# Evaluating a Requirements Engineering Framework for Serious Games for Health

Bruna M. Barbosa
Institute of Informatics, Federal University of Goiás (UFG)
Goiânia, Brazil
bruna_mota@discente.ufg.br

Carlos Henrique R. Souza
Institute for Innovation in Management, Federal University of Goiás (UFG)
Cidade Ocidental, Brazil
carlosrorato@ufg.br

Halleffy H. Reis Santos
Institute of Informatics, Federal University of Goiás (UFG)
Goiânia, Brazil
HalleffySantos@gmail.com

Pedro M. R. G. Silva
Institute of Informatics, Federal University of Goiás (UFG)
Goiânia, Brazil
pedro_ribeiro@discente.ufg.br

Luciana O. Berretta
Institute of Informatics, Federal University of Goiás (UFG)
Goiânia, Brazil
luciana.berretta@ufg.br

Sergio T. Carvalho
Institute of Informatics, Federal University of Goiás (UFG)
Goiânia, Brazil
sergiocarvalho@ufg.br

## Abstract

The development of Serious Games for Health (SGH) requires balancing playful engagement with clinical effectiveness. However, many approaches rely on ad hoc practices that lack governance, traceability, and integration between clinical knowledge and game design. This paper evaluates a Requirements Engineering (RE) framework designed to support SGH development by translating clinical needs into auditable and traceable technical specifications. The framework is based on the People-Processes-Products triad and integrates RE practices with process maturity. We conducted a case study involving adults with Autism Spectrum Disorder (ASD) supported by a multidisciplinary team to evaluate the framework in practice. The process employs a Requirements Traceability Matrix (RTM) to connect clinical rules, implementation decisions, and individualized Key Performance Indicators (KPIs). The results indicate that the framework supports the systematic translation of clinical requirements into implemented game mechanics while preserving traceability throughout the development process. The use of structured artifacts, including the RTM and ScriptableObjects, enables traceability across development stages, supports the analysis of implementation decisions in relation to clinical objectives, and facilitates communication among multidisciplinary stakeholders. The findings provide evidence of the framework's applicability and suggest that traceability mechanisms contribute to a more structured and repeatable development process for serious games.

## Keywords

Requirements Engineering, Serious Games for Health, Case Study, Autism Spectrum Disorder.

## 1 Introduction

Serious games (SG) act as multifaceted intervention tools that extend beyond entertainment by leveraging game infrastructure such as narrative, rewards, and interactivity to achieve specific objectives [20]. In healthcare, developers apply these games across a wide range of contexts, including physical rehabilitation and exergaming, cognitive training, mental health support, and professional medical simulation [6, 10, 13]. By transforming repetitive or exhausting treatments into motivating activities, these systems create safe, controlled environments for skill development while enabling healthcare professionals to monitor treatment outcomes quantitatively.

The development of Serious Games for Health (SGH) poses an engineering challenge that requires balancing playful engagement with clinical effectiveness [7]. While the entertainment industry has established frameworks that support immersive experiences, researchers and developers often find that directly applying these approaches to the health domain is insufficient [19]. To achieve its objectives, SGH must operate as benefit-delivery systems in which ludic mechanics coexist with scientific rigor and therapeutic fidelity [6, 16]. Achieving this balance demands engineering a complete, cohesive game experience rather than merely applying gamification elements to existing clinical tasks.

Despite these advances, the current landscape reveals a methodological gap. Existing research shows that many groups develop health-oriented games using ad hoc or empirical approaches and rely on isolated methodologies tailored to specific clinical scenarios or technologies, without formal governance or traceability [5, 13, 14]. This fragmentation often creates a communication gap between healthcare professionals and software engineers, limiting the reuse of components and the transfer of clinical knowledge across different medical contexts [2, 8, 9, 12]. As a consequence, SGH projects often lack structured engineering processes that ensure requirements consistency, process repeatability, and end-to-end traceability throughout development.

To address this gap, this paper evaluates a Requirements Engineering (RE) framework whose theoretical conceptualization was established in previous research [4]. The present work expands upon that foundation by providing its first practical validation and methodological detailing. Specifically, it operationalizes the framework by systematically generating engineering artifacts, including the Software Requirements Specification (SRS), the Serious-GDD [17], and a complete Requirements Traceability Matrix (RTM), demonstrating how clinical requirements are translated into auditable and traceable technical specifications. The paper further analyzes how these artifacts support requirements traceability, consistency, communication among multidisciplinary stakeholders,

and alignment between clinical objectives and implemented game mechanics throughout the development process.

To evaluate the framework in practice, we conducted a case study involving the development of a vertical-slice prototype for adults with Autism Spectrum Disorder (ASD). The case study evaluates the proposed engineering process by examining how clinical knowledge is systematically translated into software artifacts and implementation decisions. The clinical effectiveness of the resulting game is outside the scope of this study. The results indicate that the proposed framework, through its traceability mechanisms, preserves clinical intent, reduces communication barriers among stakeholders, and provides a measurable foundation for validating the implemented mechanisms using individualized Key Performance Indicators (KPIs).

This paper is organized as follows. Section 2 presents the theoretical background and positions this work within the literature. Section 3 describes the architecture of the proposed framework, based on the triad of People, Processes, and Products. Section 4 details the case study methodology involving adults with ASD, including ethical protocols and stakeholder roles. Section 5 presents and discusses the results of applying the framework, with a focus on traceability, engineering artifacts, and process characteristics. Finally, Section 6 concludes the paper and outlines directions for future research.

