TYPE Original Research
PUBLISHED 06 January 2025
DOI 10.3389/frvir.2024.1505630

OPEN ACCESS

EDITED BY
Mercedes Inda-Caro,
University of Oviedo, Spain

REVIEWED BY
Javier Spector,
Austral University, Argentina
Seyeoung Chun,
Chungnam National University,
Republic of Korea
James Ko,
The Education University of Hong Kong,
Hong Kong SAR, China

*CORRESPONDENCE
Imanol Ceberio,
imanolceberio@opendeusto.es

RECEIVED 03 October 2024
ACCEPTED 28 November 2024
PUBLISHED 06 January 2025

CITATION
Ceberio I, Al-Rashaida M, García M, Lopez Paz JF, Salgueiro M, Passi N, Pavel H and Amayra I (2025) Content and face validity in virtual reality with children: a validation in five steps+1 of a wheelchair basketball game. *Front. Virtual Real.* 5:1505630. doi: 10.3389/frvir.2024.1505630

Content and face validity in virtual reality with children: a validation in five steps+1 of a wheelchair basketball game

Imanol Ceberio[1]*, Mohammad Al-Rashaida[2], Maitane García[1], Juan Francisco Lopez Paz[1], Monika Salgueiro[3], Nicole Passi[1], Humberto Pavel[4] and Imanol Amayra[1]

[1]Department of Psychology, Faculty of Health Sciences, University of Deusto, Bilbao, Spain, [2]Department of Special Education (CEDU), United Arab Emirates University, Al-Ain, United Arab Emirates, [3]Department of Clinical and Health Psychology, and Research Methodology, Faculty of Psychology, University of the Basque Country UPV/EHU, Donostia, Spain, [4]Sociedad Vasca de Minusválidos Bidaideak, Bilbao, Spain

Objective: The present study is aimed to elaborate and determinate the content and face validity of a virtual reality program attending the perspective of children. This simulation is designed to promote empathy and understanding towards children with motor disabilities through adapted sport scenes. This study proposed a validation approach with six phases to assess technical and aesthetic aspects.

Method: Sample: a) Content validity study, 20 children (11–18 years old) were recruited as lay and content experts, who assessed the properties of grade of realism and physical fidelity of wheelchair basketball scenarios. b) Face validity study, 395 children were recruited as lay experts and divided into two groups (7–9 years old and 10–12 years old), or into ten subgroups according to Age × Gender interaction. The face validity sample assessed the psychological fidelity and the presence of wheelchair basketball scenarios. *Instruments:* Virtual Reality Content Validity Questionnaire, GAMEX questionnaire and Simulator Sickness Questionnaire (SSQ).

Results: The content validity study showed preference for technical aspects (music, colors and degree of realism). Therefore, modifications in the design were made. In the content validity study, the high agreement level was influenced by previous sport experiences. In the face validity study, the cognitive development of children determined the differences in agreement levels in some virtual properties (absorption and cybersickness). In this sixth step, the ages 7–8 years *versus* 10–11 years showed significant differences in validity. The study also criticized the face validity cut-offs often used in adult-focused research, emphasizing the need to adapt them for children´s developmental stages.

Conclusion: This study proposes a sixth step not traditionally included in content and face validity processes, specially focusing on the child user. The suitability of content and scenes should follow the same principles of standardization as other methodologies, such as psychometric tests, considering age and gender.

KEYWORDS

content validity, face validity, children, experts, virtual reality

1 Introduction

The present study examines the originality of a virtual reality program aimed at fostering inclusion for children with motor disabilities through an innovative approach that extends the traditional validation phases by introducing a sixth step. This additional step focuses on analyzing the influence of developmental stages through the interaction of Age × Gender.

Virtual reality (VR) is becoming an increasingly popular tool for promoting empathy and understanding in environments in which individuals face some form of oppression. According to a growing body of research, immersive technology has the potential to foster empathy and perspective-taking (Herrera et al., 2018; Irom, 2018; Miner, 2022). Specifically, VR has been shown to be effective in promoting empathy and understanding among individuals with disabilities (Ahn et al., 2013; Herrera et al., 2018; Herrera and Bailenson, 2021; Jiang et al., 2022; Lázaro et al., 2020; Maister et al., 2013; Matera et al., 2021; Neyret et al., 2020; Oh et al., 2016; Pan and Hamilton, 2018; Peck et al., 2013; Reinhard et al., 2020; Rosenberg et al., 2013; Slater and Banakou, 2021; Yee and Bailenson, 2006). This is due to the capability of virtual reality to generate simulated environments in which individuals can experience the challenges faced by people with disabilities (Kang and Kang, 2019; Pinto-Coelho et al., 2023; Slater and Banakou, 2021). Through these experiences, individuals can gain a deeper understanding of and respect for the obstacles faced by people with disabilities. Additionally, VR can provide an opportunity for non-disabled individuals to develop empathy and perspective-taking, ultimately fostering a deeper understanding and respect of the experiences of individuals with disabilities.

Virtual reality possesses various characteristics that can be assessed using validation procedures (Harris et al., 2020). The first is the immersion, in which the person should analyze whether the sounds or images with which the user is going to interact are appropriate (Burdea and Coiffet, 2003). Second, presence understood from subjective experiences, such as the illusion of being there and the perception of realism or plausibility of the environment (Slater and Sanchez-Vives, 2016). Third, fidelity or the degree to which the simulation recreates the real world in terms of physical fidelity (visual information, congruence with the laws of physics or functionality) or psychological fidelity (demand for a similar degree of attention, effort, or generation of affective states) (Harris et al., 2020).

In the realm of virtual reality, it is crucial to conduct extensive testing prior to implementation to ensure that it accurately conveys the necessary perceptual, cognitive, and emotional aspects of the task (Harris et al., 2020; Runswick, 2023; Thomas et al., 2003). According to the methodological guidelines followed in behavioral sciences, the opinion of experts and users is integral to both content and face validity (Allen et al., 2023). These methodologies are well-suited for assessing the specific properties of virtual reality. Content validity examines the extent to which stimuli represent a representative sample of the theoretical domain of the construct (Handage and Chander, 2021; Hardesty and Bearden, 2004; Harris et al., 2020; Nunnally and Bernstein, 1994), allowing us to investigate the level of physical or psychological fidelity of the simulation. By contrast, face validity aims to enhance virtual items, arrange them appropriately, and evaluate their appeal, sensory appearance, or degree of realism (Babar et al., 2023; Feeley et al., 2021; Harris et al., 2020).

Designing and validating a virtual reality program presents a challenge for replicating objective reality in all its perceptual and cognitive dimensions. One approach to mitigate this issue is to involve the target population or lay experts in the development of the virtual construct. This ensures that the most important aspects of the virtual world are represented accurately. Professionals, experts, and psychometricians (content experts) cannot replace lay experts' unique perspectives and insights, as has been demonstrated in fields such as Psychology and Medicine (Allen et al., 2023; Connell et al., 2018; Handage and Chander, 2021; Zamanzadeh et al., 2015).

While a clinician may have a different opinion on what constitutes a good outcome compared to a user or lay expert, it is crucial to consider the perspective of the potential subject of research (Connell et al., 2018; Zamanzadeh et al., 2015). In this context, there are two possible solutions. The first would be to consider the results obtained in the design of serious games with children in order to improve the realness of the interface, the way in which the information is shown to the user, to allow the user to explore the simulated system, and to present the information according to their level of development (Valenza et al., 2019). The second solution would be the inclusion of children as experts in assessing the validity of measurement methods is becoming more frequent (Marlenga et al., 2017; Valentini et al., 2018; Robinson and Palmer, 2017). Considering the unique perspective of lay experts and incorporating elements from children's games can help ensure that virtual reality programs are as realistic and effective as possible.

Conducting research with children as experts offers a unique opportunity to present their lives and experiences from their perspective (Danby, 2009; Stewart, et al., 2005; Mason and Danby, 2011). Consequently, children have been included in studies to assess their perceptions of physical competencies through images (pictorially) (Lopes et al., 2016; Morgado et al., 2023; Valentini et al., 2018; Venetsanou et al., 2018), the usefulness of clinical questionnaires, evaluating the ambiguity and clarity of items and instructions (de Oliveira et al., 2022; Manan et al., 2024; Ryberg et al., 2023), and the quality of virtual reality content and imagery (Marlenga et al., 2017; Schwebel et al., 2008). The validity perspective of children appears to differ significantly from that of adults. While children tend to focus on first and third perspective-taking in virtual environments, this perception diminishes among adults (Choudhury et al., 2006). A similar trend is observed in the aesthetic perception of image quality, where children assign greater importance to realism, while adults emphasize expressiveness, originality, and creativity (Almeida-Rocha et al., 2020; Marlenga et al., 2017; Schwebel et al., 2008). Additionally, face validity in children is influenced by their lower levels of attention, particularly in younger children, which tends to improve with age (Hoyer et al., 2021).

In the context of children's involvement in gaming, including serious games and virtual reality, the validity of game design is influenced by the failure of practitioners such as researchers, teachers, and scientists to prioritize the enjoyment of the child (Bossavit and Parsons, 2018; Granic et al., 2014). One reason to involve children and adolescents directly in user-centered design processes is that they frequently engage in video gameplay. On average, children spend 50–300 h playing video games annually (VanDeventer and White, 2002). In addition, a significant proportion of children are familiar with Virtual Reality technology. In the United States, only 19% of the children were unfamiliar with this technology in 2017 (Yamada-Rice et al., 2017). Given that children are immersed in a technological environment, they are considered experts (Hague and Payton, 2010; Plowman et al., 2012).

The active engagement of child experts in the developmental stages of psychological, medical, and virtual measures can enhance their acceptance. This entails gathering assessments of a program's clarity, completeness, and representativeness of its domains (Bernstein and Belicki, 1996; Holden and Passey, 2010; Nevo, 1985). When examining the validity of judgments of typically developing children, it is crucial to consider the cognitive processes involved at specific ages. The quality of judgment may depend on the type and amount of information processing related to attention. One primary process is sustained attention to visual stimuli, which is influenced by tonic and phasic arousal. According to various studies, tonic arousal maturity occurs between the ages of 6 and 13 years. This process is related to the ability to maintain attention on the content of a task for an extended period (Hoyer et al., 2021). Between the ages of 9 and 13 years, there is a decrease in phasic arousal, or the vigilance state associated with sensory stimulus conditions of the task (e.g., sounds) (Hoyer et al., 2021), as well as changes in the autonomic nervous system, which consists of the sympathetic and parasympathetic nervous systems related to neurovegetative symptoms (headache, stomach-ache, or nausea).

The development of cognitive processes in children has been demonstrated to result in an increase in sustained attention levels between the ages of 8 and 12, as verified by research conducted on children, adolescents, and young adults in the presence of virtual stimuli (Baumgartner et al., 2008). According to a study by Baumgartner et al. (2008), children aged 8–11 years exhibit lower levels of activation in the prefrontal cortex than adolescents aged 13–17 years. This distinction may account for the fact that, when exposed to virtual environments, children report higher levels of presence and realness than adults (Bailey and Bailenson, 2017). The degree to which children consider certain aspects of virtual reality (such as sounds and images) to be most relevant may be influenced by the cognitive processes that develop in early childhood and the stimuli to which they pay the most attention.

The period from 7 to 11 years of age has been described by Inhelder and Piaget (1958) as being characterized by phases of egocentrism during infancy. During this stage, children may have difficulty differentiating their own mental constructs from the observed phenomena, such as images or movies. Egocentrism can impact the rotational images and the visual perspective of the observer, and can lead to confusion (Blakemore and Choudhury, 2006). The critical period for improving spatial perspective-taking is approximately 8 years of age (Beatini et al., 2024). At 10 years of age, inhibitory control in situations begins to improve (Sheridan et al., 2017), which may explain why children at this age have fewer difficulties ignoring distractions and disregarding irrelevant information. As previously mentioned, the type of information processed, degree of attention, egocentrism, or inhibitory control can all influence the judgments of face validity in children when evaluating the quality or objectives of a virtual reality program.

There is a dearth of literature on the subject of face validity or content validity studies of virtual reality applications involving children, specifically concerning their expertise in various areas. Marlenga et al. (2017) investigated the simulation of actions on a tractor, while Valentini et al. (2018) studied perceived movement competence. Robinson and Palmer (2017) assessed perceived motor competency, Schwebel et al. (2008) explored traffic situations, and Shen et al. (2022) evaluated the validity of a virtual reality cognitive assessment tool for children with traumatic brain injury. However, few studies have involved children in the design of serious games (Sim et al., 2015) or their aesthetic features (Javora et al., 2019). To our knowledge, no study has analyzed content and face validity in relation to the evolutionary cognitive development of children. In contrast, the development of virtual reality simulation programs for adults in specific domains, such as supervisory training programs in medicine or improving special clinical skills, has been well-documented (Babar et al., 2023; Bright et al., 2012; Chuan et al., 2023; Dorozhkin et al., 2017; Feeley et al., 2021; Runswick, 2023).

In clinical studies, face and content validity are typically considered important aspects of evaluation methodology. However, the studies reviewed differed in the proportion of participants with varying levels of expertise and number of participants (Alsalamah et al., 2017; Alvarez-Lopez et al., 2020; Babar et al., 2023; Bright et al., 2012; Chuan et al., 2023; Dorozhkin et al., 2017; Feeley et al., 2021; Runswick, 2023). The quality of content and face validity are quantified using different criteria that do not seem to follow the standards proposed by Lawshe (1975) or Lynn (1985) and developed by others in the application of other disciplines (Handage and Chander, 2021; Romero Jeldres et al., 2023; Zamanzadeh et al., 2015). According to these criteria, there are five phases in the assessment of content and face validity: 1) identification of the domain, 2) generation of virtual items, 3) determination of the number of raters, 4) establishment of content validity, and 5) establishment of face validity. Face validity in child age groups may be determined by subjective variables such as motivation, mood, or attitudes towards the subject being studied.

