Jeffrey Harris
Jules van den Berg
Department of Experimental Psychology
Utrecht University
Heidelberglaan 1
3584 CS Utrecht
The Netherlands

Ralf F.A. Cox
Department of Developmental Psychology
University of Groningen
Grote Kruisstraat 2/1
9712 TS Groningen
The Netherlands

Stefan Van der Stigchel
Tanja Nijboer
Department of Experimental Psychology
Utrecht University
Heidelberglaan 1
3584 CS Utrecht
The Netherlands

Complex Dynamical Systems Analysis as a Potential Measure for Presence

Abstract

The practical value of virtual reality applications appears to be linked to the user's subjective sense of being in a virtual space, also called presence. Here we examine using the framework of complex dynamical systems analysis to continuously and objectively assess presence without the need for additional hardware by assessing different subcomponents of presence. Two experiments were conducted, where participants played a game that involved tossing colorful balls into designated baskets. In the first experiment, changes and differences in baselines in the subcomponent of coherence were assessed. In a second experiment, next to baseline coherence, baseline immersion was manipulated. The second experiment compared the method's ability to respond to temporary changes with a continuous and subjective self-report method. In both experiments, a mixed model design was applied to assess the within-subject effect of temporary changes and the between-subject effect of baseline differences. Results of the first experiment showed that the method successfully responded to temporary changes and measured baseline differences in presence via its subcomponent coherence. Results of the second experiment showed differences between conditions with different levels of coherence and immersion but also demonstrated that both subcomponents had a similarly strong effect on the outcome measure. The subjective method was only partially able to predict the within-subject and between-subject effects. In an attempt to propose a new framework that addresses the lack of continuous and objective measures, complex dynamical systems analysis demonstrates a promising ability to provide a solution but requires validation in additional scenarios.

1 Introduction

For users of virtual worlds, presence is loosely perceived to be a subjective quality of their virtual experience. In scientific literature, presence is commonly described as a quale, an "internal feeling elicited by sense perceptions" (Skarbez, Brooks, & Whitton, 2018, p. 3). Due to presence, the user can then live through an experience as if it were real. According to Bailenson (2018), the practical value of providing virtual experiences lies in the fact that virtual experiences can simulate scenarios that are either too dangerous, impossible, counterintuitive, or expensive (summarized by the acronym DICE). At the same time, these scenarios contribute no added benefit over two-dimensional counterparts if they don't feel realistic (Slater, Banakou, Beacco, Gallego, Macia-Varela, & Oliva, 2022). Hence, if the aim is to simulate a scenario to the degree that it

Correspondence to j.j.harris@uu.nl.

Presence, *Vol. 34, Winter 2025*, 361–393
https://doi.org/10.1162/PRES.a.10

feels like a real experience, optimizing for presence is arguably a requirement for delivering practical value.

In operationalizing presence, coherence and immersion play an important role. Immersion “provides the boundaries” (Skarbez, Brooks, & Whitton, 2018, p. 3) for the aspect of presence that is responsible for the feeling of being located in an environment. These boundaries of immersion encompass all actions that a user can perform to perceive the virtual environment (also referred to as sensorimotor contingencies) (Slater, 2009). Therefore, the boundaries of immersion are wider in situations in which more such actions can be successfully performed, providing a higher likelihood for users to feel located in the virtual space. This experience is also referred to by Slater (2009) as the place illusion.

While the place illusion refers to the experience that emerges within the boundaries of immersion, the plausibility illusion is the context-dependent counterpart (Skarbez et al., 2018; Slater, 2009), defined by the boundaries of “coherence” (Skarbez, 2016). Compared to immersion, coherence has subjective rather than objective boundaries. According to Skarbez (2016, p. 7), coherence is “the set of reasonable circumstances that can be demonstrated by the scenario without introducing unreasonable circumstances, where a reasonable circumstance is a state of affairs in a virtual scenario that is self-evident given prior knowledge.” In this sense, influences such as prior knowledge and reasonable circumstances dictate the parameters of the boundaries. Together, immersion and coherence help categorize how different factors influence presence.

Logically, the question of what happens when these boundaries are crossed arises. While experiencing virtual reality (VR), the user might perform actions that, while normal in physical reality, are not feasible in VR due to the limits imposed by boundaries of immersion and/or coherence. Performing these actions risks a so-called break in presence. While the literature mostly refers to breaks in presence, breaks are considered to have the ability of exclusively affecting the place illusion (Skarbez et al., 2018) or plausibility illusion (Brubach, Westermeier, Wienrich, & Latoschik, 2022). During such breaks, the boundaries of immersion or coherence are exceeded and the place or plausibility illusion is temporarily broken as the user transitions from illusion to non-illusion (Liebold, Brill, Pietschmann, Schwab, & Ohler, 2017). Mapping the boundaries throughout virtual experiences and understanding the consequences of exceeding these boundaries is therefore important to ensure a stable experience of presence (Bowman, Gabbard, & Hix, 2002). By narrowing the boundaries of coherence or immersion, one can increase the likelihood of these boundaries being crossed and purposefully induce breaks to study the effects. In a nonlinear virtual world where boundaries can be dynamic and difficult to predict, there are a vast number of possible boundaries. As an example, turning one’s head too quickly may lead to noticeable latency due to lagging image frames. In this case, quick movements would be considered beyond the boundaries of immersion. In contrast, if one were to toss an object into the air and see it float without previously establishing a scenario that involves zero gravity, this would count as an “unreasonable” circumstance and therefore be beyond the boundaries of coherence. In both examples, a break would occur, resulting in a decrease in overall presence. In the case of the two experiments presented in the current paper, breaks in presence will only affect the plausibility illusion and hereafter only be referred to as “breaks” for simplicity.

To assess breaks during nonlinear, virtual experiences, a continuous, versatile, and objective measurement approach is necessary. Such an approach would need to have sufficient temporal resolution to identify breaks, while also providing valid baseline measures. Arguably, if the manipulations that are known to influence subcomponents of presence are reflected in such an approach, and the approach is also applicable to a variety of settings, then that approach should be suitable for standardized presence measurement. In this paper, it is our objective to propose a candidate for standardized measurement that does not rely on additional hardware but instead relies on body movement. We aim to assess the feasibility by manipulating subcomponents of presence in predictable ways and measuring the accuracy of the method’s response.

A promising methodological approach for investigating presence using only body movement is behavioral measures that observe the user’s body language. When implementing such behavioral measures, there is the assumption that the more a user feels present in a virtual environment, the more responses to stimuli will be similar to those expected in a similar real environment (Berkman & Akan, 2019; IJsselsteijn, Bouwhuis, & Ridder, 2004). Examples include changes in behavior after embodying avatars that either do or do not resemble the user (Fox, Bailenson, & Tricase, 2013), or avoidance responses to perceived physical threats (González-Franco, Peck, Rodríguez-Fornells, & Slater, 2014). The advantage of behavioral measures lies in the fact that body language tends to occur spontaneously, implicitly, and directly, and is continuously influenced by the immediate environment (Skarbez et al., 2018). The disadvantage lies in the potential bias introduced when experimenters categorize and rate behavior (Schirm, Tullius, & Habgood, 2019). To solve this, we propose assessing presence via complex dynamical systems analysis.

Complex dynamical systems analysis is a framework to understand how a system's quality of interaction changes over time and is based on four core characteristics. First, the complex dynamical system itself must be made up of components that interact with one another and exist on different levels (Den Hartigh, Cox, Gernigon, Van Yperen, & Van Geert, 2015). In the case of user movement in VR, the "brain in a body in an environment [. . .] constitutes a heterogeneous, complex dynamical system" (Richardson & Chemero, 2010, p. 39) that can be broken down into components and assessed using this framework.

The second required characteristic, emergence, occurs when the interaction of components results in a higher-order effect (Richardson & Chemero, 2010). When creating a VR experience to elicit the subjective feeling of presence, achieving individual sensory perceptions by themselves does not suffice. Instead, presence only emerges if the system allows for a strong interconnection between perception (receiving and interpreting information from the environment) and action (execution of motor responses) within the unique boundaries of immersion and coherence (Sadeghipour & Kopp, 2011; Schoner, Dijkstra, & Jeka, 1998). This interconnection, also referred to as perception-action coupling (Warren, 1990), must be unperturbed to allow for presence to emerge through adaptive and dynamic behavior exhibited within the environment. In nature, examples include the flocking behavior of fish and birds which need to all behave in a synchronized manner to create this temporary, grouped behavior.

Third is self-organization. Through self-organization, components adapt their interactions according to external and internal constraints. In the case of presence, tighter constraints around a system's ability to establish a place illusion may be compensated by the system's ability to establish a plausibility illusion. For instance, consider a virtual environment that has sophisticated virtual agents or non-player characters. Such an environment might also constrain users from naturally walking. This virtual environment may fail to induce a strong place illusion due to a player's inability to walk around. However, due to a strong plausibility illusion, enabled by the interaction with realistic virtual agents, a user may still report high presence. This aspect of "plasticity" regarding presence, has also been referred to as a "function" from which the qualia of presence develops (Latoschik & Wienrich, 2022, p. 4; Skarbez et al., 2018, p. 16) which is related to the self-organization aspect of complex dynamical systems. In biological systems, examples of self-organization are commonly seen in colonies of ants that form trails to access and transport food.

The final quality is multi-scaledness. Multi-scaledness explains the fact that a system, in this case, the system that leads to presence, consists of interacting components or processes that are operating on multiple timescales. Importantly, the relationship of such components and processes remains constant across scales (Richardson & Chemero, 2010), which is assumed to be a result of movement coordination within a complex dynamical system (Renaud, Bouchard, & Proulx, 2002). To illustrate this, one can analyze human walking, which requires coordination on the scale of micro-movements such as balancing, larger limb movements to control things such as the swinging of the arms and legs, and whole body movements to do things like navigate the surroundings and adapt walking speed depending on the environment.

While complex dynamical systems analysis provides the framework to describe a system's changes over time, in the context of presence in VR, such changes are subtle and reflected in the user's movements. To pick up on these subtle changes, the vast amount of native movement data generated by user interactions and required to render an accurate virtual environment will be utilized. This continuous stream of movement data, analyzed using time-series analysis, provides a new opportunity to assess changes in presence and its subcomponents during VR experiences.

For the reasons outlined above, complex dynamical systems analysis deserves a closer look. To our knowledge, the ability of complex dynamical systems analysis to analyze dynamic and modular processes has only rarely been applied to VR, specifically presence (Renaud et al., 2002; Renaud, Chartier, Albert, Décarie, Cournoyer, & Bouchard, 2007), but may qualify as a promising solution to several unsolved problems in presence conceptualization and measurement.

In this study, two experiments were conducted to investigate the feasibility of using complex dynamical systems analysis theory and related data-analysis techniques to conceptualize and assess different aspects of presence. In Experiment 1, the ability of nonlinear time-series analysis, specifically detrended fluctuation analysis (DFA), to detect breaks was assessed together with its ability to detect baseline differences in coherence between different conditions. In Experiment 1, the hypothesis was that the complex dynamical systems outcome measure is significantly affected by within-subject breaks and between-subject baseline differences in coherence.

In Experiment 2, the experimental design was extended to include the factor of immersion when testing between subject conditions. Additionally, the complex dynamical systems outcome measure was compared to the results of a continuous self-report method to contrast objective with subjective measures. In Experiment 2, the hypothesis was that mean differences in the complex dynamical systems outcome measure are significantly affected by baseline differences in immersion, and that the complex dynamical systems outcome measure correlates with the results of the continuous self-report method.

In both experiments, we also assessed the performance of using the complex dynamical systems analysis framework by testing whether the presence questionnaire correlates with the complex dynamical system analysis outcome measure.

2 Experiment 1

2.1 Participants

We allowed participants between the ages of 18 and 65 years, who were self-reportedly fluent in English, had no cognitive impairments, no mental illness, no hearing difficulties, normal or corrected acuity, required no walking aid, and had access to a Meta Quest 2 headset and a Windows PC. Data was collected between October 31st, 2021, and January 2nd, 2022.

2.2 Materials

Development of the virtual environment was done through the Unity 3D game engine. The visual elements used to construct the virtual environment were purchased via the Unity Asset Store and included the products Amazing Assets Wireframe shader, Archanor VFX Confetti FX 2, Finward Studios Office and Police station pack, BNG VR Interaction Framework, Oculus Integration, and LuckiestGuy-Regular SDF Font. The virtual environment was presented using the Meta Quest 2 VR headset from Meta, which uses a Qualcomm Snapdragon XR2 CPU, 6 GB of RAM, a resolution of 1832 $\times$ 1920 pixels (LCD), and a refresh rate of 72 Hz. To run our application on the headset, participants were required to install the software SideQuest on their Windows PCs.

Movement data was collected through a C# script within the final build of the virtual environment which recorded three-dimensional positional coordinates data from the headset and two handheld controllers at a frequency of 72 Hz. In addition, we recorded adjacent data and automatically saved it to a CSV file on the Meta Quests internal storage (see Appendix A for the full list).

