Journal on Interactive Systems, 2025, 16:1, doi: 10.5753/jis.2025.4785

Self-concepts applied as a strategy to measure the level of acceptance of a Serious Game

Anthony Lins [**Universidade Católica de Pernambuco** | ***anthony.lins@unicap.br***]
Walter Franklin Marques Correia [**Universidade Federal de Pernambuco** | ***wfmc@cin.ufpe.br***]
Fábio Ferreira da Costa Campos [**Breda University of Applied Sciences** | ***ferreira.f@buas.nl***]
Marnix van Gisbergen [**Breda University of Applied Sciences** | ***gisbergen.m@buas.nl***]
João Marcelo Xavier Natário Teixeira [**Universidade Federal de Pernambuco** | ***jmxnt@cin.ufpe.br***]

Universidade Católica de Pernambuco, Rua do Príncipe, 526, Boa Vista, Recife, PE, 50050-900, Brazil.

Received: 05 August 2024 • **Accepted:** 10 December 2024 • **Published:** 01 January 2025

Abstract This paper explores the application of Self-Concepts, a Semantic-Differential Scale-based methodology, to evaluate a Serious Game designed for cognitive impairment assessment using the MoCA test and augmented reality (AR) for mobility analysis. A case study involved seven elderly participants. Despite some challenges in navigating the AR environment and cognitive tests, self-concept evaluations consistently demonstrated high user expectations and positive experiences with the digital artifact. This study validates the process and highlights the potential of Self-Concepts in assessing product acceptance within user groups.

Keywords: Self-concepts, Serious games, User acceptance

1 Introduction

The digital games market has witnessed exponential growth, establishing itself as the most lucrative entertainment industry worldwide. In 2023, the global market value of digital games was estimated at approximately $220 billion, with projections suggesting this figure will reach $500 billion by 2027 [Statista, 2023]. This growth reflects not only the increasing popularity of games but also their diversification across applications and audiences. As illustrated in Figure 1, the decade-long revenue forecast highlights the continuous expansion of different market segments, underscoring the industry's broad appeal and sustained evolution.

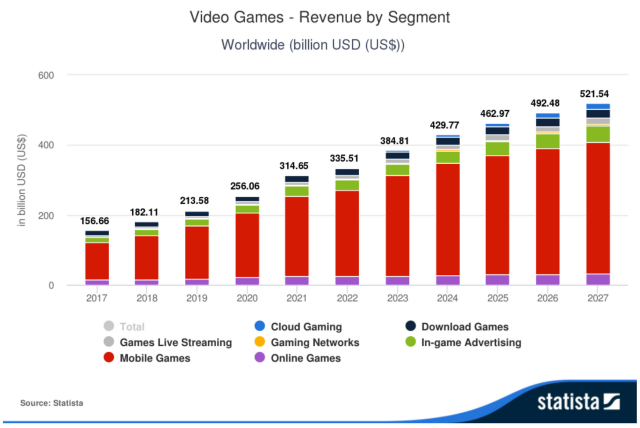

Figure 1. Revenue forecast by segment in the Gaming market between 2017 and 2027 (Source: Statista, 2023).

Within this dynamic context, Serious Games have emerged as a significant subsegment, distinguished by their ability to integrate engaging gameplay with practical objectives such as education, training, and healthcare interventions. These games are increasingly recognized for their potential to enhance cognitive, emotional, and motor functions, particularly in older adults and individuals undergoing rehabilitation [Lin *et al*., 2023]. Moreover, Serious Games presents cost-effective solutions in healthcare, improving adherence to diagnostic protocols and treatments while reducing associated costs [Carvalho *et al*., 2024; Funabashi *et al*., 2018; Amaral *et al*., 2019].

The evaluation of user experience (UX) is crucial in understanding how Serious Games are perceived and adopted by their target audiences. Methodologies such as semantic differential scales and self-concept analysis provide robust mechanisms for measuring emotional and cognitive responses, offering valuable insights for developing and refining digital artifacts [Borges *et al*., 2019].

This study explores the application of Self-Concepts as a methodology for evaluating a Serious Game designed for early cognitive impairment diagnosis and mobility assessment using augmented reality. By leveraging this approach, the study aims to validate the effectiveness of Self-Concepts in user experience evaluation and identify key factors that influence the acceptance and success of Serious Games within the healthcare domain.

2 Theoretical context

2.1 Serious Games

Serious Games are a distinct genre of digital games that transcend entertainment by integrating gamified learning processes and interactive environments aimed at addressing real-world challenges. These games are increasingly used in various sectors, including healthcare, education, and organizational training, to foster engagement and knowledge transfer Gazis and Katsiri [2023]. Within the healthcare domain, Se-

Table 1. Bipolar scale examples.

Unhappy	Happy
Annoyed	Pleased
Unsatisfied	Satisfied
Melancholic	Contented
Despairing	Hopeful
Bored	Relaxed
Relaxed	Stimulated
Calm	Excited

rious Games have demonstrated their potential to improve patient motivation and adherence to various interventions, combining engaging gameplay with practical therapeutic or diagnostic objectives [Lin *et al*., 2023].

The appeal of Serious Games lies in their ability to evoke strong emotional and cognitive responses, fostering a unique interplay between learning, behavior modification, and engagement. Recent studies emphasize their role in improving cognitive and motor skills in older adults, as well as enhancing adherence to rehabilitation programs, underscoring their transformative potential in modern healthcare systems [Lin *et al*., 2023].

