digital health and innovation

Article

A Normative Analytics Approach to Functional Component Assessment: Identifying VR Efficacy Within the Video Game Therapy® Methodology

Marcello Sarini [1,*] and Francesco Bocci [2]

[1] Department of Psychology, University of Milano-Bicocca, 20126 Milano, MI, Italy
[2] PlayBetter Association, 51031 Agliana, PT, Italy; fbocci80@gmail.com
* Correspondence: marcello.sarini@unimib.it; Tel.: +39-02-6448-3746

Abstract

Background/Objectives: Single-case studies represent a sophisticated and rigorous methodological framework, widely established in clinical research for providing high-resolution data on individual functional responses. This study evaluates the clinical utility of integrating immersive Virtual Reality (VR) gaming as a novel "functional ingredient" within the Video Game Therapy (VGT) protocol. Given the exploratory single-case nature of this intervention, clinical state-modulations cannot be rigorously validated using standard aggregated group statistics. Therefore, the core objective of this paper is to investigate the therapeutic potential of the VR session on psychological state-modulation, introducing the Single-Case Normative Analytics (SCNA) framework as the mandatory statistical vehicle required to validate individual longitudinal shifts against normative data. **Methods:** The study treats individual VR exposures as independent, short-term clinical probes embedded within a real-world clinical journey. The SCNA framework was deployed by integrating Crawford's modified *t*-tests with longitudinal percentile tracking against an empirical normative reference group ($n = 20$). Acute state-anxiety variations (STAI-Y1), psychological well-being (PGWBI), and flow dynamics were tracked across three distinct sessions to monitor the patient's relative repositioning within the normative distribution. **Results:** The inferential analysis indicates that the immersive 20-min environment facilitated reliable, statistically significant changes in acute state anxiety and flow dimensions, systematically exceeding standard measurement error boundaries and successfully moving the patient's psychometric profile toward healthy normative ranges. **Conclusions:** While these findings focus on individual, idiographic reactivity, they demonstrate the utility of the SCNA framework in providing clinicians with objective, evidence-based feedback on the clinical viability of specific VR-based functional units. This approach allows for a rigorous evaluation of standalone digital tools independently of a full, holistic VGT protocol, offering a structured alternative to traditional designs focused on identifying general patterns across groups.

Keywords: therapeutic gaming; virtual reality; single case; normative analytics

Academic Editor: Bijan Najafi

Received: 7 May 2026
Revised: 6 June 2026
Accepted: 12 June 2026
Published: 16 June 2026

1. Introduction

Video games have increasingly gained attention as tools that can support psychological therapy, extending beyond recreation to therapeutic engagement, emotional regulation, and cognitive transformation [1]. Research suggests that video games can promote engagement, immersion, and sustained attention, qualities that traditional therapeutic methods may lack, and that they can be used to improve skills such as problem-solving, decision-making, and coping strategies within safe, controlled virtual environments [2].

Within this context, Video Game Therapy® (VGT®) has been proposed as an integrative psychotherapeutic approach that places the gaming experience at the center of the therapeutic process. VGT® was formally introduced by Bocci et al. [3] and draws inspiration from Geek Therapy and classic psychodrama to create a structured framework where commercially available video games are used not as distractions, but as therapeutic tools embedded within the therapeutic relationship.

The core tenet of VGT® is that video games provide interactive, immersive environments in which individuals can express salient emotions, experiment with personal and interpersonal dynamics, and access emotional states with reduced defenses compared to traditional verbal therapy alone. Gaming experiences naturally engage cognitive processes such as attention, imagery, decision-making, and emotional regulation, which are central to psychological well-being [4].

Commercial video games have been successfully employed for cognitive rehabilitation, attention training, and emotional symptom reduction in conditions including attention-deficit/hyperactivity disorder (ADHD), depression, anxiety, and post-traumatic stress disorder (PTSD) [2,5].

Despite the promising potential of VGT®, to date none of the games integrated into VGT® protocols have been developed specifically for virtual reality (VR) platforms. Unlike traditional Virtual Reality Exposure Therapy (VRET) [6], which relies on habituation through the replication of specific pathological triggers (e.g., in PTSD or phobias), here, VR is employed as a standardized 'functional ingredient' to modulate psychological states using commercial software.

In fact, VR offers additional therapeutic affordances that may be particularly suited to the goals of VGT®:

- Enhanced immersion and presence: VR environments strongly engage sensorimotor and perceptual systems, creating a deeper sense of "being there," which is associated with improved emotional engagement and attentional focus [7,8].
- Multisensory feedback and embodiment: Through head-mounted displays and motion tracking, VR enables embodied interaction that can intensify experiential learning and self-reflection [9].
- Safe exposure and graduated challenge: VR can simulate anxiety-provoking scenarios in controlled ways, making it useful for exposure-based treatments and graded engagement (e.g., as in [10]).
- Personalized therapeutic environments: VR worlds can be tailored to individual needs, optimizing challenge and support to facilitate flow states—optimal psychological states characterized by deep engagement and absorption [11].
- High Reproducibility: VR environments deliver highly standardized and perfectly replicable digital stimuli, ensuring identical environmental and baseline conditions across different sessions;
- Strict Clinical Facilitation: While some digital health tools explore unfacilitated applications, the VGT® framework strictly mandates the continuous presence and mediation of the therapist, as the software interaction is structurally bound to real-time clinical monitoring and subsequent relational debriefing.

Over the past two decades, clinical research has demonstrated the efficacy of VR across a spectrum of psychological conditions. In anxiety and stress-related disorders, immersive environments have proven effective by acting as powerful tools for emotional regulation and cognitive distraction [9]. VR has been also employed successfully for phobia treatment, anxiety reduction, and cognitive rehabilitation, and it enhances engagement and adherence compared to non-VR platforms [10]. More recently, the scope of VR interventions has extended to severe mental health conditions, including schizophrenia and psychosis spectrum disorders. In these domains, VR is increasingly deployed to facilitate social skills training, treat paranoid ideation, and improve neurocognitive functioning within safe and controlled digital environments [12,13]. Rather than replacing traditional therapy, modern digital health research views these immersive experiences as modular functional components that complement established clinician-led protocols.

These characteristics suggest that incorporating VR games within VGT® may potentiate its therapeutic effects by increasing engagement, facilitating flow experiences, and providing highly controlled immersive environments for psychological exploration.

The primary clinical objective of the present study is to investigate the therapeutic potential of an immersive Virtual Reality commercial game, evaluating its viability as a novel 'functional ingredient' to modulate acute psychological states within the Video Game Therapy (VGT®) protocol.

While existing literature has extensively documented that VR represents a high-potential therapeutic environment due to its ability to elicit presence and emotional engagement [9,10], evaluating such an idiographic intervention within the flux of a real-world clinical journey introduces a significant methodological challenge: individual state-modulations cannot be rigorously validated using standard aggregated group statistics.

Building on this clinical question, the development and deployment of the Single-Case Normative Analytics (SCNA) framework is introduced not as a competing goal, but as the necessary statistical vehicle and rigorous inferential tool required to validate these individual longitudinal shifts against normative data.

By evaluating a single case's psychological repositioning relative to a normative control group across domains such as anxiety, well-being, and flow, this work serves as a proof-of-concept. It demonstrates how integrating single-case statistics can provide the clinician with objective, data-driven feedback on the functional impact of individual VR sessions—treated here as standalone functional ingredients. This approach allows for assessing how an individual's position within normative distributions evolves across sessions, providing empirical evidence on the specific contribution of the gaming tool before its full integration into the broader, holistic psychotherapeutic process of Video Game Therapy®.

2. Background

Single-case designs have a long and well-established tradition across neuropsychology, clinical psychology, and cognitive science as a rigorous methodological framework for studying individual functioning in relation to a normative reference population (e.g., Refs. [14,15]). Contrary to the misconception that single-case studies are anecdotal or purely descriptive, several authors have demonstrated that, when supported by appropriate statistical procedures, single-case analyses allow robust inferential conclusions about abnormality, deficit, and change relative to control groups, as described in [16–18].

A central issue in single-case methodology concerns the comparison between an individual and a small normative sample. Standard parametric statistics are not directly applicable because the case represents a sample of size one. To address this limitation, Crawford and Howell (1998) [16] introduced a modified *t*-test specifically designed for single-case studies, allowing researchers to determine whether an individual's score significantly deviates from that of a control group while properly accounting for sample size and variance. This approach was further extended by Crawford and Garthwaite (2007) [17], who proposed additional indices such as the standardized case-control difference (Z_{CC}), the Point Estimate of Abnormality (PEA), and confidence intervals for abnormality estimates. These tools enable researchers not only to test whether a deficit exists, but also to

https://doi.org/10.3390/dhi1010004

estimate the proportion of the normative population expected to score more extremely than the case.

In parallel, the Reliable Change Index (RCI) [19] has been widely adopted in clinical psychology to determine whether a change observed over time exceeds measurement error. RCI allows researchers to distinguish true psychological change from fluctuations attributable to instrument unreliability, and has become a standard tool for evaluating therapeutic outcomes at the individual level.

