UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Mixed Realities and Product Perception: Experimental analysis of the relationship between visual quality and interaction, and the user's emotional and perceptual response.

Author: Almudena Palacios Ibáñez
Advisors: Manuel Contero González
Jorge Dorribo Camba

Septiembre, 2023

Esta tesis doctoral se ha realizado bajo la financiación del Ministerio de Universidades e Investigación con una beca de Formación de Profesorado Universitario (FPU 2019/03878)

-

This doctoral thesis has been carried out under the funding of the Ministry of Universities and Research with a University Teacher Training grant (FPU 2019/03878).

Agradecimientos

Esta tesis doctoral es la primera semilla que planto en el terreno de la investigación con la esperanza de que florezca como nuevas y emocionantes formas de conocimiento. Más allá de todo lo que he logrado académicamente, esta tesis me ha regalado un respiro en medio de la complejidad de la vida y un espacio para crecer y desarrollarme como persona. He tenido la suerte de contar con personas increíbles durante este arduo camino, y por ello quiero dedicarles unas palabras de agradecimiento.

A mi tutor, Manuel. Siempre he creído que, en el camino de la vida, tener a alguien que actúe como una brújula puede marcar la diferencia entre perderse y llegar a puerto seguro. En este sentido, has sido para mí un verdadero faro. Gracias a tu sabiduría, dedicación y apoyo incondicional, por creer y confiar en mí desde el momento inicial, por alentarme en los momentos difíciles y por brindarme la oportunidad de trabajar a tu lado. Tu presencia en mi vida ha sido un regalo inestimable que recordaré siempre.

A mi cotutor, Jorge. No habría sido posible llegar hasta aquí sin tu ayuda y apoyo. Además, contar contigo durante mi estancia en Estados Unidos hizo que todo fuera mucho más llevadero. Espero que podamos seguir colaborando en el futuro y que esta amistad y relación de trabajo perdure por mucho tiempo.

A mis padres, quienes han creído en mi desde el inicio de mi carrera académica hasta el final de este proyecto de investigación. Me enseñaron el valor del esfuerzo, la perseverancia y el compromiso, y gracias a ellos hoy tengo la dicha de ver este sueño hecho realidad.

A Belén, quien siempre ha celebrado mis logros como suyos. Eres un verdadero ejemplo de lo que significa ser una hermana y espero que nuestra relación siga fortaleciéndose con los años.

A Melanie, mi confidente. Gracias por tu paciencia durante los momentos de estrés, por tus ánimos y tus palabras de aliento, por tu apoyo incondicional. Eres un verdadero pilar en mi vida y un apoyo inquebrantable en los momentos difíciles. Gracias por ser mi compañera de vida.

A las amigas y familiares que siempre están a mi lado y que hacen que siempre me sienta acompañada.

Acknowledgements

This doctoral thesis is the first seed I sow in the field of research, with the hope that it will bloom as new and exciting forms of knowledge. Beyond all that I have achieved academically, this thesis has given me a break in the midst of life's complexity, a space to grow and develop as a person. I have been fortunate to have incredible people by my side during this difficult journey, and I want to dedicate a few words of thanks to them.

To my advisor, Manuel. I have always believed that, in the journey of life, having someone who acts as a compass can make the difference between getting lost and reaching a safe harbor. In this sense, you have been a true lighthouse for me. Thank you for your wisdom, dedication, and unconditional support, for believing and trusting in me from the very beginning, for encouraging me in difficult times, and for giving me the opportunity to work alongside you. Your presence in my life has been an invaluable gift that I will always remember.

To my co-advisor, Jorge. It would not have been possible to get here without your help and support. Also, having you by my side during my stay in the United States made everything much more bearable. I hope we can continue to collaborate in the future and that this friendship and working relationship will last for a long time.

To my parents, who have believed in me from the beginning of my academic career to the end of this research project. They taught me the value of effort, perseverance, and commitment, and thanks to them, today I have the joy of seeing this dream come true.

To Belén, who has always celebrated my achievements as her own. You are a true example of what it means to be a sister, and I hope that our relationship continues to strengthen over the years.

To Melanie, my confidante. Thank you for your patience during stressful moments, for your words of encouragement and for your unwavering support. You are a true pillar in my life and an unbreakable support in difficult times. Thank you for being my life partner.

To the friends and family members who are always by my side and who make me feel accompanied.

Resumen

Las técnicas de visualización innovadoras están transformando la presentación de productos y desempeñando un papel crucial en un mercado altamente competitivo. Las tecnologías de Realidad Virtual (RV) y Mixta (RM) están en constante evolución en términos de hardware, software, ergonomía, usabilidad, calidad y eficiencia, y se han convertido en medios efectivos para representar modelos virtuales en diversas aplicaciones de diseño. La disponibilidad y accesibilidad de estas tecnologías, tanto en términos de hardware (por ejemplo, Meta Quest 2 o Pico 4) como de software (por ejemplo, el "Metaverso"), están impulsando su adopción en entornos de desarrollo de productos y están influyendo en la forma en que trabajamos y colaboramos. Además, los avances continuos en la tecnología de los smartphones están permitiendo cada vez más el acceso a entornos de realidad mixta. De hecho, muchas empresas han utilizado la RM en sus catálogos en línea como una herramienta efectiva para presentar sus productos (por ejemplo, Ikea, Sephora o L'Oreal).

En este contexto, es crucial comprender cómo estas tecnologías afectan las impresiones subjetivas de los usuarios sobre un producto en particular, es decir, cómo un producto es percibido, interpretado e interiorizado por el consumidor, ya que esto puede variar significativamente según la plataforma de presentación. Esta variación perceptual puede generar errores significativos durante el proceso de diseño, lo que a su vez puede aumentar los costos del producto.

Esta tesis doctoral presenta tres estudios que investigan el efecto de diferentes técnicas de visualización en las impresiones subjetivas del sujeto cuando evalúa un producto. Se seleccionaron diferentes tipos de productos (sillas, paragüeros, cafeteras y teléfonos de sobremesa) que fueron evaluados utilizando un Diferencial Semántico específico para cada uno de ellos. Para obtener resultados más sólidos, las escalas semánticas se clasificaron según las cuatro categorías del placer de Jordan.

Los resultados obtenidos indican que el medio utilizado para presentar un producto tiene un impacto en cómo lo percibimos y en nuestra confianza en las evaluaciones que realizamos sobre él. En este contexto, las características relacionadas con el placer físico de Jordan fueron las más afectadas por el cambio de medio. Sin embargo, el uso de RV con hápticos pasivos ayuda a minimizar estas diferencias gracias al sentido del tacto, mientras que realizar evaluaciones conjuntas de productos también reduce las diferencias causadas por la técnica de visualización. Estos hallazgos tienen un valor significativo para los desarrolladores de productos, los especialistas en marketing y los diseñadores que se esfuerzan por optimizar los beneficios de la realidad extendida y crear productos más cautivadores y efectivos.

Abstract

Innovative visualization techniques are transforming product presentation and playing a crucial role in a highly competitive market. Virtual Reality (VR) and Mixed Reality (MR) technologies are constantly evolving in terms of hardware, software, ergonomics, usability, quality, and efficiency, and they have become effective means of representing virtual models in various design applications. The availability and accessibility of these technologies, both in terms of hardware (e.g., Meta Quest 2 or Pico 4) and software (e.g., the "Metaverse"), are driving their adoption in product development environments and influencing the way we work and collaborate. Furthermore, continuous advancements in smartphone technology are increasingly enabling access to mixed reality environments. In fact, many companies have utilized MR in their online catalogs as an effective tool for showcasing their products (e.g., Ikea, Sephora, or L'Oreal).

In this context, it is crucial to understand how these technologies impact users' subjective impressions of a particular product, i.e., how a product is perceived, interpreted, and internalized by the consumer, as this can vary significantly depending on the presentation platform. This perceptual variation can lead to significant errors during the design process, which, in turn, can increase product costs.

This doctoral thesis presents three studies that investigate the effect of different visualization techniques on users' subjective impressions when evaluating a product. Different types of products (chairs, umbrella stands, coffee makers, and desktop phones) were selected and evaluated using product-specific Semantic Differentials. To obtain more robust results, the semantic scales were classified according to Jordan's four pleasure categories.

The obtained results indicate that the medium used to present a product has an impact on how we perceive it and our confidence in the evaluations we make about it. In this context, characteristics related to Jordan's physical pleasure were the most affected by the medium change. However, the use of VR with passive haptics helps minimize these differences through the sense of touch, while conducting joint product evaluations also reduces the differences caused by the visualization technique. These findings have significant value for product developers, marketing specialists, and designers who strive to optimize the benefits of extended reality and create more engaging and effective products.

Resum

Les tècniques innovadores de visualització estan transformant la presentació del producte i jugant un paper crucial en un mercat altament competitiu. Les tecnologies de Realitat Virtual (RV) i Realitat Mixta (RM) estan evolucionant constantment en termes de maquinari, programari, ergonomia, usabilitat, qualitat i eficiència, i s'han convertit en mitjans eficaços de representar models virtuals en diverses aplicacions de disseny. La disponibilitat i accessibilitat d'aquestes tecnologies, tant en termes de maquinari (per exemple, Meta Quest 2 o Pico 4) com de programari (per exemple, el "Metaverse"), està impulsant la seva adopció en entorns de desenvolupament de productes i influïnt en la forma en què treballem i col·laborem. A més, els avenços contínus en la tecnologia dels telèfons intel·ligents estan possibilitant cada vegada més l'accés a entorns de realitat mixta. De fet, moltes empreses han utilitzat la RM en els seus catàlegs en línia com a eina eficaç per exhibir els seus productes (per exemple, Ikea, Sephora o L'Oreal).

En aquest context, és crucial comprendre com aquestes tecnologies afecten les impressions subjectives dels usuaris sobre un producte en particular, és a dir, com un producte és percebut, interpretat i interioritzat pel consumidor, ja que això pot variar significativament en funció de la plataforma de presentació. Aquesta variació perceptual pot provocar errors importants durant el procés de disseny, cosa que, a la seva vegada, pot augmentar els costos del producte.

Aquesta tesi doctoral presenta tres estudis que investiguen l'efecte de diferents tècniques de visualització en les impressions subjectives dels usuaris en avaluar un producte. Es van seleccionar diferents tipus de productes (cadires, suports per a paraigües, cafeteres i telèfons de sobretaula) i es van avaluar utilitzant Diferencials Semàntics específics del producte. Per obtenir resultats més robustos, les escales semàntiques es van classificar segons les quatre categories de plaer de Jordan.

Els resultats obtinguts indiquen que el mitjà utilitzat per presentar un producte té un impacte en com el percebem i en la nostra confiança en les avaluacions que fem sobre aquest. En aquest context, les característiques relacionades amb el plaer físic de Jordan van ser les més afectades pel canvi de mitjà. No obstant això, l'ús de la RV amb haptics passius ajuda a minimitzar aquestes diferències mitjançant el sentit del tacte, mentre que la realització d'avaluacions conjuntes de productes també redueix les diferències causades per la tècnica de visualització. Aquests resultats tenen un valor significatiu per a desenvolupadors de productes, especialistes en màrqueting i dissenyadors que aspiren a optimitzar els avantatges de la realitat estesa i crear productes més atractius i efectius.

Related publications

This section presents the works that have been published in journal or conferences in relation to the results obtained during this doctoral research.

Journal papers

- Palacios-Ibáñez, A., Pirault, S., Ochando-Martí, F., Contero, M., & Camba, J. D. (2023). An Examination of the Relationship between Visualization Media and Consumer Product Evaluation. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2023.3238428

- Palacios-Ibáñez, A., Alonso-García, M., Contero, M., & Camba, J. D. (2023). The influence of hand tracking and haptic feedback for virtual prototype evaluation in the product design process. Journal of Mechanical Design, 145(4), 041403. https://doi.org/10.1115/1.4055952

- Palacios-Ibáñez, A., Navarro-Martínez, R., Blasco-Esteban, J., Contero, M., & Camba, J. D. (2023). On the application of extended reality technologies for the evaluation of product characteristics during the initial stages of the product development process. Computers in Industry, 144, 103780. https://doi.org/10.1016/j.compind.2022.103780

Conference papers

- Palacios-Ibáñez, A., Ochando-Martí, F., Camba, J. D., & Contero, M. (2022). The Influence of the Visualization Modality on Consumer Perception: A Case Study on Household Products. In 13th International Conference on Applied Human Factors and Ergonomics, New York, July (pp. 24-28). https://doi.org/10.54941/ahfe1002050

Table of contents

List of figures

List of tables

Chapter 1

1. Introduction and objectives

1.1. Context and motivation

Current markets are highly competitive (Roy et al., 2009) and people are often faced with a variety of options that can satisfy their basic needs in terms of quality, price, and function (Perez Mata et al., 2017). In this regard, emotions are playing an important role for standing out from competitors, as they can establish a profound connection between products and customers (Pieter Desmet et al., 2001).

Various emotional models have been proposed to characterize product emotion. For example, Jordan proposed an approach with four different pleasure categories (Tiger, 1992): physical (pleasures deriving from sensory organs), social (pleasures deriving from relationships with others), psychological (pleasures related to people's cognitive and emotional reactions) and ideological (pleasures related to people's values). Desmet (Pieter Desmet, 2002b) applied cognitive appraisal theory to explain the process of product emotion, and Norman (Norman, 2004) described the emotional aspects of the product using a neurobiological framework for emotions composed by three levels: (1) visceral, (2) behavioral, and (3) reflective level.

In this context, some studies have shown that the way we visualize products can also have an impact on the subject's emotions (Galán, Felip, et al., 2021). In other words, these emotions can be influenced by the mode of representation used to communicate or display a new product. Thus, the way products are presented is also becoming one of the key factors for differentiating oneself from the competition.

In the realm of industrial design and product development, the mode of presentation has demonstrated significant economic advantages, especially regarding the evaluation process. Product evaluation is an essential activity in the early stages of the product development (Ozer, 1999), something that some authors have discussed to be closely related to the success of the product (Cooper, 2019). These evaluations are commonly conducted by building physical prototypes, which incur additional costs both financially and temporally (Söderman, 2005). However, emerging visualization technologies such as Virtual Reality (VR) and Mixed Reality (MR) have the potential to expedite the design process without necessarily increasing the design cost, being an affordable and versatile alternative to physical prototyping (Cecil & Kanchanapiboon, 2007). These benefits can lead to more effective and efficient product design, as well as increased competitiveness in the market.

Focusing on online retail, e-commerce is becoming increasingly important, so traditional ways of interacting with products in physical stores are being replaced by digital means (Jeong et al., 2009; Jiang & Benbasat, 2007a), providing information to customers to evaluate the suitability of the product and make decisions. However, many product features are not easy to evaluate using these means due to their two-dimensional and visual limitations.

In this sense, different studies discuss the main factors that influence consumers when making purchasing decisions, such as being familiar with the product (Unal, 2017), the environment in which it is displayed (Jalil et al., 2016), its aesthetics (Jalil et al., 2016) or the influence of how a product is presented in digital media (Yoo & Kim, 2014). Considering the last-mentioned factor, the product representation method should be as accurate as possible to facilitate decision-making, and in any case, the representation used should generate perceptual and emotional responses as close as possible between the real product when a virtual prototype.

However, various studies indicate that the presentation medium often alters the consumer's perceptual response. Given this background, Wu et al. (Wu et al., 2016) argue that image quality affects how a product is understood, while other authors (Flavián et al., 2011) affirm that image size, quality, and motion influence how consumers perceive its usability. Other works (Artacho-Ramírez et al., 2008) found a significant influence of the representation method on product perception, while Vergara et al. (Vergara et al., 2011) showed how the level of interaction may have a significant effect on the user's subjective impressions of a product. In this context, the use of haptics during the virtual experience has proven to be quite beneficial in minimizing the disparities that exist among visualization mediums (Galán, Felip, et al., 2021). While there are controlled by computer devices (referred to as active haptics) that represent a widely used form of haptics, there are also cost-effective solutions such as passive haptics. These involve the utilization of low-cost physical prototypes that are synchronized with the virtual prototypes to deliver tactile sensations (Jerald, 2015).

At this point, it is important to note the spectacular development in recent years of VR and MR headsets. The availability and affordability of these technologies, both in terms of hardware and software have fueled their adoption in product development settings (Aurora Berni & Borgianni, 2020; Meta, 2021). In this context, low-cost VR headsets such as the Meta Quest 2 or Pico 4 stand out, which are standalone devices that are democratizing the use of VR for the average user. Devices such as the HP Omnicept Reverb G2 or Meta Quest Pro are also becoming more prevalent, as they can capture physiological and behavioral measures that allow for more precise information about the user's emotions. Furthermore, Apple's recent contribution to the market is the innovative Vision Pro, a cutting-edge MR device that holds the potential to revolutionize this field.

But VR is just one of the options that current technology provides us. If we consider Milgram and Kishino's taxonomy (Milgram & Kishino, 1994), in the context of product perception, one extreme would represent a real situation in which users can touch and see a real product (Fig. 1). At the other extreme, users could see the virtual product in a completely synthetic environment (VR). However, there are very interesting options between these extremes, such as Augmented Reality (AR), which allows for the insertion of a virtual representation into a real environment for evaluation (Craig, 2013). It is important to note that AR became particularly popular thanks to the technological advances of smartphones. This has brought AR closer to the average user, so many companies are introducing it into their online catalogs (e.g., Ikea Place) or during the product evaluation process.

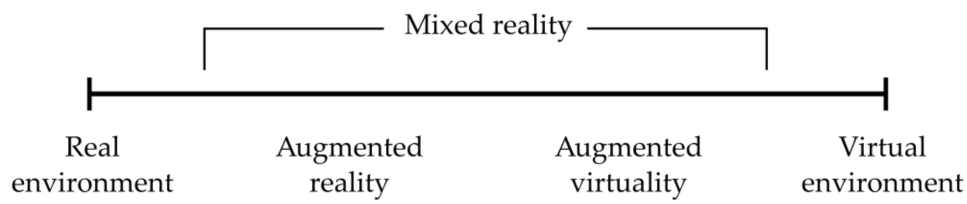

Fig. 1: Simplified representation of the virtuality continuum (Milgram & Kishino, 1994).

The Extended Reality technologies (XR)—an umbrella term for VR, AR and MR, where the "X" in XR acts as a flexible placeholder that can connect these different realities or expand beyond them (Rauschnabel et al., 2022)—are rapidly changing the paradigm in product design. They allow us to go beyond the limitations of two-dimensional screens by offering a more immersive and interactive experience. However, the different levels of reality-virtuality can potentially influence the perception of a product (Felip et al., 2019).

After conducting a deep literature review, it was found that there is a very limited number of studies that analyze in-depth how the different levels of Milgram's continuum affect the subject's perceptual response when interacting with a virtual product prototype. The rapid technological development offers VR solutions with very different levels of graphic quality and interactivity. MR devices with very interesting interaction capabilities and good graphic quality are also starting to be commercialized. However, there are very few published studies on how user perception is conditioned by the different forms of presentation and interaction made possible by these new VR and AR technologies.

The disruptive advance of new visualization media has revolutionized the way products are presented to users, both online and during the design process, but their value assumes that our subjective impressions and emotional responses to a virtual prototype are like those generated by the real product, which is not necessarily true. Hence, this research project has the potential to provide valuable insights into the ways in which XR can be used to enhance the product development and marketing process.

1.2 Research objectives and scope

The primary objective of this research was to contribute to a deeper understanding of the ways in which VR and MR technologies can be leveraged to improve the product development and marketing process by analyzing how these technologies condition our perceptual and emotional response when interacting with a virtual prototype of a product based on Jordan's emotional approach (pleasure categories).

As secondary objectives, this research aimed to identify the specific factors that may influence the user's response, such as visual quality (resolution, field of view), interaction capabilities (i.e., navigation paradigm or haptic capability), and Milgram's continuum level.

In the present context, numerous studies have been conducted which have culminated in the publication of three contributions in high-impact journals. All of them investigated the impact of various visualization techniques on product evaluation, ranging from traditional representations such as 2D images to more advanced visualization techniques such as VR or MR. The following is a brief description of the pursued objectives in each of them:

- (1) The research titled "On the Application of Extended Reality Technologies for the Evaluation of Product Characteristics During the Initial Stages of the Product Development Process" presents two case studies in which a group of individuals evaluated chair models with varying geometric characteristics (i.e., traditional and abstract designs) using distinct visualization techniques. Moreover, two assessment methods (i.e., joint and individual) were employed. The primary objective of this research was to examine the potential influence of the visual medium utilized to present these products on user perception, and to assess whether the selected evaluation methods could mitigate discrepancies resulting from changes in media.
- (2) On the other hand, in the publication "The Influence of Hand Tracking and Haptic Feedback for Virtual Prototype Evaluation in the Product Design Process," two studies were also conducted with three models of umbrella stands. Various users evaluated these products in an immersive VR setting to investigate the impact of tactile feedback into the virtual environment on the subjective impressions of participants concerning the actual product. Additionally, it was studied whether the use of a complementary item (i.e., an umbrella) could affect the evaluation of semantic scales and whether design training could influence such assessment.
- (3) Finally, in the article "An Examination of the Relationship between Visualization Media and Consumer Product Evaluation," two case studies were performed with two product typologies (i.e., three desktop phone designs and three coffee maker

designs) to examine the influence of media on the evaluation of semantic scales. Moreover, it was investigated whether the geometry of the product itself and user's gender could affect the evaluation of product features.

In these manuscripts, several hypotheses were formulated, which will be coded differently from the original papers to allow for subsequent discussion in this doctoral thesis. All of them share one general hypotheses: the medium used to present a product significantly influences the Semantic Scales regardless of their classification according to Jordan's pleasure categories (H0.1). Additionally, it was postulated that the Purchase Decision (H0.2) and the Confidence in the Response (H0.3) could be influenced by the change of medium for research (1) and (2), and that the Overall Evaluation of the product (H0.4) could be also influenced by the visualization technique for research (1) and (3).

Table 1: Approach to primary and secondary objectives for each paper.

<table>
<tr><th rowspan="3">Paper</th><th colspan="4">Objectives</th></tr>
<tr><th colspan="2">Primary</th><th colspan="2">Secondary</th></tr>
<tr><th colspan="2">How do visual media condition our perceptual and emotional response?</th><th colspan="2">Visual quality, interaction and Milgram's continuum level</th></tr>
<tr><td>(1)</td><td>- Use of different media (including VR and AR) for product observation.</td><td rowspan="3">The Semantic Differential technique was used to obtain quantitative data about product evaluation.</td><td>Interaction: traditional interfaces for image visualization, smartphones for AR, and natural walking with VR.

Milgram's continuum level: The left end was approached using images of the product, the middle point with augmented reality (AR), and the right end with virtual reality (VR).</td><td rowspan="3">Visual quality: the use of different visualization techniques implied the use of different levels of visual quality.</td></tr>
<tr><td>(2)</td><td>- Use of VRPH to study the influence of touch in product assessment.</td><td>Interaction: introduction of haptics in the VR experience.

Milgram's continuum level: Only the right end was approached.</td></tr>
<tr><td>(3)</td><td>- Use of different media (including VR and AR) for product observation.</td><td>Interaction: traditional interfaces for image visualization, smartphones for AR, and natural walking with VR.

Milgram's continuum level: The left end was approached using images of the product, the middle point with augmented reality (AR), and the right end with virtual reality (VR).</td></tr>
</table>

Moreover, specific hypotheses were formulated for each contribution: in the article titled "On the Application of Extended Reality Technologies for the Evaluation of Product Characteristics During the Initial Stages of the Product Development Process," it was postulated that a joint evaluation of products could help minimize existing differences due to changes in the visualization medium (H1.1); on the other hand, in the publication "The Influence of Hand Tracking and Haptic Feedback for Virtual Prototype Evaluation in the Product Design Process," we proposed that user expertise and design background could affect product evaluation (H2.1) and that the use of a complementary item could influence the evaluation of semantic scales (H2.2); finally, in the article "An Examination of the Relationship between Visualization Media and Consumer Product Evaluation," it was hypothesized that a particular design within the same product typology could influence the user's subjective impressions of the product (H3.1) and that gender differences could exist in the evaluation of a product and how it is perceived (H3.2).

To the objectives described above, this research was conducted within a defined geographical location, time frame, and with a specific population or sample utilizing a quantitative research design, with data collected through surveys and user testing of virtual product prototypes. For data analysis, an inferential statistical analysis has been conducted using various multivariable statistical techniques, including group comparison methods and non-parametric techniques to assess significant differences among different experimental conditions. Table 1 describes how each paper approached the objectives described.

1.3 Contribution to knowledge

This research contributes to the knowledge in the field of product development and design by examining the impact of novel visualization techniques, such as VR and MR, on users' emotional responses when assessing a virtual prototype of a product (Table 2).

The research presented in this document demonstrates that the medium used to present a product can significantly affect users' perceptual responses in different ways, and that there are factors that can help minimize these differences. We emphasize that the results detailed below are valuable for the product design and development process, but also for online and physical store retailers.

Firstly, not all product features are influenced in the same way by the medium used for product assessment. In the conducted research, a classification of product characteristics was made based on Jordan's pleasure categories (Jordan, 2002). This helped us conclude that adjectives belonging to the category of physical pleasure are the most affected by the visualization technique (i.e., those related to the five senses). Additionally, user confidence in the evaluation is also influenced by the medium, being higher using techniques that provide greater immersion. In this regard, using immersive VR can provide certain advantages, such as cost savings during the design process or the ability to showcase a large catalog of products without the need to have the actual product available now. However, it is important to consider the emergence of these differences and, above all, to understand how to minimize them to ensure that the product evaluation is as accurate as possible.

Hence, our research has shown that there are ways to minimize the differences that arise due to the visualization medium, which may be useful for product developers, marketers, and designers. One approach is the use of tactile feedback during a virtual experience, which provides more information about the product and helps the user form a better impression of the real product (this is the use of VRPH). Additionally, this also helps increase user confidence in their response, which has a positive impact on the evaluation. Furthermore, presenting products together for assessment (e.g., conducting ranking evaluations) also helps dissipate these differences, as ranking products can be an easier task for the user when having a reference to compare with, rather than doing it individually.

Therefore, we recommend using physical prototypes to complete the virtual experience and facilitate the understanding of the product. In this case, it is not necessary for the physical prototypes to be highly faithful to the virtual prototype since their main purpose is to provide the tactile experience. However, many of their characteristics can be altered in the virtual environment without affecting the physical product.

This doctoral thesis highlights the importance of using visualization techniques that promote the greatest possible interaction with the product to obtain accurate evaluations. By understanding the impact of different visualization techniques on users' emotional responses, product developers, marketers, and designers can create more engaging and effective products that meet users' needs and preferences.

This information is also important for the product evaluation process. Specifically, it can inform the design process and aid in the creation of products that are more likely to be well-received by the intended user base. Additionally, it can assist in the evaluation of products prior to purchase by providing insight into how different users may perceive a product based on their level of expertise or preferred evaluation method.

Table 2: Contribution of each paper to knowledge production.

<table>
<tr><th rowspan="2">Paper</th><th colspan="2">Contribution to knowledge</th></tr>
<tr><th>Common</th><th>Specific</th></tr>
<tr><td>(1)</td><td rowspan="3">- The visual medium used to present a product influence how it is perceived.
- Not all pleasure categories in Jordan's classification are influenced equally by the display medium, but Jordan's physio-pleasure category is the most affected.
- Using more sophisticated visualization techniques can enhance the evaluation of the product.</td><td>- A joint evaluation can help minimize perceptual differences elicited by the visualization technique.</td></tr>
<tr><td>(2)</td><td>- Design expertise may influence product evaluation.
- The Confidence in the response is influenced by the display medium and the use of haptics.
- The use a complementary item (when needed) can offer more information about the functioning of a product.</td></tr>
<tr><td>(3)</td><td>- A particular design within the same product typology influence product assessment.
- Gender differences appeared for some semantic scales, but no pattern was observed to draw general conclusions.</td></tr>
</table>

Overall, a comprehensive understanding of the factors that can influence product evaluation is essential for the successful design and evaluation of a product. By taking these variables into account, designers and evaluators can ensure that the product meets the needs and expectations of the target audience, thereby improving the product's overall success.

1.4 Research methodology

The general research methodology followed for this doctoral thesis (Fig. 2) was based on a strategy inspired by the DSRM (Design Science Research Methodology). This approach proposes the design of new tools/artifacts that allow for the resolution of an identified problem. Once designed and developed, these tools should be tested and evaluated with the aim of verifying their effectiveness in solving these problems, so that they can become part of a new set of knowledge through the communication of contributions.

Thus, the evaluation of the project was carried out through various field studies, initially in a controlled academic environment, where quasi-experimental designs with control group were used. This allowed the established iteration cycles in the DSRM methodology to refine the prototypes, in order to advance in the research process. In the case of this work, the prototypes consisted of applications in VR and MR environments that allowed the visualization of different types of products to users through different navigation and interaction techniques. The main objective during this research was to develop at least three iterations according to the DSRM methodology, which was successfully achieved.

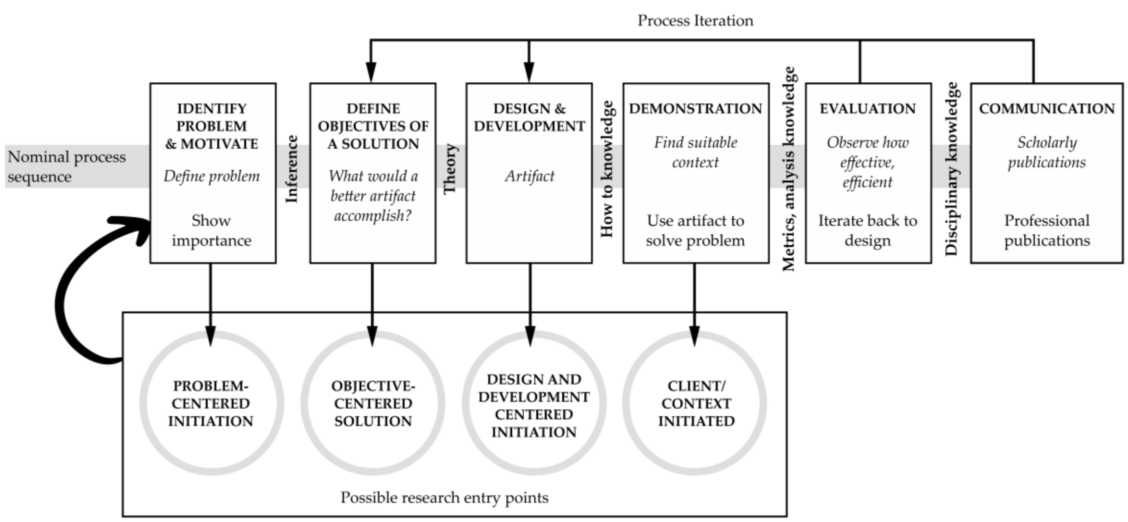

Fig. 2: Graphic scheme of the DSRM methodology (Peffers et al., 2007).

Additionally, a specific methodology was followed for the experiments detailed in this document. This methodology consisted of several steps that were carefully planned and executed:

- The first step was to define the primary objectives of the study, ensuring that they were clear, specific, and measurable. This was crucial for ensuring the validity and reliability of the study results. In general, this consisted of defining which factors were considered as those that can affect the evaluation of a product, such as visual quality, interaction capabilities or Milgram's continuum level.

- Next, the appropriate stimuli were selected for the study. This involved a careful review of existing research and consideration of factors such as the product type, the target audience, and the desired outcome. The study environments were then designed and created, which could include 3D modeling of prototypes, the generation of photorealistic images or other multimedia to ensure the accurate representation of the stimuli. In this context, it was decided to select at least three designs of the same product typology for each of the proposed case studies. This makes it easier to extrapolate the results obtained to a specific product typology, something that other authors did not take into account in their research (Artacho-Ramírez et al., 2008; Felip et al., 2019; Galán, Felip, et al., 2021; Galán, García-García, et al., 2021)
- To assess the users' perceptions, the Semantic Differential technique was used, which allowed for the evaluation of different aspects of the stimuli. The semantic differential consists of a set of semantic scales (Osgood et al., 1957) serving as product descriptors. To determine the bipolar pairs of adjectives for each product typology, we compiled a list of adjectives from similar studies that uses the same (or similar) product as stimulus if this was posible, or we carry out the procedure from scratch following the methodology carried out by other authors (Felip et al., 2019; Galán, García-García, et al., 2021).
- The sample size was also carefully considered, so an a priori power analysis with G*Power (Faul et al., 2007) was conducted to estimate the minimum sample size. Our results estimated a total sample size of 28.
- The experimental phase involved the presentation of the stimuli to the study participants in a controlled environment, where various measures were taken to eliminate extraneous variables that could potentially influence the results. All experiments carried out were within-subject studies and data were collected during this phase and subsequently analyzed using appropriate statistical methods to assess the significance of the findings.

Finally, the conclusions were drawn based on the results obtained, and recommendations for further research were made. It is worth noting that this methodology was tailored to the specific needs of the study and could be adapted to future work.

1.5 Thesis structure

This report reflects in detail the research conducted during the development of the doctoral thesis. The thesis is presented in the form of a compendium of articles, that is, the inclusion in the report of articles published and accepted in peer-reviewed prestigious indexed journals, whose main author is the doctoral candidate.

In this context, the present document has been divided into six chapters to facilitate its reading and comprehension. These are detailed below:

- *Chapter 1: Introduction and objectives*

  An overview of the objectives and intention of this doctoral thesis is presented. To this end, the research is contextualized and the fundamental reasons for addressing its subject matter are explained. Additionally, the aspects in which knowledge is sought to be advanced are detailed, and the research methodology used is described in a general manner.

- *Chapter 2: On the application of extended reality technologies for the evaluation of product characteristics during the initial stages of the product development process*

  This chapter presents the first experimental work published during this research. Two case studies were carried out to observe the influence of the visualization technique on the user's subjective impressions and the effect of the presentation format (i.e., joint or individual) on product perception. For this purpose, a group of participants evaluated different design options of a chair using the Semantic Differential. Throughout the chapter, an introduction to the main topic of study will be provided, contextualizing the reader through the section of previous works before analyzing the results obtained.

- *Chapter 3: The influence of hand tracking and haptic feedback for virtual prototype evaluation in the product design process*

  The present study showcases the second work that was developed as part of this research. This paper reports the results of a study where a group of participants evaluated three designs of an umbrella stand when viewed in a real setting, Virtual Reality (VR), and VR with passive haptics. Our goal was to observe the influence of visual media in product perception, and how the use of a complementary item (i.e., a physical umbrella) for interaction as well as user design expertise influence product assessment.

- *Chapter 4: An examination of the relationship between visualization media and consumer product evaluation*

This chapter presents the third experimental work published during this research. In line with previous cases, we report two case studies in which a group of participants evaluated three designs of two product typologies (i.e., a desktop telephone and a coffee maker) as presented in three different visual media (i.e., photorealistic renderings, AR, and VR for the first case study; and photographs, a non-immersive virtual environment, and AR for the second case study) using eight semantic scales. The aim is to observe the effect of visualization technique on the evaluation of products with significantly different formal characteristics.