## 2 Background

The development of SGH requires a convergence of Software Engineering (SE) and Health sciences to ensure digital interventions are both clinically safe and entertaining [7]. Terminological and conceptual differences between medical experts and software engineers often lead to misinterpretations of clinical protocols [16]. To mitigate this, recent literature emphasizes the need to move beyond ad hoc approaches and transition to disciplined engineering processes that treat clinical guidelines as foundational system rules [6].

The current state of the art presents various methodologies aimed at bridging this interdisciplinary gap, primarily through adaptations of Requirements Engineering (RE). Frameworks such as the MECHA model seek to map behavior change techniques directly to specific game mechanics, providing a theoretical foundation for clinical game design [12]. On a procedural level, approaches such as CaSEJ utilize visual canvases for collaborative requirements elicitation [14], while Agile Storyboarding serves as a visual contract between clinicians and developers to represent physical constraints before coding begins [10]. Furthermore, studies focusing on interactive floors [15] and wheelchair mobility [11] demonstrate the necessity of participatory design and iterative validation with domain specialists.

Despite these valuable contributions, systematic reviews indicate that most approaches remain strongly tied to specific clinical niches or lack formal process maturity [1, 2, 5, 8]. A significant gap remains regarding engineering processes that provide end-to-end traceability, governance, and repeatability throughout the software lifecycle. While existing models facilitate ideation and participatory design, they frequently omit the systematic production of auditable engineering artifacts, such as the SRS and the RTM, limiting process consistency, maintainability, and project scalability.

Table 1 compares representative SGH development approaches in terms of their RE practices, adherence to process maturity models, and central focus.

**Table 1: Comparison of representative SGH development approaches.**

| Work | Domain | RE | Main Focus |
|---|---|---|---|
| [15] | ASD | Yes | Hardware and Software Requirements for interactive floors. |
| [11] | Motricity | Yes | Alignment between clinical protocols and User Experience (UX). |
| [14] | General Health | Yes | Collaborative brainstorming via the CaSEJ Canvas. |
| [6] | General Health | Yes | Design reusability and personalization modules. |
| [12] | Behavior | Partial | Behaviour Change Wheel into game design. |
| [3] | ASD | No | Ad-hoc development focused on theoretical grounding (DiagnosTEA). |
| **Proposed Framework** | **General Health** | **Yes** | **Methodological rigor, quality, and repeatability.** |

Addressing the limitations highlighted in Table 1, the framework evaluated in this study adopts a software engineering perspective grounded in the Requirements Engineering process defined by Wiegers and Beatty [18]. By integrating a disciplined RE lifecycle with standardized engineering artifacts, including the SRS, Serious-GDD, and an explicit RTM, the framework establishes bidirectional traceability from high-level clinical objectives to technical implementation and validation. This process-oriented approach improves communication among multidisciplinary stakeholders while supporting governance, repeatability, and maintainability throughout SGH development.

## 3 Framework Overview

The framework provides a structured methodological guide that facilitates the communication between healthcare specialists and software developers. The framework establishes a repeatable engineering workflow for SGH development, replacing ad hoc creative development practices. By integrating the iterative Requirements Engineering (RE) lifecycle defined by [18], the architecture supports the convergence of multidisciplinary perspectives and the systematic production of auditable software engineering artifacts throughout development. The framework, illustrated in Figure 1, is structured around three interdependent pillars: People (Stakeholders), Processes (RE Lifecycle), and Products (Documentation and Traceability Artifacts), providing governance and end-to-end traceability across the software development lifecycle. Figure 1 depicts the complete engineering workflow and the information flow among stakeholders, engineering activities, and software artifacts.

### 3.1 Multidisciplinary Stakeholder Architecture

To prevent the loss of clinical intent and ensure a highly engaging experience, the framework relies on a modular stakeholder architecture that remains constant across different health domains. It establishes a collaborative dynamic to mediate the translation between clinical rules and interactive entertainment:

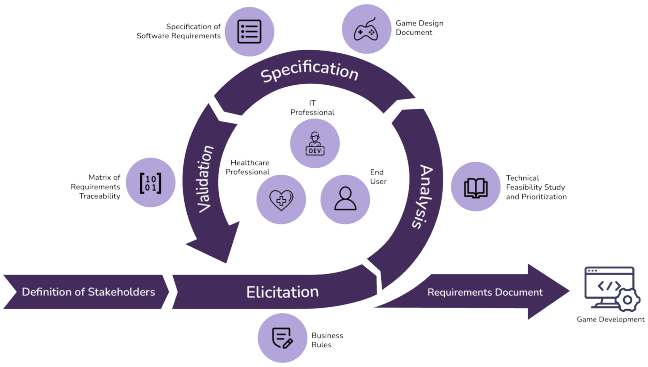


**Figure 1: Operational Architecture of the Requirements Engineering framework for SGH development.**

- **Healthcare Professional (Domain Authority):** Defines the business rules, therapeutic objectives, and non-negotiable safety protocols. They act as the primary source of clinical evidence.
- **End User (Patient):** Included as an active stakeholder from the outset to define preferences, accessibility needs, and sensory barriers, ensuring the system is tailored to the target demographic's reality.
- **IT Professional (Technical and Experience Architect):** Responsible for the system's software architecture, evaluating technological feasibility, and defining functional and non-functional requirements. In the context of SGH, this professional acts as both the software engineer and the game designer. Once the clinical base is validated, they structure the playful layer, translating clinical constraints into engaging mechanics, dynamics, and aesthetics.

Although these roles define clear responsibilities and decision-making authority, development teams remain multidisciplinary and collaborative. Multiple professionals may contribute to each role, requiring continuous interaction between health and technology specialists.