Despite the availability of these robust statistical procedures, most single-case studies apply them in isolation. Typically, the Crawford test is used to assess baseline deficit, while RCI is used to evaluate change across time. However, what remains largely underexplored is the dynamic normative positioning of the case over multiple time points relative to the same reference distribution. That is, beyond asking whether a case is significantly different from controls at baseline, and whether change is reliable over time, it is possible to ask a more informative question: how does the individual's position within the normative distribution evolve across sessions?

This perspective shifts the focus from deficit detection to normative trajectory analysis. By computing Z-scores and percentile positions at each time point using the same control group parameters, it becomes possible to track the individual's movement within the normative space, observing transitions from pathological tails of the distribution toward central or superior normative regions.

We refer to this integrated approach as Single-Case Normative Analytics. To adequately evaluate the methodological contribution of the Single-Case Normative Analytics (SCNA) framework, its architecture must be precisely positioned against the established tradition of longitudinal single-case designs in neuropsychological rehabilitation. For decades, clinical neuropsychology has relied on robust inferential methods—most notably those operationalized by Crawford, Howell, and Garthwaite [16,20]—to rigorously compare an individual patient's performance against a small normative sample while controlling for Type I error inflation. However, in classical neurorehabilitation, these techniques are predominantly restricted to assessing stable cognitive deficits (e.g., localized brain injury impairments) or evaluating macro-longitudinal treatment effects across expansive baseline-to-post-intervention windows spanning several months [21]. The SCNA framework diverges from this traditional application by re-engineering these foundational inferential tools for a micro-longitudinal, state-dependent digital health paradigm. Rather than monitoring chronic cognitive capacities, SCNA applies normative analytics to track highly volatile, transient psychological states—specifically state anxiety, general well-being, and flow dimensions—across compressed, repeated experimental windows (i.e., immediate pre- and post-assessments encompassing a 20-min digital intervention). Consequently, the novelty of SCNA does not lie in the mathematical generation of the underlying single-case *t*-tests, but in its operationalization as a dynamic, real-time framework capable of quantifying whether an immersive technological 'ingredient' can reliably shift a psychiatric patient's affective state toward healthy normative boundaries session by session. Normative analytics combines:

- Crawford and Howell's modified *t*-test for baseline deficit assessment,
- standardized case-control positioning (Z_{CC}, percentiles, PEA),
- longitudinal percentile tracking across sessions using a fixed normative reference,
- and Reliable Change Index to confirm that observed repositioning exceeds measurement error.

Within this framework, psychological change is not described solely in terms of score variation, but in terms of relative repositioning within a normative distribution. This approach provides a more interpretable description of how an individual's psychological state is modulated by specific gaming components, allowing researchers to isolate the functional properties of a single tool—the 'ingredient'—from the complex clinical architecture in which it is embedded.

Such an approach is especially suited to investigating innovative therapeutic resources, such as Virtual Reality commercial games. In these contexts, normative analytics provides a quantitative lens to capture whether and how isolated gaming sessions possess the necessary 'active principles' to move an individual from a baseline deficit toward a normative functional range. By focusing on these discrete sessions, the framework identifies the specific potential of the VR experience to act as a catalyst for state-level shifts in flow, well-being, and anxiety.

The present study applies this framework to a single case undergoing three VR gaming sessions. These are analyzed not as a full psychotherapeutic intervention, but as independent clinical probes conducted within a broader therapeutic journey. This allows for a rigorous, data-driven evaluation of the gaming tool's independent impact, demonstrating how single-case statistics can validate individual therapeutic ingredients before their formal integration into a holistic Video Game Therapy® protocol.

3. Materials and Methods

3.1. Study Rationale and Analytic Scope

Building on the framework of Single-Case Normative Analytics, this study operationalizes a longitudinal evaluation of individual psychological states during specific VR gaming sessions. These sessions are analyzed as standalone therapeutic ingredients to determine their functional impact before their formal integration into the broader Video Game Therapy® protocol. Rather than treating the case as an isolated observation, the participant's state is continuously evaluated against a fixed normative anchor (a control group of $n = 20$) across multiple domains: state anxiety, flow experience, and psychological well-being. This design allows for three complementary levels of analysis:

- Baseline Deficit Assessment: Using Crawford and Howell's modified t-test and Z_{CC} indices to determine the subject's initial positioning relative to the normative distribution.
- Normative Trajectory Analysis: Tracking the subject's 'movement' within the normative space across three sessions, observing transitions from the clinical tails toward functional ranges.
- Reliable Change Assessment: Utilizing the Reliable Change Index (RCI) to ensure that the observed repositioning reflects genuine psychological modulation rather than measurement error.

By integrating these levels, the analysis seeks to provide a data-driven account of how specific VR gaming probes can modulate psychological states, offering a rigorous method for validating the components of a complex intervention.

3.2. Design

The study adopted a single-case design with normative comparison, aimed at validating the sensitivity of the Single-Case Normative Analytics framework. The psychological state of one participant (the case) was repeatedly assessed across three standalone VR gaming sessions. To ensure consistency and monitor longitudinal changes reliably, these sessions were conducted on a strict weekly schedule, with a standardized time gap of exactly seven days between consecutive sessions (i.e., seven days between Session 1 and Session 2, and seven days between Session 2 and Session 3). All sessions took place at the same time of day to minimize diurnal variations in psychological states. These sessions were analyzed as discrete experimental probes to evaluate the immediate impact of the gaming tool. This approach follows the methodological framework proposed by Crawford and Howell (1998) and Crawford and Garthwaite (2007) [16,17] for Single-Case Analysis against small control samples. Unlike group-comparison designs, this methodology uses a control group ($n = 20$) not as a direct clinical match, but as a fixed normative reference distribution (a 'normative anchor'). By using the same control group parameters (means and standard deviations) across all sessions, the design allows for the longitudinal tracking of the case's position within the normative space, determining how the 'VR ingredient' modulates the subject's profile relative to a functional baseline. It is important to clarify that, following Crawford and Howell's methodology, the control group serves as a psychometric benchmark rather than a clinical match. The goal of this normative comparison is to provide a standardized 'anchor' to evaluate the extremity of the case's scores relative to a healthy population distribution. While this allows for a precise estimation of abnormality (PEA), we acknowledge that demographic or clinical differences between the patient and the normative group must be considered when interpreting percentile positioning. For instance, the case's initial positioning in the extreme tails reflects his functional distance from a non-clinical population, providing a rigorous baseline for tracking longitudinal shifts toward a more normative range.

3.3. Setting and Analytic Scope

To ensure a rigorous evaluation of the technology's specific contribution, the three VR gaming sessions analyzed in this study were conducted and measured as standalone clinical probes. While the participant was concurrently engaged in a broader psychotherapeutic journey, this research adopts a component-based approach: it focuses exclusively on the VR gaming experience as an independent 'therapeutic ingredient' to assess its immediate functional impact on the subject's psychological state. This design intentionally brackets the holistic Video Game Therapy (VGT)® process to isolate the specific affordances of the VR environment.

Data collection took place between February and March 2024 at CPM Arcobaleno, a medium-support psychiatric community managed by the Cooperativa Fraternità (Brescia, Italy). All gaming sessions were conducted in a dedicated, quiet room within the facility to minimize external distractions and ensure the consistency of the experimental conditions.

Safety Monitoring and Adverse Events Protocol: Given the clinical vulnerability of the participant, a rigorous safety and ethical monitoring protocol was systematically enforced during and after each of the three experimental VR sessions. The participant was continuously accompanied by a trained clinician who conducted real-time visual, behavioral, and verbal monitoring throughout the 20-min immersive gaming window to detect any potential adverse events. This screening specifically targeted signs of simulator sickness (e.g., nausea, vertigo, visual fatigue), acute dissociative phenomena (e.g., transient depersonalization or derealization induced by the immersive medium), or any immediate exacerbation of psychiatric symptoms associated with the underlying schizoaffective disorder. Standardized pause criteria were established beforehand, allowing for the immediate termination of the session upon any sign of clinical distress. Over the course of the entire experimental protocol, the participant demonstrated excellent tolerance to the immersive VR ingredient; no somatic side effects, cybersickness, or psychological adverse events were either observed by the clinician or self-reported by the patient during or after the 20-min sessions.

Procedural Fidelity and Standardization: To ensure high internal validity and facilitate experimental replication, a strict procedural fidelity protocol was enforced across all sessions for both the participant and the normative reference group. All experimental sessions were facilitated by the same licensed psychologist, who was specifically trained in the Video Game Therapy (VGT®) framework. The interaction followed a standardized operational script divided into three consecutive steps. First, the participant received identical verbal instructions regarding hardware setup, involving the precise calibration of the PlayStation VR2 headset (including eye-tracking configuration; Sony Interactive Entertainment, San Mateo, CA, USA) and a brief description of the ergonomic controller layout. Second, the gameplay objective within Horizon Call of the Mountain was restricted to a standardized initial exploration phase, ensuring that the technical and narrative demands remained equivalent. During the 20-min immersive gaming window, the facilitator maintained a non-directive posture, providing technical troubleshooting if required but strictly refraining from any clinical or verbal intervention that could bias the user's psychological state. Finally, environmental factors—such as room illumination, acoustics, ambient temperature, and a standardized seated posture—were kept constant across all sessions to minimize external confounding noise.