- *Chapter 5: General discussion*

  This chapter carries out the analysis and interpretation of the most relevant findings obtained in the research in relation to the established objectives. Additionally, the implications of these findings are examined within the framework of the existing literature. This chapter provides an opportunity for a critical reflection that allows for the extraction of the most significant conclusions from the research and successfully conclude the doctoral work.

- *Chapter 6: Conclusions*

  This chapter provides a final summary of the results obtained, analyzing the limitations and presenting possible future lines of research that have emerged from this study.

Chapter 2

2. On the application of extended reality technologies for the evaluation of product characteristics during the initial stages of the product development process

Palacios-Ibáñez, A., Navarro-Martínez, R., Blasco-Esteban, J., Contero, M., & Camba, J. D. (2023). On the application of extended reality technologies for the evaluation of product characteristics during the initial stages of the product development process. Computers in Industry, 144, 103780. https://doi.org/10.1016/j.compind.2022.103780

This chapter presents one of the experimental works developed during this research. Two case studies were carried out to observe the influence of the visualization technique on the user's subjective impressions and the effect of the presentation format (i.e., joint or individual) on product perception. For this purpose, a group of participants evaluated different products belonging to the same product typology (i.e., a chair) using the Semantic Differential, whose scales have been grouped into categories based on Jordan's emotional approach to obtain more robust conclusions. Throughout the chapter, an introduction to the main topic of study will be provided, contextualizing the reader through the section of previous works before analyzing the results obtained.

2.1. Introduction

Fast-growing global markets are forcing companies to continuously re-evaluate consumer needs when designing new products (Coutts et al., 2019). Consumers must decide among a wide range of functional products, where market saturation has led to an increasing supply of products with high emotional value (Aftab & Rusli, 2017).

Although the level of innovation of these products can affect consumer choice, it is not the main factor for product success. In fact, many innovative products fail when they reach the market (Marquis & Deeb, 2018; van Kleef et al., 2005a). Some researchers have suggested that success is strongly linked to product evaluation throughout the design process (Cooper, 2019).

Physical prototyping is a popular tool for evaluating design concepts. Several authors have studied how product-user interaction influences product success (P. M. A. Desmet et al., 2008). However, the physical prototyping process may involve significant investments in terms of time and money, with limited flexibility to modifications, and even efficient techniques such as 3D printing and rapid prototyping do not significantly reduce the amount of time invested in many cases.

The availability and affordability of XR technologies such as VR and AR have fueled their adoption in a number of sectors, including product development (Aurora Berni & Borgianni, 2020). New devices with improved quality, usability, ergonomics, and efficiency have made virtual prototyping an effective tool to represent products in various industrial and design applications (A. Berni et al., 2020). Today, XR is used throughout the design process, particularly in the early stages of development where many design variations need be produced quickly (Cecil & Kanchanapiboon, 2007).

The cost of making a design change increases dramatically as the product moves through its lifecycle (van Kleef et al., 2005b). In this regard, XR technologies can help reduce design costs by enabling engineers to study and improve the product in virtual space before it is passed on to manufacturing. XR technologies have also proven to be an effective alternative for involving final users in the design process by facilitating the collaboration between designers and users (Bruno & Muzzupappa, 2010). Finally, although the creation of a Virtual Environment (VE) may require a considerable time investment depending on the desired level of realism, XR technologies allow the real-time modification of certain product features (i.e. textures, colors or materials), which enables the exploration of a large number of design alternatives without the need for physical prototypes. In this regard, many authors have employed these techniques in studies on design evaluation (Arbeláez & Osorio-Gómez, 2018; Cascini et al., 2020; De Crescenzio et al., 2019; Song et al., 2018).

XR have also had an effect on online retail and e-commerce, whose popularity (Y. Wang et al., 2020) has been increasing steadily in recent years, especially during the COVID-19 pandemic (Tran, 2021). Today, products are presented in online platforms in ways that go beyond traditional 2D images (Galán, Felip, et al., 2021), making the presentation medium a key differentiating factor to make a product successful. Some authors have discussed the physical barrier that exists in online platforms between user and product, which can be overcome by making attractive high-quality product presentations that allow consumers to evaluate the product accurately (Bleier et al., 2019). In this regard, new visualization technologies have been gaining traction in recent years as mechanisms to enable richer user-product interactions. They are in high demand (Statista, 2021) and concepts such as the "metaverse" are becoming more popular.

Current technologies enable a wide range of possibilities for presenting a product, including non-immersive VEs, where content is displayed through traditional interfaces (Pleyers &

Poncin, 2020), immersive VEs experienced through Head Mounted Displays (HMD's) (Jerald, 2015), or technologies such as AR (Arbeláez & Osorio-Gómez, 2018), which integrate virtual products into real environments. However, the level of fidelity and realism of the prototypes in different stages of the design process may vary depending on the specific needs (C. H. Chu et al., 2022). Modern devices can even collect physiological data such as eye-tracking, heart rate, or cognitive load, thus allowing subconscious user opinions to be translated into new design requirements in a non-invasive manner.

In this context, a critical factor in the decision-making process that occurs early in the product development process is that the evaluation provided by test users using XR technologies must be as accurate and reliable as possible. These technologies enable the creation of high-fidelity geometric representations to evaluate a product (Bordegoni, 2011a), positively affecting the user's confidence and accuracy in the assessment (Hannah et al., 2012). While some aspects such as aesthetic features or visual quality can be effectively assessed using mixed realities, product features that require physical interaction still rely on physical prototypes (Bordegoni, 2011a). It is generally assumed that our perceptual and emotional response to a product perceived using XR technologies is comparable to that of the physical product. However, research has shown that this is necessarily not the case (Felip et al., 2019; Galán, Felip, et al., 2021).

Product form can also influence product perception (Achiche et al., 2014). Form plays a significant role in the aesthetics of a product and it is a critical aspect during the design process. Form also usually conveys the first impression about the product to the user. In this context, XR technologies can be leveraged to understand the influence of geometry on product perception. Products with atypical and complex geometrical shapes may be more likely to elicit perceptual differences on features related to product aesthetics when viewed in different visual media (Palacios-Ibáñez et al., 2022). Determining whether these results are generalizable to other product typologies or more traditional designs can provide valuable information to inform product design processes and decisions at the point of sale.

In addition, conjoint analysis (Eggers et al., 2022) can be used to study whether the emotional responses elicited by a product are comparable when the product is viewed separately or in context when surrounded by other designs. Generally, in a retail environment (both physical and online) different designs of the same product typology are available to the user at any given time, so a joint evaluation is a more prevalent and realistic scenario. Authors Hsee et al. (Hsee et al., 1999) reported that when people evaluate products separately, relevant attributes that are difficult to evaluate are likely to be neglected in favor of attributes that are irrelevant but salient. Therefore, it should be more difficult for consumers to assess certain product attributes when evaluating one product at a time rather than when multiple products are evaluated simultaneously (Christopoulos et al., 2011).

Since new product designs are often evaluated individually throughout the design process, it could be argued that these assessments may not be entirely accurate compared to joint evaluations in VEs, which can help minimize perceptual differences.

The present study contributes to our understanding of how XR technologies can affect the various dimensions of the perceptual space linked to a product using different evaluation methods. The study examines how XR technologies affect the user's emotional response during product evaluation, assessing whether these technologies can make the design process less time-consuming and more cost-effective. Furthermore, we investigate how the evaluation format (individual or joint) affects product assessment, and whether providing several concept options to the users can yield more accurate evaluations. We applied the semantic differential technique (Osgood et al., 1957) to evaluate the characteristics of two sets of chairs with very different characteristics.

2.2. Background

The influence of the presentation medium on user perception has been studied in different fields (Bordegoni, 2011b; C. H. Chu et al., 2022). The effect of visual presentation media on product perception was first studied by Söderman (Söderman, 2005) who compared assessments of cars in a non-immersive VR environment and 2D sketches to evaluations of physical cars. His results showed that the perception of certain product features was not affected by the display medium, which the author attributed to the user's prior knowledge of the product.

Artacho-Ramírez et al. (Artacho-Ramírez et al., 2008) studied the perceptual variations of two loudspeaker designs in five different media, concluding that the type of media significantly influenced product perception. Other studies incorporated more sophisticated technologies (Kato, 2019). For example, Galán et al. (Galán, Felip, et al., 2021) used household products that included passive haptics in which physical objects were synchronized with virtual counterparts to allow users to physically feel some of the virtual objects they saw and interacted with. The authors found perceptual differences due to the variation of the medium, especially for those in Jordan's category of physical pleasure (Jordan, 2002).

In the realm of AR, Samantak and Mi (Ray & Choi, 2017) investigated how AR affects product assessment and Agost et al., (M. J. Agost, 2020) reported a study in which a sideboard were evaluated using 360-degree visualizations, AR, and VR.

Few authors have employed conjoint product evaluation in their studies on perceptual analysis with different means of evaluation. Chuang et al. (Chuang et al., 2001) used this methodology to jointly evaluate 26 cell phones using a semantic differential scale to examine the relationship between users' preferences and the geometric design elements of the products. Some researchers have also studied the perceptual differences that arise when changing the presentation format of a product, but observations have only been made on a case-by-case basis. Furthermore, consumers often evaluate different designs of the same product typology before making a purchasing decision. Therefore, studying how a user's perception of a single product varies individually when changing the display medium may not provide results that are applicable to an everyday situation. Instead, simultaneous evaluation methods may provide more accurate insights.

Although new visualization methods are gradually reducing the physical barriers between user and product in virtual platforms, it is unclear the extent to which the user's perception of the product is influenced by the presentation media. In this paper, we present two studies with different chair designs as stimuli. In our first study, users evaluated four common chairs through 2D photographs of the product (IMG), a non-immersive environment (3D), and AR using a conjoint evaluation method. In the second study, we used four concept chair designs submitted to the 2014 Annual Wilsonart Student Chair Design Competition for the evaluation (the chair designs can be viewed at http://www.blogtour.co/wilsonart-does-design-proud-student-chair-design-competition/). In this case, product assessments were done individually using non-photorealistic renderings (NPR), AR, and VR.

2.3. Hypotheses

We postulated the following hypotheses: The presentation media influences the user's perception of the product (H0.1); The presentation media influences the user evaluation of semantic scales independently of their classification in Jordan's categories (H0.2); The overall evaluation of the product (H0.3), the confidence in the user’s response (H0.4) and the purchase decision (H0.5) is influenced by the presentation medium; A joint evaluation can minimize perceptual differences for the semantic scales between visual media (H1.1, hypotheses postulated only for the first study).

2.4. Materials and Methods

2.4.1. Case Study

To validate the previous hypotheses, we designed two experimental studies. The main purpose was to examine the interaction between product aesthetic features and modes of representation. To draw robust conclusions, we selected different designs of the same product typology (i.e. a chair) with notable aesthetic differences. Each set of chairs was studied separately. Because of their aesthetic differences, various means of presentation were selected to facilitate the understanding of each product.

In our first study, a group of 40 participants evaluated four common everyday chairs as shown in Fig. 1. Participants ranked the chairs according to eight semantic differential scales in three different means of presentation:

- Photographs of the product (Fig. 3) taken from multiple points of view and on a white background to avoid any interferences with external stimuli. Pictures were displayed using a computer screen, and the participants were allowed to use the mouse to zoom in and out if needed, as well as the arrows keys in the keyboard to switch pictures.

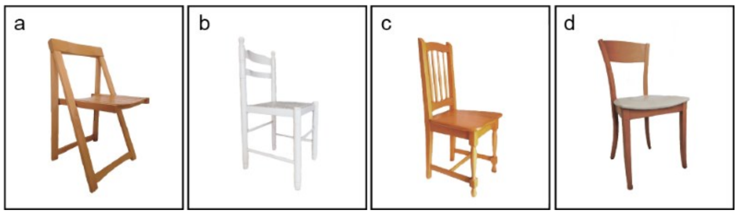

Fig. 3: Photographs of the chairs used in the first study.

- A non-immersive environment (Fig. 4) where virtual products were placed at the center of a virtual room. To minimize the impact of the environment on the product's evaluation or attention, the environment consisted of a simple shape and large room with neutral colors. This setting was displayed on a computer monitor, and no interaction with the virtual product was allowed, but the user was able to navigate the space by using their mouse and different keys on the keyboard to look at the product from any angle.

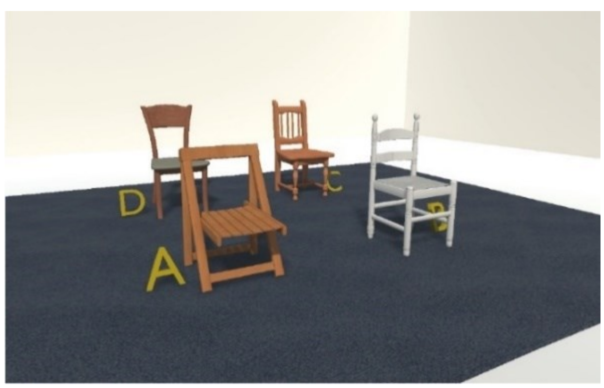

Fig. 4: Interactive 3D environment used in the first study.

- Augmented reality (Fig. 5), where the virtual products were placed in the real world. Since there was limited control over the surrounding environment in this setting, the participant was asked to display the objects in a common usage environment with a clear area for the models to have enough space. The user was allowed to move around the objects to view them from any angle. As in the previous environment, interaction with the products was not allowed.

Fig. 5: Virtual products placed in a real environment (AR) for the first study.

The means of representation selected for the first study were considered to provide enough information to the participants, as they were all common designs.

In the second study, 32 participants evaluated four unconventional chairs individually using a seven-level semantic differential scale. The media used included:

- A set of 2D images synthesized through non-photorealistic rendering techniques (Fig. 6). The images were generated from different points of view and on a white background to avoid any interferences with external stimuli during the product assessment. As in the first study, these pictures were displayed on a computer screen, and the participant was allowed to use the mouse to zoom in and out if needed, and the arrow keys in the keyboard to switch pictures.
- AR (Fig. 7), where the virtual product was placed in a real environment with similar characteristics to the first study. Interaction with the virtual product was not allowed, but the user was allowed to move freely.
- VR (Fig. 8), where the virtual product was placed in a VE experienced via an HMD. Product interaction was not allowed, but the user was allowed to move around the object and view it from different points of view.

We considered VR for the second study as a medium that could provide valuable information to compensate for the fact that participants may not have relevant previous knowledge about these chairs and thus not be as familiar with the characteristics of the products.

For both studies, participants were asked to rate how much they liked the product as well as their level of confidence in their responses using a 5-point Likert scale (where 1 represented "Dislike" and "no confidence," and 5 represented "Like" and "total confidence") and indicate their purchasing decision with a "Yes" or "No" answer.

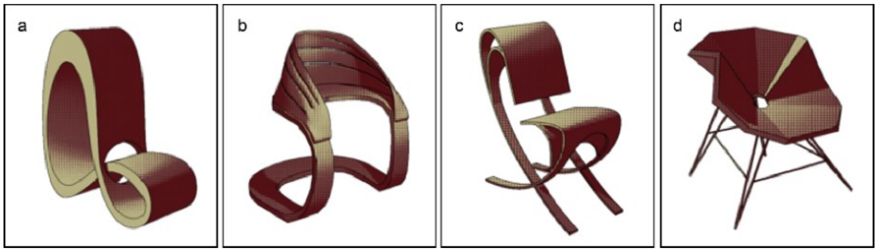

Fig. 6: NPR of the chairs used in the second study (Credit: Jenny Trieu (A), Abizer Raja (B), Arturo Barrera (C), and Carrah Kaijser (D)).

Fig. 7: Virtual product placed in a real environment (AR) for the second study.

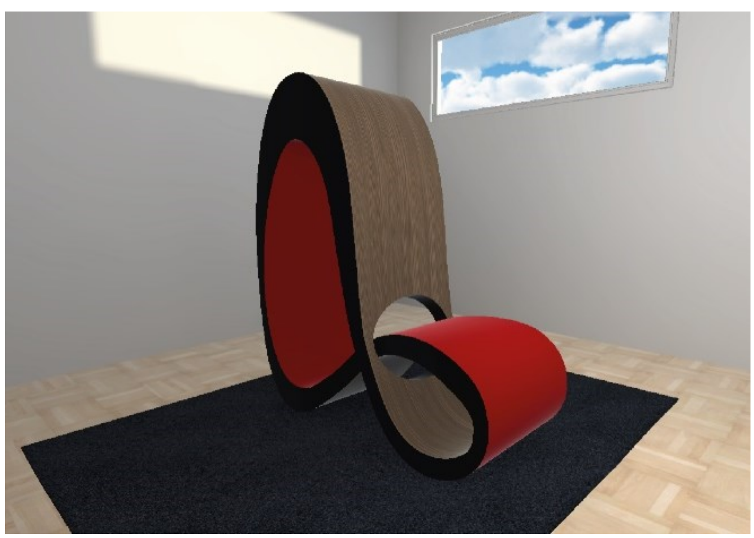

Fig. 8: Virtual product placed in a virtual environment (VR) for the second study.

2.4.2. Semantic Differential Scales for Product Evaluation

For our two studies, a set of semantic differential scales (Osgood et al., 1957) composed of eight bipolar pairs of adjectives serving as chair descriptors were used for product evaluation. This rating scale is a common method of product assessment (Pieter Desmet, 2002a) that does not force respondents to discriminate between items and allows them to state that several items are of similar importance (Flynn & Marley, 1992). To determine the semantic differential scales for our experiment, we compiled a list of adjectives from similar studies that uses the same (or similar) product as stimulus (Felip et al., 2019; Galán, García-García, et al., 2021). The final bipolar pairs of adjectives are shown in Tables 1 – 2 and are classified according to the four categories described by Jordan (Jordan, 2002). Due to the significant geometrical differences between the selected stimuli (a classical and homogeneous design in the first case study vs. an atypical design in the second), a decision was made to generate a slightly different semantic differential. In the first case study, a rank method was used to evaluate each product. For the second study, we opted for seven-level scales.

Table 3: List of bipolar scales used in study 1 classified by Jordan's pleasure categories.

Physio-pleasure	Psycho-pleasure	Socio-pleasure	Ideo-pleasure
Comfortable – Uncomfortable	Complex – Simple	Classic – Modern	Industrial – Handmade
Well proportioned – Unproportioned	Minimalistic – Overelaborated	Elegant – Conventional	Fun – Serious

Table 4: List of bipolar scales used in study 2 classified by Jordan's pleasure categories.

Physio-pleasure	Psycho-pleasure	Socio-pleasure	Ideo-pleasure
Comfortable - Uncomfortable	Complex – Simple	Classic – Modern	Industrial – Handmade
Light – Heavy	Practical – Impractical	Attractive – Unattractive	Fun – Serious

2.4.3. Materials

All the chairs in our studies were modeled in Blender 2.93.0. Textures had an image size of 512px and were acquired from Adobe Substance 3D Assets. The VR environment and 3D were designed using Unity 2019.4.14f1 with baked lights. The AR environment was created using the online resource Clon Digital (https://clondigital.es), which enables the integration of 3D models in a real environment without the need to develop a custom application.

To correctly visualize the non-immersive environment, we used Simmer (https://simmer.io), an online repository for Unity WebGL games. The VR environment was displayed using the Oculus Quest 2 HMD, a standalone immersive VR device with a Single Fast-Switch LCD of 1832×1920 pixels per eye and a refresh rate of 72Hz. For the AR environment in the first study, we used a wide range of smartphones with a minimum API level of 7 for Android devices, and iOS 11 version and an A9 processor as minimum specifications for iOS devices. For the second study, we used a OnePlus 7T, a smartphone with a screen size of 6,55in and a 1080x2400 screen resolution.

Due to COVID-19 restrictions, some interviews were conducted online for the first experiment via the Discord platform (instant messaging service for voice chat, video, and text chat). IBM SPSS Statistics 22 and Microsoft Excel were used for the inferential statistical analysis.

2.4.4. Sample

An a priori power analysis with G*Power (Faul et al., 2007) was conducted to estimate the minimum sample size, assuming an ANOVA with repeated measures with the following input parameters: effect size: 0.25, α=.05, (1-β)=.80 and 1 group. Our results estimated a total sample size of 28. To guarantee a power of .80, a total of 40 volunteers participated in the first experiment (25 men and 15 women, average age: 32 years old). The experiment was conducted both in person and online (62,50% of the participants in person, and 37.50% online). Before the experiment, users were asked to rate their experience with AR using a four-point Likert scale from 0 to 3 (0 = no experience, and 3 = significant experience). A total of 42.50% of the participants had no previous experience with AR, 37.50% rated their experience as limited, 17.50% stated having a lot of experience with AR, and 2.50% rated their experience as significant.

A total of 32 participants took part in the second study: 25% were male and 75% female with an average age of 24,88 years old. 50% of the male participants and 53,1% of the female participants had no previous experience with VR and AR, 34.4% of males and 28.1% of females rated their experience as limited with VR and AR, 15.6% claimed to have vast experience with VR and AR, and 3.1% rated their experience as significant with AR.

2.4.5. Experimental protocol

For both studies, all volunteers were over 18 years old. Before starting the experiment, verbal consent was obtained from each participant as well as basic demographics information (gender and age), data about the user experience in AR (for both studies) and VR (only for the second). Participants experienced the three experimental conditions (the viewing media) in a random sequence to minimize the effect of the presentation order. No interaction with the product was allowed in any of the visual media.

In the first study, participants were exposed to the four chairs simultaneously in each media and asked to rank them using the eight bipolar pairs in Table 3. They also evaluated how much they liked each chair and rated the level of confidence in their responses using a 5-point Likert scale. Finally, they were asked to make a purchasing decision ("Yes" or "No"). Each participant spent an average of 22 minutes per interview.

The second experiment was conducted entirely in person and involved the use of an HMD. In this case, participants were exposed to each chair individually in each media and asked to evaluate them using a eight seven-level semantic scales. Each participant completed a total of 12 evaluations. They also evaluated how much they liked each chair and rated the level of confidence in their responses using a 5-point Likert scale. Finally, they were asked to make a purchasing decision ("Yes" or "No"). The experiment took an average duration of 45 minutes per interview.

2.5. Results

In order to obtain more robust and reliable results, participants who experienced difficulties interpreting and applying the semantic scales were identied as outliers. As an additional factor, we decided to eliminate only those data points that appeared as outliers in at least 3 chairs. As a result, 5 data points were deleted for the first study.

2.5.1. First study results

Chairs were ranked according each bipolar pair presented on Table 3. The score obtained by each chair was equivalent to its position on the bipolar pair rank. Each end of the scale was associated to a particular adjective of a pair, so that a score closer to 1 indicated a greater correspondence with that adjective, and a score closer to 4 indicated a greater correspondence with the opposite adjective.

Four different data sets were obtained: the semantic scales, the overall evaluation, the purchase decision, and finally, the level of confidence in the response for each media. A normality test was performed on each data set to select the appropriate statistical test. As the sample size was less than 50 participants, we used a Shapiro-Wilks's normality test (significance level of .05). Results showed that the data was not normally distributed, so parametric tests were unsuitable.

We applied the Aligned Rank transform (ART) procedure (Higgins et al., 1990) as it provides a powerful and robust nonparametric alternative to other traditional techniques (Mansouri et al., 2004). It relies on a preprocessing step that "aligns" data before applying averaged ranks. After this step, common ANOVA procedures can be applied. The descriptive statistics for our four data sets are shown in Tables 3 – 5 and semantic scales stacked bar charts are shown in Fig. 7. The adjective in bold corresponds to a score of 4, whereas the adjective in italics(Ant Ozok & Komlodi, 2009) corresponds to a score of 1.

Table 5: Descriptive statistics for bipolar scales (study 1).

Semantic scales		Chair A-1			Chair B-1			Chair C-1			Chair D-1		
		IMG	3D	AR	IMG	3D	AR	IMG	3D	AR	IMG	3D	AR
Comfortable – *Uncomfortable*	M	**1.63**	1.57	*1.49*	*2.31*	2.69	**2.89**	2.31	**2.37**	*2.29*	3.74	**3.37**	*3.34*
	Mdn	1.00	1.00	1.00	2.00	3.00	3.00	2.00	2.00	2.00	4.00	4.00	4.00
	SD	.81	.79	.66	.83	.87	.99	.99	1.14	1.02	.61	.88	.84
Well proportioned – *Unproportioned*	M	**1.60**	*1.31*	1.40	*2.51*	2.89	**3.00**	3.34	**3.40**	*3.31*	**2.54**	2.40	*2.34*
	Mdn	1.00	1.00	1.00	3.00	3.00	3.00	4.00	4.00	4.00	3.00	3.00	3.00
	SD	.78	.63	.55	1.04	.90	.80	.80	.81	.83	1.12	.91	1.11
Simple – *Complex*	M	*2.86*	2.89	**2.97**	**2.89**	*2.57*	2.69	**1.63**	1.49	*1.43*	*2.63*	3.06	**3.37**
	Mdn	3.00	3.00	3.00	3.00	2.00	3.00	2.00	1.00	1.00	3.00	3.00	4.00
	SD	1.22	1.13	1.12	.99	.95	.87	1.02	.66	.65	1.14	1.00	.88
Minimalistic – *Overelaborated*	M	**3.34**	*3.14*	3.31	*2.46*	**2.57**	**2.57**	1.57	**1.63**	*1.54*	2.60	**2.66**	*2.57*
	Mdn	4.00	3.00	3.00	2.00	2.00	3.00	1.00	1.00	1.00	3.00	3.00	2.00
	SD	.90	.97	.80	.98	1.04	.95	.85	.94	.92	1.04	1.00	1.07
Classic – *Modern*	M	*2.29*	**2.49**	*2.29*	**3.46**	*3.37*	3.43	1.94	*1.89*	**1.97**	*2.31*	**2.26**	*2.31*
	Mdn	2.00	2.00	2.00	4.00	4.00	4.00	1.00	1.00	2.00	2.00	2.00	2.00
	SD	.86	.98	1.10	.70	.84	.78	1.19	1.11	1.01	1.08	1.01	1.02
Elegant – *Conventional*	M	1.54	*1.46*	**1.60**	*2.11*	**2.31**	2.14	*3.29*	**3.34**	*3.29*	**3.09**	*2.89*	3.06
	Mdn	1.00	1.00	1.00	2.00	2.00	2.00	3.00	4.00	4.00	3.00	3.00	3.00
	SD	.85	.74	.85	.99	.90	.88	.79	.91	1.02	.85	.96	.84
Industrial – *Handmade*	M	3.31	**3.40**	*3.29*	*1.49*	**1.57**	1.54	2.57	**2.63**	*2.54*	**2.63**	*2.40*	**2.63**
	Mdn	4.00	4.00	4.00	1.00	1.00	1.00	2.00	3.00	2.00	3.00	3.00	3.00
	SD	.90	1.04	1.07	.78	.66	.82	.98	.84	.980	1.00	1.09	.88
Serious – *Fun*	M	**2.77**	*2.69*	**2.77**	**2.77**	*2.49*	2.54	*1.91*	**2.11**	2.09	2.54	**2.63**	*2.51*
	Mdn	3.00	3.00	3.00	3.00	2.00	3.00	1.00	2.00	2.00	2.00	3.00	2.00
	SD	1.00	.93	1.06	1.00	1.04	1.04	1.15	1.21	1.12	1.15	1.24	1.20

Highest values and corresponding adjective in bold. Lowest values and corresponding adjective in italics.

Table 6: Descriptive statistics for overall evaluation and purchasing decision (study 1).

		Chair A-1			Chair B-1			Chair C-1			Chair D-1		
		IMG	3D	AR	IMG	3D	AR	IMG	3D	AR	IMG	3D	AR
Like/ Dislike	M	2.31	**2.34**	*2.14*	2.69	**3.03**	*2.57*	*3.49*	**3.51**	**3.51**	**3.37**	3.06	*2.74*
	Mdn	2.00	2.00	2.00	3.00	3.00	3.00	4.00	4.00	4.00	3.00	3.00	3.00
	SD	.72	.85	.85	.87	.82	.78	.74	.85	.70	.77	.87	.95
Purchasing decision	M	**.06**	**.06**	*.03*	*.06*	**.11**	.09	*.26*	*.26*	**.29**	*.17*	**.20**	*.17*
	Mdn	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
	SD	.236	.236	.169	.236	.323	.284	.443	.443	.458	.382	.406	.382

Highest values in bold. Lowest values in italics

Table 7: Descriptive statistics for the response confidence (study 1).

		IMG	3D	AR
Response confidence	M	**3.74**	3.63	*3.43*
	Mdn	4.00	4.00	3.00
	SD	.61	.65	.66

Highest values in bold. Lowest values in italics.

Next, differences between the display techniques were analyzed. One factor repeated measures ANOVAs and post-hoc tests with Bonferroni correction were performed for the semantic scales (Table 8 – 7). Although the p-value of the bipolar pair "Well proportioned - Unproportioned" was .036 for Chair B-1, post-hoc tests did not find significant differences in the pairwise comparisons.

We also performed a one-factor repeated measures ANOVA for the overall evaluation (Table 10), and a post-hoc analysis for Chairs B-1 and D-1 are shown in Table 11. Cochran's Q test was performed for the purchase decision (Table 12), but no significant differences were found between means.

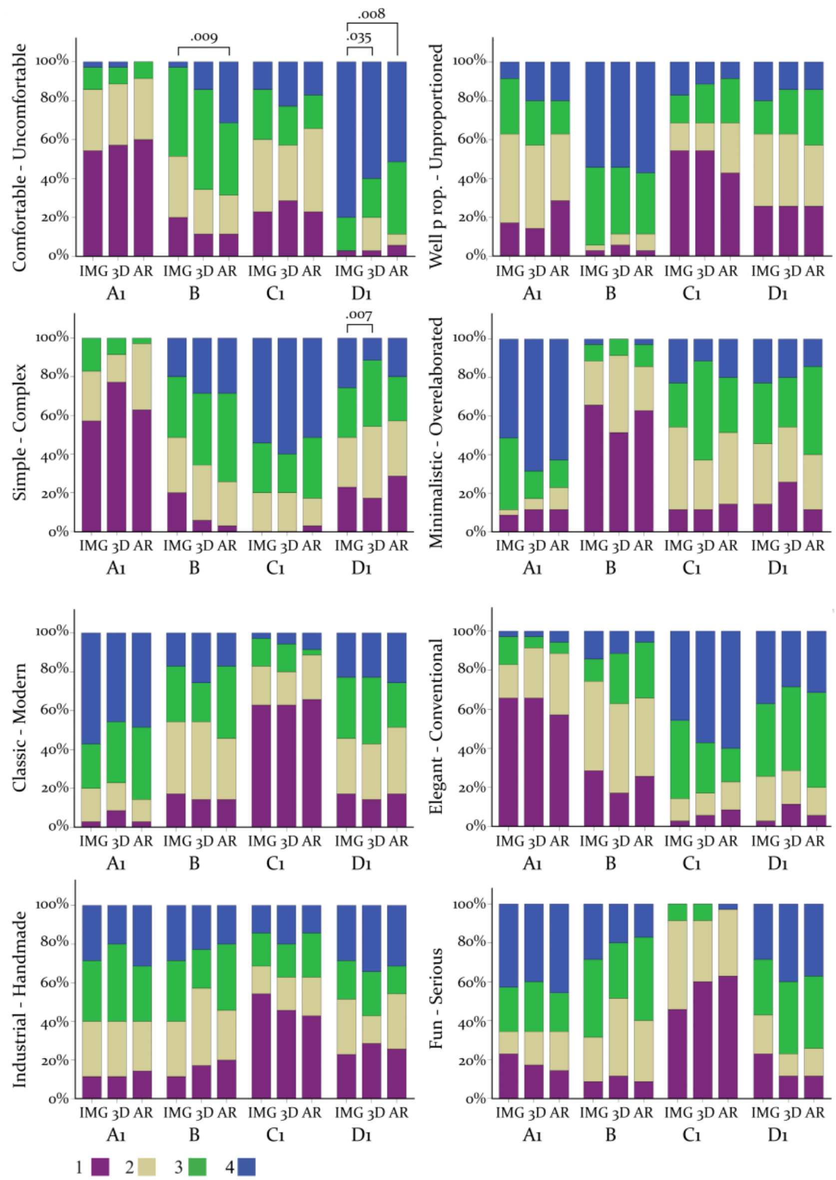

Fig. 9: Stacked bar charts for the semantic differential scales (study 1).

Table 8: Repeated measures ANOVA for the bipolar scales (study 1).

	Semantic scales	df	Chair A-1		Chair B-1		Chair C-1		Chair D-1	
			F	Sig.	F	Sig.	F	Sig.	F	Sig.
PHYSIO	Comfortable – Uncomfortable	2	2.166	p=.122	6.567	**p=.002**	.069	p=.933	6.088	**p=.004**
	Well prop. – Unproportioned		2.166	p=.122	3.502	**p=.036**	.175	p=.840	.578	p=.564
PSYCHO	Simple – Complex		.262	p=.770	1.732	p=.185	2.764	p=.070	3.992	**p=.023**
	Minimalistic – Overelaborated		.851	p=.432	.315	p=.731	.194	p=.824	.150	p=.861
SOCIO	Classic – Modern		.962	p=.387	.051	p=.950	.486	p=.617	.086	p=.918
	Elegant – Conventional		.626	p=.538	2.123	p=.128	.550	p=.580	.594	p=.555
IDEO	Industrial – Handmade		1.559	p=.218	1.211	p=.304	.317	p=.729	1.482	p=.235
	Serious – Fun		.147	p=.864	1.180	p=.313	.726	p=.488	.185	p=.831

p-values less than .05 are shown in bold.

Table 9: Post-hoc tests for the bipolar scales (chair B-1 and D-1).

Semantic scales		Chair B-1	Chair D-1
	Means	Sig.	Sig.
Comfortable – Uncomfortable	IMG – 3D	p=.054	**p=.035**
	IMG - AR	**p=.009**	**p=.008**
	3D - AR	p=.563	p=1.000
Well prop. – Unproportioned	IMG – 3D	p=.175	
	IMG - AR	p=.085	
	3D - AR	p=1.000	
Simple – Complex	IMG – 3D		**p=.047**
	IMG - AR		p=.169
	3D - AR		p=1.000

p-values less than .05 are shown in bold.

Table 10: Repeated measures ANOVA for the overall evaluation (study 1).

	df	Chair A-1		Chair B-1		Chair C-1		Chair D-1	
		F	Sig.	F	Sig.	F	Sig.	F	Sig.
Overall evaluation	2	1.486	p=.234	5.940	**p=.004**	0.034	p=.967	11.180	**p<.001**

p-values less than .05 are shown in bold.

Table 11: Post-hoc tests for the overall evaluation (chair B-1 and D-1).

		Chair B-1	Chair D-1
Like/Dislike	Means	Sig.	Sig.
	IMG – 3D	p=.123	p=.096
	IMG - AR	p=.622	**p<.001**
	3D - AR	**p=.011**	p=0.64

p-values less than .05 are shown in bold.

Table 12: Cochran's Q test for the purchasing decision (study 1).

	df	Chair A-1		Chair B-1		Chair C-1		Chair D-1	
		Q	Sig.	Q	Sig.	Q	Sig.	Q	Sig.
Overall evaluation	2	2.000	p=.368	1.500	p=.472	.290	p=.867	.500	p=.779

Finally, no statistically significant differences were found for the response confidence, F(2, 68) = 2.474, p = .092.