In this study, we propose the use of a virtual reality program, BSR-RV 2®, based on the five established validation steps, to which we add a new step not previously considered in earlier studies. This additional step evaluates the influence of developmental progression by analyzing the interaction of Age x Gender during the validation process. To this end, the BSR-RV 2® program has been designed to enhance inclusive attitudes towards individuals with disabilities. This technology can help to address the lack of knowledge regarding disabilities. It is essential to acknowledge the role of sports, which embody values such as camaraderie, support, and empathy. Consequently, to cultivate values such as solidarity, honesty, and responsibility in children through sporting activities (Mendoza Estor, 2017), this type of activity should be encouraged, as it favors the development of positive attitudes, promoting the teaching of values and a willingness to get to know others (Ruiz Amayas, 2019). By experiencing a virtual body, individuals may modify their attitudes and behaviors to align with what they perceive as societal expectations of individuals with that particular body type (Yee et al., 2009). Virtual reality can simulate the challenges and limitations faced by individuals with physical disabilities, ultimately leading to more positive attitudes towards those without disabilities

(Ceberio, 2022; Chowdhury et al., 2021; Jiang et al., 2022; Matera et al., 2021; Tammy Lin and Wu, 2021).

Moreover, to our knowledge, no study has specifically validated a virtual reality environment for transforming the attitudes of children toward individuals with disabilities via wheelchair basketball simulation. However, few studies have analyzed the change in attitudes in children using virtual reality scenarios aimed at analyzing the role of perspective taking (Matera et al., 2021). Therefore, the primary objective of this study was to determine the content and face validity of the BSR-RV 2® with children as experts following the five-phase procedure described above, with the addition of a new phase focused on developmental evolution. One of the most innovative aspects of this study is the transparent way the entire validation process is presented. Typically, this process is described briefly and often lacks clear and concise explanations. For this reason, we believe that this study may serve as a valuable resource for those interested in gaining a deeper understanding of each step in the validation process of a scale.

This game was designed to foster positive attitudes toward individuals with disabilities, and the properties related to the virtual reality program in content and face validity were linked to a) technical considerations, b) number and type of scenes of wheelchair basketball, c) sounds and images that evoke the sports environment, d) level of difficulty of the scenes, e) congruence with the laws of physics through chair and ball movements, f) adverse effects due to the use of the simulation, g) enjoyment, h) absorption, i) intrinsic motivation, j) absence of negative affect, k) activation, and l) adverse effects due to the use of the simulation.

In contrast, this study employs a method of examining virtual properties (Harris et al., 2020) that adhere to quality standards for content analysis and face validity, considering variables such as gender, age, area and level of expertise, and the number of experts involved. The program is intended for children between the ages of 7 and 12 and stands out for assessing expertise based on developmental psychology criteria (Staudinger et al., 2003). These criteria have been utilized in other studies that have adopted the same theoretical framework to differentiate their samples (Bart et al., 2008; Hoyer et al., 2021).

2 Methodology

2.1 Participants

For the present study, three distinct sample recruitment processes were conducted. First, a focus group was assembled to identify the main domains of the virtual reality program. Second, a sample was recruited for a content validity study, and finally, a sample was gathered for the face validity study.

2.1.1 Focus group

The focus group was recruited by convenience and completed by four experts (N = 4) in adapted basketball from two clubs (Bidaideak Bilbao Basket and Fundación Vital Zuzenak) who participated in this phase, which allowed us to identify the main domains of the virtual reality program. These individuals met the inclusion criterion of having a minimum of 8 years of experience in national and international competitions.

2.1.2 Content validity study

A total of 20 participants (N = 20) were selected, including 14 lay experts (conventional basketball players) and six content experts (wheelchair basketball players), aged 11–18 years (Table 1). The participants were recruited from the summer programs of Bilbao Basket and Bidaideak Bilbao BSR in Bilbao, Biscay. Before initiating the study, informed consent was obtained from the parents of all participants. Lay experts were defined as adapted basketball players with a minimum of 2 years of experience, while content experts were defined as adapted basketball players with a minimum of 3 years of experience. The participants were recruited from the summer programs of Bilbao Basket and Bidaideak Bilbao BSR in Bilbao, Biscay. Before initiating the study, informed consent was obtained from the parents of all participants.

The inclusion and exclusion criteria were as follows: a) competitive basketball players, b) schooling, and d) informed parental consent. The exclusion criteria were as follows: (a) photosensitivity reaction, (b) refusal of consent, and (c) visual functional diversity.

2.1.3 Face validity study

In the face validity study, we carefully selected a diverse sample of 395 children aged 7 to 12, none of whom had prior experience in wheelchair basketball, ensuring gender and age parity (χ^2 (1) = 0.375; p = 0.540) (Table 1). The children were given the opportunity to experience virtual sports challenges such as moving around a field or making baskets or passes while using a virtual wheelchair. The validation process took 3 months to complete (September to November of 2023) and involved children from schools in Basque Country (*Colegio Alazne* N = 191), Cantabria (*Colegio Riomar* N = 53), and Castilla and Leon (*Colegio Nuestra Señora de las Altices* N = 105; *Colegio Santa Cecilia* N = 65). The children participated in two trials, 1 week apart.

The inclusion criteria were as follows: a) gender parity, b) schooling, and c) informed parental consent. The exclusion criteria were as follows: (a) photosensitivity reaction, (b) refusal of consent, and (c) visual functional diversity.

2.1.4 Ethical considerations

In both studies, the ethical principle of beneficence was upheld. Children were informed of the possibility of experiencing mild physical side effects (e.g., cybersickness) and their right to withdraw from the trial at any time. Children with medical contraindications (e.g., neurological conditions) were advised not to participate in the study. The same considerations applied to potential emotional reactions stemming from visual stimulation (e.g., stress, anxiety).

Both studies ensured the confidentiality of participants and their data in accordance with the Organic Law on Personal Data Protection (15/1999, December 13) and international standards of the Declaration of Helsinki. This study was approved by the Ethics Committee of the University of Deusto (ETK-4/21-22).

2.2 Study design and procedure

This study involved two interconnected studies: the first focuses on content validity, and the second on face validity.

TABLE 1 Demographic distribution of participants in both studies.

	N	Age			Gender			
		Range (years)	Mean (years)	SD	Male		Female	
					N	%	N	%
Content validity study	20	11 to 18	14,95	3,63	16	80%	4	20%
Face validity study	395	7 to 12	9,58	1,304	206	52,2%	189	47,8%
	176	7 to 9	8,28	0,641	99	56,3%	77	43,8%
	219	10 to 12	10,62	0,548	107	48,9%	112	51,1%

The overall objective was to assess the validity of the BSR-RV 2® virtual reality program. A summary of all phases is shown in Figure 1 (Appendix).

2.2.1 Step 1. Identifying of the domain

The first phase involved identifying the domain by conducting a literature review on terms related to virtual reality, content and face validation of simulated environments, and disability. The content was determined by reviewing literature from the Web of Science, Scopus, and Proquest databases. Constructs related to attitudes towards people with disabilities and wheelchairs, skill acquisition through virtual reality, and content and face validity in virtual reality were examined.

2.2.2 Step 2. Generation of virtual items: focus group

Virtual items were generated using the following properties: scenes of adapted wheelchair sports, sounds and images evoking a sports environment, level of difficulty related to wheelchair basketball, and congruence with the laws of physics through chair and ball movements. Four experts in wheelchair basketball from two basketball clubs helped on the identification of these aspects. Three focus group sessions were conducted between January and February of 2023. In the second phase of our study, virtual items such as scenes of adapted wheelchair sports were meticulously crafted. These included auditory and visual elements that were designed to emulate a realistic sports environment. Each scene varied in difficulty to adequately challenge the participants, and all movements of the chairs and balls strictly adhered to the laws of physics for authenticity. A systematic procedure to assess these virtual items was employed: children interacted with the VR environment under controlled conditions, and their interactions were monitored and recorded.

The guidelines provided by Hollis et al. (2002) and Dahlin Ivanoff and Hultberg (2006) were followed for the recruitment of a focus group.

2.2.3 Step 3. Determining the number of raters

The number of raters who participated in the content and face validity study was determined according to the guidelines provided by Lynn (1985) and Romero Jeldres et al. (2023). To study the content validity, both lay and content experts in the field of basketball were chosen. To study the face validity, lay experts in the field of new technologies were chosen.

2.2.4 Step 4. Establishing content validity

Content validity was assessed by experts (lay and content experts), who determined the elimination of virtual items based on the proportion of judges in agreement. The criteria of Lawshe (1975) and Zamanzadeh et al. (2015) were followed. The Virtual Reality Content Validity Questionnaire, a 9-item five-point scale (1 = strongly disagree to 5 = strongly agree), was used. Scores of 4 (agree) and 5 (totally agree) were considered for item inclusion.

For the side effects study the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993), a four-point scale was applied: 0 = absence, 1 = slight, 2 = moderate, and 3 = severe. The last two scores, 0 (absent) and 1 (slight), were considered for validity purposes for item inclusion.

The Content Validity Ratio (CVR) was then applied, where Ne is the number of panelists indicating agree or strongly agree, and N is the number of panelists.

$$\text{CVR} = \frac{NE - N/2}{N/2}$$

The numerical value of the CVR according to Lawshe's table for the 20 panelists was 0.42 (Lawshe, 1975; Romero Jeldres et al., 2023).

The Content Validity Index (CVI) can be categorized into two types: Item-level CVI (I-CVI) and Scale-level CVI (S-CVI). The S-CVI can be measured in two different ways: one method calculates the average of the I-CVI scores across all items on the scale (S-CVI/Ave), while the other evaluates the proportion of items rated as either 3 or 4 in relevance by all experts (S-CVI/UA). In this study, the Item-level Content Validity Index (I-CVI) was applied to the proportion of raters giving an item agreement rating of 4 or 5 on the Virtual Content Validity Questionnaire and 0 or 1 on the SSQ. The numerical value for the I-CVI as acceptable is 0.79 as proposed by Yusoff (2019a).

$$\text{I-CVI} = \frac{Raters\,in\,agreement}{Number\,of\,total\,raters}$$

The choice to use the Item-level Content Validity Index (I-CVI) over the Scale-level Content Validity Index (S-CVI) stems from the ability of the I-CVI to focus on the relevance of individual items. The objective is to identify the most valid items in the scale (I-CVI) rather than initially considering the proportion of total items deemed valid (S-CVI/Ave).

In this study, participants were exposed to a virtual reality program for approximately 10–15 min. The study lasted 1 week in July 2023 and covered various aspects, including wheelchair

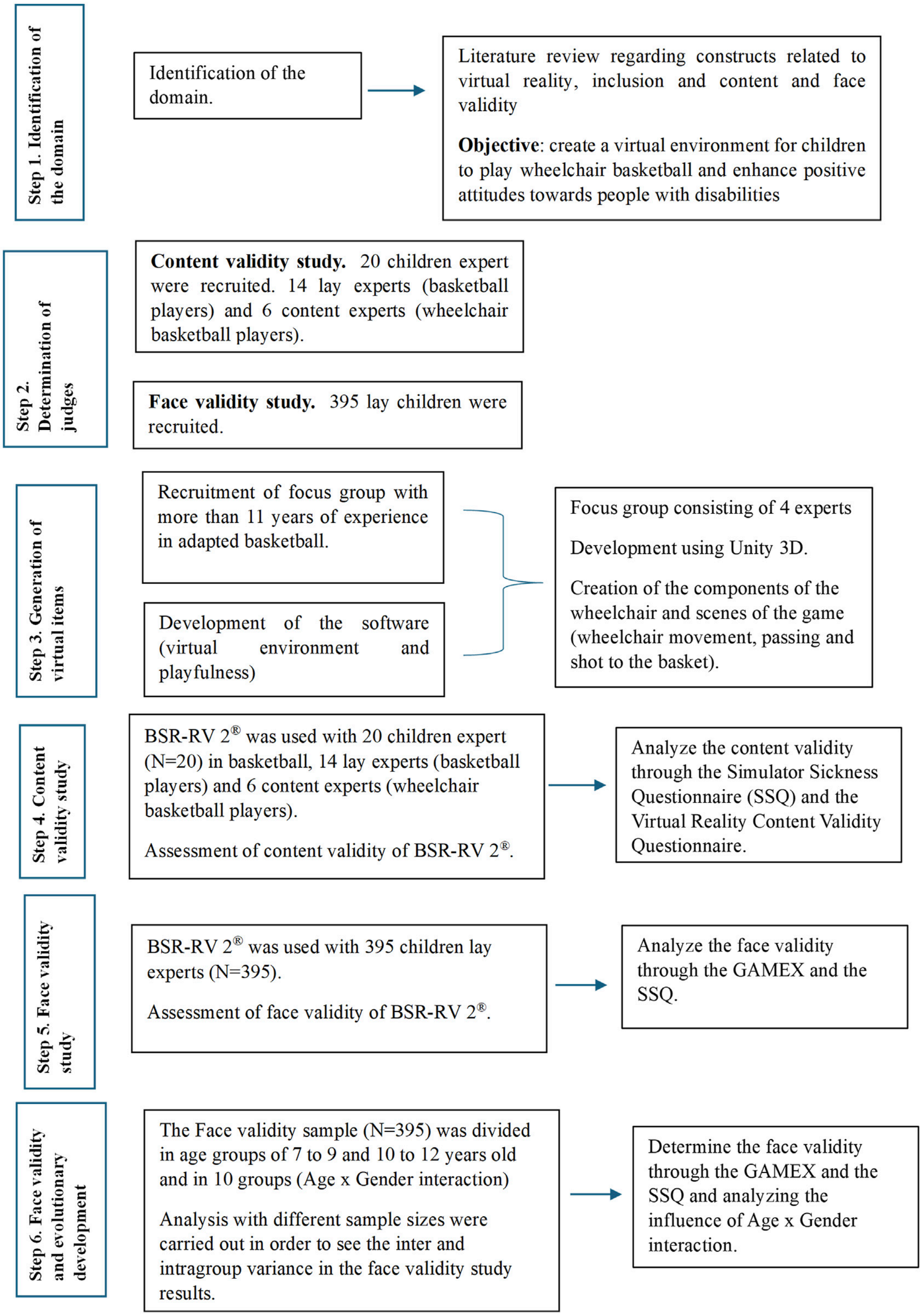

FIGURE 1
BSR-RV 2 Validation Process.

movement, passing and shooting at the basket. Participants were given the opportunity to experience challenges commonly faced in wheelchair basketball, such as moving around the field, passing, and shooting baskets while seated in a virtual wheelchair. The items evaluated by both groups of experts were related to the understanding of instructions, identification of the elements of the scenario, learning to recognize the mobility difficulties of wheelchair use, and discovering the perceptual limitations experienced by people during the adapted game, regarding appearance (sensory appearance, attractiveness, physical appearance) and the degree of realism.

The participants' expertise was also valuable in assessing the degree of realism, instruction comprehension of the items of all the scenes, capacity to acknowledge objects, level of difficulty, and aesthetic values, such as images and sound quality, measured using the Virtual Reality Content Validity Questionnaire. Cybersickness in the virtual reality game was measured using the Simulator Sickness Questionnaire (SSQ).

