



The Plausibility Paradox for Resized Users in Virtual Environments

Matti Pouke*, Katherine J. Mimnaugh, Alexis P. Chambers, Timo Ojala and Steven M. LaValle

Center for Ubiquitous Computing, Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland

This paper identifies and confirms a perceptual phenomenon: when users interact with simulated objects in a virtual environment where the users' scale deviates greatly from normal, there is a mismatch between the object physics they consider realistic and the object physics that would be correct at that scale. We report the findings of two studies investigating the relationship between perceived realism and a physically accurate approximation of reality in a virtual reality experience in which the user has been scaled by a factor of ten. Study 1 investigated perception of physics when scaled-down by a factor of ten, whereas Study 2 focused on enlargement by a similar amount. Studies were carried out as within-subjects experiments in which a total of 84 subjects performed simple interaction tasks with objects under two different physics simulation conditions. In the *true physics* condition, the objects, when dropped and thrown, behaved accurately according to the physics that would be correct at that either reduced or enlarged scale in the real world. In the *movie physics* condition, the objects behaved in a similar manner as they would if no scaling of the user had occurred. We found that a significant majority of the users considered the *movie physics* condition to be the more realistic one. However, at enlarged scale, many users considered *true physics* to match their expectations even if they ultimately believed *movie physics* to be the realistic condition. We argue that our findings have implications for many virtual reality and telepresence applications involving operation with simulated or physical objects in abnormal and especially small scales.

Keywords: virtual reality, perception, scaling, plausibility, human factors

1 INTRODUCTION

Many studies have confirmed the so-called “body-scaling effect”: if presented with mismatching size cues, humans tend to use their visible body as the dominant cue when perceiving sizes and distances (van der Hoort et al., 2011; Banakou et al., 2013; Linkenauger et al., 2013; Langbehn et al., 2016; Ogawa et al., 2017). For example, if a person was somehow shrunk to the size of a doll, the person would be inclined to regard the world as scaled-up and him/herself as normal-sized (van der Hoort et al., 2011). In this paper, we investigate the human perception of physics, specifically when subjects have been either scaled down or up by a significant amount. We believe this relatively underrepresented topic has implications to various virtual reality (VR) and telepresence applications. More specifically, we focus on the plausibility of rigid body dynamics when subjects are presented with realistic and unrealistic approximations of object motions when either scaled down or scaled up by a factor of ten. Previously, we investigated the perception of physics in VR when subjects were scaled down by a factor of ten (Pouke et al., 2020). We found that subjects considered a physics model of regular human scale to be more realistic than an accurate approximation of physics in the scaled down environment. This offered additional proof of humans

OPEN ACCESS

Edited by:

Ronan Boulic,
École Polytechnique Fédérale de
Lausanne, Switzerland

Reviewed by:

Niels Christian Nilsson,
Aalborg University Copenhagen,
Denmark

Richard Skarbez,
La Trobe University, Australia

*Correspondence:

Matti Pouke
matti.pouke@oulu.fi

Specialty section:

This article was submitted to
Virtual Reality and Human Behavior,
a section of the journal
Frontiers in Virtual Reality

Received: 19 January 2021

Accepted: 08 April 2021

Published: 30 April 2021

Citation:

Pouke M, Mimnaugh KJ,
Chambers AP, Ojala T and LaValle SM
(2021) The Plausibility Paradox for
Resized Users in Virtual Environments.
Front. Virtual Real. 2:655744.
doi: 10.3389/fvир.2021.655744

being oriented to Newtonian physics taking place at human scale, and anything deviating much from that scale appears unnatural. In this paper we present again the work of Pouke et al. (2020), while extending our prior work with new results and discussion considering scaled-up subjects. This allows us to compare the perception of physics in VR both in small and large scales.

Currently, not much is known about how scaling a person would affect their perception of physical phenomena, such as accelerations. Interestingly, if we consider the interaction of scaled-down characters with their surroundings in many works of fiction, the tendency to represent the world as scaled up in comparison to normal-sized protagonists can be observed. Early examples can be seen in the classic film *The Incredible Shrinking Man*. When the main character throws grains of sand off the table while insect-sized, the grains accelerate and fall as if they were boulders - when they should be falling down instantly. Similarly, when the character is awash with rainwater holding onto a pencil, the water and the pencil act more akin to a river and a log when the pencil should be bobbing with few waves and no visible whitewater should be apparent. Although the deficiencies in the realism of the *Incredible Shrinking Man* can be attributed to 1950s technologies, similar inaccuracies still remain in modern movies from *Honey I Shrunk the Kids* to *Downsizing*. These inaccuracies are not necessarily resulting from directors' lack of understanding of physics, but might be conscious choices to represent what the viewers would expect.