2.2 User Experience

User Experience (UX) design plays a pivotal role in analyzing and enhancing user interactions with digital artifacts, particularly in the realm of Serious Games. UX encompasses a wide range of dimensions, including usability, aesthetics, engagement, and emotional impact, which collectively shape the overall player experience. As Serious Games grows in narrative complexity, gameplay diversity, and applicability, evaluating UX through psychological and emotional lenses—such as immersion, flow, and enjoyment—has become increasingly critical [Borges *et al*., 2019][Michailidis *et al*., 2018].

Research in Human-Computer Interaction (HCI) and design underscores the importance of UX in addressing users' motivations, expectations, and interactions with products and services. Beyond usability, UX integrates hedonistic and affective aspects, aligning with Maslow's hierarchy of needs to address not only physiological but also psychological and self-actualization needs [Neves, 2021]. This holistic approach is particularly significant in gaming, where decision-making influenced by emotional aspects is a primary driver of enjoyment [Nacke, 2015].

2.2.1 Methods for Evaluating UX

UX evaluation methods for Serious Games often leverage structured tools such as semantic differential scales, questionnaires, and other HCI-driven approaches to capture user perceptions effectively. These methods assess critical dimensions, including cognitive and emotional responses, which directly influence user engagement and satisfaction. For example, semantic differential scales utilize bipolar adjective pairs, such as "boring–engaging" or "simple–complex", to quantify emotional and cognitive dimensions of gameplay experiences [Bernhaupt *et al*., 2008]. Examples of commonly used bipolar scales are shown in Table 1.

Questionnaire-based methods offer a quick way to obtain quantitative analyses of players' feelings and attitudes. While these tools may lack the depth of interviews or the objectivity of physiological metrics, they become increasingly reliable with larger participant numbers [Nacke, 2015]. Complementary techniques, such as behavioral observations and heuristic evaluations, provide richer insights into user interactions.

2.2.2 Applications in Serious Games

In Serious Games, UX evaluation extends to psychological, physical, and emotional dimensions, making it possible to assess the holistic impact of gameplay. Methods like semantic differential scales and self-assessment questionnaires have proven effective in measuring pleasure, arousal, and mastery—core aspects of human judgment across perceptual and symbolic stimuli [Borges *et al*., 2019]. Moreover, technological applications, including facial expression analysis, gesture recognition, and neurophysiological monitoring, offer promising avenues to correlating gameplay experiences with emotional responses [Lin *et al*., 2023].

By focusing on the nuances of user interactions, these evaluation methods enable developers to refine Serious Games iteratively. In healthcare and education, this refinement ensures that games not only meet functional objectives but also foster meaningful and engaging experiences that align with user expectations and broader goals.

2.3 Semantic Differential Scale

The Semantic Differential Scale (SDS), originally developed by Osgood in 1952, remains a versatile and effective tool for evaluating user perceptions across diverse domains. SDS employs bipolar adjective pairs, such as "unsatisfied–satisfied" or "confident–insecure," to measure individuals' reactions to stimuli, providing a nuanced understanding of their emotional and cognitive responses [Stoklasa *et al*., 2019] [Garcia, 2023]. Initially conceived for psychological research, SDS has since been adapted for various fields, including marketing, education, and digital interaction design Osgood [1952].

In the context of Serious Games, SDS is particularly valuable for evaluating critical design elements such as gameplay, narrative coherence, and aesthetic quality. It has been shown to outperform traditional Likert scales in specific contexts by capturing complex constructs like ambivalence and emotional engagement more effectively [Bernhaupt *et al*., 2008][Borges *et al*., 2019]. For instance, SDS can quantify user experiences through scales like "happy–sad," "relaxed–stimulated," or "simple–complex," enabling a balanced view of both positive and negative perceptions. Recent studies demonstrate that the SDS is effective in evaluating design elements in serious games, such as gameplay, narrative, and aesthetic quality. For example, Nakamura et al. (2021) [Nakamura *et al*., 2021] developed pairs of adjectives to assess audiovisual effects in video games, highlighting the utility of the SDS in quantifying user impressions.

2.3.1 Key Dimensions and Methodological Considerations

SDS measures perceptions by assigning values on a numerical scale with contrasting adjectives at each end. Typical coding includes neutral (0), slightly negative or positive (-1 and 1), moderately negative or positive (-2 and 2), and extremely negative or positive (-3 and 3). Through factor analysis, three principal dimensions—evaluation (E), power (P), and activity (A)—have been identified, providing a robust framework for interpreting semantic space [Stoklasa *et al*., 2019]. For example, "calm–excited" may represent the activity dimension, while "attractive–unappealing" might correspond to evaluation. An example of the scale applied in SDS is illustrated in 2.

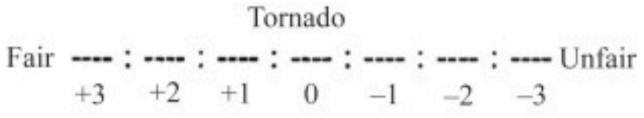

Figure 2. Example of semantic differential scaling (Adapted from Ploder and Eder [2015]).

