productos sobre un mostrador, entorno grande con 24 productos repartidos en secciones de moda, tecnología, decoración, juguetes y alimentación) con una duración media de 28,5 minutos por sesión. Tras la sesión, los participantes respondieron a una batería de cuestionarios estandarizados: UEQ-S (User Experience Questionnaire), SUS (System Usability Scale) y MSAQ (Motion Sickness Assessment Questionnaire), todos en escala Likert de 1 a 5.

Los resultados confirmaron la viabilidad del enfoque (Figura 6.9). La preferencia por las interacciones manuales sobre los controladores es clara y consistente: M=3,5 (SD=0,81) para las manos frente a M=2,6 (SD=0,49) para los controladores, una diferencia de casi un punto en una escala de cinco. La fluidez y naturalidad de las interacciones es positiva (M=3,5, SD=0,92), igual

que la facilidad de teletransportación (M=3,7, SD=0,64) y la usabilidad del carrito de compra, que alcanza M=3,6 tanto en su versión de interacción manual como con botón gestual. La inmersión percibida es la dimensión mejor evaluada (M=4,1, SD=0,54). Las puntuaciones de motion sickness son bajas en las tres dimensiones medidas: malestar general M=1,6, fatiga M=1,9 y mareo M=1,5 (todas en escala 1–5 donde el valor bajo es el deseable), lo que respalda la decisión de combinar caminata natural con teletransportación gestual frente al desplazamiento continuo.

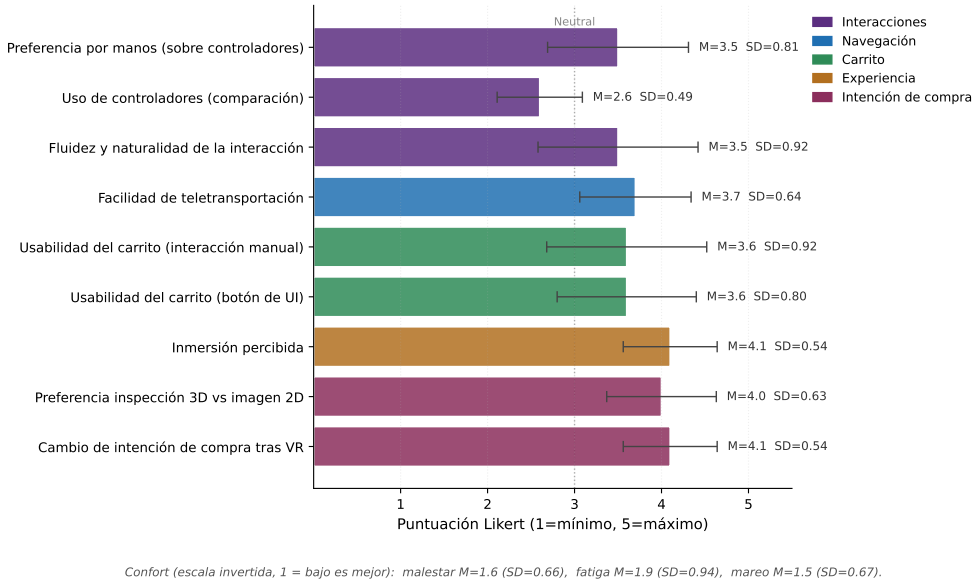


Figura 6.9: **Validación empírica de los mecanismos de interacción y navegación de VRSI.** Puntuaciones medias y desviaciones estándar de las dimensiones evaluadas en el estudio de [92] con 10 participantes sobre Meta Quest Pro. Los colores agrupan las dimensiones en cinco categorías. Las interacciones con las manos se prefieren claramente a los controladores (3,5 frente a 2,6), la teletransportación se evalúa como fácil de usar (3,7) y la inmersión percibida es la dimensión mejor puntuada (4,1). Las puntuaciones de motion sickness (no representadas en el gráfico al usar escala invertida) son consistentemente bajas: malestar M=1,6, fatiga M=1,9, mareo M=1,5.

Dos hallazgos secundarios refuerzan el resto de la discusión. El primero es la preferencia por entornos grandes sobre pequeños (M=1,7 en una escala donde 1 indica preferencia por grande), un dato que apoya la elección de estructura espacial multisección de SO3.3. El segundo es que cuando los participantes

seleccionaron los cinco productos más realistas del experimento, el 50 % de las respuestas correspondieron a modelos generados con Luma AI (NeRF). Este dato conecta directamente con la conclusión de RQ2 (Sección 6.2) sobre la superioridad de NeRF en calidad visual percibida y refuerza la coherencia transversal de las decisiones técnicas tomadas a lo largo del compendio.

A esta validación controlada se añade una segunda capa de evidencia, esta vez en condiciones de campo. El artículo IV del compendio [4] despliega las interacciones VRSI en un evento comercial real (Serendipia, Ciudad Real) con 48 participantes. Aunque el objetivo central del artículo IV no es RQ3 sino RQ6 (impacto en la compra real), el simple hecho de que VRSI haya sostenido un despliegue de N=48 en una feria abierta al público sin que se reportaran fallos significativos o rechazos de los usuarios constituye una validación operativa de los mecanismos aquí discutidos.

Una tercera capa de evidencia, indirecta pero significativa, proviene del artículo III del compendio [3], que extiende las interacciones de VRSI para usuarios con limitaciones motoras. Lo relevante para RQ3, más allá de lo que se discute en detalle en RQ4 (Sección 6.4), es que esa extensión fue posible sin reescribir el framework. Los gestos adaptados se insertan como variantes paralelas en el mismo componente Retail Environment Setup, sin tocar el núcleo de los mecanismos de SO3.1.

6.3.7 | Limitaciones reconocidas

Conviene reconocer las limitaciones de esta validación. La principal es el tamaño muestral del estudio controlado: 10 participantes son insuficientes para reclamar resultados con potencia estadística robusta, y la mayoría son estudiantes universitarios jóvenes con poca experiencia previa en VR. El despliegue del artículo IV con N=48 en una feria, amplía la base poblacional, pero introduce otras variables como la autoselección de público interesado y la instrumentación de campo más limitada que la disponible en laboratorio. Una segunda limitación es la dependencia del SDK de Meta para el hand-tracking, lo que constriñe la portabilidad a dispositivos de la familia Quest. La adaptación a otros sistemas (Pico, HTC Vive XR, Apple Vision Pro) requeriría reescribir el componente de captura, además de disponer de estos, teniendo en cuenta las restricciones de financiación existentes en el mundo de la investigación. Una

tercera limitación, ya identificada en el artículo II, es la ausencia de una API externa que permita integrar VRSI con servicios complementarios (sistemas de recomendación, pasarelas de pago, gestión de inventario).

6.3.8 | Síntesis de la respuesta a RQ3

La respuesta a RQ3 se construye sobre tres niveles complementarios. En el nivel de diseño, las cinco interacciones con productos (agarre uni/bimanual, examen libre, rescalado, soltar, añadir al carrito), los dos mecanismos de navegación (teletransportación gestual y rotación corporal) y las tres dimensiones de la estructura espacial (estanterías configurables, áreas de interés y zonas de interés) constituyen un conjunto suficiente para sostener una experiencia de VR-shopping cercana a la compra física. En el nivel de infraestructura, VRSI los empaqueta como un framework open-source publicado en SoftwareX bajo licencia MIT, reutilizable y extensible por terceros y reproducible a partir del código publicado. En el nivel empírico, los resultados de [92] con N=10 (Likert promedio entre 3,5 y 4,1 en las dimensiones positivas, motion sickness consistentemente por debajo de 2) y la robustez operativa demostrada en el estudio de campo del artículo IV con N=48 sostienen empíricamente que estas interacciones funcionan con usuarios reales. La triangulación entre diseño formalizado, infraestructura abierta y validación empírica en dos contextos distintos responde a RQ3 de manera sustantiva, dejando como continuación natural la ampliación de la base muestral y la integración con extensiones recientes del catálogo gestual de los SDK comerciales.

6.4 | RQ4: adaptabilidad de mecanismos de interacción

> ***Pregunta de investigación 4.*** *«¿Pueden este tipo de entornos inmersivos, con interacciones basadas en hand-tracking, ser accesibles a todo tipo de personas, especialmente a aquellas con limitaciones motoras en miembros superiores?»*

La tabla 6.4 resume las contribuciones por publicación para responder esta pregunta de investigación.

Tabla 6.4: **Mapeo de contribuciones a RQ4.** El artículo III del compendio es la pieza central: cubre los cuatro subobjetivos con un caso de estudio en el Hospital Nacional de Parapléjicos de Toledo (HNPT) con 10 pacientes con lesión medular y 2 sujetos de control sanos. El artículo presentado en el congreso CEIG en 2024 [97] sirve de antecedente metodológico: introduce el conjunto de mecanismos adaptados y realiza una validación preliminar con sanos y un voluntario con secuelas de ictus. La arquitectura VR-ZOCO declara la accesibilidad como funcionalidad nativa.

Publicación	Venue / año	Tipo	O4			
			SO4.1	SO4.2	SO4.3	SO4.4
Enhancing Hand Interactions and Accessibility	IEEE Access, 2025	AC	●	●	●	●
Adaptation of Interaction Mechanisms	CEIG, 2024	CN	●	●	○	○
VR-ZOCO architectural proposal	Appl. Syst. Innov., 2024	AR	●	●	○	○

Contribución: ● principal ● parcial ○ no contribuye **Tipo:** AC compendio AR revista CL cap. libro CI congr. int. CN congr. nac.

6.4.1 | La accesibilidad como problema técnico

Los SDK disponibles para desarrollar aplicaciones para los headsets VR comerciales asumen por defecto que el usuario es capaz de realizar gestos finos: aproximar pulgar e índice para realizar el denominado *pinch*, mantener una pose estable de la palma para invocar la teletransportación, realizar gestos con los dedos de la mano que suponen mantener una pose determinada durante cierto tiempo o cerrar la mano con fuerza suficiente para activar el agarre. Estas asunciones excluyen a una población de usuarios con limitaciones motoras en miembros superiores: pacientes con lesión medular, supervivientes de ictus, personas con esclerosis múltiple, parálisis cerebral o secuelas postquirúrgicas que afectan a la fuerza, la precisión o el rango de movimiento de las manos. Las interacciones de VRSI discutidas anteriormente (Sección 6.3) heredan estas limitaciones por defecto al apoyarse sobre el SDK de Meta. La arquitectura VR-ZOCO reconoce explícitamente esta brecha y declara la funcionalidad FE10 (Virtual Reality Environment Adaptability) como elemento nativo del sistema, pero esa declaración arquitectónica solo se materializa al diseñar y validar mecanismos concretos con personas con limitaciones reales. Eso es lo que las dos publicaciones que se discuten en esta sección han abordado.

Más allá del enfoque técnico, hay un componente ético: en un hipotético

escenario futuro donde el comercio se desplaza progresivamente a entornos inmersivos, dejar la accesibilidad fuera del diseño consolida una nueva forma de exclusión digital.

6.4.2 | Análisis de las limitaciones motoras y diseño de los mecanismos adaptados (SO4.1 y SO4.2)

El primer paso del trabajo, abordado en colaboración directa con la Unidad de Biomecánica del HNPT, es identificar qué gestos de los SDK estándar resultan inviables para personas con lesión medular cervical o dorsal alta. La caracterización clínica destaca tres limitaciones recurrentes: la espasticidad (contracciones involuntarias que impiden mantener una pose estable), la atrofia muscular en flexores y extensores del antebrazo (que reduce la fuerza necesaria para cerrar la mano contra resistencia) y la deformidad conocida como mano en garra (donde los dedos quedan en flexión permanente y el usuario no puede extenderlos de manera voluntaria). Estas tres limitaciones afectan de manera directa a las tres operaciones nucleares de hand-interaction (interacción con las manos) estándar: el *pinch*, el cierre del puño y la apertura completa de la palma. El análisis cumple SO4.1 y se realizó dentro de los dos artículos de RQ4: en [97] a nivel preliminar y en [3] con el detalle clínico completo sobre una muestra con pacientes.

A partir de este análisis, el artículo de CEIG 2024 [97] introduce el conjunto de cinco mecanismos adaptados que dan respuesta a SO4.2. La razón de adaptar los mecanismos de interacción nace de la imposibilidad por personas con movilidad reducida en miembros superiores de realizar los gestos, movimientos o ejercer la fuerza necesaria para que las cámaras de las gafas de VR actuales puedan detectarlos, imposibilitando de forma completa su interacción con las aplicaciones. Los controladores no son una alternativa viable, pues su uso requiere una forma específica y fuerza constante para poder presionar los botones de los mismos. Las adaptaciones diseñadas superan estas dificultades mediante la eliminación de gestos complicados, adaptando los umbrales de fuerza ejercida o automatizando interacciones mediante una intención clara del usuario (si el usuario quiere agarrar un objeto pero no puede ejercer la fuerza necesaria, que esté un tiempo la mano con la que lo quiere agarrar en el volumen del objeto). La Fig. 6.10 muestra la complejidad del gesto por

defecto de teletransporte, indicando los distintos ángulos ejes tridimensionales en los que las Meta Quest se basa para reconocer acciones determinadas con las manos.

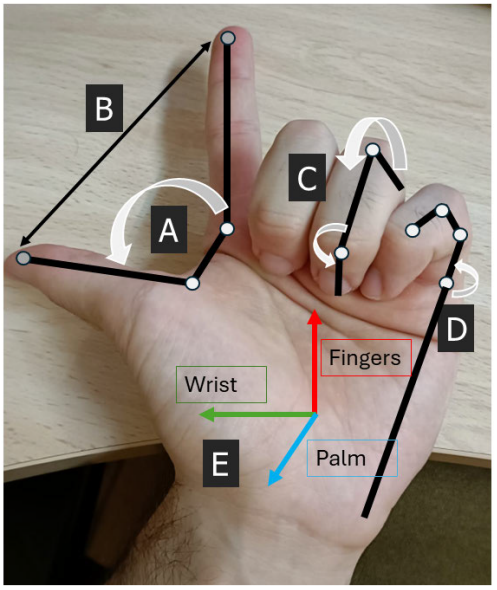


Figura 6.10: **Marco anatómico y geométrico empleado en el diseño de los mecanismos adaptados de hand-tracking.** Las etiquetas A y B identifican los puntos del pulgar e índice sobre los que se calcula el ángulo del gesto de pinza; C y D, los puntos de flexión de los dedos relevantes para el agarre adaptado y la detección de la postura de la mano; E identifica la articulación de muñeca cuya rotación activa los mecanismos de navegación. Los ejes coloreados Fingers (dedos), Wrist (muñeca) y Palm (palma) corresponden al sistema de coordenadas local sobre el que se evalúan los umbrales relajados que hacen viables las interacciones para personas con espasticidad, atrofia muscular o deformidad en mano en garra. Adaptado del artículo III del compendio [3].

La adaptación se diseña con dos restricciones explícitas: ser lo más genérica posible para cubrir un espectro amplio de grados de lesión, y mantener retroalimentación visual constante para que el usuario tenga siempre una señal de progreso. El primer mecanismo adapta la **teletransportación**: en lugar de exigir un *pinch* para confirmar el destino, basta con apuntar con la palma hacia arriba durante 1,3 segundos sobre marcadores circulares que el entorno coloca en el suelo. Un círculo de carga visual rellena de forma progresiva el tiempo restante. El segundo adapta la **rotación corporal**: cada mano se asocia

a una flecha (izquierda o derecha) que rota el cuerpo virtual 45° al detectar simplemente un movimiento del antebrazo o de la muñeca en la dirección señalada, sin requerir gestos finos de los dedos. El tercero, el **agarre adaptado**, modifica la API de Meta para reducir los umbrales de fuerza necesarios para iniciar y mantener un *grasp*, e incorpora un script de C# adicional que activa un *auto-grasp* cuando la mano del usuario permanece sobre la superficie de un objeto durante 1 segundo. El cuarto, el **agarre a distancia**, permite traer objetos lejanos hasta la mano del usuario apuntando con un *raycast* originado en la articulación Metacarpophalangeal/Metacarpofalángica (MCP) del dedo índice (no en la punta del dedo, que requeriría mayor movilidad) durante 1,5 segundos. El quinto adapta la **liberación del objeto**: en lugar de exigir abrir la mano por completo (gesto inviable para mano en garra), el usuario apunta durante 2 segundos a una papelera de reciclaje y el objeto vuelve por sí solo a su posición original. Las dos vías para añadir productos al carrito (aproximación con partículas verdes de confirmación, o apuntado al carrito durante 2 segundos) completan el conjunto. La validación preliminar en CEIG 2024 [97] con 4 estudiantes sanos (instruidos para no usar sus dedos) más un voluntario con secuelas de ictus confirmó la viabilidad del diseño y motivó el paso al estudio formal en HNPT que constituye el núcleo del artículo III.

6.4.3 | Métricas derivadas del hand-tracking: DI y ED (SO4.3)

Un aspecto distintivo del artículo III [3] es que no se limita a reportar tasas de completación y cuestionarios subjetivos, que es lo habitual en la literatura sobre accesibilidad en VR. Define dos métricas objetivas calculadas íntegramente a partir de las trazas de hand-tracking grabadas durante las sesiones, lo que permite cuantificar el esfuerzo del usuario sin recurrir a sensores adicionales ni a percepción subjetiva.