VR and telepresence applications allow humans to live through experiences such as the *Incredible Shrinking Man* through the eyes of a scaled-down entity. A specific category of virtual environments (VEs) providing such experiences are multiscale collaborative virtual environments (mCVEs), in which multiple users can collaborate in, for example, architectural or medical visualizations across multiple, nested levels of scale (e.g., Kopper et al., 2006; Zhang and Furnas 2005). In addition, the scaling of users has been utilized in several collaborative mixed reality (MR) systems (e.g., Billingham et al., 2001; Piumsomboon et al., 2018b; Piumsomboon et al., 2018a). Teleoperation of robots can allow humans to interact with the physical world at micro- and nanoscale. Similar to mCVEs, robotic teleoperation systems using multiple scales are beginning to emerge (Izumihara et al., 2019). Although teleoperation in the physical world can leverage stereoscopic camera systems resembling immersive VR applications (Hatamura and Morishita, 1990), purely virtual representations leveraging computer graphics can be used in, for example, educational and training systems for micro- and nanoscale tasks (Boloipon and Régner, 2013; Millet et al., 2008). Robotic surgery systems can perform operations at a microscopic level (Hongo et al., 2002) whereas stereoscopic VR can be utilized in telesurgery (Shenai et al., 2014). The benefits of VEs have been identified in various design and prototyping processes (Mujber et al., 2004) that can be extended into small-scale VEs, as well. Already 2 decades ago, both the design (Li and Sitte, 2001) and assembly (Alex et al., 1998) of microelectromechanical systems (MEMS) were prototyped through desktop VEs. Recent studies have also investigated self-scaling as a method to help with aspects related to architectural and interior design (Zhang et al., 2020a, Zhang et al., 2020b). Finally, self-scaling has also

been utilized as a locomotion method in VR, allowing one to cross large distances by physical walking (e.g., Krekhov et al., 2018; Abtahi et al., 2019).

Understanding human perception of scale-varying phenomena will be useful for the future design of applications such as those listed above. Although existing research has addressed many perceptual questions, such as the perception of distance and dimensions after altering one's virtual size (e.g., van der Hoort et al., 2011; Banakou et al., 2013; Kim and Interrante, 2017), the perception of the behavior of physical objects has received relatively little attention. There are many potential future use cases for user scaling that might require interaction with physical or physically simulated objects. We argue that it is not intuitive for humans to correctly perceive physical phenomena, such as rigid body dynamics, in scales that differ greatly from a normal human scale. An object dropped from 20 cm takes significantly less time to fall than an object dropped from 2 m, and their perceived accelerations are different. Additional physical phenomena, such as fluid dynamics, frictions, and static electricity might affect interactions even further as the scale of the operations becomes smaller. For this reason, additional consideration is required when designing systems in which real or virtual interactions take place on atypical scales, and thus it is important to understand human perception of physical phenomena on those atypical scales.

In this paper, we present our results on human perception of physics at abnormal scales. First, we focus on the mismatch between perceived realism and a physically accurate approximation of reality when interacting in a VE while scaled down by a factor of ten. Then, we present the results of a similar study where subjects were scaled up by a factor of ten and compare the results between the two studies. Based on previous research, we believe that humans generally perceive themselves at the correct scale when presented with mismatching size cues, as long as visual body cues are present (Langbehn et al., 2016). We also believe humans are generally accustomed to rigid body dynamics taking place at a human scale and under normal gravity conditions (McIntyre et al., 2001). Therefore, we hypothesize that subjects neither accept the realistic approximation of physics at an abnormal scale, nor are they blind to changes in scale. Instead, when presented with two different scale-dependent rigid body dynamics, they are more likely to consider the physically inaccurate one to be the more perceptually realistic one.

This paper is structured as follows. **Section 2** reviews previous research related to this work. **Section 3** presents the research method, experimental setup and the results of Study 1. **Section 4** similarly reports Study 2 and also compares the results of both studies. **Section 5** discusses our findings and **Section 6** concludes the paper.