## 3.2 The Requirements Lifecycle and Specification Formats

The framework follows an iterative and incremental RE lifecycle organized into four stages: Elicitation, Analysis, Specification, and Validation. This structure introduces a bidirectional validation mechanism to maintain consistency between clinical objectives and implemented features. The phases are iterative by design, allowing the workflow to function as a continuous loop that returns to earlier stages for refinement, as shown in the framework's architecture, while preserving clinical intent throughout development.

During Elicitation, the team captures raw clinical evidence and user feedback to generate high-level Business Rules (BR) and Stakeholder Requirements. This is immediately followed by the Analysis phase, in which the IT Professional serves as a technical filter. In this stage, the team produces a Technical Feasibility Study to evaluate technological constraints, define the executable scope, and prevent scope creep, ensuring that only viable clinical demands advance to the design phase.

The Specification phase then formalizes the project by synchronizing two core artifacts: the SRS and the Serious Game Design Document (Serious-GDD). To improve consistency and reduce ambiguity, it is suggested that the specification relies on standardized textual formats rather than free-form descriptions. For example, teams can adopt structured textual requirements using formal syntax and unique identifiers (e.g., STK-REQ-01 for Stakeholder Requirements, RN-01 for Business Rules, and RF-01 for Functional Requirements) to facilitate automated tracking. Furthermore, functional requirements can be broken down into user stories and acceptance criteria (e.g., "As a patient with ASD, I want visual indicators of emotion so that I can understand the NPC's intent without ambiguity"). While the SRS establishes a holistic product view detailing technical capabilities, the Serious-GDD translates these requirements into technical mechanics and playful sensations (e.g., a "Social Thermometer" UI).

Finally, Validation concludes the increment. In this stage, the multidisciplinary team evaluates the implemented features against the predefined acceptance criteria and KPIs, thereby transforming subjective feedback into quantitative data to support informed decision-making. If a requirement is found to be clinically valid but technically infeasible or insufficiently engaging during this phase, the iterative loop forces the team to return to the Elicitation or Analysis stages to adjust the design while preserving the clinical goals. This iterative workflow also ensures that every implementation decision remains explicitly connected to its originating requirement and corresponding validation artifact, preserving consistency throughout the development lifecycle.

## 3.3 Artifacts and Traceability Analysis

To elevate SGH development from an ad hoc activity to a mature engineering process, the framework institutionalizes software quality through a standardized documentation baseline. This centralized artifact acts as a comprehensive product contract, seamlessly integrating clinical objectives with engineering goals. It demonstrates how the conceptual framework is materialized into auditable engineering artifacts, ensuring that technical implementations do not deviate from medical protocols. The framework mandates the progressive construction of a unified Requirements Document structured into nine interdependent modules. This model ensures that no code is written without prior clinical justification:

(1) **Stakeholder Profiles:** Formalizes the multidisciplinary triad (Healthcare Professional, IT Professional, and End User) and defines responsibilities to prevent validation gaps.
(2) **Business Requirements:** Defines the high-level strategic and operational goals of the software (e.g., cost reduction, scalability).
(3) **Stakeholder Requirements:** Captures the direct needs, pain points, and expectations of the users and clinicians, elicited before any technical discussion.
(4) **Business Rules (Clinical Protocols):** Establishes the clinical rules and safety limits that govern system behavior. These are non-negotiable constraints translating medical knowledge into logical sentences.

(5) **Technical Feasibility and Prioritization:** Evaluates technological feasibility, defines the project's actual scope based on available resources, preventing scope creep and ensuring executable boundaries.
(6) **Serious Game Design Document (Serious-GDD):** Maps the ludic elements (mechanics, dynamics, aesthetics) directly linked to health objectives.
(7) **Software Requirements Specification (SRS):** Details the Functional and Non-Functional Requirements necessary for implementation.
(8) **Acceptance Criteria (KPIs):** Defines quantifiable metrics to validate whether the implemented system successfully meets the clinical and technical baselines.
(9) **Requirements Traceability Matrix (RTM):** Establishes a bidirectional map linking the clinical origin of a requirement to its technical component and validation metric.

To prevent communication gaps, the framework avoids isolated documentation. Instead, it synchronizes the SRS with the Serious-GDD. While the SRS details the system's technical boundaries and constraints, the Serious-GDD translates these boundaries into ludic experiences and mechanics. A significant risk in SGH development is the loss of clinical fidelity when projects are transferred between different multidisciplinary teams or third-party developers who lack medical backgrounds. The framework ensures autonomy and mitigates the risk of this protocol breach by treating the RTM as a standalone technical contract, supported by strict ID standardization.

During the handover process, the receiving team not only receives the source code but also inherits the complete documentation baseline, in which traceability IDs are directly embedded in the game engine's architecture (e.g., variable names or inspector tags within Unity). For example, if a third-party developer needs to alter the Social Filter System dialogue tree, the engine's interface explicitly references a documentation requirement. Before modifying the code, the developer must cross-reference the RTM, verify that the requirement is strictly bound to RN (Denotative feedback) and validate it against KPI.

This closed-loop methodology minimizes the need for direct supervision from the original healthcare professionals during technical updates. The standardized artifacts provide the necessary guardrails, ensuring that any external team can safely scale or maintain the serious game without distorting the original clinical intent or compromising the statistical validity of patients' reports. Consequently, the framework supports software evolution while preserving the consistency between clinical requirements, implementation decisions, and validation criteria.