La primera métrica es el *Difficulty Index* (DI), que captura la dificultad efectiva con la que el usuario realiza cada tarea a partir de cuatro componentes normalizadas en el intervalo $[0, 1]$: el volumen del *bounding box* que envuelve el rango de movimiento de las manos (NV_{BX}, indicador del esfuerzo espacial), la distancia total euclidiana recorrida por las manos (N_{TD}), la eficiencia del movimiento entendida como la razón entre la distancia recta entre dos puntos y la distancia efectivamente recorrida (N_{ME}), y el número de movimientos por

segundo (N_{MPS}, indicador de inestabilidad del agarre). El DI resultante es la media aritmética de los cuatro componentes:

$$DI = \frac{NV_{BX} + N_{TD} + N_{ME} + N_{MPS}}{4} \tag{6.1}$$

Las señales en crudo del hand-tracking se preprocesan con un filtro de Kalman para suavizar las fluctuaciones de alta frecuencia propias del temblor o de imprecisiones de las cámaras del headset, conservando solo los desplazamientos significativos. La normalización por máximo absoluto asegura que ningún componente domine sobre los demás.

La segunda métrica es el *Effort Degree* (ED), que integra el DI con dos magnitudes adicionales de rendimiento de tarea (número de intentos y tiempo empleado), penaliza la no completación de la tarea, y produce un valor agregado de esfuerzo por usuario y tarea:

$$ED = \left(\frac{Attempts_T}{Attempts_{max}} \times \frac{Time_{Ti}}{Time_{max}} \right) \times (1 + DI_{U,env}) + 1 \times (NotCompleted(T, U, i)) \tag{6.2}$$

Donde el término final añade una unidad cuando la tarea T no se completa por el usuario U usando las interacciones i. Esta penalización resulta crítica para diferenciar entre tareas técnicamente completadas con esfuerzo alto y tareas abandonadas. Ambas métricas constituyen una contribución metodológica que va más allá del trabajo descriptivo habitual sobre accesibilidad en VR: permiten comparaciones cuantitativas pareadas entre condiciones experimentales sin depender de cuestionarios subjetivos. SO4.3 queda cubierto íntegramente por el artículo III, sin precedente en las publicaciones previas del trabajo.

6.4.4 | Caso de estudio en el HNPT (SO4.4)

El núcleo experimental del artículo III es un caso de estudio realizado en el Hospital Nacional de Parapléjicos de Toledo con 12 participantes: 10 pacientes con lesión medular (U1-U10, edades entre 14 y 54 años, media 32,0 $\pm$ 13,2; todos varones; nivel de lesión entre C2 y D8; escala AIS entre A y D; UEMS del brazo tratado de 14,14 $\pm$ 8,62 sobre 25 máximo; SCIM de 48,0 $\pm$ 25,3 sobre 100) y 2 sujetos sanos de control reclutados entre el personal del propio HNPT (U11, ingeniero biomecánico de 44 años; U12, terapeuta de 38 años).

El protocolo experimental sigue un diseño *within-subjects*: cada participante completa primero las diez tareas del estudio en el entorno no adaptado (con un breve entrenamiento previo de 1 a 2 minutos) y a continuación las mismas diez tareas en el entorno adaptado (sin entrenamiento previo, para evaluar la facilidad de aprendizaje de los mecanismos nuevos).

Las diez tareas representan acciones de un escenario de VR-shopping: navegar a secciones concretas mediante teletransportación, girar el cuerpo virtual para alinearse con productos, agarrar objetos de distintas alturas y materiales, devolverlos en caso necesario, añadirlos al carrito, y una tarea final de exploración libre. La Figura 6.11 resume el comportamiento cuantitativo por usuario en las dos dimensiones críticas: la tasa de completación de tareas y el Difficulty Index.

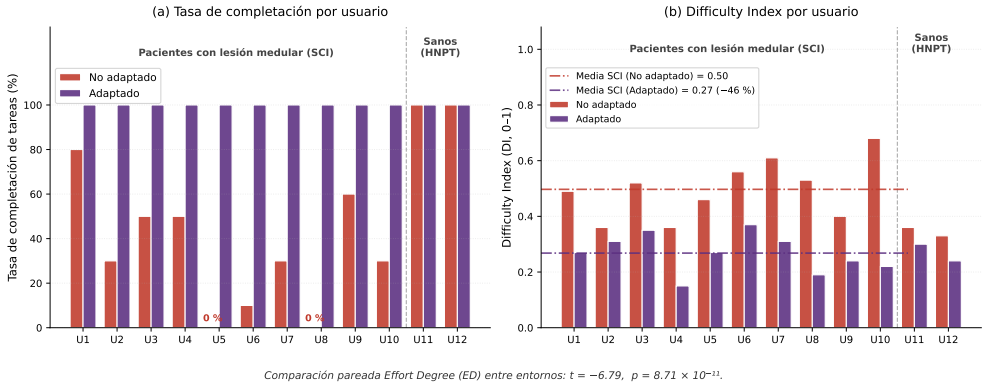


Figura 6.11: **Resultados cuantitativos del caso de estudio HNPT.** (a) Tasa de completación de tareas por usuario en el entorno no adaptado (rojo) y en el adaptado (morado). Todos los pacientes con lesión medular (U1–U10) alcanzan el 100 % de completación en el entorno adaptado, frente a tasas que oscilan entre el 0 % (U5 y U8, los casos más severos) y el 80 % (U1) en el entorno no adaptado. Los dos controles sanos (U11, U12) completan el 100 % en ambos entornos. (b) Difficulty Index por usuario y entorno. Las líneas discontinuas marcan las medias de los pacientes con lesión medular (0,50 no adaptado, 0,27 adaptado, reducción del 46 %). La comparación pareada Effort Degree entre los dos entornos arroja $t = -6{,}79$ y $p = 8{,}71 \times 10^{-11}$, indicador de diferencia estadísticamente significativa. Datos extraídos de las Tablas 4 y 5 del artículo III [3].

La tasa de completación ilustra el resultado más contundente del estudio: los diez pacientes con lesión medular completaron el 100 % de las tareas en el

entorno adaptado, mientras que sin adaptaciones la tasa media de completación fue del 33 % y dos de ellos (U5 y U8) no consiguieron completar ni una sola tarea por imposibilidad física. El artículo III subraya que esta diferencia no es solo cuantitativa sino cualitativa: las adaptaciones permiten realizar tareas que sin ellas son sencillamente imposibles. Para los sujetos de control sanos, ambos entornos producen 100 % de completación, lo que confirma que los mecanismos adaptados no introducen fricción que impida su uso por una población general (los tiempos sí aumentan ligeramente, principalmente por los segundos de espera de cada gesto adaptado y por la costumbre de interactuar de forma natural).

El Difficulty Index ofrece una lectura complementaria. La media del DI sobre los pacientes con lesión medular pasa de 0,497 en el entorno no adaptado a 0,268 en el adaptado, una reducción del 46,1 %. La comparación pareada del Effort Degree entre los dos entornos a lo largo de los doce participantes y las diez tareas arroja $t = -6{,}79$ con $p = 8{,}71 \times 10^{-11}$, valor muy por debajo del umbral $\alpha = 0{,}05$ y compatible incluso con criterios mucho más estrictos. La diferencia en esfuerzo entre los dos entornos queda, por tanto, contrastada estadísticamente. La Fig. 6.12 sirve de ejemplo mostrando los volúmenes de los movimientos con los miembros superiores realizados por algunos de los participantes en los entornos con interacciones adaptadas y no adaptadas.

El análisis de correlaciones de Spearman entre DI y métricas de rendimiento aporta una capa adicional de interpretación. En el entorno no adaptado, el DI correlaciona de forma significativa y de manera negativa con la tasa de completación ($\rho = -0{,}578$, $p = 0{,}049$) y de manera positiva con el tiempo de completación ($\rho = 0{,}711$, $p = 0{,}009$): los pacientes que perciben más dificultad completan menos tareas y emplean más tiempo en las que sí completan. En el entorno adaptado, la correlación con la tasa de completación no se puede computar porque todos los usuarios alcanzan el 100 % (no hay varianza), y la correlación con el tiempo de completación es estadísticamente no significativa ($\rho = -0{,}084$, $p = 0{,}794$). Esta disociación es esperable de unas adaptaciones efectivas: el esfuerzo subjetivo deja de determinar el resultado de la tarea.

Más allá del caso de uso específico del VR-shopping, estas interacciones adaptadas son suficientemente generales como para ser usadas en otras aplicaciones de VR que hagan uso de interacción con manos, independientemente del campo de aplicación, pues el conjunto es modificable y ampliable.

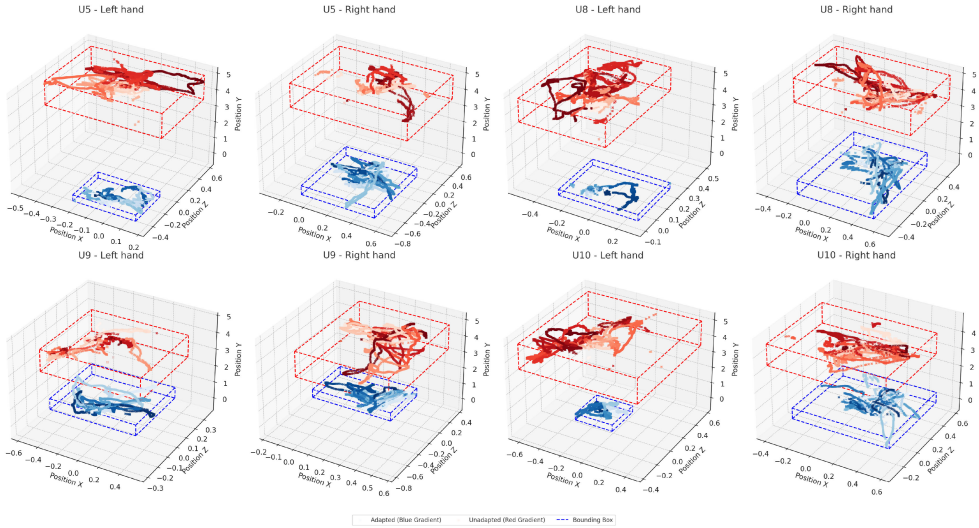


Figura 6.12: **Visualización tridimensional del movimiento de manos en cuatro pacientes con lesión medular.** Trayectorias completas de las manos izquierda y derecha de los usuarios U5, U8, U9 y U10 durante las diez tareas del caso de estudio en el HNPT, separadas por condición experimental: gradiente rojo para el entorno no adaptado y gradiente azul para el adaptado. Las cajas envolventes (*bounding boxes*) discontinuas marcan el volumen mínimo que contiene todas las posiciones registradas en cada condición. La reducción visible del volumen del entorno no adaptado al adaptado se corresponde con el componente NV_{BX} del Difficulty Index. U5 y U8 son los dos casos más severos del estudio, en los que la condición no adaptada arrojó tasas de completion del 0 %. Imagen extraída del artículo III del compendio [3].

6.4.5 | Validación subjetiva: NASA-TLX, SUS, UEQ y MSAQ

Junto con las métricas objetivas, el artículo III recoge un cuestionario post-sesión que combina preguntas de cuatro instrumentos estandarizados: NASA-TLX para carga de trabajo percibida, SUS para usabilidad, UEQ para experiencia de usuario y MSAQ para motion sickness. Cada participante respondió las mismas preguntas para las dos condiciones (adaptada y no adaptada) sobre una escala Likert de 1 a 6, lo que permite una comparación pareada. Conviene tener en cuenta que la polaridad de la escala depende de la pregunta: en algunas un valor bajo es lo deseable (esfuerzo, dificultad, frustración) y en otras lo es un valor alto (facilidad de uso, seguridad percibida).

Las diferencias entre ambas condiciones son grandes y consistentes en

las dimensiones de carga, esfuerzo y usabilidad. La exigencia física pasa de una media de 4,33 (SD 1,15) en la condición no adaptada a 1,91 (SD 0,79) en la adaptada; el esfuerzo percibido, de 4,33 (SD 1,70) a 1,75 (SD 0,97); y la frustración, de 3,50 (SD 1,40) a 1,50 (SD 0,67). La facilidad de uso de las interacciones mejora en la misma línea, de 4,16 (SD 1,40) a 1,67 (SD 0,89) en una pregunta donde el valor bajo indica mayor facilidad, y la dificultad para agarrar, soltar, desplazarse y añadir productos cae de 4,41 (SD 1,57) a 1,50 (SD 0,67). La seguridad percibida durante el uso mejora de 3,67 (SD 1,15) a 1,33 (SD 0,49), y la intención de uso recurrente, de 4,75 (SD 1,42) a 1,67 (SD 0,65). Estas magnitudes muestran que las adaptaciones no solo hacen las tareas posibles, sino que las hacen percibir como notablemente menos costosas.

Las diferencias se contrastaron con un test de Wilcoxon de rangos con signo, apropiado para datos ordinales de tipo Likert, sobre las preguntas que presentaban variabilidad entre condiciones. El artículo III adopta un umbral conservador de $\alpha = 0{,}005$ para reducir la probabilidad de falsos positivos ante comparaciones múltiples. Bajo ese criterio, las diferencias son estadísticamente significativas en la práctica totalidad de las preguntas evaluadas (por ejemplo $p = 0{,}0005$ en exigencia física, $p = 0{,}0045$ en esfuerzo, $p = 0{,}0041$ en seguridad percibida e intención de uso). La diferencia porcentual media entre las respuestas a las condiciones adaptada y no adaptada, considerando todos los encuestados, es del 40,15 % con una desviación estándar del 8,60 %, cifra que converge de forma notable con la reducción cercana al 40 % observada en las métricas objetivas de esfuerzo.

Las dimensiones de motion sickness se mantienen bajas y, lo que es más relevante, sin diferencia apreciable entre condiciones: malestar o mareo general en 1,00 (SD 0,00), dolor de cabeza en 1,17 (SD 0,39) y visión borrosa o fatiga ocular en 1,25 (SD 0,45), valores idénticos en los dos entornos. La inmersión percibida tampoco varía entre condiciones (media de 1,50 en una pregunta donde el valor bajo indica alta inmersión), lo que indica que la adaptación de los gestos no degrada la sensación de presencia. Estos niveles son compatibles con sesiones de duración limitada a 25-30 minutos, una decisión deliberada del diseño experimental para evitar fatiga acumulada en una población más propensa a distintos tipos de fatiga física y mental. La fatiga autorreportada sí se separa con claridad entre condiciones (de 3,67 a 1,17), lo que confirma que el cansancio del entorno no adaptado proviene del esfuerzo interactivo y no del

dispositivo. La combinación de métricas objetivas y cuestionarios subjetivos respalda así la conclusión común desde dos ángulos independientes.

6.4.6 | Limitaciones reconocidas

La principal limitación es el tamaño muestral: 10 pacientes con lesión medular constituyen una muestra pequeña en términos estadísticos absolutos, aunque moderadamente relevante en el contexto de los estudios clínicos en VR aplicada, donde disponer de un número elevado de pacientes es complicado debido a las dificultades logísticas que conlleva el trabajo con poblaciones hospitalizadas o con seguimiento rehabilitador intenso. El propio artículo III justifica este tamaño con el hecho de que las tendencias observadas son consistentes y los efectos estadísticos muy fuertes ($p < 10^{-10}$ en el ED). La segunda limitación es la ausencia de un seguimiento longitudinal: el estudio mide la diferencia entre adaptado y no adaptado en una única sesión, pero no evalúa cómo evoluciona el desempeño de cada paciente a lo largo de varias sesiones, lo que sería relevante para juzgar si las adaptaciones favorecen el aprendizaje o introducen dependencias indeseadas. La tercera limitación es que el estudio se ejecuta en un único entorno virtual. La cuarta limitación es la dependencia de Meta SDK para el hand-tracking, ya señalada en RQ3 y heredada por RQ4 al construirse las adaptaciones sobre la misma base. La compatibilidad con el estándar OpenXR mitiga parcialmente este punto.

6.4.7 | Síntesis de la respuesta a RQ4

La respuesta a RQ4 se construye sobre cuatro contribuciones diferenciadas. Primero, la caracterización clínica de las limitaciones motoras (espasticidad, atrofia, mano en garra) que excluyen a las personas con lesión medular del uso de los SDK de hand-tracking estándar, realizada en colaboración con la Unidad de Biomecánica y Ayudas Técnicas del HNPT. Segundo, el diseño e implementación de cinco mecanismos adaptados (teletransportación por pose mantenida, rotación corporal por movimiento del antebrazo, agarre con umbrales modificados y auto-grasp, *distance grab* por raycasting desde la articulación MCP, liberación por apuntado a papelera) que cubren las acciones nucleares de un escenario de VR-shopping y se construyen sobre la misma base de VRSI sin alterar su núcleo. Tercero, las métricas objetivas *Difficulty Index* y *Effort*

Degree, derivadas íntegramente del hand-tracking, que permiten comparar de manera cuantitativa el esfuerzo entre condiciones experimentales sin depender de cuestionarios subjetivos. Cuarto, la validación experimental en el HNPT con 10 pacientes con lesión medular y 2 sujetos de control sanos, que confirma con tasas de completación del 100 % (frente al 33 % medio del entorno no adaptado), una reducción del 46 % en DI y una diferencia altamente significativa en el ED ($t = -6{,}79$, $p = 8{,}71 \times 10^{-11}$) que las adaptaciones permiten realizar tareas que sin ellas son imposibles para esta población. La triangulación entre análisis clínico, diseño técnico, métricas objetivas y validación con usuarios reales responde a RQ4 de manera sustantiva y consolida la dimensión de accesibilidad como elemento de primer orden de la arquitectura VR-ZOCO, en consistencia con la declaración FE10 del artículo arquitectónico.

6.5 | RQ5: recolección y análisis de datos comportamentales

> ***Pregunta de investigación 5.*** *«¿Puede el comportamiento del usuario en este tipo de entornos comerciales ser recopilado de manera no invasiva y analizado para optimizar la experiencia de compra, incrementar las ventas, personalizar la oferta y tomar decisiones estratégicas basadas en datos reales?»*

La tabla 6.5 resume las contribuciones por publicación para responder esta pregunta de investigación.