Beyond movement-related data, we employed the ITC-Sense of Presence Inventory (ITC-SOPI) (Lessiter, Freeman, Keogh, & Davidoff, 2001) as a subjective measure of presence to see how objective and subjective measures are compared. The ITC-SOPI was divided into four subscales, namely, spatial presence, engagement, ecological validity, and negative effects. The demographic data that was collected included age, gender, the number of hours played per week, the number of hours played per session, the years of VR experience, and the size of the play area in meters. Questionnaire data was collected via the online survey platform Qualtrics which participants were able to access only via a Windows PC. All the above data were analyzed using either MATLAB (version 2021b) or IBM SPSS Statistics (version 28.0.1.1.).

2.3 Procedure

The experiment was conducted online due to the coronavirus pandemic. Recruitment was done through Prolific, an online participant platform. After registering for the study, participants received a PDF file with a guide on how to proceed through the experiment. First, the participants started preparing their headset and their PCs. Preparation involved adjusting pre-specified settings within their headset, installing the software SideQuest on their PC that allowed them to install non-native applications including the VR application needed for the environment. The instructions also included requirements like being alone in the room, having a minimum play area of 1×2 meters of free space, and initiating a screen recording before launching the virtual environment. Afterward, the instructions on how to perform the task were introduced. Once the task was completed, participants were asked to fill out the ITC-SOPI via another link. Finally, participants emailed their files to the researcher and were debriefed.

2.4 Stimuli

The virtual environment consisted of two rooms, connected by an elevator. The first room (hereafter referred to as the "practice room") was designed for participants to get accustomed to interacting with the virtual environment and practicing the experimental task. The practice room was a square room, with a single gray basket and a virtual display above the basket on the wall (see Figure 1a). There were three balls with randomized colors on the floor. The possible colors were randomly generated from the HEX codes #FF0000, #00FF00, #0000FF, #ffd000, #d000ff, #ff8800, #00ffee, #0a0a0a, or #fcfcfc. On the wall behind the participant was a virtual mirror and a generic landscape painting. On the far end was an elevator door with a green button on the right side. In two of the diagonal corners of the room, there were CCTV cameras on the ceiling. In one corner of the ceiling, where there was no CCTV camera, was a translucent pipe from which balls (re-)spawned if they were incorrectly tossed into a basket. The elevator, connecting the practice room with the room in which the task was performed (hereafter referred to as the "task room"), included a virtual mirror at head height and a single CCTV camera in the corner of the ceiling (see Figure 1b). Next to the door were buttons. The task room was a square room with three baskets and displays, arranged in front of the remaining three walls, similar to the practice room. The task room also included two cameras and a translucent pipe for (re-)spawning (see Figure 1c).

The texture of the walls, ceiling, and floor in each room was designed to resemble the look of an average office building. The ceiling lighting within the virtual environment was dynamic, creating realistic shadows of the virtual objects. The sound within the virtual environment was designed to give participants the sense of 3D sound, becoming louder or quieter depending on the participant's location. Sound sources included the elevator, which produced a dull mechanical sound when in use, the elevator button, which produced clicks when pressed, and the ceiling lights, which produced a soft buzzing sound. The baskets produced a low-pitched thump or a high-pitched ring, depending on what ball was tossed into it (correct or incorrect). All sounds were subject to the condition which the participant was currently in.

(a)	(b)	(c)
Practice room with one active basket	*Elevator with mirror and buttons*	*Task room with one active and two inactive baskets*

Figure 1. *Virtual environment stimuli showing all three rooms that participants experience.*

2.5 Task

For the practice room, the participants were instructed to become familiar with the look and feel of the virtual environment by walking around and looking at the room but also interacting with the balls, trying out the tossing task, and standing in front of the mirror and pushing the buttons (see Appendix B for full instructions). When trying out the tossing task, the virtual display above the basket would light up in the requested color indicating to the participant which ball to toss next. If the color of the ball was correct, the display would then indicate that a successful toss was performed. This response differed depending on the manipulation. If the toss was incorrect, the ball would respawn from the tube in the ceiling. Once all three balls had been correctly tossed, the participants were supposed to use the elevator to get to the task room. In the task room, there were three baskets and nine balls on the floor. Only one of the baskets was active at any given time as indicated by a glowing ring on the baskets' edge and the virtual display. When inactive, the ring remained gray with the display displaying three gray z's. Once a successful toss was made, the basket would become inactive and another basket would activate. If an incorrect ball was entered into the basket, the ball would respawn and the basket would remain active. The task ended when all the balls on the floor were gone by displaying a menu with an option to quit.

2.6 Manipulations

To establish baselines and create breaks, several types of manipulations were used to alter the likelihood of a plausibility illusion being established (see Figure 2). Manipulations used to set baseline levels of coherence were present in the practice room, the elevator, and the task room. Manipulations aimed at producing breaks only occurred in the task room and were activated at predetermined moments during the experimental task. Concerning the baseline manipulations, there were different categories of manipulations used. Each targeted different aspects of coherence.

Priming was used as it established a context for the environment's behavior, increasing familiarity and limiting inconsistencies in mental models (Cerda, Fauvarque, Graziani, & Del-Monte, 2021) through self-evident prior knowledge (Skarbez et al., 2018). For this, CCTV cameras were used to remind the participants of their tasks and reinforce the overarching narrative (Gorini, Capideville, De Leo, Mantovani, & Riva, 2011; Park, Min Lee, Annie Jin, & Kang, 2010). In the high coherence condition, the participants were told, "There are cameras on the ceiling which are watching you and are actively tracking your movement in the rooms. The video they are capturing will analyze your behavior and determine whether you are following the instructions correctly." In this condition, the cameras tracked the participant's position in the room, providing a

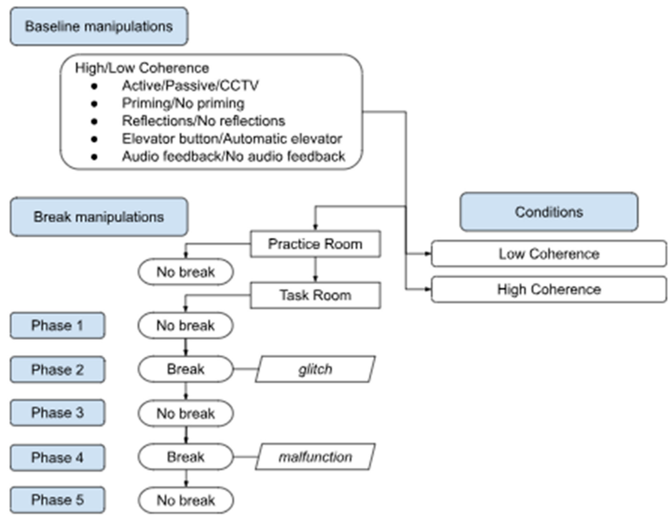


Figure 2. *Flowchart depicting the manipulations and sequence of steps in Experiment 1.*

reasonable circumstance and strengthening the plausibility illusion during the experimental task. In the low coherence condition, the participants were told, "There are cameras on the ceiling which are panning back and forth in each of the rooms. They are meant to provide a sense of scale and prevent nausea." Here, the cameras only panned back and forth. This introduced an unreasonable circumstance contrasting prior knowledge of how CCTV cameras work, weakening the plausibility illusion.

The locus of control manipulated coherence via auditory (Kern & Ellermeier, 2020) and visual feedback (Gonzalez-Franco, Perez-Marcos, Spanlang, & Slater, 2010). The virtual environment offered opportunities for the participants to interact with and receive feedback from elements in each of the rooms. Participants primarily interacted with the environment by picking up and tossing balls into the baskets. In the high coherence condition, high-pitched or low-pitched sounds were emitted from the basket as a response to the correct or incorrect ball being tossed. In the low coherence condition, there was no auditory feedback. Instead, the basket simply deactivated or remained open. Therefore, only participants in the high coherence condition received auditory feedback from the virtual environment as a result of their interaction, strengthening their plausibility illusion by demonstrating reasonable circumstances and consistency with other interactable objects. Participants in the low coherence condition encountered an unreasonable circumstance as the lack of auditory response from the basket contrasted with the behavior of other virtual elements in the environment. Another opportunity to interact with the environment was the elevator and the elevator buttons. In the high coherence condition, pushing the buttons resulted in the button visually depressing, a clicking sound being emitted, and the elevator arriving at the appropriate floor. In the low coherence condition, the button did not depress, there was no corresponding sound, and the arrival of the elevator was triggered automatically by the proximity of the participant to the elevator doors. The button's auditory and visual feedback, or lack thereof, was designed to strengthen or weaken the plausibility illusion similar to how the auditory feedback of the baskets worked. In the high coherence condition, having control over the elevator's behavior provided a reasonable circumstance that was self-evident given prior knowledge of how elevators work. The low coherence condition lacked this sense of control.

Lastly, coherence was also manipulated through embodiment. In the high coherence condition, the virtual mirror showed a reflection of the participant's virtual hands and the room providing a sense of visual feedback (Gonzalez-Franco et al., 2010) and embodiment (Kalina & Johnson-Glenberg, 2020). In the low coherence condition, only a reflection of the room was visible. In the case of the high coherence condition, a reasonable circumstance was demonstrated by showing a part of the participant's body reflected in the mirror similar to how the rest of the environment was reflected, demonstrating an internal consistency.

Next to the persistent manipulations aimed at manipulating the baseline coherence, the breaks were aimed at disrupting the participants' plausibility illusion temporarily. There were two breaks triggered at predetermined moments during the task (see Figure 2). The first break was triggered after one-third of the balls had been tossed into the baskets. To the participant, this break looked as if the colored balls on the floor of the virtual environment instantaneously shifted positions without any visible external force, resembling a "glitch." This introduced an unreasonable circumstance as all the previous tosses (both in the practice room and the task room) did not result in this type of behavior. After two-thirds of the balls had been tossed, the second break occurred, resembling a malfunction or feedback error (Si-Mohammed, Lopes-Dias, Duarte, Argelaguet, Jeunet, Casiez, et al., 2020). In this break, the basket, after tossing the correctly colored ball, gave incorrect feedback, namely keeping the basket active and re-spawning the ball, prompting the participant to try again. This behavior persisted until the participant tossed an incorrectly colored ball into the basket, demonstrating incorrect behavior by the basket since a wrong ball led to the "correct" feedback. This also introduced an unreasonable circumstance. In summary, the breaks were both meant to specifically target the plausibility illusion, whereas the baseline manipulations narrowed the boundaries of coherence for the entire duration of the virtual experience.

2.7 Data Analysis

The movement data collected during these phases resulted from the positional coordinates of the head, the left controller, and the right controller. Each of the three tracked points was defined by three variables, namely the x, y, and z positional coordinates. In total, this resulted in nine variables which made up the initial raw signal of the movement data before it was processed further. The nine variables were recorded continuously throughout the experimental task, producing a time-series of positional coordinates. This raw time-series of positional coordinates however still needed to be segmented for later analysis. The time-series of the experimental task was split into smaller time-series which either included breaks or did not include breaks. This way differences could be detected by comparing the mean levels of the respective time-series. These smaller time-series will hereafter be referred to as phases, numbered 1 through 5. Phases 2 and 4 contained the above mentioned breaks, and phases 1, 3, and 5 did not contain breaks. This pattern ensured that there was space before, between, and after breaks to allow for the plausibility illusion to start from and return to the baseline.

To capture the participants' immediate response to the breaks, the beginning of each phase was set to start the moment after the ball entered the basket. To ensure reliable outcome measures, each participant's phases were cut to a minimum length of 512 data points (5.66 seconds) across all phases. This meant that if the shortest length was 512 data points long, all other phases were also cut to that same length. If the shortest of the five phases was longer than 512 data points, then the other four phases were cut to that length. Five hundred twelve data points were chosen to provide the highest degree of validity when applying the time-series analysis method DFA to the time-series. This allowed for the shortest possible phases and the highest temporal resolution (Sekine, Akay, Tamura, Higashi, & Fujimoto, 2004). To complete the data preparation, each phase needed to be checked for outliers. Outliers can significantly distort the temporal structure that the time-series analysis method DFA is trying to capture. Hence, values above three median absolute deviations were removed from the time-series.

To capture the degree to which participants interact with the environment in a well-functioning, adaptive, and deliberate way, variability changes in volitional movement were assessed. To operationalize this movement, we calculated velocity changes from the raw time series of positional coordinates. The first step involved measuring the Euclidean distance between the three-dimensional coordinate of the hands at time t and the corresponding three-dimensional coordinate of the head at the same time t. This measurement was repeated for each successive time point, resulting in two time series of distances between the head and each respective hand. The next step involved calculating the velocity by finding the rate of change in distance over time to better detect trends or patterns in the steadiness or variability of movement.