2.5.2. Second study results

Descriptive statistics for our second study are shown in Tables 11 – 13. A value closer to -3 represents a closer correspondence with the adjective to the left, and a value closer to 3 indicates a closer correspondence with the adjective to the right. Stacked bar charts for the semantic scales are shown in Fig. 8. A Shapiro-Wilk's normality test (significance level of .05) revealed that the data was not normally distributed, so parametric tests proved unsuitable. Therefore, the ART procedure was applied once again.

Table 13: Descriptive statistics for the bipolar scales (study 2).

Semantic scales		Chair A-2			Chair B-2			Chair C-2			Chair D-2		
		NPR	AR	VR	NPR	AR	VR	NPR	AR	VR	NPR	AR	VR
Light – **Heavy**	M	*1.41*	1.97	**2.31**	*-1.63*	**-0.91**	-1.47	-1.06	**-0.81**	*-1.12*	*-0.50*	-0.34	**-0.22**
	Mdn	2.00	2.00	2.50	-2.00	-2.00	-2.00	-1.00	-1.50	-2.00	-1.00	-1.00	0.00
	SD	1.54	1.45	0.78	1.21	1.92	1.63	1.60	1.82	1.45	1.70	1.91	1.91
Comfortable – **Uncomfortable**	M	**1.19**	0.25	*0.03*	1.09	*1.00*	**1.13**	**-1.34**	*-1.38*	*-1.38*	1.69	**1.75**	*1.66*
	Mdn	1.00	1.00	0.50	2.00	1.00	2.00	-2.00	-2.00	-2.00	2.00	3.00	3.00
	SD	1.51	1.76	1.91	1.87	1.88	1.93	1.47	1.50	1.74	1.51	1.81	2.06
Practical – **Impractical**	M	1.63	**1.69**	*1.38*	0.00	*-0.16*	**0.25**	-1.06	**-0.87**	*-1.16*	*0.88*	1.16	**1.28**
	Mdn	2.00	2.00	2.00	-0.50	-0.50	1.00	-2.00	-1.00	-2.00	1.00	2.00	2.00
	SD	1.36	1.65	1.75	1.80	1.94	2.00	1.62	1.64	1.65	1.74	1.89	1.85
Simple - **Complex**	M	*1.34*	1.56	**1.78**	*-0.22*	**0.13**	-0.16	0.66	**0.75**	*0.19*	*0.09*	**0.91**	0.69
	Mdn	2.00	2.00	2.00	-0.50	0.00	0.00	1.00	1.00	0.50	0.50	1.00	1.00
	SD	1.45	1.58	1.21	1.84	1.81	1.72	1.54	1.70	1.42	1.92	1.63	1.65
Modern – **Classic**	M	**-2.38**	*-2.72*	-2.63	**-0.94**	*-1.63*	-1.31	**-1.37**	*-1.87*	-1.50	**-2.00**	-2.28	*-2.47*
	Mdn	-2.50	-3.00	-3.00	-1.00	-2.00	-2.00	-2.00	-2.00	-1.50	-2.00	-2.00	-3.00
	SD	0.75	0.52	0.70	1.54	1.21	1.38	1.50	1.24	1.27	0.95	0.81	0.76
Attractive – **Unattractive**	M	**-0.22**	-1.09	*-1.41*	0.25	*-0.16*	**0.31**	-1.66	**-1.62**	*-1.75*	0.00	*-0.19*	**0.16**
	Mdn	-1.00	-1.50	-2.00	-1.00	-1.00	0.00	-2.00	-2.00	-2.00	0.00	-1.00	0.50
	SD	1.79	1.69	1.54	1.74	1.83	1.59	1.28	1.38	1.27	1.88	2.12	2.16
Fun – **Serious**	M	-1.69	*-1.91*	**-1.62**	*-0.87*	**-0.69**	-0.72	*-1.44*	-1.41	**-1.38**	**-0.62**	*-1.06*	-0.69
	Mdn	-2.00	-2.00	-2.00	-1.00	-1.00	-1.00	-2.00	-2.00	-2.00	-1.00	-1.00	-1.00
	SD	1.09	1.17	1.31	1.07	1.35	1.35	1.43	1.60	1.60	1.56	1.46	1.65
Handmade – **Industrial**	M	**0.47**	*0.28*	*0.28*	**0.31**	*-0.75*	-0.16	0.41	**0.50**	*0.13*	*0.88*	**1.13**	1.06
	Mdn	1.00	1.00	1.00	1.00	-1.00	-1.00	1.00	1.00	0.00	2.00	2.00	1.50
	SD	2.29	2.14	2.27	1.79	1.93	1.74	2.11	1.97	1.88	2.06	2.14	1.92

Highest values and corresponding adjective in bold. Lower values and corresponding adjective in italics.

Table 14: Descriptive statistics for overall evaluation and purchase decision (study 2).

		Chair A-2			Chair B-2			Chair C-2			Chair D-2		
		NPR	AR	VR	NPR	AR	VR	NPR	AR	VR	NPR	AR	VR
Like/Dislike	M	*2.81*	3.09	**3.22**	*2.16*	2.25	**2.31**	3.72	*3.56*	**3.75**	*2.37*	**2.59**	2.41
	Mdn	3.00	3.00	3.50	2.00	2.00	2.00	4.00	4.00	4.00	2.00	2,50	2.00
	SD	0.96	1.09	1.13	0.99	1.05	1.15	0.96	1.10	1.05	1.21	1.36	1.32
Purchase decision	M	*0.22*	0.34	**0.53**	*0.13*	**0.22**	0.19	**0.72**	*0.66*	**0.72**	*0.19*	**0.25**	*0.19*
	Mdn	0.00	0.00	1.00	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00
	SD	0.42	0.48	0.51	0.34	0.42	0.40	0.46	0.48	0.46	0.40	0.44	0.40

Highest values in bold. Lower values in italics.

Table 15: Descriptive statistics for response confidence (study 2).

		Chair A-2			Chair B-2			Chair C-2			Chair D-2		
		NPR	AR	VR	NPR	AR	VR	NPR	AR	VR	NPR	AR	VR
Response confidence	M	*3.47*	3.87	**4.09**	*3.50*	3.81	**3.94**	4.09	*3.94*	**4,13**	*3.72*	3.94	**4.06**
	Mdn	4.00	4.00	4.00	3.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
	SD	0.80	0.66	0.59	0.84	0.82	0.91	0.86	0.80	0.79	0.92	0.88	0.88

Highest values in bold. Lower values in italics.

We performed one-factor repeated measures ANOVA and post-hoc tests (Bonferroni adjustment was applied) for each data set. Although the p-value of the bipolar pair "Modern-Classic" was .046 for Chair B-2, post-hoc tests did not find significant differences in the pairwise comparison. For the purchasing decision, we performed Cochran's Q and McNemar tests. Results are shown in Tables 14 – 19.

Table 16: Repeated measures ANOVA for the bipolar scales (study 2).

	Semantic scales	df	Chair A-2		Chair B-2		Chair C-2		Chair D-2	
			F	Sig.	F	Sig.	F	Sig.	F	Sig.
PHYSIO	Light – Heavy	2	5.024	**p=.010**	1.855	p=.165	.162	p=.850	.336	p=.716
	Comfortable – Uncomfortable		3.668	**p=.031**	.153	p=.859	.202	P=.817	1.446	p=.243
PSYCHO	Practical – Impractical		.709	p=.496	.724	p=.489	.656	p=.523	2.134	p=.127
	Simple – Complex		1.469	p=.238	.476	p=.624	3.117	p=.051	2.726	p=.073
SOCIO	Modern – Classic		3.445	**p=.038**	3.235	**p=.046**	3.697	**p=.030**	4.884	**p=.011**
	Attractive – Unattractive		12.830	**p<.001**	1.519	p=.227	.205	p=.815	1.236	p=.297
IDEO	Fun – Serious		1.239	p=.297	.072	p=.930	.146	p=.865	1.560	p=.218
	Handmade – Industrial		.577	p=.564	5.492	**p=.006**	.893	p=415	.477	p=.623

p-values less than .05 are shown in bold.

Table 17: Post–hoc tests for the semantic scales (study 2).

Semantic scales	Means	Chair A-2 Sig.	Chair B-2 Sig.	Chair C-2 Sig.	Chair D-2 Sig.
Light - Heavy	NPR – AR	p=.072			
	NPR - VR	**p=.016**			
	AR - VR	p=1.000			
Comfortable - Uncomfortable	NPR – AR	p=.083			
	NPR - VR	**p=.040**			
	AR - VR	p=1.000			
Modern – Classic	NPR – AR	**p=.048**	p=.080	p=.137	p=.129
	NPR - VR	p=.197	p=.267	p=1.000	**p=.030**
	AR - VR	p=1.000	p=.972	**p=.032**	p=.530
Attractive - Unattractive	NPR – AR	**p=.001**			
	NPR - VR	**p<.001**			
	AR - VR	p=.716			
Handmade - Industrial	NPR – AR		**p=.008**		
	NPR - VR		p=.554		
	AR - VR		p=.155		

p-values less than .05 are shown in bold.

Table 18: Repeated measures ANOVA for the overall evaluation (study 2).

	df	Chair A-2		Chair B-2		Chair C-2		Chair D-2	
		F	Sig.	F	Sig.	F	Sig.	F	Sig.
Overall evaluation	2	2.780	p=.070	.300	p=.734	.800	p=.455	.860	p=.428

Table 19: Cochran's Q test for the purchasing decision (study 2).

	df	Chair A-2		Chair B-2		Chair C-2		Chair D-2	
		Q	Sig.	Q	Sig.	Q	Sig.	Q	Sig.
Purchase decision	2	15.200	**p<.001**	2.800	p=.247	0.800	p=.670	1.330	p=513

p-values less than .05 are shown in bold.

Table 20: McNemar test for the purchasing decision (chair A-2).

	Media	Chair A-2 Sig.
Purchase Decision	NPR – AR	**p=.001**
	NPR – VR	**p<.001**
	AR - VR	p=.207

p-values less than .05 are shown in bold.

Finally, the response confidence, statistically significant differences were found for Chairs A-2 (F(2, 62)=8.56, p=.001) and D-2 (D-2 F(2,62)=3.73, p.030). Post-hoc tests are shown in Table 239.

Table 21: Post–hoc tests for response confidence.

	Media	Chair A-2	Chair D-2
		Sig.	Sig.
Response Confidence	NPR – AR	**p=.020**	p=.298
	NPR – VR	**p=.003**	**p=.031**
	AR - VR	p=.433	p=.931

p-values less than .05 are shown in bold.

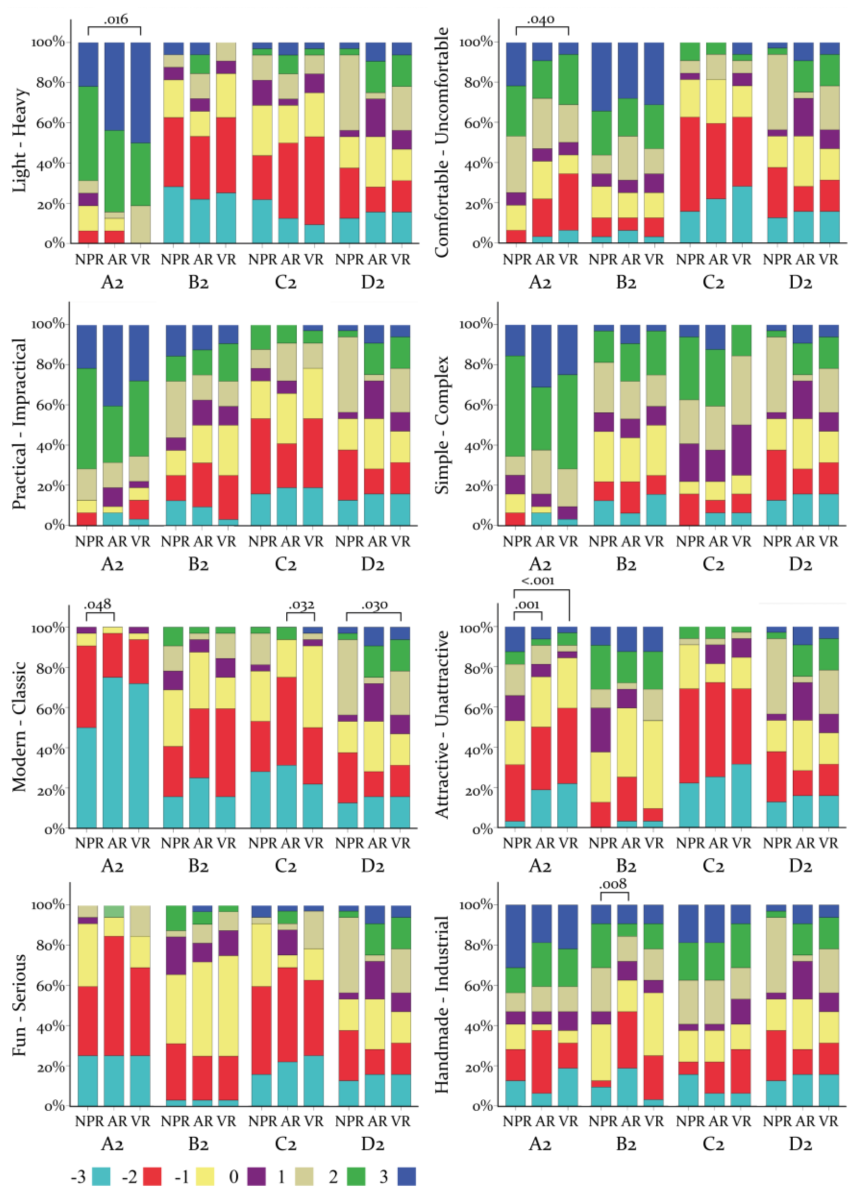

Fig. 10: Stacked bar charts for the semantic differential scales (study 2).

2.6. Discussion

In this paper, we examined the influence of the presentation medium on the perception of a set of chairs using different evaluation methods. We conducted two different studies: in the first case, four chairs with a classic and homogeneous design were selected to be viewed and evaluated simultaneously using photographs of the actual product, a non-immersive virtual environment, and AR. In the second case study, a more diverse set of four chairs with an atypical design were presented to participants and evaluated individually using a set of 2D images synthesized through non-photorealistic rendering techniques, AR, and VR. In both studies, participants used the semantic differential technique to evaluate the products. In addition, participants were asked to provide an overall evaluation of each chair ("Like/Dislike"), a purchase decision, and a rating of the level of confidence in their responses for each visual medium.

Our results show that the purchase decision is not influenced by the visual medium. Differences between media for this dataset were found only for Chair A-2 (H0.5 is rejected). Therefore, it could be argued that although access to information is critical to make a purchase decision (O'Keefe & McEachern, 1998), the medium used to present this information may not be a determining factor. These results agree with other studies that concluded that 2D media may offer sufficient information to reliable assess a product (Ant Ozok & Komlodi, 2009).

In our first hypothesis (H0.1), we speculated that the presentation media could influence the user's perception of the products. In the first case study, perceptual differences were found for certain bipolar pair of adjectives for Chairs B-1 and D-1 (Table 8). Differences were also found in all chairs for certain bipolar pairs of adjectives in our second study (Table 16). These results confirm H0.1 for both studies, which agree with the results obtained by other authors (Bleier et al., 2019; Palacios-Ibáñez et al., 2022) and contributes to expand the scope of product typologies. Similar to Artacho-Ramírez et al., who studied the influence of the graphical representation in the evaluation of different models of a loudspeaker (Artacho-Ramírez et al., 2008), and Agost et al., who reached similar conclusions when assessing two different type of furniture (a sideboard and a lamp) (M.-J. Agost et al., 2021), our study confirms that these results can be extrapolated to other types of products.

In H0.2, we questioned whether the presentation media influences the user evaluation of semantic scales independently of their classification in Jordan's categories. In the case of Chair B-1, differences were found for the bipolar pair "Comfortable - Uncomfortable" (physiological pleasure category), and for Chair D-1, differences were found for "Comfortable - Uncomfortable" and "Simple – Complex" (psychological pleasure category). In this case, the physio-pleasure category was the most affected by the change of visual medium, which aligns with other authors who confirmed the importance of haptics for the evaluation of these characteristics (Galán, Felip, et al., 2021). The absence of touch in our study may have had a negative influence on the evaluation of certain product attributes.

Although some adjectives from the physio-pleasure category were affected by the change of medium in the second case study, the socio-pleasure category (closely related to product aesthetics) was the most influenced. These results may have been affected by the geometric characteristics of the chairs, as these products had a higher aesthetic value . It has been argued that the aesthetic elements of a product's shape can influence user perceptions (Achiche et al., 2014) so a less typical design of a chair may have influenced perception regardless of the medium in which it was presented. Our results align with those obtained by Palacios-Ibáñez et al. , who demonstrated that the sociological pleasure category was the most affected by the change of media for the case of products with high aesthetic value (i.e., coffee makers). For our study, we also used stimuli with high aesthetic value, but significantly different features (i.e., chairs). Differences were mostly found between IMG – AR and IMG – 3D, but it is important to emphasize that different representation methods do not necessarily have the same interaction capabilities. More sophisticated media provide higher levels of interaction, so it can be expected that these perceptual differences stem from this, as demonstrated by Ozok and Komlodi (Ant Ozok & Komlodi, 2009).

According to the descriptive statistics for the semantic scale in our second study (Table 13), it is important to highlight how all four chairs appeared heavier to participants in the more immersive media (Chairs B-2 and C-2 when presented in AR, and Chairs A-2 and D-2 when presented in VR), whereas three of the four chairs (A-2, B-2, and D-2) appeared lighter when viewed by participants in NPR. This result could be attributed to the fact that the chairs were displayed in their actual sizes in AR and VR compared to flat images shown on a 2D screen (with dimensional limitations) which may have made the product appear smaller (as in (Galán, Felip, et al., 2021)), and thus lighter. In addition, the NPR medium did not accurately represent the actual material of the chair, so some participants may have perceived the product as heavier in a more interactive medium. Our analyses confirm H0.2 for both studies, which agree with the studies discussed earlier.

We also speculated that, for both studies, the overall product evaluation and response confidence could be influenced by the presentation medium (H0.3 and H0.4). For the first study, our results showed that AR was the medium in which the product was less liked, and where the user felt less confident about their response, followed by 3D and IMG. Although we expected the opposite effect for the response confidence (as 3D representations provide more information to the user), our results could be explained by the user's limited experience with technology as well as the online interview conducted with some users, which could lead to lower levels of confidence when performing evaluations (42.5% had no previous experience with AR, and 37.5% rated their experience as limited). The lack of experience plus the pressure of taking part of an online study may have led some users to experience technology related anxiety: an individual's concern about being able to use a technological device correctly (Arvanitis et al., 2009). Users who have this anxiety are not likely to experience the AR in mobile applications effectively (Oyman et al., 2022), which may have influenced our results. The overall evaluation data set also showed statistically significant differences between means. These differences where found between IMG-AR and 3D-AR, expected result as AR was the only medium showing possible usage environment (the real world). Our results contradict those obtained by (Galán, García-García, et al., 2021), who compared the overall evaluation of one chair in a real setting, VR and VR with passive haptics. In our case, the visual media used to display the stimuli presented greater interaction differences, which may explain the results, which are similar to those obtained by (Palacios-Ibáñez et al., 2022).

For the second study, results showed that participants were generally more confident in their responses in the VR environment. This result agrees with previous studies: while 2D media may offer sufficient information to evaluate a product, 3D representations provide richer information to the user (Ant Ozok & Komlodi, 2009), which helps them assess the appearance and features more directly (S. Liu, 2017), resulting in greater levels of certainty and confidence. Forbes et al. confirmed this result in their study with armchairs, where a more interactive and immersive medium helped to increase the user response confidence (Forbes et al., 2018a). Although the overall evaluation data set did not show statistically significant differences between means, we observe that three of the four chairs were rated more favorably in VR, and all the designs were rated less favorably in NPR, as shown in the descriptive statistics for the overall evaluation. Therefore, we can confirm H0.3 and H0.4.

Finally, in H1.1, we speculated that a joint assessment could minimize the differences between the assessments made by the participants in different media. In the case of Chair C-1, the variation in the evaluation of the product is minimal after the change of media, according to (Galán, García-García, et al., 2021) (same score in both 3D and RA), as shown in Table 5 and Fig. 9. Likewise, variations in the purchasing decision responses are also minimal, as shown in Table 6. For example, for Chair A-1, the scores are the same for all three media, and for Chairs B-1 and D-1, the same scores were obtained for two media (IMG – 3D and IMG – AR, respectively). Although the standard deviation in this dataset is not small and mean values may be misleading, they are smaller than those obtained in the second case study (Table 13). Our results thus confirm hypothesis H1.1.

We highlight that differences were found in Jordan's physio- pleasure category for both studies, and that the sociological pleasure category may also be affected by the change of medium if the products are highly aesthetical. It is interesting to note that for a simultaneous evaluation, users take less time to complete the task compared to an individual evaluation, and that the results obtained can be similar in both methods. Moreover, a simultaneous evaluation helps to minimize these differences. Therefore, we propose this method of evaluation as an effective alternative to evaluate aspects related to people's cognitive and emotional reactions (psycho and ideo pleasure categories), which was confirmed by other studies, such as Lee, Kim, Chen et al. (Lee et al., 2004), who showed that the relative evaluation of the selected stimulus was the same regardless of the media used.

2.7. Conclusions

Being able to present a product effectively and understanding how it is perceived and assessed by users are critical factors for its success. Our study demonstrates that the visual medium used to present a product can influence how it is perceived and evaluated. Our results contribute to the literature of product design and engineering by empirically assessing the reliability of XR as a tool for product evaluation in the early stages of the product development process.

Certain characteristics such as comfort and size are particularly significant, as the perceptual differences elicited by different media are more pronounced. Although some studies highlight the importance of touch on the evaluation of a product, our results also show how other features that do not require haptics may also be affected by the influence of media or geometric product features (such as Jordan's socio-pleasure category). Our results have also revealed how these perceptual differences can be minimized, to a certain extent, by using joint product evaluation. In addition, certain product attributes can be emphasized in more immersive media (such as AR or VR), which is useful for both product development and point-of-sale presentation. We also emphasize the importance of having experience with AR and VR if the user will be using these technologies alone for product assessment, so that technology related anxiety does not negatively influence the evaluation.

We acknowledge the limitations of our study. First, all participants had limited experience in the use of AR, which may have influenced the evaluation of the products, especially in cases where the debriefing and follow-up interview was conducted online, instead of in person. Second, although our findings could potentially be extrapolated to similar products of the same typology, additional tests with other types of products are recommended to obtain more conclusive results. In future studies, we plan to conduct a similar experiment by changing the evaluation method (an individual evaluation for the first case study, and a simultaneous for the second). We also plan to use physiological measures such as eye-tracking technologies to analyze the user’s gaze and overall behavior more accurately and objectively during product evaluation.

Chapter 3

3. The influence of hand tracking and haptic feedback for virtual prototype evaluation in the product design process

Palacios-Ibáñez, A., Alonso-García, M., Contero, M., & Camba, J. D. (2023). The influence of hand tracking and haptic feedback for virtual prototype evaluation in the product design process. Journal of Mechanical Design, 145(4), 041403. https://doi.org/10.1115/1.4055952

The present study showcases the second work that was developed as part of this research. Product evaluation throughout the design process is a fundamental task for product success, which also helps to reduce design related costs. Physical prototyping is a common method to assess design alternatives, but often requires significant amounts of time and money. Extended Reality (XR) technologies are changing how products are presented to the user, making virtual prototyping an effective tool for product evaluation. However, it is generally assumed that our perceptual and emotional responses to a product viewed in an XR modality are comparable to those elicited by the physical product. This paper reports the results of a study where a group of participants evaluated three designs of a product (i.e., umbrella stands) when viewed in a real setting, Virtual Reality (VR), and VR with passive haptics. Our goal was to observe the influence of visual media in product perception, and how the use of a complementary item (i.e., a physical umbrella) for interaction as well as user design expertise influence product assessment.

3.1. Introduction

Conceptual design is a critical phase in the product development process to determine the cost and quality of a new product. Getting feedback from potential users during this phase is important to identify the needs and issues that must be addressed in concept validation (Coutts et al., 2019) and to fulfill design goals (Tiainen et al., 2014), as design changes are relatively inexpensive and easy to perform at this stage. The cost of making a design change increases dramatically as the product moves through its lifecycle (Ye et al., 2007).

During the product development process, the ability to fully understand a form and accurately interpret its geometry is crucial (Lau et al., 2003). Design engineers use CAD tools to conceptualize, design, visualize, and validate certain parts of the design, but the inherent limitations of displaying 3D geometry on a 2D screen make it difficult to quickly and fully understand the product features (Evans et al., 2020).

Many researchers agree that the success of a product is highly dependent on the evaluation process that is conducted throughout design, even at the early stages (Cooper, 2019). Both the designer (who can test aesthetic, functional, technical, and performance aspects of the product) and the end user (who can reveal design errors or misinterpretations of the initial requirements) are part of the evaluation process (Bordegoni, 2011a). Depending on the testing purpose, we can choose among various types of prototypes: visual prototypes, which simulate the final aesthetics of the product (e.g., sketches or photorealistic renders); form prototypes, which highlight the shape and size of the product (and are generally made with rapid prototyping or hand-made techniques); functional prototypes, which simulate product performance and can be correctly assessed using a CAD model; fully physical prototypes, which simulate product final design, aesthetics, materials, and functionality (Bordegoni, 2011a).

In this regard, physical prototyping is the most common method to obtain user feedback, but traditional physical prototyping techniques may involve large financial and time investments, and costs can increase even higher when modifications and adjustments need to be made to the prototypes. Because of this, understanding the utility of different product representation methods can help reduce time and cost (i.e., sketches, 3D representations or physical prototypes) (Hannah et al., 2012). During the design process, fidelity of the prototype may vary depending on the stage of product development (C.-H. Chu & Kao, 2020) and depends on the tests purpose (Virzi et al., 1996), as manufacturing high-fidelity prototypes has a significant impact on product development costs. This factor may significantly affect the user’s confidence and accuracy in their product evaluation (Hannah et al., 2012).

Recent advances in visualization technologies have made virtual prototyping and effective and sustainable tool for design evaluation (Aurora Berni & Borgianni, 2020; Gibson et al., 2004), especially during the early stages of development where many design variations must be produced (Cecil & Kanchanapiboon, 2007). These technologies enable the creation of high-fidelity geometric representations in a rapid and cost-effective manner for the evaluation of product aesthetics, ergonomics, and usability aspects (Bordegoni, 2011a). However, certain product features can be difficult to evaluate when using virtual prototyping techniques, particularly if physical interaction with the product is required (C.-H. Chu & Kao, 2020). In design evaluation, a key advantage of physical prototyping over virtual prototyping is the ability to feel and interact with the physical product (Kent et al., 2021). Indeed, the sense of touch is critical when evaluating a product because it provides a direct way to obtain information that could not be acquired otherwise (Ranaweera et al., 2021). Some new technologies are attempting to fill this gap in virtual environments.

The term "haptics" refers to the artificial forces between virtual objects and the user's body. They are commonly classified as passive or active. Active haptics are controlled by a computer, and forces can be dynamically manipulated to provide a wide range of feelings for simulated virtual objects (Jerald, 2015). Although some authors have argued that active haptic devices can be a good option for haptic feedback (J. Wang & Lu, 2012), others have pointed out the large financial investment that is required and the need for a much larger workspace to fit some devices to the user's hand (Kreimeier et al., 2019). Researchers have also mentioned the discomfort of wearing haptic devices which may reduce the feeling of presence and immersion in the virtual environment (Stamer et al., 2020) and have a negative impact on the user's emotional response.

Alternatively, passive haptics, provide a sense of touch in VR by synchronizing physical objects to virtual assets (Lindeman et al., 1999), significantly reducing costs without the use of intrusive devices, just the user's hands. This has been shown to increase the sense of presence, improve cognitive mapping of the environment, and improve training performance (Insko, 2001). Interactions with VR with Passive Haptics setting (VRPH) can enhance the overall user experience along with the possibility of interacting and modifying the virtual setting in real time (e.g., the environment or the product aesthetics). Although a main drawback of passive haptics has traditionally been the absence of the hands in the virtual environment, low-cost VR devices such as the Oculus Quest have introduced non-intrusive hand-tracking in their virtual experiences, where people can see a virtual representation of their hands in the digital environment that is updated in real-time. Although proper calibration is required to ensure the virtual hands are displayed correctly in the virtual environment, effective experiences can be delivered at a very low cost.

The use of VR has proven to be an effective alternative for product evaluation during the early stages of development. However, its value is predicated on the assumption that our perceptual and emotional responses to a product that is perceived through a VR environment are similar to those elicited by the actual product, which is not necessarily the case(Artacho-Ramírez et al., 2008; Palacios-Ibáñez et al., 2022), particularly when evaluating product features that rely heavily on our sense of touch (Felip et al., 2019; Galán, Felip, et al., 2021; Galán, García-García, et al., 2021). Although some researchers have incorporated mechanisms to simulate the sense of touch in VR to make product evaluations more accurate, the effect of integrating non-intrusive virtual models of the user's hands have not yet been considered.

In this paper, we report the results of an experiment where a group of participants were asked to evaluate three designs of a product (i.e., umbrella stands) in three visual media: a real setting (R), a VR environment, and a VR environment with Passive Haptics (VRPH).

Participants were also asked to indicate their intended purchasing decision and to rate their level of confidence in their response. We considered the user's background (trained in design vs. not trained in design) to determine the influence on product evaluation and the existence (or lack thereof) of a combined effect between this factor and the medium used to view the product. Finally, we examined how the use of a physical item can affect product perception, not only by helping users understand the function of the product under evaluation, but also by providing interaction opportunities to generate an experience in which users feel more engaged (Serrano et al., 2013).

3.2. Background

According to Steuer (1992), VR is "a real or simulated environment in which a perceiver experiences telepresence." It can also be defined as a "computer-generated digital environment that can be experienced and interacted with as if that environment were real" (Jerald, 2015). Although VR technologies have been around for years, recent advances have made them more affordable. VR has been successfully applied in a variety of areas, such as entertainment (Zubair et al., 2022), healthcare (Aziz, 2018), education (Camba et al., 2017), psychology and marketing (J. Park et al., 2005), architecture (Kuliga et al., 2015), and industrial design and product development (Kent et al., 2021).

In design disciplines, market competitiveness and saturation have driven companies to emphasize product attributes that address meaning and emotion (Kamil & Abidin, 2013), and how these can be expressed to positively influence the user's experience, as well as how consumers think, feel, and act (Aftab & Rusli, 2017). In current markets, the added value enabled by these attributes can significantly influence consumer choice (Li et al., 2021). Although critical, innovation is not the only factor for product success. In fact, some of the most innovative products fail when they reach the market (Kuliga et al., 2015; J. Park et al., 2005). The subtle differences between products in terms of technical characteristics, quality, and price, often make product differentiation a difficult task. In this context, consumer emotions play an important role (Pieter Desmet et al., 2001) and can significantly influence the purchasing decision (Holbrook, 1986).

Various models have been proposed to characterize product emotion (Pieter Desmet, 2007). Jordan proposed an approach with four different pleasure categories (Tiger, 1992): physical (pleasures deriving from sensory organs), social (pleasures deriving from relationships with others), psychological (pleasures related to people's cognitive and emotional reactions) and ideological (pleasures related to people's values). Alternatively, Desmet (Pieter Desmet, 2002b) applied cognitive appraisal theory to explain the process of product emotion, and Norman (Norman, 2004) explained product emotion through a neurobiological emotion-framework that distinguishes several levels of information processing: visceral, behavioral, and reflective.

Numerous studies have examined the influence of the presentation medium on product evaluation (e.g., 2D images, interactive 3D models, AR, or VR). Söderman offered some initial insights with a study that examined perceptual differences elicited by a car when viewed in a non-immersive VR environment and as a set of sketches, compared to reality (Söderman, 2005). Reid, MacDonald and Du performed a similar experiment with different 2D visual media (Reid et al., 2013), observing that the medium used to present the product could influence customer judgments. Their conclusion was confirmed by authors Artacho-Ramírez et al., who conducted an experiment (Artacho-Ramírez et al., 2008) where they used two models of loudspeakers in five different media (photographs, static image, 3D models navigable with a computer mouse, and 3D model navigable with stereoscopic images) and compared user evaluations with the corresponding real products. Their findings showed that the type of representation significantly influences product perception, which was corroborated by (M. J. Agost et al., 2021; C.-H. Chu & Kao, 2020; Sylcott et al., 2016) in their studies.

Despite the fact that researchers have investigated the influence of media into product evaluation (Ray & Choi, 2017; Tiainen et al., 2014), only recently have they begun to incorporate immersive VR headsets (e.g., Oculus Rift or HTC Vive) in their experiments (Felip et al., 2019; Palacios-Ibáñez et al., 2023). For example, Galán et al. (Galán, García-García, et al., 2021) made one of the first contributions using VRPH. Although users were not able to see their hands, their results demonstrated that the medium used to display the product influences how users perceive it, and that the use of VRPH settings could influence the product assessment positively. They also found that certain product features are more sensitive to perceptual differences, such as those related to the sense of touch. The authors also proposed a division of the semantic differential into Jordan's pleasure categories to obtain more specific results. They concluded that Jordan's approach was an effective method to evaluate product perception. For this reason, we have adopted this approach in our study.

Researchers have also explored the influence of the user's background on product evaluation when using different visual media (Forbes et al., 2018a). Design training on product evaluation or the fine arts has also been extensively studied (Whitfield & Wiltshire, 1982). Solso explored the influence of user expertise on brain activity (Solso, 2001), concluding that trained participants employed more higher-level cognitive abilities than untrained participants. Other authors have suggested that participants trained in design disciplines are more capable of identifying the relationships between the composition of an entire form and its elements (Nodine et al., 1993; Solso et al., 2007). We speculate that this factor may also have an impact on the perception of a product and should be considered when using VR, as there could be a combined effect between background and medium that affects product assessment.

3.3. Research goals and hypotheses

The main goal of our study was to determine whether the medium used to present a product influence how it is perceived and evaluated, and whether the integration of haptics can help reduce perceptual differences. To this end, we selected three different designs of a product (i.e., an umbrella stand) and asked a group of participants to assess the products as presented in R, VR, and VRPH using the Semantic Differential technique.

Some studies suggest that a simple product representation provides enough information to perform a reasonable evaluation (Ant Ozok & Komlodi, 2009). Therefore, the purchasing decision is not necessarily influenced by the change of medium. For this, users were asked to indicate their intended purchasing decision for each product. Our first hypothesis is stated as:

H1: The purchasing decision is not affected by the change of medium.

Although similar studies have shown that Jordan's physiological-pleasure category is the most affected by the change of medium (Galán, Felip, et al., 2021), other authors have argued that other categories can be significantly influenced by the medium (Palacios-Ibáñez et al., 2022). Therefore, the following hypothesis is stated:

H2: The visual medium used to present a product influences user perception regardless of the classification in Jordan's category.

Furthermore, some researchers have claimed that haptic feedback can affect the users' confidence in their evaluations (Grohmann et al., 2007), (Peck & Childers, 2003). Our third hypothesis is stated as:

H3: The level of Confidence in the Response is affected by the change of visual medium.