6.5.1 | Utilidad de los datos comportamentales

Los entornos de VR-shopping generan señales de comportamiento que son inherentemente más ricas que las disponibles en una web de comercio electrónico convencional. Esta pregunta tiene, por tanto, dos planos: uno técnico (qué señales son capturables y qué análisis admiten) y uno aplicado (por qué cada señal importa para el comercio y qué decisión de negocio o de diseño habilita). Los SDK de los headsets comerciales registran la posición y orientación tridimensional de la cabeza y de ambas manos del usuario, además de los eventos discretos de interacción (agarrar, soltar, teletransportarse o pulsar elementos de

Tabla 6.5: **Mapeo de contribuciones a RQ5.** El artículo II del compendio (VRSI) es la única publicación que aborda RQ5 de manera integral. El artículo de Arcos et al. [6] es el precursor cronológico: formaliza los cuatro tipos de acción que un usuario realiza en un entorno de VR-shopping (manipulación, visión, teletransportación, uso) y diseña la interfaz VR-IA que después VRSI generaliza. Artículo IV y el artículo [92] aportan validación de uso: el primero a escala (N=48 en feria comercial), el segundo a menor escala (N=10).

Publicación	Venue / año	Tipo	O5		
			SO5.1	SO5.2	SO5.3
Virtual Reality Shopping-Insights (VRSI)	SoftwareX, 2024	AC	●	●	●
Gamifying VR-shopping	Virtual Reality, 2026	AC	●	●	●
User Activity Monitoring & Recommendation	IEEE AIxVR, 2024	CI	○	●	○
Exploring Interaction Mechanisms	ICEIS, 2024	CI	○	●	●
VR-ZOCO architectural proposal	Appl. Syst. Innov., 2024	AR	●	●	○

Contribución: ● principal ● parcial ○ no contribuye **Tipo:** AC compendio AR revista CL cap. libro CI congr. int. CN congr. nac.

la interfaz). Estas señales, en su conjunto, permiten reconstruir con un grado de detalle alto el recorrido del usuario por la tienda, qué productos ha mirado y durante cuánto tiempo, qué objetos ha tomado, en qué orden, con qué mano, desde qué distancia, y qué decisión ha tomado finalmente.

El problema no es la disponibilidad de la señal, sino estandarizar su recolección y análisis en un ámbito determinado; en este caso, el VR-shopping. Los estudios previos en la literatura tienden a uno de tres patrones poco satisfactorios para la investigación reproducible. Algunos se apoyan exclusivamente en cuestionarios estandarizados (UEQ, SUS, MSAQ) y dejan fuera el dato comportamental fino. Otros sí utilizan datos de interacción, pero no publican las herramientas software con las que los obtuvieron, lo que impide replicar el estudio o aplicar el mismo método a un nuevo dominio. Y un tercer grupo propone marcos conceptuales para el análisis sin proporcionar la implementación que los soporte. La capa de interacción y recolección de datos declarada en la arquitectura VR-ZOCO [5] reconoce esta necesidad a nivel funcional, pero, como en los demás casos arquitectónicos, su materialización experimental queda como contribución separada. Esta pregunta de investigación es, en

consecuencia, cómo construir una infraestructura única y abierta que cubra los tres pasos del ciclo (configurar el entorno, recolectar la actividad y analizar el comportamiento) y permita a otros investigadores replicar y extender el trabajo, incluso iniciar diferentes estudios a partir de la información extraída, tal y como se realizó en el artículo IV del compendio. Esa es la contribución del artículo II del compendio [2].

6.5.2 | Formalización de los eventos de actividad

Antes de discutir VRSI en detalle conviene situar su precursor directo. En el contexto del proyecto VR-ZOCO, el artículo de Arcos et al. presentado en IEEE AIxVR 2024 [6] aborda una pregunta más acotada: si vamos a alimentar un sistema de recomendación basado en IA con datos de actividad del usuario en VR, ¿qué tipo de eventos conviene registrar y cómo formalizarlos? Su respuesta propone cuatro acciones nucleares formalizadas:

- $M(c_i)$: manipulación de un producto de categoría c_i (agarre, transformación, gestos, pulsaciones), capturando una tupla rica con *timestamp*, ID del usuario, del producto, de la tienda y del tipo de interacción.
- $V(c_i)$: visión de productos de categoría c_i, implementada mediante *ray-casting* desde el punto entre las cejas hacia la dirección de mirada cuando el dispositivo no dispone de eye-tracking real.
- $T(c_i)$: teletransportación a una zona expositora de productos de categoría c_i.
- $U(m, o, R, a)$: uso explícito de un objeto, definido por la mano, el objeto, la orientación y la acción aplicada.

Los eventos se acumulan en lotes mediante llamadas a una API y se persisten en una base de datos PostgreSQL. Esta formalización tiene dos virtudes que después VRSI hereda y generaliza: primero, define un vocabulario común para hablar de actividad en VR-shopping; y segundo, separa la captura del evento de su explotación por parte del sistema de IA, lo que permite reutilizar los datos para cualquier otro análisis.

6.5.3 | VRSI como marco de trabajo

VRSI (Virtual Reality Shopping-Insights) organiza la respuesta a RQ5 en tres componentes. El primero, *Retail Environment Setup*, ofrece *prefabs* de Unity reutilizables y atiende al SO5.1. El segundo, *Data Collection*, instrumenta las señales de los SDK y aborda al SO5.2. El tercero, *Behaviour Statistics*, transforma los datos en métricas y reportes estandarizados y atiende al SO5.3. Los tres componentes están publicados en un repositorio público con documentación, ejemplos de uso, vídeo demostrativo y un proyecto Unity completo a modo de plantilla de partida. La Fig. 6.13 muestra un diagrama de bloques que representa la arquitectura de este marco de trabajo.

6.5.4 | Configuración del entorno de compra (SO5.1)

El componente *Retail Environment Setup* establece la base sobre la que se construye una aplicación de VR-shopping. Aporta *prefabs* reutilizables para la creación del espacio tridimensional, configurados de tal forma que: (i) son compatibles con funciones de hand-interaction del Meta SDK y (ii) poseen componentes nativos del motor de Unity ajustados y configurados para la captura y análisis de datos por defecto. En particular para las estanterías (Top, Middle y Bottom como tres alturas de exposición que se pueden modificar si la estantería tiene otra distribución), para las áreas de interés (AOIs, asociadas a productos o secciones concretas), para las zonas de interés (ZOIs, anillos concéntricos sobre el plano del suelo: la propia estantería, adyacente, cercana y lejana), las manos y cabeza virtuales del usuario y un envoltorio para los productos que el usuario desee añadir al entorno. El sistema integra también la configuración de la navegación por teletransportación con sus zonas de aterrizaje y las opciones de hand-tracking discutidas en RQ3 (Sección 6.3). Toda la configuración es accesible desde el editor de Unity mediante propiedades expuestas, lo que permite a un investigador o desarrollador construir una tienda virtual completa sin escribir código de bajo nivel. Esta capa de abstracción es lo que distingue a VRSI de una colección de scripts ad-hoc: el investigador que quiere estudiar el comportamiento del consumidor en un layout concreto solo tiene que diseñar el layout, no las primitivas técnicas que lo soportan. La elección de tres alturas de exposición no es casual: la altura del producto en el lineal condiciona la atención y la probabilidad de compra en él, de modo que

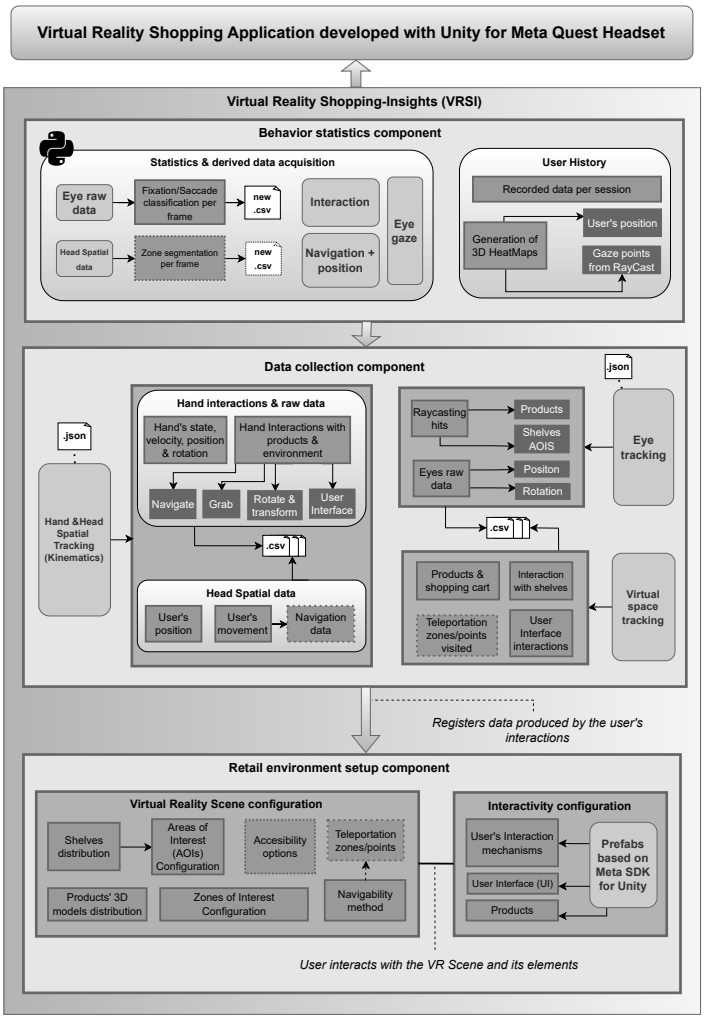


Figura 6.13: **Arquitectura general del framework VRSI organizada en sus tres componentes.** De arriba a abajo: el componente *Retail Environment Setup* (panel superior) proporciona los prefabs de Unity reutilizables para construir la tienda virtual y configura las áreas y zonas de interés; el componente *Data Collection* (panel intermedio) instrumenta los SDK de Meta y captura los veintiún campos por frame que alimentan a los componentes posteriores; el componente *Behaviour Statistics* (panel inferior, implementado en Python como herramienta de línea de comandos) procesa los datos crudos y genera las métricas espacio-temporales, las fijaciones oculares con el algoritmo I-DT y los reportes en PDF. Adaptada de la Figura 1 del Artículo II del compendio [2].

registrar el AOI activo permite contrastar de forma cuantitativa el efecto de la colocación.

6.5.5 | Recolección de datos no invasiva (SO5.2)

El componente *Data Collection* es el que captura la riqueza comportamental que motiva el trabajo. Su instrumentación se apoya en las tecnologías ya integradas en los headsets autónomos comerciales sin requerir hardware adicional. La recolección es pasiva por diseño: la experiencia del usuario no se ve afectada por la instrumentación. La Tabla 6.6 resume los veintiún campos que VRSI registra por cada frame.

Tabla 6.6: **Datos capturados por el componente Data Collection de VRSI por cada frame.** Cada fila representa un campo del registro temporal que VRSI escribe en CSV. Detalle exhaustivo de tipos y rangos en la Tabla 1 del artículo II [2].

Categoría	**Campos capturados por frame**
Temporales	Número de frame, marca de tiempo (segundos desde el inicio de la aplicación).
Cinemática de manos	Posición, rotación y velocidad de mano izquierda y derecha en el espacio 3D.
Cinemática de cabeza	Posición y rotación del HMD.
Eye-tracking	Posición y rotación de ambos ojos, punto de colisión promediado del *raycast* de mirada.
Estructura espacial	AOI activo (Top/Middle/Bottom), nombre del producto, estado del producto en el ciclo Interactor-Interactable (Disabled, Normal, Hover, Select), sección temática, identificador de estantería.
Eventos de interacción	Estado del Interactor de mano izquierda y derecha, acción de carrito de compra (Add, Remove), elemento de UI y acción asociada.
Eventos de navegación	Zona de teletransporte objetivo, duración del apuntado, éxito o no del teletransporte, ZOI activo (Start, Adjacent, Near, Far), distancia recorrida o distancia al estante más cercano.

Cada familia de la Tabla 6.6 responde a una pregunta distinta sobre el comprador. La mirada y la cinemática de cabeza informan de qué capta su atención y durante cuánto tiempo; la cinemática de manos y los eventos de interacción,

de qué productos toca, examina y descarta, y en qué orden; los eventos de navegación, de cómo recorre el espacio y dónde se detiene. En conjunto representan información que una web convencional, limitada a métricas como número de clics y tiempos de página, no puede capturar.

Los datos en crudo se persisten en un conjunto de archivos CSV agrupados por familia (eye-tracking, cinemática, interacciones, navegación, actividad de compra, ...) y dos archivos JSON adicionales que sirven como entrada al componente de estadísticas para la generación de mapas de calor volumétricos (ver Fig. 6.14) y de fijaciones de mirada. Esta separación entre datos crudos y datos derivados es deliberada: un investigador con experiencia en análisis de datos puede ignorar el componente de estadísticas y procesar los CSV con sus propias herramientas, mientras que un usuario con menos experiencia técnica puede consumir directamente los reportes generados por VRSI para iniciar de manera explorativa un primer análisis de los datos.

6.5.6 | Estadísticas de comportamiento (SO5.3)

El componente *Behaviour Statistics* es una herramienta de línea de comandos escrita en Python que procesa los datos en crudo y genera tres tipos de salida. La primera son métricas espacio-temporales derivadas: tiempo de presencia por sección, número y duración de visitas a cada estantería y a cada AOI, distancia recorrida total, mapas de calor volumétricos visualizables en Unity con paleta de color configurable. La segunda son métricas oculares basadas en el algoritmo **I-DT!** (**I-DT!**) [70] adaptado a entornos VR inmersivos, que segmenta las trayectorias de mirada en fijaciones y sacadas según un umbral de dispersión configurable, criterio adecuado para el ruido inherente al eye-tracking de los headsets comerciales. La tercera salida son tres reportes PDF (uno por familia de datos: mirada, interacción, navegación con posición) que combinan gráficas y tablas numéricas según un layout estandarizado. Estos reportes proporcionan un análisis inicial sin requerir programación adicional por parte del usuario, lo que sitúa a VRSI en el espectro del software *out-of-the-box* sin perder la flexibilidad para análisis avanzados.

Una decisión de diseño que conviene subrayar es la segmentación por ZOI. Las cuatro zonas (Start, Adjacent, Near, Far) [28] permiten reescribir cualquier métrica espacio-temporal en función de la proximidad del usuario

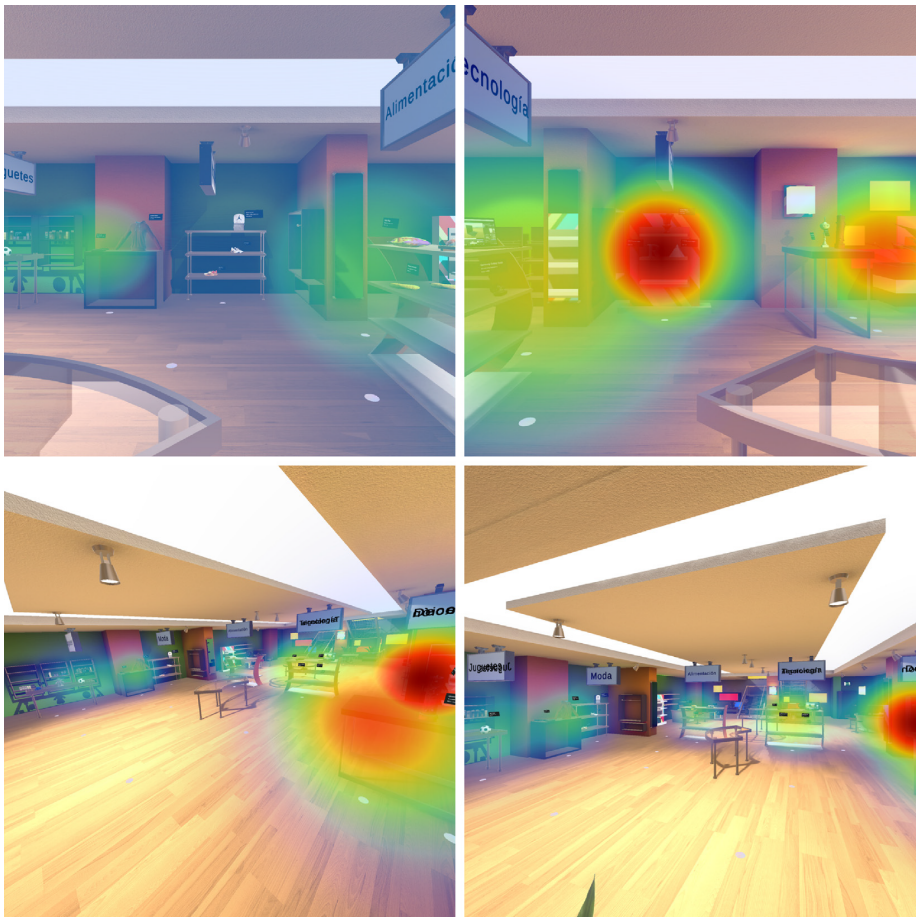


Figura 6.14: **Ejemplos de mapas de calor volumétricos generados por el componente Behaviour Statistics de VRSI.** Los cuatro paneles muestran distintos puntos de vista de la atención agregada de los usuarios dentro de la tienda virtual utilizada como prototipo en distintos trabajos de la tesis. Los colores cálidos (rojo, naranja) corresponden a zonas donde los usuarios fijaron la mirada con mayor frecuencia o duración; los colores fríos (azul, verde) corresponden a zonas de menor atención visual. Estos mapas permiten al investigador identificar de un vistazo qué productos y qué secciones de la tienda captan el interés del comprador y rediseñar el *layout* en consecuencia. La paleta de color es configurable desde la propia herramienta. Imagen obtenida del artículo II del compendio [2].

a la estantería, lo que da pie a análisis comparativos. Estas comparativas son posibles porque la ZOI activa se etiqueta por frame, y son las que sostienen análisis posteriores como los del artículo IV del compendio [4], que se discute con detalle en RQ6.