2.2.5 Step 5. Establishing face validity

Once the total number of items constituting the virtual reality program was selected, the fifth step involved assessing face validity based on criteria, such as psychological fidelity (enjoyment, intrinsic motivation, absence of negative affect, and activation), presence (absorption), and side effects generated using virtual reality. Aspects such as program suitability, attractiveness, sensory appearance, and the degree of realism have been considered (Holden and Passey, 2010; Nevo, 1985; Yusoff, 2019b). The items were measured on two scales: the GAMEX scale, which evaluates acceptance, understanding, and familiarity with the virtual experience using a five-point scale (1 = strongly disagree to 5 = strongly agree), and the SSQ, which assesses potential side effects using a four-point scale (0 = absence to 3 = severe).

Two formulas were applied for this purpose: the impact score (Zamanzadeh et al., 2015) and the Item-Level Face Validity Index (I-FVI) (Yusoff, 2019b). The impact score of each item was determined based on the proportion of patients who identified it as important and the mean importance score attributed to the item. In the quantities method, for calculating the item impact score, the first is calculated as the percentage of patients who scored 4 or 5 in the GAMEX and 0 or 1 in the SSQ to item importance (frequency), and the mean importance score of items (importance). Items associated with an impact score ≥1,5 (which corresponds to a mean frequency of 50%) were retained for further analysis (Zamanzadeh et al., 2015).

$$Item\ Impact\ Score = \text{Frequency} \times \text{Importance}$$

Finally, the I-FVI index was applied to the proportion of raters giving an item agreement rating of 4 or 5 in the GAMEX and 0 or 1 in the SSQ. The cut-off was 0.79 according to Yusoff (2019b).

$$\text{I-FVI} = \frac{Raters\ in\ agreement}{Number\ of\ total\ raters}$$

Informed consent was obtained from parents and school administrators, after which the children were informed of the objectives of the study. The participants played in groups of 15 children each. Each child had their own virtual reality glasses during the trial period. Before the trials began, the experimenters provided instructions on how to play the game, and the children were seated in a chair fitted with the necessary devices, including glasses and joysticks, connected to the program. Each participant played for approximately 15 min per trial, and after completing both trials, the SSQ was administered to assess any potential side effects. The GAMEX Questionnaire, which was completed on paper to facilitate answers, was administered at the end of the second trial.

2.2.6 Step 6. Face validity and developmental evolution

The sixth phase focuses on the developmental evolution of children, considering their cognitive and psychological developmental stages. The sample was divided into two groups, 7–9 years old and 10–12 years old, according to developmental psychology (Staudinger et al., 2003). Thus, both groups were categorized as lay experts because of their experience in new technologies. In addition, 10 subgroups were created from the Age × Gender interaction.

This involves using well-established frameworks, such as Piaget's stages of cognitive development and the age groups proposed in developmental psychology (Staudinger et al., 2003). Feedback was gathered from the children after their interaction with the program to identify any issues and make necessary modifications. This comprehensive approach guarantees that the virtual reality program is effective, engaging, and tailored to the needs of the children.

2.3 Instruments

To determine the content validity (Step 4) of the BSR-RV 2® game scenes, a 9-item questionnaire, the Virtual Reality Content Validity Questionnaire, was created. This questionnaire measures both objective and subjective aspects of the game, such as gameplay, understanding of instructions, and usability of controls. The Cronbach's alpha for this scale was α = 0.727, indicating acceptable internal consistency. This questionnaire evaluates the domains explained in Step 1.

For both the first and second studies, the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993) was administered using Spanish validation (Campo-Prieto et al., 2021) to assess the potential side effects of the game. The SSQ consists of three subscales: oculomotor, disorientation, and nausea. In the content validity study, the overall alpha was α = 0.886, with subscale alphas ranging from α = 0.593 (nausea) to α = 0.864 (disorientation). In the face validity study, the overall alpha was α = 0.818, with subscale alphas between α = 0.641 (nausea) and α = 0.755 (oculomotor). This questionnaire was applied in Steps 4 and 5 of this study.

To measure the degree of acceptance, understanding, and familiarity of users with the virtual experience, the GAMEX Index (Gameful Experience in Gamification), developed and validated in English by Eppmann et al. (2018) and further validated in Spanish by Parra-González and Segura-Robles (2019), was used. The scale consists of 27 items, with response options ranging from "strongly disagree" (1) to "strongly agree" (5) on a Likert-type scale. The questionnaire included five subscales: enjoyment (items 1–6), absorption (items 7–12), intrinsic motivation (items 13–16 and 24–27), activation (items 17–20), and absence of negative affect (items 20–23). The answers for items in the "absence of negative affect" subscale were recoded, meaning that higher scores on this factor indicated a lower incidence of negative affect. The original alpha value for the scale was above α = 0.90, and in the Spanish version, the alpha was above α = 0.80. In the face validity study, the alpha was α = 0.875. This questionnaire was applied in Step 5 of this study.

2.4 Statistical analysis

Descriptive statistics, including mean and standard deviation, were calculated for variables such as participant age, gender, and level of expertise in both the content and face validity studies. The Kolmogorov-Smirnov test was used to assess the normality of the distributions. Once normality was established, tests for mean differences and ranges were conducted using the *t*-test and Mann-Whitney U test.

For the content validity study (Step 3), the Content Validity Ratio (CVR) and the Item-Level Content Validity Index (I-CVI) were computed using data from the Virtual Reality Content Validity Questionnaire and the Simulator Sickness Questionnaire (SSQ). Non-parametric tests, specifically the Mann-Whitney U test, were applied to analyze differences by level of expertise.

In the face validity study, both the impact score (Zamanzadeh et al., 2015) and the Item-Level Face Validity Index (I-FVI) (Yusoff, 2019b) were calculated using data from the GAMEX and SSQ. Parametric t-tests, including independent sample t-tests and paired sample t-tests, were used to examine differences across gender, age, Age × Gender interaction and number of trials. Chi-square tests were conducted to analyze the influence of age groups on face validity outcomes. The effect sizes were calculated using Cohen's d and Cramer's V was calculated to determine the effect size for chi-square tests.

All inferential analyses were conducted with the significance level set at $p < 0.05$. The statistical software SPSS version 27 was used for all analyses.

3 Results

The entire procedure for the study, comprising the development of the game and the subsequent application of both content and face validity studies, is illustrated in Figure 1 (Appendix). The outcomes of the six phases are as follows.

3.1 Step 1. Identification of the domain: literature review and domain-specific terms

Following a comprehensive literature review, we have selected the following domain-specific terms for further exploration: "attitudes towards disability and wheelchairs" (195 articles), "skill acquisition through virtual reality" (1169 articles), "content validity and face validity in virtual reality" (414 articles), "virtual simulation skill acquisition" (43 articles), and "advanced methodology in virtual reality and disability" (34 articles).

We have also reviewed 11 articles related to the constructs of "attitudes towards disability of people and wheelchairs" (Carter et al., 2014; Evans et al., 2015; Grenier et al., 2014; Lundberg et al., 2008; Matera et al., 2021; Medland and Ellis-Hill, 2008; Moss et al., 2020; Sapey et al., 2005; Sofokleous and Stylianou, 2023; Tamm and Prellwitz, 2001; Vilchinsky et al., 2010).

In addition, 11 studies on "skill acquisition through virtual reality" were reviewed (Alsalamah et al., 2017; Alvarez-Lopez et al., 2020; Babar et al., 2023; Bright et al., 2012; Chuan et al., 2023; Dorozhkin et al., 2017; Feeley et al., 2021; Harris et al., 2021; Runswick, 2023; Shen et al., 2022; Willaert et al., 2012).

We have also examined 15 articles on "content validity and face validity in virtual reality" (Carter et al., 2005; Dulan et al., 2012; Feeley et al., 2021; Harris et al., 2020; Hung et al., 2011; Kelly et al., 2012; Kenney et al., 2009; Marlenga et al., 2017; McDougall et al., 2006; Morgado et al., 2023; Rosenberg et al., 2013; Schwebel et al., 2008; Seixas-Mikelus et al., 2010; Verdaasdonk et al., 2006; Xiao et al., 2014). These articles discussed the perceptual limitations, sensory appearance, and degree of realism associated with virtual reality.

In conclusion, the literature on content and face validity in virtual reality in the educational and psychological fields is limited, although there are more references in the medical field.

3.2 Step 2: Generating virtual items

3.2.1 Focus group

In the second phase, a focus group consisting of an incidental sample was assembled. The participants had an average age of 39.25 years and had more than 12 years of experience in wheelchair basketball. Three focus group sessions were conducted with two designers of the simulation program. The objective was to apply a deductive-inductive method to analyze the results and conceptualize and generate items. The items were then organized and sequenced to provide the user with a coherent and suitable order throughout the simulation.

The focus group identified several themes: a) technical considerations, b) number and type of scenes of adapted wheelchair sport, c) sounds and images that evoke a sports environment, d) level of difficulty of the scenes, e) congruence with the laws of physics through chair and ball movements.

3.2.2 Development of the virtual environment of BSR-RV 2®

The virtual environment used in BSR-RV 2® for playing wheelchair basketball was created using Unity 3D, a versatile gaming development platform that is compatible with both PC and MAC systems. The game was validated using the Oculus Quest 2 device and the study was conducted in a school setting. The BSR-RV 2® VR game was developed using the Unity 3D graphics platform, which is widely utilized in the video game development industry and in various other applications such as simulations, augmented reality (AR), and virtual reality (VR). To collect user interaction statistics for the application, the API Rest was developed using Spring Boot and MongoDB. Database creation was integral in the deployment of the service on "railway.app" because MongoDB databases do not require a relational design, unlike other SQL alternatives.

The incorporation of co-routines into Unity is facilitated by the use of C# and the "IEnumerator" return type, which is managed by the Unity engine. This approach enables the acquisition of reaction times at each level, as it waits for the user to initiate the first action of the level and records the time taken to complete the action during the game session. The server responsible for managing the data input from the BSR-RV 2® application and the database where the records are stored are both hosted on a web service called "Railway." The scores for each scene were uploaded to the database and assigned to an identifier in the questionnaire.

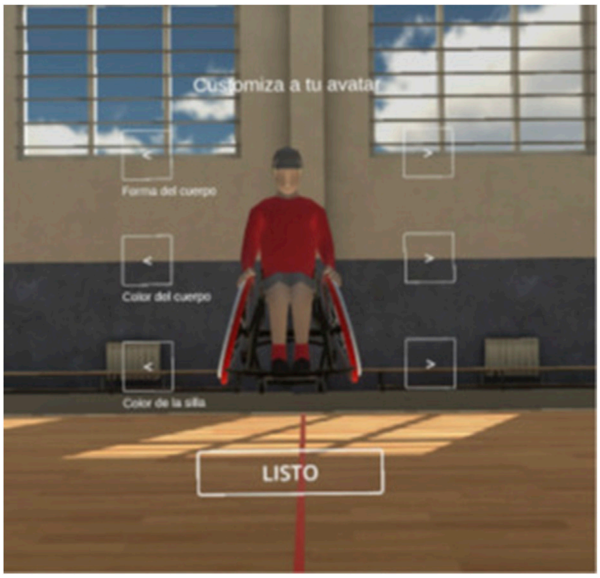

FIGURE 2
Avatar customization.

FIGURE 4
Hit the goal targets.

FIGURE 3
Wheelchair Zig-Zag movement.

FIGURE 5
Shooting to the basket.

The following text provides an overview of the components required to facilitate the application of the BSR-RV 2®. It is essential to develop a series of components to create realistic and enjoyable experience. The 3D models were created using the Blender software. Four distinct models were developed to design the wheelchair, each representing a chair used by an individual with a different level of mobility.

3.2.3 Scenes

Initially, the user has the flexibility to personalize the avatar in the wheelchair by modifying gender, body shape, and other visual attributes (Figure 2). Following personalization, the user commences by reading instructions on the usage of the controls necessary for maneuvering the chair and practicing the skills necessary to navigate the game (Figure 3). Subsequently, the user encounters a goal and attempts to score it by hitting the targets with a basketball (Figure 4). The subsequent scene entails practicing shooting, in which the user's distance from the basket is adjusted (Figure 5). Finally, the user experienced a realistic basketball game setting to simulate a genuine match experience (Figure 6). Upon the completion of these actions, the user is presented with their final score and thanked for participating in the game. In the creation of these scenes, the program adhered to the criteria specified in the Focus Group and addressed the following domains:

a) Technical aspects to consider include the realism of objects, images, and movements, as well as the degree of negative consequences, level of embodiment, and clarity of basketball court visuals (Figure 2). In addition, it is important to consider

FIGURE 6
Wheelchair movement through rivals.

the sharpness of the images, such as the ball, wheelchair, basket, and avatar.

b) The number and type of adapted wheelchair sports scenes presented primarily focused on wheelchair basketball exercises, and the specific order in which these scenes were presented to the user was not clearly defined. The user was able to experience a sense of being in a wheelchair during these scenes, which showcased a variety of wheelchair movements, including passing and basket throwing (Figures 3, 4).
c) Sounds and images that elicit a sports environment: the sound of the audience (whether low, intermediate, or high), the sound of the ball bouncing on the floor, the type of soundtrack for the simulation, the realism of the basketball court, the realism of the avatars (in terms of color, shape, and movement), and congruence with the real wheelchair basketball visual perspective.
d) The degree of difficulty for each scene was determined by classifying the scenes into two levels. This classification considers the number of instructions per scene, clarity of the instructions, and their position in the user's visual field (as shown in Figure 5).
e) Adherence to the laws of physics can be demonstrated through realistic movements of the chair and ball, including friction, displacement on the court, and players' ability to catch the ball through simulation. Furthermore, the realism of the ball's angle when thrown towards the basket (as depicted in Figure 6) must be considered.

3.3 Step 3. Determination of raters

A table of critical CVR values computed by Lawshe (1975) was adopted, in which it was assumed that the level of agreement exceeded that of chance, with values ranging from 0 to 1. A critical CVR of 0.42 is assumed for a value of 20 experts (Lawshe, 1975; Romero Jeldres et al., 2023). A critical CVR value of 0.42 is assumed for 20 experts (Lawshe, 1975; Romero Jeldres et al., 2023).