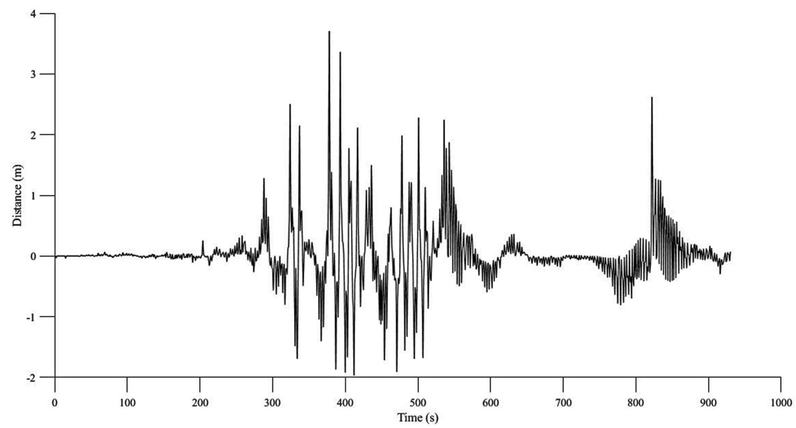
Velocity time-series capture aspects of a person's active control of movement as changes in velocity require the volitional application of force through muscle activation (Dotov, Nie, & Chemero, 2010). Velocity has been used to indicate motor control (Simperingham, Cronin, & Ross, 2016) or the loss of active motor control (Sekine et al., 2004). In presence research, velocity time-series deriving from head-tracking or eye-tracking movements have been used to assess presence only twice (Renaud, Bouchard, & Proulx, 2002; Renaud, Chartier, Albert, Décarie, Cournoyer, & Bouchard, 2007), according to our knowledge. In the current study, the distances were measured with the positional coordinates of the head as the reference point and the positional coordinates of the right and left hand as the second point. This allowed for dynamic reference points relative to the participant's general position. As a result, the nine variables resulted in two velocity time-series based on how the right and left hands moved while the participant was interacting with the virtual environment. See Figure 2 for an example time-series plotted, from the right hand of a participant during the second break.

As a time-series analysis method, detrended fluctuation analysis (DFA) was applied due to its ability to detect fluctuations in long-range correlations during head-tracking behavior in virtual environments (Renaud et al., 2007). Using DFA, we analyzed the temporal structure of variation in the positional coordinates time-series (Peng, Havlin, Stanley, & Goldberger, 1995; Wijnants, Hasselman, Cox, Bosman, & Van Orden, 2012). This resulted in an outcome measure that quantifies each time-series' degree of (temporal) self-similarity, providing a way of assessing whether long-range correlations existed within our signal. Generally speaking, the existence of long-range correlations has been associated with well-functioning, adaptive, and healthy biological and nonbiological systems (Grönlund, Yi, & Kim, 2012; Mansfield, Roy, & Shiratori, 2001; Peng et al., 1995; Werner, 2010; Wijnants, Hasselman et al., 2012). In movement science, long-range correlations differentiated between healthy and unhealthy participants (Hausdorff, 2007) or professional athletes and unskilled participants (Den Hartigh et al., 2015; Phillips, Portus, Davids, & Renshaw, 2012). In the context of a complex dynamical systems framework, DFA is a practical time-series analysis method for assessing the occurrence of long-range correlations during virtual experiences with differing degrees of presence.

DFA involves three main steps, namely removing any linear trends (also referred to as systematic variation) from our velocity time-series data, examining the changes in fluctuations from the mean in the context of differently sized windows (or boxes) (see Figure 3b), and assessing the relationship between the fluctuations and the window size. If the relationship is linear on a log-log scale, this points toward a power-law relation, with the slope of the regression line being the DFA exponent. Such a scaling relation in the time-series typically reveals that there are long-range correlations, and points toward a type of complex dynamical system referred to as an interaction-dominant system (Diniz, Wijnants, Torre, Barreiros, Crato, Bosman, et al., 2011; Van Orden, Holden, & Turvey, 2003). Interaction-dominant systems tend to exhibit the above-mentioned three qualities of a well-functioning complex dynamical system, namely interacting components, self-organization, emergence, and multiscaledness. The degree to which these qualities are present is quantified by the score of the DFA outcome measure, also called the DFA exponent. DFA exponents (α) describe the temporal structure of time-series variations. Values around $\alpha = 0.5$ indicate that the time-series derives from more random and unpredictable movements and values around $\alpha = 1.5$ indicate that the time-series derives from more structured and rigid movements. The optimal

(a)

Sample time-series of right-hand velocity for a random participant, showing velocity variability over time during the first phase with a break prior to detrended fluctuation analysis.



(b)

Sample DFA plot depicting the relationship between fluctuation function and box size on a log-log scale after processing time-series of right-hand velocity of the same randomly selected participant.

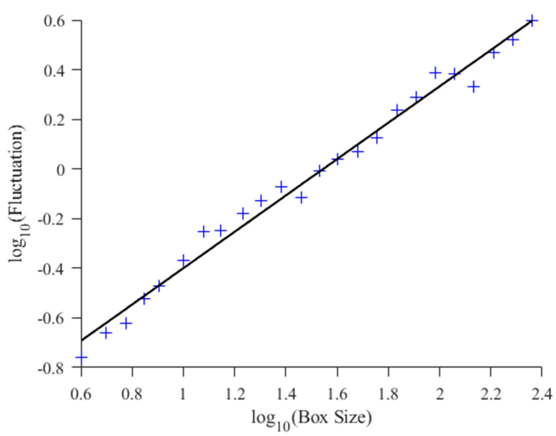


Figure 3.

NOTE: (b) F(n) (on the vertical axis) is the fluctuation function, which is the "average" of the root mean square (RMS) in each bin, at each particular bin size.

range lies around the value of $\alpha = 1$ as these values are indicative of dynamic movements, an interaction-dominant system, and long-range correlations. Noise that results from a time-series is frequently also defined as white ($\alpha = 0.5$), pink ($\alpha = 1$), or Brownian ($\alpha = 1.5$) noise. In other words, experiences with varying levels of coherence and immersion are hypothesized to generate different noise patterns during temporary changes and baseline differences.

Outputs resulting from the right-hand and left-hand acceleration time-series were averaged to create a single DFA exponent for each phase of each participant.

The reason for this was the small sample size which prevented us from assessing the between-subject effect and within-subject effects of each participant's hand individually. Assessing each participant's results exclusively based on right- or left-hand DFA exponents would not have yielded enough statistical power to be valid. The results of each hand are reported in Appendix C.

With the DFA exponent as the dependent variable and the experimental condition/level as an independent variable, the statistical analysis was performed. By applying a repeated measures analysis of variance (RM ANOVA), within-subject, between-subject, and interaction effects were calculated. RM ANOVA results were compared with the results of the questionnaires.

Missing values were imputed using SPSS via multiple imputations with 10 imputations, using the fully conditional imputation method of 25 iterations, a predictive mean matching model, and the constraints of imputing missing values and using them as a predictor. There were a total of $n = 4$ missing mean values in the positional coordinates data set with $n = 3$ being in phase 2 and $n = 1$ being in phase 4. We interpreted these values to be missing at random as the values were only missing during phases that involved breaks and the probability of them missing did depend on the way that participants progressed through the virtual environment. Moving through the experimental task too quickly resulted in participants not meeting the required minimum duration within a measurement window. In addition, we also found $n = 4$ instances where demographic data was missing completely due to technical issues with the questionnaire tool, which were imputed in the same way.

2.8 Results

Two sets of $n = 17$ (low coherence) and $n = 17$ (high coherence) participants per condition were collected for a total of $n = 34$ participants. The average age was $M = 27.50$, $SD = 7.16$ years, with $n = 26$ males and $n = 4$ females. From $n = 4$ participants the responses regarding their sex were lost due to technical errors in the survey application. Participants indicated to play VR for an average of $M = 2.83$, $SD = 2.27$ hours per week with an average session length of $M = 1.09$, $SD = .41$ hours. Participants had an average of $M = 1.50$, $SD = 1.07$ years of experience using VR. The average room size was $M = 3.37$, $SD = 1.47$ meters^2.

With the assumption of sphericity met, $p = .239$, there was a significant difference between the means of the phases, $F(4, 128) = 6.778, p = < .001$ for both groups, indicating that the DFA exponents differentiated between two or more of the five phases. This points toward the ability of DFA to respond to breaks in the plausibility illusion during continuous measurement. As there was also a significant interaction effect $F(4,128) = 3.357, p = .012,$ we looked further into where these effects exist and confirmed whether the differences reflected the manipulations of coherence.

We compared the means between each of the five phases in every possible combination for the high coherence condition. The high coherence condition was assessed because it was the only condition in which significant differences existed between the phases. We previously assumed that the environment providing the highest coherence would also allow for the strongest breaks and biggest mean differences between phases. The results supported the assumption. Significant differences were indeed most common among phases that included breaks, namely phases 2 and 4, $p < .001$, and phases that did not include breaks, namely phases 1 and 3, $p = .005$. One exception was the lack of significant difference between the two phases which did include breaks and the centermost phase, namely phases 2, 4, and 3, $p > .05$ (see Figure 4).

When visually inspecting the line graph (see Figure 4) it is also noticeable how the middle data points representing phases 2–4 on the high coherence condition (solid line) are roughly the same height, compared to the first and fifth data points for phases 1 and 5. While the DFA exponent of the centermost phase is slightly lower, it is not statistically different from the DFA exponents of the neighboring data points. The exact reason for this is unclear but there is a possibility that after the first break during phase 2, the plausibility illusion did not yet recover to where it was before the break in phase 2, leading to a similarly high DFA exponent. If there would have been more time between phase 2 and phase 4, the plausibility illusion may have recovered fully

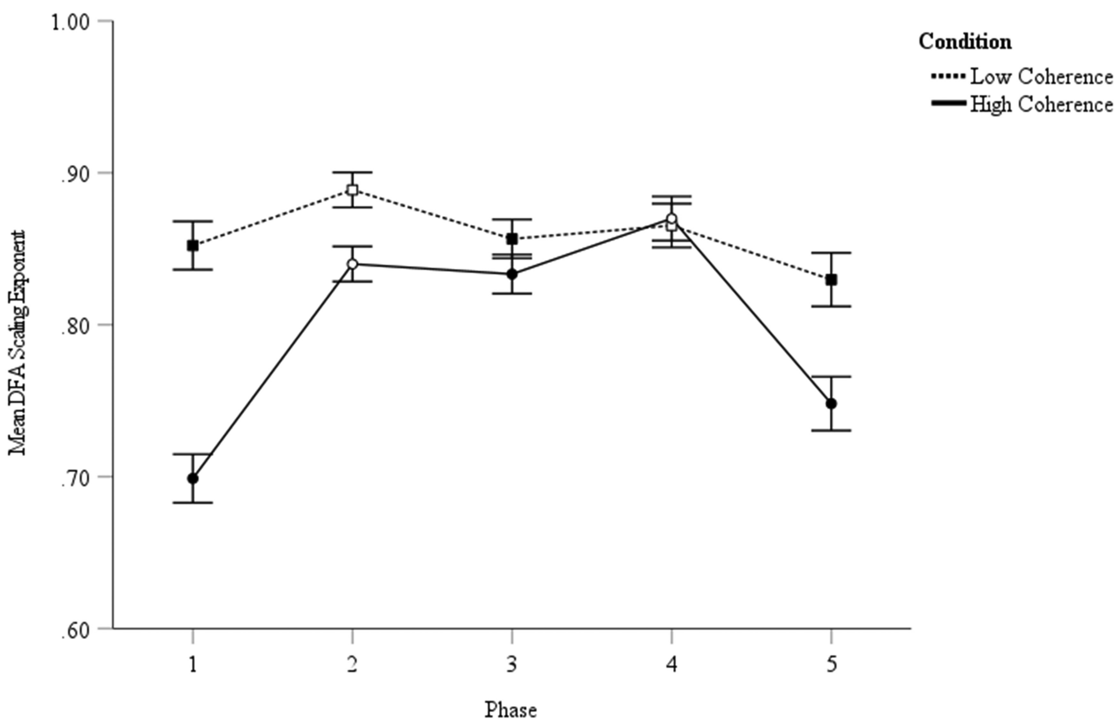


Figure 4. *Mean DFA phase outputs per condition.*

NOTE. White markers represent phases that include breaks; black markers represent phases that do not include breaks.

so that the middle data point representing phase 3 may have been even lower.

However, compared to phase 5, this is not the case. While one can say that phase 3 has a similar DFA exponent to the preceding phase, which includes a break, phase 5, which also follows a phase with a break, behaves differently. The data point is visibly lower (see Figure 4) and the difference in DFA exponents between phases 4 and 5 is significant, $p = .009$. This means that the phases following each break are potentially affected differently by each respective break. A reason for this might be that the breaks are different, one resembling a sudden glitch and the second one resembling a malfunction. The malfunction took participants longer to resolve compared to the first break which essentially resolved itself instantly. Participants might have already rebounded during this time, returning to the plausibility illusion quicker than in the first break. Overall, we still considered these findings to support the validity of the breaks since the largest mean differences between DFA exponents were between phases with breaks and phases without breaks, with the exception of phase 3.

Most apparently, when examining Figure 4, it becomes evident that phases involving break manipulations result in outcome measures closer to pink noise, indicating stronger long-range correlations compared to phases without manipulations. This is apparent from the data points at phases 2 and 4, which align more closely with the DFA exponent value of 1. This trend appears consistent across both groups, suggesting that the manipulations effectively and consistently altered the fluctuation dynamics in the velocity time-series. The observed increase in DFA exponent values implies a shift from less predictable, more random-like behavior to more constrained, repetitive patterns, as breaks actively limited the boundaries of coherence during these phases compared to phases without breaks.