6.5.7 | Validación de uso por terceros

VRSI se publicó como Original Software Publication: el proceso de revisión por pares de SoftwareX evalúa no solo el artículo descriptivo sino también el código y su capacidad para ser ejecutado por terceros. La aceptación de la publicación constituye, por tanto, una primera capa de validación operativa. A esta validación se añaden tres usos efectivos del framework, dos de ellos dentro de la tesis y uno externo.

El primero es el estudio empírico de mecanismos de interacción [92] discutido en RQ3, ejecutado con 10 participantes. Los nueve archivos CSV de VRSI fueron la materia prima del análisis cuantitativo del trabajo, en particular para los gráficos de tiempo y número de interacciones por sección y por producto, y para los datos de tracking ocular utilizados en la discusión de qué productos llamaron más la atención del usuario. El estudio publica además todos los archivos generados en un repositorio público.

El segundo es el caso de estudio en el HNPT del artículo III del compendio [3] discutido en RQ4. Las métricas DI y ED que ese artículo define, se calculan íntegramente a partir de los datos de cinemática de manos que VRSI captura, en concreto los campos de posición y velocidad.

El tercero, y empíricamente el más exigente, es el caso de estudio del artículo IV del compendio [4], que se discute en detalle en RQ6 (Sección 6.6). En éste, se implementó VRSI para capturar datos comportamentales completos de 48 participantes durante una feria comercial real. Para lo que aquí concierne, la conclusión relevante es que VRSI sostuvo, sin reportes de fallos significativos, un despliegue con público no entrenado, en condiciones ambientales diversas (ruido, distracciones, sesiones cortas, iluminación que cambiaba levemente con el paso de las horas) y con productos digitalizados a partir de modelos reales. La validación de dicho estudio de N=48 es complementaria a la validación controlada de N=10 del párrafo anterior: la primera mide robustez, la segunda mide profundidad analítica, y juntas cubren ambas dimensiones de evidencia

esperables de un framework de investigación.

6.5.8 | Posicionamiento frente a la literatura y limitaciones

La aportación de VRSI se sitúa en una región poco poblada de la literatura sobre VR-shopping. Frameworks comerciales como Cognitive3D[1] ofrecen análisis de comportamiento en VR con calidad de producto, pero son privativos, de pago y no están diseñados específicamente para escenarios de compra. Otras publicaciones suelen abordar parte del ciclo (instrumentación de *eye-tracking*, o de interacciones, pero rara vez todo el conjunto) [28, 29, 70] y casi nunca se publican como software reutilizable. VRSI cubre esta brecha al unir tres condiciones a la vez: alcance completo del ciclo (configuración, recolección, análisis), publicación abierta y validación tanto por revisión por pares como en casos de estudio empíricos.

Conviene reconocer las limitaciones que el propio artículo II identifica de manera explícita. La primera es que VRSI captura los datos brutos de hand-tracking, pero no calcula métricas derivadas de los gestos concretos del usuario, sino solo los estados del Interactor-Interactable de Meta SDK. Un análisis fino del repertorio gestual (por ejemplo, contar cuántos *pinches* fallan, cuántos *poke* se realizan sobre superficie equivocada) requiere por ahora análisis ad-hoc sobre los CSV. La segunda, ligada a la anterior, es la dependencia del SDK de Meta y la limitación a la familia de dispositivos Quest. La compatibilidad con OpenXR existe pero no está implementada en la versión publicada del framework. La tercera es que el componente *Behaviour Statistics* se limita a métricas espacio-temporales y oculares clásicas. Otras dimensiones de análisis relevantes para VR-shopping (análisis de patrones temporales mediante series, segmentación no supervisada de comportamientos mediante técnicas de aprendizaje no supervisado) quedan fuera del alcance de la herramienta de línea de comandos y se realizan, cuando son necesarias, mediante scripts externos sobre los CSV.

[1] `https://cognitive3d.com`

6.5.9 | Síntesis de la respuesta a RQ5

Más allá de demostrar que las señales son capturables y analizables, los tres estudios que usaron VRSI produjeron conclusiones comportamentales concretas que ilustran el tipo de conocimiento que la instrumentación habilita y que un canal web convencional no podría haber generado. Conviene recogerlos aquí, porque son la evidencia de que el dato registrado se traduce en valor y no en mera telemetría.

Primero, el comportamiento del usuario depende del tipo de escena, y la instrumentación lo captura de forma diferenciada. En el estudio de campo del artículo IV [4], las mismas personas se comportan de manera distinta en la escena de tienda (centrada en tomar e inspeccionar productos) y en la escena gamificada (centrada en locomoción y resolución de tarea), lo que confirma que registrar las métricas por escena, y no de forma agregada, es una decisión de diseño acertada para entender el comportamiento de compra.

Segundo, y más relevante para el comercio, la intensidad de interacción resultó ser un mejor predictor de la conversión que la eficiencia en la tarea. El perfil de usuario que más manipulaba productos (más agarres, más objetos distintos examinados) y más se desplazaba por el entorno convirtió a compra en mayor proporción que el perfil eficiente y orientado a completar la tarea con el mínimo de acciones. Es un hallazgo con valor directo para el diseño de tiendas virtuales: fomentar la interacción directa y exploratoria con los productos parece un motor de venta más eficaz que optimizar un recorrido ágil y sin fricción, lo que invierte la intuición habitual del diseño de interfaces orientado a la eficiencia.

Tercero, el comportamiento observado aportó información que el cuestionario no recogía. En ese mismo estudio, los dos perfiles comportamentales identificados no mostraron diferencias estadísticamente significativas en ninguna de las ocho medidas actitudinales del cuestionario post-sesión (por ejemplo, atractivo de la VR con $p = 0{,}344$ o usabilidad e inmersión con $p = 0{,}785$), pese a comportarse de forma claramente distinta dentro del entorno y a convertir a compra en proporciones muy diferentes. Esto respalda con evidencia empírica la premisa que motiva RQ5: la conducta registrada de forma no invasiva captura aspectos del comprador que las encuestas tradicionales, base habitual de los estudios de VR-shopping, no alcanzan a reflejar.

Cuarto, los datos de mirada se mostraron útiles incluso para validar decisiones técnicas ajenas a la propia recolección. En el estudio controlado del artículo paralelo de congreso [92], cuando se pidió a los participantes que señalaran los productos que percibían como más realistas, la mitad de las respuestas correspondió a modelos generados con NeRF, dato que conecta con la conclusión de RQ2 sobre la superioridad perceptual de esta tecnología y que solo fue posible contrastar gracias a la instrumentación de seguimiento ocular del framework.

6.6 | RQ6: VR y gamificación en compras reales

> ***Pregunta de investigación 6.*** *«¿Influye positivamente la VR, combinada con estrategias de gamificación, en el comportamiento de compra real del consumidor?»*

La tabla 6.7 resume las contribuciones por publicación para responder esta pregunta de investigación.

Tabla 6.7: **Mapeo de contribuciones a RQ6.** RQ6 es la pregunta de cierre del compendio: integra los componentes desarrollados en las RQ anteriores y los pone a prueba en un escenario comercial real. El artículo IV del compendio es la única publicación que aborda los cuatro subobjetivos. Artículo I (digitalización 3D con NeRF) y artículo II (VRSI) contribuyen instrumentalmente al diseño de los escenarios (SO6.1) y a la captura de datos en el campo (SO6.2), pero no son contribuciones directas a SO6.3 ni SO6.4, que son novedosos del artículo IV. La arquitectura VR-ZOCO declara la integración de gamificación a nivel funcional.

Publicación	Venue / año	Tipo	O6			
			SO6.1	SO6.2	SO6.3	SO6.4
Gamifying VR-shopping	Virtual Reality, 2026	AC	●	●	●	●
Performance Evaluation of low-cost 3D scanning	Appl. Sci., 2024	AC	●	○	○	○
Virtual Reality Shopping-Insights	SoftwareX, 2024	AC	●	●	○	○
VR-ZOCO architectural proposal	Appl. Syst. Innov., 2024	AR	●	○	○	○

Contribución: ● principal ● parcial ○ no contribuye **Tipo:** AC compendio AR revista CL cap. libro CI congr. int. CN congr. nac.

6.6.1 | El vacío en la literatura: VR, gamificación y compra real

La hipótesis general de la tesis afirma que los entornos de VR-shopping basados en digitalización de bajo coste e interacción gestual inclusiva pueden generar un impacto positivo y cuantificable en la intención de compra del consumidor. Las RQ1 a RQ5 cubren la viabilidad técnica (arquitectura, digitalización, interacciones, accesibilidad, infraestructura de datos). RQ6 cierra el ciclo al someter la propuesta completa a un examen más exigente: medir si, en un contexto comercial real, con público no entrenado y con dinero propio, los usuarios se ven influenciados positivamente por el uso de estas tecnologías en su decisión de compra.

La literatura previa sobre VR-shopping ha estudiado de manera amplia la intención de compra y la satisfacción percibida mediante cuestionarios estandarizados aplicados en laboratorio. La gamificación, por su parte, se ha estudiado en contextos de aprendizaje, salud y entretenimiento, con resultados consistentes sobre su capacidad para aumentar la motivación y el tiempo de permanencia. La intersección de los dos campos, sin embargo, está poco explorada y, lo que es más relevante, casi siempre se aborda en condiciones controladas donde la compra es hipotética, con valores monetarios ficticios. La conexión entre el comportamiento del usuario en un entorno VR-shopping gamificado y su decisión de compra real (con dinero propio, en un evento abierto al público, sobre productos físicos que recibirá) es exactamente la brecha que el artículo IV del compendio [4] aborda.

6.6.2 | Diseño de los dos escenarios experimentales (SO6.1)

El diseño correspondiente a SO6.1 se compone de dos escenarios complementarios accesibles desde la misma aplicación VR. El primero es un entorno de tienda virtual con productos físicos digitalizados mediante Luma AI, lo que cierra de forma natural el ciclo con RQ2. Los productos escaneados fueron productos reales de un pequeño comercio, *Librería Serendipia*, dedicado a artículos de colección, literatura, juegos de mesa y artículos relacionados. Cabe destacar que el entorno virtual construido para su exposición fue una representación minimalista de su tienda. El segundo es un escenario gamificado basado en

la narrativa de una novela gráfica: el usuario entra en un mundo distinto a través de una puerta dimensional, recorre las primeras páginas de la novela gráfica a modo de escenografía, y debe resolver un puzzle de piezas dispersas por el entorno antes de completar la experiencia. La gamificación no es un elemento decorativo añadido al final del flujo de compra; es la estructura narrativa que sostiene todo el segundo escenario y exige al usuario locomoción, manipulación de objetos y resolución de tarea para avanzar.

La elección de los dos escenarios responde a una hipótesis de diseño relevante: si la gamificación favorece la involucración y motivación del usuario, y estas favorecen la decisión de compra, entonces se esperaría que la conducta del usuario en el escenario gamificado fuera predictiva de la compra final, y no solo la conducta en el escenario de tienda. Sostener esta hipótesis exige medir el comportamiento en ambos escenarios con la misma infraestructura, lo que el componente Data Collection de VRSI [2] permite directamente.

6.6.3 | Caso de estudio (SO6.2)

El caso de estudio se llevó a cabo durante la feria anual *Manchacómic!*, un evento comercial abierto al público en Ciudad Real, con 48 participantes voluntarios reclutados de manera orgánica en el propio evento, sin que estos conociesen de qué trataba la experiencia. Cada voluntario era instruido brevemente en cómo interactuar dentro de la aplicación, antes de comenzar cada sesión. La duración media por sesión fue de aproximadamente 20 minutos, con la primera escena (tienda) limitada a 5 minutos para garantizar que todos los participantes pudieran completar ambos escenarios dentro del tiempo asignado por los organizadores del evento. Tras la experiencia, los participantes podían comprar con su propio dinero (en el puesto de *Librería Serendipia* situado en el propio evento, mediante pago directo, sin que ningún investigador interviniera en la transacción) los productos físicos que habían explorado en la tienda virtual o la novela gráfica sobre la que estaba basada la segunda escena.

La población se caracteriza por una diversidad demográfica amplia y por una experiencia previa muy limitada con dispositivos VR, dato que el artículo IV reporta mediante un cuestionario post-sesión y que es valioso desde un punto de vista de validez ecológica: el experimento mide cómo se comportan

compradores ordinarios. La muestra (n=48) tiene una edad media de 23,7 años (SD 13,6; mediana 21, rango 8 a 55), un 68,8 % de hombres y un 29,2 % de mujeres, y una experiencia previa con VR autodeclarada baja (media 2,58 sobre 5). Durante la sesión se capturaron 21 métricas espacio-temporales y cinemáticas por participante, derivadas íntegramente de los CSV que VRSI produce. El cuestionario, de 14 preguntas Likert de 1 a 5, agrega composiciones temáticas validadas (atractivo de la VR, usabilidad e inmersión, intención de compra y reuso, y enganche con la gamificación).

La tasa de conversión global a compra real fue del 35,42 %, notablemente superior a las tasas típicas del e-commerce convencional, que rara vez exceden el 4 %. Conviene desglosar esa cifra, porque mezcla dos productos de naturaleza distinta. De los 48 participantes, 9 (18,75 %) compraron un producto que habían examinado en la tienda virtual de la primera escena, y 8 (16,67 %) compraron un ejemplar de la novela gráfica *Calatrava: Pedro Girón* sobre la que se construyó la segunda escena gamificada, una novela de 20 €, precio superior al habitual de entre 5 y 10 € del género, y que ninguno de los compradores conocía previamente. La intensidad de la respuesta comercial fue también reseñable: el volumen de ventas de una sola mañana de sábado igualó al de todo el día anterior.

6.6.4 | Determinantes actitudinales de la intención de compra

Antes de analizar la conducta observada, el artículo IV explota los datos del cuestionario para identificar qué actitudes se asocian con la disposición a comprar. El hallazgo principal es contraintuitivo y tiene valor directo para el diseño: el motor de la intención de compra no es la novedad de la tecnología en sí. El atractivo general de la VR fue alto y homogéneo (media de 4,17 sobre 5 en interés por comprar con esta tecnología, 4,44 en novedad percibida), pero precisamente por ser homogéneo no discrimina entre quienes manifiestan alta y baja intención de compra. Los factores que sí discriminan son el enganche con la gamificación y la percepción de una interacción usable e inmersiva.

La evidencia más clara procede de comparar a los participantes con experiencia previa en VR baja (n=27) y alta (n=21) mediante pruebas t de Welch sobre los composites del cuestionario. El enganche con la gamificación es la

dimensión que más separa a ambos grupos, con una diferencia muy significativa y un tamaño de efecto grande (4,21 frente a 4,74; $p = 0{,}0001$; $d = 1{,}00$), seguida de la usabilidad e inmersión ($p = 0{,}02$; $d = 0{,}67$) y de la intención de compra y reuso ($p = 0{,}03$; $d = 0{,}63$). En cambio, el atractivo de la VR ($p = 0{,}87$; $d = 0{,}05$) y la incomodidad o mareo ($p = 0{,}96$) no diferencian a los grupos. La lectura de conjunto es que la conversión de la curiosidad en compra depende sobre todo de un diseño gamificado motivador y de una interacción percibida como natural e inmersiva, y no de la novedad del medio ni de la ausencia de cybersickness. Para una PYME que quiera adoptar VR-shopping, la implicación práctica es que invertir en la calidad y el enganche de la experiencia rinde más que apoyarse en el efecto novedad, que se desvanece.

6.6.5 | Análisis exploratorio: perfiles comportamentales y compra (SO6.3)

El análisis de SO6.3 procede en dos pasos. El primero identifica perfiles comportamentales sin recurrir a información demográfica ni a la decisión final de compra. Las 21 métricas extraídas de los CSV se reducen primero con UMAP a dos dimensiones y se agrupan después con HDBSCAN. El resultado de este proceso de aprendizaje no supervisado son dos clusters bien definidos (Figura 6.15b): un Cluster 0 con 20 participantes (42 % de la muestra) y un Cluster 1 con 28 participantes (58 %). Tres tests sobre variables demográficas (Mann-Whitney sobre edad $U = 238{,}0$, $p = 0{,}457$; chi-cuadrado sobre género $\chi^2(2) = 2{,}33$, $p = 0{,}311$; Mann-Whitney sobre experiencia previa en VR $p = 0{,}369$) descartan que los clusters sean explicables por demografía. La estructura emerge únicamente del comportamiento.