In the face validity study (Step 5), 395 children (lay experts) were recruited to evaluate the items. The formula proposed by Yusoff (2019b) was applied, with a criterion of 0.79 as a minimum value to maintain the item. Subsequently, the results were compared with those of other randomly selected small samples (N = 30, N = 60, N = 90, N = 300 and N = 360) to determine if sample size affected I-FVI regarding the Age × Gender interaction across the GAMEX questionnaire according the criteria of the original authors (Eppmann et al., 2018).

3.4 Step 4: Content validity study

Before establishing the Content Validity Ratio (CVR), the impact of the level of expertise (lay experts *versus* content experts), gender, and age was evaluated. The responses from the content and lay experts groups were compared, and no differences were observed in the "Virtual Reality Content Validity Questionnaire."

The responses between the content and lay experts groups were analyzed comparatively and no differences were found in the nine variables except for the item "the objects in the game seemed real to me," which had a better rate of the content experts *versus* the lay experts ($U = 19{,}5$, $Z = -2.193$, $p = 0.028$).

In relation to side effects, no relationship was observed between the level of expertise and side effects, except for item 7 "*sweating*" ($U = 28.00$, $Z = -2.21$, $p = 0.027$) which affects the content expert group. Age correlated with item 10 "*headache*" ($Rho = 0.46$, $p = 0.041$). Finally, female gender correlates with item 7 "*sweating*" ($U = 16.00$, $Z = -2.902$, $p = 0.004$).

Items 1, 2, 5, 6, 7, and 8 had CVR values greater than 0.42 (Lawshe, 1975; Romero Jeldres et al., 2023) and I-CVI values higher than 0.79 (Yusoff, 2019a). The primary outcomes of this study are presented in Tables 2, 3, respectively.

The decision was made to evaluate the items that impacted the "Virtual Reality Content Validity Questionnaire" scale, which had a validity of less than 0.42 in the CVR values (Lawshe, 1975; Romero Jeldres et al., 2023) and less than 0.79 in the I-CVI values (Yusoff, 2019a). Consequently, modifications were implemented to the sound and music of the various scenes (volume of the soundtrack, type of music presented), as well as the instructions (visual position and information displayed to the user, type of instructions given to the user), adaptations related to the operation of the controllers to improve the users' intuitiveness of the simulation (ease of moving the wheelchair, the way the user could take and pass the ball, ease of throwing the ball into the basket, and natural movement of the controls to enhance the user's understanding of the simulation's functioning) (Figures 3, 4).

3.5 Step 5. Face validity study

In step five, a face validity study was conducted to evaluate the Validity Index using 395 lay experts. The sample size was reduced by 17 individuals in the GAMEX study due to unforeseen circumstances beyond the control of the researchers. The results of the study indicate that the GAMEX scale exhibits strong face validity in terms of enjoyment and absence of negative affect, but lower validity in factors such as absorption, intrinsic motivation, and

TABLE 2 Virtual Reality Content Validity Questionnaire in BSR-RV 2®.

	Fr (N= 20)	CVR	I-CVI
1. The scenes of the game seemed real to me.	16	0.6	0.8
2. The objects in the game seemed real to me.	17	0.7	0.85
3. I found the sound effects of the game to be entertaining.	3	−0.7*	0.15*
4. I liked the game's music.	3	−0.7*	0.15*
5. I found the colors of the game attractive.	18	0.8	0.9
6. I have been able to see the objects and instructions inside the game.	14	0.4	0.7*
7. I found the instructions of the game to be clear.	20	1	1
8. I felt excited while playing.	17	0.7	0.85
9. I found it easy to understand how the controls work	12	0.2*	0.6*

*CVR Values <0.42; I-CVI Values <.79.

TABLE 3 BSR-RV 2® side effects through SSQ in the pilot study.

	NSlight/ Absent	NTotal	CVR	I-CVI
General discomfort	20	20	1	1
Fatigue	16	20	0,6	0,8
Headache	19	20	0,9	0,95
Eyestrain	18	20	0,8	0,9
Difficulty focusing	18	20	0,8	0,9
Increased salivation	20	20	1	1
Sweating	19	20	0,9	0,95
Nausea	19	20	0,9	0,95
Difficulty concentrating	19	20	0,9	0,95
Fullness of head	19	20	0,9	0,95
Blurred vision	19	20	0,9	0,95
Dizzy (eyes open)	19	20	0,9	0,95
Dizzy (eyes closed)	19	20	0,9	0,95
Vertigo	20	20	1	1
Stomach awareness	19	20	0,9	0,95
Burping	20	20	1	1

negative activation (Table 4). The measurement of activation presents a methodological issue as it contains contradictory meanings in its items. The two positive activation items ("While playing I felt activated" and "While playing I felt excited") and the two negative activation items ("While playing I felt jittery," and "While playing I felt frenzied") are problematic. If the I-FVI criterion proposed by Yusoff (2019b) is applied, the absorption, intrinsic motivation, and negative activation factors would not be considered adequate. However, if the Impact Score criterion is used, only the negative activation factor is considered inadequate. The low levels of agreement in negative activation observed are a positive finding, as they suggest low levels of jittery and frenzy.

The relationship between age groups and face validity assessed through the GAMEX was examined, revealing a correlation between age and absorption, intrinsic motivation (Items 13 and 24), and negative activation (Item 18) (Table 4).

3.6 Step 6. Face validity and evolutionary development

A subsequent within-group analysis examining the Age × Gender interaction found variations within the 7–9 and 10–12 age groups. However, due to the small sample size of 12 years (N = 7), the 12-year-old group was excluded from the analysis (Figure 7). The chi-square analysis of Age × Gender interaction revealed that the highest degree of agreement in absorption occurred in the 8-year-old male group within the 7–9 years age group, and in the 11-year-old female group within the 10–12 years age group for item 7 ($\chi 2$ (36) = 53,641; V = 0.186, $p = 0.030$), item 8 ($\chi 2$ (36) = 55,768; V = 0.194, $p = 0.019$), Item 9 ($\chi 2$ (36) = 54,229; V = 0.192, $p = 0.026$), item 10 ($\chi 2$ (36) = 66,673; V = 0.212, $p = 0.001$), item 11 ($\chi 2$ (36) = 57,102; V = 0.197, $p = 0.014$), item 12 ($\chi 2$ (36) = 57,301; V = 0.197, $p = 0.013$), and item 24 ($\chi 2$ (36) = 56,417; V = 0.195, $p = 0.016$).

According to the criteria established by GAMEX Eppmann et al. (2018), we assessed the face validity based on the sample size. In our sample of 376 cases, we selected 30, 60, 90, 300 and 360 cases to determine if sample size affected I-FVI regarding the Age × Gender interaction across the GAMEX questionnaire (Table 5), with the 8 years old male age group showing the most notable differences in absorption.

For the purpose of analyzing the SSQ questionnaire, the impact score and the I-FVI were assessed using the criteria of Zamanzadeh et al. (2015) and Yusoff (2019b), respectively. In each instance, the scores were close to the minimum range, as indicated in Table 6.

The extent of agreement regarding side effects was assessed during the two trials. A *t*-test for paired samples indicated variations between the first and second trial in item 3 "headache" (t (393) = −2.73; $p = 0.007$, d = 0.812; 0.43 *versus* 0.38), item 5 "difficulty focusing" (t (388) = 3.509; $p < 0.001$, d = 0.881; 0.53 *versus* 0.38), item 9 "difficulty concentrating" (t (392) = 4.938; $p < 0.001$, d = 0.919; 0.55 *versus* 0.32), and item 11 "blurred vision" (t (392) = 2.743; $p = 0.006$, d = 0.846; 0.52 *versus* 0.40).

TABLE 4 Perception of the quality of the BSR-RV 2® tool through GAMEX items and factors.

Items	Fr	I-FVI /Item	Factor	I-FVI/Factor	Impact Score	χ2 (%) 7-9 vs 10-12	Cramer's V	p
1. Playing the game was fun.	358/378	0,95	Enjoyment	0,91	4,22			
2. I liked playing the game	363/378	0,96						
3. I enjoyed playing the game very much	350/378	0,93						
4. My game experience was pleasurable.	335/374	0,90						
5. I think playing the game is very entertaining	355/377	0,94						
6. I would play this game for its own sake, not only when being asked to	304/376	0,81	Absorption	0,49*	1,59	36,9% vs 35,1%	0,225	0,017
7. Playing the game made me forget where I am.	156/375	0,41						
8. I forgot about my immediate surroundings while I played the game	173/376	0,46						
9. After playing the game, I felt like coming back to the "real world" after a journey	235/376	0,62				69,1% vs 55,9%	0,202	0,045
10. Playing the game "got me away from it all."	155/377	0,41				45,4% vs 35,1%	0,237	0,009
11. While playing the game I was completely oblivious to everything around me	177/376	0,47						
12. While playing the game I lost track of time	228/378	0,60						
13. Playing the game sparked my imagination.	234/376	0,62	Intrinsic motivation	0,53*	1,83	39,9% vs 28,6%	0,209	0,034
14. While playing the game I felt creative.	244/378	0,64						
15. While playing the game I felt that I could explore things.	244/376	0,66						
16. While playing the game I felt adventurous	220/375	0,58						
24. While playing the game I felt dominant/I had the feeling of being in charge.	137/378	0,36						
25. While playing the game I felt influential	87/378	0,23						
26. While playing the game I felt autonomous	170/378	0,44						
27. While playing the game I felt confident	218/378	0,77						

(Continued on following page)

TABLE 4 (*Continued*) Perception of the quality of the BSR-RV 2® tool through GAMEX items and factors.

Items	Fr	I-FVI /Item	Factor	I-FVI/Factor	Impact Score	χ2 (%) 7-9 vs 10-12	Cramer's V	p
17. While playing the game I felt activated.	291/377	0,77	Activation (positive)	0,76*	3,12	75.3% vs 68,8%	0,200	0,049
20. While playing the game I felt excited	284/376	0,75						
18. While playing the game I felt jittery	100/370	0,27	Activation (negative)	0,30**	0,77*			
19. While playing the game I felt frenzied.	129/374	0,34						
21. While playing the game I felt upset.	309/377	0,82	Absence of negative affect	0,8	2,67			
22. While playing the game I felt hostile	297/378	0,78						
23. While playing the game I felt frustrated	309/378	0,81						

* I-FVI Values <0.79; ** Low levels of agreement in negative activation reveal low degree of symptoms.

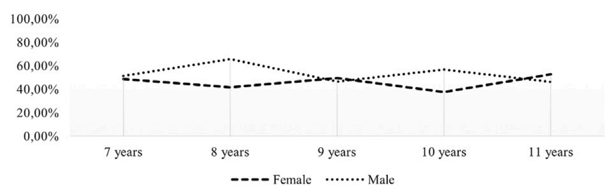

FIGURE 7
Intragroup variance on the absorption factor.

Table 7 displays the differences between trials and age groups (7–9 years old *versus* 10–12 years old) in trials 1 and 2. Although the differences were statistically significant, the level of symptomatology in both the trials was relatively low.

Analysis of Age × Gender interaction revealed significant differences in fatigue (χ2 (27) = 41,763; V = 0.189, p = 0.035) in the first trial, as well as in eyestrain (χ2 (27) = 45,458; V = 0.199, p = 0.015), difficulty in focusing (χ2 (27) = 40,283; V = 0.186, p = 0.048), increased salivation (χ2 (27) = 44,385; V = 0.195, p = 0.019), and dizziness with eyes open (χ2 (27) = 46,785; V = 0.200, p = 0.010) in the second trial. The differences were particularly pronounced among individuals aged 7 years, as shown in Figure 8. However, it should be noted that the values for all other Age × Gender interactions were very close to zero, indicating a negligible effect.

4 Discussion

The main objective of this study was to develop and validate an adapted basketball virtual reality program aimed at promoting the inclusion of children with motor disabilities who use wheelchairs. The novelty and truly pioneering aspect of this study lies in the validation of a tool exclusively from the perspective of children. Furthermore, the creation of a virtual reality game aimed at enhancing inclusivity for children aged 7 to 12 presents an additional challenge, as, to our knowledge, no other virtual reality games have been validated in children with this purpose. Additionally, this study has emphasized the necessity of conducting validity analyses aligned with age groups, demonstrating variability in results based on age groups and the interaction Age × Gender.

Another primary objective of this study was to assess the content and face validity of the BSR-RV 2® virtual reality game in children aged 11–18 years (content validity study) and 7–12 years (face validity study) divided into two age groups (7–9 years and 10–12 years). In addition, the content validity study included an analysis of the perceptions of content and lay experts. The content validity of the game was evaluated based on the model proposed by Harris et al. (2020), which considers the properties of immersion, presence, psychological fidelity, and physical fidelity. The face validity study primarily focused on the properties of presence and psychological fidelity (attention and affective states). Analysis of the five phases proposed by various authors (Lynn, 1985; Yusoff, 2019a; 2019b; Zamanzadeh et al., 2015) was conducted as follows: 1) identification of the domain, 2) generation of virtual items, 3) determination of the number of judges, 4) establishment of content validity, and 5) establishment of face validity. As a novelty and given that the study involved a child population, a sixth stage was added: analyzing the role of developmental evolution in the judgments of the properties.

4.1 Step 1. Identification of the domain

In our examination of the domains involved in constructing and validating virtual environments, we did not find any literature analyzing the content and face validity of virtual tools designed to change attitudes towards individuals with disabilities among typically developing children. However, several studies have focused on changing attitudes towards people with disabilities through virtual reality simulations (Ceberio, 2022; Chowdhury et al., 2017; Chowdhury and Quarles, 2022; Matera et al., 2021;

TABLE 5 Chi-square analysis depending on the Age x Gender interaction in different sample sizes.