## 4 Case Study Methodology

This case study evaluates the proposed Requirements Engineering (RE) framework in practice by developing a serious game for adults with Autism Spectrum Disorder (ASD)[1]. The evaluation focuses on how the proposed engineering process systematically transforms clinical knowledge into consistent, traceable, and auditable software artifacts. The resulting prototype serves as the vehicle for evaluating the framework.

To execute the proposed engineering process, the project established a multidisciplinary volunteer team comprising representatives from each stakeholder group defined by the framework.

- **UF-01 (End User):** An adult diagnosed with ASD and with a high level of digital literacy, representing the patient's voice to confirm accessibility and engagement.
- **PS-01 (Healthcare Professional):** A psychiatrist specializing in neurodevelopment, focused on identifying clinical objectives and support for digital interventions.
- **PS-02 (Healthcare Professional):** A neuropsychologist specialized in cognitive evaluation and social movement, defining the rules for pragmatic communication and metaphor interpretation.
- **PS-03 (Healthcare Professional):** A psychiatrist focused on behavioral modeling, scenario predictability, and preventing infantilization in adults.
- **DEV-01 (IT Professional):** Researcher and developer responsible for technical architecture, requirements engineering, and implementation.
- **DEV-02 (IT Professional):** Software engineer responsible for auditing and validating the generated artifacts.

The execution of the framework was analyzed across its four iterative phases to examine how each stage contributes to requirements refinement, artifact generation, traceability, and validation throughout the development lifecycle.

**Phase 1 - Elicitation:** Utilized structured interviews to identify clinical guidelines and user barriers. UF-01 identified sensory triggers and the need for objective indicators of intent to reduce social "paralysis". PS-01 established the need for multimodal synchrony (speech, face, and text). PS-02 proposed training for "social filters" and metaphor interpretation. PS-03 emphasized scenario predictability and the elimination of punitive "Game Over" states. This phase established the initial documentation baseline by producing Business Rules (BR) and Stakeholder Requirements, which served as the foundation for all subsequent engineering artifacts.

**Phase 2 - Analysis:** DEV-01 evaluated clinical demands for technical feasibility. This phase filtered requirements to ensure they were executable within the Unity engine, for example, by replacing real video assets with controlled digital animations to maintain frame stability and prevent sensory overload. The outcome of this phase was a refined and technically consistent requirements baseline. DEV-02 audited the artifacts and identified missing traceability details required to support third-party development. Based on this process, the team developed the Technical Feasibility Study and prioritized functionalities, such as the "Social Thermometer" and "Loop of Correction", based on impact and implementation effort.

**Phase 3 - Specification:** DEV-01 formalized the engineering baseline for implementation. DEV-01 generated the SRS, detailing technical requirements like telemetry and audio control, and the Serious Game Design Document (Serious-GDD) [17], which translates clinical needs into narrative and procedural mechanics. Table 2 presents a simplified template fragment that demonstrates how these clinical rules are mapped to both technical and playful documentation before any code is written. These artifacts collectively

[1]This study was approved by the Ethics Committee of the participating institution. The details of this approval can be found on the Brazil Platform (plataformabrasil.saude.gov.br) under the Certificate of Presentation for Ethical Review (CAAE) number 8442725.6.0000.5083, with approval opinion number 7.615.894.

established the documentation baseline used throughout implementation and validation.

**Table 2: Simplified Documentation Baseline Template**

| Dimension | Identifier | Description |
|---|---|---|
| **Clinical Origin** | STK-REQ-03 RN-01 | **Pedagogical Feedback:** The user requires immediate, clear explanations. **Business Rule:** Prohibition of punitive "Game Over" states. |
| **Technical Spec. (SRS)** | RF-02 | **Functional Req.:** The system must implement modular progression allowing an immediate retry loop upon error. |
| **Design Spec. (Serious-GDD)** | GDD-3.3 | **Procedural Dimension:** Implementation of a "Correction Loop" that pauses gameplay to provide explicit feedback without resetting progress. |

**Phase 4 - Validation:** Focused on validating both the implemented prototype and the engineering artifacts produced throughout development. This phase generated a set of KPIs established by the clinical stakeholders (PS-01, PS-02, PS-03) and UF-01, including the accuracy rate for NPC intention interpretation (KPI-01) and session engagement time (KPI-03). These KPIs also allowed the multidisciplinary team to verify that the engineering artifacts consistently upheld the intended clinical objectives throughout implementation. The RTM linked each implemented mechanic to its clinical origin and corresponding validation KPI. Table 3 illustrates the core structure of the RTM applied to the case study.

## 5 Framework Instantiation: A Vertical Slice

To validate the practical applicability of the RE framework, a functional prototype was developed as a vertical slice. The prototype was built in the Unity Engine[2], using C# to implement procedural logic. This resulting prototype serves as a proof of concept for the proposed RE framework, demonstrating how clinical requirements elicited during the early phases are systematically translated into executable software components while preserving traceability to the corresponding engineering artifacts.

The resulting artifact is a narrative simulator designed to train social and pragmatic skills for adults with ASD. Set in a highly predictable simulated office environment, the game presents players with everyday workplace interactions and social dilemmas. Players engage in branching dialogues with Non-Player Characters (NPCs) and must navigate figurative language, sarcasm, and social nuances. To ensure cognitive accessibility, the game features a Low-Poly aesthetic with a muted color palette to prevent sensory overload, alongside dynamic visual aids that translate subjective emotional cues into objective data. All implemented features, technical constraints, and ludic mechanics instantiate the requirements and traceability links previously established in Table 2 and Table 3.