Another factor that has not been sufficiently studied is the influence of the user's background on product assessment. Some studies have reported that in product evaluation individuals with a design-related background tend to employ more higher-level cognitive abilities than untrained individuals (C.-H. Chu & Kao, 2020). These differences may influence the evaluation of the product. Our fourth hypothesis is states as:

H4: User expertise and design background affects product evaluation.

Finally, in some product evaluation scenarios, it is necessary to use a complementary item to correctly understand how a product performs. Since the use of complementary items (when needed) is not feasible in online channels or even in some physical stores, an important part of the information of the product is lost. We considered this factor as an interesting research goal, as more information can improve product evaluation (Hannah et al., 2012). Our fifth hypothesis is stated as:

H5: The use of a complementary item to assess a product influences the evaluation of the product semantic scales.

3.4. Materials and methods

To test our hypotheses, we structured our study into two analyses A1 and A2. In the first analysis (A1), we compared the user's perception of three umbrella stand designs as presented in three different settings (R, VR, and VRPH) and determined the influence of the user's previous design expertise on product assessment. In the second analysis (A2), we examined the perceptual differences in product evaluation elicited by the presence or absence of a complementary physical item (i.e., an umbrella) in two different media (R and VR).

A total of 91 users participated in our experiment. All users gave informed consent to participate in our study, which was approved by the CEENB-OMGs section of the Bioethics Committee of the University of Cadiz (Ref. 006/202). To describe our experimental conditions, we next discuss (1) our case study, (2) materials, (3) the semantic scale established for the evaluation of the different products, (4) the description and justification of the sample, and (5) the protocol followed for the development of the case study.

3.4.1. Case study

The different umbrella stands selected for our study are shown in Fig. 11. We selected this type of product for three main reasons: 1) it is an easy-to-use product, which avoids frustration and prevents it from affecting our results (C.-H. Chu & Kao, 2020); 2) thanks to the geometry of the design options selected, hand-tracking was not lost, which guaranteed a consistently good experience; and 3) it lends itself to the use of a complementary item (i.e., un umbrella), as this is one of the research goals of our study.

Although similar studies have limited the number of options, we decided to use three prototypes to obtain more general results, since more products meant an increase in the time of the experiment, and people's fatigue could affect the evaluation.

Each stand has an opening to accommodate both long and short umbrellas (at least four) and is equipped with a water tray. All three models had neutral colors to mitigate potential perceptual differences elicited by color (Hagtvedt & Adam Brasel, 2017). We note that umbrella stand B was displayed in the university hall for 30 days as a part of a student's product design exhibition. For A2, two different umbrellas were selected (a long and a short umbrella), as shown in Fig. 12. The three umbrella stands were used for both analyses, but the umbrellas were only used in A2 to determine the impact of user-object interaction in the perception of the umbrella stands.

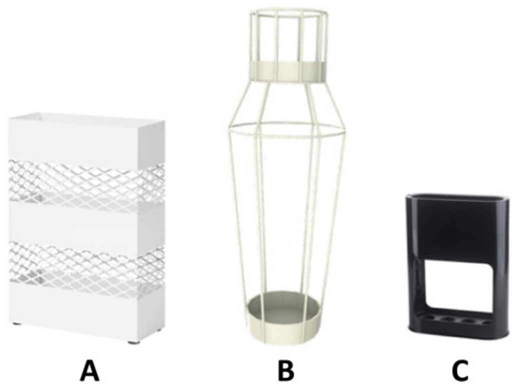

Fig. 11: Umbrella stands used as stimuli. **Fig. 12**: Umbrellas used for the second case study.

For the three experimental conditions (the viewing media), each product was arranged in the same manner and in identical spaces. Since only one physical model of each umbrella stand was available, two physical rooms were built to present the three scenarios: one empty room and one with the physical products, as shown in Fig. 13. The interior area of the rooms (both physical and virtual) was 5 square meters. Each physical room was built using seven movable panels positioned contiguously and attached to a 3-metre-long wall.

Fig. 13: Exterior view of the rooms (1): Interior as perceived by users in R (2) and interior as perceived by users in VR and VRPH (3).

The scenarios are described below:

1 - Scenario 1 (R): This room consisted in a real environment with the products on the floor. Users were able to view and touch the real objects, but not allowed to change their position. For A2, two physical umbrellas were placed on a small table. The user was allowed to grab and move the umbrellas to interact with the umbrella stands.

2 - Scenario 2 (VR). This room consisted in a VR simulated environment on an HMD. Oculus Quest hand-tracking interaction was enabled, so the users were able to see a non-intrusive virtual representation of their hands in real-time while interacting with the products as no external devices were used apart from the HDM (Fig. 14), which allowed a more natural interaction with the objects and the environment. The products were anchored to the virtual floor, so they could not be moved. In the case of the A2, virtual replicas of the two umbrellas were added to the scene and placed on a virtual table. The umbrellas could be grabbed and moved to interact with the umbrella stands.

3 - Scenario 3 (VRPH): This scenario consisted in a VR environment displayed on an HMD where the position of the virtual products was synchronized with the position of real products, so haptic feedback was enabled. In this scenario, the physical and virtual products were fixed to the floor. Hand-tracking interaction was enabled, so the users were able to see a non-intrusive virtual representation of their hands in real-time while interacting with the products.

It is important to note that scenarios 1 and 3 were displayed in the same physical room (as shown in Fig. 13.1), but never at the same time since only one participant was allowed to be present during the evaluation of the products in one medium. The use of the physical room alternated between the two scenarios.

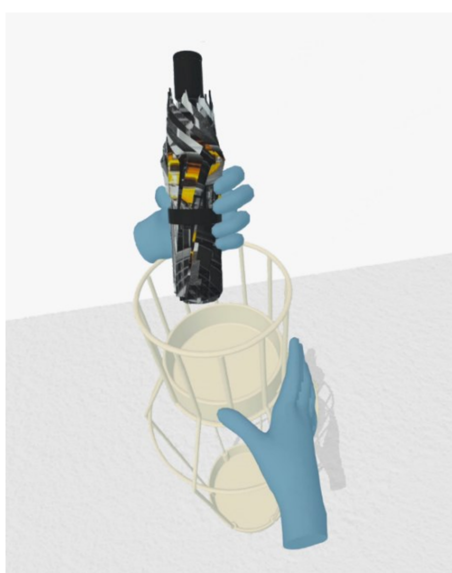

Fig. 14: User touching an umbrella stand and grabbing the small umbrella.

3.4.2. Materials

The VR environments were experienced using an Oculus Quest 2 HMD upgraded to version 36.0, a stand-alone immersive VR device with a Single Fast-Switch LCD display of 1832×1920 pixels per eye and a refresh rate of 72Hz. The virtual environment was designed using Unity 2020.3.11f1. We used the Oculus Integration asset (version 36.0) and HPTK Posing and Snapping 2.0.0. asset for the hand tracking interaction (as the Oculus Interaction SDK was not available when the experiment was run). The Passthrough Capability was enabled for the calibration of the virtual objects before starting the experiment, and the hand tracking interaction capability was enabled to provide the non-intrusive virtual models of the user's hands. The scene used a Realtime light with hard shadows enabled, and materials were built using a Standard Shader. The virtual objects were modeled in SolidWorks 2020, and UV mapping was completed in Blender 2.93.0. Hygiene and disinfection material were also provided to ensure optimal sanitary conditions.

For product evaluation, the participants completed a questionnaire comprised of 12 semantic differential scales. Participants were asked to indicate their intended purchasing decision with a "Yes" or "No" answer and rate their level of confidence in their response. To measure user presence in the VR environments, we used the Slater-Usoh-Steed (SUS) presence questionnaire (Slater et al., 1998). This instrument is comprised of six 7-point Likert scale questions where a higher score indicates higher levels of presence.

3.4.3. Semantic differential

The Semantic Differential (Osgood et al., 1957) approach is a common method for product evaluation (S. H. Hsu et al., 2000; Vergara et al., 2011) and an effective procedure to obtain consumer opinion and preferences. It is comprised of Likert-type scales of 5 or 7 points that typically use bipolar pairs of adjectives to describe the product that is being evaluated. In our study, we used the procedure by Hsu et al. (S. H. Hsu et al., 2000) to generate the semantic space for our products, as described next.

Information was collected from four different sources (professional designers, design-related users, average users, and manufacturers) to match a general criterion. The sample included 28 volunteers (20 female and 8 male) with an average age of 33.4 years. The group of professional designers (with at least 5 years of experience) consisted of 8 people; the design-related users' group (people with a design background, such as industrial design students, design researchers or professors) was comprised of 12 participants; the average user's group consisted of 8 volunteers. Product adjectives were selected from various sites such as Ikea, Amazon, B-line, Systemtronic, or Mox.

Participants were shown a set of 12 images of various umbrella stands (Fig. 15) and asked to evaluate each product using an online form. Next, we conducted a keyword analysis. First, the frequency with which each adjective was repeated was counted, and those with the same meaning were grouped together (e.g., "big" and "large"). Antonyms were also grouped to build the most frequent bipolar pairs of adjectives (for terms where no antonym was available, they were added by the authors to complete the definition of the bipolar pair). We classified the bipolar pairs based on the four categories of pleasure defined by Jordan (Jordan, 2002), and the two most frequent bipolar pairs were selected for each category, which resulted in a total of eight bipolar pairs of adjectives, as shown in Table 22.

Fig. 15: Umbrella stands used to define the semantic differential scale.

Table 22: List of bipolar pairs of selected adjectives classified by Jordan's pleasure categories.

Physio	Psycho	Socio	Ideo
Light – Heavy	Simple – Complex	Attractive – Unattractive	Inexpensive - Expensive
Large – Small	Practical – Impractical	Traditional – Modern	Elegant - Ordinary
Stable - Unstable	Functional - Decorative	Minimalist - Overelaborated	Common - Original

3.4.4. Sample

Two different samples were used in our study. To estimate the minimum sample size, we performed an a priori power analysis with G*Power (Faul et al., 2007). For A1, a repeated measures ANOVA test was applied with the following input parameters: effect size: 0.25, α=0.05, (1-β)=0.80 and 1 group. For A2, a repeated measures ANOVA (within-between interaction) was applied with the following input parameters: effect size: 0.25, α=0.05, (1-β)=0.80, 2 group. Our results estimated a total sample size of 28 in both cases.

3.4.4.1. Sample for A1 analysis

Prior to data processing, an outlier study was conducted to obtain more robust and reliable results. Users with low confidence levels in their responses (a score with a mean value <.60) in at least one medium for each umbrella stand who also appeared as an outlier in the Semantic Differential Scale data set were excluded from the analysis. Our final sample size was 58 users, so a power of 0.80 was guaranteed.

The mean age of the sample was 20.9 years old. The sample consisted of 39 men and 19 women, where 41 of the volunteers were involved in industrial design disciplines either academically and/or professionally. Approximately half of the participants (50.8%) had no experience with VR before participating in the study, 37.3% of the participants claimed to have limited experience with VR, and only 12% rated their experience with VR as significant. 6.7% of the participants used VR environments frequently and 3.4% very frequently. Finally, 55.9% of the participants stated that they have visual problems. 52.5% reported having myopia, 11.8% astigmatism and 3.4% hyperopia and other problems. 35.5% wore glasses and 18.6% wore contact lenses.

3.4.4.2. Sample for A2 analysis

Since the goal of this analysis was to compare the user's perception of a product in the presence or absence of a complementary item, some participants from A1 were also included in A2. A total of 23 participants who completed product evaluations in R and VR media (regardless of order) in the first or second place were randomly selected. This way, we ensured that the evaluation of products through VRPH did not affect our results. In addition, 23 new participants evaluated the products through R and VR using the complementary item.

The 46 participants in A2 guarantee a power of 0.80. The mean age was 20.63 years (56.5% were men and 43.5% were women); 56.5% were involved in industrial design disciplines either academically and/or professionally; 58.7% of the participants had never used VR devices before the study and 32.6% claimed to have vision problems; 54.4% reported myopia; 17.4% astigmatism and 6.5% hyperopia and other problems; 26.1% wore glasses and 26.1% contact lenses.

3.4.5. Experimental Protocol

The complete procedure followed by the participants in our study is illustrated in Fig. 16, differentiating between those participating in A1 and A2. The experiment was conducted in 5 consecutive days (22.5 hours). Twenty-seven 50-minute shifts were established, allowing the participation of up to 3 users per shift. Participants accessed the area reserved for the experiment in pre-established groups. This area consisted of two physical rooms (described above) and a table with chairs reserved for completing the required documentation and questionnaires. Data for A1 was collected during the first four days, and for A2 on the last day.

Before the experiment, participants were provided with the necessary information about the study and given the opportunity to ask questions. All participants gave informed consent. A member of the research team positioned the virtual objects for Scenarios 2 and 3 using the Passthrough Capability of the Oculus Quest 2 ensuring an accurate overlay of the virtual and real objects for Scenario 3.

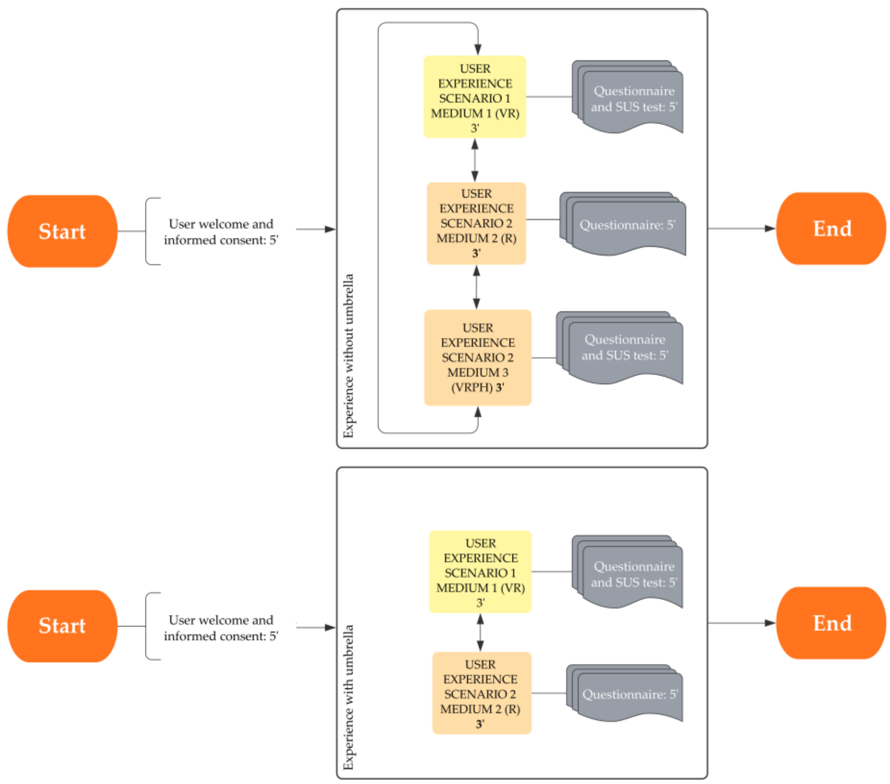

Fig. 16: Complete cycle performed by each participant in both studies.

Each participant viewed the products in the different proposed media. Before entering Scenarios 2 and 3, a member of the research team confirmed that the user was correctly seeing the virtual models of their hands. After each medium, participants were asked to fill out the evaluation questionnaire before moving to the next experimental condition (Fig. 17.1). The order of the viewing media was randomized for each participant to minimize any potential unwanted effect on the results. Two participants completed the activities in each session, each accompanied by a member of the research team. Inside the room, participants were allowed to interact with the umbrella stands (Fig. 17.2 and 17.3). In the second study, participants were allowed to use two physical umbrellas to interact with the products (Fig. 17.4.)

Fig. 17: Participants filling out questionnaires (1), participant experiencing the VR environment for A1 (2), participant experiencing the VRPH environment in A1 (3), and participant evaluating products in R using the umbrellas in A2 (4).

3.5. Results

This section has been divided into two sections in terms of the analyses performed (A1 and A2). A schematic diagram is shown in Fig. 18 to facilitate the understanding of the section.

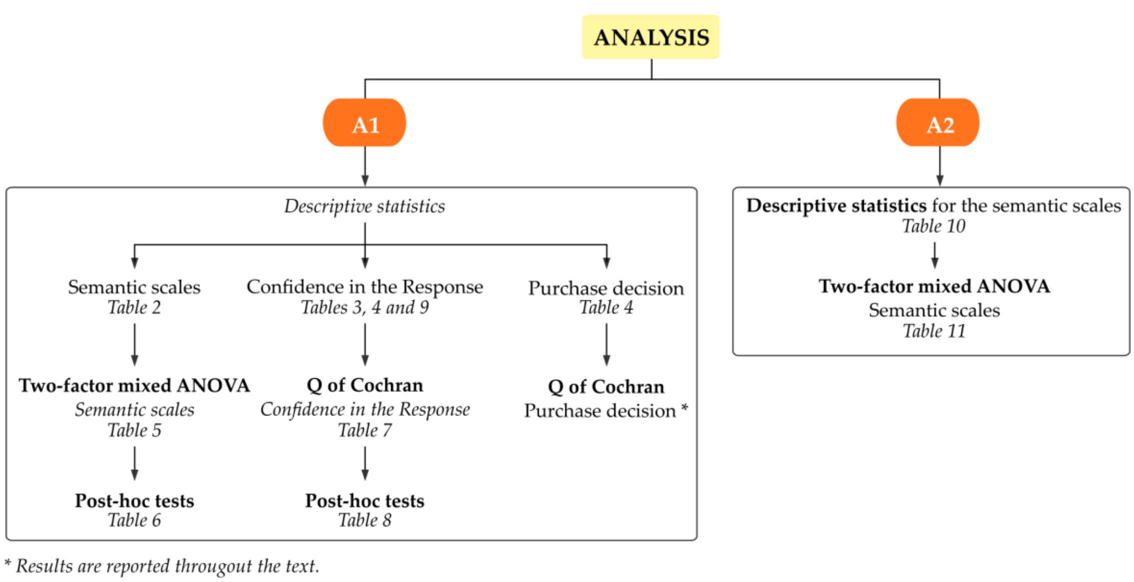

Fig. 18: Schematic diagram for the data analysis.

3.5.1. Analysis A1 results

Four data sets were obtained: the (1) Semantic Scales, (2) Confidence in the Response, (3) Purchase Decision, and (4) Presence test scores. The descriptive statistics for Semantic Scales, Purchase Decision, and the Confidence in the Response are shown in Tables 21 – 23. Boxplots for the Semantic Scales are shown in Fig. 19.

For the Semantic Scales, a value closer to -3 represents a closer correspondence with the adjective in italics, and a value closer to 3 indicates a closer correspondence with the adjective in bold. For the Purchase Decision and the Confidence in the Response, a value closer to 0 indicates "not buying the product" and a low Confidence in the Response, whereas a value closer to 1 represents "buying the product" and a high Confidence in the Response. Descriptive statistics for the SUS presence questionnaire were $M_{VR} = 4.701$, $Mdn_{VR} = 5$, $SD_{VR} = 1.62$ for VR, and $M_{VRPH} = 5.10$, $Mdn_{VRPH} = 5$, $SD_{VRPH} = 1.63$ for VRPH.

Table 23: Descriptive statistics for the semantic scales.

	Semantic scales		A			B			C		
			R	VR	VRPH	R	VR	VRPH	R	VR	VRPH
PHYSIO	*Heavy* – **Light**	M	**.36**	.16	*.03*	-.47	*-.51*	**-.41**	.67	**.90**	*.74*
		Mdn	1.00	.00	.00	-1.00	-1.00	-1.00	1.00	1.00	1.00
		SD	1.66	1.42	1.34	1.76	1.66	1.59	2.13	1.55	1.75
	Small – **Large**	M	**.21**	*-.16*	.17	**2.22**	2.00	*1.98*	-1.74	*-1.88*	**-1.66**
		Mdn	.00	.00	.00	2.00	2.00	2.00	-2.00	-2.00	-2.00
		SD	1.02	1.28	1.03	.73	.70	.91	1.05	.82	1.10
	Unstable - **Stable**	M	**1.47**	1.43	*1.38*	*.97*	*.85*	**1.28**	.76	*.67*	**.88**
		Mdn	2.00	2.00	2.00	2.00	1.00	2.00	1.00	1.00	1.00
		SD	1.37	1.30	1.31	1.57	1.42	1.36	1.66	1.42	1.46
PSYCHO	*Simple* – **Complex**	M	*-.59*	**-.41**	-.55	*-.33*	-.31	**-.05**	-.40	*-.83*	**-.38**
		Mdn	-1.00	.00	-1.00	-1.00	.00	.00	-1.00	-1.00	-1.00
		SD	1.27	1.44	1.22	1.66	1.66	1.46	1.62	1.43	1.61
	Impractical – **Practical**	M	**1.07**	*.66*	1.02	**1.53**	**1.53**	*1.52*	.79	*.52*	**.91**
		Mdn	1.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00	1.00
		SD	1.15	1.18	1.19	1.30	1.25	1.22	1.44	1.31	1.33
	Decorative – **Functional**	M	**1.02**	*.59*	.95	*.12*	**.31**	.19	**1.02**	*.69*	.86
		Mdn	1.00	1.00	1.00	.00	.00	.00	1.00	1.00	1.00
		SD	1.33	1.44	1.42	1.68	1.64	1.53	1.32	1.33	1.23
SOCIO	*Unattractive*– **Attractive**	M	.17	*.05*	**.19**	**1.26**	1.16	*1.10*	.26	*.10*	**.41**
		Mdn	.50	.00	.50	1.00	1.00	1.00	.00	.00	1.00
		SD	1.60	1.54	1.55	1.32	1.36	1.39	1.43	1.55	1.35
	Traditional – **Modern**	M	*.26*	**.53**	.28	*.62*	**1.00**	.78	**1.38**	*1.07*	1.28
		Mdn	.00	1.00	1.00	1.00	2.00	1.00	2.00	1.50	1.00
		SD	1.38	1.56	1.45	1.76	1.73	1.64	1.47	1.57	1.17
	Minimalist – **Overelaborated**	M	*-.76*	-.86	**-.69**	*-1.19*	**-1.10**	-1.12	**-1.28**	**-1.28**	**-1.28**
		Mdn	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00	-1.00
		SD	1.05	1.25	1.31	1.29	1.47	1.35	1.20	.99	1.23
IDEO	*Inexpensive* – **Expensive**	M	-.76	**-.69**	*-.78*	*-.16*	**-.03**	-.05	*-.53*	-.43	**-.38**
		Mdn	-1.00	-1.00	-1.00	.00	.00	.00	-1.00	.00	.00
		SD	1.14	1.19	1.24	1.25	1.21	1.23	1.50	1.27	1.50
	Ordinary – **Elegant**	M	**.05**	*-.12*	-.05	**1.29**	*1.10*	*1.10*	**.53**	*.36*	.47
		Mdn	.00	.00	.00	1.50	1.00	1.50	1.00	.00	.00
		SD	1.29	1.44	1.30	1.06	1.29	1.40	1.42	1.27	1.27

Semantic scales		A			B			C		
		R	VR	VRPH	R	VR	VRPH	R	VR	VRPH
Common – **Original**	M	**-.28**	*-.14*	**-.28**	1.48	*1.10*	**1.50**	**.76**	.69	*.67*
	Mdn	.00	.00	.00	2.00	2.00	2.00	1.00	1.00	1.00
	SD	1.40	1.50	1.62	1.51	1.72	1.40	1.57	1.54	1.48

For each umbrella stand (A, B, C) and semantic scale, the highest value for mean is shown in bold and the lowest one in italics.

Table 24: Descriptive statistics for the confidence in the response by semantic scale.

	Semantic scales		A			B			C		
			R	VR	VRPH	R	VR	VRPH	R	VR	VRPH
PHYSIO	*Heavy* – **Light**	M	**.86**	*.52*	.77	**.91**	*.60*	.79	**.88**	*.55*	.79
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.35	.50	.42	.28	.49	.41	.33	.50	.41
	Small – **Large**	M	**1.00**	*.97*	**1.00**	**1.00**	*.95*	*.95*	**1.00**	*.98*	*.98*
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.00	.18	.00	.00	.22	.22	.00	.13	.13
	Unstable - **Stable**	M	**.95**	*.72*	.90	.90	*.76*	**.95**	**.90**	*.69*	.84
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.22	.45	.31	.31	.43	.22	.31	.47	.37
PSYCHO	*Simple* – **Complex**	M	**.98**	.97	*.95*	.97	*.93*	**.98**	**.95**	*.91*	**.95**
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.13	.18	.23	.18	.26	.13	.22	.28	.22
	Impractical – **Practical**	M	**.86**	*.78*	.83	**.97**	.90	*.88*	**.79**	*.67*	.78
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.35	.42	.38	.18	.31	.33	.41	.47	.42
	Decorative – **Functional**	M	**.95**	*.83*	.84	**.95**	.88	*.84*	**.90**	*.66*	.76
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.22	.38	.37	.22	.33	.37	.31	.48	.43
SOCIO	*Unattractive* – **Attractive**	M	**1.00**	.97	*.95*	*.95*	.98	**.97**	**.97**	*.91*	.95
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.00	.18	.23	.22	.13	.18	.18	.28	.22
	Traditional – **Modern**	M	.90	**.97**	*.90*	.97	**.98**	*.93*	**.97**	*.90*	.91
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.31	.18	.31	.18	.13	.26	.18	.31	.28
	Minimalist – **Overelaborated**	M	**.98**	*.93*	.97	**.97**	.95	*.93*	**.98**	.95	*.93*
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.13	.26	.19	.18	.22	.26	.13	.22	.26
IDEO	*Inexpensive* – **Expensive**	M	.48	*.41*	**.49**	**.52**	*.40*	**.52**	**.47**	*.41*	.45
		Mdn	.00	.00	.00	1.00	.00	1.00	.00	.00	.00
		SD	.50	.50	.50	.50	.49	.50	.50	.50	.50
	Ordinary – **Elegant**	M	*.90*	.91	**.97**	**.97**	*.91*	**.97**	.89	*.84*	**.90**
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.31	.28	.27	.18	.28	.18	.31	.37	.31
	Common – **Original**	M	**.91**	*.85*	**.91**	.93	*.91*	**.95**	**.86**	*.81*	.84
		Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		SD	.28	.37	.34	.26	.28	.22	.35	.40	.37

For each umbrella stand (A, B, C) and semantic scale, the highest value for the mean is shown in bold and the lowest one in italics.

Table 25: Descriptive statistics for the purchase decision and the confidence in the response.

		A			B			C		
		R	VR	VRPH	R	VR	VRPH	R	VR	VRPH
Purchase decision	M	.34	*.31*	**.41**	*.67*	.74	**.76**	.52	*.43*	**.53**
	Mdn	.00	.00	.00	1.00	1.00	1.00	1.00	.00	1.00
	SD	.48	.47	.50	.47	.44	.43	.50	.50	.50
Confidence in the Response	M	**.90**	*.82*	.87	**.92**	*.85*	.89	**.88**	*.77*	.84
	Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	SD	.30	.39	.34	.28	.36	.32	.33	.42	.37

For each umbrella stand (A, B, C) and data set, the highest value for the mean is shown in bold and the lowest one in italics.

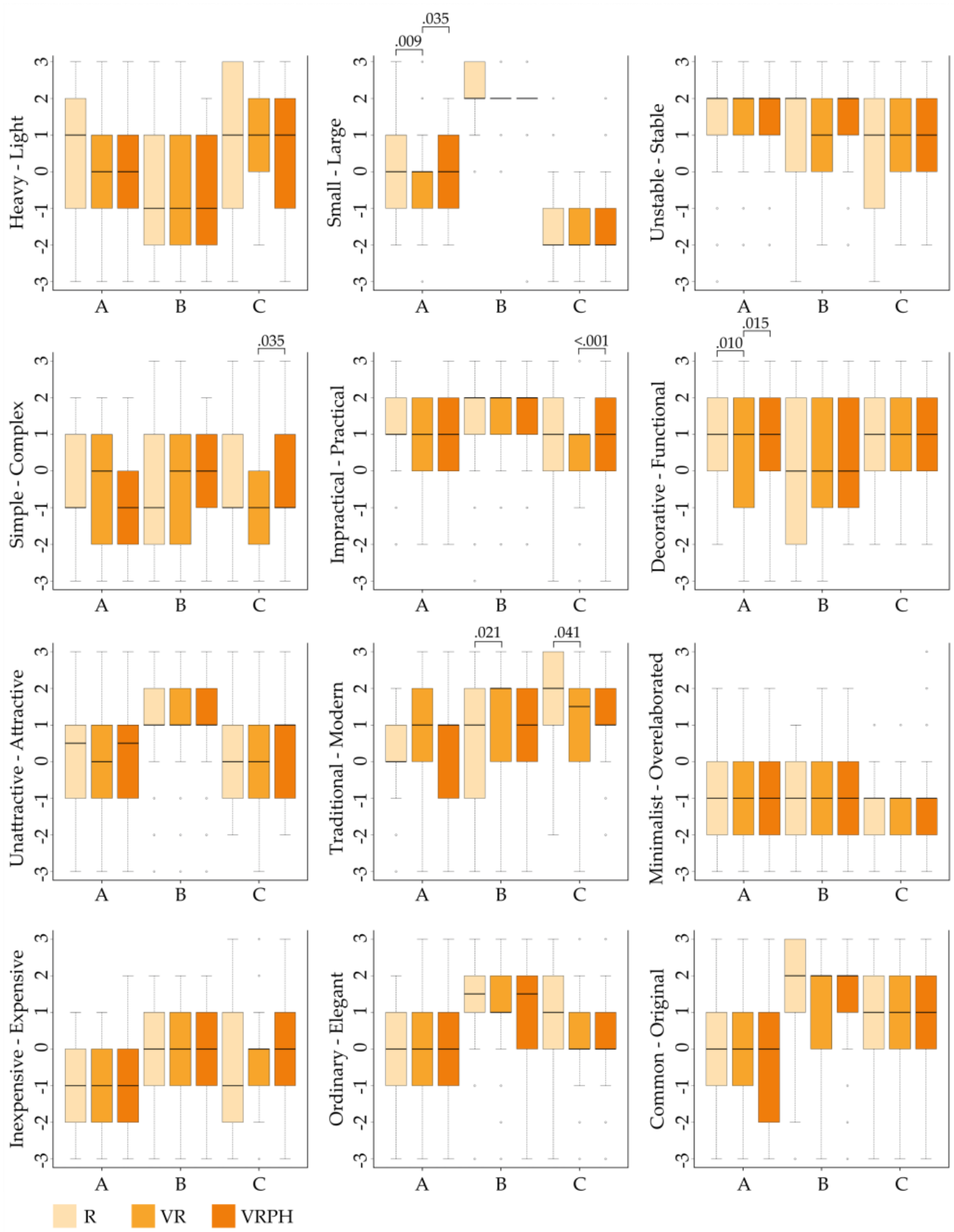

Fig. 19: Boxplots for the Semantic Scales.

A normality test was performed to select the appropriate statistical test to determine differences between means for data sets (1), (2), and (3). We used a Kolmogorov–Smirnov normality test (significance level of .05), as the sample size was >50. Our results showed that the data was not normally distributed, so parametric tests proved unsuitable. For the case of the Semantic Differential, we decided to apply the Aligned Rank transform (ART) procedure (Higgins et al., 1990), a powerful and robust nonparametric alternative to other traditional techniques (Mansouri et al., 2004). ART relies on a preprocessing step that "aligns" data before applying averaged ranks. After this, common ANOVA procedures can be applied. For the Confidence in the Response and the Purchase Decision, Q of Cochran test was used (dichotomous data).

The results of the Two-factor Repeated Measures ANOVA test (with media as within subjects' factor and user background as between subject's factor) for the semantic scales are shown in Table 26. Statistically significant differences were found for "Small - Large" and "Decorative – Functional" adjectives for design A. The bipolar pair "Traditional – Modern" showed significant differences for B and C. Finally, design C showed significant differences for "Simple – Complex" and "Practical – Impractical."

Umbrella stand A was the only design for which the influence of the participant's background was significant in some adjectives. Also, a combined factor of medium and background was obtained for "Simple – Complex" and "Ordinary – Elegant" for design A, for "Inexpensive – Expensive" for design B and C, and for "Decorative – Functional" for C.

To determine the conditions between which these differences were found for the factor of the presentation medium, we conducted pairwise comparisons for the semantic scales (Table 27) using the Tukey's honestly significant difference (HSD) test. To use this test, each group must have approximately equal variance (homogeneity of variances). In our case, no between subject factors were specified, so the assumption is always met when using the Levene's test to assess equality of variances.

Table 26: Two-factor mixed ANOVA for the semantic scales with media and participant's design expertise (exp) as main factors.

		A			B			C		
	Semantic Scales	Media	Exp	Mixed	Media	Exp	Mixed	Media	Exp	Mixed
		Sig.	Sig.	Sig.	Sig.	Sig.	Sig.	Sig.	Sig.	Sig.
PHYSIO	Heavy – Light	p=.080	p=.178	p=.050	p=.677	p=.682	p=.637	p=.957	p=.875	p=.290
	Small – Large	**p=.002**	p=.683	p=.687	p=.234	p=.807	p=.581	p=.144	p=.511	p=.510
	Unstable - Stable	p=.241	p=.775	p=.288	p=.078	p=.578	p=.326	p=.655	p=.317	p=.647
PSYCHO	Simple – Complex	p=.932	p=.685	**p=.003**	p=.153	p=.914	p=.470	**p=.018**	p=.706	p=.393
	Impractical – Practical	p=.037*	p=.569	p=.271	p=.112	p=.736	p=.256	**p=.001**	p=.233	p=.098
	Decorative – Functional	**p=.004**	p=.215	p=.128	p=.218	p=.129	p=.586	p=.285	p=.669	**p=.002**

SOCIO	Unattractive – Attractive	p=.467	p=.055	p=.792	p=.641	p=.210	p=.470	p=.042*	p=.342	p=.450
	Traditional – Modern	p=.724	**p=.009**	p=.559	**p=.014**	p=.226	p=.319	**p=.046**	p=.081	p=.532
	Minimalist – Overelaborated	p=.545	p=.275	p=.075	p=.104	p=.795	p=.774	p=.161	p=.897	p=.364
IDEO	Inexpensive – Expensive	p=.545	p=.222	p=.084	p=.121	p=.284	**p<.001**	p=.339	p=.809	**p=.036**
	Ordinary – Elegant	p=.193	**p=.038**	**p=.049**	p=.662	p=.305	p=.270	p=.220	p=.687	p=.585
	Common – Original	p=.161	p=.263	p=.137	p=.332	p=.786	p=.391	p=.672	p=.355	p=.576

*Factor p value for each umbrella stand (A, B, C) and semantic scale in which perceptual differences were found are shown in bold. Values with * showed no differences in the post-hoc analysis.*

Table 27: Post-hoc tests for the semantic scales.