3.4. Participants

The Case. The case study involved a male participant (born in 1998) with a diagnosis of schizoaffective disorder, depressive type (ICD-10: F25.1). At the time of the study, he had been residing in the psychiatric community since 2021 and was engaged in a long-term weekly psychotherapeutic journey. This concurrent verbal psychotherapy was a stable, routine treatment that remained unchanged in frequency and therapeutic orientation throughout the study, thereby acting as a constant baseline factor. For the purposes of this research, he completed three standalone VR gaming sessions, conducted on a strict weekly schedule (with a standardized 7-day interval between consecutive sessions). To isolate the direct psychological impact of the immersive experience from the broader effects of the ongoing psychotherapy, these sessions were monitored via acute pre- and post-session assessments, capturing real-time state modulations within a restricted 20-min experimental window.

The Normative Reference Group. Following the methodology for single-case inferential statistics [16], a normative reference group of $n = 20$ healthy participants was recruited from the staff and trainees of the same facility. This specific convenience sampling strategy was adopted to maximize contextual and environmental standardization, ensuring that the control sample experienced the VR module under identical physical and logistical conditions as the clinical case. Furthermore, from a statistical perspective, an $n = 20$ reference size is fully optimized for single-case inferential methodologies—such as Crawford's t-tests—which are mathematically built to operate robustly with small normative samples (typically between 10 and 50 participants) while maintaining strict control over Type I error rates, offering a superior empirical benchmark compared to generic population data from past literature. This group consisted of:

- 9 males (mean age = 47.55, SD = 10.4)
- 11 females (mean age = 38.09, SD = 12.1)

Consistent with the 'normative anchor' design, the control group completed a single VR session to establish the baseline parameters (means and standard deviations) of functional response to the gaming tool. It is important to note that, in Crawford-style analytics, the control group does not serve as a clinical match, but as a psychometric benchmark to estimate the rarity or abnormality of the case's scores. The use of a small but well-characterized normative sample is a validated practice in neuropsychological single-case research when large-scale normative data for specific innovative tasks (like VR gaming) are unavailable [18].

3.5. Experimental Rationale: Testing the VR Ingredient

To evaluate the functional properties of the VR gaming 'ingredient', three independent sessions were conducted using a Sony PlayStation VR2 headset and the commercial videogame Horizon: Call of the Mountain. The selection of this specific hardware platform was based on both practical deployment needs and technical specifications. Methodologically, the PS VR2 serves as a highly standardized, high-tier commercial off-the-shelf (COTS) system. It delivers a deeply immersive experience via its OLED displays (2000 × 2040 pixels per eye), a 110-degree field of view, and advanced sensory haptics. Crucially for this study, the console-based architecture (PlayStation 5) provided a stable, plug-and-play environment that bypassed the configuration complexities and software instabilities often associated with high-end PC-based VR setups, making it ideally suited for the logistical constraints of a residential community setting. This hardware choice guaranteed that the technological 'functional ingredient' remained perfectly uniform across all experimental sessions. Each session lasted approximately 20 min. It is crucial to specify that these gaming sessions were not part of a standard Video Game Therapy® (VGT®) protocol. Instead, they were designed as independent experimental probes to investigate whether VR gaming possesses the necessary affordances—such as flow induction and anxiety modulation—to be potentially integrated into the VGT® framework in the future.

While the case was undergoing concurrent psychotherapy, the gaming sessions were 'bracketed' and measured as standalone experiences. This design ensures that the observed psychological states (pre- and post-session) are attributable to the immersive technological experience itself, rather than to the psychotherapeutic elaboration typical of the full VGT® 'menu'. In this sense, the study acts as a pre-integration validation, testing the raw 'active principles' of the VR tool before it is officially formulated as a component of the VGT® methodology.

Functional Characteristics of the Selected VR Ingredient

The videogame Horizon: Call of the Mountain (Firesprite/Guerrilla Games) was selected for its specific functional affordances, which provide a rich interactive context suitable for exploring comprehensive shifts in psychological states. To date, this commercial off-the-shelf (COTS) AAA title has not been documented in previous psychological or clinical research. The choice to implement a cutting-edge commercial video game, rather than a traditional, simplified laboratory-developed VR environment, represents a deliberate methodological decision. This was done to evaluate the therapeutic potential of high-budget digital products characterized by deep narrative engagement, complex spatial interaction, and advanced visual fidelity, treating them as complex 'functional ingredients' whose transient psychological effects can be rigorously captured at a single-case level by our analytics framework. The game combines physically-driven climbing mechanics, exploration of high-fidelity naturalistic environments, and archery-based interaction. Rather than serving as isolated experimental variables, these interconnected features represent the clinical rationale for choosing this title, offering a multifaceted experiential baseline for:

- Attentional Engagement and Flow: The requirement for precise sensorimotor coordination and spatial orientation is particularly suited to inducing flow states and a high sense of presence.
- Environmental Modulation: Exposure to panoramic, hyper-realistic natural landscapes provides a stimulus for emotional regulation and relaxation, acting as a potential modulator of state anxiety.
- Agency and Mastery: The progressive challenge-skill balance inherent in the climbing and combat mechanics allows for the assessment of self-efficacy and agency within a safe virtual space.

https://doi.org/10.3390/dhi1010004

- Arousal Regulation: The narrative structure alternates between contemplative exploration and high-activation sequences, providing a dynamic range to test the sensitivity of normative analytics in capturing arousal fluctuations.

By evaluating the holistic impact of this interactive experience, the study investigates whether the 'VR ingredient' as a unified intervention can reliably elicit the broad psychological responses necessary to support its future exploratory evaluation within the VGT® protocol.

3.6. Measures

To capture the multi-dimensional impact of the VR sessions, the assessment was divided into baseline trait measures and session-specific state measures. Baseline Trait Assessment. The following instruments were administered once to establish the initial psychological profile of the case and the normative group:

- I-TIPI-R: The Italian version [22] of the Ten-Item Personality Inventory to assess Big Five traits [23].
- STAI-Y2: The trait scale of the State-Trait Anxiety Inventory [24].
- GFPS: The General Flow Proneness Scale to measure the predisposition to flow states [25].
- Compin: A scale assessing the clinical index of feelings of inferiority [26].

Session-Specific State Assessment. The following scales were administered pre- and post-session to the case (and once to the control group) to monitor the immediate modulation induced by the VR 'ingredient':

- STAI-Y1: The state scale of the State-Trait Anxiety Inventory [24].
- FSS: The Flow State Scale [27,28], encompassing nine dimensions (e.g., Balance between challenge and skills, Concentration, Sense of control). Dimensions D1, D3, and D4 were specifically monitored as functional preconditions for the flow experience [29].
- PGWBI: The Psychological General Well-Being Index [30], assessing six domains (Anxiety, Depression, Vitality, Positive Well-being, Self-control, and General Health).

3.7. Procedure and Assessment Schedule

The experimental procedure followed a standardized three-phase protocol for each VR session: (1) a Pre-session phase (20 min) for baseline state assessment; (2) the VR gaming phase (20 min); and (3) a Post-session phase (15 min) for immediate impact evaluation (see Figure 1).

To monitor the psychological modulation induced by the VR 'ingredient', the assessment was scheduled as follows:

- Baseline (Session 1)—Case and Normative Group: Participants completed a comprehensive assessment. Pre-session measures included trait scales (I-TIPI-R, STAI-Y2, GFPS, Compin) to establish the initial profile, and the state-anxiety scale (STAI-Y1). Post-session measures (FSS, STAI-Y1, PGWBI) evaluated the immediate response to the first VR exposure.
- Longitudinal Tracking (Sessions 2 and 3)—Case Only: To observe the trajectory of psychological states, the case underwent two additional sessions. These followed the same structure but focused exclusively on state-level changes (Pre-session: STAI-Y1; Post-session: FSS, STAI-Y1, PGWBI).