El segundo paso compara la tasa de compra entre clusters. El Cluster 0 alcanza una tasa del 55,0 % (11 compras sobre 20 participantes) y el Cluster 1 una tasa del 21,4 % (6 compras sobre 28). La asociación entre pertenencia a cluster y compra es estadísticamente significativa: la prueba chi-cuadrado arroja $\chi^2(1, N = 48) = 4{,}37$ con $p = 0{,}036$, el tamaño del efecto medido por Cramér's V es de 0,30 (efecto medio según los criterios convencionales), y el test exacto de Fisher confirma la significancia con $p = 0{,}031$. La interpretación es directa: la conducta del usuario en el entorno VR predice de manera estadísticamente fiable y con tamaño de efecto medio si comprará o no, sin necesidad

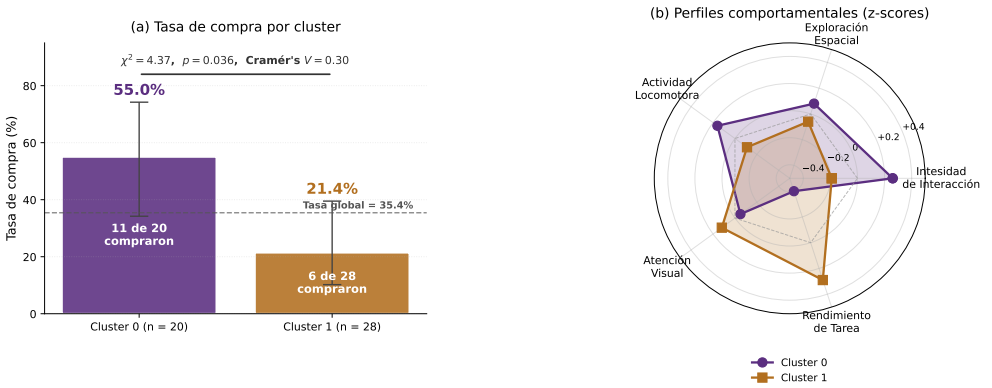


Figura 6.15: **Tasa de compra y perfiles comportamentales por cluster.** (a) Tasa de compra de cada uno de los dos clusters identificados por UMAP+HDBSCAN, con barras de error que representan los intervalos de confianza Wilson al 95 %. El Cluster 0 (n=20, etiquetado *Active engaged*) presenta una tasa del 55,0 % y el Cluster 1 (n=28, *Focused performer*) del 21,4 %. La línea discontinua marca la tasa global de compra del 35,4 %. La asociación entre cluster y compra es estadísticamente significativa ($\chi^2 = 4{,}37$, $p = 0{,}036$, Cramér's $V = 0{,}30$). (b) Perfiles comportamentales de cada cluster en las cinco dimensiones de la taxonomía de experiencia en VR de Hameed et al. (2024), representados como medias de *z-score* normalizado. El Cluster 0 destaca en Interaction Intensity y Locomotion Activity y queda por debajo en Task Performance; el Cluster 1 muestra el patrón inverso, con Task Performance y Visual Attention elevados a costa de menor interacción y locomoción. Datos extraídos del artículo IV [4].

de preguntarle por sus intenciones.

La caracterización cuantitativa de los dos clusters es lo que permite extraer la lección de diseño. Mapeando los 21 indicadores sobre las dimensiones de la taxonomía de experiencia en VR de Hameed et al. [76], respaldada por Croucher et al. [100] para entornos comerciales, los dos clusters se describen como dos perfiles bien diferenciados. El Cluster 0, etiquetado *Active engaged user*, presenta una firma comportamental dominada por una intensidad de interacción y una actividad de locomoción significativamente más altas (diferencias de $\Delta z = 0{,}26$ en interacción y $\Delta z = 0{,}16$ en locomoción respecto al otro grupo): agarra más objetos, manipula un mayor número de productos distintos y se desplaza más por el entorno, a costa de una menor eficiencia en la tarea ($\Delta z = -0{,}40$ en *Task Performance*). El Cluster 1, *Focused performer*,

muestra el patrón inverso: puntúa más alto en rendimiento de tarea ($\Delta z = 0{,}29$) y atención visual ($\Delta z = 0{,}12$), completando el puzzle de forma más eficiente, pero con menor interacción ($\Delta z = -0{,}19$) y locomoción ($\Delta z = -0{,}11$), lo que define a un usuario más observacional y estático.

La asociación de la mayor compra al Cluster 0 y no al Cluster 1 es coherente con la psicología clásica del comercio: el comprador que explora libremente, toma los productos y los examina sin prisa es más permeable a la compra que el comprador eficiente que entra con un objetivo y sale apenas lo cumple. La novedad del trabajo es que esta intuición del retail físico se confirma cuantitativamente y con tamaño de efecto medible en un entorno VR-shopping con productos digitalizados y narrativa gamificada. El compromiso central entre ambos perfiles, la relación inversa entre intensidad de interacción y eficiencia en la tarea, sugiere una conclusión de diseño relevante: fomentar la interacción directa y manual con los productos es un motor de conversión más eficaz que optimizar un recorrido ágil y sin fricción. Es la firma comportamental, y no la eficiencia, lo que mejor anticipa la compra.

Un resultado adicional refuerza el valor del dato comportamental sobre el autorreporte. Pese a comportarse de forma tan distinta y a convertir en proporciones muy diferentes, los dos clusters no mostraron diferencias estadísticamente significativas en ninguna de las ocho medidas actitudinales del cuestionario (por ejemplo, atractivo de la VR con $p = 0{,}344$, o usabilidad e inmersión con $p = 0{,}785$). Dicho de otro modo, la encuesta no habría distinguido a estos dos tipos de comprador; solo la conducta observada los separa. Es la confirmación empírica de que la instrumentación comportamental captura señales que el cuestionario tradicional no alcanza.

6.6.6 | Predicción exploratoria de la compra (SO6.4)

El artículo IV aborda esta cuestión con clasificadores Random Forest entrenados y evaluados mediante validación cruzada estratificada repetida (5 particiones, 3 repeticiones, 15 evaluaciones por configuración), con métrica principal AUC y consciencia explícita del desbalance de clases (17 compradores frente a 31 no compradores). El objetivo declarado es exploratorio y de prueba de concepto, no construir un clasificador generalizable, dada la limitada potencia estadística de la muestra. Se comparan tres modelos de complejidad

creciente.

El primer modelo, entrenado con las 21 métricas comportamentales, alcanza un AUC promedio de 0,655, modesto pero por encima del azar, con un F1 medio de 0,315 y una recall de 0,256. La importancia de las características se distribuye entre numerosas variables: la más influyente, la duración media de fijación ocular en la escena gamificada (`AvgFixDurS2`), solo concentra el 7,6 % del peso del modelo. Esta dispersión es informativa: la decisión de compra no se explica por una única acción, sino por un patrón complejo y multivariado, coherente con la idea de que el engagement es una propiedad agregada y no atribuible a un gesto concreto.

El segundo modelo aplica selección de características y se queda con las 11 más predictivas. La elección de $k = 11$ no es arbitraria: se alineó con el número de componentes que UMAP proporcionó para el mejor índice DBCV en el proceso de clustering. Reducir el espacio de características casi a la mitad mejora el rendimiento, con un AUC promedio de 0,702 (una ganancia de 0,046 sobre el modelo completo) y un F1 de 0,412. Las variables que emergen como más importantes en este modelo refinado, `AvgFixDurS1`, `GrabsPerMinS1` y `TeleportCountPerMinS1`, reafirman que la atención visual, la interacción y la locomoción son los predictores comportamentales de mayor peso.

El tercer modelo amplía el espacio incorporando los datos actitudinales del cuestionario (los 21 indicadores comportamentales más cuatro composites y cuatro ítems individuales, 29 variables en total) y mejora de forma drástica: AUC promedio de 0,849. Las dos características más influyentes con diferencia son los composites de intención de compra y reuso (`PurchaseReuse` y `PurchaseIntention`), lo cual es esperable, ya que la intención declarada es un fuerte predictor de la acción. Pero el modelo no se apoya solo en la actitud: métricas puramente comportamentales como `TeleportCountPerMinS1`, `GrabsPerMinS2` y `AvgFixDurS1` se mantienen en posiciones altas y confirman que la conducta observada aporta información complementaria al autorreporte. La conclusión metodológica que el artículo IV deriva es que el enfoque holístico, datos implícitos del comportamiento más datos explícitos de la actitud, es superior a cualquiera de los dos por separado para predecir la compra en VR-shopping.

6.6.7 | Hallazgos secundarios y consistencia transversal con otras RQ

Más allá de los resultados principales sobre clusters y predicción, el artículo IV reporta varios hallazgos que enriquecen la discusión y refuerzan la coherencia transversal del compendio. Primero, los productos digitalizados que componen la escena de tienda funcionaron sin incidencias en las tres unidades de Meta Quest 3 durante toda la duración del evento, lo que confirma la conclusión de RQ2 (Sección 6.2) sobre la viabilidad técnica de la digitalización de bajo coste en hardware comercial. Segundo, los participantes no recibieron entrenamiento específico en hand-interaction más allá de una instrucción breve sobre los gestos básicos, lo que confirma la conclusión de RQ3 (Sección 6.3) sobre la naturalidad de las interacciones diseñadas. Tercero, VRSI sostuvo un despliegue con público no técnico, sin reportes de fallos significativos, lo que valida operativamente la conclusión de RQ5 (Sección 6.5) sobre la robustez del framework. La integración de estos tres elementos en un único caso de estudio es, en sí misma, una contribución metodológica: hasta donde alcanza la revisión bibliográfica del artículo IV, ningún estudio previo sobre VR-shopping combina digitalización de productos reales, hand-interaction puro, gamificación narrativa y compra real con dinero propio en una misma sesión de campo.

6.6.8 | Limitaciones reconocidas

Conviene reconocer las limitaciones del estudio. La principal es el tamaño muestral: 48 participantes son suficientes para detectar efectos medios o grandes (como la asociación cluster-compra con $V = 0{,}30$ que se observa) pero insuficientes para detectar efectos pequeños o para análisis estratificados robustos por subgrupos demográficos. Una réplica con $N > 200$ permitiría análisis más granulares y una posible identificación de más de dos clusters. La segunda limitación es el sesgo de autoselección inherente a un evento comercial: los participantes voluntarios pueden estar predispuestos a la novedad tecnológica, lo que sesga la generalización a poblaciones menos curiosas. La tercera limitación es la duración acotada de cada sesión (20 minutos) impuesta por la logística del evento: una sesión más larga podría revelar patrones de

engagement diferentes, en particular efectos de saturación que en 20 minutos no llegan a manifestarse. La cuarta es que el estudio compara dos escenarios pero no contrasta con un grupo de control sin VR (compradores que examinen el producto físico solamente o mediante una web), por lo que el efecto neto atribuible a la VR sobre la compra no se aísla con la limpieza que un diseño con condición control permitiría. Conviene además recordar que el análisis predictivo es explícitamente exploratorio: con 48 muestras y 21 a 29 características, los modelos sirven como prueba de concepto sobre la viabilidad de predecir la compra, no como clasificadores generalizables, y así lo declara el propio artículo IV.

6.6.9 | Síntesis de la respuesta a RQ6

La respuesta a RQ6 es afirmativa y se sostiene sobre evidencia empírica cuantificada en cuatro frentes. Primero, los entornos de VR-shopping que combinan digitalización de bajo coste, interacción gestual natural y narrativa gamificada generan una tasa de conversión a compra real del 35,42 % (18,75 % en productos de la tienda y 16,67 % en la novela gráfica de 20 €), muy por encima de las tasas habituales del e-commerce convencional, en un caso de estudio con 48 participantes, dinero propio y productos físicos. Segundo, los determinantes de la intención de compra no son la novedad de la tecnología, sino el enganche con la gamificación ($d = 1{,}00$) y la usabilidad e inmersión percibidas ($d = 0{,}67$), lo que orienta de forma precisa dónde debe invertir una PYME que adopte estos entornos. Tercero, la conducta observable del usuario identifica de manera no supervisada dos perfiles comportamentales que se asocian con la compra de forma estadísticamente significativa y con tamaño de efecto medio ($\chi^2 = 4{,}37$, $p = 0{,}036$, Cramér's $V = 0{,}30$; tasas del 55,0 % frente al 21,4 %), pese a no diferenciarse en ninguna medida actitudinal del cuestionario. Cuarto, la predicción individual de la compra es viable y mejora de forma monótona al combinar datos comportamentales y actitudinales (AUC de 0,655, 0,702 y 0,849 en los tres modelos). En conjunto, estas conclusiones sostienen la hipótesis general de la tesis sobre el impacto positivo y cuantificable de la VR y la gamificación en el comportamiento de compra real, y cierran de forma coherente el ciclo experimental que las RQ anteriores habilitan.

CAPÍTULO 07

07

Conclusiones y trabajo futuro

Aportaciones de la tesis y líneas abiertas para seguir explorando.

7

Conclusiones y trabajo futuro

Este último capítulo cierra el documento tomando distancia de los resultados individuales para leerlos como un todo. El comercio electrónico se ha consolidado como un canal de venta indispensable, pero su crecimiento acelerado desde el 2020 ha dejado a las PYMEs en cierta desventaja respecto a las grandes empresas, ya que carecen de los recursos necesarios para competir con ellas. La hipótesis de esta tesis sostenía que los sistemas de compra inmersiva basados en digitalización de bajo coste e interacción gestual inclusiva pueden implementarse de forma técnicamente viable para negocios con recursos limitados, que mejoran la experiencia de compra mediante una usabilidad natural cercana al entorno físico, que habilitan la recogida de datos de comportamiento más ricos que los canales web convencionales, y que generan un impacto cuantificable en el comportamiento de compra real, con aplicación directa al pequeño comercio.

Tras responder a las preguntas y revisar la hipótesis, se resume el grado de cumplimiento de los objetivos específicos y se trazan las líneas de trabajo futuro que se desprenden de los resultados.

7.1 | Respuesta a las preguntas de investigación

RQ1: ¿Cómo deben evolucionar los *e-marketplaces* actuales para integrar de forma holística tecnologías inmersivas y de IA, atendiendo a las necesidades específicas de las PYMEs y las economías locales?

Los e-marketplaces actuales no tratan la RV como un añadido ocasional, ni mucho menos un canal de primer nivel. No ofrecen una vía accesible para que las PYMEs digitalicen sus catálogos y operan al margen de las dinámicas comerciales locales. La respuesta que esta tesis sostiene es un modelo arquitectónico formalizado, VR-ZOCO, que (i) trata los entornos VR, móvil y web de manera simétrica, (ii) extiende el catálogo a entidades no estrictamente comerciales para vincular la plataforma con la economía de proximidad, (iii) modela al usuario desde cinco perspectivas comportamentales para alimentar una capa de IA que actúe como planificador y no como recomendador clásico, e (iv) incorpora la digitalización de productos y la accesibilidad como módulos nativos y no como añadidos posteriores. VR-ZOCO se especificó como una quíntupla de doce módulos funcionales; siete de sus trece módulos se han implementado y validado experimentalmente en artículos del compendio. Esta pregunta corresponde al Objetivo O1, respaldado por la publicación arquitectónica [5] y el resto de trabajos mencionados en la Tabla 6.1.

RQ2: ¿Pueden producirse modelos 3D con calidad visual y rendimiento suficientes para su integración en entornos de *VR-shopping* ejecutados en dispositivos comerciales a bajo coste?

Sí, es posible con tecnologías de bajo coste producir modelos 3D utilizables en VR-shopping comercial. De las tecnologías evaluadas, NeRF (a través de Luma AI) en su variante low-poly ofrece el mejor compromiso entre calidad visual percibida, coste de captura y rendimiento sobre headsets actuales. Superó a la fotogrametría y al LiDAR en todas las métricas perceptuales (mejoras medias del 17,78 % y del 22,73 % en NR-3DQA, respectivamente) y alcanzó un error geométrico del 1,4 % respecto al *ground truth* físico, siendo además la opción más accesible, gratuita y ejecutable desde un smartphone. Los modelos

de alta resolución son inviables en modo standalone, pero la versión low-poly cuesta solo en torno al 5 % de calidad perceptual, y un pipeline de optimización adicional sostiene más de 71 FPS en escenas de hasta 96 productos en los tres cascos Meta Quest evaluados. La digitalización de productos por parte de una PYME desde un smartphone comercial es, por tanto, técnicamente viable. Esta pregunta corresponde al Objetivo O2, respaldado por el Artículo I del compendio [1].

RQ3: ¿Cómo debe ser la interacción y navegación en entornos VR-shopping haciendo uso de *hand-tracking* para proporcionar una experiencia cercana a la compra en entornos físicos?

Un conjunto de cinco interacciones con productos (agarre uni y bimanual, examen libre, rescalado, soltar y añadir al carrito), dos mecanismos de navegación (teletransportación gestual y rotación del cuerpo virtual) y una estructura espacial configurable resultó suficiente para sostener una experiencia de VR-shopping cercana a la compra física. Todo ello se empaquetó en el framework de código abierto VRSI. En el plano empírico, los usuarios prefirieron la interacción manual a los controladores (3,5 frente a 2,6 en una escala de cinco puntos), valoraron la inmersión percibida como la dimensión mejor puntuada de todas (4,1) y reportaron niveles de motion sickness consistentemente bajos (por debajo de 2). Los mecanismos resistieron además un despliegue público abierto con 48 participantes no entrenados sin fallos significativos. Esta pregunta corresponde al Objetivo O3, respaldado por el Artículo II [2] y un estudio paralelo de congreso con diez participantes [92].

RQ4: ¿Pueden este tipo de entornos inmersivos, con interacciones basadas en *hand-tracking*, ser accesibles a todo tipo de personas, especialmente a aquellas con limitaciones motoras en miembros superiores?

Los SDK estándar de hand-interaction asumen un control motor fino y, con ello, excluyen a usuarios con condiciones como espasticidad, atrofia mus-

cular o mano en garra. No obstante, se ha demostrado que es posible una adaptación de interacciones basadas en *hand-tracking* para hacer accesibles los entornos inmersivos a personas con limitaciones motoras en miembros superiores. Nuestra respuesta se justifica en la propuesta de un conjunto de cinco mecanismos adaptados (teletransportación por pose mantenida, rotación corporal por movimiento del antebrazo, agarre con umbrales relajados y auto-grasp, *distance grab* por raycasting desde la articulación MCP y liberación por apuntado a una papelera), junto con dos métricas objetivas derivadas del hand-tracking, el *Difficulty Index* y el *Effort Degree*. Validados en el Hospital Nacional de Parapléjicos de Toledo con doce participantes, los mecanismos adaptados elevaron la tasa de completación de tareas al 100 % (frente al 33 % de media sin adaptación), redujeron el Difficulty Index en un 46 % y arrojaron una diferencia altamente significativa en el esfuerzo ($t = -6{,}79$, $p < 10^{-10}$). Las adaptaciones hacen posibles tareas que de otro modo son físicamente imposibles para esta población. Esta pregunta corresponde al Objetivo O4, respaldado por el Artículo III [3].