Next, we assessed the DFA's ability to respond to differences in baseline coherence between the high and the low coherence conditions. With the assumption of homogeneity of variances met, $p > .05$, the mean DFA exponents for each group were found to be significantly different $F(1, 32) = 10.523, p = .003$. This suggests that the manipulations were valid in creating experiences that provide different baseline levels of coherence and that DFA is suitable for assessing these differences. This is in alignment with the within-subject finding discussed above. Similarly to the within-subject results, the effect of the baseline manipulation showed a similar direction towards the DFA exponent value of $\alpha = 1$. The low coherence condition showed a higher mean DFA exponent $M = .859, SD = .013$ than the high coherence condition $M = .798, SD = .013$. When considered alongside the within-subject results, there is a clear pattern in the directionality of the outcome measure between breaks and low coherence baselines.

We also wanted to understand how the mean values of individual phases differentiated themselves from one condition to the next. For this comparative assessment of corresponding phases, different contrasts were examined. In the pairwise comparison it was found that only phase 1 was significantly different, $p < .001$, in the high coherence condition $M = .697, SD = .116$, compared to the low coherence condition $M = .851, SD = .109$. In all other instances, the difference between the DFA exponents in the high coherence phases and the corresponding low coherence phases was not significant. From this, we concluded that the baseline manipulations were successful as participants appeared to have different starting levels of coherence in each condition.

Multiple linear regressions were carried out to investigate whether the ITC-SOPI subscales could significantly predict the DFA exponents for the phases in both conditions. We defined the results of the questionnaire scales as independent variables and DFA exponents for the five phases as dependent variables. In terms of self-reported presence, only one subscale from the ITC-SOPI predicted the DFA exponent, namely negative effects $M = 1.892, SD = .945$. Negative effects describe "adverse physiological reactions" (Lessiter et al., 2001, p. 9) and predicted two phases without breaks, namely phase 1, $R^2 = .171, F(1, 32) = 6.599, p = .015$ and phase 5 $R^2 = .361, F(1, 32) = 4.804, p = .036$. The subscale negative effects are related to the assessment of nausea which is negatively correlated to presence (Weech, Kenny, & Barnett-Cowan, 2019). Interestingly, the subscale negative effects did not predict the DFA exponent for the phases that include breaks. This is surprising because those phases would be expected to have a higher likelihood of causing nausea due to their unexpected nature. Hence, based on the above results, only the DFA exponents of phases without manipulations are predicted by the presence questionnaire subscale, showing no clear overall pattern.

Regarding the demographic data, only the hours played per week had a predictive effect on the DFA exponent. The number of hours played per week was gathered to understand if the amount of practice would affect the DFA exponent. The effect was significant for phase 1, which had no break $R^2 = .329, F(1, 32) = 15.692, p < .001$. Therefore, in regards to demographic characteristics, the amount of weekly usage of VR appears to affect the amount of coherence experienced in the initial phase.

2.9 Interim Discussion

Experiment 1 was carried out to understand the suitability of assessing breaks and baseline differences in coherence by using DFA in the context of a complex dynamical systems analysis framework. The primary inquiry was to test whether induced breaks would correspond to changes in DFA exponents in a velocity time-series calculated from positional coordinates. Additionally, we tested whether differences between high coherence and low coherence baselines would correspond to detectable differences in DFA exponents. Finally, the DFA method was also contextualized by testing if DFA exponents would be predictable through presence questionnaire subscales or demographics.

Results reveal that using DFA is an effective approach for assessing breaks and differences in baseline coherence. The finding that phases with breaks can be distinguished from those without breaks aligns with the assumption that changes in the participant's perception-action coupling influence the temporal structure of their movements, which we infer to be closely linked to coherence. We therefore conclude that applying DFA to movement time-series, captured via positional coordinates, is an appropriate way to assess when changes occur in the user's plausibility illusion. We also conclude that baseline differences in coherence can be assessed by applying DFA to the entirety of the velocity time-series. The finding regarding baseline differences also provides validity to the baseline manipulations we employed to create differences in coherence throughout the task. In Experiment 2, we further tested the robustness of our approach by examining the effect of alternative manipulations. Specifically, we investigated whether manipulating immersion, rather than coherence, would yield comparable baseline differences.

A further notable finding is that the plausibility illusion seems to be disrupted more strongly in the high coherence compared to the low coherence condition. The low coherence condition was expected to enable less severe changes because it was assumed that there was "less plausibility to break." This assumption is based on the idea that there is a sort of minimum coherence (and minimum immersion) beyond which a break always occurs.

Lastly, additional validation of our method comes from the finding that the DFA exponents of phases are partly predicted by self-report measurements, specifically those regarding negative effects. However, a major limitation regarding such self-report questionnaires is the lack of continuity. Therefore, to create a more detailed comparison between temporary changes in DFA outputs and self-report ratings, we used a different self-report approach in Experiment 2 that allows for continuous responses.

3 Experiment 2

3.1 Methods

Experiment 2 examined how immersion would affect the DFA exponent. The perception-action coupling process during the experience of virtual environments may be affected by both the boundaries of coherence and the boundaries of immersion. To assess the effect of immersion, we included immersion in our between-subject analysis. To further contextualize the method with subjective methods, the DFA outputs were correlated with outputs from a continuous response measurement tool (IJsselsteijn et al., 2004).

In this experiment, the materials, stimuli, and tasks remained the same. There were additional materials used for the additional measurement tool. In the data analysis, all steps including the imputation method also remained the same. Only the data from the continuous response measurement tool was added and integrated into the analysis as described below.

3.2 Participants

The second experiment was conducted at Utrecht University in Utrecht, the Netherlands, and involved students who were recruited via an internal participant recruitment system called Sona Systems. The study requirements were identical to Experiment 1 except for requiring access to a Quest 2 headset. Data collection took place between September 19 and 23, 2022.

3.3 Materials

The questionnaires used for data collection remained the same but were filled out by the participants using an iPad. In contrast to Experiment 1, we implemented a continuous self-report method called continuous response measure (CRM) (IJsselsteijn et al., 2004). The software used was an open-source tool named *CARMA: Software for Continuous Affect Rating and Media Annotation* (version 14.08), developed by Girard (2014). The tool is designed as a modernized version of the "continuous assessment methodology" (Freeman, Avons, Pearson, & IJsselsteijn, 1999; IJsselsteijn, De Ridder, Hamberg, Bouwhuis, & Freeman, 1998) sometimes also referred to as the "affect rating dial" (Gottman & Levenson, 1985). The computer that CARMA ran on was a Windows PC. A joystick connected via USB was used by the participant to indicate the degree of presence. Participants watched a video recording of the first-person view of the virtual experience while pushing the joystick forward or pulling it backward, resulting in the slider on the screen moving up or down. If there was no force applied by the participant, the joystick would default back to the middle resulting in the slider on the screen moving to 0. Avidemux (version 2.8.0) was used to edit the screen recording files and split them into two parts, namely the part containing the practice room and the part containing the task room. The lab's physical dimensions were 4.6 × 6 meters.

3.4 Procedure

When participants entered the lab, they were greeted and asked to read and sign the informed consent. After filling out the demographics questionnaire, participants read the instructions on how to perform the experimental task in VR. The headset was then prepared by the researcher by starting the screen recording and selecting the correct environment depending on the participant's condition. The headset was then mounted by the participant. The practice room loaded once the participant pressed start using the controller. In the practice room, the participants had time to test out the look and feel of VR and ask questions if anything was unclear. These instructions were the same as in Experiment 1. Participants then used the elevator to get to the task room.

After the experimental task ended, the participants were prompted to remove the headset, take a seat at a table, and fill out the ITC-SOPI. Subsequently, participants read through the instructions on how to use the CRM. The instructions emphasized that the participants should focus on the video itself and not on the on-screen slider or the joystick. In the instructions, presence was described as "the feeling of being in a virtual space." To illustrate what presence means, two examples were given. The participants were told the following. "For example, if you feel a strong sense of presence, then you forget the physical world around you and experience the virtual world as if it were real. If you feel a weaker sense of presence, then you are more aware of the physical world around you and notice the artificial nature of the virtual world. In the experience that you just had, you might have felt a stronger or weaker sense of presence. Your feeling of presence may have also changed throughout the experience. This change or lack of change is what we are interested in."

While participants were reading the instructions, the researcher connected the headset to the PC and saved the file containing the positional coordinates and the video file from the headset before editing a copy of the video recording and splitting it into two separate files as mentioned above. The first file (a first-person video recording of the practice room) was then loaded into the CARMA software and prepared for the participants. The settings of the software were set so that the axis labels were labeled "Strong Presence" (upper axis label) and "Weak Presence" (lower axis label). The axis maximum and minimum values were set to 100 and −100 with nine steps in between and an axis starting value of 0. The color map and bin size (1 second) were left in default and the sampling rate was set to the maximum of 30 Hz. After the participants finished reading the instructions, they completed a practice round using footage from the practice room. After finishing and saving the CARMA output file, the software was then loaded with the video from the task room and the participant repeated the CRM rating for the experimental task. Participants were then debriefed.

3.5 Manipulations

The within-subject manipulations remained the same. The experimental baseline conditions, however, were extended to include immersion (see Figure 5). Immersion was set by manipulating four factors, namely image resolution (Duh, Lin, Kenyon, Parker, & Furness, 2002), refresh rate (Bowman & McMahan, 2007), spatial sound (Poeschl, Wall, & Doering, 2013), and locomotion (Boletsis & Cedergren, 2019; Sayyad, Sra, & Hollerer, 2020). All are objective characteristics of a VR system (Slater, 1999) that determine the degree to which a vivid illusion of reality can be delivered (Slater & Wilbur, 1997). The resolution settings were toggled between 512 (low immersion) and 2,560 pixels (high immersion) as these were the lowest and highest settings that did not affect the performance. The refresh rate was

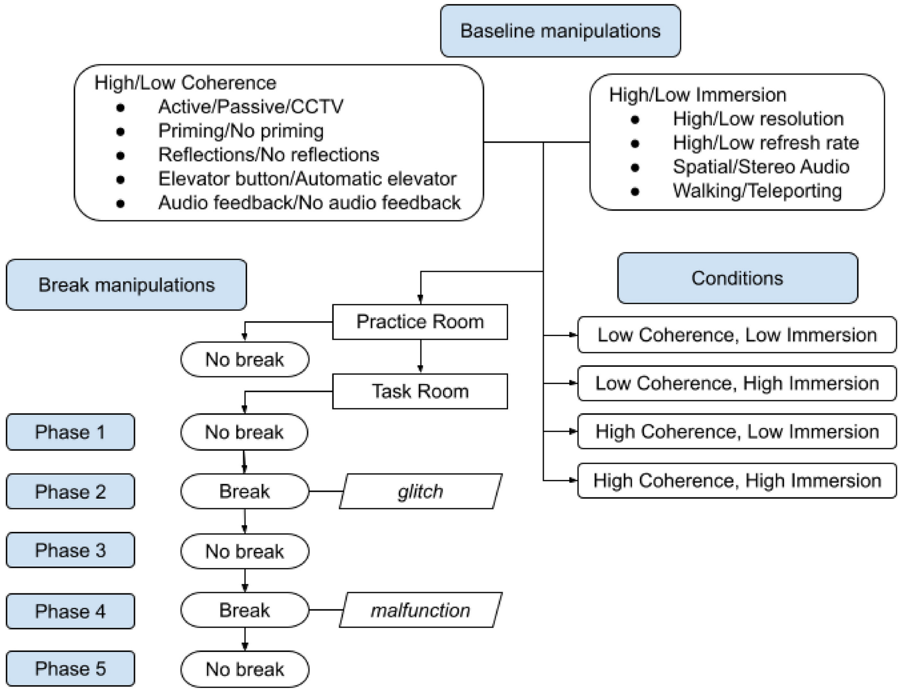


Figure 5. *Flowchart depicting the manipulations and sequence of steps in Experiment 2.*

toggled between 60 and 72 Hz. According to the literature, a refresh rate of 60 Hz has been shown to cause significantly lower quality experiences due to higher average nausea ratings (Weech, Kenny, & Barnett-Cowan, 2019) and user performance decreases (Wang, Shi, Zheng, Xie, Kao, & Liang, 2023). Locomotion was manipulated by instructing participants to either use only teleportation (low immersion) or only free walking (high immersion) as these two are commonly used in VR applications (Boletsis & Cedergren, 2019). Walking is widely regarded as providing the best locomotion experience compared to teleportation, which results in lower quality experiences, higher reported nausea, and lower immersion (Boletsis & Cedergren, 2019; Sayyad, et al., 2020). The sound was manipulated by toggling between spatial and mono. Spatial sound that creates a feeling of 3-dimensional audio by using head-tracked binaural audio has been found to improve immersion and navigation abilities in VR users compared to mono (Chandrasekera, Yoon, & D'Souza, 2015; Potter, Cvetković, & De Sena, 2022). The sum of these manipulations allowed us to adjust the boundaries of immersion reliably.