[2]https://unity.com/

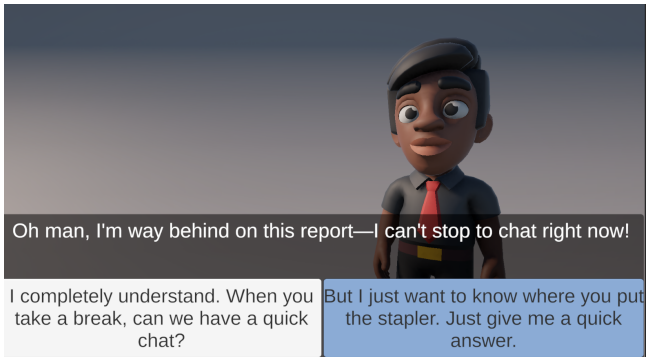


**Figure 2: Gameplay overview of the narrative simulator, showcasing the predictable office environment and the interactive dialogue interface designed for pragmatic training.**

Beyond the visual and interactive layer, a fundamental software engineering decision in this project was the strict separation of procedural game logic from clinical content. In SGH development, clinical protocols are frequently revised by healthcare professionals, making hardcoded dialogue or scenario logic highly inefficient and error-prone. To resolve this, the architecture leverages Unity's `ScriptableObjects` as data containers to structure the narrative and clinical scenarios. This isolation ensures that clinical guidelines such as the requirement to train social nuances (STK-REQ-05) using denotative feedback (RN-06) are treated as independent data assets rather than being embedded within C# execution scripts. Furthermore, the `DialogueNode` data structure allows developers to directly link each dialogue branch to its specific functional requirements, such as the dialogue-branching logic (RF-03), in the Unity Inspector. This explicit linkage, illustrated in Figure 3, ensures that every interaction presented to the player serves a documented clinical purpose, facilitating both technical maintenance and clinical auditing.

The framework mandates that non-negotiable clinical rules govern system behavior. For adults with Autism Spectrum Disorder (ASD), traditional punitive mechanics like a "Game Over" state can trigger frustration and anxiety, violating the principle of sensory safety. Consequently, the prototype replaces punitive states with a non-punitive "Correction Loop", directly fulfilling the clinical business rule that strictly prohibits "Game Over" states (RN-01) through modular progression (RF-02). When a user selects a literal, socially inappropriate dialogue option, the system pauses the narrative progression. Instead of terminating the session, the system triggers a pedagogical feedback panel (STK-REQ-03), explicitly detailing the social nuance of the situation and allowing for an immediate re-attempt. As seen in Figure 4, this mechanic directly translates the clinical requirement for modeling-based learning into an executable software flow without causing cognitive overload. This mechanic demonstrates how a documented clinical requirement can be consistently propagated from specification to executable behavior.

To transform the ludic experience into a legitimate diagnostic and therapeutic support tool, data collection must be transparent,

**Table 3: Requirements Traceability Matrix (RTM) Core Structure**

| Clinical Origin (STK-REQ / RN) | Technical Spec. (SRS / GDD) | Implemented Game Mechanic | Validation (KPI) |
|---|---|---|---|
| STK-REQ-01: Support for social reading; RN-04: Multimodal synchrony. | RF-01: Emotion indicator. GDD-3.2: Visual UI. | **Social Thermometer UI** linked to dynamic NPC facial expressions. | KPI-01: Accuracy rate in intention interpretation > 70%. |
| STK-REQ-03: Pedagogical feedback; RN-01: No "Game Over". | RF-02: Modular progression. GDD-3.3: No-punishment. | **Correction Loop:** Pauses gameplay upon an error to provide explicit feedback and allow a retry. | KPI-03: Session engagement time > 15 min. |
| STK-REQ-05: Train social nuances; RN-06: Denotative feedback. | RF-03: Dialogue branching. GDD-3.3: Semantic Sys. | **Social Filter System:** Dialogue choices distinguishing literal truth from social adequacy. | KPI-06: Accuracy in metaphor interpretation > 70%. |
| STK-REQ-04: Clinical monitoring; RN-07: Data integrity. | RF-04: Telemetry module. GDD-3.4: Data export. | **Automated Logs:** Extraction response latency and dialogue errors to CSV. | KPI-05: Integrity of the exported data. |

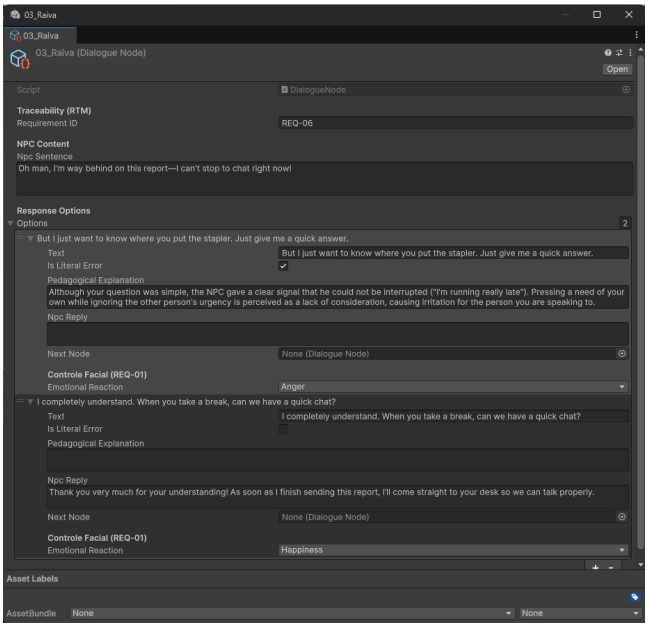


**Figure 3: Unity Inspector view of a DialogueNode showcasing the decoupling of clinical content and the direct requirement.**