	Semantic Scales	Media	A	B	C
			Sig.	Sig.	Sig.
PHYSIO	Small – Large	R – VR	**p=.009**		
		R – VRPH	p=.854		
		VR – VRPH	**p=.035**		
	Unstable – Stable	R – VR			
		R – VRPH			
		VR – VRPH			
PSYCHO	Simple – Complex	R – VR			p=.136
		R – VRPH			p=.580
		VR – VRPH			**p=.035**
	Impractical – Practical	R – VR	p=.085		p=.164
		R – VRPH	p=.980		p=.131
		VR – VRPH	p=.062		**p<.001**
	Decorative – Functional	R – VR	**p=.010**		
		R – VRPH	p=.999		
		VR – VRPH	**p=.015**		
SOCIO	Unattractive – Attractive	R – VR			p=.538
		R – VRPH			p=.255
		VR – VRPH			p=.056
	Traditional – Modern	R – VR		**p=.021**	**p=.040**
		R – VRPH		p=.257	p=.252
		VR – VRPH		p=.286	p=.696

P values for each umbrella stand (A, B, C) and semantic scale showing perceptual differences are shown in bold.

To determine differences in the Confidence in the Response, we performed the Q of Cochran test for non-parametric dichotomous data. Our results (Table 28) showed that the level of confidence is affected by the medium for adjectives “Heavy – Light” and “Unstable – Stable” for each product. Statistically significant differences were also found for Umbrella stand C for “Decorative – Functional”. Post-hoc tests are shown in Table 29.

Table 28: Q of Cochran for the level of confidence in the response.

	Semantic scales	df	A		B		C	
			Q	Sig.	Q	Sig.	Q	Sig.
PHYSIO	Heavy – Light	2	26.000	**p<.001**	22.455	**p<.001**	24.250	**p<.001**
	Small – Large		4.000	p=.125	4.500	p=.105	1.000	p=.607
	Unstable - Stable		14.623	**p=.001**	13.857	**p=.001**	13.000	**p=.002**
PSYCHO	Simple – Complex		1.200	p=.549	4.667	p=.097	.800	p=.670
	Impractical – Practical		2.111	p=348	3.500	p=.174	3.600	p=.165
	Decorative – Functional		5.733	p=.057	5.091	p=.078	11.043	**p=.004**
SOCIO	Unattractive – Attractive		3.500	p=.174	1.200	p=.549	1.400	p=.497
	Traditional – Modern		2.909	p=.234	2.333	p=.311	2.889	p=.236
	Minimalist – Overelaborated		2.800	p=.247	.857	p=.651	3.600	p=.165
IDEO	Inexpensive – Expensive		2.211	p=.331	4.455	p=.108	.824	p=.662
	Ordinary – Elegant		1.273	p=.529	2.250	p=.325	1.385	p=.500
	Common – Original		2.167	p=.338	.857	p=.651	1.385	p=.500

P values for each umbrella stand (A, B, C) and semantic scale showing perceptual differences are shown in bold.

Table 29: Post-hoc tests for level of confidence in the response for each bipolar pair of adjectives.

	Semantic Scales	Media	A	B	C
			Sig.	Sig.	Sig.
PHYSIO	Heavy – Light	R – VR	**p<.001**	**p<.001**	**p<.001**
		R – VRPH	p=.180	**p=.039**	p=.227
		VR – VRPH	**p=.001**	**p=.007**	**p=.001**
PSYCH	Unstable - Stable	R – VR	**P=.002**	P=.057	**P=.002**
		R – VRPH	p=.250	p=.250	p=.453
		VR – VRPH	**p=.013**	**p=.001**	**p=.035**
SOCIO	Decorative – Functional	R – VR			**p=.001**
		R – VRPH			**p=.039**
		VR – VRPH			p=.238

P values for each umbrella stand (A, B, C) and semantic scale showing perceptual differences are shown in bold.

To examine the influence of the participant's background on the Confidence in the Response, we report the descriptive statistics of the two groups of participants: design-trained (T) and non-design-trained (NT). The T group was comprised of 41 people, whereas NT was comprised of 17 participants. Results are shown in Table 30.

Table 30: Descriptive statistics for the confidence in the response by participant's design expertise.

	A						B						C					
	R		VR		VRPH		R		VR		VRPH		R		VR		VRPH	
	DT	NDT	DT	NDT	DT	NDT	DT	NDT	DT	NDT	DT	NDT	DT	NDT	DT	NDT	DT	NDT
M	.89	.92	.83	.79	.87	.89	.92	.91	.86	.81	.88	.91	.86	.92	.76	.80	.81	.91
Mdn	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
SD	.31	.28	.38	.41	.35	.31	.28	.28	.35	.39	.33	.29	.34	.28	.42	.40	.39	.29

Finally, no significant differences were found for the intended purchasing decision, as indicated by Cochran's Q test ($\chi2(2) = 4.67$, $p_A<.097$), ($\chi2(2) = 4.20$, $p_B<.122$), ($\chi2(2) = 3.20$, $p_C<.212$).

3.5.2. Analysis A2

To determine the influence of the use of a complementary item, i.e. a physical umbrella, on the evaluation of each umbrella stand, a Two-factor Mixed ANOVA with 2 degrees of freedom was performed, with the secondary element (Obj) as the between-subjects factor, and the presentation medium as the within-subjects factor (Table 32). Descriptive statistics for the semantic scales are shown in Table 31.

Table 31: Descriptive statistics for the semantic differential organized by presence (U) or absence (NU) of the umbrella.

		A				B				C			
		R		VR		R		VR		R		VR	
Semantic Scales		U	NU	U	NU	U	NU	U	NU	U	NU	U	NU
Heavy – Light	M	-.30	.39	.30	.65	**-.43**	-.39	**.61**	-.17	1.00	**.96**	**.35**	.91
	Mdn	-1.00	.00	1.00	1.00	-1.00	-1.00	.00	.00	1.00	2.00	.00	1.00
	SD	1.58	1.44	1.15	1.40	2.09	1.85	1.73	1.67	2.04	2.12	1.75	1.59
Small – Large	M	**-.13**	.35	**-.30**	-.17	2.26	**1.96**	2.13	**1.70**	**-2.39**	-1.52	**-1.96**	-1.52
	Mdn	.00	.00	.00	.00	2.00	2.00	2.00	2.00	-3.00	-2.00	-2.00	-2.00
	SD	1.25	.78	1.36	1.07	.92	0.71	1.06	.70	.72	1.34	1.02	.79
Unstable – Stable	M	**1.13**	1.78	1.43	**1.26**	1.78	**0.61**	1.09	**.39**	.83	**.65**	**.78**	1.13
	Mdn	1.00	2.00	2.00	2.00	2.00	1.00	1.00	1.00	2.00	1.00	1.00	1.00
	SD	1.60	1.28	1.44	1.39	1.48	1.64	1.50	1.23	1.95	1.70	1.78	1.32
Simple – Complex	M	**-1.43**	-.35	**-1.74**	-.26	**.00**	.26	**-.09**	-.04	**-.52**	-.26	**-.78**	-.39
	Mdn	-2.00	.00	-2.00	.00	.00	1.00	.00	.00	.00	-1.00	-1.00	-1.00
	SD	1.44	1.23	1.05	1.32	1.54	1.79	1.68	1.36	1.27	1.68	1.78	1.62
Impractical – Practical	M	**0.13**	1.13	1.09	**0.30**	2.39	**1.30**	2.04	**1.22**	.87	**.83**	**.78**	.83
	Mdn	.00	2.00	1.00	1.00	3.00	2.00	2.00	2.00	2.00	1.00	1.00	1.00
	SD	1.71	1.39	1.31	1.36	0.72	1.55	1.22	1.28	2.03	1.40	1.78	1.07
Decorative – Functional	M	.87	**.83**	.96	**.22**	.52	**-0.13**	0.13	**.00**	1.17	**1.13**	**1.09**	0.91
	Mdn	1.00	1.00	1.00	.00	1.00	.00	.00	.00	2.00	1.00	2.00	1.00
	SD	1.39	1.59	1.58	1.44	1.44	1.60	1.39	1.45	1.64	1.18	1.73	1.16
Unattractive – Attractive	M	**-.13**	.61	**.13**	.30	1.48	**1.39**	1.78	**1.17**	**.26**	.35	**-.22**	.17
	Mdn	.00	1.00	.00	.00	2.00	1.00	2.00	1.00	.00	.00	.00	.00
	SD	1.74	1.59	1.36	1.46	1.41	1.08	1.38	1.15	1.39	1.37	1.20	1.61

		A				B				C			
		R		VR		R		VR		R		VR	
Semantic Scales		U	NU	U	NU	U	NU	U	NU	U	NU	U	NU
Traditional – Modern	M	.35	**.30**	**.65**	**.65**	.91	**.57**	1.26	**1.09**	**1.39**	1.61	1.09	**.91**
	Mdn	.00	.00	1.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00
	SD	1.72	1.29	1.43	1.50	1.68	2.02	1.45	1.83	1.37	1.44	1.24	1.73
Minimalist – Overelaborated	M	**-1.22**	-.65	**-1.52**	-.83	**-1.61**	-.91	**-1.65**	-1.17	-1.96	**-1.70**	-1.57	**-1.39**
	Mdn	-1.00	-1.00	-2.00	-1.00	-2.00	.00	-2.00	-1.00	-2.00	-2.00	-2.00	-2.00
	SD	1.20	1.11	.99	1.34	1.27	1.41	1.34	1.40	.88	.97	.99	.84
Inexpensive – Expensive	M	**-.91**	-.87	**-1.04**	-.70	.43	**-.04**	.61	**.22**	**-.26**	-.57	-.70	**-.39**
	Mdn	-1.00	-1.00	-1.00	-1.00	.00	.00	1.00	.00	.00	-1.00	-1.00	.00
	SD	.90	1.22	1.26	1.02	1.16	1.19	.99	1.13	1.36	1.50	1.15	1.08
Ordinary – Elegant	M	.09	**.00**	.35	**.04**	1.61	**1.48**	1.43	**1.22**	.83	**.70**	.70	**.30**
	Mdn	.00	.00	.00	.00	2.00	2.00	2.00	1.00	1.00	1.00	1.00	.00
	SD	1.65	1.31	1.11	1.33	1.27	.99	.90	1.31	1.34	1.29	1.52	1.22
Common – Original	M	**-.78**	-.61	**-.74**	-.17	2.04	**1.70**	2.04	**1.35**	.96	**.91**	**.57**	.65
	Mdn	-1.00	-1.00	-1.00	.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00	1.00
	SD	1.68	1.50	1.45	1.27	1.07	1.15	.88	1.61	1.80	1.53	1.34	1.37

For each umbrella stand (A, B, C) and semantic scale, the highest value for the mean is shown in bold.

Table 32: Two-factor mixed ANOVA for the semantic scales with medium and complementary item (obj) as main factors.

		A			B			C		
	Semantic Scales	Media	Obj	Mixed	Media	Obj	Mixed	Media	Obj	Mixed
		Sig.	Sig.	Sig.	Sig.	Sig.	Sig.	Sig.	Sig.	Sig.
PHYSIO	Heavy – Light	p=.096	p=.104	p=.496	**p=.013**	p=.488	p=.118	p=.081	p=.500	p=.323
	Small – Large	**p=.034**	p=.392	p=.452	p=.153	**p=.009**	p=.583	p=.056	**p=.012**	p=.606
	Unstable - Stable	p=.610	p=.643	p=.057	**p=.013**	**p=.004**	p=.398	p=.583	p=1.000	p=.377
PSYCHO	Simple – Complex	p=.936	**p<.001**	p=.586	p=.266	p=.758	p=.603	p=.339	p=.439	p=.619
	Impractical – Practical	p=.771	p=.679	**p<.001**	p=.164	**p=.004**	p=.962	p=.675	p=.360	p=.794
	Decorative – Functional	p=.333	p=.242	p=.129	p=.530	p=.226	p=.284	p=.482	p=.476	p=.758
SOCIO	Unattractive – Attractive	p=.734	p=.345	p=.143	p=.750	p=.166	p=.061	p=.078	p=.589	p=.392
	Traditional – Modern	p=.052	p=.739	p=.731	**p=.046**	p=.771	p=.577	**p=.003**	p=.743	p=.373
	Minimalist – Overelaborated	p=.167	p=.075	p=.923	p=.266	p=.113	p=.391	**p=.044**	p=.196	p=.931
IDEO	Inexpensive – Expensive	p=.992	p=.432	p=.346	p=.169	p=.135	p=.905	p=.559	p=.976	p=.069
	Ordinary – Elegant	p=.642	p=.975	p=.618	p=.113	p=.513	p=.845	p=.245	p=.380	p=.342
	Common – Original	p=.181	p=.266	p=.406	p=.269	p=.156	p=.582	p=.064	p=.949	p=.547

Factor p value for each umbrella stand (A, B, C) and semantic scale in which perceptual differences were found are shown in bold.

3.6. Discussion

In our first hypothesis (H1), we stated that the purchasing decision was not affected by the change of medium. Based on the results of our study, it seems that the presentation medium does not have a significant effect on the intended purchasing decision for umbrella stands (no significant differences were found between means for any of the products: $p_A<.097$,

p_B<.122, p_C<.212). Our findings agree with other studies that suggest that although users need as much information as possible to make a purchasing decision (O'Keefe & McEachern, 1998), even a simple presentation medium provides sufficient details to perform a reasonable evaluation (Ant Ozok & Komlodi, 2009). In our case, considering that a 7-point Likert scale was used for product evaluation, the scores obtained between different media (Table 23) did not vary significantly. In addition, the Confidence in the Response scores (Table 24 - 23) were also high in the three visual media, which suggests that participants had sufficient information to make a purchasing decision.

Our results differ from those obtained in a similar study with coffee makers conducted by Palacios-Ibáñez et al. (Palacios-Ibáñez et al., 2022), where significant differences between means were found for the purchasing decision. These differences may be explained by the average age of the selected samples, which was very limited (mean age of 20 years old). In our study, the selected product may have not elicited a purchasing decision in any medium for umbrella stand A and C (Table 6). For the case of umbrella stand B, previous familiarity with the product could have influenced the results, as stated by Söderman (Söderman, 2005). Therefore, hypothesis H1 is rejected for the case of umbrella stands and our experimental conditions. To draw more generalizable conclusions, however, it is necessary to expand the study to other product typologies.

We note, however, that the descriptive statistics for this dataset (Table 25) show higher values in the VRPH condition, which may indicate that presenting the product in this type of setting may lead to an increase in intended purchasing decision. The user's confidence in the response also shows higher values in this medium, which may have had a positive effect on the purchasing decision.

In H2, we postulated that the medium used to display the product influences user perception. Our results (Table 26) revealed significant differences between means for a subset of the bipolar pairs of adjectives in different categories, which confirms H2 for the case of umbrella stands. However, we cannot state which specific adjectives are influenced by the medium, as none of the adjectives were affected in more than one product. In the case of the umbrella stands, we note that certain categories are more sensitive to be affected by the presentation medium than others. Our results are in line with those by other authors. Artacho-Ramírez et al. obtained similar results when studying the influence of the graphical representation in the evaluation of different models of a loudspeaker (Artacho-Ramírez et al., 2008), and Agost et al. reached similar conclusions when assessing two different types of furniture (M.-J. Agost et al., 2021). Palacios-Ibáñez et al. also reported similar results when evaluating three different designs of a coffee maker (Palacios-Ibáñez et al., 2022) and Galán et al. (Galán, Felip, et al., 2021) did the same using an armchair as the main stimuli. This set of studies demonstrate that regardless of the product selected as stimulus, some characteristics may be more difficult to evaluate depending on the means used to assess it.

The psycho-pleasure category (pleasures related to people's cognitive and emotional reactions) was highly affected by the presentation medium (Table 26). In this category, design A showed significant differences for "Decorative - Functional" between R – VR and VRPH – VR, whereas design C showed significant differences for "Simple – Complex" and "Impractical – Practical" between VR – VRPH (Table 27). In any case, scores in the VRPH setting were similar to the R setting (Table 23), which suggests that the haptic feedback may have influenced the evaluation of the product. Although previous studies have established that the sense of touch affects mostly adjectives related to the physio-pleasure category (derived from sensory organs) (Galán, García-García, et al., 2021), its positive influence on the sense of presence may have also had an impact on product evaluation, as passive haptics significantly enhance virtual environments and can affect cognitive processes (Insko, 2001). VR scored the lowest in the presence questionnaire as well as the level of confidence in the response for the affected adjectives.

Umbrella stand A also showed significant differences (Table 26) for "Small – Large" (physiological pleasure). This result is in line with other authors (Galán, Felip, et al., 2021; Galán, García-García, et al., 2021), who reported that pleasures deriving from sensory organs are better assessed with the addition of physical interaction. In this case, it was expected that differences appeared between R – VR and VRPH – VR (Table 8), as VR was the condition that offered the least amount of information due to the absence of tactile feedback. The scores for the VRPH condition were similar to those in the R setting, which was expected, since the product data collected from both means were similar (physical interaction and being able to view the product at full scale). Although we expected to find differences between umbrella stands B and C for this pleasure category and behave similarly to design A, it is important to note that the studies mentioned earlier were limited to only one product design and typology, so the findings are not necessarily generalizable to all designs or products as other factors may also impact product assessment (Achiche et al., 2014).

Finally, for B and C, significant differences were found for "Traditional – Modern" between R – VR (Table 27). This semantic scale is part of the sociological pleasure category (pleasures deriving from relationships with others), which was the second most affected category and the only one that was affected for design B. These results could be explained by the participants' previous knowledge of this product, as umbrella stand B was displayed in the university hall for several days. Previous studies reported that previous knowledge of the product can minimize perceptual differences between means (Söderman, 2005).

In H3, we speculated that the Confidence in the Response is affected by the presentation medium. Cochran's Q test results (Table 28 and 27) showed differences between means for adjectives related to the physio-pleasure category, which confirms H3. The descriptive statistics for the Confidence in the Response (Table 24) also showed higher values for R, followed by VRPH and VR, respectively, which means that haptics can help to increase the participants' Confidence in the Response. Some authors have argued that haptic feedback during product evaluation increases the user's confidence in the response (Grohmann et al., 2007), which explains why the VR setting obtained the lowest scores for this dataset, and why statistically significant differences were found for those product features that relied heavily on the sense of touch. We note that significant differences were also found for umbrella stand C for "Decorative – Practical" between R – VR and R – VRPH. Some participants admitted during the experiment that they were unsure whether the product was an umbrella stand, so this aspect could have been difficult to evaluate.

Our results show that in 64% of the cases, the mean scores obtained in the VRPH medium for the Semantic Differential were similar to the R setting. Descriptive statistics for the Confidence in the Response (Table 24 - 23) show the same result for 78% of the cases. According to Grohmann and Spangenberg (Grohmann et al., 2007), the introduction of tactile feedback has a positive impact on this factor, reducing perceptual variations. This claim is supported by the results in our two-factor ANOVA for the Semantic Differential (Table 26) and the post-hoc tests (Table 27). Regardless of the adjectives with statistical differences, no differences were found between R and VRPH. Finally, we highlight that the presentation of the product in a VRPH setting can positively influence the user's purchasing decision, for the case of umbrella stands. Considering the average age of our sample, it would be interesting to test whether these results are consistent with older adults, since technology acceptance tends to be negatively associated with age (Hauk et al., 2018).

Our results show that the use of touch during the virtual experience could be an effective alternative for product evaluation in the specific case of umbrella stands. Studies with other product typologies are required to draw more generalized conclusions.

To test H4 (how user expertise and design background affect product evaluation), we performed a Two-factor Mixed ANOVA for the Semantic Scales (Table 26). Results showed that "Traditional – Modern" and "Ordinary – Elegant" were influenced by the participant's design background for design A, which means that this factor is not associated with a specific design typology or adjectives but can affect the perception of certain product characteristics. The psycho and ideo categories showed a combined factor between the medium and the participant's design expertise. Some studies have reported that in product evaluation individuals with a design-related background tend to employ more higher-level cognitive abilities than untrained individuals (Solso, 2001). Therefore, the influence of design expertise on the psycho-pleasure category was expected. The descriptive statistics for the Confidence in the Response showed that having a design background does not increase the participant's Confidence in the Response toward product evaluation, which aligns with the study of Forbes et al. (2018) and confirms H4. Making an assessment about specific product features is a subjective task, so having a background in design does not necessarily increase this level confidence.

Finally, a Two-factor Mixed ANOVA (Table 32) was performed to test H5 (the use of a complementary item to assess a product influences the evaluation of the product semantic scales). Our goal was to create a richer and more sophisticated interactive experience which is sometimes limited in online media to determine the impact on product perception. Our results showed that the presence of the complementary item is a factor that significantly influences product perception, which confirms H5. For umbrella stands, the physio-pleasure category was highly affected by this factor. The use of a physical umbrella while evaluating the main stimuli influenced the assessment of "Small – Large" for umbrella stands B and C. The descriptive statistics (Table 31) show that designs A and C were perceived as smaller when using the umbrella, whereas product B was perceived as larger. Likewise, Jordan's psycho-pleasure category was the second most affected using the complimentary item. The psycho-pleasure category is related to people's cognitive and emotional reactions toward a product, so the complementary item may have provided additional information about the product's performance. The descriptive statistics show that the umbrella stands are perceived as more practical when the umbrella is used (67% of the cases), so a more sophisticated experience where all the necessary elements for product testing are present can improve product perception thanks to the availability of more information (Hannah et al., 2012). A summary of our results is shown in Table 33.

Table 33: Summary of results and findings of the study.

Hypothesis	Accept/Reject	Comments
H1	Rejected	The purchase decision was not influenced by the presentation medium for the case of umbrella stands for two possible reasons: the participant may have had enough information to make a purchase decision in each experimental condition, or the selected product may not have triggered a purchase decision in any medium.
H2	Accepted	User perception is influenced by the visual medium used to present the product. Significant differences between means were found for a subset of adjectives in different categories, but we cannot state which specific adjectives are influenced by the medium, as none of the adjectives were affected in more than one product.
H3	Accepted	Confidence in the Response was affected by the presentation medium. Differences between means were found for adjectives related to the physio-pleasure category, and descriptive statistics for this dataset showed that haptics can help increase the participants' confidence in the response.
H4	Accepted	Two pairs of adjectives were influenced by the participant's design background for one of the designs. Making an assessment about specific product features is a subjective task, so having a background in design does not necessarily increase this level of confidence.

Hypothesis	Accept/Reject	Comments
H5	Accepted	The use of a complementary element (when needed) improved the evaluation of the product thanks to the additional information provided.

3.6.1. Implications for design practice

Traditional prototyping processes in product development can be costly. Although new means of representation have emerged in recent years, several authors have shown that the user's perceptual response to a particular product may not be the same depending on the medium used to present it.

Some researchers have begun to explore virtual experiences that leverage the sense of touch, but little research has examined the impact of displaying non-intrusive virtual hand models in the environment. The capability is currently available on low-cost VR devices, but its impact on product evaluation has not yet been studied.

In our study, the hypothesis on the purchasing decision could not be confirmed due to certain limitations of our experiments. Nevertheless, our results highlight the need for designers to conduct user studies on potential customers. We have also shown that the user's background (i.e., designers vs. non-designers) can influence product assessment. Therefore, the characteristics of the potential buyers must be reflected in the sample.

For our dataset, the VRPH medium scored higher, which suggests that the medium is important when presenting products at the point of sale and that, in some cases, the right medium can increase purchasing decisions. On a practical level, both designers and retailers should allow users to physically interact with their products, since it can result in more favorable emotional responses.

On the other hand, the medium used to present a product may result in perceptual differences during evaluation, as some features of the real product may be difficult or impossible to replicate virtually. In our case, leveraging the sense of touch seems to minimize these differences. The use of passive haptics in the virtual experience for product evaluation may be an effective strategy. From a practical point of view, for some type of products, 3D printers can be used to create physical mockups that enable touch capabilities in the VR environment. In this regard, future development of object tracking capabilities and more accurate hand tracking performance of HMDs will facilitate the creation of VRPH experiences.

The user's confidence in the response during the virtual experience also increases with the use of tactile feedback, which can have a positive impact during the evaluation. Finally, based on the results related to the use of the complementary item, it is important for designers and marketing professionals provide consumers with all the necessary elements to interact with the product at the time of evaluation. Otherwise, valuable information may remain hidden, distorting the emotional response of the user toward the product.

3.7. Conclusions

Participants in our study were exposed to three products (i.e. umbrella stands) in different visual media: a real setting (R), VR, and VRPH. For the evaluation, participants used a Semantic Differential composed by 12 bipolar pairs of adjectives divided by Jordan's pleasure categories. In addition, they were asked to provide a purchase decision and rate the level of confidence in their responses. We also gathered information about the user presence in the virtual environments using the SUS presence questionnaire.

Our results demonstrate that the visual medium used to present a product can influence how it is perceived. We also showed that an individual's background as well as the use of a complementary item during product evaluation (when needed) can also influence product perception.

Our results showed that not all pleasure categories in Jordan's classification were influenced equally by the visual medium. Although some studies highlight the importance of touch on the evaluation of a product, our study also showed that features that do not require haptics can also be affected by the medium. For the case of umbrella stands, the psycho-pleasure category was the most affected, and the ideo-pleasure category was the only one not influenced by the medium.

Our results showed that the differences in the semantic scales between visual media were due to the absence of the sense of touch during the virtual experience. The absence of tactile feedback had a negative influence on the user's level of confidence in the response, which influenced the results. Therefore, it is proposed that the use of VRPH can be an effective tool for product evaluation for product conceptualization, particularly in experiences that leverage the sense of touch.

Our study also demonstrated that an individual's background does not influence the Confidence in the Response, but it can influence the assessment of certain product features, particularly in the psychological and ideological pleasure category. Finally, the use of a complementary item (when needed) for product evaluation can also affect product perception, as additional information on product performance becomes available.

We acknowledge the limitations of our study. First, the particular characteristics of the participants in our sample make our findings only applicable to equivalent user groups. Second, some participants were familiar with one of the designs selected as stimulus. Finally, although our findings could potentially be extrapolated to similar products of the same typology, additional tests with a more diverse set of products are recommended to draw more conclusive results.

Our work contributes to the area of product design by empirically assessing the reliability of VR and VRPH as a tool for product evaluation in the early stages of the product development. In future studies, we plan to use physiological measures such as eye-tracking technologies to analyze user behavior more accurately and objectively during product evaluation activities.

Chapter 4

4. An Examination of the Relationship between Visualization Media and Consumer Product Evaluation

Palacios-Ibáñez, A., Pirault, S., Ochando-Martí, F., Contero, M., & Camba, J. D. (2023). An Examination of the Relationship between Visualization Media and Consumer Product Evaluation. IEEE Transactions on Visualization and Computer Graphics. https://doi.org/10.1109/TVCG.2023.3238428

This chapter presents the third experimental work developed during this research. In line with previous cases, we report two case studies in which a group of participants evaluated three designs of two product typologies (i.e., a desktop telephone and a coffee maker) as presented in three different visual media (i.e., photorealistic renderings, AR, and VR for the first case study; and photographs, a non-immersive virtual environment, and AR for the second case study) using eight semantic scales. An inferential statistical method using Aligned Rank Transform (ART) proceedings was applied to determine perceptual differences between groups.

4.1. Introduction

Manufactured products play a substantial role in our daily lives (Donald, 2004). People buy, collect, and surround themselves with different objects, sometimes to express different aspects of their personalities (Ortíz Nicolás et al., 2013). Current markets are highly competitive (Roy et al., 2009) and people are often faced with a variety of options that can satisfy their basic needs in terms of quality, price, and function (Perez Mata et al., 2017). As a result, affective values (J. Singh & Sarkar, 2022), which have been extensively examined by researchers in the field of emotional design, are now becoming a product differentiation tool (Qu & Guo, 2019).

Product evaluation is an essential activity in the early stages of the development (Ozer, 1999). Obtaining feedback from potential customers is essential to identify the design issues that must be resolved before validating design concepts (Coutts et al., 2019; Tiainen et al., 2014). These processes require continuous product evaluations which are usually conducted with physical models and prototypes whose level of fidelity may vary widely depending on the design phase and testing purpose (C.-H. Chu & Kao, 2020; Virzi et al., 1996).

The cost of design changes increases dramatically as a product moves through its lifecycle (Ye et al., 2007). In this regard, effective evaluations can help identify potential issues early in the design process. Additionally, prototyping may involve considerable financial and time investments with limited flexibility to modifications (Söderman, 2005). In large scale production environments, for example, prototypes can take months to produce, and even cease to be valid representations of the product at the time of evaluation (Arrighi & Mougenot, 2019).

Virtual prototyping is an affordable and versatile alternative to physical prototyping (Cecil & Kanchanapiboon, 2007). High-fidelity virtual prototypes, which have been shown to positively influence user's confidence and accuracy in product evaluation (Hannah et al., 2012), can be produced faster and more cost-effectively than traditional methods.

In increasingly competitive markets where e-commerce is becoming more prevalent (Jeong et al., 2009), the manner in which a product is portrayed and presented to the user can be a key differentiating factor. Static images, text, and other common means of representation are often insufficient to convey all the information related to a product, especially in terms of the experience that the product can afford (Jiang & Benbasat, 2007b). Furthermore, product displays in physical stores are gradually being replaced by digital media in online platforms through which the different characteristics of the product must be conveyed (Yoo & Kim, 2014).

Emerging visualization technologies such as Virtual Reality (VR) are changing the manner in which products are presented to the user and helping consumers form a clearer understanding of complex products (Kinzinger et al., 2022). These technologies are rapidly evolving in terms of hardware, software, usability, ergonomics, quality, and efficiency. They are establishing themselves as effective mechanisms to represent virtual models in various design applications (A. Berni et al., 2020; Hoermann & Schwalm, 2015; Tesch & Dörner, 2020). Augmented Reality (AR) (Bimber & Raskar, 2005), for example, is being widely adopted in industry (Greengard, 2022) to enhance consumer experiences by combining virtual assets with real content (Suh & Prophet, 2018) to elicit specific emotions (Beck & Crié, 2018).

The availability and affordability of extended reality technologies, both in terms of hardware (e.g. Quest 2, Pico 4) and software (e.g. the "Metaverse") have fueled their adoption in product development settings (Aurora Berni & Borgianni, 2020; Meta, 2021) and are shaping the way we work and collaborate. Likewise, continuous advances in smartphone technology have favored access to mixed reality environments. Indeed, many companies have implemented AR in their online catalogs as an effective product representation tool (e.g., Ikea, Sephora, or L'Oreal) (Kim & Choo, 2021).

From an impact standpoint, it is important to understand how these technologies influence the subjective impressions of users about a particular product, i.e. the manner in which a product is perceived, interpreted, and internalized by the consumer, as they may vary significantly depending on the presentation platform (Artacho-Ramírez et al., 2008; Galán, García-García, et al., 2021). Extended reality can be an effective tool to optimize the product development process (Bordegoni, 2011a) as well as a means to provide additional information to potential customers during purchasing, particularly in online environments (Ant Ozok & Komlodi, 2009). However, its value is predicated on the assumption that our subjective impressions and emotional responses to a virtual prototype are similar to those elicited by the real product, which is not necessarily the case (Felip et al., 2019; Galán, García-García, et al., 2021). Therefore, in order to obtain the most accurate evaluation, it is critical to consider how the representation medium can influence the user's emotional response.

The present study contributes to advance our understanding of how visual media influences the various dimensions of the perceptual space linked to a product and whether consumer-grade extended reality technologies can be an effective tool for product evaluation both (1) during the NPD process (an environment in which the experimental conditions and the context in which the product is displayed are controlled), and (2) at the point of sale (where there is limited control over the user's physical environment and devices). In this paper, we discuss two experimental studies in which a group of participants used the Semantic Differential method to evaluate designs of two product typologies presented in different media (photorealistic renders, AR, and VR for the first case study; and real photographs, a non-immersive virtual environment, and AR, for the second case study).

4.2. Related work

The role of emotion is critical for providing a meaningful user experience and influencing consumer choices (Pieter Desmet et al., 2001; Li et al., 2021). Different approaches have been proposed to characterize product emotion (Pma Desmet, 2011). Most notably, Jordan suggested four pleasure categories (Tiger, 1992): physiological-pleasure (deriving from sensory organs), sociological-pleasure (deriving from relationships with others), psychological-pleasure (related to people's cognitive and emotional reactions), and ideological-pleasure (related to people's values). Alternatively, Desmet (Pieter Desmet, 2002b) applied cognitive appraisal theory to explain the process of product emotion, while Norman described product emotion by distinguishing three levels of information processing: visceral, behavioral, and reflective (Donald, 2004).

For over two decades, various studies have examined how different media can influence the user's emotional response in product evaluation, ranging from simple 2D images and interactive 3D models displayed on computer screens, to AR and VR devices with different levels of immersion and realism. For example, Söderman (Söderman, 2005) examined the perceptual differences elicited by viewing a car in non-immersive VR and as a set of sketches, versus reality. The author found no significant differences among the interaction methods and attributed this finding to the prior knowledge that the participants may have had about the product, as they were potential consumers interested in it. Karlsson et al. (Engelbrektsson et al., 2000) concluded that experience and prior knowledge of the product (or a similar product) plays a critical role in product evaluation, and Schoormans et al. (Schoormans, J. P., Ortt, R. J., & De Bont, 1995) suggested that prior knowledge of the product enables users to unconsciously fill in missing information. Similarly, Reid et al., (Reid et al., 2013) stated that prior knowledge about the product's dimensions could influence the user's decisions, a factor that was considered in other studies to minimize deviations (Artacho-Ramírez et al., 2008; C.-H. Chu & Kao, 2020).

Artacho-Ramírez et al. (Artacho-Ramírez et al., 2008) made further advances by presenting two models of loudspeakers in five different media (photographs, static infographic imagery, an interactive 3D model, and stereoscopic images) and comparing user evaluations with the corresponding real products. The authors concluded that the type of representation significantly influences the user's subjective impressions of the product.

It is important to note that different representation methods do not afford the same possibilities for interaction with a product. More sophisticated media usually provide higher levels of interaction, so it can be expected that these perceptual differences stem from the inherent differences between media, as demonstrated by Ozok and Komlodi (Ant Ozok & Komlodi, 2009). In their study, the authors found significant differences with respect to the information provided by 2D images and non-interactive and interactive CAD objects.

Various researchers have begun to incorporate immersive virtual reality headsets (e.g., Oculus Rift, HTC Vive) in their experimental studies, such as Forbes et al. (Forbes et al., 2018a), who evaluated the perception of three armchairs using prints of a rendered CAD model, a 3D interactive CAD model, AR, VR, as well as real settings with and without tactile interaction. The authors concluded that, although virtual prototyping cannot completely replace physical prototypes, it can provide sufficient information to filter out poor design concepts before producing physical prototypes. In product evaluation scenarios, immersive VR technology can also highlight aspects of the product that would go unnoticed in a real setting. Furthermore, different levels of immersion can affect how certain characteristics of the product such as size are perceived (Galán, Felip, et al., 2021; Heineken & Schulte, 2007). We highlight the study by Felip et al. (Felip et al., 2019), who observed that product evaluation in VR differed significantly from physical evaluations when using passive haptics. In their study, evaluation scores were generally higher when performed in the virtual environment, which could be explained by the novelty effect (all the participants in the study had limited to no experience with product presentations in VR). The previous work was expanded by Galán et al. (Galán, García-García, et al., 2021) who observed that although some product features were affected by the change of medium, the overall product evaluation remained unaffected by this factor. Additionally, the authors noted that the introduction of touch during the virtual experience could positively influence the user's opinion of the product.