	N = 30			N = 60			N = 90			N = 300				N = 360			
	χ2	p	V	χ2	p	V	χ2	p	V	χ2	p	V	%	χ2	p	V	%
1. Playing the game was fun.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2. I liked playing the game	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3. I enjoyed playing the game very much	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4. My game experience was pleasurable.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5. I think playing the game is very entertaining	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6. I would play this game for its own sake, not only when being asked to.	-	-	-	-	-	-	-	-	-	55,35	0,021	0,226	91,1% (8 male)	52,57	0,037	0,203	90,7% (8 male)
7. Playing the game made me forget where I am.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8. I forgot about my immediate surroundings while I played the game.	-	-	-	-	-	-	-	-	-	56,51	0,016	0,228	62,3% (8 male)	51,46	0,046	0,201	55,5% (7 female)
9. After playing the game, I felt like coming back to the "real world" after a journey.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10. Playing the game "got me away from it all."	-	-	-	-	-	-	-	-	-	60,98	0,006	0,237	75% (7 female)	60,45	0,007	0,217	61,3% (8 male)
11. While playing the game I was completely oblivious to everything around me	-	-	-	-	-	-	-	-	-	52,23	0,039	0,22	100% (7 female)	58,75	0,01	0,214	75,7% (10 male)
12. While playing the game I lost track of time	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13. Playing the game sparked my imagination.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
14. While playing the game I felt creative.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15. While playing the game I felt that I could explore things.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16. While playing the game I felt adventurous	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24. While playing the game I felt dominant/I had the feeling of being in charge	-	-	-	-	-	-	-	-	-	55,62	0,019	0,226	66,7% (7 male)	-	-	-	-
25. While playing the game I felt influential	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
26. While playing the game I felt autonomous	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
27. While playing the game I felt confident	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
17. While playing the game I felt activated.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
20. While playing the game I felt excited	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
18. While playing the game I felt jittery	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
19. While playing the game I felt frenzied.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

(Continued on following page)

TABLE 5 (*Continued*) Chi-square analysis depending on the Age x Gender interaction in different sample sizes.

	N = 30			N = 60			N = 90			N = 300				N = 360			
	χ2	p	V	χ2	p	V	χ2	p	V	χ2	p	V	%	χ2	p	V	%
21. While playing the game I felt upset.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22. While playing the game I felt hostile	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
23. While playing the game I felt frustrated	-	-	-	-	-	-	-	-	-	56,312	0,017	0,227	26,7% (8 male)	-	-	-	-

- = no statistically significant differences; V = Cramer's V.

TABLE 6 Perception of the side effects of the BSR-RV 2® tool through SSQ items and factors.

	Trial 1		Trial 2			
	NSlight/Absent	NTotal	NSlight/Absent	NTotal	I-FVI (Mean)	Impact score (Mean)
General discomfort	388	395	377	395	0,97	2,71
Fatigue	365	395	349	395	0,90	2,31
Headache	368	395	353	395	0,91	2,40
Eyestrain	362	395	361	395	0,92	2,35
Difficulty focusing	348	395	358	395	0,89	2,28
Increased salivation	372	395	384	395	0,96	2,66
Sweating	358	395	364	395	0,91	2,39
Nausea	387	395	387	395	0,98	2,83
Difficulty concentrating	342	395	368	395	0,90	2,30
Fullness of head	344	395	347	395	0,87	2,17
Blurred vision	356	395	368	395	0,92	2,31
Dizzy (eyes open)	379	395	376	395	0,96	2,68
Dizzy (eyes closed)	384	395	385	395	0,97	2,80
Vertigo	383	395	384	395	0,97	2,80
Stomach awareness	383	395	380	395	0,97	2,72
Burping	390	395	393	395	0,99	2,94

Pivik et al., 2002), but they were not designed, created and validated through the children perspective.

Regarding the simulation of sports practice, a significant challenge is the scarcity of studies assessing the accuracy of children's motor and movement skills. Most studies relate to pictorial images and do not focus on virtual reality (Lopes et al., 2016; Morgado et al., 2023; Valentini et al., 2018). The primary objective of the present study was to identify the different types of mastery. Those related to physical fidelity involve reproducing movements and motor skills related to chair movement, passing the ball between seated players, or changing the visual perspective when throwing the ball to the basket. The realism domain involves reproducing a game with two or more players. The goal is for a typically developing child to discover the limitations or barriers experienced by a disabled child when playing a wheelchair basketball game. In doing so, the child may change his or her perspective and promote inclusion in the future.

While the difficulty level of the motor and coordination activities in the different scenes of the study was inspired by general recommendations from medical research, it was based on several factors, including the sports hall environment (Kelly et al., 2022; Runswick, 2023), handling of tractors by children (Marlenga et al., 2017), and objectives and development of the present study, which appear to be novel. The wide range of elements considered in the study, such as familiarization with the movement of the wheelchair; the performance of two or more simultaneous tasks; and aesthetic aspects, sounds, images, or degree of realism adapted to the age of the participants, make it a unique contribution to the field. The target group consisted of children aged 8–12 years, and while there were few studies that analyzed domains adapted to developmental development, most notably, the works of Marlenga et al. (2017) and Schwebel et al. (2008), the age groups did not follow the developmental criteria proposed in the present study.

TABLE 7 Chi-square analysis between trials and age groups (7–9 years and 10–12 years).

		X^2 (%) 7-9 vs 10-12	Cramer's V	p
Trial 1	Eyestrain	92,7% vs 91,3%	0.178	0.006
	Difficulty focusing	89,8 % vs 88,7%	0.153	0.027
	Difficulty concentrating	81,4 vs 91,2%	0.148	0.035
	Vertigo	97,7% vs 96,7%	0.144	0.042
Trial 2	Eyestrain	88% vs 95,8%	0.179	0.006
	Blurred vision	90,4% vs 95,8%	0.151	0.030
	Dizziness with the eyes closed	94,9% vs 99,5%	0.159	0.019

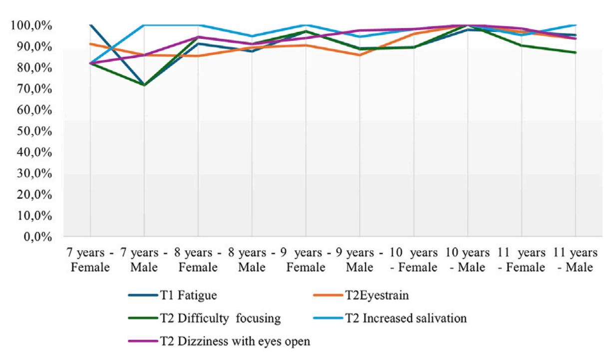

FIGURE 8
Differences on Trial 1 and Trial 2 on AgexGender interaction analysis on the incidence of side effects.

4.2 Step 2. Generation of Virtual Items

The development of virtual items in the present study involved a focus group consisting of adult players with over 12 years of experience in adapted basketball. Four participants agreed that the main thematic areas should focus on two domains: technical considerations and degree of difficulty and congruence with physical laws.

To the best of our knowledge, this is a pioneering effort to promote inclusion in children aged 8–12 years through an adapted basketball virtual reality game. The game features three main scenes, each comprising two subscenes with increasing difficulty levels. In the "Final Game," users put into practice all they have learned previously. In addition, each user was provided with various options for wheelchair avatars at the beginning of the simulation to identify someone who used a wheelchair.

The primary focal point of the initial scene is the manipulation of the wheelchair, in which the user must traverse various obstacles while navigating the wheelchair towards a designated location. The subsequent main scene entails a passing drill, where the user acquires proficiency in catching the ball and simulating a pass to a teammate in both static and dynamic settings. Third, in the basket-shooting scene, the user must attempt shots at the basket from diverse angles and ranges, requiring the user to approach closer or farther, based on the difficulty presented. Lastly, all the learned skills are integrated in the "Final Game" scene, where the user must apply the knowledge acquired in the three main scenes previously described in a simulated game scenario.

4.3 Step 3. Identifying the number and type of raters

The current study utilized children as experts for both content and face validity assessments. For content validity, their understanding of basketball and wheelchair basketball, as well as their familiarity with new technologies, was considered. In the face validity study, their expertise in new technologies was evaluated, considering their experience playing video games (VanDeventer and White, 2002) and the daily usage of these technologies (Hague and Payton, 2010; Plowman et al., 2012).

In both content validity and face validity studies, the number of participating children exceeded the number of experts usually recommended in studies with adults (Lynn, 1985). The value for content validity is of 20 experts, which suggests that the random inter-judge agreement could decrease. In our study and many other studies, the arbitrary limit proposed by Lynn (1985) has surpassed, with as many as 30 experts involved in some cases (Handage and Chander, 2021; Yadav et al., 2021). In the case of the content validity study conducted with children aged between 11 and 18 years, we agree with the proposal by Stewart et al. (2005), which suggests that the panel of experts should be expanded to avoid linguistic, conceptual, and contextual comprehension issues. Furthermore, the level of attention or concentration is influenced by the maturity of neural circuits that determine the level of attention.

The same holds true for the face validity study, which included 395 children aged between 7 and 12 years. The enlargement of the sample size was consistent with the recommendations of other authors (Stewart et al., 2005). To ensure the representativeness of the sample, it is crucial to include individuals from the target age group and gender. As cognitive and affective development can vary significantly between males and females, it is essential to account for these differences. The 395 participants were divided into two main groups: those aged 7–9 years and those aged 10–12 years. The sample was further divided into 10 subgroups, each containing no more than 60 cases, based on the Age × Gender interaction.

In contrast to studies that examine age and gender variables independently in the context of virtual reality for children (Bart et al., 2008; Marlenga et al., 2017; Schwebel et al., 2008), our study considers the Age × Gender interaction effect. This approach enabled us to examine the fluctuations in the critical values of I-FVI that depend on both variables more precisely.

4.4 Step 4. Establishing content validity

The Virtual Reality Content Validity Questionnaire and the SSQ were employed to analyze the level of agreement regarding immersion, presence, physical and psychological fidelity, degree of realism, and the side effects associated with the use of the program BSR-RV 2®. Notably, no significant differences were observed between the critical values of CVR, unless in the grade of realism, in which differences were observed between content experts (wheelchair basketball players) and lay experts (conventional basketball players) aligning with studies with experts and novices in the area of creativity (Kaufman et al., 2008). The findings of our study indicate that there are some variations in the scores obtained by the two groups in the Virtual Reality Content Validity Questionnaire, which we attribute to their level of familiarity with basketball. It is possible that some of the tasks in the program, such as chair movement, ball passing, and basket shooting, were not entirely new to some participants.

Regarding the SSQ, the CVR data showed a low incidence of cybersickness, with a high agreement on the low incidence of symptoms. The results revealed a virtual absence of general discomfort, salivation, vertigo, and burping, along with a low incidence of headache, difficulty in focusing, or eyestrain. Furthermore, the low intergroup differences in the SSQ may be explained by the results of Huppert et al. (2019), who reported that the degree of cybersickness is similar or decreases in the age group–12–16 years and older.

Overall, the CVR values were found to be lower than 0.42 in some instances, which necessitated the adjustment of the sound and music in the different scenes, as well as the instructions and adaptations related to the operation of the controllers to enhance the intuitiveness of the simulation. This is consistent with I-CVI values, although the magnitude of the critical values may be misleading in certain cases.

4.5 Step 5. Establishing face validity

Step 5 involved conducting a face validity study that involved selecting 395 children aged 7–12 years as lay experts to assess the test (Nevo, 1985). The overall face validity indices, including the presence and psychological fidelity indices, varied. The critical values for the impact score were high for enjoyment and positive activation, whereas the levels of intrinsic motivation and absorption were moderate. However, the negative activation critical value was low, since most participants did not show this type of reaction.

The impact of the Virtual Reality program's duration and number of trials on non-aesthetic domains, such as absorption and intrinsic motivation is likely to be reflected in the results obtained. To increase absorption, it is necessary to have longer sessions. The same can be said for GAMEX's measure of intrinsic motivation, in which lower values were associated with achieving a higher degree of autonomy and sense of control. Given that the BSR-RV 2® program included scenes of sporting activities with varying levels of difficulty, such as shooting a basket or moving a chair, a greater number of trials and longer duration would have been necessary to achieve a higher degree of control.

Although both the I-FVI and Impact Score indices are presented in this study, there is currently no consensus on the ideal index for assessing the face validity of a virtual tool, as there are studies that use the I-FVI proposed by Yusoff (2019b) as a reference (Alvarez-Lopez et al., 2020; Baharuddin et al., 2024; Mohamad Marzuki et al., 2018), whereas others use the impact score proposed by other authors (Abdollahipour et al., 2016; Bek et al., 2009; DunnGalvin et al., 2008; Lacasse et al., 2002; Zamanzadeh et al., 2015). However, other studies have established the mean between groups of experts to certify their validity (Alsalamah et al., 2017; Alzahrani et al., 2013; Chuan et al., 2023; Evans et al., 2015). Given this lack of consensus, the present study has decided to show both values, despite the fact that these values have been determined with a lower sample size in previous studies (Yusoff, 2019b; Zamanzadeh et al., 2015). Even so, the vast majority of the values presented in the face validity study were still acceptable. Other studies, however, have simply established the mean between groups of experts to certify their validity (Alsalamah et al., 2017; Alzahrani et al., 2013; Chuan et al., 2023; Evans et al., 2015).

Regarding the cybersickness scale SSQ, there were very few differences between T1 and T2 in terms of the symptoms, except for headache, difficulty focusing, difficulty concentrating, and blurred vision. These symptoms showed a significant decrease within the absence and mild range. However, the values obtained on the SSQ in both the I-FVI and impact score were higher than acceptable in all items. It is important to note that the neurovegetative symptomatology was practically absent in all cases.

The issue with this study lies in the analysis of the results based on the I-FVI on the GAMEX scale, which yielded values below the accepted threshold of 0.79, as proposed by Yusoff (2019b). Upon examining the studies that contributed to this figure, it is evident that the majority of them involved adult participants. The situation changes when analyzing studies involving children of a similar age and subject matter to our own study. In research on the validity of non-virtual pictorial instruments for games or sports involving children, the face validity scores are typically lower than 0.79. This is true in studies assessing sports physical competencies with samples larger than 200 cases (Lopes et al., 2016), as well as in studies on performance in aquatic physical activities (Morgado et al., 2023), movement actions (Moulton et al., 2019), and serious games (Sim et al., 2015). The same trend is observed in virtual reality studies with scores below 4 on a 5-point Likert scale, such as the study by Marlenga et al. (2017), Schwebel et al. (2008), and Veerman et al. (2024).

In certain studies, when the values were lower than anticipated (I-FVI <0.80; Kappa <0.70), some items that violated these principles had to be revised or retained (Schilling et al., 2007). This discussion has also been raised in other academic disciplines such as marketing, where the percentage of experts required to agree on retaining an item can vary significantly, ranging from 60% to 80% (Hardesty and Bearden, 2004).

To determine whether the outcomes were influenced by the sample size and to adhere to the criteria suggested for evaluating the construct validity of the GAMEX (Eppmann et al., 2018). The results reveal that the use of small sample sizes (N = 30; N = 60; N = 90) do not detect some differences in the degree of agreement, especially on cognitive variables (e.g., absorption) when considering the Age × Gender interaction. However, this issue is resolved with larger sample sizes (N = 300; N = 360). This result is crucial as it calls into question the traditional reliance on non-parametric samples to determine face validity with children. Therefore, the greater intra- and intergroup variability in their results suggests that enlarging the sample size is necessary to avoid inconsistencies and identify the developmental subgroups.