This schedule ensures that trait-level parameters remain as a fixed reference, while state-level variations are tracked longitudinally only for the case, allowing for the Normative Analytics comparison against the control group's single-session response.

https://doi.org/10.3390/dhi1010004

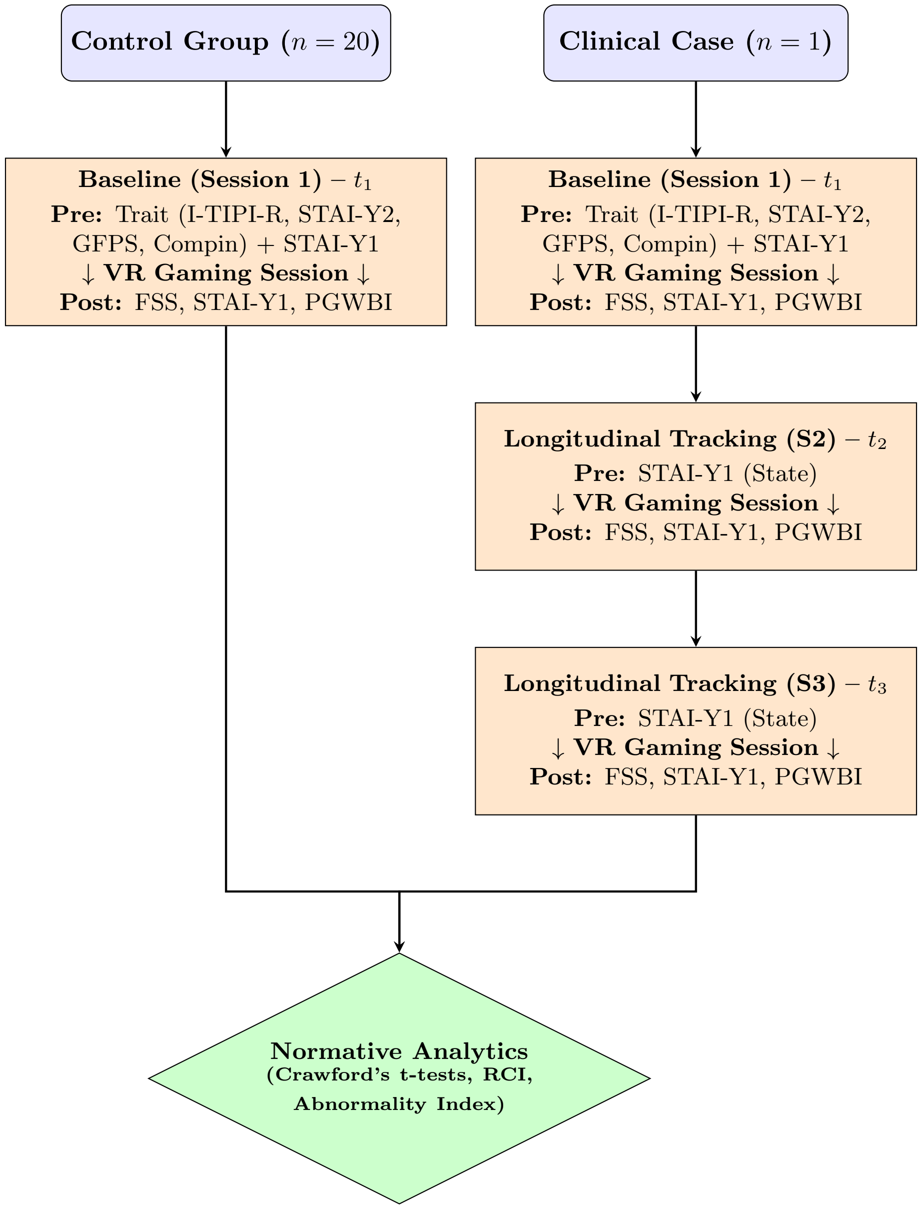

Figure 1. Flowchart of the experimental protocol with chronological time-points (t_1, t_2, t_3). The Baseline session (t_1) is identical for both groups to establish the normative benchmark. Subsequent sessions (t_2, t_3) track the longitudinal state-level trajectory of the clinical case.

3.8. Statistical Analyses

All data processing and statistical analyses were conducted in the R environment (version 4.5.1). Data manipulation was performed using the `dplyr` and `tidyr` packages to ensure a transparent and reproducible pipeline for longitudinal tracking. String processing and session identifier standardization were carried out using the `stringr` package. Data visualization was conducted using `ggplot2` to generate longitudinal normative trajectories. These plots represent the repositioning of the case across sessions relative to the 50th percentile of the control group, providing a visual 'map' of the psychological modulation induced by the VR sessions. To rigorously evaluate the impact of the VR 'ingredient' on the single case, all statistical indices were derived through custom R implementations based on the original methodological formulations. This approach ensures full adherence to the mathematical definitions of:

- Crawford and Howell's modified *t*-test, used to determine the significance of the difference between the case and the normative anchor;
- Standardized case–control difference (Z_{CC}), to estimate the magnitude of the effect;
- Point Estimate of Abnormality (PEA) and confidence intervals, to quantify the rarity of the case's scores within the normative population;
- Reliable Change Index (RCI), to confirm that shifts across sessions exceed measurement error and reflect genuine state modulation.

By using base R functions (e.g., `pt()`, `qt()`) for these calculations, we maintained complete control over the computational procedures, avoiding the limitations of external black-box packages and facilitating full transparency in the analytic focus.

3.8.1. Baseline Deficit Analysis

At Session 1, the case was compared to the control group using Crawford and Howell's modified *t*-test [16]:

$$t = \frac{X_{case} - \bar{X}_{ctrl}}{S_{ctrl}\sqrt{\frac{n+1}{n}}}$$

where n is the control group size. For each comparison, the following indices were computed: t-value, associated p-value, standardized case–control difference (Z_{CC}), percentile position, and the Point Estimate of Abnormality (PEA) with 95% confidence intervals.

3.8.2. Longitudinal Normative Positioning

For Sessions 2 and 3, Z-scores and percentiles were computed using the control group parameters from Session 1 as a fixed normative anchor:

$$Z = \frac{X_{case} - \bar{X}_{ctrl}}{S_{ctrl}}$$

This procedure allows for tracking the individual's trajectory within the same normative space, observing shifts from the clinical tails toward functional regions of the distribution.

3.8.3. Interpretation of Percentile Direction

Because instruments differ in score directionality, percentile interpretation followed this rule:

- For STAI, Neuroticism, and Compin: higher percentiles indicate worse clinical condition.
- For Big Five positive traits, GFPS, PGWBI, and all FSS dimensions: higher percentiles indicate better psychological functioning.

3.8.4. Reliable Change Index (RCI)

To determine whether the shifts observed between Session 1 and Session 3 exceeded measurement error, the Reliable Change Index (RCI) was computed [19]:

$$RCI = \frac{X_3 - X_1}{SD\sqrt{2(1-r)}}$$

where X_1 and X_3 are the scores at baseline and at the final probe, respectively.

Following a conservative and robust analytic approach, a uniform reliability coefficient of $r = 0.90$ was applied across all instruments. This choice was made for two reasons:

1. High Psychometric Threshold: 0.90 represents the stringent threshold typically required for clinical decision-making and high-stakes assessment. By applying this

value, we ensure that a change is labeled as "reliable" only if it is exceptionally robust, minimizing the risk of Type I errors (false positives) across heterogeneous scales.

2. Methodological Parity: In the absence of session-specific reliability data for innovative VR contexts, using a fixed, high-level coefficient provides a consistent and comparable benchmark for all "ingredients" (anxiety, flow, well-being) measured in the study.

From a formal psychometric perspective, this uniform application serves as a deliberate operational constraint optimized for single-case digital health evaluations. In the standard Jacobson and Truax formulation, the Standard Error of Measurement is defined as $S_E = s_1\sqrt{1 - r_{xx}}$ (where s_1 is the standard deviation of the normative group), which directly propagates into the standard error of the difference denominator:

$$S_{\text{diff}} = \sqrt{2(S_E)^2} \tag{1}$$

By mathematically fixing r_{xx} at a conservative baseline of 0.90, the framework artificially inflationizes the S_{diff} boundary for any individual instrument whose intrinsic published reliability might actually be lower than this anchor. This mathematical constraint systematically protects the Single-Case Normative Analytics (SCNA) framework against the inflation of Type I errors by forcing a significantly higher statistical clearance for any longitudinal shift to be deemed significant. While this approach introduces a conservative bias that mathematically increases the probability of Type II errors (potentially masking subtle clinical fluctuations), it guarantees that any statistical repositioning captured by the model is highly robust and entirely insulated from psychometric artifacts across heterogeneous instruments.

Absolute RCI values greater than 1.96 (corresponding to $p < 0.05$) were considered statistically reliable modulations.

4. Results

4.1. Baseline Comparison and Initial Positioning

At baseline (Session 1), the case was compared against the normative reference group ($n = 20$) to establish the initial distance from the functional anchor.

To ensure full transparency regarding the patient's initial state, Table 1 reports the raw baseline scores for the case alongside the normative parameters (Mean and Standard Deviation) of the control group used as a psychometric benchmark for Crawford's *t*-tests.

Several variables showed statistically significant deviations ($p < 0.05$), confirming that the case's psychological profile at study entry was positioned in the extreme tails of the normative distribution. Extremely elevated trait anxiety was observed on the STAI-Y2 ($t(19) = 3.58$, $p = 0.002$, $Z_{CC} = 3.67$), with a Point Estimate of Abnormality (PEA) of 0.0001, placing the case at the 99.99th percentile. Similarly, the clinical index for inferiority feelings (Compin) was markedly elevated ($t(19) = 3.17$, $p = 0.005$, $Z_{CC} = 3.25$; 99.94th percentile). Personality assessment via I-TIPI-R revealed a severe deficit in Conscientiousness ($t(19) = -2.82$, $p = 0.011$, $Z_{CC} = -2.89$; 0.19th percentile).Regarding the state-related dimensions that the VR 'ingredient' is hypothesized to modulate, the case showed a significant initial deficit in Concentration (FSS-D5: $t(19) = -2.14$, $p = 0.046$, $Z_{CC} = -2.19$; 1.43rd percentile) and in the PGWBI Depressed Mood subscale ($t(19) = -2.49$, $p = 0.022$, $Z_{CC} = -2.55$; 0.54th percentile). Beyond statistical significance, several PGWBI and Flow domains showed clinically compromised positioning (PEA < 0.10), including Anxiety (3.89th percentile), Vitality (3.31st percentile), Self-control (2.80th percentile), and Flow Dimensions D4 (9.04th percentile) and D7 (5.18th percentile). These values provide a rigorous quantitative baseline, against which the subsequent 'repositioning' induced by the VR sessions was tracked. Detailed estimates and confidence intervals are reported in

Table 2. Given the high number of variables analyzed at this stage, these baseline comparisons were treated as exploratory to identify potential areas of psychological impairment. Consequently, no formal correction for multiple comparisons (e.g., Bonferroni) was applied, which may increase the risk of Type I error; readers should therefore exercise appropriate caution in interpreting the statistical significance of individual baseline tests.