However, comparatively few studies have investigated the perceptual differences that may be elicited by AR (Banerjee et al., 2021). In this regard, Ray and Choi (Ray & Choi, 2017) investigated how AR representations differ from other kinds of concept representations. The authors identified user interface challenges but also emphasized the opportunity for studies to explore the role of AR in the design lifecycle. Other studies have suggested that younger participants tend to consider AR more helpful during the purchasing decision process than 2D renderings (M. J. Agost, 2020). Agost et al. (M.-J. Agost et al., 2021) studied how the presentation media (2D renderings, a 360-degree display technique, AR, and VR) influenced product evaluation in online shopping environments. The researchers reported that some users had difficulty using AR and VR in their experiment, but highlighted the value of AR for evaluating certain products such as large appliances. According to a recent study (Palacios-Ibáñez et al., 2023), AR can increase the user’s level of confidence in the response during product evaluation, as long as users do not experience technology-related anxiety (Arvanitis et al., 2009), and the change of medium does not generally affect the purchasing decision. Experiments, however, are often limited to a single type of product or design option.

Little is known about how gender may affect product evaluations (Lin et al., 2019; Van Slyke et al., 2002). Studies have shown that women generally have a more positive attitude toward conventional shopping than they do toward online shopping (Dittmar et al., 2004), which may be explained by the lower cognitive attitude toward this modality (Hasan, 2010). The literature discussed in this section illustrates the need for further research on the influence of the medium, particularly AR, on the evaluation of different product typologies.

4.3. Research goal and hypotheses

The goal of our study was to analyze the influence of the presentation mechanism in product evaluation to determine whether differences in subjective impressions exist when a product is presented in different media. In addition to the presentation medium, we also considered gender as a factor in perceptual variations as well as how product design within the same typology can affect the evaluation of certain bipolar pairs of the Semantic Differential. In our studies, a group of consumers was asked to view and evaluate a product in three different settings.

One main hypothesis was postulated: the medium used to present a product influences how the user evaluates the semantic scales regardless of their classification in Jordan's categories

(H1). Two complementary hypotheses were also postulated: a particular design within the same product typology influences the user's subjective impressions of the product (H2); and gender differences exist in the evaluation of a product and how it is perceived (H3).

4.4. Materials and methods

Two case studies were conducted to test the hypotheses. For each case study, three designs of a particular product typology with clear morphological differences were used: desktop telephones for the first case study and coffee makers for the second. Both case studies consisted of a within-subject study where participants were allowed to view the product in three different visual media, and asked to evaluate it using semantic scales and rate it using a 5-point Likert scale.

For the first case study (desktop telephones), our goal was to examine the influence of the representation technique on product evaluation *in the design process*. The product was displayed using photorealistic renderings, AR, and VR. The physical room was identical for all users to eliminate the potential influence of external factors on the evaluation. An attempt was made to minimize the differences in the product placement context between experimental conditions.

For the second case study (coffee makers), the goal was to examine the influence of the medium during the evaluation of the product in an online assessment scenario. Due to COVID-19 restrictions, some interviews were conducted online. Furthermore, a certain level of control over the product placement context was lost (i.e., environmental noise, the complementary objects within the evaluation scene, and the exact device used for evaluation), which is common in online shopping scenarios.

4.4.1. Case study I: desktop telephones

Three representative designs of a desktop telephone were selected for this experiment: Swissvoice Epure 2 (Fig. 20.a), Daewoo DTD-1400 W (Fig. 20.b), and Philips M110w (Fig. 20.c). To test the hypotheses, three studies were designed. For each case, a particular design of a desktop telephone was used. Each product was presented in three different media:

- Photorealistic images (Fig. 21.a), which display multiple points of view of the product. Images were displayed on a computer screen.
- AR, where the virtual product was presented in a real environment. The product can be viewed from any angle (Fig. 21.b) but no interaction was allowed. The product was displayed on a smartphone.

- VR, where the virtual product was presented in a virtual room with neutral colors and dimensions of 4 x 3.5 x 2.5 m. The product was placed on a table in the center of the room (Fig. 21.c). Interaction was not allowed.

Fig. 20: Swissvoice Epure 2 (a), Daewoo DTD-1400W (b), and Philips M110w (c) desktop telephones. Images of the actual phones from vendors' websites.

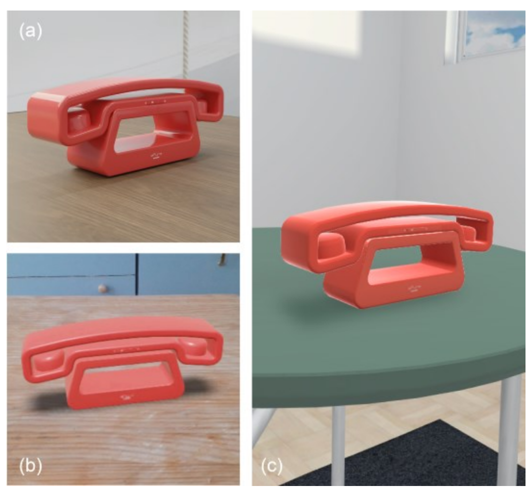

Fig. 21: Swissvoice Epure 2 displayed in different media: photorealistic render (a), AR environment (b), and VR environment (c).

4.4.2. Case study II: coffee makers

Three different designs of coffee makers were select-ed: Nespresso Essenza (Fig. 22.a), Moka Pot (Fig. 22.b), and Nespresso Inissia (Fig. 22.c). Each product was presented in three different media:

- Photographs of the product (Fig. 22), which dis-played multiple points of view of the product. Images were displayed on a computer screen.

- A non-immersive virtual environment (N-IVE), which displayed a 3D model of the product on a table in a virtual environment. User interaction with the product was not allowed but the user was allowed to navigate the environment using the computer mouse and key-board (Fig. 23.a).
- AR, where the VP was presented in a physical en-vironment. The product could be viewed from any angle (Fig. 23.b), but no interaction was allowed. The product was displayed on a smartphone (Fig. 24).

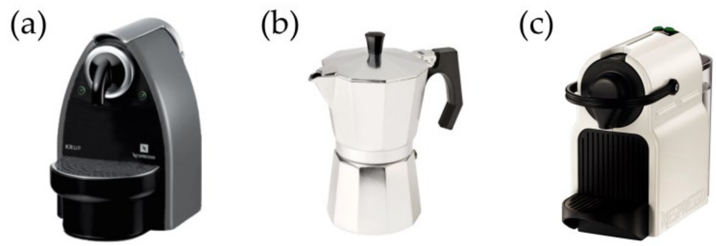

Fig. 22: Nespresso Essenza (a), Moka Pot (b), and Nespresso Inissia (c) coffee makers. Images of the physical coffee makers from vendors' websites.

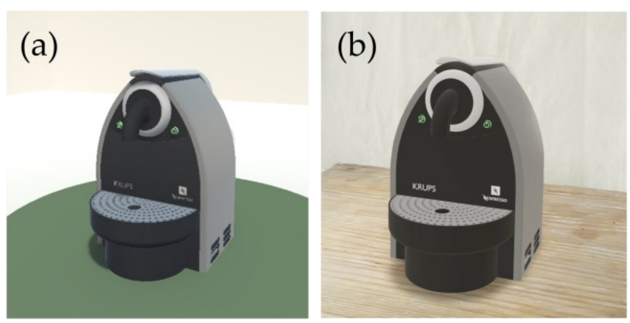

Fig. 23: Nespresso Essenza displayed in different media: N-IVE (a), and AR environment (b).

4.4.3. Semantic differential for product evaluation

The Semantic Differential is a common method of product evaluation (Osgood et al., 1957) that uses a 5- or 7-point Likert-type scale and typically includes various bipolar pairs of adjectives to describe the product that is being evaluated. For the first case study, we adopted the semantic space from the work by Hsu et al. (S. H. Hsu et al., 2000) since the product typology used to define the semantic space was also desktop telephones. We classified the 24 bipolar pairs collected by the authors according to the four categories defined by Jordan: physio, socio, psycho, and ideo (Jordan, 2002). For each of them, two pairs were selected, resulting in a total of eight bipolar pairs of adjectives.

For the second case study, we generated a semantic differential based on adjectives collected from three different sources: users, vendors, and manufacturers. As in the previous case, eight semantic scales were generated and classified according to Jordan's pleasure categories. The adjectives are listed in Table 34.

Table 34: List of the selected bipolar pairs of adjectives.

Physio	Psycho	Socio	Ideo
Desktop telephones			
Heavy/Handy	Decorative/Practical	Traditional/Modern	Childish/Mature
Large/Compact	Simple/Complex	Nostalgic/Futuristic	Handmade/High-tech
Coffee makers			
Minimalist/Overelaborated	Practical/Impractical	Traditional/Modern	Unsustainable/Sustainable
Large/Compact	Difficult/Easy to use	Unappealing/Appealing	Expensive/Inexpensive

4.4.4. Materials

The 3D models of the phones used as stimuli were created from scratch using Blender 2.93.0, a free and open-source 3D creation suite. Coffee maker models were downloaded from free online repositories and optimized with Blender 2.93.0. Photorealistic renders were generated using Blender's Cycles engine and displayed on a Fujitsu Lifebook E Serie laptop with a 15.6 inches screen size.

Virtual environments (VR and N-IVE) were designed using Unity 2019.4.14f1. For the VR setting, the Oculus Integration package (version 29.0) was used, which is freely available at the Unity asset store. The environment used baked lights and a standard shader was utilized for product materials. The VR environment was experienced on a Meta Quest 2 headset, a standalone immersive virtual reality device with a Single Fast-Switch LCD of 1832×1920 pixels per eye and a refresh rate of 72Hz. The N-IVE setting was displayed in a wide range of computer screens since the experiment was conducted online due to COVID-19 restrictions. The user was able to navigate the environment using their mouse and keyboard.

The AR environment was displayed in a Huawei P20 smartphone with a screen size of 5.8 in and a resolution of 1080x2240 pixels for desktop telephones. In the case of coffee makers, a wide range of smartphones were used since the experiment was conducted online due to COVID restrictions. To visualize the 3D model in the real environment, we used Clon Digital (https://clondigital.es), an online resource for integrating 3D products in a real space without the need to develop a specific application (Fig. 22). A texturized 3D model in a compatible format (.glTF, in our case) is required. The tool positions the 3D model on a flat surface at the pressing a button. For the AR viewer, ARCore (for Android devices) and

ARKit (for iOS devices) light estimation were used. Geometry cannot exceed 100,000 polygons, and textures are compressed to 512 pixels.

Fig. 24: Using Clon Digital with the Nespresso Essenza coffee maker (second case study).

4.4.5. Sample

We conducted an a priori power analysis with G*Power (Faul et al., 2007) assuming an ANOVA with repeated measures with the following input parameters: effect size: 0.25, α=0.05, (1-β)=0.80 and 1 group. Our results estimated a total sample size of 28. To guarantee a power of 0.80, a total of 36 volunteers participated in the first experiment (19 women and 17 men ages between 19 and 35 years old, with a mean age of 25.25 years old). 55.56% of the volunteers were from France, 25% from Spain, and the remaining 19.44% were from Belgium, England, Germany, New Caledonia, South Africa, Sweden, and Switzerland. A total of 39 participants from Spain participated in the second experiment (15 women and 24 men, with a mean age of 31.46 years old).

Prior to the experiment, users were asked to rate their previous experience with AR (for both cases) and VR (for the first case study) using a four-point Likert scale from 0 to 3 (0 = no experience with the technology, and 3 = significant experience). In the first case study, 52.8% had no previous experience with AR, and 50% had no experience with VR. 30.6% of the participants rated their experience with AR as limited and 44.4% did so with VR. 13.9% stated that they had significant experience with AR and 5.6% with VR. Only 2.8% of the participants claimed they had extensive experience with AR, but no participant claimed to have extensive experience in VR. In the second case study, 64.1% reported to have no previous experience with AR, 28.2% had some experience with AR, and 7,7% had significant experience with this technology.

Participants were recruited via web advertising in the university website. No target population was defined. Any individual was eligible to participate. Participants were required to have a computer and a smartphone for the second case study. People interested in participating signed up using an online questionnaire. The online form provided detailed information about the experiment. Participants were then contacted by a member of the research team to schedule an appointment. Participants who expressed interest in receiving the results after the experiment were contacted a second time for debriefing.

All participants provided verbal informed consent to participate in our studies. Our study was deemed exempt from IRB at our institution, since the information obtained is the result of straight-forward consumer acceptance testing which does not employ an intervention. Also, all information obtained was recorded in such a manner that the privacy of subjects is protected, and the confidentiality of data is maintained. For the online part of our study, participants provided verbal consent during a virtual meeting with one of the members of the research team prior to the study.

4.4.6. Methodology

Participants were not paid to participate in our study. Also, due to COVID-19 restrictions at the time of performing our second case study, some interviews were conducted online. Instructions were given to participants prior to starting the session. As part of the study, we also collected data on the user experience in AR and VR. Each participant went through each experimental condition of the assigned case study. To minimize the possibility of the order of presentation of the stimuli affecting the results, the presentation sequence for each participant was randomized. The physical evaluation room was the same for all experimental conditions in the first case study (desktop telephones), so the potential influence of the external environment on the evaluation was the same for each medium. Since some interviews were conducted online for the second case study, the physical room was the same between each of the experimental conditions but not between participants.

Participants were informed that interaction with the product was not allowed on any medium. In the VR and AR environments, participants were allowed to move around and perform any actions that did not involve direct interaction with the product. 2D images of the product were displayed on a computer screen. Participant could switch between images using their keyboards. Six views were provided for each product for desktop telephones (a front view, two side views, a zenithal view, and two isometric views), whereas four views were provided for each product in the case of coffee makers (a front view, a side view, a zenithal view, and an isometric view). Users pressed the left and right arrow keys to scroll through the images. In the AR environment, users were asked to hold the smartphone to examine the object until they considered they had enough information to complete the evaluation. Likewise, in the VR environment, users were allowed to adjust the headset beforehand to ensure that the image quality was acceptable. In both case studies, the evaluation was performed during the viewing of the product to avoid assessments based on recalled information. Each condition took approximately ten minutes to complete. Participants were asked to rate the product according to the eight semantic pairs using a 7-point semantic scale. In addition, they used a 5-point scale to rate how much they liked/disliked the product being displayed, as well as their intended purchasing decision.

4.5. Analysis and results

4.5.1. Case study I: desktop telephones

Descriptive statistics for each data set in our study are shown in Tables 33 and 34. The stacked bar charts for the semantic scales are shown in Fig. 25. The semantic scale data collection uses a 7-point Likert scale with a neutral value of 0 and two extreme values of 3 (-3 and 3). A higher value indicates better correspondence with the adjective represented on this end. For the "Like/Dislike" data set, a 5-point Likert scale was used with 1 as the lowest value and 5 as the highest.

Table 35: Descriptive statistics for the semantic scales (case I).

Semantic scales		Daewoo			Swissvoice			Philips		
		VR	AR	2D	VR	AR	2D	VR	AR	2D
Heavy/**Handy**	M	**2.22**	2.08	*1.31*	**1.08**	.31	*.14*	**.94**	-.06	*-.17*
	Md	2.00	2.00	2.00	2.00	1.00	-.50	2.00	.50	.00
	SD	1.17	.99	1.72	1.78	1.88	2.11	1.72	1.72	1.61
Large/**Compact**	M	**2.03**	1.94	*1.36*	**.28**	-.31	*-1.00*	**-.36**	-1.00	*-1.06*
	Md	2.00	2.00	2.00	.50	-1.00	-1.50	*-1.00*	-2.00	-2.00
	SD	1.21	.92	1.55	2.07	1.96	1.77	1.85	1.55	1.61
Decorative/**Practical**	M	1.33	*1.28*	**1.50**	-1.72	*-1.78*	**-1.56**	1.14	**1.33**	*.67*
	Md	2.00	2.00	2.00	-2.00	-2.00	-2.00	1.00	2.00	*1.00*
	SD	1.33	1.34	1.06	1.45	1.61	1.42	1.39	1.43	1.57
Simple/*Complex*	M	**-1.75**	-1.61	*-1.56*	**-1.53**	*-1.36*	-1.39	-1.36	*-1.00*	**-1.72**
	Md	-2.00	-2.00	-2.00	-2.00	-2.00	-2.00	-2.00	-2.00	-2.00
	SD	1.52	1.42	1.40	1.72	1.64	1.59	1.42	1.76	1.21
Traditional/**Modern**	M	.17	**.19**	*-.19*	**1.58**	1.22	*1.14*	-.39	*-.44*	**-.28**
	Md	.50	1.00	-.50	2.00	2.00	2.00	-.50	-1.00	.00
	SD	1.84	1.58	1.68	1.72	1.79	1.93	1.71	1.69	1.91
Nostalgic/**Futuristic**	M	**.19**	.16	*.11*	**.42**	.25	*.00*	*-.67*	-.53	**.19**
	Md	.00	.50	.00	.50	.00	.00	-1.00	.00	.00
	SD	1.34	1.34	1.47	1.99	1.90	2.11	1.39	1.38	1.53
Childish/**Mature**	M	**1.47**	*1.06*	**1.47**	*-.25*	-.47	**-.75**	**1.83**	*1.47*	1.56
	Md	2.00	1.00	2.00	.00	-1.00	-1.00	2.00	*2.00*	2.00
	SD	1.16	1.26	1.21	1.86	1.75	1.81	1.13	1.18	1.36

	M	**1.03**	.94	*.78*	**1.08**	.31	*.14*	*.69*	.72	**.83**
Handmade/**Hi-tech**	Md	1.00	1.00	1.00	1.00	.00	.50	1.00	1.00	1.00
	SD	1.42	1.29	1.31	1.78	1.88	2.11	1.23	1.37	1.50

Highest values and corresponding adjective are shown in bold, lowest values and corresponding adjective in italics.

Table 36: Descriptive statistics for the overall evaluation (case I).

		Daewoo			Swissvoice			Philips		
		VR	AR	2D	VR	AR	2D	VR	AR	2D
	M	2.72	2.44	*2.38*	4.06	*3.75*	3.89	*2.41*	2.44	2.69
Like/Dislike	Md	3.00	2.50	2.00	4.00	4.00	4.00	2.00	2.00	2.50
	SD	.91	.99	1.10	.86	1.08	.95	1.23	1.05	1.17

Highest values are shown in bold, lowest values in italics

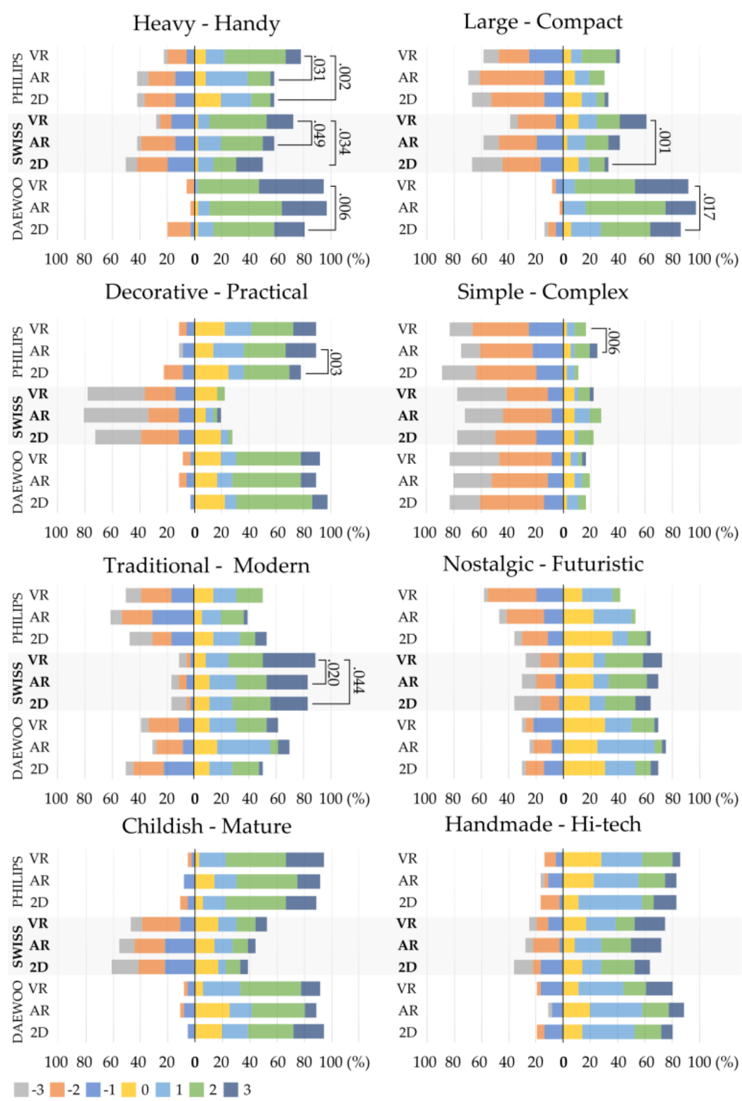

Fig. 25: Stacked bar charts for semantic scales (desktop telephones).

An inferential statistical method was applied to test the hypotheses described in Section 3 and a normality test was performed on each data set to select the appropriate statistical test. As the sample size was less than 50 participants, we used a Shapiro-Wilks's normality test (significance level of .05). Results showed that the data did not follow a normal distribution, so parametric tests proved unsuitable.

Since classic nonparametric statistical tests only allow for the analysis of a single factor, we applied the Aligned Rank Transform (ART) procedure (Higgins et al., 1990) in order to analyze multiple factors. ART is known to provide a powerful and robust nonparametric alternative to traditional techniques (Mansouri et al., 2004). It relies on a preprocessing step that "aligns" data before applying averaged ranks. After this step, common ANOVA procedures can be applied (Wobbrock et al., 2011).

In our study, we performed a series of Repeated Measures ANOVAs after the ART procedures as well as post-hoc tests (Bonferroni correction was applied) when perceptual differences were found between media to determine the exact groups involved.

Our Repeated Measures ANOVA (Tables 35 and 36) showed that Jordan's physio-pleasure category was the most influenced by the medium, as "Heavy – Handy" showed significant differences for each telephone, while "Large – Compact" showed differences phones. On the other hand, gender differences were found on "Simple – Complex" for the Daewoo DTD-1400w, and for "Large - Compact" for the Swissvoice Epure 2. The latter phone also showed a combined effect of the medium and the gender for "Handmade – Hi-tech".

Post-hoc tests for the semantic scales are shown in Table 39. It is important to note that, although the p-value of the bipolar pair "Childish – Mature" was .048 in the Daewoo DTD-1400W, post-hoc tests did not reveal any significant differences in the pairwise comparison.

Table 37: Two-factors repeated measures ANOVA for semantic scales (case I).

Semantic scales			Daewoo		Swissvoice		Philips	
		df	F	Sig.	F	Sig.	F	Sig.
Heavy – Handy	Media	2	6.590	**.002**	3.962	**.024**	7.855	**<.001**
	Gender	1	.182	.673	.299	.588	2.880	.099
	Mixed	2	.054	.948	1.352	.265	.711	.495
Large – Compact	Media	2	4.560	**.014**	7.601	**.001**	2.624	.080
	Gender	1	1.110	.299	7.520	**.010**	.791	.380
	Mixed	2	.691	.505	1.723	.186	1.054	.354
Decorative – Practical	Media	2	.197	.822	.945	.394	4.826	**.011**
	Gender	1	.262	.612	.239	.628	2.310	.138
	Mixed	2	.004	.996	.413	.663	2.631	.079
Simple – Complex	Media	2	.111	.895	.264	.768	6.200	**.003**
	Gender	1	5.600	**.024**	.459	.503	.000	.998
	Mixed	2	0.597	.942	.456	.636	2.858	.064

Semantic scales		Daewoo			Swissvoice		Philips	
		df	F	Sig.	F	Sig.	F	Sig.
Traditional – Modern	Media	2	.738	.482	5.232	**.008**	.142	.868
	Gender	1	.546	.465	.220	.642	.194	.662
	Mixed	2	.073	.929	2.561	.085	.306	.737
Nostalgic – Futuristic	Media	2	.151	.860	1.611	.207	1.429	.247
	Gender	1	.156	.695	.688	.413	.151	.700
	Mixed	2	.078	.925	.180	.835	.108	.898
Childish – Mature	Media	2	3.180	.048*	1.875	.161	1.910	.156
	Gender	1	.001	.982	1.220	.276	1.840	.184
	Mixed	2	.411	.664	.345	.710	.455	.636
Handmade – Hi-tech	Media	2	.978	.381	2.203	.118	.278	.758
	Gender	1	.096	.758	.315	.578	.004	.952
	Mixed	2	1.820	.170	3.378	**.040**	1.260	.291

*Factor value in which perceptual differences were found are shown in bold. * No significant differences were found in the pairwise comparisons.*

Table 38: Two-factors repeated measures ANOVA for overall evaluation (case I).

		Daewoo			Swissvoice		Philips	
		df	F	Sig.	F	Sig.	F	Sig.
Like/Dislike (1 – 5)	Media	2	3.21	.046	3.63	**.032**	2.06	.135
	Gender	1	.0491	.826	.181	.673	.0195	.890
	Mixed	2	.255	.776	1.72	.186	.663	.519

Factor value in which perceptual differences were found are shown in bold.

Table 39: Post-hoc test for semantic scales (case I).

Semantic scales	Condition	Daewoo	Swissvoice	Philips
		Sig.	Sig.	Sig.
Heavy – Handy	2D - AR	.150	.970	1.000
	2D – VR	**.006**	**.034**	**.002**
	VR - AR	.263	**.049**	**.031**
Large –Compact	2D - AR	.278	.060	
	2D – VR	**.017**	**.001**	
	VR - AR	.656	.302	
Decorative –Practical	2D - AR			**.003**
	2D – VR			.218
	VR - AR			1.000
Simple – Complex	2D - AR			**.006**
	2D – VR			.107
	VR - AR			.337
Traditional – Modern	2D - AR		1.000	
	2D – VR		**.044**	
	VR - AR		**.020**	
Childish – Mature	2D - AR	.233		
	2D – VR	1.000		
	VR - AR	.085		

Conditions and p-values in which perceptual differences were found are shown in bold.

The p-value for the overall evaluation was .046 (Table 7), but post-hoc tests did not reveal any significant differences in the pairwise comparison. Post-hoc tests showed that differences were found between 2D–VR (p=.028).

Finally, a repeated measures ANOVA with two within-subjects factors (product and medium) was performed to test whether the product's appearance and design could affect the perceptual variation of the semantic scales and thus explain our previous results. Our results (shown in Table 40) revealed a significant influence of the phone's design for the paired adjectives "Heavy - Handy", "Large - Compact", "Decorative - Practical", "Traditional - Modern" and "Childish - Mature". The adjectives "Decorative - Practical" and "Handmade - Hi-tech" showed a combined effect between the two factors. The adjectives in the physical pleasure category ("Heavy - Handy", "Large - Compact") as well as "Childish - Mature" were the most affected by the medium.

Table 40: Two-factors repeated measures ANOVA for semantic differential (case I).

Semantic scales	Product	Media	Mixed
	Sig.	Sig.	Sig.
Heav – /Handy	<.001	<.001	.118
Large – Compact	<.001	<.001	.233
Decorative – Practical	<.001	.192	.005
Simple – Complex	.238	.082	.064
Traditional – Modern	<.001	.521	.659
Nostalgic – Futuristic	.092	.396	.233
Childish – Mature	<.001	.045	.215
Handmade – Hi-tech	.501	.227	.031

Factor value and bipolar pairs of adjectives in which perceptual differences were found are shown in bold.

4.5.2. Case study II: coffee makers

The descriptive statistics for each data set for this case study are shown in Tables 39 and 40. The same criteria were used for data collection as in the previous case. The stacked bar charts are shown in Fig. 26.

To test our hypotheses, we applied an inferential statistical method. In this case, a Shapiro-Wilk's normality test showed that the data did not follow a normal distribution, so the ART procedure was applied.

Table 41: Descriptive statistics for the semantic scales (case II).

Semantic scales		Moka			Inissia			Essenza		
		2D	3D	AR	2D	3D	AR	2D	3D	AR
Minimalist – *Overelaborated*	M	*.72*	1.64	1.72	*0.46*	1.00	1.13	*-.15*	.59	.77
	Mdn	.00	2.00	2.00	.00	1.00	1.00	.00	.00	.00
	SD	.32	1.27	1.08	1.34	1.19	1.36	.23	1.37	2.49

Semantic scales		Moka			Inissia			Essenza		
		2D	3D	AR	2D	3D	AR	2D	3D	AR
Large – *Compact*	M	.15	-.51	*.33*	-.33	-.56	*.51*	*-.13*	-.72	-.38
	Mdn	.00	-1.00	.00	-1.00	.00	0.00	.00	-1.00	.00
	SD	.71	1.52	1.13	1.38	1.21	1.91	1.59	1.28	1.18
Practical – *Impractical*	M	1.72	*1.54*	1.72	2.13	*2.05*	2.36	2.03	*1.69*	2.03
	Mdn	2.00	2.00	3.00	2.00	2.00	3.00	2,00	2.00	3.00
	SD	1.56	1.70	1.81	1.08	1.32	.87	1.20	1.78	1.44
Difficult – *Ease to use*	M	1.79	1.87	*1.97*	2.23	2.13	*2.33*	2.03	*2.26*	1.71
	Mdn	2.00	2.00	3.00	2.00	3.00	3.00	2.00	3.00	3.00
	SD	1.79	1.61	1.46	.96	1.20	.87	1.14	1.33	1.37
Traditional – Modern	M	-2.31	*-2.56*	-1.05	*.74*	.95	1.49	*.00*	.05	1.05
	Mdn	-2.00	-3.00	-1.00	1.00	1.00	2.00	.00	.00	1.00
	SD	0.77	1.10	1.92	.91	.86	1.02	1.43	1.49	1.30
Unappealing – Appealing	M	.28	*-.28*	1.79	1.49	*1.38*	2.13	*.15*	1.08	1.49
	Mdn	.00	.00	2.00	1.00	1.00	2.00	0.00	1.00	2.00
	SD	1.19	1.61	1.17	.91	1.09	.98	1.71	1.42	1.49
Unsusainable – Sustainable	M	1.72	1.72	*1.46*	*-2.13*	-2.08	-2.03	-.19	*-1.92*	-1.82
	Mdn	2.00	2.00	1.00	-3.00	-2.00	-2.00	-.00	-2.00	-2.00
	SD	1.40	1.28	1.14	1.17	1.11	1.20	1.35	1.18	1.34
Expensive – Inexpensive	M	2.36	2.28	*1.05*	-.18	*-.37*	.05	*-.08*	.15	.00
	Mdn	3.00	3.00	1.00	.00	.00	.00	.00	.00	.00
	SD	.84	1.00	1.70	1.30	.22	1.34	.38	1.33	1.10

Highest values and corresponding adjective are shown in bold, lowest values and corresponding adjective in italics.

Table 42: Descriptive statistics for the overall evaluation (case II).

		Moka			Inissia			Essenza		
		2D	3D	AR	2D	3D	AR	2D	3D	AR
Like/*Dislike*	M	3.23	*2.82*	4.51	*3.72*	3.79	4.46	*3.10*	3.49	4.21
	Md	3.00	3.00	5.00	4.00	4.00	5.00	3.00	4.00	4.00
	SD	.93	1.28	.56	.86	.89	.68	.85	.82	.83

Highest values are shown in bold, lowest values in italics.

Once again, we performed a series of Two-factor Repeated Measures ANOVAs after the ART procedures for the semantic differential (Tables 41) and the Overall evaluation (Table 44). Post-hoc tests (with Bonferroni correction) were also performed when perceptual differences were found between media to determine the exact groups involved (Table 45).

Table 43: Two-factors repeated measures ANOVA for semantic scales (case II).

Semantic scales			Moka		Inissia		Essenza	
		df	F	Sig.	F	Sig.	F	Sig.
Minimalist – Overelaborated	Media	2	.556	.576	7.029	**.002**	8.661	**<.001**
	Gender	1	9.349	**.004**	1.440	.238	5.050	**.031**
	Mixed	2	.611	.545	.467	.629	1.654	.198
Large – Compact	Media	2	3.189	**.047**	8.623	**<.001**	1.992	.144
	Gender	1	.028	.869	7.548	**.009**	.855	.361
	Mixed	2	.552	.578	2.408	.097	.941	.395
Practical – Impractical	Media	2	1.233	.297	1.407	.251	5.985	**.004**
	Gender	1	16.841	**<.001**	.236	.630	.502	.483
	Mixed	2	3.176	**.048**	.277	.759	8.421	**.001**

Semantic scales			Moka		Inissia		Essenza	
		df	F	Sig.	F	Sig.	F	Sig.
Difficult – Ease to use	Media	2	14.079	**<.001**	.469	.627	8.115	**.001**
	Gender	1	12.607	**<.001**	2.506	.122	.521	.475
	Mixed	2	4.170	**.019**	1.004	.371	.713	.494
Traditional – Modern	Media	2	14.838	**<.001**	8.149	**.001**	10.245	**<.001**
	Gender	1	5.722	**.022**	7.773	**.008**	3.537	.068
	Mixed	2	1.336	.269	1.396	.254	1.269	.287
Unappeal – Appealing	Media	2	35.657	**<.001**	10.702	**<.001**	13.070	**<.001**
	Gender	1	1.490	.230	.245	.623	.475	.495
	Mixed	2	1.584	.212	2.866	.063	5.655	**.005**
Unsustainable – Sustainable	Media	2	1.398	.255	.681	.509	1.661	.197
	Gender	1	2.559	.118	1.060	.310	.696	.409
	Mixed	2	1.296	.280	.452	.638	1.877	.160
Expensive – Inexpensive	Media	2	13.066	**<.001**	.496	.611	1.214	.303
	Gender	1	2.219	.145	.034	.855	.664	.420
	Mixed	2	1.481	.234	.650	.525	2.116	.128

Factor value in which perceptual differences were found are shown in bold.

Table 44: Two-factors repeated measures ANOVA for overall evaluation (case II)

			Moka		Inissia		Essenza	
		df	F	Sig.	F	Sig.	F	Sig.
Like/ Dislike	Media	2	45.231	**<.001**	15.263	**<.001**	27.558	**<.001**
	Gender	1	.481	.492	2.437	.127	1.098	.301
	Mixed	2	.800	.453	.680	.510	5.498	**.006**

Factor value in which perceptual differences were found are shown in bold.

Table 45: Post-hoc test for semantic scales (case II)

Semantic scales	Condition	Moka	Inissia	Essenza
		Sig.	Sig.	Sig.
Minimalist – Overelaborated	2D – 3D		-158	**.004**
	2D – AR		**.001**	**.036**
	3D - AR		.297	.278
Large – Compact	2D – 3D	.252	.725	
	2D – AR	1.000	**.037**	
	3D - AR	**.022**	**<.001**	
Difficult – Ease to use	2D – 3D	**.001**		**.004**
	2D – AR	**.001**		**.028**
	3D - AR	**.010**		.632
Traditional – Modern	2D – 3D	**.011**	.868	1.000
	2D – AR	**.043**	**.001**	**<.001**
	3D - AR	**<.001**	**.027**	**.003**
Unappealing – Appealing	2D – 3D	.853	1.000	**.002**
	2D – AR	**<.001**	**.004**	**<.001**
	3D - AR	**<.001**	**<.001**	.323

Expensive – Inexpensive	2D – 3D	1.000
	2D – AR	**<.001**
	3D - AR	**.001**

Conditions and p-values in which perceptual differences were found are shown in bold.