3.6 Data Analysis

CRM data was used to correlate the subjective ratings of the participant's degree of presence with the DFA exponents. To compare the two measures, the CRM data, which was also in the form of a time-series, was aligned with the positional coordinates time-series and split along the same time stamps. The means of each CRM output were calculated for each respective phase and each participant and compared to the mean DFA outputs by calculating correlation coefficients.

3.7 Results

A data set of $n = 27$ participants was collected. There were four combinations of high/low coherence/immersion settings resulting in a total of four conditions (see Table 1). Of all participants, the average age was $M = 22.63$, $SD = 2.92$ with $n = 14$ female, $n = 12$ male, and $n = 1$ participant who did not wish to indicate their gender. When asked about VR use, none of the participants indicated that they own a VR headset. Still,

Table 1. *List of Between-Subject Conditions for Both Experiments with Mean DFA Results*

Experiment 1					
Condition	Coherence		N	*M* DFA	*SD* DFA
Low Coherence	low		17	.799	.014
Low Coherence	high		17	.858	.014
Experiment 2					
Condition	Coherence	Immersion	N	*M* DFA	*SD* DFA
Low Coherence, Low Immersion	low	low	7	.967	.009
High Coherence, Low Immersion	high	low	7	.978	.009
Low Coherence, High Immersion	low	high	7	.859	.009
High Coherence, High Immersion	high	high	6	.947	.010

of all participants, $n = 15$ indicated that they had tried VR before and $n = 12$ mentioned that they had never tried VR before. When asked about their experience in non-VR gaming, participants indicated an average of around $M = 3.74$, $SD = 3.10$ hours of gaming per week with an average session length of $M = 1.22$, $SD = .76$.

To determine if our method could assess baseline differences in immersion between conditions, we performed an RM ANOVA. After confirming homogeneity in variance for each condition through Levene's tests, $p > .05$, the DFA method provided a significant result, assessing a difference in mean DFA exponents between at least one pair of conditions $F(3, 23) = 6.193, p = .003$. We proceeded to conduct in-depth pairwise comparisons for further analysis. During the examination of all conditions via pairwise comparisons, significant mean differences were found in pairs that either varied solely in their degree of immersion or both their degree of immersion and coherence. In general, differences were observed due to immersion alone or due to the combined aspects of immersion and coherence. The difference in immersion was found between the low coherence, low immersion condition, and the low coherence, high immersion condition. A difference in both immersion and coherence was found between the high coherence, low immersion condition, and the low coherence, high immersion condition (see Table 2). Based on these findings it appears that in Experiment 2 immersion is responsible for most of the between-subject effect.

Because immersion appears to strongly influence mean differences, the DFA method may potentially address baseline differences in immersion as effectively, or even better, than baseline differences in coherence. However, we cannot yet know if baseline differences in immersion are separately identified by DFA or if DFA simply cannot differentiate between coherence and immersion. The mixed results suggest that the DFA method may pick up on immersion, coherence, or both. Whether one of the two has a stronger effect can therefore not be concluded from only the above results.

To corroborate the findings of Experiment 1 and test whether phases were distinguishable via DFA, a within-subject effect was tested. As Sphericity was not met, the Hyunh-Feldt correction was chosen. However, in Experiment 2, there was no within-subject effect and no interaction effect found, $p > .05$.

Similarly to the DFA method, the CRM method was also subjected to an RM ANOVA analysis to detect within-subject and between-subject effects. With Sphericity met, $p > .05$, the subjective CRM method demonstrated that it was able to detect changes in phases $F(4, 92) = 11.376, p < .001$ but not differences in baselines $F(3, 23) = 11.723, p = .134$. Hence, based on the RM ANOVA results alone, the subjective/objective

Table 2. *Pairwise Comparisons for Conditions with Coherence and Immersion Settings for DFA Results*

Condition	Condition	Mean Difference	*SE*	*p*	Lower Bound	Upper Bound
Low Coherence, Low Immersion	High Coherence, Low Immersion	− .011	.031	1.000	− .099	.078
	Low Coherence, High Immersion	.108*	.031	.011	.020	.196
	High Coherence, High Immersion	.021	.032	1.000	− .071	.113
High Coherence, Low Immersion	Low Coherence, Low Immersion	.011	.031	1.000	− .078	.099
	Low Coherence, High Immersion	.119*	.031	.005	.030	.207
	High Coherence, High Immersion	.031	.032	1.000	− .061	.123
Low Coherence, High Immersion	Low Coherence, Low Immersion	− .108*	.031	.011	− .196	− .020
	High Coherence, Low Immersion	− .119*	.031	.005	− .207	− .030
	High Coherence, High Immersion	− .087	.032	.070	− .179	.005
High Coherence, High Immersion	Low Coherence, Low Immersion	− .021	.032	1.000	− .113	.071
	High Coherence, Low Immersion	− .031	.032	1.000	− .123	.061
	Low Coherence, High Immersion	.087	.032	.070	− .005	.179

NOTE. Values that are flagged with $*$ are significant results at the level of $p < .05$ (2-tailed).

measures CRM and DFA do not appear to align in Experiment 2. The nonsignificant interaction effect *$F(12, 92) = .935$, $p < .516$* for the CRM results show that the break manipulations during the phases appeared to affect the different conditions similarly.

In terms of assessing the break manipulations, as seen in Figure 6b, the DFA outcome measures in phase 2, during the first break manipulation, on average are higher compared to phase 1. In phase 4, during the second break manipulation, only the high coherence, high immersion group exhibited the same DFA pattern seen in the DFA exponents of Experiment 1 (see Figure 4). In the same high coherence, high immersion condition, the values of phases 2 and 4 appear to be higher than in phases 1, 3, and 5. All other conditions seemed to show a drop in DFA exponent values when going from phase 3 to 4. Nonetheless, while interesting, the pairwise comparisons between the phases did not yield any significant results in mean differences (see Appendix C for a table with all combinations).

Interestingly, while looking at the drop in DFA exponent values from phase 3 to phase 4, the CRM values (Figure 6[a]) reveal a similar drop from phase 3 to phase 4. The mean difference of *$M = 32.823$ ($p < .001$)* indicates much lower values during phase 4 compared to phase 3, which is in strong contrast to the nonsignificant differences between phases 1, 2, and 3 *($p > .05$)* (see Appendix C for a table with all combinations). This suggests that, when taking the effects and the visual results together, the second break manipulation during phase 4 may have been stronger than the first break manipulation during phase 2. While the DFA results in Figure 6(b) appear to show a decrease in phase 4 for most conditions, this was not significant and subsequently, the CRM results in Figure 6a appear to reflect this distinction more clearly than the DFA results.

Based upon visual inspection alone, the responses to the phase manipulations do not seem to match between subjective and objective methods. Only the condition of high coherence, high immersion appears to show uniform changes in phases 2 and 4 during both the CRM and the DFA method. When examining the pattern closely, it looks as though the pattern is the same zig-zag pattern in both Figures 6a and 6b, but simply flipped across the horizontal axis. Interestingly, this condition, by providing both high coherence and high immersion, enabled a "best case scenario," which may be the reason why the manipulations affected objective and subjective outcome measures similarly.

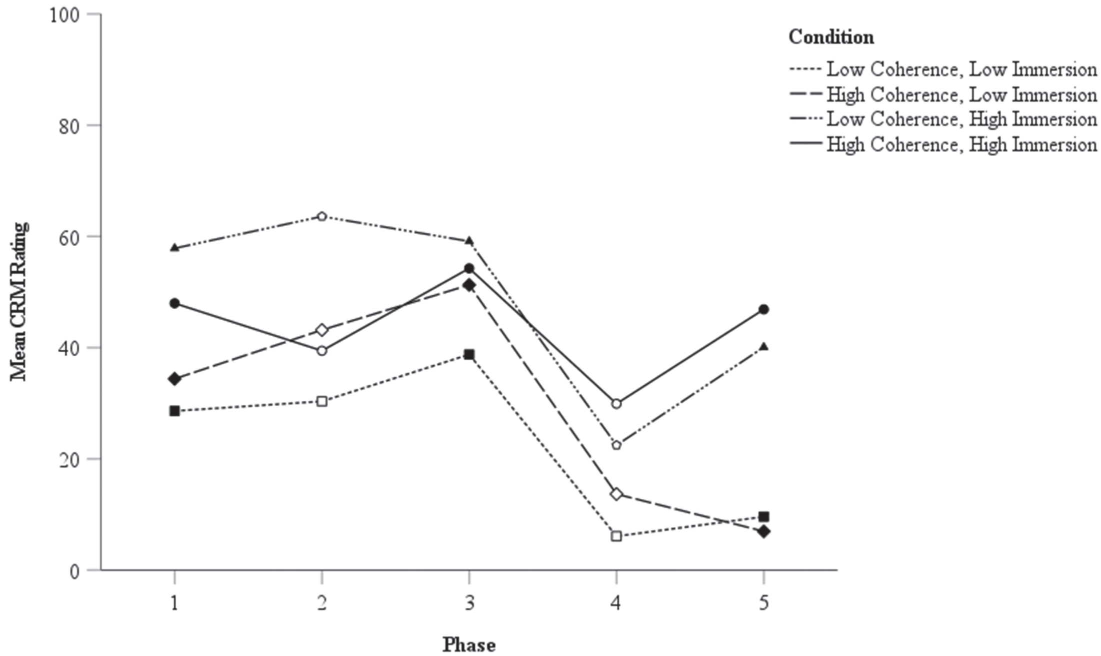
Finally, the correlation between the CRM and DFA exponents during phases was analyzed. To test this, the DFA exponents for each phase were correlated with the outcome measures of the CRM for each phase. While it was expected that all of the phases would correlate between both measures, only phases 2 (phase with break) *$r(25) = -.485$, $p = .010$,* and phases 5 (phase without a break) *$r(25) = -.444$, $p = .020$* showed correlations for corresponding phases (see Table 3). The results illustrate the difficulty of capturing the construct of presence and its subcomponents with different approaches. Next to the interpretation and directionality of the data, methods may differ in many aspects including a participant's recollection or subjective understanding of the construct, leading to results that are difficult to compare.

The presence questionnaire ITC-SOPI did not significantly predict the DFA exponents, *$p > .05$* (see Table 4). This again underlines the issue addressed in the above paragraph regarding the difficulty of capturing the construct of presence and its subcomponents by finding overlapping results from different assessment methods. It is also interesting to note that, compared to Experiment 2, in Experiment 1 the subscale negative effects from the presence questionnaire ITC-SOPI were able to predict DFA exponents in some of the phases. This illustrates the subtly different distributions between the sample with a high amount of experience versus a sample with less overall experience.

Gender predicted the DFA exponents for phase 2 (phase with first break) *$R^2 = .188$, $F(1, 25) = 5.796$, $p = .024$*, while age predicted the DFA exponents for phase 4 (phase with second break) (*$R^2 = .159$, $F(1, 25) = 6.829$, $p = .040$*). In the previous Experiment 1, only the number of hours played per week was able to predict the DFA exponents for some of the phases. Gender and age showed no predictive value in Experiment 1, which again illustrates how different the sample was. On the contrary, this may also point toward how the predictive weight of some demographics lessens or changes

(a)

Average subjective outcome measure.



(b)

Average objective outcome measure.

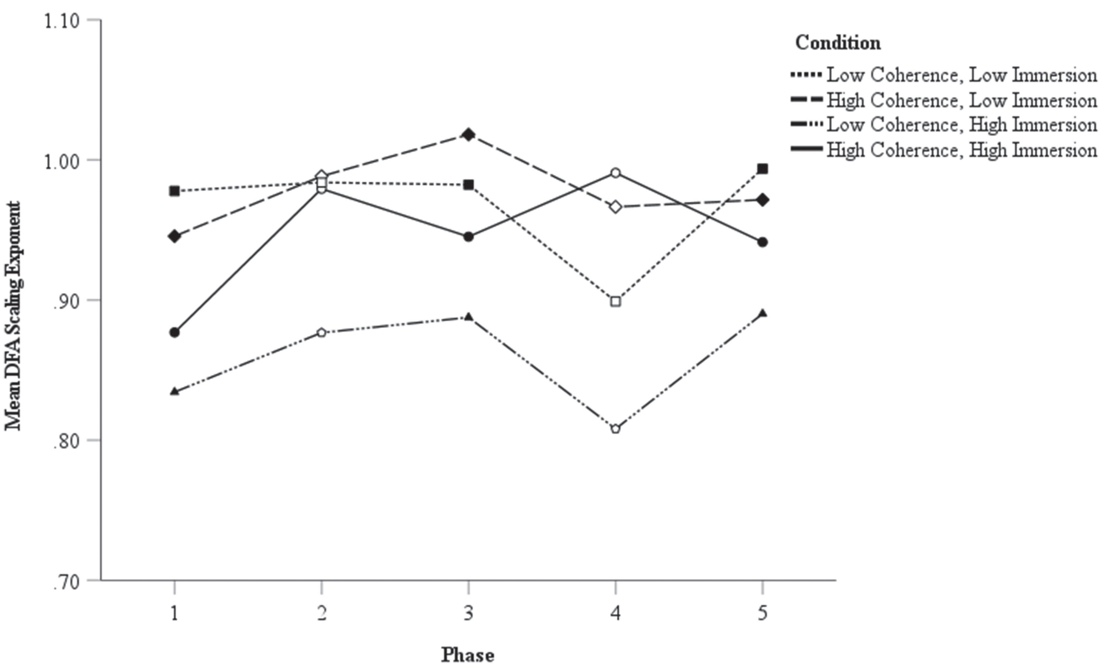


Figure 6. *Two line graphs displaying the average outcome measure for all phases in each condition.*

NOTE. White markers represent phases that include breaks; black markers represent phases that do not include breaks.