While SDS is conceptually straightforward, its application requires careful attention to cultural and linguistic contexts. Studies highlight the importance of selecting relevant bipolar adjectives and ensuring their clarity to avoid ambiguity. Additionally, issues like midpoint neutrality, where responses are interpreted as neutral but may reflect ambivalence or indifference, must be addressed [Van Harreveld *et al*., 2015]. To mitigate such biases, alternating the placement of positive and negative terms on the scale has been suggested [Friborg *et al*., 2006]. Kim and Jung (2023) [Kim and Jung, 2023a] developed a semantic differential scale to analyze human attitudes toward artificial intelligence agents, emphasizing the necessity of considering cultural and linguistic contexts when constructing bipolar adjective scales.

2.3.2 Applications in Serious Games and Healthcare

SDS has demonstrated substantial utility in healthcare-related Serious Games, where emotional and cognitive engagement are critical metrics. For instance, it has been used to evaluate virtual reality games designed for cognitive and motor rehabilitation, revealing actionable insights for iterative design improvements [Lin *et al*., 2023]. By analyzing emotional responses to game elements, researchers can identify areas for refinement, such as adjusting gameplay difficulty or enhancing the visual appeal to align with user expectations.

Furthermore, SDS offers advantages over other evaluation methods by capturing the connotative meaning of concepts, thereby elucidating the nuances of user experiences. Recent applications include measuring engagement and usability in augmented reality environments, where user feedback is pivotal for improving interaction design [Garcia, 2023]. Lin et al. (2023)[Lin *et al*., 2023] utilized the SDS to evaluate virtual reality games designed for cognitive and motor rehabilitation, providing actionable insights for iterative design improvements.

2.3.3 Challenges and Limitations

Despite its strengths, SDS is not without limitations. Cultural variations in interpreting bipolar adjectives and individual tendencies, such as acquiescence (the propensity to agree with items regardless of content), can influence results. Additionally, participants with stronger negative biases may disproportionately impact overall findings, skewing ambivalent situations more negatively [Friborg *et al*., 2006] [Van Harreveld *et al*., 2015].

To address these challenges, researchers are advised to pilot-test scales in the target population to ensure relevance and clarity. Combining SDS with complementary methods, such as behavioral observations or physiological measurements, can also provide a more comprehensive understanding of user responses.

2.4 Self-Concepts

Self-concept is a multifaceted construct encompassing an individual's perceptions, beliefs, and feelings about themselves, significantly influencing behaviors, attitudes, and reactions. It is commonly defined as the totality of self-referential thoughts and feelings, reflecting one's understanding of personal attributes and identity [McMullen, 2020].

This construct is typically measured across various domains, capturing how individuals perceive their abilities and characteristics in specific contexts [Trautwein and Möller, 2016]. Self-concept comprises two essential components: self-image, which pertains to one's perception of personal attributes, and self-esteem, relating to the evaluative aspect of self-perception [García-Martínez *et al*., 2022]. These elements are shaped through experiences and environmental interactions, forming a dynamic self-structure.

Self-concept manifests in various forms, including:

- Actual Self: How individuals currently perceive themselves.
- Ideal Self: How individuals aspire to be perceived.
- Social Self: How individuals believe they are perceived by others.
- Situational Self: How individuals perceive themselves in specific contexts.

These dimensions highlight the complexity of self-concept and its role in personal identity.

Self-Congruence Theory posits that consumer behavior is influenced by the alignment between a consumer's self-concept and the perceived image of a product or brand. High congruence between the ideal self-concept ("how I would feel or be perceived using the product") and the actual self-concept ("how I truly feel about the product") correlates with increased preference or adoption of that product [Aini and Ferdinand, 2021]. This alignment affects attitudes toward brands and purchasing decisions.

To quantify self-concept congruence, researchers often employ semantic differential scales to assess personal and product-related perceptions. The degree of congruence is calculated using metrics such as Euclidean distance between these perceptions, providing insights into consumer preferences [Hosany and Martin, 2012]. This methodology facilitates the evaluation of artifacts, assessment of modifications

or innovations, prediction of market performance, and positioning of products within specific contexts. Aguiar (2011) [Aguiar *et al*., 2011] examines the application of self-report scales, such as Likert and Semantic Differential Scales, for capturing user perceptions of engagement, satisfaction, and usability in games. It validates these methodologies as effective tools for evaluating subjective experiences. In your study, Aguiar's findings reinforce the reliability of using Semantic Differential Scales to analyze ideal and real self-concepts, providing a robust methodological framework for subjective user assessments.

Self-concept analysis is applicable across various stages of the design process, including problem exploration and solution evaluation. It serves as a valuable tool for comparing artifacts, understanding user perceptions, and guiding design decisions to enhance user engagement and satisfaction.

3 Applying Evaluation Metrics in Serious Games

Digital games are sophisticated artifacts that combine computational architecture with multimedia elements to deliver immersive, emotional, and creative experiences. As [Suryapranata *et al*., 2017] argues, games, like any other type of software, require robust evaluation metrics to assess their quality. However, traditional software quality metrics are insufficient for games due to their distinctive features, such as visual and audio design, usability, gameplay, content, and the element of fun. Consequently, the development of specialized methods for evaluating the quality of digital games, particularly Serious Games, has become a critical area of research. Fellows et al. (2022) [Fellows *et al*., 2022] explore evaluation methods emphasizing social interaction and cultural-historical contexts in serious games. The review discusses approaches that assess not only educational effectiveness but also user motivation and engagement. This study aligns with your methodology by highlighting the importance of evaluating both emotional and educational impacts, particularly in serious games that integrate health and learning objectives.