2 RELATED RESEARCH

2.1 Perspective

The manipulation of a user's scale can be accomplished by changing various properties of the virtual character the user is

controlling in the VE. Changing these properties has various subjective effects. When scaling a user's virtual size, one of the most obvious properties to change is the viewpoint height, as it defines the virtual camera origin in relation to the VE, simulating a change in physical size. Viewpoint height affects egocentric distance perception (Zhang and Furnas 2005; Leyrer et al., 2011). Interestingly, minor changes in viewpoint height might go unnoticed by users (Leyrer et al., 2011; Deng and Interrante, 2019). Users' interaction capabilities such as locomotion speed and interaction distance can be changed according to scale, depending on the purpose of the application (Zhang and Furnas, 2005). When using a head mounted display (HMD), the scaling of the user can also affect the virtual interpupillary distance (IPD), which is the distance between the two virtual cameras that are used to render the environment for the user. Changing this distance can affect the user's sense of their own size relative to the VE (Kim and Interrante, 2017; Piumsomboon et al., 2018a).

2.2 Body Scaling

As already mentioned, *body scaling* refers to humans utilizing their own body as a primary scale cue, hence the virtual representation of the user's body greatly affects their perception of sizes and distances in a VE (Ogawa et al., 2017; Ogawa et al., 2019). Linkenauger et al. (2013) studied the role of one's hand as a metric for size perception; they conducted an experiment where they scaled the users' virtual hand and found out that it had a strong correlation with perceived object size. Ogawa et al. (2019) studied the effect of hand visual fidelity on object size perception and found that the visual realism of the hand affects the extent of the body scaling effect. van der Hoort et al. (2011) embodied the entire user in a doll's body as well as in a giant's body using a stereoscopic video camera system and an HMD. They found that the embodiment significantly affected the users' distance and size perceptions, especially if the user experienced a strong body ownership illusion (Slater et al., 2009) with the virtual body. Banakou et al. (2013) compared the effects of embodying the user as a child vs. as a scaled-down adult. They found that the effect of altered size and distance perceptions was even larger when embodied as a child, and it also made the users associate themselves with childlike personality traits.

2.3 Environmental Cues

The environment, whether real or virtual, affects the perception of scale. There is evidence of humans generally underestimating egocentric distances in VEs, except when the VE is faithfully modeled to represent a real environment (Renner et al., 2013). However, if a familiar room is scaled slightly up or down, underestimations are reintroduced (Interrante et al., 2008). Familiar size cues also affect the sensitivity to eye height manipulations (Deng and Interrante, 2019). Langbehn et al. (2016) studied the effect of body and environment representations as well as the scale of external avatars on users' perception of dominant scale in mCVEs (the dominant referring to the "true" scale in an mCVE system where users can coexist in multiple scales). They found that humans tended to use

their body as the primary metric for judging their own size and the environment if the representation of one's own body was not available. In addition, an environment with familiar size cues helps in the determination of scale, whereas an abstract environment does not. They also found that the majority of subjects tended to estimate external avatars to be at the dominant scale instead of themselves.

2.4 Perception of Physics

Previous research suggests that humans have an internal physics model according to which they expect the world to function. Studies in micro- and nanoscale teleoperation have revealed that, due to changes in physics, interactions at these scales can become difficult for the human operator, but education inside virtual reality environments has been found to alleviate this drawback (Sitti, 2007; Millet et al., 2008). McIntyre et al. (2001) reported a study in which astronauts' movements to catch a vertically moving ball were more inaccurate in zero gravity (0 g) in comparison to earth gravity (1 g); this was interpreted as evidence that the central nervous system utilizes an internal model of gravity in addition to visual judgment of acceleration to synchronize movements. Senot et al. (2005) used VR to study human estimation capabilities to intercept moving balls and found further evidence on subjects being more capable of intercepting objects accelerating according to normal gravity. Yao and Hayward (2006) created a haptic illusion of an object rolling or sliding inside a cavity and studied the subjects' capability of estimating the lengths of virtual tubes. According to their results, the subjects performed better than chance in estimating the tube lengths even using reduced sensory cues, indicating a capability to estimate object movements under the influence of gravity.