The validity scores obtained from the group of child experts were likely influenced by the unique psychological processes associated with each developmental stage. As a result, it was deemed necessary to examine the values derived from the SSQ and GAMEX using methods that focused on both inter- and intra-group differences.

4.6 Step 6: Face validity and evolutionary development

When we examined the critical values (I-FVI and Impact Score) in relative terms, we observed variability in the age-gender interaction, particularly in absorption and negative activation. The comparison between the 7–9- and 10–12-year-old groups revealed slight inter-group and intra-group discrepancies. Overall, the 7–9-year-old group exhibited a higher degree of absorption (items 8–12), sense of dominance (Item 24), and negative activation (Items 18 and 19). The level of agreement was higher in 7-year-old females and 7- and 8-year-old males in terms of absorption, reaching close to 73%. This is significant because absorption is linked to the level of tonic activation and attention. Children in this age group displayed greater engrossment or preoccupation with the objects and images in the BSR-RV 2® program, often disregarding other aspects of their surroundings ("Playing the game made me forget where I am") and, in some cases, exhibiting dissociative signs.

Dissociation that is not pathological can result in experiences where children become fully engaged and focused, leading to a temporary loss of awareness and decreased monitoring of internal emotions and external activities (Eisen and Lynn, 2001). Children who score higher in this area (between 7–8 years old) exhibit characteristics of the pre-operational stage of cognitive development, a period marked by egocentrism that can make it challenging to distinguish between one's own thoughts and external reality (e.g., images and movies). Our findings align with those of Beatini et al. (2024), which suggest that perspective-taking and changes in egocentrism occur around the age of nine. As such, we propose that absorption may represent typical manifestations of infantile cognitive structures and a normative regulatory strategy, which may be influenced by neurobiological development in 7–9 to year-olds. For instance, in a study using virtual reality by Baumgartner et al. (2008), prefrontal brain activation was less pronounced in the 8–11 age group.

Few studies have examined the impact of virtual reality on child development from a comparative perspective. Our research aligns with the findings of Baughan et al. (2024), which link high absorption in VR to a lack of attention to internal sensations and external interruptions. However, they did not explore the relationship between absorption and age. In contrast, Nakayama et al. (2020) observed that children aged 5–8 years were more susceptible to problems related to video games than those aged 10 years.

Two sessions were conducted to measure cybersickness using the SSQ, separated by a 1-week interval. This approach was used to control for the effects of phasic activity or fluctuations over time that may occur spontaneously or in response to an event. These fluctuations are changes in the autonomic nervous system that were assessed through symptoms associated with cybersickness and negative emotions, such as "jittery" and "frenzied." The agreement on negative activation in the GAMEX questionnaire was higher for both females and males at ages 7 and 8 years than for the 9–11 age group. Significant intra-group variations were observed.

The findings from the SSQ displayed few inconsistencies between the two trials (trial 1 and trial 2) for the age groups of 7–9-year-olds and 10–12-year-olds. However, an analysis of the effects of age, gender, and trial revealed minor differences in the level of agreement for symptoms such as eyestrain, difficulty focusing and concentrating, vertigo, and blurred vision. Children aged 7–8 years had a higher incidence of these symptoms. The comparative analysis by Age × Gender interaction showed irregular results, demonstrating a significant decrease in agreement from the age of 9 years onwards, both in males and females. Our results align with those of other studies (Baumgartner et al., 2006; Huppert et al., 2019; Nakayama et al., 2020), which found that the age group with the highest incidence of side effects was 8 years, corroborating the data presented in our study. The reduced awareness of bodily signals in the age group of 9–11 years is similar to the results obtained in studies with adults who played videogames, which showed various forms of reduced awareness of bodily signals (Swinkels et al., 2021), and with children over 10 years old (Nakayama et al., 2020).

Generally, the findings do not support the developmental range suggested by Staudinger et al. (2003) for 7–9-year-olds and 10–12-year-olds, as the face validity values indicate greater homogeneity between the two age groups. Specifically, the highest values of I-FVI in phasic activation (cybersickness) and tonic activation (degree of absorption) suggest a critical moment at age 7–8 in males, whereas a similar trend is observed in females of the same age range, although at a lower intensity. These findings are consistent with those of studies that examine the impact of gender on video game use (Martucci et al., 2023).

Depending on age and gender, psychological transitions from sensation to perception, perception to thought, and thought to action exhibit variations. Physical sensations generally do not interfere with the children's evaluation of the scenes, except for a small subgroup of 7–8-year-old males, possibly linked to the low incidence of adverse effects (e.g., nausea, dizziness). Additionally, the limited concern associated with these effects may have contributed to improved attention levels, particularly in 10–11-year-old males. The children's identification with the scenes appears to have been enhanced by their enjoyment of the immersive environment, with a stronger relationship observed in 10–11-year-old males. This, in turn, may have positively influenced higher levels of absorption and intrinsic motivation among the children.

5 Implications

The present study's methodology offers new information on the intra- and inter-group analysis of the sample, making it one of the first studies to date to examine both Age × Gender interaction in the developmental trajectory of children aged 7–12 years.

By involving children in the content validity process and the design of the virtual tool, their opinions on a virtual reality game targeting their age group were considered. Additionally, as VR has proven effective in fostering perspective-taking and empathy, games such as BSR-RV 2®, which aim to improve attitudes toward people with disabilities, play a crucial role in promoting inclusivity from a young age.

6 Limitations

Despite the numerous strengths of this study, it had several limitations that must be acknowledged. First, the duration of exposure to the BSR-RV 2® game may not have been enough to fully assess the level of immersion in the virtual environment. Extending the length of the sessions may be necessary considering the developmental stages of each child. However, this may require adapting the length of the sessions to the specific needs of younger children, such as those aged 7–9 years, who may require less exposure to achieve a comparable level of absorption. By contrast, older children may benefit from a longer exposure time to achieve a higher level of immersion.

Second, the study experienced a slight loss of experimental participants, partly due to the inconsistency and lack of attention to the concepts addressed in the questionnaires administered.

Third, the unfamiliarity of the virtual environment presented to the children in this study may have influenced the results, as the children had no prior experience with wheelchair or adapted sports. Therefore, it is crucial to examine children's expectations and previous experiences with the content of virtual reality programs, particularly in serious games where the themes may be new to them.

Fourth, the limited number of trials in this study may have affected the scores obtained. Therefore, a larger number of trials may be necessary to determine whether the validity scores of the tool change with increased exposure to the BSR-RV 2® virtual environment.

Fifth, the absence of comparative studies with different age groups makes it difficult to evaluate the scores obtained for both content and face validity.

Finally, this study focused exclusively on content and face validity data. Construct, predictive, and/or criterion validity data will be presented in subsequent studies, which will explore the relationship between BSR-RV 2 measurement indicators and psychometric variables such as empathy and inclusion attitudes.

7 Conclusion

The present study evaluated the content and face validity of a basketball virtual reality tool adapted for children, using two samples of children. The content validity study found that wheelchair basketball players and conventional basketball players had similar opinions on the technical aspects of the game, as suggested by the focus group of adult professional players. However, the children in the face validity study rated enjoyment and positive emotions higher than the cognitive aspects of the game: absorption and intrinsic motivation were less common. The level of face validity for absorption and intrinsic motivation varied based on the child's age and gender, with males aged 7–8 years having a higher agreement, while the agreement was lower in the 10–11 age group. This study suggests that the sample size criteria and face validity cutoff scores should be adjusted for children, as the variability is greater in this population. A cutoff correction factor for age groups and Age × Gender interaction should be included in the formula for calculating content and face validity, as it is applied in cognitive psychometric tests. Furthermore, the sample size for estimating face validity should be enlarged to encompass a larger number of groups representing the Age × Gender interaction. This approach will ease the control over the influence of cognitive developmental variables across different developmental periods.

In conclusion, it is crucial to involve children in the design and development of virtual tools, taking into account their preferences and opinions, to ensure content and face validity. This aspect has not been sufficiently addressed in the design of new virtual tools. By incorporating children's perspectives, other validity analyses such as construct validity can become more accurate in measuring the concept from the child's perspective.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

IC: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. MA-R: Supervision, Writing–review and editing. MG: Supervision, Writing–review and editing. JL: Visualization, Writing–review and editing. MS: Supervision, Writing–review and editing. NP: Data curation, Investigation, Resources, Software, Writing–review and editing. HP: Visualization, Writing–review and editing. IA: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study has been supported by the "Plan de Recuperación, Transformación y Resilicencia" within the project "Ayudas europeas a proyectos de investigación aplicada a la (AFBS) y la medicina deportiva," funded by the European Union – NextGenerationEU (proyect number EXP_74959). Ministerio de Cultura y Deporte (Spain).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdollahipour, F., Alizadeh Zarei, M., Akbar Fahimi, M., and Karamali Esmaeili, S. (2016). Study of face and content validity of the Persian version of behavior rating inventory of executive function, preschool version. *J. Rehabilitation* 17 (1), 10–17. doi:10.20286/jrehab-170110

Ahn, S. J., Le, A. M. T., and Bailenson, J. (2013). The effect of embodied experiences on self-other merging, attitude, and helping behavior. *Media Psychol.* 16 (1), 7–38. doi:10.1080/15213269.2012.755877

Allen, M. S., Robson, D. A., and Iliescu, D. (2023). Face validity: a critical but ignored component of scale construction in psychological assessment. *Eur. J. Psychol. Assess.* 39 (3), 153–156. doi:10.1027/1015-5759/a000777

Almeida-Rocha, T., Peixoto, F., and Jesus, S. N. (2020). Aesthetic development in children, adolescents and young adults. *Anal. Psicol.* 38 (1), 1–13. doi:10.14417/ap.1657

Alsalamah, A., Campo, R., Tanos, V., Grimbizis, G., Van Belle, Y., Hood, K., et al. (2017). Face and content validity of the virtual reality simulator 'ScanTrainer®'. *Gynecol. Surg.*, 14(1), 14–16. doi:10.1186/s10397-017-1020-6

Alvarez-Lopez, F., Maina, M. F., and Saigí-Rubió, F. (2020). Use of a low-cost portable 3d virtual reality gesture-mediated simulator for training and learning basic psychomotor skills in minimally invasive surgery: development and content validity study. *J. Med. Internet Res.* 22 (7), e17491. doi:10.2196/17491

Alzahrani, T., Haddad, R., Alkhayal, A., Delisle, J., Drudi, L., Gotlieb, W., et al. (2013). Validation of the da vinci surgical skill simulator across three surgical disciplines:A pilot study. *J. Can. Urological Assoc.* 7 (7-8), 520. doi:10.5489/cuaj.419

Babar, Z. U. D., Max, S. A., Martina, B. G., Rosalia, R. A., Peek, J. J., van Dijk, A., et al. (2023). Virtual reality simulation as a training tool for perfusionists in extracorporeal circulation: establishing face and content validity. *JTCVS Tech.* 21 (C), 135–148. doi:10.1016/j.xjtc.2023.06.004

Baharuddin, I. H., Ismail, N., Naing, N. N., Ibrahim, K., Yasin, S. M., and Patterson, M. S. (2024). Content and face validity of workplace COVID-19 knowledge and stigma scale (WoCKSS). *BMC Public Health* 24 (1), 874. doi:10.1186/s12889-023-17614-3

Bailey, J. O., and Bailenson, J. N. (2017). "Immersive virtual reality and the developing child," in *Cognitive development in digital contexts*. Editors F. C. B. En and P. J. B. Brooks (Elsevier Inc). doi:10.1016/B978-0-12-809481-5.00009-2

Bart, O., Katz, N., Weiss, P. L., and Josman, N. (2008). Street crossing by typically developed children in real and virtual environments. *OTJR Occup. Participation Health* 28 (2), 89–96. doi:10.3928/15394492-20080301-01

Baughan, A., Izenman, E., Schwamm, S., Alsabeh, D., Arbor, A., Bickham, D., et al. (2024) "Investigating attention and normative dissociation in children ' s social video games," in *Proceedings of the 23rd annual ACM interaction design and children conference*, 30–43. doi:10.1145/3628516.3655808

Baumgartner, T., Speck, D., Wettstein, D., Masnari, O., Beeli, G., and Jäncke, L. (2008). Feeling present in arousing virtual reality worlds: prefrontal brain regions differentially orchestrate presence experience in adults and children. *Front. Hum. Neurosci.* 2 (AUG), 8–12. doi:10.3389/neuro.09.008.2008

Baumgartner, T., Valko, L., Esslen, M., and Jäncke, L. (2006). Neural correlate of spatial presence in an arousing and noninteractive virtual reality: an EEG and psychophysiology study. *Cyberpsychology Behav.* 9 (1), 30–45. doi:10.1089/cpb.2006.9.30

Beatini, V., Cohen, D., Di Tore, S., Pellerin, H., Aiello, P., Sibilio, M., et al. (2024). Measuring perspective taking with the "Virtual Class" videogame: a child development study. *Comput. Hum. Behav.* 151 (July 2023), 108012–108017. doi:10.1016/j.chb.2023.108012

Bek, N., Engin, I. E., Erel, S., Yakut, Y., and Uygur, F. (2009). Turkish version of impact on family scale: a study of reliability and validity. *Health Qual. Life Outcomes* 7, 1–7. doi:10.1186/1477-7525-7-4

Bernstein, D. M., and Belicki, K. (1996). On the psychometric properties of retrospective dream content questionnaires. *Imagination, Cognition Personality* 15 (4), 351–364. doi:10.2190/R1FR-YHF7-EVG9-UDJT

Blakemore, S. J., and Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *J. Child Psychol. Psychiatry Allied Discip.* 47 (3-4), 296–312. doi:10.1111/j.1469-7610.2006.01611.x

Bossavit, B., and Parsons, S. (2018). Outcomes for design and learning when teenagers with autism codesign a serious game: a pilot study. *J. Comput. Assisted Learn.* 34 (3), 293–305. doi:10.1111/jcal.12242

Bright, E., Vine, S., Wilson, M. R., Masters, R. S. W., and McGrath, J. S. (2012). Face validity, construct validity and training benefits of a virtual reality turp simulator. *Int. J. Surg.* 10 (3), 163–166. doi:10.1016/j.ijsu.2012.02.012