3.0.1 Quality Metrics for Serious Learning Games

Suryapranata et al. (2017) [Suryapranata *et al*., 2017] propose a quality evaluation method tailored to Serious Games, utilizing Likert scale questionnaires. These tools allow players to express their opinions on game quality by evaluating both learning aspects (attention, relevance, confidence, and satisfaction) and technical quality factors. The proposed metrics integrate perspectives from both the player (balance, usability, and learning effectiveness) and the software (performance, completeness, compatibility, maintainability, resource capacity, and portability). This dual-focus approach ensures a comprehensive game quality assessment, encompassing user experience and technical performance.

3.0.2 Evaluation Characteristics and Methodologies

A broader analysis of Serious Games evaluation is presented by Fokides *et al*. [2019], who identify eighteen critical characteristics, including game design, aesthetics, user satisfaction, performance, engagement, motivation, educational effectiveness, social impact, and entertainment. Their study highlights the dual purpose of Serious Games, combining leisure and educational objectives, with user experience shaped by both pedagogical strategies and gameplay mechanics. Fokides et al.Fokides *et al*. [2019] also propose an evaluation method incorporating exploratory and confirmatory stages, using Likert scales to examine factor relationships and consistency levels. Despite challenges in obtaining reliable data on player engagement duration, their research contributes a validated evaluation scale that could inform the development of semantic differential scales for analyzing Serious Games. Dutra et al. (2021) [Dutra *et al*., 2021] provides a comprehensive review of Human-Computer Interaction (HCI) evaluation methods in educational serious games, highlighting widely used tools such as self-report questionnaires, heuristic evaluations, and behavioral observation. This study identifies their strengths and limitations in assessing usability and user experience, it also supports the use of the Semantic Differential Scale (SDS) as a robust tool for capturing user expectations and experiences. By aligning with established HCI practices in serious games, your methodology gains further validation.

3.0.3 Gameplay and Educational Effectiveness in 2D and 3D Serious Games

Kim and Jung (2023)[Kim and Jung, 2023b] examined gameplay and educational content in 2D and 3D Serious Games used in educational contexts. Using Likert scale questionnaires, they correlated common student characteristics (e.g., gender, technological proficiency, prior gaming experience, and domain knowledge) with factors influencing gameplay and learning outcomes. Their findings suggest that 2D Serious Games are more effective in enhancing learning outcomes than their 3D counterparts, irrespective of student demographics. This observation underscores the need to align game design with pedagogical objectives to maximize educational impact.

3.0.4 Emotional Indicators in Serious Games

Emotions play a pivotal role in shaping the user experience in Serious Games. According to [Anolli *et al*., 2010], emotional indicators are instrumental in assessing the balance between learning and fun. These indicators help monitor how emotions moderate learning outcomes and identify potential design flaws before a game's release. Such metrics are particularly valuable for ensuring that Serious Games effectively integrates educational goals with engaging gameplay.

4 Proposal of Self-Concepts for Serious Games – Case Study

This case study investigates the digital artifact BrainTest, a mobile application initially developed for the Android platform. BrainTest is designed for health applications, with a primary focus on the early detection of cognitive decline associated with neurodegenerative diseases. The artifact integrates gamified functionalities to deliver a series of cognitive assessments derived from the widely used Montreal Cognitive Assessment (MoCA). MoCA is a validated screening tool for detecting cognitive impairment, evaluating cognitive domains such as memory, attention, language, visuospatial skills, and executive functions. By incorporating these assessments into a gamified framework, BrainTest aims to enhance user engagement and compliance.

In addition, BrainTest employs Augmented Reality (AR) technology to assess motor functions related to mobility and spatial awareness. This is achieved through interactive features, such as virtual markers (pinpoints), which encourage users to perform specific movements and interact with their physical environment, as depicted in Figure 3.

Figure 3. Augmented Reality feature using virtual markups (pinpoints) to identify points for the user move towards to them. (Source: The Authors).

4.1 Web platform for Self-Concepts

To better understand users' perceptions of BrainTest, two web-based forms were developed to capture self-concepts related to the user experience. These forms are integrated into a platform designed to facilitate the registration and evaluation of artifacts. The platform also identifies key factors that contribute to the Semantic Differential Scale (SDS) used in self-concept questionnaires. Figure 4) illustrates an example of the semantic differential scale employed in the study.

	+++	++	+	0	+	++	+++	
Sad	○	○	○	○	○	○	○	Happy
Insecure	○	○	○	○	○	○	○	Confident
Unsatisfied	○	○	○	○	○	○	○	Satisfied
Dull	○	○	○	○	○	○	○	Excited
Dumb	○	○	○	○	○	○	○	Smart
Awkward_Weird	○	○	○	○	○	○	○	Normal
Frustrated	○	○	○	○	○	○	○	Rewarded
Stressed	○	○	○	○	○	○	○	Peaceful

Figure 4. Semantic-Differential Scale with Sentiments' List (Source: The Authors).

The platform supports the analysis of both ideal and real self-concepts, enabling the quantification of congruence between these constructs. By doing so, it provides actionable insights into user satisfaction and potential areas for artifact improvement.