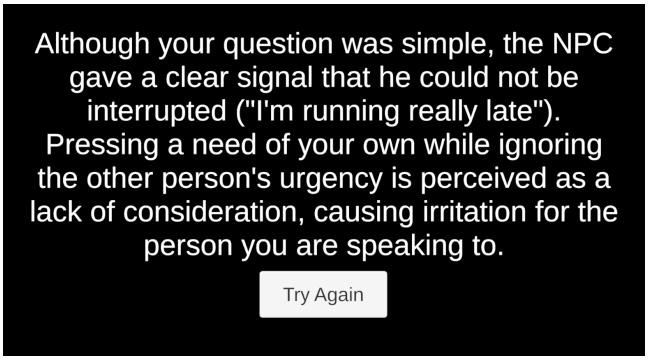


**Figure 4: Interface of the non-punitive Correction Loop displaying pedagogical explanations after a literal error.**

structured, and continuous to guarantee clinical monitoring (STK-REQ-04) and data integrity (RN-07). The architecture incorporates a dedicated telemetry module (RF-04) designed to capture clinical metadata without interfering with gameplay performance. The telemetry module continuously monitors user interactions across all scenes and dialogue nodes. The telemetry module consolidates performance data in real time, capturing critical metrics such as Response Latency (the hesitation time during social processing) and Accuracy Rates per specific dialogue node. At the conclusion of each session, the `TelemetryManager` exports the consolidated metrics to a .csv file. This structured output, shown in Figure 5, allows healthcare professionals to import the logs into external data analysis software, providing quantitative evidence for long-term clinical tracking across different simulated social environments. Because every collected metric remains linked to its originating requirements and corresponding KPIs through the RTM, the telemetry module also supports the validation of the engineering artifacts produced by the proposed RE framework.

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | Date | Patient | Scenario | ID_Requisite | NPC_Sentence | Result | Attempts |
| 2 | 2026-04-06 15:21 | Patient 01 | Scene 01 | STK-REQ-03 | So? What did you think of the draft I made? Be honest—I can handle the truth! | Needs Practice | 2 |
| 3 | 2026-04-06 15:21 | Patient 02 | Scene 01 | STK-REQ-03 | I think there are some very good points, but we can tweak a few details together to make it perfect. | Mastered | 1 |
| 4 | 2026-04-06 15:21 | Patient 03 | Scene 01 | STK-REQ-03 | I feel like I'm carrying the weight of the world on my shoulders this week. | Needs Practice | 2 |
| 5 | 2026-04-06 15:21 | Patient 04 | Scene 01 | STK-REQ-03 | It looks like you have a heavy workload. Is there anything I can help with? | Mastered | 1 |
| 6 | 2026-04-06 15:21 | Patient 05 | Scene 01 | STK-REQ-03 | Oh man, I'm way behind on this report; I can't stop to chat right now! | Needs Practice | 4 |
| 7 | 2026-04-06 15:21 | Patient 06 | Scene 01 | STK-REQ-03 | I totally understand. When you take a break, can we have a quick chat? | Mastered | 1 |
| 8 | 2026-04-06 15:22 | Patient 07 | Scene 02 | RN-01 | You did a great job on the report. You really nailed it! | Needs Practice | 5 |
| 9 | 2026-04-06 15:22 | Patient 08 | Scene 02 | RN-01 | We need this code ready by 5 PM. Time flies when we're at the end of a project! | Needs Practice | 5 |
| 10 | 2026-04-06 15:22 | Patient 09 | Scene 02 | RN-01 | Wow! Arriving 20 minutes late on a Monday... quite the start to the week, huh? | Needs Practice | 2 |
| 11 | 2026-04-06 15:22 | Patient 10 | Scene 02 | RN-01 | I'm really overwhelmed with these new requirements. Could you give me a hand with this documentation? | Mastered | 1 |
| 12 | 2026-04-06 15:22 | Patient 11 | Scene 02 | RN-01 | Your presentation was good, but I feel like you really dragged out the introduction with a lot of fluff. | Mastered | 1 |

**Figure 5: Clinical report generated by the TelemetryManager, logging patient performance metrics and response latency.**

## 6 Validation and Discussion

The case study evaluation aimed to assess whether the proposed RE framework successfully aligns technical execution with clinical intent while preserving requirements traceability throughout

development. The evaluation also verified compliance with the established Key Performance Indicators (KPIs) and examined whether the engineering artifacts consistently supported implementation and validation activities. Data collection was conducted through assisted inspection sessions and semi-structured interviews. Evaluators utilized a formal assessment instrument based on a 5-point Likert scale—ranging from Strongly Disagree (1) to Strongly Agree (5) to assess compliance with sensory, clinical, and architectural requirements. To mitigate proximity bias, the evaluation panel was divided into two distinct groups: internal stakeholders who participated in the elicitation phase (comprising the end-user UF-01, healthcare professionals PS-01 and PS-02, and IT professional DEV-02), and independent external auditors (an external psychologist and an external software engineer).

The quantitative and qualitative results indicate that the proposed engineering process consistently supported the translation of clinical requirements into implemented software components. The documentation baseline and traceability mechanisms remained consistent throughout development, and the resulting prototype demonstrated strong compliance with the established KPIs.