RQ5: ¿Puede el comportamiento del usuario en este tipo de entornos comerciales ser recopilado de manera no invasiva y analizado para optimizar la experiencia de compra, incrementar las ventas, personalizar la oferta y tomar decisiones estratégicas basadas en datos reales?

Los entornos inmersivos exponen señales que pueden ser recopiladas y analizadas en beneficio del negocio, tal y como se ha demostrado. La respuesta se justifica por el desarrollo de un framework abierto, VRSI, que cubre el ciclo completo: configuración del entorno, recolección no invasiva de veintiún campos por frame (cinemática de manos y cabeza, eye-tracking, eventos de interacción y de navegación) y análisis post hoc estandarizado (métricas espacio-temporales, fijaciones oculares con I-DT y mapas de calor volumétricos). Más allá de enumerar qué se captura, el trabajo muestra que estas señales son conocimiento accionable: cada familia de datos responde a una pregunta concreta del comercio, y los datos demostraron ser capaces de revelar diferencias entre usuarios que los cuestionarios no captaban. El framework fue

revisado por pares como Original Software Publication en SoftwareX, donde se revisa el propio código, y se utilizó en tres estudios [3, 4, 92]. Esta pregunta corresponde al Objetivo O5, respaldado por el Artículo II [2].

RQ6: ¿Influye positivamente la VR, combinada con la gamificación, en el comportamiento de compra real del consumidor?

Sí, de forma cuantificable. En un estudio de campo en un evento comercial abierto al público, con 48 participantes no entrenados que compraban con su propio dinero, la aplicación alcanzó una tasa de conversión a compra real del 35,42 % (18,75 % en productos de la tienda y 16,67 % en la novela gráfica gamificada), muy por encima de las tasas habituales del comercio electrónico convencional. Los determinantes actitudinales de la intención de compra no fueron la novedad de la tecnología, sino el enganche con la gamificación ($d = 1{,}00$) y la usabilidad e inmersión percibidas ($d = 0{,}67$). El análisis no supervisado de la conducta, por sí solo, identificó dos perfiles cuyas tasas de compra diferían de forma significativa (55,0 % frente a 21,4 %; $\chi^2 = 4{,}37$, $p = 0{,}036$, Cramér's $V = 0{,}30$), pese a no diferenciarse en ninguna medida actitudinal. Por último, predecir la compra individual resultó viable y mejoró de forma monótona al combinar datos comportamentales y actitudinales (AUC de 0,655, 0,702 y 0,849 en tres modelos). Esta pregunta corresponde al Objetivo O6, respaldado por el Artículo IV [4].

7.2 | Cumplimiento de los objetivos

Para garantizar la trazabilidad, la Tabla 8.1 relaciona cada objetivo específico con la pregunta de investigación a la que sirve, la publicación que principalmente lo respalda y el grado de cumplimiento alcanzado. Los seis objetivos se cumplen en grado principal; el matiz parcial sobre la adopción por parte del pequeño comercio afecta al objetivo general más que a un objetivo concreto y se discute bajo la respuesta a RQ6 en el Capítulo 6.

Tabla 7.1: **Cumplimiento de los objetivos específicos.** Cada objetivo específico se relaciona con su pregunta de investigación asociada, la publicación que principalmente lo respalda y su grado de cumplimiento.

Obj.	RQ	Publicación principal de respaldo	Cumplimiento
O1	RQ1	Propuesta arquitectónica VR-ZOCO [5]	Principal
O2	RQ2	Artículo I — Evaluación de escaneo 3D de bajo coste [1]	Principal
O3	RQ3	Artículo II — VRSI [2]	Principal
O4	RQ4	Artículo III — Adaptación de interacciones y accesibilidad [3]	Principal
O5	RQ5	Artículo II — VRSI [2]	Principal
O6	RQ6	Artículo IV — Gamificación en VR-shopping [4]	Principal

7.3 | Líneas de trabajo futuro

El trabajo realizado abre dos planos complementarios de continuación: uno con extensiones directas de las publicaciones del compendio, y otro con líneas emergentes que comienzan a tomar forma a partir de las herramientas y la infraestructura ya desarrolladas.

La primera línea es el **estudio de la utilidad de agentes de inteligencia artificial dentro de la arquitectura de VR-ZOCO**. No se trata de una simple extensión del sistema de recomendación y del generador de planes de ocio, sino de integrar IA agéntica a nivel arquitectónico, capaz de asistir al usuario durante la compra tanto en el entorno web como en el VR, sin limitarse a recomendar productos, sino pudiendo interactuar con el propio entorno y generar un acompañamiento social en lugar de una compra solitaria.

La segunda línea es la **generalización de las adaptaciones de accesibilidad** a otras condiciones motoras más allá de la lesión medular: ictus crónico, esclerosis múltiple, parálisis cerebral, secuelas postquirúrgicas. La modularidad del diseño del Artículo III facilita esta extensión, pero requiere una caracterización clínica específica de cada condición y validación con la población correspondiente. La colaboración ya establecida con el Hospital Nacional de Parapléjicos de Toledo proporciona la base para extender el trabajo a otras unidades del propio hospital o a otros centros de referencia. Una de las vías sería un conjunto

de mecanismos de interacción completamente basado en voz para personas con nula movilidad de miembros superiores.

La tercera línea es el **desarrollo de una herramienta de evaluación de *layouts* comerciales en VR**, dirigida a un comerciante que quiera diseñar un espacio de compra, físico o virtual, y comprobar si su distribución es efectiva antes de construirla físicamente. La idea parte de la infraestructura de captura y análisis de comportamiento ya desarrollada en esta tesis: un comerciante podría disponer los productos y las estanterías en un entorno virtual, simular el recorrido de compra y obtener métricas de atención, interacción y navegación sobre esa disposición concreta, todo ello sin el coste de reorganizar físicamente la tienda. Esto convierte la prueba de un *layout* en un experimento reproducible y de bajo coste, y abre la puerta a iterar sobre el diseño del espacio comercial antes de comprometer recursos reales.

8

Conclusions and Future Work

This final chapter closes the document by stepping back from the individual results in order to read them as a whole. E-commerce has become an indispensable sales channel, but its accelerated growth since 2020 has left SMEs at a certain disadvantage with respect to large companies, since they lack the resources needed to compete with them. The hypothesis of this thesis held that immersive shopping systems based on low-cost digitization and inclusive gesture-based interaction can be implemented in a technically feasible way for businesses with limited resources, that they improve the shopping experience through a natural usability close to the physical setting, that they enable the collection of behavioural data richer than that of conventional web channels, and that they produce a measurable impact on real purchasing behaviour, with direct application to small retail.

After answering the questions and revisiting the hypothesis, we summarize the degree of fulfilment of the specific objectives and outline the lines of future work that follow from the results.

8.1 | Answering the research questions

RQ1: How should current *e-marketplaces* evolve to integrate immersive and AI technologies in a holistic way, addressing the specific needs of SMEs and local economies?

Current e-marketplaces treat VR as an occasional add-on, far from a first-class channel. They offer no accessible path for SMEs to digitize their catalogues, and they operate detached from local commercial dynamics. The answer this thesis supports is a formalized architectural model, VR-ZOCO, which (i)

treats the VR, mobile and web environments symmetrically, (ii) extends the catalogue to non-strictly-commercial entities in order to bind the platform to the proximity economy, (iii) models the user from five behavioural perspectives to feed an AI layer that acts as a planner rather than a classic recommender, and (iv) incorporates product digitization and accessibility as native modules rather than later add-ons. VR-ZOCO was specified as a quintuple of twelve functional modules; seven of its thirteen modules have been implemented and experimentally validated in compendium articles. This question corresponds to Objective O1, supported by the architectural publication [5] and the remaining works listed in Table 6.1.

RQ2: Can 3D models be produced with visual quality and performance sufficient for their integration into *VR-shopping* environments running on low-cost commercial devices?

Yes, it is possible to produce 3D models usable in commercial VR-shopping with low-cost technologies. Of the technologies evaluated, NeRF (through Luma AI) in its low-poly variant offers the best trade-off between perceived visual quality, capture cost and performance on current headsets. It outperformed photogrammetry and LiDAR in all perceptual metrics (mean improvements of 17.78 % and 22.73 % in NR-3DQA, respectively) and achieved a geometric error of 1.4 % with respect to the physical *ground truth*, while being the most accessible option, free and runnable from a smartphone. High-resolution models are unfeasible in standalone mode, but the low-poly version costs only around 5 % of perceptual quality, and an additional optimization pipeline sustains more than 71 FPS in scenes of up to 96 products across the three Meta Quest headsets evaluated. Product digitization by an SME from a commercial smartphone is therefore technically feasible. This question corresponds to Objective O2, supported by Paper I of the compendium [1].

RQ3: How should interaction and navigation in VR-shopping environments be designed, using *hand-tracking*, to provide an experience close to shopping in physical settings?

A set of five product interactions (uni- and bimanual grab, free examination, rescaling, release and add-to-cart), two navigation mechanisms (gesture teleportation and virtual-body rotation) and a configurable spatial structure proved sufficient to sustain a VR-shopping experience close to physical shopping. All of this was packaged into the open-source framework VRSI. Empirically, users preferred hand interaction over controllers (3.5 versus 2.6 on a five-point scale), rated perceived immersion as the highest-scoring dimension of all (4.1) and reported consistently low motion sickness (below 2). The mechanisms also withstood an open public deployment with 48 untrained participants without significant failures. This question corresponds to Objective O3, supported by Paper II [2] and a companion conference study with ten participants [92].

RQ4: Can this kind of immersive environment, with *hand-tracking*-based interactions, be made accessible to all kinds of people, especially those with upper-limb motor impairments?

Standard hand-interaction SDKs assume fine motor control and, with it, exclude users with conditions such as spasticity, muscular atrophy or claw hand. Nevertheless, it has been shown that an adaptation of *hand-tracking*-based interactions is possible in order to make immersive environments accessible to people with upper-limb motor impairments. Our answer is grounded in the proposal of a set of five adapted mechanisms (teleportation by sustained pose, body rotation by forearm movement, grasping with relaxed thresholds and auto-grasp, *distance grab* by raycasting from the MCP joint, and release by pointing at a bin), together with two objective metrics derived from hand-tracking, the *Difficulty Index* and the *Effort Degree*. Validated at the Hospital Nacional de Parapléjicos de Toledo with twelve participants, the adapted mechanisms raised the task completion rate to 100 % (from a 33 % average without adaptation), reduced the Difficulty Index by 46 % and yielded a highly

significant difference in effort ($t = -6.79$, $p < 10^{-10}$). The adaptations make possible tasks that are otherwise physically impossible for this population. This question corresponds to Objective O4, supported by Paper III [3].

RQ5: Can user behaviour in this kind of commercial environment be collected non-invasively and analyzed to optimize the shopping experience, increase sales, personalize the offer and make strategic decisions based on real data?

Immersive environments expose signals that can be collected and analyzed for the benefit of the business, as has been demonstrated. The answer is grounded in the development of an open framework, VRSI, that covers the full cycle: environment configuration, non-invasive collection of twenty-one fields per frame (hand and head kinematics, eye-tracking, interaction and navigation events) and standardized post-hoc analysis (spatio-temporal metrics, I-DT eye fixations and volumetric heat-maps). Beyond enumerating what is captured, the work shows that these signals are actionable knowledge: each data family answers a concrete commercial question, and the data proved able to reveal differences between users that questionnaires did not capture. The framework was peer-reviewed as an Original Software Publication in SoftwareX, where the code itself is reviewed, and was used in three studies [3, 4, 92]. This question corresponds to Objective O5, supported by Paper II [2].

RQ6: Do VR and gamification positively influence real consumer purchasing behaviour?

Yes, in a measurable way. In a field study at a public commercial event, with 48 untrained participants buying with their own money, the application achieved a real-purchase conversion rate of 35.42 % (18.75 % for store products and 16.67 % for the gamified graphic novel), well above the typical rates of conventional e-commerce. The attitudinal determinants of purchase intention were not the novelty of the technology, but engagement with the gamification ($d = 1.00$) and perceived usability and immersion ($d = 0.67$). Unsupervised analysis of behaviour, on its own, identified two profiles whose purchase rates

differed significantly (55.0 % versus 21.4 %; $\chi^2 = 4.37$, $p = 0.036$, Cramér's $V = 0.30$), despite not differing on any attitudinal measure. Finally, predicting an individual purchase proved feasible and improved monotonically when behavioural and attitudinal data were combined (AUC of 0.655, 0.702 and 0.849 across three models). This question corresponds to Objective O6, supported by Paper IV [4].

8.2 | Fulfilment of the objectives

For traceability, Table 8.1 relates each specific objective to the research question it serves, the publication that mainly supports it and the degree of fulfilment reached. The six objectives are met at the principal level; the partial nuance regarding adoption by small retail affects the overarching objective rather than any single objective, and is discussed under the answer to RQ6 in Chapter 6.

Tabla 8.1: **Fulfilment of the specific objectives.** Each specific objective is related to its associated research question, the publication that mainly supports it and its degree of fulfilment.

Obj.	RQ	Main supporting publication	Fulfilment
O1	RQ1	VR-ZOCO architectural proposal [5]	Principal
O2	RQ2	Paper I — Evaluation of low-cost 3D scanning [1]	Principal
O3	RQ3	Paper II — VRSI [2]	Principal
O4	RQ4	Paper III — Adaptation of interactions and accessibility [3]	Principal
O5	RQ5	Paper II — VRSI [2]	Principal
O6	RQ6	Paper IV — Gamification in VR-shopping [4]	Principal

8.3 | Lines of future work

The work carried out opens two complementary planes of continuation: one with direct extensions of the compendium publications, and another with

emerging lines that begin to take shape from the tools and infrastructure already developed.

The first line is the **study of the usefulness of artificial-intelligence agents within the VR-ZOCO architecture**. This is not a mere extension of the recommender system and the leisure-plan generator, but the integration of agentic AI at the architectural level, capable of assisting the user during shopping in both the web and the VR environment, not limited to recommending products but able to interact with the environment itself and to generate social accompaniment instead of a solitary purchase.

The second line is the **generalization of the accessibility adaptations** to motor conditions beyond spinal cord injury: chronic stroke, multiple sclerosis, cerebral palsy, post-surgical sequelae. The modularity of the Paper III design facilitates this extension, but it requires a condition-specific clinical characterization of each condition and validation with the corresponding population. The collaboration already established with the Hospital Nacional de Parapléjicos de Toledo provides the basis to extend the work to other units of the hospital itself or to other reference centres. One of the avenues would be a fully voice-based interaction set for people with no upper-limb mobility.

The third line is the **development of a VR-based tool for evaluating commercial layouts**, aimed at a retailer who wants to design a shopping space, whether physical or virtual, and to check whether its arrangement is effective before building it physically. The idea builds on the behaviour capture and analysis infrastructure already developed in this thesis: a retailer could place products and shelves in a virtual environment, simulate the shopping journey and obtain attention, interaction and navigation metrics over that specific arrangement, all without the cost of physically reorganizing the store. This turns the testing of a layout into a reproducible, low-cost experiment, and opens the door to iterating over the design of the commercial space before committing real resources.

Bibliografía

[1] Rubén Grande, Javier Albusac, David Vallejo, Carlos Glez-Morcillo, and José Jesús Castro-Schez. Performance Evaluation and Optimization of 3D Models from Low-Cost 3D Scanning Technologies for Virtual Reality and Metaverse E-Commerce. *Applied Sciences*, 14(14):6037, January 2024. ISSN 2076-3417. doi: 10.3390/app14146037. URL https://www.mdpi.com/2076-3417/14/14/6037. Publisher: Multidisciplinary Digital Publishing Institute.

[2] Rubén Grande, Javier A. Albusac, David Vallejo, Carlos González-Morcillo, Santiago Sánchez-Sobrino, and José J. Castro-Schez. Virtual Reality Shopping-Insights: A data-driven framework to assist the design and development of Virtual Reality shopping environments. *SoftwareX*, 27:101874, September 2024. ISSN 2352-7110. doi: 10.1016/j.softx.2024.101874. URL https://www.sciencedirect.com/science/article/pii/S2352711024002449.

[3] Rubén Grande, Javier Albusac, Vanesa Herrera, Dorothy Monekosso, Ana De Los Reyes-Guzman, David Vallejo, and J. J. Castro-Schez. Enhancing Hand Interactions and Accessibility in Virtual Reality Environments for Users With Motor Disabilities: A Practical Case Study on VR-Shopping. *IEEE Access*, 13:44062–44089, 2025. ISSN 2169-3536. doi: 10.1109/ACCESS.2025.3549527. URL https://ieeexplore.ieee.org/document/10918638.

[4] R. Grande, J. Albusac, S. Sanchez Sobrino, D. Vallejo, and J. .J. Castro-Schez. Gamifying virtual reality shopping experiences: a case study on a real-world scenario. *Virtual Reality*, 30(2):65, February 2026. ISSN 1434-9957. doi: 10.1007/s10055-026-01323-4. URL https://doi.org/10.1007/s10055-026-01323-4.

[5] José Jesús Castro-Schez, Rubén Grande, Vanesa Herrera, Santiago Schez-Sobrino, David Vallejo, and Javier Albusac. E-marketplace state of the art and trends: Vr-zoco—an architectural proposal for the future. *Applied System Innovation*, 7(5):76, 2024. ISSN 2571-5577. doi: 10.3390/asi7050076. URL https://www.mdpi.com/2571-5577/7/5/76.

[6] María Garrido Arcos, Rubén Grande, Sergio Martínez-Cid, Carlos González Morcillo, José Jesús Castro Sánchez, and Javier Albusac. User Activity Monitoring and Personalized Recommendations for Enhancing the VR Shopping Experience. In *2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*, pages 176–184,

January 2024. doi: 10.1109/AIxVR59861.2024.00030. URL https://ieeexplore.ieee.org/document/10445571. ISSN: 2771-7453.