Ullman et al. (2017) compared humans' internal physics model to a contemporary game engine. Their findings suggest that although humans are not entirely capable of accurately predicting object motions, they are capable of making noisy, "good enough" approximations of Newtonian physics which can be compared to the physics simulations generated by physics engines that are integrated in contemporary game engines. McCoy and Ullman (2019) asked more than a thousand subjects to rate the "effort" required by various imaginary magical spells violating physics and found the subjects' responses as strikingly consistent. Despite describing completely imaginary phenomena, the subjects were very consistent in defining relative efforts that seemed to depend not only the type of the spell (such as levitate or conjure) but the size of the target as well. Although this finding is not directly related to the perception of physical phenomena, it again speaks for internal intuitive physics model that is consistent across humans.

2.5 Presence and Plausibility

The concepts of immersion (Slater and Wilbur, 1997), presence, and plausibility (Slater, 2009) are relevant for this study. In Slater's classical definition, the level of immersion refers to the level of technical fidelity of the VR system (i.e., resolution, field of view, vividness of graphics; Slater and Wilbur, 1997). In addition,

the realism of the user's response to the VR system depends on two orthogonal components, presence or place illusion (PI) and the plausibility illusion (PSI; Slater, 2009). PI refers to the sensation of being in another place, whereas PSI refers to the perceived credibility of the virtual scenario or experience (the illusion of being there vs. the realness of what is happening; Rovira et al., 2009). These illusions can occur despite the user of the VR system knowing for sure that the events are not actually taking place (Slater, 2009). PSI depends on the extent to which the VE can produce authentic responses for user actions. Rovira et al. (2009) argued that for PSI to occur, participants must perceive themselves as beings that exist in the VE; user actions must elicit actions in the VE and the VE must acknowledge the user (for example, virtual characters react to the user). In addition, the VE should match the user's prior knowledge and expectations. Skarbez et al. (2017b) used the term *coherence* to refer to the aspects of a VE that contribute to PSI, such as virtual humans and the behavior of virtual objects. They argued that although immersion is a technical attribute that affects PI, coherence is a similar technical attribute affecting PSI. A similar concept called *authenticity* was suggested by Gilbert (2016); according to Gilbert, the user's response to a VR system does not only depend on its technical attributes, but also on how the design of the VE matches the user's expectations. The components that influence PSI have been investigated by Skarbez et al. (2017b) and Bergström et al. (2017). In these studies, a methodology adopted from color matching theory was utilized to investigate the relative importance of various components, such as the behavior of virtual characters, physics, and audio.

In Study 1, we used the concept of PSI to study human perception of the behavior of physical objects while the subject was scaled down and interacting in a normal-sized environment. In Study 2, we repeated the same procedure for scaled-up subjects. However, we did not study multiple of components of PSI similarly to Skarbez et al. (2017b) or Bergström et al. (2017) in either study. Instead, we were interested in how subjects would perceive the coherence specifically in terms of the behavior of the virtual objects, when it would be reasonable to expect a mismatch between expectations and correctly simulated reality. In addition, we investigated whether the extent of the PI affected the results of our studies.

Building on the terminology discussed by Skarbez et al. (2017a), Skarbez et al. (2020), the phenomenon studied in this paper could also be referred to as *coherence - fidelity* mismatch; the logic expected by the users did not match with a more faithful representation of reality. It is expected that coherence differs from reality in, for example, fantasy games or other entertainment applications where PSI is maintained even when unearthly phenomena are taking place. In addition, Nilsson et al. (2017) discussed several cases where deliberately limiting a specific fidelity component led to an overall increase in perceived realism, compensating for the imperfect fidelity of another related component. They suggest that a phenomenon analogous to the *Uncanny Valley* (Mori et al., 2012) could exist in the relationship between fidelity and perceived realism.

3 STUDY 1

3.1 Study Design

3.1.1 Physics Conditions

The specific objective of Study 1 was to investigate the PSI of subjects in two different physics conditions. The purpose of both conditions was to visually represent a scaled-down subject in a normal-sized environment, and the physics simulations differed between the conditions as follows. In the *true physics* condition, the rigid body dynamics affect virtual objects in an approximately similar way to what would be accurate at that scale. In the *movie physics* condition (named after physical behavior as typically seen in Hollywood movies in scenes depicting scaled-down characters), rigid body dynamics behave in what would be the approximation of a normal human scale.

Our assumption was that the users would be able to distinguish the difference between *true physics* and *movie physics*, and we predicted that subjects would be more likely to expect and feel the *movie physics* condition to be the more perceptually realistic representation. This would suggest a Plausibility Paradox, a mismatch between perceived realism and the correct approximation of realism.