Campo-Prieto, P., Rodríguez-Fuentes, G., and Cancela Carral, J. M. (2021). Traducción y adaptación transcultural al español del Simulator Sickness Questionnaire (Translation and cross-cultural adaptation to Spanish of the Simulator Sickness Questionnaire). *Retos* 43, 503–509. doi:10.47197/retos.v43i0.87605

Carter, B., Grey, J., McWilliams, E., Clair, Z., Blake, K., and Byatt, R. (2014). 'Just kids playing sport (in a chair)': experiences of children, families and stakeholders attending a wheelchair sports club. *Disabil. Soc.* 29 (6), 938–952. doi:10.1080/09687599.2014.880329

Carter, F. J., Schijven, M. P., Aggarwal, R., Grantcharov, T., Francis, N. K., Hanna, G. B., et al. (2005). Consensus guidelines for validation of virtual reality surgical simulators. *Surg. Endosc. Other Interventional Tech.* 19 (12), 1523–1532. doi:10.1007/s00464-005-0384-2

Ceberio, I. (2022). "Eficacia de la realidad virtual (RV) y psicoeducación hacia la inclusión de la diversidad funcional," in *Advances in Clinical Psychology*. 1st Edn, Editor G. Buela-Casal (Dykinson), 366–367. doi:10.2307/jj.5076281.357

Choudhury, S., Blakemore, S. J., and Charman, T. (2006). Social cognitive development during adolescence. *Soc. cognitive Affect. Neurosci.* 1 (3), 165–174. doi:10.1093/scan/nsl024

Chowdhury, T. I., Ferdous, S. M. S., and Quarles, J. (2021). VR disability simulation reduces implicit bias towards persons with disabilities. *IEEE Trans. Vis. Comput. Graph.* 27 (6), 3079–3090. doi:10.1109/TVCG.2019.2958332

Chowdhury, T. I., and Quarles, J. (2022). A wheelchair locomotion interface in a VR disability simulation reduces implicit bias. *IEEE Trans. Vis. Comput. Graph.* 28 (12), 4658–4670. doi:10.1109/TVCG.2021.3099115

Chowdhury, T. I., Shahnewaz Ferdous, S. M., and Quarles, J. (2017). "Information recall in a virtual reality disability simulation. Proceedings of the ACM symposium on virtual reality software and technology,". Gothenburg, Sweden: VRST. *Part F1319*. doi:10.1145/3139131.3139143

Chuan, A., Qian, J., Bogdanovych, A., Kumar, A., McKendrick, M., and McLeod, G. (2023). Design and validation of a virtual reality trainer for ultrasound-guided regional anaesthesia. *Anaesthesia* 78 (6), 739–746. doi:10.1111/anae.16015

Connell, J., Carlton, J., Grundy, A., Taylor Buck, E., Keetharuth, A. D., Ricketts, T., et al. (2018). The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). *Qual. Life Res.* 27 (7), 1893–1902. doi:10.1007/s11136-018-1847-y

Dahlin Ivanoff, S., and Hultberg, J. (2006). Understanding the multiple realities of everyday life: basic assumptions in focus-group methodology. *Scand. J. Occup. Ther.* 13 (2), 125–132. doi:10.1080/11038120600691082

Danby, S. (2009). Childhood and social interaction in everyday life: an epilogue. *J. Pragmat.* 41 (8), 1596–1599. doi:10.1016/j.pragma.2008.06.015

de Oliveira, C. A. S., Firmino, R. T., de Morais Ferreira, F., Vargas, A. M. D., and Ferreira e Ferreira, E. (2022). Development and validation of the quality of life in the neighborhood questionnaire for children 8 to 10 Years of age (QoL-N-kids 8–10). *Child Indic. Res.* 15 (5), 1847–1870. doi:10.1007/s12187-022-09944-2

Dorozhkin, D., Olasky, J., Jones, D. B., Schwaitzberg, S. D., Jones, S. B., Cao, C. G. L., et al. (2017). OR fire virtual training simulator: design and face validity. *Surg. Endosc.* 31 (9), 3527–3533. doi:10.1007/s00464-016-5379-7

Dulan, G., Rege, R. V., Hogg, D. C., Gilberg-Fisher, K. K., Tesfay, S. T., and Scott, D. J. (2012). Content and face validity of a comprehensive robotic skills training program for general surgery, urology, and gynecology. *Am. J. Surg.* 203 (4), 535–539. doi:10.1016/j.amjsurg.2011.09.021

DunnGalvin, A., De BlokFlokstra, B. M. J., Burks, A. W., Dubois, A. E. J., and Hourihane, J. O. B. (2008). Food allergy QoL questionnaire for children aged 0-12 years: content, construct, and cross-cultural validity. *Clin. Exp. Allergy* 38 (6), 977–986. doi:10.1111/j.1365-2222.2008.02978.x

Eisen, M. L., and Lynn, S. J. (2001). Dissociation, memory and suggestibility in adults and children. *Appl. Cogn. Psychol.* 15 (7 SPEC. ISS.), 49–73. doi:10.1002/acp.834

Eppmann, R., Bekk, M., and Klein, K. (2018). Gameful experience in gamification: construction and validation of a gameful experience scale [GAMEX]. *J. Interact. Mark.* 43 (2018), 98–115. doi:10.1016/j.intmar.2018.03.002

Evans, A. B., Bright, J. L., and Brown, L. J. (2015). Non-disabled secondary school children's lived experiences of a wheelchair basketball programme delivered in the East of England. *Sport, Educ. Soc.* 20 (6), 741–761. doi:10.1080/13573322.2013.808620

Feeley, A., Feeley, I., Merghani, K., and Sheehan, E. (2021). A pilot study to evaluate the face and construct validity of an orthopaedic virtual reality simulator. *Injury* 52 (7), 1715–1720. doi:10.1016/j.injury.2021.04.045

Granic, I., Lobel, A., and Engels, R. C. M. E. (2014). The benefits of playing video games. *Am. Psychol.* 69 (1), 66–78. doi:10.1037/a0034857

Grenier, M., Collins, K., Wright, S., and Kearns, C. (2014). Perceptions of a disability sport unit in general physical education. *Adapt. Phys. Act. Q.* 31 (1), 49–66. doi:10.1123/apaq.2013-0006

Handage, S., and Chander, M. (2021). Development of an instrument for measuring the student learning outcomes: a content validation process. *Indian J. Ext. Educ.* 57 (03), 01–07. doi:10.48165/ijee.2021.57302

Hague, C., and Payton, S. (2010). Digital literacy across the curriculum. Bristol: Futurelab 4 (1), 1–63.

Hardesty, D. M., and Bearden, W. O. (2004). The use of expert judges in scale development. *J. Bus. Res.* 57 (2), 98–107. doi:10.1016/S0148-2963(01)00295-8

Harris, D. J., Bird, J. M., Smart, P. A., Wilson, M. R., and Vine, S. J. (2020). A framework for the testing and validation of simulated environments in experimentation and training. *Front. Psychol.* 11 (March), 605–610. doi:10.3389/fpsyg.2020.00605

Harris, D. J., Buckingham, G., Wilson, M. R., Brookes, J., Mushtaq, F., Mon-Williams, M., et al. (2021). Exploring sensorimotor performance and user experience within a virtual reality golf putting simulator. *Virtual Real.* 25 (3), 647–654. doi:10.1007/s10055-020-00480-4

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., and Zak, J. (2018). "Building long-term empathy: a large-scale comparison of traditional and virtual reality perspective-taking,". Editor En B. Bastian, 13, e0204494. doi:10.1371/journal.pone.0204494*PLoS ONE*Número 10

Herrera, F., and Bailenson, J. N. (2021). Virtual reality perspective-taking at scale: effect of avatar representation, choice, and head movement on prosocial behaviors. *New Media Soc.* 23 (8), 2189–2209. doi:10.1177/1461444821993121

Holden, R. R., and Passey, J. (2010). Socially desirable responding in personality assessment: not necessarily faking and not necessarily substance. *Personality Individ. Differ.* 49 (5), 446–450. doi:10.1016/j.paid.2010.04.015

Hollis, V., Openshaw, S., and Goble, R. (2002). Conducting focus groups: purpose and practicalities. *Br. J. Occup. Ther.* 65 (1), 2–8. doi:10.1177/030802260206500102

Hoyer, R. S., Elshafei, H., Hemmerlin, J., Bouet, R., and Bidet-Caulet, A. (2021). Why are children so distractible? Development of attention and motor control from childhood to adulthood. *Child. Dev.* 92 (4), e716–e737. doi:10.1111/cdev.13561

Hung, A. J., Zehnder, P., Patil, M. B., Cai, J., Ng, C. K., Aron, M., et al. (2011). Face, content and construct validity of a novel robotic surgery simulator. *J. Urology* 186 (3), 1019–1025. doi:10.1016/j.juro.2011.04.064

Huppert, D., Grill, E., and Brandt, T. (2019). Survey of motion sickness susceptibility in children and adolescents aged 3 months to 18 years. *J. Neurology* 266 (s1), 65–73. doi:10.1007/s00415-019-09333-w

Inhelder, B., and Piaget, J. (1958). The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures. *1st Edn. London: Routledge*.

Irom, B. (2018). Virtual reality and the Syrian refugee camps: humanitarian communication and the politics of empathy. *Int. J. Commun.* 12, 4269–4291.

Javora, O., Hannemann, T., Stárková, T., Volná, K., and Brom, C. (2019). Children like it more but don't learn more: effects of esthetic visual design in educational games. *Br. J. Educ. Technol.* 50 (4), 1942–1960. doi:10.1111/bjet.12701

Jiang, Z., Meltzer, A., and Zhang, X. (2022). Using virtual reality to implement disability studies' advocacy principles: uncovering the perspectives of people with disability. *Disabil. Soc.* 0 (0), 1592–1612. doi:10.1080/09687599.2022.2150601

Kang, S., and Kang, S. (2019). The study on the application of virtual reality in adapted physical education. *Clust. Comput.* 22, 2351–2355. doi:10.1007/s10586-018-2254-4

Kaufman, J. C., Baer, J., Cole, J. C., and Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Res. J.* 20 (2), 171–178. doi:10.1080/10400410802059929

Kelly, D. C., Margules, A. C., Kundavaram, C. R., Narins, H., Gomella, L. G., Trabulsi, E. J., et al. (2012). Face, content, and construct validation of the da Vinci Skills Simulator. *Urology* 79 (5), 1068–1072. doi:10.1016/j.urology.2012.01.028

Kelly, N., Stafford, J., Craig, C., Herring, M. P., and Campbell, M. (2022). Using a virtual reality cricket simulator to explore the effects of pressure, competition anxiety on batting performance in cricket. *Psychol. Sport Exerc.* 63 (July), 102244. doi:10.1016/j.psychsport.2022.102244

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* 3 (3), 203–220. doi:10.1207/s15327108ijap0303_3

Kenney, P. A., Wszolek, M. F., Gould, J. J., Libertino, J. A., and Moinzadeh, A. (2009). Face, content, and construct validity of dV-trainer, a novel virtual reality simulator for robotic surgery. *Urology* 73 (6), 1288–1292. doi:10.1016/j.urology.2008.12.044

Lacasse, Y., Godbout, C., and Sériès, F. (2002). Health-related quality of life in obstructive sleep apnoea. *Eur. Respir. J.* 19 (3), 499–503. doi:10.1183/09031936.02.00216902

Lawshe, C. H. (1975). A quantitative approach to content validity. *Pers. Psychol.* 28, 563–575. doi:10.1111/j.1744-6570.1975.tb01393.x

Lázaro, E., Amayra, I., López-Paz, J. F., Martínez, O., Alvarez, M. P., Berrocoso, S., et al. (2020). Using a virtual serious game (Deusto-e-motion1.0) to assess the theory of mind in primary school children: observational descriptive study. *JMIR Serious Games* 8 (2), 129711–e13014. doi:10.2196/12971

Lopes, V. P., Barnett, L. M., Saraiva, L., Gonçalves, C., Bowe, S. J., Abbott, G., et al. (2016). Validity and reliability of a pictorial instrument for assessing perceived motor competence in Portuguese children. *Child care, health Dev.* 42 (5), 666–674. doi:10.1111/cch.12359

Lundberg, N. R., Zabriskie, R. B., Smith, K. M., and Barney, K. W. (2008). Using wheelchair sports to complement disability awareness curriculum among college students. *SCHOLE A J. Leis. Stud. Recreat. Educ.* 23 (1), 61–74. doi:10.1080/1937156X.2008.11949610

Lynn (1985). Determination and quanfication of content validity. *Nurs. Res.*

Maister, L., Sebanz, N., Knoblich, G., and Tsakiris, M. (2013). Experiencing ownership over a dark-skinned body reduces implicit racial bias. *Cognition* 128 (2), 170–178. doi:10.1016/j.cognition.2013.04.002

Manan, N. M., Musa, S., Nor, M. M. D., Saari, C. Z., and Al-Namankany, A. (2024). The abeer children dental anxiety scale (ACDAS) cross-cultural adaptation and validity of self-report measures in the Malaysian children context (MY-ACDAS). *Int. J. Paediatr. Dent.* 34 (3), 267–276. doi:10.1111/ipd.13132

Marlenga, B., Berg, R. L., Pickett, W., Brown, T., Becklinger, N., and Schwebel, D. C. (2017). Using simulation to assess the ability of youth to safely operate tractors. *Transp. Res. Part F Traffic Psychol. Behav.* 48, 28–37. doi:10.1016/j.trf.2017.04.021

Martucci, A., Gursesli, M. C., Duradoni, M., and Guazzini, A. (2023). Overviewing gaming motivation and its associated psychological and sociodemographic variables: a prisma systematic review. *Hum. Behav. Emerg. Technol.* 2023, 1–156. doi:10.1155/2023/5640258

Mason, J., and Danby, S. (2011). Children as experts in their lives: child inclusive research. *Child Indic. Res.* 4 (2), 185–189. doi:10.1007/s12187-011-9108-4

Matera, C., Nerini, A., Di Gesto, C., Policardo, G. R., Maratia, F., Dalla Verde, S., et al. (2021). Put yourself in my wheelchair: perspective-taking can reduce prejudice toward people with disabilities and other stigmatized groups. *J. Appl. Soc. Psychol.* 51 (3), 273–285. doi:10.1111/jasp.12734

McDougall, E. M., Corica, F. A., Boker, J. R., Sala, L. G., Stoliar, G., Borin, J. F., et al. (2006). Construct validity testing of a laparoscopic surgical simulator. *J. Am. Coll. Surg.* 202 (5), 779–787. doi:10.1016/j.jamcollsurg.2006.01.004

Medland, J., and Ellis-Hill, C. (2008). Why do able-bodied people take part in wheelchair sports? *Disabil. Soc.* 23 (2), 107–116. doi:10.1080/09687590701841133

Mendoza Estor, G. C. (2017). "La lúdica y el deporte como estrategia de integración social en niños del INEM Simón Bolívar de Santa Marta," in *Repositorio institucional - Fundación Universitaria Los Libertadore*. Master's thesis. Fundación Universitaria Los Libertadores, 6–18.