Table 3. *DFA Exponent and CRM Correlation Matrix Comparing the Correlation of Outcome Measures across Phases*

	CRM phase 1		CRM phase 2		CRM phase 3		CRM phase 4		CRM phase 5	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
DFA phase 1	− .188	.349	− .150	.451	− .051	.799	− .010	.961	.005	.982
DFA phase 2	− .447*	.019	− .485*	.010	− .273	.168	− .320	.164	− .326	.097
DFA phase 3	− .335	.088	− .310	.116	− .078	.700	− .496**	.008	− .527**	.005
DFA phase 4	− .302	.126	− .240	.229	− .278	.160	.061	.764	.142	.481
DFA phase 5	− .452*	.018	− .335	.088	− .376	.053	− .172	.391	− .444*	.020

NOTE. Values that are flagged with ∗ or ∗∗ are significant results at the level of $p < .05$ (2-tailed) or $p < .01$ (2-tailed), respectively.

Table 4. *Descriptives of Questionnaire Subscales for the ITC-SOPI Presence Questionnaire*

Subscales	N	Minimum	Maximum	*M*	*SD*
Spatial presence	27	2.61	4.17	3.56	.37
Negative effects	27	1.00	4.17	2.57	.71
Engagement	27	2.62	4.38	3.50	.46
Ecological validity	27	2.00	4.00	3.11	.53

with increased experience. In Experiment 2, when asked about VR use, none of the participants indicated that they own a VR headset, with $n = 15$ indicating that they had tried VR before while $n = 12$ mentioned that they had never tried VR before. In Experiment 1 all participants owned their own VR headsets.

3.8 Interim Discussion

In Experiment 2, we investigated whether time-series analysis on movement data, registered during task execution, can assess differences in immersion. DFA showed that our method is also suited to detect baseline differences in immersion, supporting the first of our two hypotheses for Experiment 2. The pairwise comparisons between the four between-subject conditions solidified this result. In addition, we applied the same RM ANOVA to the CRM results to test whether the subjective and the objective measures displayed a similar ability in responding to manipulations in presence. The significant within-subject effect but non-significant between-subject effect of the CRM demonstrated that comparing the CRM to the DFA method yielded mixed results. Finally, we assessed the correlation between DFA exponents and CRM outcome measures to test how strongly associated the subjective and objective continuous measurements are overall and found that only some of the phases correlated in both outcome measures.

As the focus of this experiment was on how the method responds to changes in immersion, it remains unclear why immersion is a potentially stronger, or at least similarly strong differentiator between baseline conditions, compared to coherence. There may be a fundamental commonality in how immersion and coherence affect perception-action coupling through mental models. It is also possible that manipulated sensorimotor contingencies (the boundaries of immersion) affect the DFA exponents more strongly than manipulated reasonable circumstances (boundaries of coherence). This would be in line with findings from the congruence and plausibility (CaP) model (Latoschik & Wienrich, 2022), in which the bottom-up influences have a stronger effect on presence than top-down influences. One approach that may shed light on this topic would be to employ breaks in immersion that specifically target the place illusion. This could help determine if breaks in coherence and immersion lead to fundamentally different results in the context of detrended fluctuation analysis.

Our question of whether the DFA method can be an alternative to the CRM method could not be conclusively answered. While the DFA and CRM outputs correlate during some phases and both can detect significant effects for baseline differences or breaks, it is too early to infer that either method can substitute for the other. The primary reason for this is the lack of agreement across the DFA and CRM results for the phases in different conditions (see Figure 6). A potential avenue to further examine the relationship between CRM and DFA may be to increase the number of sources of positional coordinate data (e.g., torso, legs/feet) and therefore capture a participant's movement and the task space in more detail. This added data may provide a more complete picture of how the perception-action coupling affects the user, thereby increasing validity and potentially also the method's correspondence with subjective methods.

4 General Discussion

In the two experiments presented here, the main goal was to understand whether the complex dynamical systems framework, specifically DFA, was an appropriate approach to assess breaks in time-series data taken from participants' movement patterns in VR. The same approach was also used to test whether baseline differences in coherence and immersion could be assessed. Finally, the DFA exponents of within-subject levels were examined to understand whether DFA exponents could be predicted by presence ratings. The presence ratings included questionnaire data and a subjective self-rating method which produces continuous time-series data. The results showed that DFA is a promising behavioral method that responds to breaks and baseline differences, and partially corresponds to subjective measures. To our knowledge, the ability to differentiate phases that involve and don't involve breaks using a continuous measure has only ever been achieved with post-hoc self-report measures (Chung & Gardner, 2012; Garau, Friedman, Widenfeld, Antley, Brogni, & Slater, 2008), neurological (Porssut, Iwane, Chavarriaga, Blanke, Millán, Boulic et al., 2023; Si-Mohammed et al., 2020), or physiological methods (Liebold et al., 2017). Although our behavioral method demonstrates the capability to address this limitation, future iterations of this approach should be subjected to different study designs to assess performance in scenarios such as those where breaks are unpredictable, where tasks and movement are more complex, and where other measures are weighed against the DFA approach.

Considering this paper's limitations, it is important to address the choice and application of the complex dynamical systems framework and the specific focus on detrended fluctuation analysis (DFA). DFA is one of several methods applicable to analyze long-range correlations in human movements. While prior research has compared alternatives such as standardized dispersion analysis (SDA) and power spectral analysis (PSA) (e.g., Van Orden et al., 2003 and Wijnants, Cox, Hasselman, Bosman, & Van Orden, 2012), these methods have not outperformed DFA in handling non-stationary data and revealing fractal-like scaling. This robustness has led to DFA's widespread adoption in movement sciences (Den Hartigh et al., 2015; Den Hartigh, Marmelat, & Cox, 2018). Additionally, multifractal DFA has shown similarly promising results in identifying interaction-dominant dynamics in human movement analysis (Bloomfield, Lane, Mangalam, & Kelty-Stephen, 2021; Mangalam, Sadri, Hayano, Watanabe, Kiyono, & Kelty-Stephen, 2023). However, the choice for DFA aligns with the general consensus of being a robust and well-established method. Nonetheless, follow-up studies should consider applying additional methods from the nonlinear dynamics toolbox, like recurrence quantification analysis (e.g., Wijnants, Cox, et al., 2012).

Also, within the literature, the focus on detecting long-range correlations has been viewed critically due to its universal existence in nature and human movement (Ahn & Hogan, 2013), suggesting that these signals will inevitably be found regardless of any critical state. The authors would reply that DFA does not detect recurrent trends but rather captures the scaling relation that exists between fluctuations across multiple timescales (Bardet & Kammoun, 2008). This specific topic has been a subject of debate within recent years but did see a consensus amongst peers who deemed the existence of long-range dependence to be a definitive aspect of criticality within a complex dynamical system (Bak, Tang, & Wiesenfeld, 1987; Diniz et al., 2011; Van Orden et al., 2003).

Finally, the use of velocity time-series derived from distances between the head and hands may appear arbitrary, as interaction with the virtual environment could be defined by various task spaces. For example, head rotation (Maneuvrier, Nguyen, & Renaud, 2023) or gaze movements from eye-tracking (Renaud, Chartier, Albert, Décarie, Cournoyer, & Bouchard, 2007), have already successfully been used. However, this task space was chosen to represent the reaching and tossing motion central to the task at hand. While movements such as walking or picking up balls were also part of the task, they were less defining and harder to record due to the controller-based tracking limitations. Future studies could address this by incorporating additional sensors to capture movements like gait for a more comprehensive analysis.

An aim of this study, besides tackling the lack of continuity in presence measurements, was to also address the dependence on different types of virtual environments that evoke specific responses. In the case of physiological methods, an alternative continuous measurement, environments tend to rely on evoking specific responses such as arousal (Halbig & Latoschik, 2021), stress (Meehan, Insko, Whitton, & Brooks, 2002), and increased cognitive workload (Slater, 2009). This prevents the application of presence measurement to a wide range of experiences, a limitation also shared with most behavioral measures which capture only individual movements or interactions (Slater, 2004). While the DFA method addresses this limitation, more research using other environments is needed to expand on the validity of this solution.

There are only a handful of studies in the literature that utilize comparable behavioral methods and movement data either to evaluate presence alone (Renaud et al., 2002, 2007) or to examine related concepts such as the effect of nausea on postural control (Freeman, Chander, Kodithuwakku Arachchige, Turner, Jones, Pan, et al., 2023), gait (Maneuvrier, et al., 2023), or head rotation (Murata, 2004) in VR. Besides the small number of examples, studies that do employ DFA, frequently limit the complexity of interaction. Specifically, in the two studies by Renaud et al. (2002, 2007), the interaction was confined to a seated visual search task. Such restrictions, while necessary in some cases, don't reflect the vast variety of possible interactions afforded by VR environments. Nonetheless, these studies have demonstrated the promising ability to assess presence by using time-series movement data and the framework of complex dynamical systems analysis. With the experiments discussed in this paper, we sought to extend these findings by examining the approaches' applicability in more complex movements and environments without relying on specific responses.

In addition to continuity, independence of environments, and robustness, the issue of consistency between subjective and objective measures was an area of interest. The inconclusive link between DFA exponents and subjective outcome measures shows that this remains an obstacle, reflecting similar findings in the literature (Thorp, Rimol, & Grassini, 2023). This particular limitation of our study and presence research in general, is hard to solve because the validity of subjective presence measures themselves (specifically questionnaires) has been a source of debate (Graf & Schwind, 2020; Slater, 2004). Adding to the complexity, there exist numerous presence theories, definitions, and conceptualizations (Latoschik & Wienrich, 2022; Lombard & Ditton, 2006; Skarbez et al., 2018; Slater et al., 2022), overlapping and related concepts such as embodiment or flow (Forster, Karimpur, & Fiehler, 2022; Pianzola, 2021), and unclear links to physiology, cognition, and personality (Grassini & Laumann, 2020; Phillips, Interrante, Kaeding, Ries, & Anderson, 2012). We attempted to solve this by using the CRM method, which allows for continuous ratings and relies less on recollection but achieved only indefinite results. To enhance the validity of our method despite the mentioned limitations, different assessment methods, including neurological and physiological methods, should be used to corroborate findings and link subjective and objective measures.

During the study design, another important topic was the selection of break manipulations. The first break during phase 2 resembles a type of violation of the laws of physics. Similar manipulations have been employed in prior studies with success (Liebold et al., 2017; Skarbez et al., 2021). Particularly, the manipulation of gravity has been frequently employed due to its status as a so-called "strong prior," defined as an assumption about the world that has extremely high reliability (Jörges & López-Moliner, 2017). In the case of gravity, the violation of the assumption predictably breaks plausibility in a bottom-up nature. In comparison, so-called "weak priors" are assumptions that are less reliable and have less influence on perceptions and behavior. In our case, an example of a weak prior would be the second break during phase 4 in which the virtual environment demonstrates behavior that violates the priming instructions, and prior experiences provided about how the environment behaves. While the usage of strong priors to produce breaks seems almost trivial, using weak priors has been less popular in presence literature, specifically when employing breaks in presence. Brubach et al. (2022) even actively opted out of employing weak priors to increase the chances of breaking plausibility and achieving a significant result.

We found, however, that it is important to explore other priors besides gravity and pay closer attention to ecological validity since many practical VR applications would likely not involve changes in gravity or manipulations of other comparable strong priors. Hence, with the second break, the goal was to explore this gap in the literature and employ a manipulation that resembles a weak prior. To our knowledge, this has been the first time that a weak prior has been used to break the plausibility illusion during a continuous behavioral presence measure. More such tests could be immensely helpful for categorizing and defining breaks and pinpointing their respective use cases.

In upcoming studies, we aim to further explore the limits of this method and understand how the multitude of factors that influence presence are reflected in the movement data. We hope that the practicality of the framework helps move presence research closer toward standardization.