Table 1. Baseline psychometric profile: Raw scores of the case and normative parameters (Mean and Standard Deviation) of the control group.

Variable	Case (S1)	Control Mean (M)	Control SD
STAI-Y2	64	38.55	6.94
STAI-Y1 (Pre)	51	36.75	6.88
Extraversion	6	8.30	3.10
Agreeableness	12	11.00	2.38
Conscientiousness	5	11.00	2.08
Neuroticism	7	6.60	2.48
Openness	8	8.85	1.73
Compin	159	88.90	21.56
GFPS	41	47.30	4.68
STAI-Y1 (Post)	41	35.05	4.36
FSS D1	12	11.65	3.13
FSS D2	14	11.00	2.83
FSS D3	9	11.20	4.14
FSS D4	6	10.30	3.21
FSS D5	9	13.40	2.01
FSS D6	10	10.60	3.53
FSS D7	9	14.55	3.41
FSS D8	15	13.55	3.46
FSS D9	15	15.40	3.17
PGWBI Anxiety	15	22.45	4.22
PGWBI Depression	12	16.30	1.69
PGWBI Positive Well-being	13	17.40	3.36
PGWBI Self-control	10	14.45	2.33
PGWBI General Health	11	13.10	2.90
PGWBI Vitality	11	16.85	3.18

Table 2. Baseline comparison using Crawford and Howell's modified *t*-test.

Variable	t	p	Z_{CC}	Percentile	PEA	CI (Low–High)
STAI-Y2	3.58	0.002	3.67	99.99	0.0001	93.61–100
Compin	3.17	0.005	3.25	99.94	0.0006	86.59–100
Conscientiousness	−2.82	0.011	−2.89	0.19	0.0019	0.00–22.86
FSS D5	−2.14	0.046	−2.19	1.43	0.0143	0.00–48.25
PGWBI Depression	−2.49	0.022	−2.55	0.54	0.0054	0.00–34.41

4.2. Directionality and Normative Interpretation

Since the normative analytic framework relies on percentile positioning relative to the control group, the interpretation of the subject's trajectory depends on the specific directionality of each instrument. To ensure consistency across all analyses and graphical representations, we defined a functional orientation for percentile values:

- Distress-related scales (STAI-Y1, STAI-Y2, Compin, Neuroticism): Higher percentiles indicate an increased distance from the normative mean toward the pathological tail. Movement from high to low percentiles represents a shift toward the normative functional range.
- Well-being and Flow scales (PGWBI, FSS D1–D9, Big Five positive traits, GFPS): Higher percentiles reflect superior psychological functioning or higher state-resonance with the 'ingredient'. Movement from low to high percentiles represents a transition from a deficit condition toward a normative or supra-normative functional range.

This interpretative rule allows for a coherent reading of the longitudinal repositioning maps presented in the following sections, where the 50th percentile of the control group serves as the target functional anchor.

4.3. Longitudinal Normative Tracking

The longitudinal progression of the case across the three sessions is documented in Table 3. These raw scores provide the foundation for the calculation of the Reliable Change Index (RCI) and illustrate the observed psychometric shifts between the initial (t_1), intermediate (t_2), and final (t_3) assessment points.

Table 3. Longitudinal monitoring: Raw scores of the case across the three experimental VR sessions.

Variable	Session 1	Session 2	Session 3
STAI-Y1 (Pre)	51	50	41
STAI-Y1 (Post)	41	29	38
FSS D1	12	18	16
FSS D2	14	17	16
FSS D3	9	16	19
FSS D4	6	17	17
FSS D5	9	15	13
FSS D6	10	15	16
FSS D7	9	18	11
FSS D8	15	16	15
FSS D9	15	20	20
PGWBI Anxiety	15	19	17
PGWBI Depression	12	13	14
PGWBI Positive Well-being	13	12	17
PGWBI Self-control	10	10	9
PGWBI General Health	11	13	14
PGWBI Vitality	11	14	12

To evaluate the impact of the VR sessions over time, Z-scores and percentiles were computed for each session using the control group's Session 1 parameters as a fixed normative anchor. This allowed for a consistent longitudinal tracking of the subject's repositioning within the same functional space (see Figure 2).

4.3.1. Flow State Induction (FSS D1–D9)

At Session 1, the case's ability to enter a flow state was severely compromised, with scores falling in the lower tail of the normative distribution (e.g., Feedback [D4] = 9.04th percentile; Concentration [D5] = 1.43rd; Loss of Self-Consciousness [D7] = 5.18th).

By Session 2, a shift was observed: all Flow dimensions moved into the upper quartiles, exceeding the 75th percentile of the normative group (range: 76.08th–98.31st). This 'high-flow' profile was consistently maintained at Session 3 (range: 66.26th–98.15th). This suggests that the VR 'ingredient' was highly effective in bypassing initial cognitive-emotional barriers to flow, repositioning the subject from a deficit state to a supra-normative functional range.

4.3.2. Psychological Well-Being Modulation (PGWBI)

The tracking revealed a progressive trend toward the normative mean across several well-being domains. Depressed Mood moved from the extreme tail (0.54th percentile at baseline) to the 8.66th percentile at Session 3. Positive Well-being showed a substantial increase from the 9.53rd to the 45.27th percentile. Notably, General Health perception shifted from the 23.45th to the 62.18th percentile, effectively moving from a compromised state to a position above the normative median.