Regarding sensory safety (KPI-02), the end-user (UF-01) reported agreement that the neutral color palette and controlled visual stimuli prevented sensory overload, confirming that the Non-Functional Requirements directly supported cognitive accessibility. Clinically, both internal and external healthcare professionals assigned scores that validated the pedagogical flow (KPI-03). They also confirmed the non-punitive correction loop as an effective, modeling-based learning tool for adults with ASD, noting that the absence of a "Game Over" state aligns with clinical protocol. Furthermore, the specialists confirmed the clinical utility and integrity of the exported telemetry data (`KPI-05`), emphasizing that the captured response latencies provide actionable, quantitative evidence for future therapeutic planning.

From a software engineering perspective, the external technical audit validated the framework's autonomy and scalability. The independent developer confirmed that the documentation baseline—comprising the SRS, Serious-GDD, and RTM—was sufficiently complete and consistent to support implementation without requiring continuous supervision from the original clinical team. Moreover, the internal audit conducted by DEV-02 successfully identified specific traceability elements that were insufficiently specified for third-party scenarios. This finding is a positive outcome of the framework's governance, demonstrating that its review mechanisms effectively catch architectural ambiguities before full-scale implementation, thereby triggering a necessary refinement iteration.

Despite these promising outcomes, this study presents specific threats to validity that must be acknowledged. The primary limitation is the sampling bias, as the practical evaluation of the vertical slice relied on feedback from a single end-user with ASD ($N = 1$). Consequently, the current evaluation should be interpreted as a validation of the engineering process rather than evidence of the clinical effectiveness of the resulting serious game. While the multidisciplinary panel of specialists provides significant analytical depth, the limited sample size of users prevents statistical generalization of the game's clinical efficacy across the highly heterogeneous autism spectrum. The current findings therefore validate the engineering process itself. The evaluation demonstrates that the framework successfully generates auditable, traceable, and sensory-safe software artifacts, providing a structured foundation for future large-scale clinical trials.

## 7 Conclusion

The development of Serious Games for Health requires a structured integration between clinical knowledge and software engineering practices. A major limitation in current SGH literature is the lack of formal governance. This evaluated framework addresses this gap by explicitly prioritizing process maturity, standardized engineering artifacts, and rigorous traceability. This structural alignment supports the development of serious games as an institutionalized, repeatable software engineering process rather than an individual, experience-driven activity. The framework formally translates subjective clinical needs into objective, verifiable system requirements before technical implementation begins.

Furthermore, requirements management is operationalized through the RTM, which establishes bidirectional traceability between clinical objectives, software artifacts, implementation decisions, and validation criteria. This continuous mapping helps prevent uncontrolled scope changes, supports software evolution, and preserves clinical fidelity throughout the project lifecycle. By standardizing the documentation baseline—including the SRS, Serious-GDD, and RTM—and employing standardized identification across projects, the framework enables multidisciplinary teams to develop SGH solutions with improved governance, maintainability, repeatability, and technical quality.

The case study, which focused on developing a vertical slice for adults with ASD, provided practical evidence of the framework's applicability. The case study results indicate that the proposed engineering process supports the systematic generation of traceable software artifacts and facilitates the alignment between clinical objectives and implemented game mechanics. The framework's effectiveness depends on the quality of requirements elicitation and ongoing stakeholder involvement. Nevertheless, the case study demonstrates the feasibility of using it as a structured Requirements Engineering process for SGH development.

Overall, the results support the applicability of the proposed framework as a disciplined Requirements Engineering process for Serious Games for Health, while broader empirical evaluations remain necessary to establish its generalizability.

Future work will extend the evaluation to broader clinical scenarios and larger participant groups, addressing the current sampling limitations and assessing the framework's generalizability across different neurodivergent profiles. Additional studies will investigate the framework's scalability in more complex projects and its applicability to other healthcare domains, such as physical rehabilitation. Future research will also explore integrating the RTM into agile software development environments to support automated traceability management during iterative development cycles.

## Artifact Availability

Due to ethical restrictions and patient confidentiality requirements established by the Ethics Committee, the clinical datasets and telemetry logs generated in this study are not publicly available.

## References


[1] Nurul Ulfa Abdul Aziz, Roslina Ibrahim, and Firdaus Abdullah. 2024. Review of Game Features to Develop a Serious Game Model for Mental Health Awareness. In *2024 International Conference on Computing Innovation, Intelligence, Technologies and Education (CIITE)*. IEEE, Sepang, Selangor, Malaysia, 1–6. doi:10.1109/CIITE62244.2024.10987762

[2] Jorge Fernando Ambros-Antemate, María Del Pilar Beristain-Colorado, Marciano Vargas-Treviño, Jaime Gutiérrez-Gutiérrez, Pedro Antonio Hernández-Cruz, Itandehui Belem Gallegos-Velasco, and Adriana Moreno-Rodríguez. 2021. Software Engineering Frameworks Used for Serious Games Development in Physical Rehabilitation: Systematic Review. *JMIR Serious Games* 9, 4 (11 Nov 2021), e25831. doi:10.2196/25831

[3] Bruna Barbosa, Marcos W. Ribeiro, Luciana Berretta, and Sergio Carvalho. 2024. DiagnosTEA: a digital game as a tool for the diagnosis/therapy of Autism Spectrum Disorder. In *Anais do XXIII Simpósio Brasileiro de Jogos e Entretenimento Digital* (Manaus/AM). SBC, Porto Alegre, RS, Brasil, 1707–1718. doi:10.5753/sbgames.2024.241241

[4] Bruna Barbosa, Carlos Souza, Halleffy Santos, Luciana Berretta, and Sérgio Carvalho. 2026. From Clinical Requirements to Game Mechanics: A Requirements Engineering Framework for Serious Games for Health. In *Anais do XXVI Simpósio Brasileiro de Computação Aplicada à Saúde* (Ouro Preto/MG). SBC, Porto Alegre, RS, Brasil, 705–716. doi:10.5753/sbcas.2026.21474

[5] Bruna Barbosa, Carlos Henrique Rorato Souza, Halleffy Santos, Sergio T. Carvalho, and Luciana O. Berretta. 2026. The Role of Software Engineering in the Development of Serious Games for Health: A Systematic Literature Review. In *Proceedings of the IEEE International Conference on Serious Games and Applications for Health (SeGAH)*. To appear.