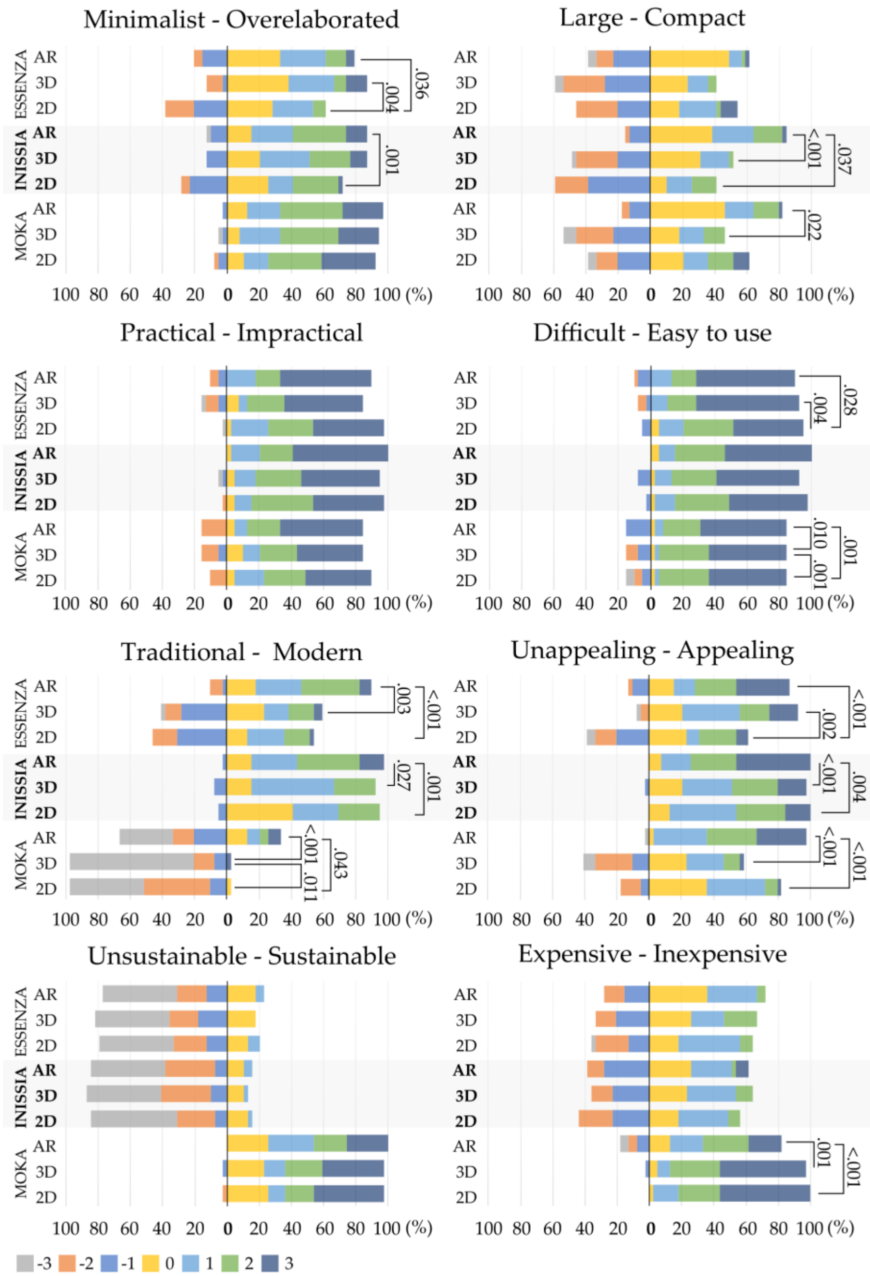

Fig. 26: Stacked bar charts for the semantic scales (coffee makers).

Our results show that the Overall Evaluation was influenced by the medium for each coffee maker. Pairwise comparisons showed that these differences were statistically significant between 2D – 3D and 2D – AR (p<.001) for each case. The Repeated Measures ANOVA for the semantic differential showed that, although some scales were influenced by the medium in certain products, the adjectives related to Jordan's sociological pleasure category were the most affected by the visual media. Post-hoc tests showed that these differences were mostly found between 2D – AR and 3D – AR.

Finally, a Repeated Measures ANOVA with two within-subjects factors (product and medium) was performed to test whether the product's appearance and design could affect the perceptual variation of the semantic scales. Results are shown in Table 46.

Table 46: Two-factors repeated measures ANOVA for semantic scales (case II)

Semantic scales	Product	Media	Mixed
	Sig.	Sig.	Sig.
Minimalist – Overelaborated	**<.001**	**<.001**	**.004**
Large – Compact	.076	**.003**	.212
Practical – Impractical	.455	**.031**	.456
Difficult to use – Ease to use	.192	**<.001**	**<.001**
Traditional – Modern	**<.001**	**<.001**	**.009**
Unappealing – Appealing	**<.001**	**<.001**	**<.001**
Unsustainable – Sustainable	**<.001**	**.031**	**.001**
Expensive – Inexpensive	**<.001**	**<.001**	**<.001**

Factor value and bipolar pairs of adjectives in which perceptual differences were found are shown in bold.

4.6. Discussion

In this paper, we presented the results of two case studies where a group of participants evaluated three different designs of a product typology using the Semantic Differential method in different visual media.

In our main hypothesis H1, we questioned whether the medium used to present a product influences how the user evaluates the semantic scales regardless of their classification in Jordan's categories. The classification of the bipolar pairs according to the pleasure categories defined by Jordan (Jordan, 2002) helped us determine which type of adjectives are most affected by the display medium.

Results of our two-factor repeated measures ANOVA for the semantic scales for desktop telephones (Table 37) showed that visual media can influence the user's subjective impression of a product, which agree with similar studies (M. J. Agost et al., 2021; Artacho-Ramírez et al., 2008). Although several bipolar pairs of adjectives were influenced by the medium, not all categories were affected in the same manner, as confirmed by Galán et al. (Galán, García-García, et al., 2021). These results align with those from similar studies such as (Galán, Felip, et al., 2021) in which the differences are justifed by the absence of touch.

Although the absence of touch may be highly relevant to the physical category (Grohmann et al., 2007) (which may also explain our results), our evaluation relied entirely on the sense of sight, so visual differences between media may have caused these variances.

In the case of "Large - Compact", the stimulus was displayed in various sizes, which could be interpreted as a limitation when presenting a product. For example, the descriptive statistics for this data set (Table 35) show that the product was perceived as larger when displayed through 2D images. Reasons could be attributed to the size of the computer screen, which may have made the object appear oversized, or to color saturation (some authors have suggested that the higher the color saturation, the greater the perception of size (Hagtvedt & Adam Brasel, 2017)). Brightness may have also affected size perception, as brighter objects often appear larger and closer (G. Singh et al., 2020) to the user. Although an attempt was made to maintain consistent levels of saturation and brightness for each medium, slight differences may have influenced the results (e.g., slightly higher color saturation or brightness may be present in the 2D images). Additionally, scale/size/distance judgements in VR are difficult (Kelly, 2022) and the fact that the product was viewed with no references to other objects may also have hindered the assessment in 2D images (in both AR and VR, a virtual table was present which could have served as a reference) (H. Park et al., 2021). The pair "Heavy - Handy" could be directly related to the perception of size (the larger the heavier). It is important to note that differences were found mainly between 2D and VR.

The fact that no differences were found in the bipolar pair "Large - Compact" for the Philips M110w phone could be explained by the possible influence of certain aspects of the product (e.g., geometric elements) on the adjectives, as reflected in Table 37. Indeed, shape can influence the user’s subjective impressions (Achiche et al., 2014), but this is further discussed in H2. For this phone, perceptual differences were also found for "Decorative - Practical" (in this category, the effect of the product and the combined effect between medium and product were also significant) and “Simple–Complex”, which are part of Jordan’s psycho-pleasure category. Our results suggest that the psycho-pleasure category can be affected by the presentation medium, not for a specific product typology (desktop telephones) but for a specific product design within a typology. To draw more generalizable conclusions, additional studies with different types of products are needed.

Finally, it is important to note that in the overall evaluation for this case study (Table 38), significant differences were found only for one of the phones (Swissvoice), which means that the results cannot be generalized to any medium. Additionally, the aesthetics of this particular design stand out from the others, which may have influenced this result. Furthermore, the highest scores for this dataset (Table 36) were generally found in the VR medium, suggesting that presenting a product in an immersive virtual medium may favor the overall evaluation of the product.

Similar results were obtained for the case of coffee makers. Several bipolar pairs of adjectives were affected, mostly within Jordan's physiological and sociological pleasure categories. Some authors have suggested that the physical pleasure category may not be the only one affected by the display method. Galán et al. (Galán, García-García, et al., 2021), for example, showed that the ideological pleasure category could also be highly influenced by the medium.

Regarding physical pleasure, the results are analogous to the case of telephones, i.e., the presence of touch could have naturally affected this category (Grohmann et al., 2007) but the visual differences between the media are likely at the source of these variances. For example, for the bipolar pair "Minimalist – Overelaborated," we speculate that the photographs could have shown the product in greater detail, which could have distorted the user's subjective impression of the product compared to 3D or AR media. Alternatively, the 3D medium could have made users perceive the product as larger compared to AR, where objects generally appear smaller and are conditioned by the size of the smartphone screen. Similarly, saturation, brightness, and product context may have also affected the results (Hagtvedt & Adam Brasel, 2017; H. Park et al., 2021; G. Singh et al., 2020). The socio-pleasure category (adjectives linked to the aesthetics of the product) may have been affected by the combination of the product appearance and the medium (Table 46). Finally, bipolar pairs where significant differences were found for the medium for only one of the designs could have been affected by the product geometry (Achiche et al., 2014).

The overall evaluation was also affected by the change of medium for all three products. In addition, the descriptive statistics (Table 42) for this dataset showed much higher values for the AR medium, suggesting that the presentation of a product in a physical context or using visualization techniques with higher levels of interaction may favor product evaluation.

In general, although the overall evaluation cannot be generalized to all the products, results related to product features (those that comprise the semantic differential) agree with previous studies in that adjectives that require sensory interaction such as touch are more sensitive to the change of display medium (Galán, Felip, et al., 2021). Therefore, the visualization medium may have an impact on the user's subjective impressions of the product, particularly those in the physical pleasure category, which derives, to a great extent, from senses such as touch. Because of this, the absence of physical interaction with the product may have influenced the evaluation, which confirms H1.

In H2, we postulated that a particular design within the same product typology may influence the user's subjective impressions (H2). Many authors have examined the relationship between a product's shape and the emotions elicited by it. Aesthetics is one of the first channels through which designers communicate with consumers (C. C. Hsu et al., 2017; X. Liu & Yang, 2022).

The results of our study (Table 40 and 44) are consistent with those obtained by other authors which suggest that the geometric features of the product (i.e. product aesthetics) may influence the user's subjective impressions (Achiche et al., 2014; Perez Mata et al., 2017). Additionally, for the case of desktop telephones (Table 40), the bipolar pair "Decorative – Practical" (Jordan's psychological-pleasure category) and "Handmade – Hi-tech" (ideo-pleasure category) revealed a combined effect between factors. For the case of coffee makers (Table 46), "Large – Compact" (physio-pleasure category) and "Practical – Impractical" (psychological-pleasure category) were the only pairs that did not show a combined effect. It is important to note that although the medium is not the only factor that may influence the user's subjective impressions, a combination of factors may cause perceptual differences. Based on the above discussion, H2 is confirmed.

In our third hypothesis (H3), we questioned the existence of gender differences in the user's subjective impression of the product. In the case of desktop telephones, although the overall evaluation did not show an influence of gender in any case (Table 38), the Essenza coffee maker did show a combined effect between gender and medium for the overall evaluation (Table 44).

Regarding the semantic scales (Table 37 and 41), different adjectives revealed an influence of gender in some of the products, but no pattern was observed to draw general conclusions. In other words, there may be an influence of gender on the evaluation of some characteristics, but only for certain designs within the same product category. For example, the Daewoo DTD-1400w, gender differences were found in "Simple-Complex", while for the Swissvoice Epure 2, they were found in "Traditional – Modern." A combined effect of Media*Gender in "Handmade–Hi-tech" was also found for this phone. The mean scores obtained for "Simple – Complex" by gender are M_{IMG}=-1.12, M_{AR}=-1.35, M_{VR}=-1,12 for males, and M_{IMG}=-1.95, M_{AR}=-1.84, M_{VR}=-2,32 for females. On the other hand, the mean scores obtained for "Traditional – Modern" are M_{IMG}=1.47, M_{AR}=1,12, M_{VR}= 2.00 for males, and M_{IMG}=.84, M_{AR}=1.32, M_{VR}=1.21 for females. In general, females scored the adjectives "Simple" and "Traditional" higher than males in all media. In this context, some authors have suggested that women generally favor the evaluation of physical products more than men (Cho, 2004; Dittmar et al., 2004), so preferring a traditional shopping method may have had an effect on some evaluations. Our results suggest that gender differences may exist in product evaluation, not for a specific type of product but for some characteristics within a specific product design, so H3 is rejected. Further research is needed for testing this hypothesy.

Our study shows that the visual medium used to present a product may significantly affect how the product is perceived. However, other factors such as geometry can also influence the user's subjective impressions of a product. Therefore, not all products will yield the exact same response when the presentation medium is changed. For certain product features (e.g., those in Jordan's ideo pleasure category) and specific evaluation purposes, technologies such as VR or AR can be effective tools, but it is important to recognize how a particular medium relates to a specific product typology.

4.7. Conclusion

Understanding how a product is perceived and how users evaluate it are important aspects to ensure a design is presented and communicated effectively.

This study demonstrates that the medium used to view a product can influence how it is perceived, as certain characteristics (such as weight and size) are particularly significant, as the perceptual differences elicited by the different media are more pronounced. By furher dividing the bipolar pairs that make up the semantic differential used for the evaluation of the product typologies used, we observed that not all of Jordan's pleasure categories are affected equally by the presentation medium.

Table 47: Summary of results for the two case studies.

<table>
<tr><th>Hypotheses</th><th>Result</th><th>Desktop telephones (case study I)</th><th>Coffee makers (case study II)</th></tr>
<tr><td>H1</td><td>Confirmed</td><td>1 – Jordan physio-pleasure category was the most affected.
2 – Psycho-pleasure category can be affected by the presentation medium, not for a specific product typology but for a specific product design within the typology.
3 – Although the overall evaluation may not be influenced by visual medium, the VR setting can positively influence user's subjective impressions.</td><td>1 – Although the physiological pleasure category was influenced by visual medium, the physiological and sociological pleasure categories were the most affected by the medium.
2 – The overall evaluation was also affected by the change of medium, whereas the AR setting can positively influence user's subjective impressions.</td></tr>
<tr><td rowspan="2">H2</td><td rowspan="2">Confirmed</td><td>"Decorative – Practical" (psychological-pleasure category) and "Handmade – Hi-tech" (ideo-pleasure category) presented a combined effect.</td><td>"Large – Compact" (physio-pleasure category) and "Practical – Impractical" (psychological-pleasure category) did not present a combined effect.</td></tr>
<tr><td colspan="2">1 – The medium is not the only factor that may influence the user's subjective impressions.
2 – A combination between factors may cause perceptual differences.</td></tr>
<tr><td>H3</td><td>Rejected</td><td colspan="2">1 – Gender differences may exist in product evaluation, not for a specific type of product but for certain characteristics within a specific product design.
2 – Further research is needed for more generalizable conclusions.</td></tr>
</table>

Other aspects may be affected by design factors, and not just by the change of medium. The presentation media can also be a powerful mechanism for highlighting certain attributes of the product, especially those that require physical interactions with the product.

Our findings are useful from a product development standpoint and identify important communication aspects that should be considerd for presentation at the point of sale. Product features in Jordan's physical pleasure category are the most difficult to evaluate with virtual prototypes. In these cases, physical prototypes can help minimize these differences (X. Liu & Yang, 2022).

VR and AR technologies can facilitate both product development processes and product presentation at physical points of sale where there is a high level of control of the evaluation context. Multiple design alternatives can be evaluated virtually without the need for physical prototypes, which can save time and costs. In online sales channels, the use of physical prototypes is not possible, but AR and VR technologies can help enhance the user's perception of the product as well as provide richer information, especially when compared to simple 2D images. Because the use context is important during product evaluation, VR can also increase the level of control over the evaluation process.

Although our study can be extrapolated to similar products of the same typology (i.e., telephones and coffee makers), additional tests with other types of products are necessary to draw more generalized conclusions. In future studies, we plan to use physiological measures such as eye-tracking technologies to analyze user behavior more accurately and objectively during product evaluation.

Chapter 5

5. General Discussion

This doctoral thesis presents the outcomes of various experimental works carried out to investigate the impact of the visualization technique on users' subjective impressions of a product. Additionally, the emotional response was analyzed concerning the evaluation method (i.e., individual or joint) and the interaction capabilities with the product, which were considered as significant factors affecting user's perception. Finally, the user's design expertise and gender along with the design geometry were also considered for the analysis of results. A summary of our results is shown in Table 48.

Table 48: Summary of results.

Hypotheses	Result	Discussion
H0.1	Accepted	• The visual medium used to present a product can influence how it is perceived. • Not all pleasure categories in Jordan's classification were influenced equally by the visual medium. • Bipolar pairs of adjectives belonging to Jordan's physio-pleasure category are the most affected by the change of medium. • The introduction of haptic feedback during the product evaluation process may help minimize the differences that arise because of the visual medium.
H0.2	Rejected	• The purchasing decision seemed not to be affected by the visual medium, but further research is needed to study the influence of the visualization technique in this dataset.
H0.3	Accepted	• The confidence in the response is influenced by the visualization technique. • The use of haptics during the evaluation process can increase the user's confidence in the response.
H0.4	Accepted	• The overall evaluation of the product is influenced by the visualization technique. • The use of more sophisticated visual media can help to improve the perception of the product.
H1.1	Accepted	• A joint evaluation can help minimize perceptual differences elicited by the visualization technique.
H2.1	Accepted	• Design expertise may influence product evaluation but not necessarily implicate an increase of confidence in the response.
H2.2	Accepted	• The use a complementary item can offer more information about the functioning of a product
H3.1	Accepted	• A particular design within the same product typology could influence product assessment. • This is something that seems to occur in all categories, except for physical pleasure, whose results seem to be more generalizable.
H3.2	Rejected	• Gender differences appeared for some semantic scales. • No pattern was observed to draw general conclusions. • Further research needs to be carried out to obtain more robust results.

The main objective of this analysis is to provide a discussion that interconnects and relates the results that have been obtained through the different experiments that were carried out during the research. By linking the results together, we aim to extract a more nuanced and robust interpretation of the research, which can ultimately provide valuable insights into the topic under scrutiny. First, an analysis of the postulated hypotheses will be conducted, followed by an assessment in the subsequent discussion to determine whether the primary and secondary objectives have been met.

In the first hypotheses, we proposed that the visual medium used to present a product could influence the user's subjective impressions of it (H0.1), which has been validated through the experimental studies conducted. In this regard, H0.1 is accepted, and these findings are consistent with those reported by other authors (M. J. Agost, 2020; Artacho-Ramírez et al., 2008; Galán, Felip, et al., 2021).

The classification of the specific bipolar pairs used on each study according to Jordan's pleasure categories (Jordan, 2002) helped us determine which type of adjectives were the most affected by the display medium. Although several adjectives can be influenced by the medium, not all categories are affected in the same manner, as confirmed also by Galán et al. (Galán, García-García, et al., 2021). This phenomenon can be attributed to the inherent characteristics of the product, which implies that certain features may be more susceptible to the influence of the shift in medium depending on the product type, and not all products will behave in the same way in response to the change in medium.

In a general sense, the characteristics associated with Jordan's physical pleasure category (related to the five senses) have been found to be the most difficult to evaluate. Based on an analysis of various case studies, it becomes evident that this category is the one most significantly impacted by the change of medium. However, it is noteworthy that the introduction of tactile feedback has proven instrumental in mitigating these disparities.

In this regard, the use of more sophisticated display media, i.e., VRPH, made it possible to observe changes in perception of certain product characteristics that may be attributed to the presence of tactile feedback during the product assessment (Galán, Felip, et al., 2021). The incorporation of tactile feedback during the evaluation of a virtual prototype may aid in minimizing the differences in most cases (Grohmann et al., 2007). This was confirmed by post-hoc testing conducted in the statistical analyses. Consequently, the use of VRPH may prove to be an effective tool in the design and development of products, as it permits user engagement throughout the design process and reduces costs during the prototyping phase in general.

Should we proceed with further discourse regarding the differences that emerge because of alterations in the medium, in H0.4 we hypothesized that the overall evaluation of the product could be affected by the visualization technique, which was accepted due to the results obtained. Within this context, it is worth noting that at least one product from the selected set for each case study demonstrated dissimilarities within this dataset due to the change in medium. However, there was an exception with regards to the second case study outlined in the article entitled "On the application of extended reality technologies for the evaluation of product characteristics during the initial stages of the product development process," wherein none of the products showcased differences in the comprehensive evaluation. This outcome can be attributed to the specific evaluation method employed, namely a joint evaluation, which will be discussed in greater depth later. Our findings suggest that the overall assessment of a product may not necessarily be impacted by the medium, but variations may arise for specific designs within a particular product typology. Furthermore, upon analyzing the outcomes of the descriptive statistics, we highlight that more conventional visualization media (usually 2D images) produced the lowest values. Thus, the incorporation of more sophisticated media, such as interactive 3D visualizations or immersive virtual reality, can make the product be perceived as more attractive. This can be explained by the fact that 3D representations are considered richer, more fun to explore, and more preferable overall (Ant Ozok & Komlodi, 2009).

In hypothesis H0.2, we postulated that the purchasing decision could be affected by the change of the product's display medium. In both cases, it was determined that the purchase decision was not a factor affected by the medium, but this result requires further investigation for this to be confidently asserted. In this case, certain limitations may cause the purchase decision not to vary depending on the medium: (1) having a simple scene in which the participant only has to visualize a product without any context may not generate the purchase intention of the product; (2) the stimulus used may not be consistent with the preferences of the participant sample of the study, causing the purchase decision to also not appear and therefore not vary between media. Although our findings can agree with other studies that suggest that although users need as much information as possible to make a purchasing decision (O'Keefe & McEachern, 1998), even a simple presentation medium provides sufficient details to perform a reasonable evaluation (Ant Ozok & Komlodi, 2009), it is necessary to design a more sophisticated shopping experience in order to verify this hypothesis.

In this context, although the hypothesis was formulated and rejected in the initial published studies, it can be concluded that a more sophisticated study is needed to verify this hypothesis (which is why it was not formulated in the last one). In this regard, works such as (Martínez-Navarro et al., 2019) or (Khatri et al., 2022) present clear examples of a real shopping experience, a methodology that could be interesting to use in order to observe the differences in the purchase intention towards a specific product.

As previously mentioned, a more comprehensive study is required to establish a definitive conclusion. Nonetheless, based on the descriptive statistics from both studies, it has been demonstrated that scores were higher in media featuring a 3D virtual prototype, such as VRPH or AR. Therefore, presenting a 3D model of a product may potentially enhance the intention to purchase by providing a greater level of information about the product than that offered by a two-dimensional medium, which typically presents static images.

On the other hand, the confidence in the response was also considered to be affected by the visual media (H0.3). This hypothesis was accepted given the results obtained from the conducted research. Thus, the statistical analysis showed that some designs exhibited differences between means for this dataset. In the article "The influence of hand tracking and haptic feedback for virtual prototype evaluation in the product design process," a more specific analysis was carried out for each of the semantic scales, with it being observed that the category of Jordan's physical pleasure was also the most affected by the medium for Confidence in Response. In this context, it is important to highlight that the VRPH medium had the highest score (after R), so the introduction of touch may increase the user's Confidence in Response, something that is in line with results obtained by other authors (Grohmann et al., 2007).

As mentioned in previous sections, a series of specific hypotheses were also formulated for each of the studies presented, which are discussed below. First, for the paper "On the application of extended reality technologies for the evaluation of product characteristics during the initial stages of the product development process" it was postulated that a joint evaluation could help minimize the perceptual differences elicited by the change of visual medium (H1.1). In this instance, a study was devised whereby four chairs were subjected to evaluation using a ranking system employing various semantic scales. The outcome of the statistical analysis (as detailed in Table 8) revealed that only two of the chairs were impacted by a change in medium, specifically for adjectives that predominantly pertained to Jordan's physical pleasure category. It has been previously demonstrated that this category is the most vulnerable to medium change. This finding contrasts with the results of the second case study carried out in this article, where all chairs were affected by the visualization medium. Additionally, each of the product typologies that were assessed during this doctoral thesis also underwent individual evaluation, and in every instance, it was observed that at least one of the product's characteristics had been affected. Consequently, it is deemed that the hypothesis has been confirmed.

In this context, other authors have demonstrated how utilizing a ranking evaluation can assist in mitigating differences that may arise due to varying levels of design expertise among users (Hu et al., 2022). This discovery offers insight: conducting a ranking evaluation has the potential to minimize perceptual differences and may be extrapolated to the evaluation of products through different visual media. Accordingly, further inquiry is necessary, such as examining products individually and conducting a ranking evaluation to determine if this phenomenon remains evident.

On the other hand, the design expertise was also considered as a factor that could have an effect into product evaluation (H2.1) for the research titled “The influence of hand tracking and haptic feedback for virtual prototype evaluation in the product design process”. Results showed that product features were influenced by the participant’s design background one umbrella stand design, which means that this factor is not associated with a specific design typology or adjectives but can affect the perception of certain product characteristics. Some studies have reported that in product evaluation individuals with a design-related background tend to employ more higher-level cognitive abilities than untrained individuals (Solso, 2001), so the influence of design expertise on the psycho-pleasure category was expected. The descriptive statistics for the confidence in the response showed that having a design background does not increase the participant’s confidence in the response toward product evaluation, which aligns with the study of (Forbes et al., 2018a) and confirms H4. Making an assessment about specific product features is a subjective task, so having a background in design does not necessarily increase this level confidence.

Additionally, we postulated for this research that the use of a complementary item could influence product perception (H2.2), which was also accepted due to the results obtained. In general, individuals require comprehensive information when evaluating a product. However, in certain instances, particularly in the context of online assessments, users may face limitations in their ability to fully explore a product's functionality, owing to the absence of a complementary element - such as an umbrella in this case. Statistical analysis revealed significant differences in some product characteristics due to the introduction of this missing element. Specifically, these differences were also evident in Jordan's psychological pleasure category, which pertains to the product's functional performance.

Finally, in the article "An Examination of the Relationship between Visualization Media and Consumer Product Evaluation," we suggested that a particular design within the same product typology could influence the user's subjective impressions of the product (H3.1) and that gender differences could exist in the evaluation of a product and how it is perceived (H3.2). Thanks to the selection of more than one product within the same typology we were able to test and accept H3.1. It was possible to observe that some of the results showed that certain bipolar pairs of adjectives were individually affected for one of the typologies, which was attributed to the product's own design. In this context, our results suggest that some bipolar pairs of adjectives can be affected by the presentation medium, not for a specific product typology but for a specific product design within a typology. This is something that seems to occur in all categories, except for physical pleasure, whose results seem to be more generalizable. This shows that the perception that a person is going to have about certain characteristics can be unpredictable, and for that reason, we strongly advise that this type of study be carried out prior to designing or launching a product. Finally, although gender differences appeared for some semantic scales, no pattern was observed to draw general conclusions. Further research needs to be carried out to obtain more robust results, so H3.2 is rejected.

Now that the hypotheses have been analyzed, a brief discussion about the achievement of the previously postulated objectives (primary and secondary) is conducted.

As the primary objective, this research aimed to understand how VR and MR technologies could enhance product development and marketing by analyzing their impact on our emotional and perceptual responses to virtual prototypes based on Jordan's emotional approach. In this regard, the hypotheses analysis has proven that the medium used to display a product has an impact on our subjective impressions of the product and in our confidence in our responses (H0.1, H0.3 and H0.4), something that must be consider (1) during the product evaluation process to obtain reliable feedback of the product, and (2) at the point of sale, since increasing response confidence can have a positive impact on purchasing decisions (Zheng & Bensebaa, 2022). However, while some perceptual differences may persist—it is believed that many more analyses should be conducted on this topic regarding different types of products—, the utilization of more advanced visualization mediums—such as immersive VR—holds the potential to refine our product perceptions, aligning them more closely with the actual product.

VR techniques confer significant advantages within the design process (Aurora Berni & Borgianni, 2020). The low cost of the HMD allows the design process to be more cost-effective thanks to virtual prototyping (Cecil & Kanchanapiboon, 2007), while also minimizing the time investment in the evaluation process where potential errors need to be identified and corrected. High-fidelity virtual prototypes and immersive environments can be created quickly; thus, we can have controlled environments for product evaluation, which are also highly adaptable to changes without the need to increase the product cost. (Hannah et al., 2012)

On the other hand, the introduction of passive haptics (also a low-cost solution to introduce tactile sensations during the experience) proves to be an effective approach to minimize differences between visualization methods and make our subjective impressions more similar to the real product. The easy access to rapid prototyping—such as 3D printers which are also very affordable devices—, we can manufacture low-cost prototypes that can provide haptic sensation during the evaluation process. This alignment is essential for obtaining reliable feedback to make more informed decisions. Moreover, incorporating the sense of touch during the virtual experience also positively influences response confidence, empowering users with greater trust in their opinions and yielding more dependable feedback. Although this is still difficult to achieve during online shopping, this allows to minimize the physical product catalog on display at the physical point of sale, also saving on costs.

When aligning these findings with our secondary goals, we can observe that progressing towards the more immersive end of the Milgram and Kishino continuum (i.e., VR) is associated with improved evaluations, more closely resembling those evaluations conducted when we position ourselves at the left end of this continuum (i.e., the real product). Similarly, enhancements in visual quality yield positive ramifications for our subjective assessments. As an illustration, the Meta Quest 2, a pivotal hardware component in these investigations, boasts an impressive resolution of 1832×1920 pixels per eye, a marked improvement from the 1920 x 1080 pixels common to computer LCD screens, the mediums through which the images are typically viewed. If the visual quality is the same, the introduction of haptic feedback can enhance the virtual experience and fortify user response confidence.

Chapter 6

6. Conclusions

This section presents the primary findings derived from the conducted research. The investigation has focused on examining the impact of novel and sophisticated visualization techniques, namely VR and AR, on the emotional response of users when assessing virtual prototypes. In this regard, several experimental studies were conducted, whereby various subjects evaluated different products utilizing diverse visualization methods.

New means of product representation have emerged in recent years to remplace traditional prototyping processes in product development, which can still be quite costly. Although these new technologies have proven to be quite efficient, this research has indicated that the user's perceptual response to a particular product may vary depending on the medium used to present it, which is in line with other studies.

Our results suggest that the medium used to present a product is important, and that the use of VRPH can increase purchasing decisions in some cases. Additionally, allowing users to physically interact with products can lead to more favorable emotional responses, thus trying to use means that favor the greatest possible interaction with the product can help to obtain more accurate evaluations regarding the actual product. However, it's important to note that presenting a product in a virtual medium can result in perceptual differences during evaluation, as some features of the real product may be difficult or impossible to replicate virtually, so it is essential to determine which medium to use when analyzing certain product characteristics.

In our case, using passive haptics in the virtual experience seemed to minimize these differences, and the use of 3D printers to create physical mockups that enable touch capabilities in the VR environment could be an effective strategy for some products. Future developments in object tracking capabilities and more accurate hand tracking performance of HMDs will facilitate the creation of VRPH experiences. Additionally, it is noteworthy that displaying products together can facilitate the evaluation of a specific product, as ranking products can be an easier task for the user when having a reference to compare with, rather than doing it individually.

Hence, presenting products jointly may reduce the perceptual differences resulting from the visualization medium employed, and integrating touch in virtual environments may provide a more authentic experience that elicits a comparable emotional response when assessing a product.

Within this context, we highlight that these findings hold significant value for product developers, marketers, and designers who endeavor to optimize the benefits of XR and create more captivating and effective products.

6.1. Limitations

This section describes several factors that limit the scope and applicability of the research work. These factors pertain to the selection of stimuli, the composition of the sample, the methods of measurement, and the software development.

The first limitation pertains to the selection of stimuli used in experimental studies. The stimuli chosen for this research work have largely consisted of everyday products such as furniture and small appliances. While different concepts have been selected to enable generalization of the results to different product typologies, many of the conclusions derived may not be extrapolated to other product categories. Therefore, it is essential to conduct experiments with different products to gauge the emotional response of the subjects regardless of the type of product selected.

The composition of the sample is tentatively considered a limiting factor in this research. As this work has been carried out within the academic framework, the volunteers recruited primarily consist of university students between the ages of 20 and 30 years old. Consequently, the results of the research may not be generalizable to older subjects due to the latter's potential reluctance towards new technologies, which may negatively impact the research findings.

Furthermore, the use of self-report questionnaires is deemed to be a limiting factor of this research work. Self-report questionnaires have been shown to be highly subjective, making the task of obtaining accurate results challenging. Future research should utilize more objective techniques, such as physiological measures like eye-tracking, to obtain more reliable results.

Finally, the development of software for this research may also limit the study design. The interaction of virtual prototypes with users is constrained to the tools currently available, which may limit precision in certain situations. It is expected that advancements in the field of virtual environment interaction will enable the development of more sophisticated experiments, leading to more robust results.

6.2. Future work

The present doctoral thesis has led to the emergence of new research studies focused on the emotional evaluation of products. In this regard, it has been observed that the self-report questionnaires used until now are highly subjective and do not offer completely reliable data on the user's emotional response in some cases. The release of new VR devices (i.e., Meta Quest Pro or HP Reverb Omnicept G2) capable of capturing physiological and behavioral measures of the user offers new opportunities to investigate the subject's perceptual response to a product in virtual environments.

In this context, product evaluation largely depends on the sense of sight. Thus, eye-tracking is a technique that has great potential to provide valuable data on the user's emotions or preferences (Marshall et al., 2014). Although this technique has been widely used during product evaluation (Kukkonen, 2005), new studies have emphasized its use in virtual environments during the evaluation of a virtual prototype. On the other hand, face- tracking is a tool that is seldom used as a measurement technique to evaluate affective responses during product evaluation, since its usefulness for measuring typically mild emotions generated by interaction with a product has yet to be established (Laurans et al., 2009). Therefore, great interest has been aroused in its study, as contributing to this field that has been so little explored can bring significant advances within product development. In this context, the Facial Action Coding System (FACS) will be used, which refers to a set of facial muscle movements that correspond to an emotion displayed, which allows determining the emotion shown by a participant (Farnsworth, 2022).

This research work has generated new collaborations with other universities. Specifically, a collaboration is being carried out with students from Purdue University to conduct experiments with eye-tracking. The objective is to compare the subject's emotional response to different stimuli presented jointly or individually between a VR environment and 2D images. At the same time, collaboration is taking place with researchers from Universitat Jaume I on the use of eye tracking, phase tracking, and different locomotion techniques to observe their effect on user perception in virtual environments.