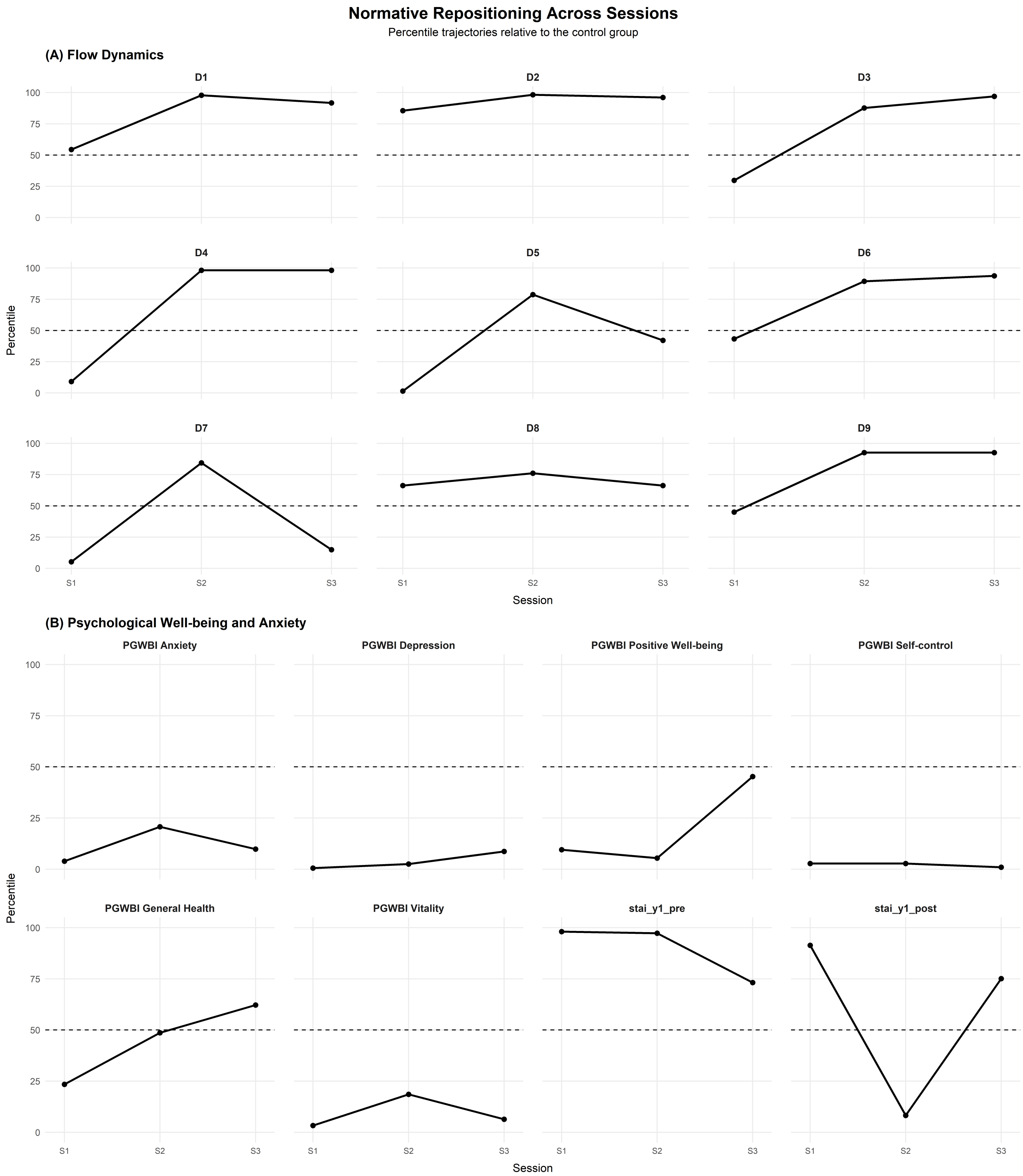

Figure 2. Normative repositioning across sessions. Panel (**A**) displays the individual's percentile trajectories across the nine dimensions of Flow (D1–D9), while Panel (**B**) illustrates the trajectories for Psychological Well-being (PGWBI subscales) and State Anxiety (STAI-Y1). The horizontal dashed line fixed at the 50th percentile represents the median score of the normative reference group ($n = 20$), serving as a baseline psychometric benchmark to evaluate the patient's positioning and trajectory relative to the central tendency of the healthy control population.

4.3.3. State Anxiety Reduction (STAI-Y1)

The most striking evidence of the 'active principle' in the VR sessions was found in anxiety modulation. Pre-session anxiety decreased from the 98.08th percentile (extreme distress) to the 73.16th. Even more significant was the Post-session drop: in Session 2, the subject's anxiety plummeted from the 91.39th percentile to the 8.26th percentile. This repositioning indicates that, following the VR experience, the subject's state was perceived as significantly less anxious than the average of the healthy normative group.

4.4. Reliable Change Assessment

To distinguish genuine psychological modulation from measurement error, the Reliable Change Index (RCI) was computed between Session 1 (baseline) and Session 3 (final probe). By applying a stringent and uniform reliability coefficient of $r = 0.90$, the analysis imposed a high bar for statistical significance, requiring an absolute RCI value > 1.96 to confirm a reliable change. As shown in Table 4, the VR 'ingredient' elicited a robust and statistically reliable improvement across multiple functional domains. Specifically:

- State Anxiety: A reliable reduction was confirmed for Pre-session STAI-Y1, indicating a stable downward shift in distress levels.
- Flow Dynamics: Six out of nine flow dimensions (D1, D3, D4, D5, D6, and D9) showed reliable increases, confirming that the gaming sessions consistently induced a state of high attentional engagement and mastery that transcends random fluctuation.
- Psychological Well-being: Significant reliable changes were observed in the PGWBI domains of Depressed Mood, Positive Well-being, and General Health.

These results provide the final validation for the Single-Case Normative Analytics framework: even under the most conservative psychometric assumptions, the data capture a clear and reliable transition from a clinically impaired baseline toward a normative functional range.

Table 4. Reliable Change Index between Session 1 and Session 3.

Variable	RCI	Reliable
STAI-Y1 pre	−3.25	Yes
D1	2.85	Yes
D3	5.41	Yes
D4	7.65	Yes
D5	4.45	Yes
D6	3.80	Yes
D9	3.53	Yes
PGWBI Depression	2.65	Yes
PGWBI Positive Well-being	2.66	Yes
PGWBI General Health	2.31	Yes

4.5. Synthesis of Results: Methodological Convergence

The results from Crawford's inferential test, longitudinal percentile tracking, and Reliable Change Index (RCI) converge into a highly coherent pattern, validating the sensitivity of the Single-Case Normative Analytics framework. As synthesized in Tables 2 and 4, the analysis captures a three-stage functional process:

1. Initial Deviation: The baseline assessment (Table 2) statistically confirms a profile positioned in the pathological tails of the normative distribution, particularly regarding anxiety and depressed mood.
2. Normative Repositioning: The session-by-session tracking (Figure 2) shows a rapid 'migration' of the subject's state toward the normative mean, with the VR 'ingredient' consistently eliciting supra-normative flow states (exceeding the 75th percentile).
3. Statistical Robustness: The RCI analysis (Table 4) confirms that these shifts are not artifacts of measurement noise but represent stable, reliable modulations of the subject's psychological state.

This convergence provides robust evidence that the integrated analytic approach can effectively map the 'active principles' of a digital intervention, providing a rigorous 'proof-of-concept' for the integration of VR gaming into structured clinical protocols.

5. Discussion

The present study investigated the psychological impact of Virtual Reality (VR) gaming sessions through the lens of Single-Case Normative Analytics. The results demonstrate that the VR 'ingredient'—operationalized through Horizon: Call of the Mountain—elicited a robust and reliable normative repositioning of the subject. Moving from a clinically compromised baseline, the participant reached functional and, in some cases, supra-normative levels across dimensions of flow, anxiety modulation, and psychological well-being.

The VR Ingredient as a Flow Catalyst: A central finding is the shift in Flow dimensions (FSS). At baseline, the subject's concentration (D5) was in the extreme lower tail (1.43rd percentile), reflecting the cognitive interference typical of schizoaffective distress. Following the VR probes, all flow dimensions surged above the 75th percentile.

This suggests that the immersive nature of VR acts as a powerful catalyst for Deep Effortless Concentration [31]. Unlike traditional media, the VR environment scaffolds the "preconditions of flow" (D1: Challenge-Skill Balance; D3: Clear Goals; D4: Immediate Feedback). Our data show that while the subject had a low trait-predisposition for flow (GFPS at the 8th percentile), the VR tool was able to facilitate the opening of these cognitive gates. This indicates that VR does not just wait for flow to happen; it actively structures the experience to make flow inevitable, even for individuals with high baseline deficits.

Emotional Regulation and Well-being: The reliable improvement (RCI) in PGWBI domains and the reduction of state anxiety (STAI-Y1) post-session provide evidence of the "active principles" of the intervention. The alternation between high-arousal climbing and contemplative panoramic views appears to have functioned as a digital emotional regulator.

The subject did not merely feel "better"; he moved from a position where only 0.5% of the healthy population would be found (Depressed Mood) toward a functional range. This repositioning suggests that the VR experience provides a "safe haven" of mastery and agency that can momentarily override pathological trait characteristics (such as high Neuroticism and low Conscientiousness identified at baseline).

From Experimental Probe to VGT® Integration: While these sessions were conducted as independent experimental probes, the results have direct implications for the Video Game Therapy® (VGT®) protocol. The high resonance of the subject with the VR tool suggests that:

- Phase Potentiation: VR can significantly potentiate the 'Immersive Phase' of VGT®, reaching therapeutic levels of absorption faster than standard gaming.
- Symbolic Priming: The metaphor of "climbing the mountain" provided rich symbolic material for the subsequent debriefing, facilitating themes of personal redemption and agency.

Therefore, we propose that VR-based gaming should be considered an ingredient to be integrated into the VGT® menu for patients with severe deficits in attentional control or high trait anxiety.

External Validity and Framework Transferability: A critical consideration in single-case digital health research concerns external validity and the extent to which these findings can be translated to other clinical ecosystems. Structurally, an *N*-of-1 design does not aim for the statistical generalizability of clinical efficacy; rather, its external validity resides in the transferability of its methodological and analytical architecture. The Single-Case Normative Analytics (SCNA) framework presented here is inherently scalable across three distinct dimensions. First, regarding patient populations and diagnoses, the mathematical reliance on

Crawford's *t*-tests and Reliable Change Indices (RCI) allows clinicians to export this protocol to diverse psychiatric profiles (e.g., generalized anxiety, major depression, or neurodevelopmental conditions), provided that appropriate context-standardized normative reference groups are established. Second, the framework is highly adaptable to other commercial off-the-shelf (COTS) VR games. While this study utilized a fantasy-adventure title, any software can be integrated as a digital 'ingredient' if its specific structural affordances—whether contemplative and relaxing or highly activating and goal-oriented—are systematically mapped onto the patient's therapeutic targets. Lastly, concerning therapeutic settings, the plug-and-play nature of modern console-based VR systems (such as the PS VR2) ensures high logistical flexibility, demonstrating that this methodology can easily transition from highly structured residential psychiatric communities to standard outpatient clinics or private psychological practices. Future research should prioritize multi-center single-case experimental designs (SCEDs) to empirically map this multi-dimensional scalability.