REFERENCES

Ahn, J., & Hogan, N. (2013). Long-range correlations in stride intervals may emerge from non-chaotic walking dynamics. *PLOS One*, *8*(9), e73239. 10.1371/journal.pone.0073239

Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do* (First edition). W. W. Norton & Company, Inc.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/*f* noise. *Physical Review Letters*, *59*(4), 381–384. 10.1103/PhysRevLett.59.381

Bardet, J.-M., & Kammoun, I. (2008). Asymptotic properties of the detrended fluctuation analysis of long-range-dependent processes. *IEEE Transactions on Information Theory, 54*(5), 2041–2052. 10.1109/TIT.2008.920328

Berkman, M. I., & Akan, E. (2019). Presence and immersion in virtual reality. In N. Lee (Ed.), *Encyclopedia of computer graphics and games* (pp. 1–10). Springer International Publishing. 10.1007/978-3-319-08234-9_162-1

Bloomfield, L., Lane, E., Mangalam, M., & Kelty-Stephen, D. G. (2021). Perceiving and remembering speech depend on multifractal nonlinearity in movements producing and exploring speech. *Journal of The Royal Society Interface*, *18*(181), 20210272. 10.1098/rsif.2021.0272

Boletsis, C., & Cedergren, J. E. (2019). VR locomotion in the new era of virtual reality: An empirical comparison of prevalent techniques. *Advances in Human-Computer Interaction*, *2019*, 1–15. 10.1155/2019/7420781

Bowman, D. A., Gabbard, J. L., & Hix, D. (2002). A survey of usability evaluation in virtual environments: Classification and comparison of methods. *Presence: Teleoperators and Virtual Environments*, *11*(4), 404–424. 10.1162/105474602760204309

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, *40*(7), 36–43. 10.1109/MC.2007.257

Brubach, L., Westermeier, F., Wienrich, C., & Latoschik, M. E. (2022). Breaking plausibility without breaking presence—Evidence for the multi-layer nature of plausibility. *IEEE Transactions on Visualization and Computer Graphics*, *28*(5), 2267–2276. 10.1109/TVCG.2022.3150496

Cerda, L., Fauvarque, A., Graziani, P., & Del-Monte, J. (2021). Contextual priming to increase the sense of presence in virtual reality: Exploratory study. *Virtual Reality*, *25*(4), 1105–1112. 10.1007/s10055-021-00515-4

Chandrasekera, T., Yoon, S.-Y., & D'Souza, N. (2015). Virtual environments with soundscapes: A study on immersion and effects of spatial abilities. *Environment and Planning B: Planning and Design*, *42*(6), 1003–1019. 10.1068/b130087p

Chung, J., & Gardner, H. J. (2012). Temporal presence variation in immersive computer games. *International Journal of Human–Computer Interaction*, *28*(8), 511–529. 10.1080/10447318.2011.627298

Den Hartigh, R. J. R., Cox, R. F. A., Gernigon, C., Van Yperen, N. W., & Van Geert, P. L. C. (2015). Pink noise in rowing ergometer performance and the role of skill level. *Motor Control*, *19*(4), 355–369. 10.1123/mc.2014-0071

Den Hartigh, R. J. R., Marmelat, V., & Cox, R. F. A. (2018). Multiscale coordination between athletes: Complexity matching in ergometer rowing. *Human Movement Science*, *57*, 434–441. 10.1016/j.humov.2017.10.006

Diniz, A., Wijnants, M. L., Torre, K., Barreiros, J., Crato, N., Bosman, A. M. T., Hasselman, F., Cox, R. F. A., Van Orden, G. C., & Delignières, D. (2011). Contemporary theories of 1/f noise in motor control. *Human Movement Science*, *30*(5), 889–905. 10.1016/j.humov.2010.07.006

Dotov, D. G., Nie, L., & Chemero, A. (2010). A demonstration of the transition from ready-to-hand to unready-to-hand. *PLOS One*, *5*(3), e9433. 10.1371/journal.pone.0009433

Duh, H. B.-L., Lin, J. J. W., Kenyon, R. V., Parker, D. E., & Furness, T. A. (2002). Effects of characteristics of image quality in an immersive environment. *Presence: Teleoperators and Virtual Environments*, *11*(3), 324–332. 10.1162/105474602317473259

Forster, P.-P., Karimpur, H., & Fiehler, K. (2022). Why we should rethink our approach to embodiment and presence. *Frontiers in Virtual Reality*, *3*, 838369. 10.3389/frvir.2022.838369

Fox, J., Bailenson, J. N., & Tricase, L. (2013). The embodiment of sexualized virtual selves: The Proteus effect and experiences of self-objectification via avatars. *Computers in Human Behavior*, *29*(3), 930–938. 10.1016/j.chb.2012.12.027

Freeman, H. R., Chander, H., Kodithuwakku Arachchige, S. N. K., Turner, A. J., Jones, J. A., Pan, Z., Hudson, C., & Knight, A. C. (2023). Postural control behavior in a virtual moving room paradigm. *Biomechanics*, *3*(4), 539–551. 10.3390/biomechanics3040043

Freeman, J., Avons, S. E., Pearson, D. E., & IJsselsteijn, W. A. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoperators and Virtual Environments*, *8*(1), 1–13. 10.1162/105474699566017

Garau, M., Friedman, D., Widenfeld, H. R., Antley, A., Brogni, A., & Slater, M. (2008). Temporal and spatial variations in presence: Qualitative analysis of interviews from an experiment on breaks in presence. *Presence: Teleoperators and Virtual Environments*, *17*(3), 293–309. 10.1162/pres.17.3.293

Girard, J. M. (2014). CARMA: Software for continuous affect rating and media annotation. *Journal of Open Research Software*, *2*, e5. 10.5334/jors.ar

González-Franco, M., Peck, T. C., Rodríguez-Fornells, A., & Slater, M. (2014). A threat to a virtual hand elicits motor cortex activation. *Experimental Brain Research*, *232*(3), 875–887. 10.1007/s00221-013-3800-1

Gonzalez-Franco, M., Perez-Marcos, D., Spanlang, B., & Slater, M. (2010). The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. *2010 IEEE Virtual Reality Conference (VR)*, 111–114. 10.1109/VR.2010.5444805

Gorini, A., Capideville, C. S., De Leo, G., Mantovani, F., & Riva, G. (2011). The role of immersion and narrative in mediated presence: The virtual hospital experience. *Cyberpsychology, Behavior, and Social Networking*, *14*(3), 99–105. 10.1089/cyber.2010.0100

Gottman, J. M., & Levenson, R. W. (1985). A valid procedure for obtaining self-report of affect in marital interaction. *Journal of Consulting and Clinical Psychology*, *53*(2), 151–160. 10.1037/0022-006X.53.2.151

Graf, S., & Schwind, V. (2020). Inconsistencies of presence questionnaires in virtual reality. *26th ACM Symposium on Virtual Reality Software and Technology*, 1–3. 10.1145/3385956.3422105

Grassini, S., & Laumann, K. (2020). Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology*, *11*, 349. 10.3389/fpsyg.2020.00349

Grönlund, A., Yi, I. G., & Kim, B. J. (2012). Fractal profit landscape of the stock market. *PLOS One*, *7*(4), e33960. 10.1371/journal.pone.0033960

Halbig, A., & Latoschik, M. E. (2021). A systematic review of physiological measurements, factors, methods, and applications in virtual reality. *Frontiers in Virtual Reality*, *2*, 694567. 10.3389/frvir.2021.694567

Hausdorff, J. M. (2007). Gait dynamics, fractals and falls: Finding meaning in the stride-to-stride fluctuations of human walking. *Human Movement Science*, *26*(4), 555–589. 10.1016/j.humov.2007.05.003

IJsselsteijn, W., Bouwhuis, D. G., & De Ridder, H. (2004). *Presence in depth*. Technische Universiteit Eindhoven.

IJsselsteijn, W., De Ridder, H., Hamberg, R., Bouwhuis, D., & Freeman, J. (1998). Perceived depth and the feeling of presence in 3DTV. *Displays*, *18*(4), 207–214. 10.1016/S0141-9382(98)00022-5

Jörges, B., & López-Moliner, J. (2017). Gravity as a strong prior: Implications for perception and action. *Frontiers in Human Neuroscience*, *11*, 203. 10.3389/fnhum.2017.00203

Kalina, E., & Johnson-Glenberg, M. C. (2020). Presence and platform: Effects of embodiment comparing a 2D computer and 3D VR game. *2020 6th International Conference of the Immersive Learning Research Network*, pp. 31–37. 10.23919/iLRN47897.2020.9155160

Kern, A. C., & Ellermeier, W. (2020). Audio in VR: Effects of a soundscape and movement-triggered step sounds on presence. *Frontiers in Robotics and AI*, *7*, 20. 10.3389/frobt.2020.00020

Latoschik, M. E., & Wienrich, C. (2022). Congruence and plausibility, not presence: Pivotal conditions for XR experiences and effects, a novel approach. *Frontiers in Virtual Reality*, *3*, 694433. 10.3389/frvir.2022.694433

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, *10*(3), 282–297. 10.1162/105474601300343612

Liebold, B., Brill, M., Pietschmann, D., Schwab, F., & Ohler, P. (2017). Continuous measurement of breaks in presence: Psychophysiology and orienting responses. *Media Psychology*, *20*(3), 477–501. 10.1080/15213269.2016.1206829

Lombard, M., & Ditton, T. (2006). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, *3*(2), 0. 10.1111/j.1083-6101.1997.tb00072.x

Maneuvrier, A., Nguyen, N.-D.-T., & Renaud, P. (2023). Predicting VR cybersickness and its impact on visuomotor performance using head rotations and field (in)dependence. *Frontiers in Virtual Reality*, *4*, 1307925. 10.3389/frvir.2023.1307925

Mangalam, M., Sadri, A., Hayano, J., Watanabe, E., Kiyono, K., & Kelty-Stephen, D. G. (2023). Multifractal foundations of biomarker discovery for heart disease and stroke. *Scientific Reports*, *13*(1), 18316. 10.1038/s41598-023-45184-2

Mansfield, G., Roy, T. K., & Shiratori, N. (2001). Self-similar and fractal nature of Internet traffic data. *Proceedings of the 15th International Conference on Information Networking*, pp. 227–231. 10.1109/ICOIN.2001.905432

Meehan, M., Insko, B., Whitton, M., & Brooks, F. P. (2002). Physiological measures of presence in stressful virtual environments. *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, p. 645. 10.1145/566570.566630

Murata, A. (2004). Effects of duration of immersion in a virtual reality environment on postural stability. *International Journal of Human-Computer Interaction*, *17*(4), 463–477. 10.1207/s15327590ijhc1704_2

Park, N., Min Lee, K., Annie Jin, S.-A., & Kang, S. (2010). Effects of pre-game stories on feelings of presence and evaluation of computer games. *International Journal of Human-Computer Studies*, *68*(11), 822–833. 10.1016/j.ijhcs.2010.07.002

Peng, C.-K., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *5*(1), 82–87. 10.1063/1.166141

Phillips, E., Portus, M., Davids, K., & Renshaw, I. (2012). Performance accuracy and functional variability in elite and developing fast bowlers. *Journal of Science and Medicine in Sport*, *15*(2), 182–188. 10.1016/j.jsams.2011.07.006

Phillips, L., Interrante, V., Kaeding, M., Ries, B., & Anderson, L. (2012). Correlations between physiological response, gait, personality, and presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, *21*(2), 119–141. 10.1162/PRES_a_00100

Pianzola, F. (2021). *Presence, flow, and narrative absorption questionnaires: A scoping review*. 10.31234/osf.io/8xvtp

Poeschl, S., Wall, K., & Doering, N. (2013). Integration of spatial sound in immersive virtual environments: An experimental study on effects of spatial sound on presence. *2013 IEEE Virtual Reality*, 129–130. 10.1109/VR.2013.6549396

Porssut, T., Iwane, F., Chavarriaga, R., Blanke, O., Millán, J. D. R., Boulic, R., & Herbelin, B. (2023). EEG signature of breaks in embodiment in VR. *PLOS One*, *18*(5), e0282967. 10.1371/journal.pone.0282967

Potter, T., Cvetković, Z., & De Sena, E. (2022). On the relative importance of visual and spatial audio rendering on VR immersion. *Frontiers in Signal Processing*, *2*, 904866. 10.3389/frsip.2022.904866

Renaud, P., Bouchard, S., & Proulx, R. (2002). Behavioral avoidance dynamics in the presence of a virtual spider. *IEEE Transactions on Information Technology in Biomedicine*, *6*(3), 235–243. 10.1109/TITB.2002.802381

Renaud, P., Chartier, S., Albert, G., Décarie, J., Cournoyer, L.-G., & Bouchard, S. (2007). Presence as determined by fractal perceptual-motor dynamics. *CyberPsychology & Behavior*, *10*(1), 122–130. 10.1089/cpb.2006.9983

Richardson, M. J., & Chemero, A. (2010). Complex dynamical systems and embodiment. In *The Routledge handbook of embodied cognition*. Routledge. 10.4324/9781315775845.ch4

Sadeghipour, A., & Kopp, S. (2011). Embodied gesture processing: Motor-based integration of perception and action in social artificial agents. *Cognitive Computation*, *3*(3), 419–435. 10.1007/s12559-010-9082-z

Sayyad, E., Sra, M., & Hollerer, T. (2020). Walking and teleportation in wide-area virtual reality experiences. *2020 IEEE International Symposium on Mixed and Augmented Reality*, 608–617. 10.1109/ISMAR50242.2020.00088