In summary, the use of these new measures will allow for more rigorous and objective research on the subject's emotions, which can have important implications for both the industry and research in the field of psychology and marketing.

References

Achiche, S., Maier, A., Milanova, K., & Vadean, A. (2014). Visual Product Evaluation: Using the Semantic Differential to Investigate the Influence of Basic Geometry on User Perception. *Volume 11: Systems, Design, and Complexity*, 1–10. https://doi.org/10.1115/IMECE2014-40443

Aftab, M., & Rusli, H. A. (2017). Designing Visceral, Behavioural and Reflective Products. *Chinese Journal of Mechanical Engineering (English Edition)*, *30*(5), 1058–1068. https://doi.org/10.1007/s10033-017-0161-x

Agost, M. J. (2020). Analysis of perceptions according to the technique of product display. Comparative among 360º rotation, virtual and augmented reality. *24th International Congress on Project Management and Engineering*.

Agost, M. J., Vergara, M., & Bayarri, V. (2021). The Use of New Presentation Technologies in Electronics Sales Enviroments and Their Influence on Product Perception. In S. Yamamoto & H. Mori (Eds.), *Human Interface and the Management of Information. Information Presentation and Visualization* (Vol. 12765, pp. 3–15). Springer International Publishing. https://doi.org/10.1007/978-3-030-78321-1

Ant Ozok, A., & Komlodi, A. (2009). Better in 3D? an empirical investigation of user satisfaction and preferences concerning two-dimensional and three-dimensional product representations in business-to-consumer e-commerce. *International Journal of Human-Computer Interaction*, *25*(4), 243–281. https://doi.org/10.1080/10447310802546724

Arbeláez, J. C., & Osorio-Gómez, G. (2018). Crowdsourcing Augmented Reality Environment (CARE) for aesthetic evaluation of products in conceptual stage. *Computers in Industry*, *99*(April), 241–252. https://doi.org/10.1016/j.compind.2018.03.028

Arrighi, P. A., & Mougenot, C. (2019). Towards user empowerment in product design: a mixed reality tool for interactive virtual prototyping. *Journal of Intelligent Manufacturing*, *30*(2), 743–754. https://doi.org/10.1007/s10845-016-1276-0

Artacho-Ramírez, M. A., Diego-Mas, J. A., & Alcaide-Marzal, J. (2008). Influence of the mode of graphical representation on the perception of product aesthetic and emotional features: An exploratory study. *International Journal of Industrial Ergonomics*, *38*(11–12), 942–952. https://doi.org/10.1016/j.ergon.2008.02.020

Arvanitis, T. N., Petrou, A., Knight, J. F., Savas, S., Sotiriou, S., Gargalakos, M., & Gialouri, E. (2009). Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities. *Personal and Ubiquitous Computing*, *13*(3), 243–250.

https://doi.org/10.1007/s00779-007-0187-7

Aziz, H. (2018). Virtual Reality Programs Applications in Healthcare. *Journal of Health and Medical Informatics*, *9*(1), 305. https://doi.org/10.4172/2157-7420.1000305

Banerjee, N. T., Baughman, A. J., Lin, S., Witte, Z. A., Klaus, D. M., & Allison, P. (2021). Side-by-Side Comparison of Human Perception and Performance using Augmented, Hybrid, and Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*. https://doi.org/10.1109/TVCG.2021.3105606

Beck, M., & Crié, D. (2018). I virtually try it … I want it ! Virtual Fitting Room: A tool to increase on-line and off-line exploratory behavior, patronage and purchase intentions. *Journal of Retailing and Consumer Services*, *40*(December 2015), 279–286. https://doi.org/10.1016/j.jretconser.2016.08.006

Berni, A., Maccioni, L., & Borgianni, Y. (2020). An Eye-Tracking Supported Investigation Into the Role of Forms of Representation on Design Evaluations and Affordances of Original Product Features. *Proceedings of the Design Society: DESIGN Conference*, *1*, 1607–1616. https://doi.org/10.1017/dsd.2020.296

Berni, Aurora, & Borgianni, Y. (2020). Applications of Virtual Reality in Engineering and Product Design: Why, What, How, When and Where. *Electronics*, *9*(7), 1064. https://doi.org/10.3390/electronics9071064

Bimber, O., & Raskar, R. (2005). Spatial Augmented Reality: Merging Real and Virtual Worlds. In *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A K Peters/CRC Press. https://doi.org/10.1201/b10624

Bleier, A., Harmeling, C. M., & Palmatier, R. W. (2019). Creating effective online customer experiences. *Journal of Marketing*, *83*(2), 98–119. https://doi.org/10.1177/0022242918809930

Bordegoni, M. (2011a). Innovation in Product Design. In M. Bordegoni & C. Rizzi (Eds.), *Innovation in Product Design*. Springer London. https://doi.org/10.1007/978-0-85729-775-4

Bordegoni, M. (2011b). Product Virtualization: An Effective Method for the Evaluation of Concept Design of New Products. In M. Bordegoni & C. Rizzi (Eds.), *Innovation in Product Design* (pp. 117–141). Springer London. https://doi.org/10.1007/978-0-85729-775-4

Bruno, F., & Muzzupappa, M. (2010). Product interface design: A participatory approach based on virtual reality. *International Journal of Human-Computer Studies*, *68*(5), 254–269. https://doi.org/10.1016/j.ijhcs.2009.12.004

Camba, J. D., Soler, J. L., & Contero, M. (2017). Immersive Visualization Technologies to Facilitate Multidisciplinary Design Education. In P. Zaphiris & A. Ioannou (Eds.), *International conference on learning and collaboration* (Vol. 10295, pp. 3–11). Springer International Publishing. https://doi.org/10.1007/978-3-319-58509-3_1

Cascini, G., O'Hare, J., Dekoninck, E., Becattini, N., Boujut, J.-F., Ben Guefrache, F., Carli, I., Caruso, G., Giunta, L., & Morosi, F. (2020). Exploring the use of AR technology for co-creative product and packaging design. *Computers in Industry*, *123*, 103308. https://doi.org/10.1016/j.compind.2020.103308

Cecil, J., & Kanchanapiboon, A. (2007). Virtual engineering approaches in product and process design. *International Journal of Advanced Manufacturing Technology*, *31*(9–10), 846–856. https://doi.org/10.1007/s00170-005-0267-7

Cho, J. (2004). Likelihood to abort an online transaction: influences from cognitive evaluations, attitudes, and behavioral variables. *Information & Management*, *41*(7), 827–838. https://doi.org/10.1016/j.im.2003.08.013

Christopoulos, G., Kokkinaki, F., Harvey, N., & Sevdalis, N. (2011). Paying for no reason? (Mis-)perceptions of product attributes in separate vs. joint product evaluation. *Journal of Economic Psychology*, *32*(5), 857–864. https://doi.org/10.1016/j.joep.2011.05.003

Chu, C.-H., & Kao, E.-T. (2020). A Comparative Study of Design Evaluation with Virtual Prototypes Versus a Physical Product. *Applied Sciences*, *10*(14), 1–20. https://doi.org/10.3390/app10144723

Chu, C. H., Chen, Y. A., Huang, Y. Y., & Lee, Y. J. (2022). A Comparative Study of Virtual Footwear Try-On Applications in Virtual and Augmented Reality. *Journal of Computing and Information Science in Engineering*, *22*(4). https://doi.org/10.1115/1.4053328

Chuang, M. C., Chang, C. C., & Hsu, S. H. (2001). Perceptual factors underlying user preferences toward product form of mobile phones. *Internation Journal of Industrial Ergonomics*, *27*(4), 247–258. https://doi.org/10.1016/S0169-8141(00)00054-8

Cooper, R. G. (2019). The drivers of success in new-product development. *Industrial Marketing Management*, *76*, 36–47. https://doi.org/10.1016/j.indmarman.2018.07.005

Coutts, E. R., Wodehouse, A., & Robertson, J. (2019). A Comparison of Contemporary Prototyping Methods. *Proceedings of the Design Society: International Conference on Engineering Design*, *1*(1), 1313–1322. https://doi.org/10.1017/dsi.2019.137

Craig, A. B. (2013). *Understanding Augmented Reality: Concepts and applications.* Elsevier. https://doi.org/10.1016/C2011-0-07249-6

De Crescenzio, F., Bagassi, S., Asfaux, S., & Lawson, N. (2019). Human centred design and evaluation of cabin interiors for business jet aircraft in virtual reality. *International Journal on Interactive Design and Manufacturing*, *13*(2), 761–772. https://doi.org/10.1007/s12008-019-00565-8

Desmet, P. M. A., Ortíz Nicolás, J. C., & Schoormans, J. P. (2008). Product personality in physical interaction. *Design Studies*, *29*(5), 458–477. https://doi.org/10.1016/j.destud.2008.06.003

Desmet, Pieter. (2002a). Designing emotions, reviewed by Norman. *The Design Journal*, *6*(2), 2–4.

Desmet, Pieter. (2002b). *Designing Emotions*. Delft University of Technology.

Desmet, Pieter. (2007). Nine sources of product emotion. In *International Association of Societies of Design Research* (pp. 1–15).

Desmet, Pieter, Overbeeke, K., & Tax, S. (2001). Designing Products with Added Emotional Value: Development and Appllcation of an Approach for Research through Design. *The Design Journal*, *4*(1), 32–47. https://doi.org/10.2752/146069201789378496

Desmet, Pma. (2011). Nine sources of product emotion. In *Proceedings of the Conference: Interfejs użytkownika - Kansei w praktyce* (pp. 8–16). Wydawnictwo PJWSTK. http://www.sd.polyu.edu.hk/iasdr/proceeding/papers/Nine Sources of Product Emotion.pdf

Dittmar, H., Long, K., & Meek, R. (2004). Buying on the Internet: Gender Differences in On-line and Conventional Buying Motivations. *Sex Roles*, *50*(5/6), 423–444. https://doi.org/10.1023/B:SERS.0000018896.35251.c7

Donald, N. (2004). *Why we love (or hate) everyday things* (Vol. 66). Basic Books.

Eggers, F., Sattler, H., & Teichert, T. (2022). Choice-Based Conjoint Analysis. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of Market Research*. Springer International Publishing. https://doi.org/10.1007/978-3-319-57413-4

Engelbrektsson, P., Yesil, Ö., & Karlsson, I. C. M. (2000). Data collection method Participants Environment Mediating tools. *Th International Product Development Management Conference*, 29–30.

Evans, G., Hoover, M., & Winer, E. (2020). Development of a 3D Conceptual Design Environment Using a Head Mounted Display Virtual Reality System. *Journal of Software Engineering and Applications*, *13*(10), 258–277. https://doi.org/10.4236/jsea.2020.1310017

Farnsworth, B. (2022). *Facial Action Coding System (FACS) – A Visual Guidebook*. Research Fundamentals. https://imotions.com/blog/learning/research-fundamentals/facial-action-coding-system/

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Felip, F., Galán, J., García-García, C., & Mulet, E. (2019). Influence of presentation means on industrial product evaluations with potential users: a first study by comparing tangible virtual reality and presenting a product in a real setting. *Virtual Reality*, *24*(3), 439–451. https://doi.org/10.1007/s10055-019-00406-9

Flavián, C., Gurrea, R., & Orús, C. (2011). The Impact of Online Product Presentation on

Consumers' Perceptions. *International Journal of E-Services and Mobile Applications*, *1*(3), 17–37. https://doi.org/10.4018/jesma.2009070102

Flynn, T. N., & Marley, A. A. J. (1992). *Worst Scaling : Theory and Methods*. 178–201.

Forbes, T., Barnes, H., Kinnell, P., & Goh, M. (2018a). A Study into the Influence of Visual Prototyping Methods and Immersive Technologies on the Perception of Abstract Product Properties. In *DS 91: Proceedings of NordDesign 2018*.

Forbes, T., Barnes, H., Kinnell, P., & Goh, M. (2018b). A Study into the Influence of Visual Prototyping Methods and Immersive Technologies on the Perception of Abstract Product Properties. In O. Philip, E., Simon, S. and Johan (Ed.), *DS 91: Proceedings of NordDesign 2018*.

Galán, J., Felip, F., García-García, C., & Contero, M. (2021). The influence of haptics when assessing household products presented in different means: a comparative study in real setting, flat display, and virtual reality environments with and without passive haptics. *Journal of Computational Design and Engineering*, *8*(1), 330–342. https://doi.org/10.1093/jcde/qwaa081

Galán, J., García-García, C., Felip, F., & Contero, M. (2021). Does a presentation Media Influence the Evaluation of Consumer Products? A Comparative Study to Evaluate Virtual Reality, Virtual Reality with Passive Haptics and a Real Setting. *International Journal of Interactive Multimedia and Artificial Intelligence*, *6*(6), 196. https://doi.org/10.9781/ijimai.2021.01.001

Gibson, I., Gao, Z., & Campbell, I. (2004). A comparative study of virtual prototyping and physical prototyping. *International Journal of Manufacturing Technology and Management*, *6*(6), 503–522. https://doi.org/10.1504/IJMTM.2004.005931

Greengard, S. (2022). *Where Augmented Reality Is Going - And Why You Should Care*. https://gritdaily.com/where-augmented-reality-is-going-and-why-you-should-care/

Grohmann, B., Spangenberg, E. R., & Sprott, D. E. (2007). The influence of tactile input on the evaluation of retail product offerings. *Journal of Retailing*, *83*(2), 237–245. https://doi.org/10.1016/j.jretai.2006.09.001

Hagtvedt, H., & Adam Brasel, S. (2017). Color Saturation Increases Perceived Product Size. *Journal of Consumer Research*, *44*(2), ucx039. https://doi.org/10.1093/jcr/ucx039

Hannah, R., Joshi, S., & Summers, J. D. (2012). A user study of interpretability of engineering design representations. *Journal of Engineering Design*, *23*(6), 443–468. https://doi.org/10.1080/09544828.2011.615302

Hasan, B. (2010). Exploring gender differences in online shopping attitude. *Computers in Human Behavior*, *26*(4), 597–601. https://doi.org/10.1016/j.chb.2009.12.012

Hauk, N., Hüffmeier, J., & Krumm, S. (2018). Ready to be a Silver Surfer? A Meta-analysis on the Relationship Between Chronological Age and Technology Acceptance.

Computers in Human Behavior, *84*, 304–319. https://doi.org/10.1016/j.chb.2018.01.020

Heineken, E., & Schulte, F. P. (2007). Seeing Size and Feeling Weight: The Size-Weight Illusion in Natural and Virtual Reality. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(1), 136–144. https://doi.org/10.1518/001872007779598028

Higgins, J. J., Blair, R. C., & Tashtoush, S. (1990). THE ALIGNED RANK TRANSFORM PROCEDURE. *Conference on Applied Statistics in Agriculture*. https://doi.org/10.4148/2475-7772.1443

Hoermann, M., & Schwalm, M. (2015). Evaluation of a Vehicle Exterior's Sportiness Under Real vs. Virtual Conditions. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9179, pp. 470–479). https://doi.org/10.1007/978-3-319-21067-4_48

Holbrook, M. B. (1986). Emotion in the consumption experience: toward a new model of the human consumer. In *The role of affect in consumer behavior: emerging theories and applications* (Vol. 6, Issue 23). Lexington Books.

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin*, *125*(5), 576–590. https://doi.org/10.1037/0033-2909.125.5.576

Hsu, C. C., Fann, S. C., & Chuang, M. C. (2017). Relationship between eye fixation patterns and Kansei evaluation of 3D chair forms. *Displays*, *50*, 21–34. https://doi.org/10.1016/j.displa.2017.09.002

Hsu, S. H., Chuang, M. C., & Chang, C. C. (2000). A semantic differential study of designers' and users' product form perception. *International Journal of Industrial Ergonomics*, *25*(4), 375–391. https://doi.org/10.1016/S0169-8141(99)00026-8

Hu, L., Boggess, M., & Shepley, M. M. (2022). Comparing Expert, Quasi-Expert, and Novice Evaluations of Award-Winning Design Products Using the Consensual Assessment Technique. *Creativity Research Journal*, *00*(00), 1–19. https://doi.org/10.1080/10400419.2022.2157591

Insko, B. E. (2001). Passive haptics significantly enhances virtual environments. In *The University of North Carolina at Chapel Hill ProQuest*.

Jalil, N. A. A., Fikry, A., & Zainuddin, A. (2016). The Impact of Store Atmospherics, Perceived Value, and Customer Satisfaction on Behavioural Intention. *Procedia Economics and Finance*, *37*(16), 538–544. https://doi.org/10.1016/S2212-5671(16)30162-9

Jeong, S. W., Fiore, A. M., Niehm, L. S., & Lorenz, F. O. (2009). The role of experiential value in online shopping: The impacts of product presentation on consumer responses towards an apparel web site. *Internet Research*, *19*(1), 105–124. https://doi.org/10.1108/10662240910927858

Jerald, J. (2015). What Is Virtual Reality? In *The VR Book* (p. 9). Association for Computing Machinery. https://doi.org/10.1145/2792790.2792793

Jiang, Z., & Benbasat, I. (2007a). Investigating the influence of the functional mechanisms of online product presentations. *Information Systems Research*, *18*(4), 454–470. https://doi.org/10.1287/isre.1070.0124

Jiang, Z., & Benbasat, I. (2007b). The effects of presentation formats and task complexity on online consumers' product understanding. *MIS Quarterly: Management Information Systems*, *31*(3), 475–500. https://doi.org/10.2307/25148804

Jordan, P. W. (2002). *Designing pleasurable products: An introduction to the new human factors*. CRC press.

Kamil, M. J. M., & Abidin, S. Z. (2013). Unconscious Human Behavior at Visceral Level of Emotional Design. *Procedia - Social and Behavioral Sciences*, *105*, 149–161. https://doi.org/10.1016/j.sbspro.2013.11.016

Kato, T. (2019). Verification of perception difference between actual space and VR space in car design. *International Journal on Interactive Design and Manufacturing*, *13*(3), 1233–1244. https://doi.org/10.1007/s12008-019-00568-5

Kelly, J. W. (2022). Distance Perception in Virtual Reality: A Meta-Analysis of the Effect of Head-Mounted Display Characteristics. *IEEE Transactions on Visualization and Computer Graphics*, 1–13. https://doi.org/10.1109/TVCG.2022.3196606

Kent, L., Snider, C., Gopsill, J., & Hicks, B. (2021). Mixed reality in design prototyping: A systematic review. *Design Studies*, *77*, 101046. https://doi.org/10.1016/j.destud.2021.101046

Khatri, J., Marín-Morales, J., Moghaddasi, M., Guixeres, J., Giglioli, I. A. C., & Alcañiz, M. (2022). Recognizing Personality Traits Using Consumer Behavior Patterns in a Virtual Retail Store. *Frontiers in Psychology*, *13*(March), 1–17. https://doi.org/10.3389/fpsyg.2022.752073

Kim, T. H., & Choo, H. J. (2021). Augmented reality as a product presentation tool: focusing on the role of product information and presence in AR. *Fashion and Textiles*, *8*(1). https://doi.org/10.1186/s40691-021-00261-w

Kinzinger, A., Steiner, W., Tatzgern, M., & Vallaster, C. (2022). Comparing low sensory enabling (LSE) and high sensory enabling (HSE) virtual product presentation modes in e-commerce. *Information Systems Journal*, *32*(5), 1034–1063. https://doi.org/10.1111/isj.12382

Kreimeier, J., Hammer, S., Friedmann, D., Karg, P., Bühner, C., Bankel, L., & Götzelmann, T. (2019). Evaluation of different types of haptic feedback influencing the task-based presence and performance in virtual reality. *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 289–298. https://doi.org/10.1145/3316782.3321536

Kukkonen, S. (2005). Exploring eye tracking in design evaluation. *Joining Forces*, 119–126.

Kuliga, S. F., Thrash, T., Dalton, R. C., & Hölscher, C. (2015). Virtual reality as an empirical research tool — Exploring user experience in a real building and a corresponding virtual model. *Computers, Environment and Urban Systems*, *54*, 363–375. https://doi.org/10.1016/j.compenvurbsys.2015.09.006

Lau, H. Y. K., Mak, K. L., & Lu, M. T. H. (2003). A virtual design platform for interactive product design and visualization. *Journal of Materials Processing Technology*, *139*(1–3), 402–407. https://doi.org/10.1016/S0924-0136(03)00510-7

Laurans, G., Desmet, P. M. a, & Hekkert, P. (2009). Assessing Emotion in Interaction: Some problems and a new approach. *Proc. DPPI 2009*, *OCTOBER*, 1–10.

Lee, S., Chen, T., Kim, J., Kim, G. J., Sungho Han, & Pan, Z. (2004). Affective property evaluation of virtual product designs. *IEEE Virtual Reality 2004*, 207–292. https://doi.org/10.1109/VR.2004.1310076

Li, S. M., Chan, F. T. S., Tsang, Y. P., & Lam, H. Y. (2021). New product idea selection in the fuzzy front end of innovation: A fuzzy best-worst method and group decision-making process. *Mathematics*, *9*(4), 1–18. https://doi.org/10.3390/math9040337

Lin, X., Featherman, M., Brooks, S. L., & Hajli, N. (2019). Exploring Gender Differences in Online Consumer Purchase Decision Making: An Online Product Presentation Perspective. *Information Systems Frontiers*, *21*(5), 1187–1201. https://doi.org/10.1007/s10796-018-9831-1

Lindeman, R. W., Sibert, J. L., & Hahn, J. K. (1999). Hand-held windows: towards effective 2D interaction in immersive virtual environments. *Proceedings - Virtual Reality Annual International Symposium*, 205–212. https://doi.org/10.1109/vr.1999.756952

Liu, S. (2017). The Effect of On-line Product Presentation : A Comparative Study between 3D and General Product Presentation. *Journal of Industrial Distribution & Business*, *8*(5), 15–21. https://doi.org/10.13106/ijidb.2017.vol8.no5.15

Liu, X., & Yang, S. (2022). Study on product form design via Kansei engineering and virtual reality. *Journal of Engineering Design*, *33*(6), 412–440. https://doi.org/10.1080/09544828.2022.2078660

Mansouri, H., Paige, R. L., & Surles, J. G. (2004). Aligned rank transform techniques for analysis of variance and multiple comparisons. *Communications in Statistics - Theory and Methods*, *33*(9 SPEC.ISS.), 2217–2232. https://doi.org/10.1081/STA-200026599

Marquis, J., & Deeb, R. S. (2018). Roadmap to a successful product development. *IEEE Engineering Management Review*, *46*(4), 51–58. https://doi.org/10.1109/EMR.2018.2884275

Marshall, B. H., Sareen, S., Springer, J. A., & Reid, T. (2014). Eye tracking data understanding for product representation studies. *RIIT 2014 - Proceedings of the 3rd Annual Conference on Research in Information Technology*, 3–8. https://doi.org/10.1145/2656434.2656439

Martínez-Navarro, J., Bigné, E., Guixeres, J., Alcañiz, M., & Torrecilla, C. (2019). The influence of virtual reality in e-commerce. *Journal of Business Research*, *100*(January 2018), 475–482. https://doi.org/10.1016/j.jbusres.2018.10.054

Meta. (2021). *Introducing Meta: A Social Technology Company*. Meta Press Release. https://about.fb.com/news/2021/10/facebook-company-is-now-meta/

Milgram, P., & Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE TRANSACTIONS on Information and Systems*, *77*(12), 1321–1329.

Nodine, C. F., Locher, P. J., & Krupinski, E. A. (1993). The Role of Formal Art Training on Perception and Aesthetic Judgment of Art Compositions. *Leonardo*, *26*(3), 219. https://doi.org/10.2307/1575815

Norman, D. A. (2004). *Emotional Design*. Basic Books.

O'Keefe, R. M., & McEachern, T. (1998). Web-based Customer Decision Support Systems. *Communications of the ACM*, *41*(3), 71–78. https://doi.org/10.1145/272287.272300

Ortíz Nicolás, J. C., Aurisicchio, M., & Desmet, P. M. A. (2013). How users experience great products. *5th International Association of Societies of Design Research*, 5546–5557.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of the meaning*. Champaign, IL: University of Illinois press.

Oyman, M., Bal, D., & Ozer, S. (2022). Extending the technology acceptance model to explain how perceived augmented reality affects consumers' perceptions. *Computers in Human Behavior*, *128*(December 2021), 107127. https://doi.org/10.1016/j.chb.2021.107127

Ozer, M. (1999). A Survey of New Product Evaluation Models. *Journal of Product Innovation Management*, *16*(1), 77–94. https://doi.org/10.1111/1540-5885.1610077

Palacios-Ibáñez, A., Navarro-Martínez, R., Blasco-Esteban, J., Contero, M., & Camba, J. D. (2023). On the application of extended reality technologies for the evaluation of product characteristics during the initial stages of the product development process. *Computers in Industry*, *144*, 103780. https://doi.org/10.1016/j.compind.2022.103780

Palacios-Ibáñez, A., Ochando-Martí, F., D. Camba, J., & Contero, M. (2022). The influence of the visualization modality on consumer perception: A case study on household products. In In: Amic G. Ho (eds) (Ed.), *13th International Conference on Applied Human Factors and Ergonomics* (Vol. 49, pp. 160–166). https://doi.org/10.54941/ahfe1002050

Park, H., Faghihi, N., Dixit, M., Vaid, J., & McNamara, A. (2021). Judgments of object size and distance across different virtual reality environments: A preliminary study. *Applied Sciences (Switzerland)*, *11*(23). https://doi.org/10.3390/app112311510

Park, J., Lennon, S. J., & Stoel, L. (2005). On-line product presentation: Effects on mood, perceived risk, and purchase intention. *Psychology and Marketing*, *22*(9), 695–719.

https://doi.org/10.1002/mar.20080

Peck, J., & Childers, T. L. (2003). To have and to hold: The influence of haptic information on product judgments. *Journal of Marketing*, *67*(2), 35–48. https://doi.org/10.1509/jmkg.67.2.35.18612

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Perez Mata, M., Ahmed-Kristensen, S., Brockhoff, P. B., & Yanagisawa, H. (2017). Investigating the influence of product perception and geometric features. *Research in Engineering Design*, *28*(3), 357–379. https://doi.org/10.1007/s00163-016-0244-1

Pleyers, G., & Poncin, I. (2020). Non-immersive virtual reality technologies in real estate: How customer experience drives attitudes toward properties and the service provider. *Journal of Retailing and Consumer Services*, *57*, 102175. https://doi.org/10.1016/j.jretconser.2020.102175

Qu, Q. X., & Guo, F. (2019). Can eye movements be effectively measured to assess product design?: Gender differences should be considered. *International Journal of Industrial Ergonomics*, *72*(195), 281–289. https://doi.org/10.1016/j.ergon.2019.06.006

Ranaweera, A. T., Martin, B. A. S., & Jin, H. S. (2021). What you touch, touches you: The influence of haptic attributes on consumer product impressions. *Psychology & Marketing*, *38*(1), 183–195. https://doi.org/10.1002/mar.21433

Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior*, *133*(May 2021), 107289. https://doi.org/10.1016/j.chb.2022.107289

Ray, S., & Choi, Y. M. (2017). Employing design representations for userfeedback in the product design lifecycle. *21st International Conference on Engineering Design, ICED 2017*, *4*, 563–572.

Reid, T. N., MacDonald, E. F., & Du, P. (2013). Impact of product design representation on customer judgment. *Journal of Mechanical Design*, *135*(9), 1–12. https://doi.org/10.1115/1.4024724

Roy, R., Goatman, M., & Khangura, K. (2009). User-centric design and Kansei Engineering. *CIRP Journal of Manufacturing Science and Technology*, *1*(3), 172–178. https://doi.org/10.1016/j.cirpj.2008.10.007

Schoormans, J. P., Ortt, R. J., & De Bont, C. J. (1995). *Enhancing concept test validity by using expert consumers* (pp. 153–162). An International Publication of the Product Development & Management Association, 12(2).

Serrano, B., Botella, C., Baños, R. M., & Alcañiz, M. (2013). Using virtual reality and mood-induction procedures to test products with consumers of ceramic tiles. *Computers in Human Behavior*, *29*(3), 648–653. https://doi.org/10.1016/j.chb.2012.10.024

Singh, G., Ellis, S. R., & Swan, J. E. (2020). The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching. *IEEE Transactions on Visualization and Computer Graphics*, *26*(2), 1385–1398. https://doi.org/10.1109/TVCG.2018.2869729

Singh, J., & Sarkar, P. (2022). Visual Product Assessment by Using the Eye-Tracking Equipment to Study the Effect of Product Shapes on consumer's Thinking. In *Lecture Notes in Mechanical Engineering* (pp. 149–158). Springer Singapore. https://doi.org/10.1007/978-981-19-0676-3_12

Slater, M., McCarthy, J., & Maringelli, F. (1998). The Influence of Body Movement on Subjective Presence in Virtual Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(3), 469–477. https://doi.org/10.1518/001872098779591368

Söderman, M. (2005). Virtual reality in product evaluations with potential customers: An exploratory study comparing virtual reality with conventional product representations. *Journal of Engineering Design*, *16*(3), 311–328. https://doi.org/10.1080/09544820500128967

Solso, R. L. (2001). Brain Activities in a Skilled versus a Novice Artist: An fMRI Study. *Leonardo*, *34*(1), 31–34. https://doi.org/10.1162/002409401300052479

Solso, R. L., Vogt, S., & Magnussen, S. (2007). Expertise in pictorial perception: Eye-movement patterns and visual memory in artists and laymen. *Leonardo*, *36*(1), 91–100. https://doi.org/10.1162/002409401300052479

Song, H., Chen, F., Peng, Q., Zhang, J., & Gu, P. (2018). Improvement of user experience using virtual reality in open-architecture product design. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *232*(13), 2264–2275. https://doi.org/10.1177/0954405417711736

Stamer, M., Michaels, J., & Tümler, J. (2020). Investigating the Benefits of Haptic Feedback During In-Car Interactions in Virtual Reality. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 404–416). Springer International Publishing. https://doi.org/10.1007/978-3-030-50523-3_29

Statista. (2021). *Augmented reality (AR) and virtual reality (VR) market size worldwide from 2016 to 2024*. https://www.statista.com/statistics/591181/global-augmented-virtual-reality-market-size/

Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, *42*(4), 73–93. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x

Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior*, *86*, 77–90. https://doi.org/10.1016/j.chb.2018.04.019

Sylcott, B., Orsborn, S., & Cagan, J. (2016). The effect of product representation in visual conjoint analysis. *Journal of Mechanical Design, Transactions of the ASME*, *138*(10), 1–8. https://doi.org/10.1115/1.4034085

Tesch, A., & Dörner, R. (2020). Expert Performance in the Examination of Interior Surfaces in an Automobile: Virtual Reality vs. Reality. *Proceedings of the 28th ACM International Conference on Multimedia*, 2673–2681. https://doi.org/10.1145/3394171.3413980

Tiainen, T., Ellman, A., & Kaapu, T. (2014). Virtual prototypes reveal more development ideas: comparison between customers' evaluation of virtual and physical prototypes: This paper argues that virtual prototypes are better than physical prototypes for consumers-involved product development. *Virtual and Physical Prototyping*, *9*(3), 169–180. https://doi.org/10.1080/17452759.2014.934573

Tiger, L. (1992). *The Pursuit of Pleasure*. Little, Brown & Company.

Tran, L. T. T. (2021). Managing the effectiveness of e-commerce platforms in a pandemic. *Journal of Retailing and Consumer Services*, *58*(September 2020), 102287. https://doi.org/10.1016/j.jretconser.2020.102287

Unal, S. (2017). The Mediating Role of Product Familiarity in Consumer Animosity. *Journal of Accounting & Marketing*, *06*(04). https://doi.org/10.4172/2168-9601.1000257

Van Kleef, E., van Trijp, H. C. M., & Luning, P. (2005a). Consumer research in the early stages of new product development: a critical review of methods and techniques. *Food Quality and Preference*, *16*(3), 181–201. https://doi.org/10.1016/j.foodqual.2004.05.012

Van Kleef, E., van Trijp, H. C. M., & Luning, P. (2005b). Consumer research in the early stages of new product development: a critical review of methods and techniques. *Food Quality and Preference*, *16*(3), 181–201. https://doi.org/10.1016/j.foodqual.2004.05.012

Van Slyke, C., Comunale, C. L., & Belanger, F. (2002). Gender differences in perceptions of web-based shopping. *Communications of the ACM*, *45*(8), 82–86. https://doi.org/10.1145/545151.545155

Vergara, M., Mondragón, S., Sancho-Bru, J. L., Company, P., & Agost, M.-J. (2011). Perception of products by progressive multisensory integration. A study on hammers. *Applied Ergonomics*, *42*(5), 652–664. https://doi.org/10.1016/j.apergo.2010.09.014

Virzi, R. A., Sokolov, J. L., & Karis, D. (1996). Usability problem identification using both low- and high-fidelity prototypes. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Common Ground - CHI '96*, 236–243. https://doi.org/10.1145/238386.238516

Wang, J., & Lu, W. F. (2012). A haptics-based virtual simulation system for product design. *International Journal of Computer Applications in Technology*, *45*(2–3), 163–170. https://doi.org/10.1504/IJCAT.2012.050705

Wang, Y., Jia, F., Schoenherr, T., Gong, Y., & Chen, L. (2020). Cross-border e-commerce firms as supply chain integrators: The management of three flows. *Industrial Marketing Management*, *89*, 72–88. https://doi.org/10.1016/j.indmarman.2019.09.004

Whitfield, A., & Wiltshire, J. (1982). Design training and aesthetic evaluation: An intergroup comparison. *Journal of Environmental Psychology*, *2*(2), 109–117. https://doi.org/10.1016/S0272-4944(82)80043-1

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*, 143. https://doi.org/10.1145/1978942.1978963

Wu, K., Vassileva, J., Zhao, Y., Noorian, Z., Waldner, W., & Adaji, I. (2016). Complexity or simplicity? Designing product pictures for advertising in online marketplaces. *Journal of Retailing and Consumer Services*, *28*, 17–27. https://doi.org/10.1016/j.jretconser.2015.08.009

Ye, J., Badiyani, S., Raja, V., & Schlegel, T. (2007). Applications of Virtual Reality in Product Design Evaluation. In *Human-Computer Interaction. HCI Applications and Services* (pp. 1190–1199). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-73111-5_130

Yoo, J., & Kim, M. (2014). The effects of online product presentation on consumer responses: A mental imagery perspective. *Journal of Business Research*, *67*(11), 2464–2472. https://doi.org/10.1016/j.jbusres.2014.03.006

Zheng, L., & Bensebaa, F. (2022). Need for touch and online consumer decision making: the moderating role of emotional states. *International Journal of Retail & Distribution Management*, *50*(1), 55–75. https://doi.org/10.1108/IJRDM-04-2020-0158

Zubair, S., Ansari, A., Shukla, V. K., Saxena, K., & Filomeno, B. (2022). *Implementing Virtual Reality in Entertainment Industry* (J. M. R. S. Tavares, P. Dutta, S. Dutta, & D. Samanta (eds.); Vol. 291). Springer Singapore. https://doi.org/10.1007/978-981-16-4284-5