4.2 Participant Demographics

The study engaged a group of seven participants aged 61 to 78, all of whom held higher education degrees. Participants were encouraged to engage in social and cultural activities within an academic environment. Ethical considerations were observed throughout the study; no sensitive data, such as names, social security numbers, or addresses, were collected. Only scores from the cognitive activities were recorded, ensuring privacy and compliance with ethical research guideline

4.3 Testing Experience and User Feedback

To document responses from the ideal and real evaluation questionnaires regarding the Augmented Reality (AR) application for cognitive assessment, participants accessed the dedicated web platform at `https://selfconceptsdesign.000webhostapp.com/`, using the access keys *eval* for the evaluation module and *land* for the product registration and questionnaire configuration module.

The testing experience provided valuable insights from the target audience, highlighting their engagement and adaptability. Despite anticipated challenges, users demonstrated enthusiasm and skill in navigating the application. The primary usability issues identified included accessibility challenges and confusion related to the AR feature. Specific feedback pointed to the small size of image buttons and the complexity of interacting with AR components. These observations were crucial in identifying areas for improvement, aiding in the refinement of the overall user experience.

4.4 Step-by-Step Application of the Test

1. Introduction and Consent: Participants were informed about the purpose of the study, and their informed consent was obtained prior to participation.
2. Initial Questionnaire (Ideal Self-Concept): Participants completed a questionnaire designed to capture their Ideal self-concepts concerning the use of the artifact. The Semantic Differential Scale (SDS) was employed with pairs of adjectives representing opposing attributes.
3. Interaction with BrainTest:
    - Participants engaged with the BrainTest application, completing cognitive tests and AR-based tasks.
    - Scores and performance metrics were recorded, ensuring no personal or sensitive information was collected.
4. Post-Interaction Questionnaire (Real Self-Concept): Following the interaction, participants completed a sec-

ond questionnaire to capture their Real self-concepts, again using the Semantic Differential Scale.

5. Data Collection and Storage: All collected data were anonymized and securely stored to ensure confidentiality and integrity.

4.5 Analysis of Self-Concepts

To assess the degree of user interest and adoption of the artifact, a series of bipolar adjective pairs (e.g., Insecure–Confident, Happy–Sad, Unsatisfied–Satisfied) was used. These pairs were analyzed through Ideal and Real self-concept questionnaires, providing a quantitative measure of user perceptions. The results, as presented in Figures 5 and 6, reveal insightful trends regarding user expectations and actual experiences.

The analysis showed that the Real perception (mean score: 3.58) exceeded the Ideal perception (mean score: 2.97). This suggests that participants' actual experiences with the artifact were generally more positive than their initial expectations. Individual results varied, with some evaluators displaying significantly higher Real perceptions than others. This variability indicates that personal factors, such as familiarity with technology or previous exposure to similar applications, may influence user engagement and satisfaction.

4.6 Validation of Results

This study assessed the applicability of the Self-Concepts methodology for evaluating user experience in a serious game designed for cognitive assessment. The participant group consisted of seven individuals aged between 61 and 78 years, predominantly female, all with higher education levels. This homogeneity ensured a consistent baseline in cognitive abilities and familiarity with digital interfaces, facilitating a focused evaluation of the game's design and usability.

The total performance score, derived as a cumulative measure of individual task scores, revealed distinct patterns in user performance. Participants consistently excelled in naming tasks, achieving the maximum score, and demonstrated strong performance in attention-related tasks, particularly those involving numbers. Conversely, tasks related to memory recall and ordering exhibited lower scores, indicating potential usability issues or task complexity. These findings suggest the need for targeted refinements in these areas to improve accessibility and engagement for future iterations.

Preliminary correlation analysis suggested meaningful relationships between specific task scores and overall performance. Attention-related tasks showed a likely positive correlation with total scores, highlighting attention as a foundational cognitive function that impacts performance across various domains. Conversely, tasks with low scores, such as memory recall and ordering, appeared less correlated with overall cognitive ability, pointing to task-specific design challenges rather than broader cognitive deficits. While the limited sample size restricts generalization, exploring gender-related trends in performance could yield further insights into task-specific engagement and usability disparities.

The analysis supported two key hypotheses. First, higher attention scores are expected to correlate with higher total scores, as attention plays a critical role in cognitive task performance. Second, memory-related scores likely contribute significantly to variability in total scores, given that memory tasks are often challenging for older populations. These hypotheses provide a foundation for future statistical validation using regression analysis or other hypothesis testing methods.

Although specific metrics for the augmented reality (AR) feature were not isolated, qualitative feedback highlighted its impact on user experience. Participants reported difficulties navigating the AR environment, particularly in interpreting visual markers and completing spatial tasks. Despite these usability challenges, participants' real self-concept scores exceeded their ideal self-concept scores, indicating that the artifact successfully engaged users and met their overall expectations.

Future studies should build on these findings by expanding the participant sample to enhance generalizability and addressing usability issues in memory recall and ordering tasks through iterative design improvements. Isolating metrics for AR-based tasks will also clarify their contribution to user engagement and task performance. Additionally, testing hypotheses regarding the relationship between specific task scores and overall performance will further validate the methodology's applicability.

The findings of this study demonstrate that the Self-Concepts methodology is a valuable tool for evaluating serious games, particularly in healthcare contexts. By addressing identified challenges and refining the approach in future iterations, this methodology has the potential to significantly improve the design and adoption of serious games for diverse healthcare applications.