Miner, J. D. (2022). Informatic tactics: indigenous activism and digital cartographies of gender-based violence. *Inf. Commun. Soc.* 25 (3), 431–448. doi:10.1080/1369118X.2020.1797851

Mohamad Marzuki, M. F., Yaacob, N. A., and Yaacob, N. M. (2018). Translation, cross-cultural adaptation, and validation of the Malay version of the system usability scale questionnaire for the assessment of mobile apps. *JMIR Hum. Factors* 5 (2), 103088–e10317. doi:10.2196/10308

Morgado, L. D. S., Martelaer, K., Sääkslahti, A., Howells, K., Barnett, L. M., D'Hondt, E., et al. (2023). Face and content validity of the pictorial scale of perceived water competence in young children. *Children* 10 (1), 2. doi:10.3390/children10010002

Moss, P., Lim, K. H., Prunty, M., and Norris, M. (2020). Children and young people's perspectives and experiences of a community wheelchair basketball club and its impact on daily life. *Br. J. Occup. Ther.* 83 (2), 118–128. doi:10.1177/0308022619879333

Moulton, J. M., Cole, C., Ridgers, N. D., Pepin, G., and Barnett, L. M. (2019). Measuring movement skill perceptions in preschool children: a face validity and reliability study. *Aust. Occup. Ther. J.* 66 (1), 13–22. doi:10.1111/1440-1630.12485

Nakayama, H., Matsuzaki, T., Mihara, S., Kitayuguchi, T., and Higuchi, S. (2020). Relationship between problematic gaming and age at the onset of habitual gaming. *Pediatr. Int.* 62 (11), 1275–1281. doi:10.1111/ped.14290

Nevo, B. (1985). Face validity revisited. *J. Educ. Meas.* 22 (4), 287–293. doi:10.1111/j.1745-3984.1985.tb01065.x

Neyret, S., Navarro, X., Beacco, A., Oliva, R., Bourdin, P., Valenzuela, J., et al. (2020). An embodied perspective as a victim of sexual harassment in virtual reality reduces action conformity in a later milgram obedience scenario. *Sci. Rep.* 10 (1), 6207–6218. doi:10.1038/s41598-020-62932-w

Nunnally, J. C., and Bernstein, I. H. (1994). Psychometric theory. 3rd Edn. New York: McGraw-hill series.

Oh, S. Y., Bailenson, J., Weisz, E., and Zaki, J. (2016). Virtually old: embodied perspective taking and the reduction of ageism under threat. *Comput. Hum. Behav.* 60, 398–410. doi:10.1016/j.chb.2016.02.007

Pan, X., and Hamilton, A. F. d. C. (2018). Why and how to use virtual reality to study human social interaction: the challenges of exploring a new research landscape. *Br. J. Psychol.* 109 (3), 395–417. doi:10.1111/bjop.12290

Parra-González, M. E., and Segura-Robles, A. (2019). Traducción y validación de la escala de evaluación de experiencias gamificadas (GAMEX). *Bordon. Rev. Pedagog.* 71 (4), 87–99. doi:10.13042/Bordon.2019.70783

Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious. Cognition* 22 (3), 779–787. doi:10.1016/j.concog.2013.04.016

Pinto-Coelho, L., Laska-Leśniewicz, A., Pereira, E. T., and Sztobryn-Giercuszkiewicz, J. (2023). Inclusion and adaptation beyond disability: using virtual reality to foster empathy. *Med. Pr.* 74 (3), 171–185. doi:10.13075/mp.5893.01386

Pivik, J., McComas, J., Macfarlane, I., and Laflamme, M. (2002). Using virtual reality to teach disability awareness. *J. Educ. Comput. Res.* 26 (2), 203–218. doi:10.2190/WACX-1VR9-HCMJ-RTKB

Plowman, L., Stevenson, O., Stephen, C., and McPake, J. (2012). Preschool children's learning with technology at home. *Comput. Educ.* 59 (1), 30–37. doi:10.1016/j.compedu.2011.11.014

Reinhard, R., Shah, K. G., Faust-Christmann, C. A., and Lachmann, T. (2020). Acting your avatar's age: effects of virtual reality avatar embodiment on real life walking speed. *Media Psychol.* 23 (2), 293–315. doi:10.1080/15213269.2019.1598435

Robinson, L. E., and Palmer, K. K. (2017). Development of a digital-based instrument to assess perceived motor competence in children: face validity, test-retest reliability, and internal consistency. *Sports* 5 (3), 48. doi:10.3390/sports5030048

Romero Jeldres, M., Díaz Costa, E., and Faouzi Nadim, T. (2023). A review of Lawshe's method for calculating content validity in the social sciences. *Front. Educ.* 8 (November), 1–8. doi:10.3389/feduc.2023.1271335

Rosenberg, R. S., Baughman, S. L., and Bailenson, J. N. (2013). Virtual superheroes: using superpowers in virtual reality to encourage prosocial behavior. *PLoS ONE* 8 (1), e55003–e55010. doi:10.1371/journal.pone.0055003

Ruiz Amayas (2019). "Desarrollo de la empatía a través de la actividad física," in *Repositorio institucional - Universidad de Valladolid*. Master's thesis. Universidad de Valladolid, 1–58.

Runswick, O. R. (2023). Player perceptions of face validity and fidelity in 360-video and virtual reality cricket. *J. Sport Exerc. Psychol.* 1 (8), 347–354. doi:10.1123/jsep.2023-0122

Ryberg, A. M., Nielsen, P. B., Graarup, K. S., Ingeman, K., Thellefsen, M. R., and Jensen, C. S. (2023). Danish translation and cultural adaptation of the 'What do you think of hospital' patient reported experience measure for children and adolescents in outpatient settings. *J. Pediatr. Nurs.* 68, e36–e42. doi:10.1016/j.pedn.2022.10.016

Sapey, B., Stewart, J., and Donaldson, G. (2005). Increases in wheelchair use and perceptions of disablement. *Disabil. Soc.* 20 (5), 489–505. doi:10.1080/09687590500156162

Schilling, L. S., Dixon, J. K., Knafl, K. A., Grey, M., Ives, B., and Lynn, M. R. (2007). Determining content validity of a self-report instrument for adolescents using a heterogeneous expert panel. *Nurs. Res.* 56 (5), 361–366. doi:10.1097/01.nnr.0000289505.30037.91

Schwebel, D. C., Gaines, J., and Severson, J. (2008). Validation of virtual reality as a tool to understand and prevent child pedestrian injury. *Accid. Analysis Prev.* 40 (4), 1394–1400. doi:10.1016/j.aap.2008.03.005

Seixas-Mikelus, S. A., Kesavadas, T., Srimathveeravalli, G., Chandrasekhar, R., Wilding, G. E., and Guru, K. A. (2010). Face validation of a novel robotic surgical simulator. *Urology* 76 (2), 357–360. doi:10.1016/j.urology.2009.11.069

Shen, J., Koterba, C., Samora, J., Leonard, J., Li, R., Shi, J., et al. (2022). Usability and validity of a virtual reality cognitive assessment tool for pediatric traumatic brain injury. *Rehabil. Psychol.* 67 (4), 587–596. doi:10.1037/rep0000464

Sheridan, M., Kharitonova, M., Martin, R. E., Chatterjee, A., and Gabrieli, J. D. (2017). Neural substrates of the development of cognitive control in children ages 5-10 years. *J. cognitive Neurosci.* 26 (8), 1840–1850. doi:10.1162/jocn_a_00597

Sim, G., Read, J. C., Gregory, P., and Xu, D. (2015). From England to Uganda: children designing and evaluating serious games. *Human-Computer Interact.* 30 (3-4), 263–293. doi:10.1080/07370024.2014.984034

Slater, M., and Banakou, D. (2021). The golden rule as a paradigm for fostering prosocial behavior with virtual reality. *Curr. Dir. Psychol. Sci.* 30 (6), 503–509. doi:10.1177/09637214211046954

Slater, M., and Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Front. Robot. AI* 3 (DEC), 1–47. doi:10.3389/frobt.2016.00074

Sofokleous, R., and Stylianou, S. (2023). Effects of exposure to medical model and social model online constructions of disability on attitudes toward wheelchair users: results from an online experiment. *J. Creative Commun.* 18 (1), 61–78. doi:10.1177/09732586221136260

Staudinger, U. M., Werner, I., Valsiner, J., and Valsiner, K. J. (2003). "Wisdom: Its social nature and lifespan development," in *Handbook of developmental psychology*. Editors J. Valsiner and K. J. Connolly (London), 584–602.

Stewart, J. L., Mary, R., and Lynn, M. H. M. (2005). Evaluating content validity for children´s self-report instrument using children as experts. 54(6), 43–62.

Swinkels, L. M. J., Veling, H., Dijksterhuis, A., and van Schie, H. T. (2021). Availability of synchronous information in an additional sensory modality does not enhance the full body illusion. *Psychol. Res.* 85 (6), 2291–2312. doi:10.1007/s00426-020-01396-z

Tamm, M., and Prellwitz, M. (2001). 'If I had a friend in a wheelchair': children's thoughts on disabilities. *Child Care, Health Dev.* 27 (3), 223–240. doi:10.1046/j.1365-2214.2001.00156.x

Tammy Lin, J. H., and Wu, D. Y. (2021). Exercising with embodied young avatars: how young vs. Older avatars in virtual reality affect perceived exertion and physical activity among male and female elderly individuals. *Front. Psychol.* 12 (October), 693545–693614. doi:10.3389/fpsyg.2021.693545

Thomas, S., Bardy, B., and Smart, L. J. (2003). "On the nature and evaluation of fidelity in virtual environments," in *Virtual and adaptive environments. Applications, implications and human performance issues*. Editors L. J. H. En and M. Haas doi:10.1201/9781410608888

Valentini, N. C., Barnett, L. M., Bandeira, P. F. R., Nobre, G. C., Zanella, L. W., and Sartori, R. F. (2018). The pictorial scale of perceived movement skill competence: determining content and construct validity for Brazilian children. *J. Mot. Learn. Dev.*, 6, S189-204. doi:10.1123/jmld.2016-0043

Valenza, M. V., Gasparini, I., and Hounsell, M. da S. (2019). Serious game design for children: a set of guidelines and their validation. *Educ. Technol. Soc.* 22 (3), 19–31. doi:10.30191/ETS.201907_22(3).0003

VanDeventer, S. S., and White, J. A. (2002). Expert behavior in children's video game play. *Simul. Gaming* 33 (1), 28–48. doi:10.1177/1046878102033001002

Veerman, L. K. M., Fjermestad, K. W., Vatne, T. M., Sterkenburg, P. S., Derks, S. D. M., Brouwer-van Dijken, A. A. J., et al. (2024). Cultural applicability and desirability of 'Boodles': the first serious game intervention for siblings of children with disabilities. *PEC Innov.* 4 (October 2023), 100277. doi:10.1016/j.pecinn.2024.100277

Venetsanou, F., Kossyva, I., Valentini, N., Afthentopoulou, A. E., and Barnett, L. (2018). Validity and reliability of the pictorial scale of perceived movement skill competence for young Greek children. *J. Mot. Learn. Dev.* 6, S239–S251. doi:10.1123/jmld.2017-0028

Verdaasdonk, E. G. G., Stassen, L. P. S., Monteny, L. J., and Dankelman, J. (2006). Validation of a new basic virtual reality simulator for training of basic endoscopic skills: the SIMENDO. *Surg. Endosc. Other Interventional Tech.* 20 (3), 511–518. doi:10.1007/s00464-005-0230-6

Vilchinsky, N., Werner, S., and Findler, L. (2010). Gender and attitudes toward people using wheelchairs: a multidimensional perspective. *Rehabil. Couns. Bull.* 53 (3), 163–174. doi:10.1177/0034355209361207

Willaert, W. I. M., Aggarwal, R., Herzeele, I. V., Cheshire, N. J., and Vermassen, F. E. (2012). Recent advancements in medical simulation: patient-specific virtual reality simulation. *World J. Surg.* 36 (7), 1703–1712. doi:10.1007/s00268-012-1489-0

Xiao, D., Jakimowicz, J. J., Albayrak, A., Buzink, S. N., Botden, S. M. B. I., and Goossens, R. H. M. (2014). Face, content, and construct validity of a novel portable ergonomic simulator for basic laparoscopic skills. *J. Surg. Educ.* 71 (1), 65–72. doi:10.1016/j.jsurg.2013.05.003

Yadav, S., Chakraborty, P., Meena, L., Yadav, D., and Mittal, P. (2021). Children's interaction with touchscreen devices: performance and validity of Fitts' law. *Hum. Behav. Emerg. Technol.* 3 (5), 1132–1140. doi:10.1002/hbe2.305

Yamada-Rice, D., Mushtaq, F., Woodgate, A., Bosmans, D., Douthwaite, A., Douthwaite, I., et al. (2017)2017) *Children and virtual reality: emerging possibilities and challenges*, 1-23. Available at: http://digilitey.eu.

Yee, N., and Bailenson, J. (2006). Walk a mile in digital shoes: the impact of embodied perspective-taking on the reduction of negative stereotyping in immersive virtual environments. *Proc. PRESENCE*, 147–156.

Yee, N., Bailenson, J. N., and Ducheneaut, N. (2009). The Proteus effect. *Commun. Res.* 36 (2), 285–312. doi:10.1177/0093650208330254

Yusoff, M. S. B. (2019a). ABC of content validation and content validity index calculation. *Educ. Med. J.* 11 (2), 49–54. doi:10.21315/eimj2019.11.2.6

Yusoff, M. S. B. (2019b). ABC of response process validation and face validity index calculation. *Educ. Med. J.* 11 (3), 55–61. doi:10.21315/eimj2019.11.3.6

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., and Nikanfar, A.-R. (2015). Design and implementation content validity study: development of an instrument for measuring patient-centered communication. *J. Caring Sci.* 4 (2), 165–178. doi:10.15171/jcs.2015.017