Cross-Diagnostic Scalability and Clinical Contraindications: The clinical architecture of the VGT® methodology is fundamentally transdiagnostic, designed to target core mechanisms of emotional dysregulation and experiential avoidance across diverse psychological conditions. While this case study focused on a patient with schizoaffective disorder to evaluate the framework within a high-complexity residential environment, the underlying protocol of alignment between software affordances and therapeutic objectives is routinely applicable to other diagnostic populations, including generalized anxiety, major depression, and post-traumatic stress profiles. However, the broader scaling of immersive VR ingredients necessitates a strict demarcation of clinical safety boundaries and population-specific risks. Immersive hardware introduces profound alterations in sensorimotor and spatial feedback, which may present specific clinical contraindications. In psychiatric settings, the deployment of high-fidelity VR must be heavily cautioned or restricted during acute, unmedicated psychotic episodes due to the risk of delusional assimilation (where the virtual environment becomes incorporated into persecutory or grandiose belief systems). Furthermore, patients presenting with severe dissociative disorders, such as profound depersonalization or derealization, may experience an exacerbation of symptom severity due to the abrupt post-session transition back to physical reality. Lastly, standard somatic risks—including photo-sensitive epilepsy, severe vestibular dysfunctions, and pronounced simulator sickness—must be systematically screened. By establishing rigorous clinical inclusion and exclusion criteria, the VGT® framework protects patient safety while leveraging the therapeutic potential of commercial digital environments.

Methodological Rationale of the Normative Benchmark and State Dynamics: A noteworthy structural characteristic of the current design is that the normative reference group completed the VR session only once, whereas the patient underwent three longitudinal sessions. In the context of Crawford's single-case methodology, this asymmetry does not introduce a confounding artifact, as the normative sample ($n = 20$) is mathematically operationalized as a static psychometric benchmark of healthy functioning rather than a longitudinal tracking control. Furthermore, the psychometric architectures of the primary instruments utilized—specifically tracking state anxiety, general well-being (PG-WBI), and flow dimensions—are explicitly designed to capture transient, state-dependent psychological configurations rather than cumulative trait-like shifts. From a theoretical perspective, state-dependent metrics like flow are structurally insulated from artificial inflation via simple user familiarization. According to flow theory, prolonged exposure and technical familiarization with a digital task would naturally decrease the perceived challenge; if the experience failed to maintain an optimal challenge-skill balance, this mechanism would trigger boredom, leading to a systematic collapse of flow scores rather than their longitudinal optimization or maintenance. Consequently, tracking these state variations across repeated sessions against a stable normative cross-section provides a rigorous method to isolate real-time psychological adjustments within the single-case paradigm.

Limitations and Future Directions

While the present study provides a robust proof-of-concept for the proposed analytic framework, some limitations inherent to its methodological paradigm should be considered:

- Single-Case Generalizability: The use of a single-case methodology allowed for high-resolution tracking of individual trajectories; however, from an idiographic standpoint, these findings reflect a tailored clinical response. Further research utilizing replication protocols across diverse clinical profiles is necessary to confirm the external validity of the observed normative repositioning within the VGT® framework.
- Absence of a Parallel Control Condition and the Novelty Confound: In accordance with the epistemological foundations of Single-Case Research Designs (SCRDs), this study does not feature a traditional parallel control group, as the experimental paradigm relies on the patient serving as their own internal control through high-frequency longitudinal tracking. Consequently, the observed pre-to-post within-session improvements in state anxiety and psychological well-being cannot be uniquely disaggregated from generic novelty effects, attentional engagement, or non-specific therapeutic factors inherent to any stimulating digital activity. Within this tradition focusing on individuals, however, the primary clinical objective was not to establish randomized, group-level causal efficacy, but to explore the feasibility and evaluability of immersive VR as a tailored "functional ingredient" for a complex psychiatric profile. By applying the Single-Case Normative Analytics (SCNA) framework, this study demonstrates how real-time individual fluctuations can be rigorously monitored and contextualized against an empirical benchmark, providing a structured, data-driven alternative to traditional designs aimed at discovering general rules applicable to groups. While the framework successfully confirms that these individual variations systematically exceeded standard measurement error boundaries, labeling VR as an isolated "active principle" remains exploratory. Future investigations seeking to definitively disentangle the specific causal weight of the medium from the novelty effect should transition to formal Single-Case Experimental Designs (SCEDs)—such as multiple-baseline or A-B-A-B reversal protocols—rather than traditional group-level trials.
- Normative Sample Composition and Selection Bias: The empirical normative benchmark ($n = 20$), while statistically appropriate for Crawford-type single-case inferential tests, presents clear demographic and clinical asymmetries compared to the patient. Recruiting from the facility's staff and trainees introduces a selection bias (e.g., higher average educational levels, absence of severe psychiatric conditions, and lower mean age). Consequently, this reference group represents a benchmark of 'healthy, context-standardized functioning' rather than a demographically matched control. While this allows us to map the patient's repositioning toward a normative standard, future studies should aim to implement stratified normative groups to account for specific age, cognitive, and clinical socio-demographic variables.
- Reliability Coefficients: A limitation of our Reliable Change Index (RCI) analysis is the application of a uniform reliability coefficient ($r = 0.90$) across all instruments. This choice was made to impose a stringent threshold for identifying "reliable" change and to protect the analysis from measurement error. However, we acknowledge that this may introduce bias: for instruments with lower published reliability, our approach is highly conservative (increasing Type II error risk), whereas for those with exceptionally high reliability, it remains a robust benchmark. By using this uniform

high-reliability anchor, we ensure that reported improvements represent substantial shifts that exceed typical measurement fluctuations across all administered scales.

- Scope of Measurement: This research focused on psychometric and behavioral indicators of state modulation (e.g., flow and anxiety). While the results are highly consistent, the integration of objective physiological data—such as autonomic markers—could provide a more granular understanding of the mechanisms underlying the impact of immersive VR.
- Temporal Dynamics: The study evaluated the immediate impact of the VR 'ingredient' across three sessions. A relevant methodological question is whether repeated exposure, learning effects, or familiarity with the VR hardware and gameplay could have artificially inflated the flow-related measures over time. However, according to established psychological frameworks on the Flow experience, flow is fundamentally non-linear and requires a dynamic balance between contextual challenges and individual skills. If the observed shifts were a mere artifact of habituation or task learning, the progressive mastery of the game without a corresponding increase in perceived challenge would have driven the patient into a state of boredom, thereby decreasing flow scores. The progressive optimization of flow metrics suggests instead that the commercial AAA title successfully sustained the patient within the interactive 'Flow Channel,' adjusting to their emerging familiarity. Nonetheless, future studies with extended timelines are required to map these dynamics over longer periods.
- Regression to the Mean: Given that the case entered the study with scores at the extreme tails of the normative distribution, the potential influence of regression to the mean must be considered. While a portion of the observed shifts could theoretically be attributed to this statistical phenomenon, we believe its impact is partially mitigated by the methodological convergence of our analytic framework. The consistency of improvements across three independent approaches—Crawford's *t*-tests (assessing abnormality), longitudinal percentile trajectories, and the RCI (filtering measurement error)—strengthens the interpretation that these changes reflect a genuine psychological modulation rather than mere statistical fluctuation. However, we acknowledge that this effect cannot be fully excluded in a single-case context.

In conclusion, this study validates the use of single-case inferential statistics to map the impact of digital health innovations. The VR gaming ingredient proved to be a reliable driver of psychological change, offering a rigorous data-driven rationale for its formal inclusion in complex psychotherapeutic protocols like VGT®.

6. Conclusions

In conclusion, this study provides a rigorous methodological and clinical proof-of-concept for the evolution of the Video Game Therapy® framework. Through the application of Single-Case Normative Analytics, we have demonstrated that VR gaming is not merely a technological variation of traditional gaming, but an 'ingredient' capable of inducing profound normative repositioning.

The evidence of statistically reliable improvements (RCI) in flow dimensions, anxiety reduction, and psychological well-being confirms that the immersive VR environment—exemplified by Horizon: Call of the Mountain—possesses the necessary affordances to meet the therapeutic goals of VGT®. Specifically, the ability of the VR tool to facilitate the opening of flow preconditions even in a compromised baseline profile suggests its elective suitability for the more complex phases of the protocol, such as emotional training and the enhancement of agency.

Based on these preliminary outcomes, we do not argue for an immediate, formal integration of VR into the VGT® methodology; rather, we propose these data as an initial exploratory proof-of-concept. This study evaluates the integration from a singular methodological angle, and its findings cannot be generalized. Before VR components can be systematically adopted within the VGT® protocol, more extensive testing beyond a single case is strictly required. Future research must address the larger picture, including multi-center clinical trials, an analysis of development and hardware deployment costs, and a formal evaluation of patient embracement, compliance, and long-term engagement with immersive technologies. When this integration would be possible, clinicians would be allowed to leverage the deep attentional absorption and the rich symbolic-embodied material provided by VR to accelerate the transition from pathological distress toward a functional normative range.

Future research should now focus on establishing standardized VR-based VGT® guidelines, comparing different game genres, and utilizing the normative analytic framework proposed here to monitor trajectories across larger clinical cohorts. This study marks a step toward a more technologically advanced and data-driven Video Game Therapy®, where the precision of normative analytics meets the transformative power of virtual immersion.