4.6.1 Potential Applicability of the Approach

Although the dataset is limited, the methodology demonstrates significant potential for broader application:

- Scalability for Larger Studies: The use of Semantic Differential Scales and the analysis of Ideal vs. Real self-concepts provide a scalable framework for evaluating digital artifacts. Future studies with larger participant groups could refine the findings and validate the robustness of the approach.
- Design Optimization: The insights gained from self-concept analysis can guide iterative design processes. For instance, understanding discrepancies between Ideal and Real perceptions enables developers to prioritize enhancements that align with user expectations.
- Generalization Across Domains: While this study focused on a cognitive assessment application, the approach can be applied to a variety of digital artifacts, including educational tools, serious games, and health applications. The methodology's adaptability ensures its relevance across diverse user contexts.
- Integration with Advanced Analytics: Combining self-concept data with behavioral analytics, such as task completion rates and time spent on activities, could pro-

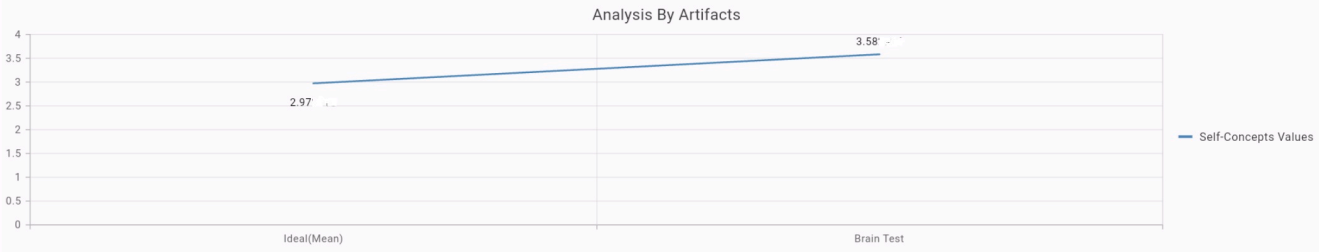

Figure 5. Performance analysis by artifact (Source: The authors).

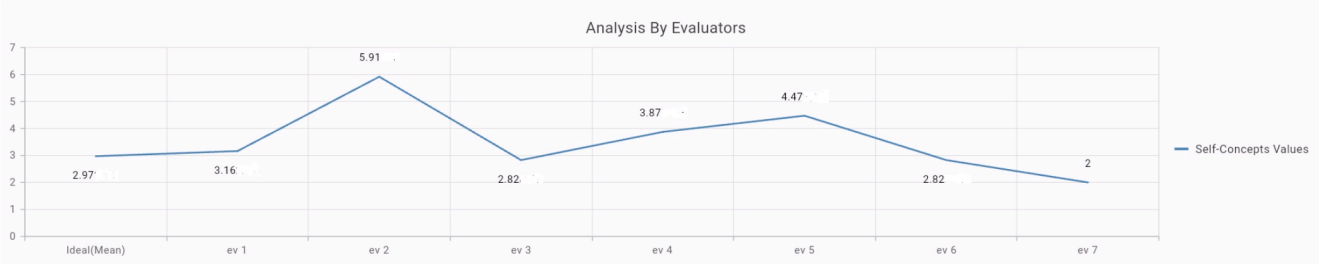

Figure 6. Performance analysis by evaluators (Source: The authors).

vide a more comprehensive understanding of user engagement. This integrated approach would enhance the predictive power of the methodology for user adoption.

4.6.2 Limitations and Future Directions

The primary limitation of this study is the small sample size, which may affect the generalizability of the results. Additionally, while the self-concept questionnaires provided valuable insights, integrating complementary methods, such as interviews or usability testing, could enrich the data and provide a deeper understanding of user experiences. Future research should focus on:

- Expanding the sample size to include a more diverse demographic.
- Exploring longitudinal studies to assess how user perceptions evolve over time.
- Investigating the impact of specific features, such as AR, on user engagement and satisfaction.

4.7 Data Transparency and Reproducibility

To ensure transparency and facilitate future research, the collected data has been made available through the following link: http://dx.doi.org/10.13140/RG.2.2.17988.95360. Researchers interested in accessing additional data for further analysis are encouraged to contact the authors directly. The repository (csv file available online) includes:

- Raw Data: Scores from cognitive tests, logs of AR feature usage, and Semantic Differential Scale responses.
- Questionnaires: Copies of the Ideal and Real self-concept questionnaires used in the study.

This approach promotes collaborative research while maintaining ethical data-sharing practices.

4.8 Ethical Considerations

This study adhered to rigorous ethical standards to ensure the protection and welfare of all participants. Participants were fully informed about the study's purpose, procedures, and their role in the research. Informed consent was obtained from each participant before their involvement. To safeguard participants' privacy, no sensitive personal data were collected; only anonymized data related to activity scores were recorded. Participation was entirely voluntary, and participants were assured of their right to withdraw from the study at any time without repercussions. All collected data were securely stored and accessible only to the research team, ensuring compliance with data protection protocols. The authors acknowledge the importance of submitting research involving human participants to an ethics committee for review. Such oversight ensures that ethical guidelines are met, and participants' rights are protected. In anticipation of future research phases involving larger participant groups, the following actions are planned:

- Ethics Committee Submission: The research protocol will be submitted to an ethics committee for formal review. This submission will include a comprehensive description of the study, detailed procedures, and the measures taken to protect participant welfare.
- Protocol Adjustments: Based on feedback and recommendations from the ethics committee, the research protocol will be revised as necessary to ensure full compliance with ethical standards.
- Adherence to Ethical Standards: All future phases of the research will strictly adhere to the highest ethical standards, with an emphasis on protecting participants' privacy and rights at every stage of the study.