3.1.2 Hypotheses

We hypothesized that in the *true physics* conditions, the behavior of physical objects would feel incorrect for subjects despite their knowledge of being virtually shrunk down. More specifically, our hypotheses were as follows:

H1: For a scaled-down user, movie physics is more likely to feel realistic than true physics.

H2: For a scaled-down user, movie physics is more likely to match the user's expectations than true physics.

3.1.3 Virtual Environment

We designed a VE for the two physics conditions described above using Unreal Engine 4.22 (UE). In both conditions, the scaling operations took place in one order of magnitude, giving the impression of a doll-sized perspective. We did not use full body tracking or attempt to induce a strong body ownership illusion (Slater et al., 2009), so there was no visualization of any body parts in the VE other than the subject's hands. We used the default UE VR hand visualization for interaction and to present a medium-fidelity body size cue (Ogawa et al., 2019). There was no difference between the conditions regarding how the hands functioned or how the user was able to move.

To help in providing accurate size cues, we modeled the VE to resemble a location in the main corridor of the campus in which the study took place. The dimensions and materials of the VE were modeled after the real environment. In addition, we took measurements of various real-world objects, such as chairs, tables, and leaflets, which we modeled and scaled accordingly and placed in the VE as static objects.

The scaling of the user in the *true physics* condition was achieved by shrinking the user with the UE's built-in *World to Meters* parameter, which automatically scales the player



FIGURE 1 | First person perspective from the subject's point of view in the VE at the start of the experiment in Study 1 (Pouke et al., 2020).

character's height, virtual IPD and interaction distance. The skeletal meshes representing the player character's virtual hands were scaled down manually. In the *movie physics* condition, the player character properties were kept as default and the VE was scaled up instead. The purpose for this approach was to give the visual illusion of a scaled-down user, while retaining physics conditions that correspond to the normal human scale. The sizes and relative distances of scene objects were increased by a factor of ten. In addition, the properties of lights and reflection capture objects were adjusted so that the overall visual appearance of both conditions were kept as similar as possible.

3.1.4 Interaction Task

The interaction task consisted of the manipulation of virtual soda can pull tabs approximately 3 cm in length and 1.9 cm in width (as presented in **Figure 1**). The tabs were chosen for the experiment both for their small, consistent mass as well as for being a reasonably authentic object that could be seen in the simulated VE. We considered a lightweight object to be most practical for simulating throwing in VR so that we would not have to simulate the decrease in hand acceleration due to increased inertia at the end of the arm or limitations due to arm strength (Cross, 2004). In both conditions, the subjects would try dropping and throwing five tabs. Picking up and throwing the tabs took place utilizing the default mechanism in UE, similar to contemporary VR applications in general. The subjects simulated grabbing objects by squeezing the trigger of the motion controller and dropping them by releasing the trigger. Virtual throwing took place by swinging the motion controller and then releasing the trigger, and the object thrown retained its velocity at the moment of release, simulating throwing.

In the *true physics* condition, the tabs would drop down fast, similarly as to if they were dropped from the height of 15–20 cm (simulated falling speed approximately 0.175 s at 20 cm in UE). In addition, the throwing distances would appear short because of the limited velocity that can be actuated due to real hand movements scaled down by an order of magnitude. The *movie*

physics condition, on the other hand, simulated the tabs as falling down more slowly, similarly to an object dropped from human height (simulated falling speed approximately 0.6375 s at 2 m in UE). In addition, the throwing distances were much larger in the *movie physics* condition due to the larger velocity that the subjects were able to actuate on the tabs by virtual throwing.

Due to the simulated size, the tabs were also different between conditions in terms of their bounciness (there were no changes in physics simulation properties, such as restitution). In the *movie physics* condition, the tabs bounced visibly off surfaces, or jittered slightly after being dropped. However, in the *true physics* condition, there was little to no visible bounciness.

The tabs were placed on top of a large book so that the subjects would not have to pick them up from the floor. The book also provided an additional size cue. We gave the book a neutral, non-distracting appearance and a general title so that it was recognizable as a book, but would not otherwise draw too much attention. A Coca-Cola can was placed as a familiar sized cue on the left side of the book. **Figure 1** shows the book and the tabs as seen in the beginning of the simulation. **Figure 2 A** and **B** show the scene as seen at the beginning of the simulation when looking forward (A) and left (B).