[7] Arman Haji Ahmad, Rizal Fauzi, and Izian Idris. Evolution of Technology and Consumer Behaviour: The Unavoidable Impacts. *Critical Review*, 7:3206–3217, July 2020. doi: 10.31838/jcr.07.19.457.

[8] Dr Naveed Aslam. Consumer Behavior in the Age of E-Commerce. *Journal for Business Research Review*, 1(01):31–40, 2022. ISSN 3006-712X. URL https://journalbrr.com/index.php/Journal/article/view/4.

[9] Dr B.Nagarjuna, Faran Izhar, Dr Anu G. S, Dr.R.Angayarkanni, and Dr.M.Indrapriya. E-COMMERCE EVOLUTION: ANALYZING THE IMPACT OF TECHNOLOGY ON RETAIL STRATEGIES AND CONSUMER BEHAVIOR. *Decision Making: Applications in Management and Engineering*, 7(1):230–242, January 2024. ISSN 2620-0104. URL https://dmame-journals.org/index.php/dmame/article/view/24.

[10] Mahdi Choshin and Ali Ghaffari. An investigation of the impact of effective factors on the success of e-commerce in small- and medium-sized companies. *Computers in Human Behavior*, 66:67–74, January 2017. ISSN 0747-5632. doi: 10.1016/j.chb.2016.09.026. URL https://www.sciencedirect.com/science/article/pii/S0747563216306537.

[11] Pedro Costa and Helena Rodrigues. The ever-changing business of e-commerce-net benefits while designing a new platform for small companies. *Review of Managerial Science*, 18(9):2507–2545, September 2024. ISSN 1863-6691. doi: 10.1007/s11846-023-00681-6. URL https://doi.org/10.1007/s11846-023-00681-6.

[12] Mohamed Reda Soltani. Adapting to the new normal: E-commerce and the evolution of consumer behavior in the pandemic era and beyond. *Journal of Security & Strategic Affairs*, 1(4):28–42, 2024. ISSN 2992-0957. URL https://www.ajol.info/index.php/jssaf/article/view/292209.

[13] Leslie Klieb. Impact of Social Media on Consumer Behaviour. *International Journal of Information and Decision Sciences*, 11, July 2018. doi: 10.1504/IJIDS.2019.10014191.

[14] Jarosław Brodny and Magdalena Tutak. Digitalization of Small and Medium-Sized Enterprises and Economic Growth: Evidence for the EU-27 Countries. *Journal of Open Innovation: Technology, Market, and Complexity*, 8(2):67, June 2022. ISSN 2199-8531. doi: 10.3390/joitmc8020067. URL https://www.mdpi.com/2199-8531/8/2/67. Publisher: Multidisciplinary Digital Publishing Institute.

[15] Agata Marcysiak and Żanna Pleskacz. Determinants of digitization in SMEs. *Entrepreneurship and Sustainability Issues*, 9(1):300–318, September 2021. ISSN 2345-0282. doi: 10.9770/jesi.2021.9.1(18). URL https://jssidoi.org/jesi/article/871.

[16] Norraidah Abu Hasan, Memiyanty Abd Rahim, Saidah Hamizah Ahmad, and Meliza Meliza. Digitization of Business for Small And Medium-Sized Enterprises (SMEs). *Environment-Behaviour Proceedings Journal*, 7(19):11–16, March 2022. ISSN 2398-4287. doi: 10.21834/ebpj.v7i19.3270. URL https://ebpj.e-iph.co.uk/index.php/EBProceedings/article/view/3270.

[17] Schulze Brock P, Katsinis A, Lagüera González J, Di Bella L, Odenthal L, Hell M, Lozar B, and Secades Casino B. Annual report on european smes 2024/2025, sme performance review. Technical Report KJ-01-25-292-EN-N (online), European Union, Luxembourg (Luxembourg), 2025.

[18] Joris Beckers, Simon Weekx, Philippe Beutels, and Ann Verhetsel. COVID-19 and retail: The catalyst for e-commerce in Belgium? *Journal of Retailing and Consumer Services*, 62: 102645, September 2021. ISSN 0969-6989. doi: 10.1016/j.jretconser.2021.102645. URL https://www.sciencedirect.com/science/article/pii/S0969698921002113.

[19] Tomoya Kawasaki, Hisayuki Wakashima, and Ryuichi Shibasaki. The use of e-commerce and the COVID-19 outbreak: A panel data analysis in Japan. *Transport Policy*, 115:88–100, January 2022. ISSN 0967-070X. doi: 10.1016/j.tranpol.2021.10.023. URL https://www.sciencedirect.com/science/article/pii/S0967070X2100305X.

[20] Dorel Mihai Paraschiv, Emilia Titan, Manea Daniela Ioana, Ionescu Crina-Dana, Mihaela Mihai, and Octavian Serban. The change in e-commerce in the context of the Coronavirus pandemic. *Management & Marketing. Challenges for the Knowledge Society*, 17(2):220–233, June 2022. doi: 10.2478/mmcks-2022-0012.

[21] Sukisno S. Riadi, Ariesta Heksarini, Dirga Lestari, Siti Maria, Saida Zainurossalamia Za, and Rizky Yudaruddin. The Benefits of e-Commerce before and during the Covid-19 Pandemic for Small Enterprises in Indonesia. *WSEAS TRANSACTIONS ON ENVIRONMENT AND DEVELOPMENT*, 18:69–79, January 2022. doi: 10.37394/232015.2022.18.8.

[22] Timothy Yaw Acheampong. Impact of Covid-19 on e-Commerce in the European Union. *ENTRENOVA - ENTerprise REsearch InNOVAtion Journal*, 7(1):86–95, 2021. ISSN 2706-4735. doi: 10.54820/VPRD4547. URL https://ojs.srce.hr/entrenova/article/view/20239.

[23] R.Y. Kim. Retail After COVID-19: Use Virtual Reality to Enhance Ecommerce. In *2022 IEEE Technology and Engineering Management Conference (TEMSCON EUROPE)*, pages 118–123. IEEE, 2022. ISBN 978-1-66548-313-1. doi: 10.1109/TEMSCONEUROPE54743.2022.9801972.

[24] Maria Rybaczewska and Leigh Sparks. Ageing consumers and e-commerce activities. *Ageing & Society*, 42(8):1879–1898, August 2022. ISSN 0144-686X, 1469-1779. doi: 10.1017/S0144686X20001932. URL https://www.cambridge.org/core/journals/ageing-and-society/article/ageing-consumers-and-ecommerce-activities/A839050DDBF817172F8FA81FB35155ED.

[25] Wala Alkarney and Riyad Almakki. Factors Affecting the Intention to Use Virtual Stores: Perspectives of Consumers in Saudi Arabia. *Mobile Information Systems*, 2022(1):8340406, 2022. ISSN 1875-905X. doi: 10.1155/2022/8340406. URL https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/8340406. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2022/8340406.

[26] Jun-Phil Uhm, Sanghoon Kim, Chanwook Do, and Hyun-Woo Lee. How augmented reality (AR) experience affects purchase intention in sport E-commerce: Roles of perceived diagnosticity, psychological distance, and perceived risks. *Journal of Retailing and Consumer Services*, 67:103027, July 2022. ISSN 0969-6989. doi: 10.1016/j.jretconser.2022.103027. URL https://www.sciencedirect.com/science/article/pii/S0969698922001205.

[27] Shobhit Kakaria, Farzad Saffari, Thomas Z. Ramsøy, and Enrique Bigné. Cognitive load during planned and unplanned virtual shopping: Evidence from a neurophysiological perspective. *International Journal of Information Management*, 72:102667, October 2023. ISSN 0268-4012. doi: 10.1016/j.ijinfomgt.2023.102667. URL https://www.sciencedirect.com/science/article/pii/S0268401223000488.

[28] Jaikishan Khatri, Javier Marín-Morales, Masoud Moghaddasi, Jaime Guixeres, Irene Alice Chicchi Giglioli, and Mariano Alcañiz. Recognizing Personality Traits Using Consumer Behavior Patterns in a Virtual Retail Store. *Frontiers in Psychology*, 13, March 2022. ISSN 1664-1078. doi: 10.3389/fpsyg.2022.752073. URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.752073/full. Publisher: Frontiers.

[29] Cristina Gil-López, Jaime Guixeres, Masoud Moghaddasi, Jaikishan Khatri, Javier Marín-Morales, and Mariano Alcañiz Raya. Recognizing shopper demographics from behavioral responses in a virtual reality store. *Virtual Reality*, 27(3):1937–1966, March 2023. doi: 10.1007/s10055-023-00767-2.

[30] Virginie Lavoye, Anssi Tarkiainen, Jenni Sipilä, and Joel Mero. More than skin-deep: The influence of presence dimensions on purchase intentions in augmented reality shopping. *Journal of Business Research*, 169(C):114247, 2023. doi: 10.1016/j.jbusres.2023.114247. URL https://ideas.repec.org//a/eee/jbrese/v169y2023ics0148296323006069.html. Publisher: Elsevier.

[31] Cindy Lombart, Elena Millan, Jean-Marie Normand, Adrien Verhulst, Blandine Labbé-Pinlon, and Guillaume Moreau. Effects of physical, non-immersive virtual, and immersive virtual store environments on consumers' perceptions and purchase behavior. *Computers in Human Behavior*, 110:106374, September 2020. ISSN 0747-5632. doi: 10.1016/j.chb.2020.106374. URL https://www.sciencedirect.com/science/article/pii/S0747563220301278.

[32] Sang-Lin Han, Myounga An, Jerry J. Han, and Jiyoung Lee. Telepresence, time distortion, and consumer traits of virtual reality shopping. *Journal of Business Research*, 118:311–320, September 2020. ISSN 0148-2963. doi: 10.1016/j.jbusres.2020.06.056. URL https://www.sciencedirect.com/science/article/pii/S0148296320304318.

[33] Nannan Xi and Juho Hamari. Shopping in virtual reality: A literature review and future agenda. *Journal of Business Research*, 134:37–58, September 2021. ISSN 0148-2963. doi: 10.1016/j.jbusres.2021.04.075. URL https://www.sciencedirect.com/science/article/pii/S0148296321003209.

[34] Marina Ricci, Alessandro Evangelista, Annalisa Di Roma, and Michele Fiorentino. Immersive and desktop virtual reality in virtual fashion stores: a comparison between shopping experiences. *Virtual Reality*, 27(3):2281–2296, May 2023. ISSN 1434-9957. doi: 10.1007/s10055-023-00806-y. URL https://doi.org/10.1007/s10055-023-00806-y.

[35] S. Wang and N. P. Archer. Electronic marketplace definition and classification: literature review and clarifications. *Enterprise Information Systems*, 1(1):89–112, February 2007. ISSN 1751-7575. doi: 10.1080/17517570601088380. URL https://doi.org/10.1080/17517570601088380. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17517570601088380.

[36] Christopher P. Holland and Manuela Gutiérrez-Leefmans. A Taxonomy of SME E-Commerce Platforms Derived from a Market-Level Analysis. *International Journal of Electronic Commerce*, 22(2):161–201, April 2018. ISSN 1086-4415. doi: 10.1080/10864415.2017.1364114. URL https://doi.org/10.1080/10864415.2017.1364114. Publisher: Routledge _eprint: https://doi.org/10.1080/10864415.2017.1364114.

[37] Hepu Deng, Sophia Xiaoxia Duan, and Feng Luo. Critical determinants for electronic market adoption: Evidence from Australian small- and medium-sized enterprises. *Journal of Enterprise Information Management*, 33(2):335–352, November 2019. ISSN 1741-0398. doi: 10.1108/JEIM-04-2019-0106. URL https://doi.org/10.1108/JEIM-04-2019-0106.

[38] Jose Alejandro Cano, Abraham Londoño-Pineda, Maria Fanny Castro, Hugo Bécquer Paz, Carolina Rodas, and Tatiana Arias. A Bibliometric Analysis and Systematic Review on E-Marketplaces, Open Innovation, and Sustainability. *Sustainability*, 14(9):5456, January 2022. ISSN 2071-1050. doi: 10.3390/su14095456. URL https://www.mdpi.com/2071-1050/14/9/5456. Publisher: Multidisciplinary Digital Publishing Institute.

[39] Carlos Moreno, Zaheer Allam, Didier Chabaud, Catherine Gall, and Florent Pratlong. Introducing the "15-Minute City": Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities. *Smart Cities*, 4(1):93–111, March 2021. ISSN 2624-6511. doi: 10.3390/smartcities4010006. URL https://www.mdpi.com/2624-6511/4/1/6. Publisher: Multidisciplinary Digital Publishing Institute.

[40] Steven Michael LaValle. *Virtual Reality*. Cambridge University Press, 2023. ISBN 978-1-108-18287-4. Google-Books-ID: patE0AEACAAJ.

[41] Yeonju Jang and Eunil Park. An adoption model for virtual reality games: The roles of presence and enjoyment. *Telematics and Informatics*, 42:101239, September 2019. ISSN 0736-5853. doi: 10.1016/j.tele.2019.101239. URL https://www.sciencedirect.com/science/article/pii/S0736585319301315.

[42] Anouk de Regt, Kirk Plangger, and Stuart J. Barnes. Virtual reality marketing and customer advocacy: Transforming experiences from story-telling to story-doing. *Journal of Business Research*, 136:513–522, November 2021. ISSN 0148-2963. doi: 10.1016/j.jbusres.2021.08.004. URL https://www.sciencedirect.com/science/article/pii/S0148296321005579.

[43] Wai Han Lo and Ka Lun Benjamin Cheng. Does virtual reality attract visitors? The mediating effect of presence on consumer response in virtual reality tourism advertising. *Information Technology & Tourism*, 22(4):537–562, December 2020. ISSN 1943-4294. doi: 10.1007/s40558-020-00190-2. URL https://doi.org/10.1007/s40558-020-00190-2.

[44] Marie-Stéphanie Bracq, Estelle Michinov, and Pierre Jannin. Virtual Reality Simulation in Nontechnical Skills Training for Healthcare Professionals: A Systematic Review. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, 14(3):188–194, June 2019. ISSN 1559-713X. doi: 10.1097/SIH.0000000000000347.

[45] Dieter Lerner, Stefan Mohr, Jonas Schild, Martin Göring, and Thomas Luiz. An Immersive Multi-User Virtual Reality for Emergency Simulation Training: Usability Study. *JMIR serious games*, 8(3):e18822, July 2020. ISSN 2291-9279. doi: 10.2196/18822.

[46] Christian Peukert, Jella Pfeiffer, Martin Meißner, Thies Pfeiffer, and Christof Weinhardt. Shopping in Virtual Reality Stores: The Influence of Immersion on System Adoption. *Journal of Management Information Systems*, 36(3):755–788, July 2019. ISSN 0742-1222. doi: 10.1080/07421222.2019.1628889. URL https://doi.org/10.1080/07421222.2019.1628889.

[47] Nannan Xi, Juan Chen, Filipe Gama, Henry Korkeila, and Juho Hamari. Acceptance of the metaverse: a laboratory experiment on augmented and virtual reality shopping. *Internet Research*, 34(7):82–117, January 2024. ISSN 1066-2243. doi: 10.1108/INTR-05-2022-0334. URL https://doi.org/10.1108/INTR-05-2022-0334.

[48] Marco Speicher, Sebastian Cucerca, and Antonio Krüger. VRShop: A Mobile Interactive Virtual Reality Shopping Environment Combining the Benefits of On- and Offline Shopping. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):102:1–102:31, September 2017. doi: 10.1145/3130967. URL https://doi.org/10.1145/3130967.

[49] Marco Speicher, Philip Hell, Florian Daiber, Adalberto Simeone, and Antonio Krüger. A virtual reality shopping experience using the apartment metaphor. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, AVI '18, pages 1–9, New York, NY, USA, May 2018. Association for Computing Machinery. ISBN 978-1-4503-5616-9. doi: 10.1145/3206505.3206518. URL https://doi.org/10.1145/3206505.3206518.

[50] Erica van Herpen, Eva van den Broek, Hans C. M. van Trijp, and Tian Yu. Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. *Appetite*, 107:196–207, December 2016. ISSN 1095-8304. doi: 10.1016/j.appet.2016.07.033.

[51] Lorenzo Donatiello, Elena Morotti, Gustavo Marfia, and Sonia Di Vaio. Exploiting Immersive Virtual Reality for Fashion Gamification. In *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 17–21, Bologna, Italy, September 2018. IEEE Press. doi: 10.1109/PIMRC.2018.8581036. URL https://doi.org/10.1109/PIMRC.2018.8581036.

[52] Gabriele Pizzi, Daniele Scarpi, Marco Pichierri, and Virginia Vannucci. Virtual reality, real reactions?: Comparing consumers' perceptions and shopping orientation across physical and virtual-reality retail stores. *Computers in Human Behavior*, 96:1–12, July 2019. ISSN 0747-5632. doi: 10.1016/j.chb.2019.02.008. URL https://www.sciencedirect.com/science/article/pii/S0747563219300603.

[53] Hyo Jeong Kang, Jung-hye Shin, and Kevin Ponto. How 3D Virtual Reality Stores Can Shape Consumer Purchase Decisions: The Roles of Informativeness and Playfulness. *Journal of Interactive Marketing*, 49(1):70–85, February 2020. ISSN 1094-9968. doi: 10.1016/j.intmar.2019.07.002. URL https://www.sciencedirect.com/science/article/pii/S1094996819300921.

[54] Monica Cortinas, Carmen Berne, Raquel Chocarro, Frode Nilssen, and Natalia Rubio. Editorial: The Impact of AI-Enabled Technologies in E-commerce and Omnichannel Retailing. *Frontiers in Psychology*, 12, July 2021. ISSN 1664-1078. doi: 10.3389/fpsyg.2021.718885. URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.718885/full. Publisher: Frontiers.