By implementing these measures, the research team reaffirms its commitment to maintaining the integrity of the research process and upholding ethical principles that safeguard participant welfare. These steps are essential for ensuring transparency, accountability, and the validity of the research outcomes.

5 Conclusion

The gaming industry has experienced substantial diversification and growth in recent years, with serious games gaining prominence due to their wide-ranging applications across various domains. Within this industry, the development of tools and methodologies to evaluate user expectations and adoption potential among diverse user groups is critical. The Self-Concepts methodology, utilizing the Semantic Differential Scale, emerges as a valuable tool for User Experience (UX) design. This approach effectively measures users' ideal expectations before engaging with a game and their real perceptions after interacting with it, offering insights into user satisfaction and alignment with their initial expectations.

This study demonstrates the potential of the Self-Concepts methodology as a robust framework for evaluating user experience in serious games, particularly within healthcare contexts. The methodology, which employs Semantic Differential Scales to capture both ideal and real self-concepts, provides valuable insights into user expectations and their alignment with actual experiences. The findings highlight its efficacy in identifying strengths and areas for improvement in game design, ensuring that serious games can better meet the needs of target audiences.

Through an experimental case study involving elderly participants, this research validated the use of a serious game designed for cognitive assessment using the Montreal Cognitive Assessment (MoCA) framework. Despite the challenges observed in tasks related to memory recall and augmented reality (AR) features, the overall user experience was positive, with participants' real self-concept scores exceeding their ideal expectations. These results underscore the game's capacity to engage users and support its applicability in healthcare-focused serious games.

The study also revealed important performance trends. Tasks related to naming and attention showed consistently high scores, while memory and ordering tasks highlighted usability challenges. These findings indicate the need for iterative improvements in task design to enhance accessibility and engagement, particularly for older populations with varying levels of digital proficiency. Furthermore, the identification of potential correlations between attention scores and overall performance highlights key areas for further exploration in future studies.

The implications of this research extend beyond the validation of the artifact. The Self-Concepts methodology offers a scalable approach for assessing user satisfaction and adoption across diverse serious games. Future research should focus on expanding participant demographics to ensure broader generalizability, refining tasks to address usability issues, and isolating metrics specific to AR features to understand their role in user engagement better. Additionally, statistical validation of hypothesized relationships between task-specific scores and total performance will strengthen the methodology's applicability.

In conclusion, the Self-Concepts methodology has the potential to significantly advance the design and evaluation of serious games in healthcare. By bridging the gap between user expectations and real-world performance, it provides actionable insights for improving user engagement and satisfaction. These contributions position the methodology as a valuable tool for fostering innovation and effectiveness in serious games, particularly for applications that require user-centered designs in healthcare and beyond.

Acknowledgements

We thank CNPq for its financial support for this research.

References

Aguiar, B., Correia, W., and Campos, F. (2011). Uso da escala likert na análise de jogos. *Salvador: SBC-Proceedings of SBGames Anais*, 7(2).

Aini, N. and Ferdinand, A. (2021). Self-congruence theory: Factors affecting brand loyalty in fast-moving consumer goods industry. *Jurnal Inovasi Ekonomi*, 7:1–12. DOI: https://doi.org/10.22219/jiko.v7i01.18111.

Amaral, L. A. d., Fortes, R. P. d. M., Chagas, M. H. N., Castro, P. C., Bittar, T. J., and Lobato, L. L. (2019). Towards adopting serious games for cognitive tests with brazilian older people. *Journal on Interactive Systems*, 10(1). DOI: https://doi.org/10.5753/jis.2019.715.

Anolli, L., Mantovani, F., Confalonieri, L., Ascolese, A., and Peveri, L. (2010). Emotions in serious games: From experience to assessment. *International Journal of Emerging Technologies in Learning (iJET)*, 5:3–11.

Bernhaupt, R., Ijsselsteijn, W., Mueller, F., Tscheligi, M., and Wixon, D. (2008). Evaluating user experiences in games. In *CHI'08 extended abstracts on Human factors in computing systems*, pages 3905–3908.

Borges, B., Sidarta, I., de Souza, A., Coelho, B., and Darin, T. (2019). Experiência do usuário em jogos digitais: Uma catalogação de instrumentos de avaliação. In *Anais do I Workshop sobre Interação e Pesquisa de Usuários no Desenvolvimento de Jogos*, pages 69–78, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/wiplay.2019.7841.

Carvalho, A. P. d., Braz, C. S., and Prates, R. O. (2024). An analysis of the evaluation methods being applied to serious games for autistic children. *Journal on Interactive Systems*, 15(1):55–78. DOI: https://doi.org/10.5753/jis.2024.3288.

Dutra, T., Tondorf, D., Zils, T., Ferreira, A., Gasparini, I., Hounsell, M., and Maschio, E. (2021). Métodos de avaliação de ihc no contexto de jogos sérios educacionais: Um mapeamento sistemático. In *Anais do XXXII Simpósio Brasileiro de Informática na Educação*, pages 564–575, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/sbie.2021.217450.

Fellows, R., Braga, J., and Dotta, S. (2022). Revisão sistemática de literatura sobre métodos, técnicas e critérios de avaliação de aprendizagem em jogos sérios sociointeracionistas. In *Anais do XXXIII Simpósio Brasileiro de Informática na Educação*, pages 528–539, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/sbie.2022.225236.