[6] Stéphanie Carlier, Vince Naessens, Femke De Backere, and Filip De Turck. 2023. A Software Engineering Framework for Reusable Design of Personalized Serious Games for Health: Development Study. *JMIR Serious Games* 11 (March 2023), e40054. doi:10.2196/40054

[7] Polona Caserman, Katrin Hoffmann, Philipp Müller, Marcel Schaub, Katharina Straßburg, Josef Wiemeyer, Regina Bruder, and Stefan Göbel. 2020. Quality Criteria for Serious Games: Serious Part, Game Part, and Balance. *JMIR Serious Games* 8, 3 (July 2020), e19037. doi:10.2196/19037

[8] David Fang, Eliana Silva, and Luís Paulo Reis. 2025. Designing Serious Games to Support Teachers' Emotion Regulation: A Systematic Review. In *2025 IEEE Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, Manchester, United Kingdom, 1–6. doi:10.1109/SeGAH65397.2025.11168429

[9] Raluca Ionela Maxim and Joan Arnedo-Moreno. 2025. Identifying Key Principles and Commonalities in Digital Serious Game Design Frameworks: Scoping Review. *JMIR Serious Games* 13 (March 2025), e54075. doi:10.2196/54075

[10] Bárbara Ramalho, Marta Vicente, Hugo Escobar, Sandra Gama, José Diniz, Vasilis Charisis, Leontios Hajileontiadis, and Sofia Dias. 2024. Developing Sensorimotor Art Games for Psoriatic Arthritis using Agile Storyboarding and Game Co-Design Processes. In *Proceedings of the 11th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*. ACM, Abu Dhabi United Arab Emirates, 260–267. doi:10.1145/3696593.3696603

[11] Douwe Ravers, Dmitry Alexandrovsky, Mari Naaris, Marco J. Konings, Elegast Monbaliu, Hans Hallez, and Kathrin Gerling. 2025. RollAbility: A Case Study of Aligning the Design of Game-Based Wheelchair Skills Training with Clinical Protocols and Player Experience Goals. *Proceedings of the ACM on Human-Computer Interaction* 9, 6 (Oct. 2025), 756–785. doi:10.1145/3748621

[12] Michael C Robertson, Tom Baranowski, Debbe Thompson, Karen M Basen-Engquist, Maria Chang Swartz, and Elizabeth J Lyons. 2021. Using the Behaviour Change Wheel Program Planning Model to Design Games for Health: Development Study. *JMIR Serious Games* 9, 4 (Dec. 2021), e29964. doi:10.2196/29964

[13] Bill Roungas. 2021. A Game Design Framework for Alleviating Negative Emotions of Cancer Patients. In *2021 IEEE 9th International Conference on Serious Games and Applications for Health(SeGAH)*. IEEE, Dubai, United Arab Emirates, 1–8. doi:10.1109/SEGAH52098.2021.9551911

[14] Emmanuel Gomes De Souza, Tadeu Moreira De Classe, and Ronney Moreira De Castro. 2023. CaSEJ: Game Design Canvas para Brainstorming de Jogos para Saúde. In *Anais Estendidos do XXII Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames Estendido 2023)*. Sociedade Brasileira de Computação, Brasil, 1223–1234. doi:10.5753/sbgames_estendido.2023.233288

[15] André Bonetto Trindade, Gabriel Brunelli Pereira, and Marcelo Da Silva Hounsell. 2022. Requisitos para Jogos Sérios Ativos em Chão Interativo para o Público Autista. In *Anais Estendidos do XXI Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames Estendido 2022)*. SBC, 542–551. doi:10.5753/sbgames_estendido.2022.224921

[16] Gary Ushaw, Janet Eyre, and Graham Morgan. 2017. A paradigm for the development of serious games for health as benefit delivery systems. In *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)*. 1–8. doi:10.1109/SeGAH.2017.7939264

[17] Marcelo Vasconcellos and Flávia Carvalho. 2023. "GDD-Sério": uma Proposta de Game Design Document (GDD) para desenvolvimento de jogos sérios. In *Anais Estendidos do XXII Simpósio Brasileiro de Jogos e Entretenimento Digital* (Rio Grande/RS). SBC, Porto Alegre, RS, Brasil, 169–178. doi:10.5753/sbgames_estendido.2023.233261

[18] Karl Wiegers and Joy Beatty. 2013. *Software Requirements* (3rd ed.). Microsoft Press.

[19] Dai-Jie Yang, Meng-Yao Lu, Chi-Wen Chen, Pei-Ching Liu, and I-Ching Hou. 2022. Development of a Therapeutic Video Game With the MDA Framework to Decrease Anxiety in Preschool-Aged Children With Acute Lymphoblastic Leukemia: Mixed Methods Approach. *JMIR Serious Games* 10, 3 (2022), e37079. doi:10.2196/37079

[20] Zhaoyi Yang, Pengcheng An, Jinchen Yang, Samuel Strojny, Zihui Zhang, Dongsheng Sun, and Jian Zhao. 2021. *Designing Mobile EEG Neurofeedback Games for Children with Autism*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3447527.3477522