Schirm, J., Tullius, G., & Habgood, J. (2019). Towards an objective measure of presence: Examining startle reflexes in a commercial virtual reality game. *Extended Abstracts of the Annual Symposium on Computer–Human Interaction in Play Companion Extended Abstracts*, 671–678. 10.1145/3341215.3356263

Schoner, G., Dijkstra, T. M. H., & Jeka, J. J. (1998). Action-perception patterns emerge from coupling and adaptation. *Ecological Psychology*, *10*(3–4), 323–346. 10.1080/10407413.1998.9652688

Sekine, M., Akay, M., Tamura, T., Higashi, Y., & Fujimoto, T. (2004). Fractal dynamics of body motion in patients with Parkinson's disease. *Journal of Neural Engineering*, *1*(1), 8–15. 10.1088/1741-2560/1/1/002

Si-Mohammed, H., Lopes-Dias, C., Duarte, M., Argelaguet, F., Jeunet, C., Casiez, G., Muller-Putz, G. R., Lecuyer, A., & Scherer, R. (2020). Detecting system errors in virtual reality using EEG through error-related potentials. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces*, 653–661. 10.1109/VR46266.2020.00088

Simperingham, K. D., Cronin, J. B., & Ross, A. (2016). Advances in sprint acceleration profiling for field-based team-sport athletes: Utility, reliability, validity and limitations. *Sports Medicine*, *46*(11), 1619–1645. 10.1007/s40279-016-0508-y

Skarbez, R. (2016). *Plausibility illusion in virtual environments.* The University of North Carolina at Chapel Hill University Libraries. 10.17615/2MG3-GH93

Skarbez, R., Brooks, F. P., & Whitton, M. C. (2021). Immersion and coherence: Research agenda and early results. *IEEE Transactions on Visualization and Computer Graphics*, *27*(10), 3839–3850. 10.1109/TVCG.2020.2983701

Skarbez, R., Brooks, Jr., F. P., & Whitton, M. C. (2018). A survey of presence and related concepts. *ACM Computing Surveys*, *50*(6), 1–39. 10.1145/3134301

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *8*(5), 560–565. 10.1162/105474699566477

Slater, M. (2003). A note on presence terminology. *Presence Connect*, *3*(3).

Slater, M. (2004). How colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *13*(4), 484–493. 10.1162/1054746041944849

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557. 10.1098/rstb.2009.0138

Slater, M., Banakou, D., Beacco, A., Gallego, J., Macia-Varela, F., & Oliva, R. (2022). A separate reality: An update on place illusion and plausibility in virtual reality. *Frontiers in Virtual Reality*, *3*, 914392. 10.3389/frvir.2022.914392

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. 10.1162/pres.1997.6.6.603

Thorp, S. O., Rimol, L. M., & Grassini, S. (2023). Association of the big five personality traits with training effectiveness, sense of presence, and cybersickness in virtual reality. *Multimodal Technologies and Interaction*, *7*(2), 11. 10.3390/mti7020011

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, *132*(3), 331–350. 10.1037/0096-3445.132.3.331

Wang, J., Shi, R., Zheng, W., Xie, W., Kao, D., & Liang, H.-N. (2023). Effect of frame rate on user experience, performance, and simulator sickness in virtual reality. *IEEE*

Transactions on Visualization and Computer Graphics, *29*(5), 2478–2488. 10.1109/TVCG.2023.3247057

Warren, W. H. (1990). The perception-action coupling. In H. Bloch & B. I. Bertenthal (Eds.), *Sensory-motor organizations and development in infancy and early childhood* (pp. 23–37). Springer Netherlands. 10.1007/978-94-009-2071-2_2

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology*, *10*, 158. 10.3389/fpsyg.2019.00158

Werner, G. (2010). Fractals in the nervous system: Conceptual implications for theoretical neuroscience. *Frontiers in Physiology*. 10.3389/fphys.2010.00015

Wijnants, M. L., Cox, R. F. A., Hasselman, F., Bosman, A. M. T., & Van Orden, G. (2012). A trade-off study revealing nested timescales of constraint. *Frontiers in Physiology*, *3*. 10.3389/fphys.2012.00116

Wijnants, M. L., Hasselman, F., Cox, R. F. A., Bosman, A. M. T., & Van Orden, G. (2012). An interaction-dominant perspective on reading fluency and dyslexia. *Annals of Dyslexia*, *62*(2), 100–119. 10.1007/s11881-012-0067-3

Appendix A

Table 5. *List of Variables Recorded in CSV Output File*

Column Name	Description
date	The date of the participant's timeslot
time	The time of the participant's timeslot
runtime	The live runtime during the VR experiment
status	The events of the game including the game starting, the color of the ball that has been tossed, and which break occurred
presence state	The condition that the participant is assigned to
room ID	The room (practice room, task room, or elevator) that the participant is currently inside
camera transform position x	x-axis coordinate of the participant's head
camera transform position y	y-axis coordinate of the participant's head
camera transform position z	z-axis coordinate of the participant's head
left hand transform position x	x-axis coordinate of the participant's left hand
left hand transform position y	y-axis coordinate of the participant's left hand
left hand transform position z	z-axis coordinate of the participant's left hand
right hand transform position x	x-axis coordinate of the participant's right hand
right hand transform position y	y-axis coordinate of the participant's right hand
right hand transform position z	z-axis coordinate of the participant's right hand

Appendix B

The instructions below reflect the most important parts of the instructions that are relevant for the subjective experience and self-reporting. In the initial section, the instructions for how to experience and navigate the virtual environment of Experiments 1 and 2 are provided. In the following section, the instructions for completing the CRM task are provided.

Experiment 1

B1:

Introducing the VR game

The game you are about to experience is a simple mini-game similar to classical arcade games like "Simon Says." The virtual world is divided into two parts. The first part is a room for you to get acquainted with the game and the look and feel of VR. In this first room, you will find a basket on the left side, some colorful balls on the floor, and a mirror on the right side. Have a look around the room, get a feel for the virtual space, and try out the game of tossing the balls into the basket.

At the far end of the room, you will find an elevator. The elevator will take you downstairs to the next room, the second part.

The second part is where you will play the game and where your arcade skills will be put to the test. Only your performance in the second room will count. Here you will see three baskets instead of one, with more balls on the floor.

B1.1: High coherence condition priming instruction

There are cameras on the ceiling that are watching you and are actively tracking your movement in each of the rooms. The video they are capturing will be used to analyze your behavior and determine whether you are following the instructions correctly.

B1.2: Low coherence condition priming instruction

There are cameras on the ceiling which are panning back and forth in each of the rooms. They are meant to provide a sense of scale and prevent nausea.

Playing the game

To play the game in the downstairs room, look around the room to find an active basket. The basket is active if the screen above the basket and the basket itself light up in the same color. Then, find a ball with a matching color and toss it into the appropriately colored basket.

When tossing the ball please do so from the opposite side of the room so that we can see how good your aim is. The goal is to throw the balls into the baskets as accurately as possible.

Note that the baskets will activate and deactivate once the correct ball enters the basket, so you need to pay close attention and see which basket is currently active. If you toss a ball with an incorrect color into a basket, then the ball will respawn and you can try again.

Once all the balls are gone the game will end and a menu will appear. Please select "Quit Game" and remove your headset. Removing your headset will then automatically end the video recording.

Navigating and interacting

Please walk around the virtual world naturally (using your feet) as long as your playspace allows. If your playspace is too small, use the joystick on the controller to move around. Use the trigger buttons to grab and let go of a ball. When grabbing a ball keep the trigger button pressed. When tossing a ball, release the trigger button at the appropriate moment.

B2:

In the following section, the instructions explain the construct of presence for the participant. This section is relevant for the participant to understand and correctly reflect on their own past virtual experience.

What we are measuring

We are going to measure a feeling called "presence." Presence is described as the feeling of being in a virtual space.

For example, if you feel a strong sense of presence, then you forget the physical world around you and experience the virtual world as if it were real. If you feel a weak sense of presence, then you are aware of the physical world around you and notice the artificial nature of the virtual world.

In the experience that you just had, you might have felt a stronger or weaker sense of presence. Your feeling of presence may have also changed throughout the experience. This change or lack of change is what we are interested in.

Appendix C

Table 6. *Left-Hand Within-Subject Effect*

	Sum of Squares	df	Mean Square	F	p	$\eta 2$
Phases	.478	4	.119	8.391	$<$ .001	.207
Phases × Condition	.127	4	.031	2.227	.069	.065
Error	1.822	128	.014			

Table 7. *Left-Hand Between-Subject Effect*

	Sum of Squares	df	Mean Square	F	p	$\eta 2$
Phases	130.950	1	130.949	4375.424	$<$ .001	.992
Phases × Condition		1	.142	4.761	.036	.129
Error	.958	32	.029			

Table 8. *Right-Hand Within-Subject Effect*

	Sum of Squares	df	Mean Square	F	p	$\eta 2$
Phases	.112	4	.028	2.191	.073	.064
Phases × Condition	.134	4	.033	2.626	.037	.075
Error	.075	128	.012			

Table 9. *Right-Hand Within-Subject Effect*

	Sum of Squares	df	Mean Square	F	p	$\eta 2$
Phases	103.228	1	103.228	6790.483	$<$ .001	.995
Phases × Condition	.150	1	.150	9.870	.004	.235
Error	.486	32	.015			

Table 10. *Pairwise Comparisons for Phases Averaged across All Four Conditions of CRM Results*

Phase	Phase	Mean Difference	*SE*	*p*	Lower Bound	Upper Bound
1	2	−1.926	2.632	1.000	−10.096	6.245
	3	−8.646*	2.101	.004	−15.167	−2.124
	4	24.178*	6.889	.019	2.795	45.560
	5	16.327	6.331	.196	−3.345	35.999
2	1	1.926	2.632	1.000	−6.245	10.096
	3	−6.720	3.603	.750	−17.902	4.463
	4	26.103*	6.417	.005	6.185	46.022
	5	18.253	6.684	.119	−2.495	39.000
3	1	8.646*	2.101	.004	2.124	15.167
	2	6.720	3.603	.750	−4.463	17.902
	4	32.823*	7.090	.001	10.815	54.832
	5	24.973*	6.581	.009	4.545	45.400
4	1	−24.178*	6.889	.019	−45.560	−2.795
	2	−26.103*	6.417	.005	−46.022	−6.185
	3	−32.823*	7.090	.001	−54.832	−10.815
	5	−7.851	6.199	1.000	−27.093	11.392
5	1	−16.327	6.338	.169	−35.999	3.345
	2	−18.253	6.684	.119	−39.000	2.495
	3	−24.973*	6.5818	.009	−45.400	−4.545
	4	7.851*	6.199	1.000	−11.392	27.093

Table 11. *Pairwise Comparisons for Conditions of CRM Results*

Condition	Condition	Mean Difference	*SE*	*p*	Lower Bound	Upper Bound
Low Coherence, Low Immersion	High Coherence, Low Immersion	−7.215	4.914	.866	−20.379	5.948
	Low Coherence, Low Immersion	−25.937*	4.914	.000	−39.101	−12.774
	High Coherence, High Immersion	−21.011*	5.114	.000	−34.712	−7.310
High Coherence, Low Immersion	Low Coherence, Low Immersion	7.215	4.914	.866	−5.948	20.379
	Low Coherence, Low Immersion	−18.722*	4.914	.001	−31.885	−5.559
	High Coherence, High Immersion	−13.795*	5.114	.047	−27.496	−.095
Low Coherence, High Immersion	Low Coherence, Low Immersion	25.937*	4.914	.000	12.774	39.101
	High Coherence, Low Immersion	18.722*	4.914	.001	5.559	31.885
	High Coherence, High Immersion	4.927	5.114	1.000	−8.774	18.628
High Coherence, High Immersion	Low Coherence, Low Immersion	21.011*	5.114	.000	7.310	34.712
	High Coherence, Low Immersion	13.795*	5.114	.047	0.095	27.496
	High Coherence, High Immersion	−4.927	5.114	1.000	−18.628	8.774

Table 12. *Pairwise Comparisons for Phases in the High Coherence Condition Comparing All Possible Combinations with and without Breaks*

High Coherence Phases					95% Confidence Interval	
Phase	Phase	Mean Difference	*SE*	*p*	Lower	Upper
1	2	−.154	.039	.004	−.271	−.037
	3	−.135	.035	.005	−.242	−.029
	4	−.165	.038	.001	−.281	−.050
	5	−.050	.037	1.000	−.163	.062
2	1	.154	.039	.004	.037	.271
	3	.018	.027	1.000	−.062	.099
	4	−.012	.029	1.000	−.100	.076
	5	.104	.039	.131	−.015	.222
3	1	.135	.035	.005	.029	.242
	2	−.018	.027	1.000	−.099	.062
	4	−.030	.027	1.000	−.111	.051
	5	.085	.032	.113	−.010	.181
4	1	.165	.038	.001	.050	.281
	2	.012	.029	1.000	−.076	.100
	3	.030	.027	1.000	−.051	.111
	5	.115	.035	.023	.011	.220
5	1	.050	.037	1.000	−.062	.163
	2	−.104	.039	.131	−.222	.015
	3	−.085	.032	.113	−.181	.010
	4	−.115	.035	.023	−.220	−.011