Fokides, E., Atsikpasi, P., Kaimara, P., and Deliyannis, I. (2019). Let players evaluate serious games. design and validation of the serious games evaluation scale. *ICGA Journal*, 41(3):116–137. DOI: https://doi.org/10.3233/ICG-190111.

Friborg, O., Martinussen, M., and Rosenvinge, J. H. (2006). Likert-based vs. semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience. *Personality and Individual Differences*, 40(5):873–884. DOI: https://doi.org/10.1016/j.paid.2005.08.015.

Funabashi, A. M. M., Aranha, R. V., Silva, T. D., Monteiro, C., Silva, W. S., and Nunes, F. (2018). A serious game for virtual rehabilitation: evaluation with patients and physiotherapists. *Journal on Interactive Systems*, 9(2). DOI: https://doi.org/10.5753/jis.2018.698.

Garcia, M. (2023). A systematic review of serious games for health education: Technology acceptance and design considerations. *Publication Platform*. DOI: https://doi.org/10.4018/979-8-3693-3661-8.ch002.

García-Martínez, I., Augusto-Landa, J. M., Quijano-López, R., and León, S. P. (2022). Self-concept as a mediator of the relation between university students' resilience and academic achievement. *Frontiers in Psychology*, 12. DOI: https://doi.org/10.3389/fpsyg.2021.747168.

Gazis, A. and Katsiri, E. (2023). Serious games in digital gaming: A comprehensive review of applications, game engines and advancements. *WSEAS Transactions on Computer Research*, 11:10–22. DOI: https://doi.org/10.37394/232018.2023.11.2.

Hosany, S. and Martin, D. (2012). Self-image congruence in consumer behavior. *Journal of Business Research*, 65(5):685–691. DOI: https://doi.org/10.1016/j.jbusres.2011.03.015.

Kim, S. and Jung, Y. (2023a). Development of semantic differential scales for artificial intelligence agents. *International Journal of Social Robotics*, 15(5):1155–1167.

Kim, S. and Jung, Y. (2023b). Development of semantic differential scales for artificial intelligence agents. *International Journal of Social Robotics*, 15(5):1155–1167. DOI: https://doi.org/10.1007/s12369-023-01010-3.

Lin, C., Ren, Y., and Lu, A. (2023). The effectiveness of virtual reality games in improving cognition, mobility, and emotion in elderly post-stroke patients: a systematic review and meta-analysis. In *Neurosurgical review*, volume 46, page 167. DOI: https://doi.org/10.1007/s10143-023-02061-w.

McMullen, T. (2020). *Self-Concept*, pages 1978–1979. Springer International Publishing, Cham. DOI: https://doi.org/10.1007/978-3-030-39903-0_1497.

Michailidis, L., Balaguer-Ballester, E., and He, X. (2018). Flow and immersion in video games: The aftermath of a conceptual challenge. *Frontiers in Psychology*, 9:1682. DOI: https://doi.org/10.3389/fpsyg.2018.01682.

Nacke, L. E. (2015). Games user research and physiological game evaluation. In *Game user experience evaluation*, pages 63–86. Springer.

Nakamura, T., Miyata, K., Yamamoto, H., and Sato, H. (2021). A study on scale construction of adjective pairs for evaluating audiovisual effects in video games. In *Design, User Experience, and Usability: UX Research and Design*, pages 278–291. Springer.

Neves, C. E. (2021). Self-concept as an artifact evaluation method: generating evidence of the use of self-concept as an artifact evaluation method trhough the application in feature films. Master's thesis, Universidade Federal de Pernambuco.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological bulletin*, 49(3):197. DOI: https://doi.org/10.1037/h0055737.

Ploder, A. and Eder, A. (2015). Semantic differential. In Wright, J. D., editor, *International Encyclopedia of the Social Behavioral Sciences (Second Edition)*, pages 563–571. Elsevier, Oxford, second edition edition. DOI: https://doi.org/10.1016/B978-0-08-097086-8.03231-1.

Statista (2023). Video games - worldwide. Last accessed: 2024-12-23. https://www.statista.com/outlook/dmo/digital-media/video-games/worldwide.

Stoklasa, J., Zvarikova, K., and Poliak, M. (2019). Semantic differential for the twenty-first century: scale relevance and methodological improvements. *Quality & Quantity*, 53(1):435–448.

Suryapranata, L. K. P., Soewito, B., Kusuma, G. P., Gaol, F. L., and Warnars, H. L. H. S. (2017). Quality measurement for serious games. In *2017 International Conference on Applied Computer and Communication Technologies (ComCom)*, pages 1–4. IEEE. DOI: https://doi.org/10.1109/COMCOM.2017.8167098.

Trautwein, U. and Möller, J. (2016). *Self-Concept: Determinants and Consequences of Academic Self-Concept in School Contexts*, pages 187–214. DOI: https://doi.org/10.1007/978-3-319-28606-8_8.

Van Harreveld, F., Nohlen, H. U., and Schneider, I. K. (2015). The abc of ambivalence: Affective, behavioral, and cognitive consequences of attitudinal conflict. In *Advances in experimental social psychology*, volume 52, pages 285–324. Elsevier. DOI: https://doi.org/10.1016/bs.aesp.2015.01.002.