Author Contributions: Conceptualization, M.S. and F.B.; methodology, M.S.; validation, M.S.; formal analysis, M.S.; data curation, M.S.; writing—original draft preparation, M.S.; writing—review and editing, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki. The research consisted exclusively of non-invasive behavioral observations through pseudonymized self-reported questionnaires and Virtual Reality (VR) gaming sessions. These sessions were conducted for research purposes only and were entirely independent of, and disconnected from, any psychological or psychotherapeutic support the participant might have been receiving elsewhere. The experimental protocol did not constitute a clinical intervention or a medical treatment. In accordance with Italian national legislation and the ethical code of the Italian Association of Psychology (AIP), formal prior approval from a Research Ethics Committee is not mandated for non-clinical, non-invasive behavioral research involving pseudonymized data collection that does not fall under medical/pharmacological trial regulations. Written informed consent was obtained from all participants, and data were handled in full compliance with the EU GDPR.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Acknowledgments: We thank Cooperativa Comunità Fraternità, specifically the Comunità Psichiatrica a Media Protezione (Ospitaletto, BS) and President Alberto Festa, for their collaboration. We are grateful to the Video Game Therapy scientific committee—Elena Del Fante, Matteo Stefano Zanon, and Marco Chiapparino—for their continuous support. Finally, we thank Marco Lazzeri, Eleonora Stingone, and Sony Italia for providing the equipment used during this study. During the preparation of this manuscript, the authors used ChatGPT version GPT-5.2 solely for language editing and formatting support. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Alexiou, A.; Schippers, M.; Oshri, I. Positive Psychology and Digital Games: The Role of Emotions and Psychological Flow in Serious Games Development. *Psychology* **2012**, *3*, 1243–1247. [CrossRef]
2. Ruiz, M.; Moreno, M.; Girela-Serrano, B.; Díaz-Oliván, I.; Muñoz, L.J.; González-Garrido, C.; Porras-Segovia, A. Winning The Game Against Depression: A Systematic Review of Video Games for the Treatment of Depressive Disorders. *Curr. Psychiatry Rep.* **2022**, *24*, 23–35. [CrossRef]

3. Bocci, F.; Ferrari, A.; Sarini, M. Putting the Gaming Experience at the Center of the Therapy—The Video Game Therapy® Approach. *Healthcare* **2023**, *11*, 1767. [CrossRef]
4. Pallavicini, F.; Pepe, A.; Mantovani, F. Commercial Off-The-Shelf Video Games for Reducing Stress and Anxiety: Systematic Review. *JMIR Ment. Health* **2021**, *8*, e28150. [CrossRef]
5. Kollins, S.H.; DeLoss, D.J.; Cañadas, E.; Lutz, J.; Findling, R.L.; Keefe, R.S.E.; Epstein, J.N.; Cutler, A.J.; Faraone, S.V. A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): A randomised controlled trial. *Lancet Digit. Health* **2020**, *2*, e168–e178. [CrossRef]
6. Boeldt, D.; McMahon, E.; McFaul, M.; Greenleaf, W. Using Virtual Reality Exposure Therapy to Enhance Treatment of Anxiety Disorders: Identifying Areas of Clinical Adoption and Potential Obstacles. *Front. Psychiatry* **2019**, *10*, 773. [CrossRef]
7. Biocca, F.; Delaney, B. Immersive virtual reality technology. In *Communication in the Age of Virtual Reality*; L. Erlbaum Associates Inc.: Hillsdale, NJ, USA, 1995; pp. 57–124.
8. Witmer, B.G.; Singer, M.J. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence Teleoper. Virtual Environ.* **1998**, *7*, 225–240. [CrossRef]
9. Riva, G.; Wiederhold, B.K.; Mantovani, F. Neuroscience of Virtual Reality: From Virtual Exposure to Embodied Medicine. *Cyberpsychol. Behav. Soc. Netw.* **2019**, *22*, 82–96. [CrossRef]
10. Pallavicini, F.; Ferrari, A.; Pepe, A.; Garcea, G.; Zanacchi, A.; Mantovani, F. Effectiveness of Virtual Reality Survival Horror Games for the Emotional Elicitation: Preliminary Insights Using Resident Evil 7: Biohazard. In *Universal Access in Human-Computer Interaction. Virtual, Augmented, and Intelligent Environments: 12th International Conference, UAHCI 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, 15–20 July 2018*; Proceedings, Part II; Springer: Berlin/Heidelberg, Germany, 2018; pp. 87–101. [CrossRef]
11. Riva, G.; Baños, R.M.; Botella, C.; Mantovani, F.; Gaggioli, A. Transforming Experience: The Potential of Augmented Reality and Virtual Reality for Enhancing Personal and Clinical Change. *Front. Psychiatry* **2016**, *7*, 164. [CrossRef] [PubMed]
12. Freeman, D.; Reeve, S.; Robinson, A.; Ehlers, A.; Clark, D.; Spanlang, B.; Slater, M. Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychol. Med.* **2017**, *47*, 2393–2400. [CrossRef] [PubMed]
13. Ghorayeb, J.; Kalahasthi, R.; Hosseini-Kamkar, N. Virtual reality in treatment of psychological disorders: A systematic review. *Front. Digit. Health* **2026**, *8*, 1736381. [CrossRef] [PubMed]
14. Kazdin, A.E. *Single-Case Research Designs: Methods for Clinical and Applied Settings*, 2nd ed.; Oxford University Press: New York, NY, USA, 2011.
15. Hersen, M.; Barlow, D. *Single Case Experimental Designs: Strategies for Studying Behavior Change*; Pergamon General Psychology Series; Elsevier Science & Technology Books: Amsterdam, The Netherlands, 1976.
16. Crawford, J.; Howell, D. Comparing an Individual's Test Score Against Norms Derived from Small Samples. *Clin. Neuropsychol.* **1998**, *12*, 482–486. [CrossRef]
17. Crawford, J.R.; Garthwaite, P.H. Comparison of a single case to a control or normative sample in neuropsychology: Development of a Bayesian approach. *Cogn. Neuropsychol.* **2007**, *24*, 343–372. [CrossRef] [PubMed]
18. Crawford, J.R.; Garthwaite, P.H.; Porter, S. Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cogn. Neuropsychol.* **2010**, *27*, 245–260. [CrossRef]
19. Jacobson, N.S.; Truax, P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *J. Consult. Clin. Psychol.* **1991**, *59*, 12–19. [CrossRef]
20. Crawford, J.R.; Garthwaite, P.H.; Howell, D.C. On comparing a single case with a control sample: An alternative perspective. *Neuropsychologia* **2009**, *47*, 2690–2695. [CrossRef]
21. Crawford, J.R.; Garthwaite, P.H. Single-case research in neuropsychology: A comparison of five forms of *t*-test for comparing a case to controls. *Cortex* **2012**, *48*, 1009–1016. [CrossRef]
22. Chiorri, C.; Bracco, F.; Piccinno, T.; Modafferi, C.; Battini, V. Psychometric properties of a revised version of the Ten Item Personality Inventory. *Eur. J. Psychol. Assess.* **2015**, *31*, 109–119. [CrossRef]
23. Gosling, S.D.; Rentfrow, P.J.; Swann, W.B. A very brief measure of the Big-Five personality domains. *J. Res. Personal.* **2003**, *37*, 504–528. [CrossRef]
24. Ilardi, C.R.; Gamboz, N.; Iavarone, A.; Chieffi, S.; Brandimonte, M.A. Psychometric properties of the STAI-Y scales and normative data in an Italian elderly population. *Aging Clin. Exp. Res.* **2021**, *33*, 2759–2766. [CrossRef] [PubMed]
25. Elnes, M.; Sigmundsson, H. The General Flow Proneness Scale: Aspects of Reliability and Validity of a New 13-Item Scale Assessing Flow. *Sage Open* **2023**, *13*, 21582440231153850. [CrossRef]
26. Čekrlija, Đ.; Djuric, D.; Mirkovic, B. Validation of Adlerian inferiority (COMPIN) and superiority (SUCOMP) complex shortened scales. *Civitas* **2017**, *7*, 13–35. [CrossRef]
27. Jackson, S.A.; Marsh, H.W. Development and Validation of a Scale to Measure Optimal Experience: The Flow State Scale. *J. Sport Exerc. Psychol.* **1996**, *18*, 17–35. [CrossRef]

28. Diana, B.; Villani, D.; Muzio, M.; Riva, G. La validazione italiana della Flow State Scale—FSS. In *Flow, Benessere e Prestazione Eccellente. Dai Modelli Teorici alle applicazioni Nello Sport e in Azienda*; Muzio, M., Riva, G., Argenton, L., Eds.; FrancoAngeli: Milano, Italy, 2012; pp. 123–142.
29. Nakamura, J.; Csikszentmihalyi, M. The concept of flow. In *Handbook of Positive Psychology*; Snyder, C.R., Lopez, S.J., Eds.; Oxford University Press: New York, NY, USA, 2002; pp. 89–105.
30. Grossi, E.; Mosconi, P.; Groth, N.; Niero, M.; Apolone, G. *Il Questionario Psychological General Well Being. Questionario per la Valutazione Dello Stato Generale di Benessere Psicologico. Versione Italiana*; Istituto di Ricerche Farmacologiche "Mario Negri": Milano, Italy, 2002.
31. Marty-Dugas, J.; Smilek, D. Deep, effortless concentration: Re-examining the flow concept and exploring relations with inattention, absorption, and personality. *Psychol. Res.* **2019**, *83*, 1760–1777. [CrossRef] [PubMed]

https://doi.org/10.3390/dhi1010004