[55] Hossam El-din Fawzy. THE ACCURACY OF MOBILE PHONE CAMERA INSTEAD OF HIGH RESOLUTION CAMERA IN DIGITAL CLOSE RANGE PHOTOGRAMMETRY. 2015. URL https://www.semanticscholar.org/paper/THE-ACCURACY-OF-MOBILE-PHONE-CAMERA-INSTEAD-OF-HIGH-Fawzy/6d69d9f4a1799139dad07542f8bc36427fbcefa0.

[56] Gregor Luetzenburg, Aart Kroon, and Anders A. Bjørk. Evaluation of the Apple iPhone 12 Pro LiDAR for an Application in Geosciences. *Scientific Reports*, 11(1):22221, November 2021. ISSN 2045-2322. doi: 10.1038/s41598-021-01763-9. URL https://www.nature.com/articles/s41598-021-01763-9. Publisher: Nature Publishing Group.

[57] Seth Miller, Alireza Hashemian, Robert Gillihan, and Ethan Helms. *A Comparison of Mobile Phone LiDAR Capture and Established Ground based 3D Scanning Methodologies*. March 2022. doi: 10.4271/2022-01-0832.

[58] Kyle Gao, Yina Gao, Hongjie He, Dening Lu, Linlin Xu, and Jonathan Li. NeRF: Neural Radiance Field in 3D Vision, A Comprehensive Review. arXiv preprint arXiv:2210.00379, May 2023. URL http://arxiv.org/abs/2210.00379.

[59] Nianchen Deng, Zhenyi He, Jiannan Ye, Budmonde Duinkharjav, Praneeth Chakravarthula, Xubo Yang, and Qi Sun. FoV-NeRF: Foveated Neural Radiance Fields for Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3854–3864, November 2022. ISSN 1941-0506. doi: 10.1109/TVCG.2022.3203102. URL https://ieeexplore.ieee.org/document/9872532.

[60] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4):1–14, July 2023. doi: 10.1145/3592433. URL https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/.

[61] Antonios Kratimenos et al. Vrsplat: Fast and robust gaussian splatting for virtual reality. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 8(1), 2025. doi: 10.1145/3728311. URL https://doi.org/10.1145/3728311.

[62] Yulia Skrupskaya, Valeriia Skibina, Victor Taratoukhine, and Elvira Kozlova. The Use of Virtual Reality to Drive Innovations. VRE-IP Experiment. pages 336–345. January 2022. ISBN 978-3-030-95493-2. doi: 10.1007/978-3-030-95494-9_28.

[63] Guillaume Lavoué. A Multiscale Metric for 3D Mesh Visual Quality Assessment. *Comput. Graph. Forum*, 30(5):1427–1437, August 2011. doi: 10.1111/j.1467-8659.2011.02017.x.

[64] Zicheng Zhang, Wei Sun, Xiongkuo Min, Tao Wang, Wei Lu, and Guangtao Zhai. No-Reference Quality Assessment for 3D Colored Point Cloud and Mesh Models. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(11):7618–7631, November 2022. ISSN 1051-8215, 1558-2205. doi: 10.1109/TCSVT.2022.3186894. URL http://arxiv.org/abs/2107.02041. arXiv:2107.02041 [cs].

[65] Zicheng Zhang, Wei Sun, Xiongkuo Min, Quan Zhou, Jun He, Qiyuan Wang, and Guangtao Zhai. MM-PCQA: Multi-Modal Learning for No-reference Point Cloud Quality Assessment. arXiv preprint arXiv:2209.00244, April 2023. URL http://arxiv.org/abs/2209.00244.

[66] Michel Wedel, Enrique Bigné, and Jie Zhang. Virtual and augmented reality: Advancing research in consumer marketing. *International Journal of Research in Marketing*, 37(3):443–465, September 2020. ISSN 0167-8116. doi: 10.1016/j.ijresmar.2020.04.004. URL https://www.sciencedirect.com/science/article/pii/S0167811620300380.

[67] Masoud Moghaddasi, Javier Marín-Morales, Jaikishan Khatri, Jaime Guixeres, Irene Alice Chicchi Giglioli, and Mariano Alcañiz. Recognition of Customers' Impulsivity from Behavioral Patterns in Virtual Reality. *Applied Sciences*, 11(10):4399, January 2021. ISSN 2076-3417. doi: 10.3390/app11104399. URL https://www.mdpi.com/2076-3417/11/10/4399. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.

[68] Lin Liwan, Wu Zongyu, Yijun Lu, Zhong Chen, and Guo Weijie. Recent progress on eye-tracking and gaze estimation for ar/vr applications: A review. *Electronics*, 14(17):3352, 2025. doi: 10.3390/electronics14173352. URL https://www.proquest.com/scholarly-journals/recent-progress-on-eye-tracking-gaze-estimation/docview/3249684526/se-2. Copyright - © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License; Last updated - 2025-09-15.

[69] Masoud Moghaddasi, Jaikishan Khatri, Jose Llanes-Jurado, Luciano Spinella, Javier Marín-Morales, Jaime Guixeres, and Mariano Alcañiz. Segmentation of Areas of Interest Inside a Virtual Reality Store. In Constantine Stephanidis and Margherita Antona, editors, *HCI International 2020 - Posters*, pages 92–98, Cham, 2020. Springer International Publishing. ISBN 978-3-030-50729-9. doi: 10.1007/978-3-030-50729-9_13.

[70] Jose Llanes-Jurado, Javier Marín-Morales, Jaime Guixeres, and Mariano Alcañiz. Development and Calibration of an Eye-Tracking Fixation Identification Algorithm for Immersive Virtual Reality. *Sensors*, 20(17):4956, January 2020. ISSN 1424-8220. doi: 10.3390/s20174956. URL `https://www.mdpi.com/1424-8220/20/17/4956`. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute.

[71] Alaa Elboudali, Améziane Aoussat, Fabrice Mantelet, Julien Bethomier, and Florian Leray. A customised virtual reality shopping experience framework based on consumer behaviour: 3DR3CO. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 14(2):551–563, June 2020. ISSN 1955-2505. doi: 10.1007/s12008-020-00645-0. URL `https://doi.org/10.1007/s12008-020-00645-0`.

[72] Kuo-Wei Su, Shih-Chih Chen, Po-Hung Lin, and Ching-I Hsieh. Evaluating the user interface and experience of VR in the electronic commerce environment: a hybrid approach. *Virtual Reality*, 24(2):241–254, June 2020. ISSN 1434-9957. doi: 10.1007/s10055-019-00394-w. URL `https://doi.org/10.1007/s10055-019-00394-w`.

[73] Ron Kohavi, Diane Tang, and Ya Xu. *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing*. March 2020. ISBN 978-1-108-72426-5. doi: 10.1017/9781108653985.

[74] Mustafa Ayobami Raji, Hameedat Bukola Olodo, Timothy Tolulope Oke, Wilhelmina Afua Addy, Onyeka Chrisanctus Ofodile, and Adedoyin Tolulope Oyewole. BUSINESS STRATEGIES IN VIRTUAL REALITY: A REVIEW OF MARKET OPPORTUNITIES AND CONSUMER EXPERIENCE. *International Journal of Management & Entrepreneurship Research*, 6(3):722–736, March 2024. ISSN 2664-3596, 2664-3588. doi: 10.51594/ijmer.v6i3.883. URL `https://fepbl.com/index.php/ijmer/article/view/883`.

[75] Nima Jamalian, Marco Gillies, Frederic Leymarie, and Xueni Pan. The effects of hand tracking on user performance: an experimental study of an object selection based memory game. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 768–776, 2022. doi: 10.1109/ISMAR55827.2022.00095.

[76] Asim Hameed, Andrew Perkis, and Sebastian Möller. Evaluating hand-tracking interaction for performing motor-tasks in VR learning environments. In *13th International Conference on Quality of Multimedia Experience (QoMEX)*, pages 219–224, 2021. doi: 10.1109/QoMEX51781.2021.9465407.

[77] Alexander Masurovsky, Paul Chojecki, Detlef Runde, Mustafa Lafci, David Przewozny, and Michael Gaebler. Controller-Free Hand Tracking for Grab-and-Place Tasks in Immersive Virtual Reality: Design Elements and Their Empirical Study. *Multimodal Technologies and Interaction*, 4(4):91, December 2020. doi: 10.3390/mti4040091.

[78] Huiyue Wu, Yu Wang, Jiali Qiu, Jiayi Liu, and Xiaolong (Luke) Zhang. User-defined gesture interaction for immersive VR shopping applications. *Behaviour & Information Technology*, 38(7):726–741, July 2019. ISSN 0144-929X. doi: 10.1080/0144929X.2018.1552313. URL https://doi.org/10.1080/0144929X.2018.1552313.

[79] Xiaolong Lou, Xiangdong Li, Preben Hansen, and Peng Du. Hand-adaptive user interface: improved gestural interaction in virtual reality. *Virtual Reality*, 25(2):367–382, June 2021. doi: 10.1007/s10055-020-00461-7.

[80] John Dudley, Lulu Yin, Vanja Garaj, and Per Ola Kristensson. Inclusive Immersion: a review of efforts to improve accessibility in virtual reality, augmented reality and the metaverse. *Virtual Reality*, 27(4):2989–3020, December 2023. ISSN 1434-9957. doi: 10.1007/s10055-023-00850-8. URL https://doi.org/10.1007/s10055-023-00850-8.

[81] Matt Dombrowski, Peter A. Smith, Albert Manero, and John Sparkman. Designing Inclusive Virtual Reality Experiences. In Jessie Y.C. Chen and Gino Fragomeni, editors, *Virtual, Augmented and Mixed Reality. Multimodal Interaction*, pages 33–43, Cham, 2019. Springer International Publishing. ISBN 978-3-030-21607-8. doi: 10.1007/978-3-030-21607-8_3.

[82] Vanesa Herrera, David Vallejo, José J. Castro-Schez, Dorothy N. Monekosso, Ana de los Reyes, Carlos Glez-Morcillo, and Javier Albusac. Rehab-Immersive: A framework to support the development of virtual reality applications in upper limb rehabilitation. *SoftwareX*, 23:101412, July 2023. ISSN 2352-7110. doi: 10.1016/j.softx.2023.101412. URL https://www.sciencedirect.com/science/article/pii/S2352711023001085.

[83] John Brooke. SUS: A quick and dirty usability scale. *Usability Evaluation in Industry*, 189: 4–7, 1996. doi: 10.1201/9781498710411-35.

[84] Sandra G. Hart and Lowell E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Peter A. Hancock and Najmedin Meshkati, editors, *Human Mental Workload*, volume 52 of *Advances in Psychology*, pages 139–183. North-Holland, January 1988. doi: 10.1016/S0166-4115(08)62386-9. URL https://www.sciencedirect.com/science/chapter/bookseries/abs/pii/S0166411508623869. ISSN: 0166-4115.

[85] Muhamad Adhytia Wana Putra Rahmadhan, Dana Sensuse, Ryan Suryono, and Kautsarina Kautsarina. Trends and Applications of Gamification in E-Commerce: A Systematic Literature Review. *Journal of Information Systems Engineering and Business Intelligence*, 9(1): 28–37, April 2023. doi: 10.20473/jisebi.9.1.28-37.

[86] Manuela Aparicio, Carlos J. Costa, and Rafael Moises. Gamification and reputation: key determinants of e-commerce usage and repurchase intention. *Heliyon*, 7(3):e06383, March 2021. ISSN 2405-8440. doi: 10.1016/j.heliyon.2021.e06383. URL https://www.sciencedirect.com/science/article/pii/S2405844021004886.

[87] Jengchung Chen, Quang An Ha, and Minh Vu. The Influences of Virtual Reality Shopping Characteristics on Consumers' Impulse Buying Behavior. *International Journal of Human-Computer Interaction*, 39(17):3473–3491, July 2022. doi: 10.1080/10447318.2022.2098566.

[88] Zhenda Wei, Jingru Zhang, Xiaoting Huang, and Hanqin Qiu. Can gamification improve the virtual reality tourism experience? Analyzing the mediating role of tourism fatigue. *Tourism Management*, 96:104715, June 2023. doi: 10.1016/j.tourman.2022.104715.

[89] Ayman Alzayat and Seung Lee. Virtual products as an extension of my body: Exploring hedonic and utilitarian shopping value in a virtual reality retail environment. *Journal of Business Research*, 130:348–363, June 2021. doi: 10.1016/j.jbusres.2021.03.017.

[90] Rubén Grande, Santiago Sánchez-Sobrino, David Vallejo, José Jesús Castro-Schez, and Javier Albusac. Supporting Small Businesses and Local Economies Through Virtual Reality Shopping and Artificial Intelligence: A Position Paper. In *Proceedings of the 25th International Conference on Enterprise Information Systems (ICEIS 2023)*, pages 312–319. SciTePress, May 2023. ISBN 978-989-758-648-4. doi: 10.5220/0011964600003467. URL `https://www.scitepress.org/PublicationsDetail.aspx?ID=1RZg1pY7NjM=&t=1`.

[91] Rubén Grande, J. Albusac, J.J Castro-Schez, D. Vallejo, and S. Sánchez-Sobrino. A Virtual Reality Shopping platform for enhancing e-commerce activities of small businesses and local economies. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 793–796, October 2023. doi: 10.1109/ISMAR-Adjunct 60411.2023.00175. URL `https://ieeexplore.ieee.org/document/10322220`. ISSN: 2771-1110.

[92] Rubén Grande, Javier Albusac, Santiago Sánchez-Sobrino, David Vallejo, José Jesús Castro-Schez, and Carlos González. Exploring Interaction Mechanisms and Perceived Realism in Different Virtual Reality Shopping Setups. In *Proceedings of the 26th International Conference on Enterprise Information Systems (ICEIS 2024) - Volume 2*, pages 505–512. SciTePress, March 2024. ISBN 978-989-758-692-7. doi: 10.5220/0012626000003690. URL `https://www.scitepress.org/Link.aspx?doi=10.5220/0012626000003690`.

[93] María Garrido, José Jesús Castro-Schez, Rubén Grande, Santiago Schez-Sobrino, and David Vallejo. Enhancement of the Online Presence of Small and Medium Sized Enterprises with Minimum Impact on Traditional Business Activities in Towns and Cities. In *Proceedings of the 26th International Conference on Enterprise Information Systems (ICEIS 2024)*, pages 989–996. SciTePress, March 2024. ISBN 978-989-758-692-7. doi: 10.5220/0012630600003690. URL `https://www.scitepress.org/Link.aspx?doi=10.5220/0012630600003690`.

[94] Javier Albusac, Diego Cordero, Mario Jiménez, Rubén Grande, Vanesa Herrera, Raquel Perales, M. Felix, and Ana Reyes. VR-ADAPT: An Immersive Learning and Training Environment for Wheelchair Users with Recent Spinal Cord Injuries. In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025)*, pages 573–580.

SciTePress, March 2026. ISBN 978-989-758-749-8. doi: 10.5220/0013352900003929. URL https://www.scitepress.org/Link.aspx?doi=10.5220/0013352900003929.

[95] Ruben Grande, Diego Cordero, David Vallejo, Carlos González, Santiango Schez-Sobrino, Jose Castro-Schez, and Javier Albusac. Pilot Study on the Effects of Gamification and Virtual Reality on the Shopping Experience. In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025)*, pages 482–493. SciTePress, March 2025. ISBN 978-989-758-749-8. doi: 10.5220/0013423500003929. URL https://www.scitepress.org/Link.aspx?doi=10.5220/0013423500003929.

[96] Rafael Echevarría, Rubén Grande, Vanesa Herrera, Ana de los Reyes, and Javier Albusac. Occlusion-resilient mixed reality: A smart object architecture for upper limb rehabilitation. In *International Conference on eXtended Reality (XR Salento)*, Salento, Italy, 2026.

[97] Rubén Grande, Vanesa Herrera, Carlos Glez-Morcillo, Ana de los Reyes, José J. Castro-Schez, and Javier Albusac. Adaptation of Interaction Mechanisms in Virtual Reality Shopping Environments for People with Upper Limb Motor Difficulties. In *Spanish Computer Graphics Conference (CEIG 2024), Congreso Español de Informática (CEDI 2024)*. The Eurographics Association, 2024. ISBN 978-3-03868-261-5. URL https://doi.org/10.2312/ceig.20241143.

[98] D. Cordero, R. Grande, V. Herrera, M. Jiménez, and J. Albusac. VR-ADAPT: A virtual reality environment for skill development in new wheelchair users with spinal cord injuries. In *VII Congreso XoveTIC 2025*, pages 287–295. Universidade da Coruña, 2025. doi: 10.17979/spu.23.c47.

[99] Rubén Grande, Santiago Sánchez-Sobrino, David Vallejo, José Jesús Castro-Schez, and Javier A. Albusac. Narrowing the Technological Gap by Promoting Small Commerce Through VR and AI for a Lifelike E-Commerce Experience: Needs and Solutions. In Joaquim Filipe, Michał Śmiałek, Alexander Brodsky, and Slimane Hammoudi, editors, *Enterprise Information Systems*, pages 92–112, Cham, 2024. Springer Nature Switzerland. ISBN 978-3-031-64755-0. doi: 10.1007/978-3-031-64755-0_5.

[100] Charlotte Croucher, Wendy Powell, Brett Stevens, Matt Miller-Dicks, Vaughan Powell, Travis J Wiltshire, and Pieter Spronck. Locomote–a framework for classification of natural locomotion in vr by task, technique and modality. *IEEE transactions on visualization and computer graphics*, 30(8):5765–5781, 2023. doi: 10.1109/tvcg.2023.3313439/mm1.