The virtual mass of the tabs was set at 1 g in both conditions. Default physics settings in UE were utilized, with the exception of turning on the physics sub-stepping for additional physics accuracy by enabling physics engine updates between frames. Drag by air resistance was set to zero in both conditions since we considered the default damping approach in UE to be imprecise for our object type, and not crucial in our simulation. The simulation itself ran at stable 80 FPS which is the maximum frame rate of Oculus Rift S.

3.1.5 Participants

The experiment was carried out as a within-subjects experiment, in which 44 subjects (23 females and 21 males) performed both conditions during one session. Two participants were excluded due to issues with the functionality of the VR equipment or due to vision impairments. The order of the conditions was



FIGURE 2 | A screenshot of the VE from Study 1 from the subject's perspective when looking forward and upward with book and tabs below the line of sight **(A)** and when looking left **(B)** (Pouke et al., 2020).

counterbalanced so that there was an equal number of male and female participants starting with each condition. The subjects' ages ranged from 19 to 66, mean and median ages being 30 and 26, respectively. The standard deviation for the ages was 10.4. The study was conducted either in English (12 females and 7 males) or in Finnish (11 females and 14 males), depending on the preference of the subject. Each participant was rewarded with a gift voucher of two euros.

3.1.6 Experimental Procedure

The experiment was set in a laboratory in which the subjects used the Oculus Rift S system with provided Oculus Touch controllers for the experiment. The Rift S has a variable IPD software setting, so the IPD was set to 62.5 for females and 64.5 for males, the closest approximation available based on the averages reported for adults by Dodgson (2004). In the beginning of a session, the subject read through a written *Information for Subjects* document and signed an informed consent sheet. The subject was then instructed on using the VR hardware, specifically how to use the Rift S Touch motion controllers for picking up and throwing objects. Next, the subject was instructed to stand on a particular starting spot in the laboratory marked with a masking tape. When the subject was wearing the HMD and the motion controllers comfortably, an instruction script was read in English or Finnish. The script stated that the subjects were at the university central hallway, shrunk down 10-fold to a size of a doll, and were to drop and throw the tabs placed on top of a book in front of them.

Active noise-cancelling headphones were placed on the subject to block out any potential external noise from other rooms in the building, and then the experiment began. After performing both conditions, the headphones and the VR hardware were removed and the subject was asked to respond to a post-experiment questionnaire as well as a background questionnaire on a different laptop. The subject was asked for any additional comments or questions, and if he/she could be contacted for future studies, and then given her/his gift voucher. The average duration of the session was 20 min per subject.

3.1.7 Questionnaires

We collected plausibility related data using two forced choice questions (*realism* and *expectations*), two open-ended questions

(O1 and O2), and a 7-point Likert scale questionnaire regarding the behavior of the tabs (L1–L5). In addition, the subjects filled out the extended version of the Slater-Ussoh-Steed (SUS) Presence questionnaire (Slater et al., 1994; Ussoh et al., 2000), as well as a background information questionnaire. The SUS questionnaire was used to estimate presence in respect to the entire experiment, not per condition. The questions on *realism* and *expectations* were as follows:

1. Thinking back how the pull tabs were behaving in the experiment, which felt more realistic (like what would happen in the real world if you had been shrunk down), the first or the second time?
2. Thinking back how the pull tabs were behaving in the experiment, which matched your expectations (similar to what would happen in the real world if you had been shrunk down), the first or the second time?

We used two separate questions *realism* and *expectations* because we assumed some subjects would find the behavior of objects surprising in the *true physics* condition, even if they would consider it objectively realistic. The questions *realism* and *expectations* were coupled with open-ended questions (O1 and O2), that were simply stated as “Why?”. The purpose of the open-ended questions was to evaluate to what extent the subjects' responses were related to the physics or other reasons.

The forced-choice and open-ended questions were followed by a 7-point Likert scale questionnaire asking subjects to judge how they perceived various aspects related to the behavior of the tabs. Each question was stated twice in the questionnaire, referring to the first time and the second time subject interacted with the tabs (either using the *true physics* and then the *movie physics* or vice versa). The first three questions (L1–L3) were bipolar, whereas the last two (L4, L5) were unipolar. The Likert questions L1–L5 and their associated scales were as follows:

- L1 The falling speed of pull tabs (too slow, too fast).
- L2 The speed of pull tabs when thrown (too slow, too fast).
- L3 The distance of pull tabs when thrown (too close, too far).
- L4 The way the pull tabs were bouncing when thrown (incorrect, correct).
- L5 The impact of gravity on the pull tabs (incorrect, correct).

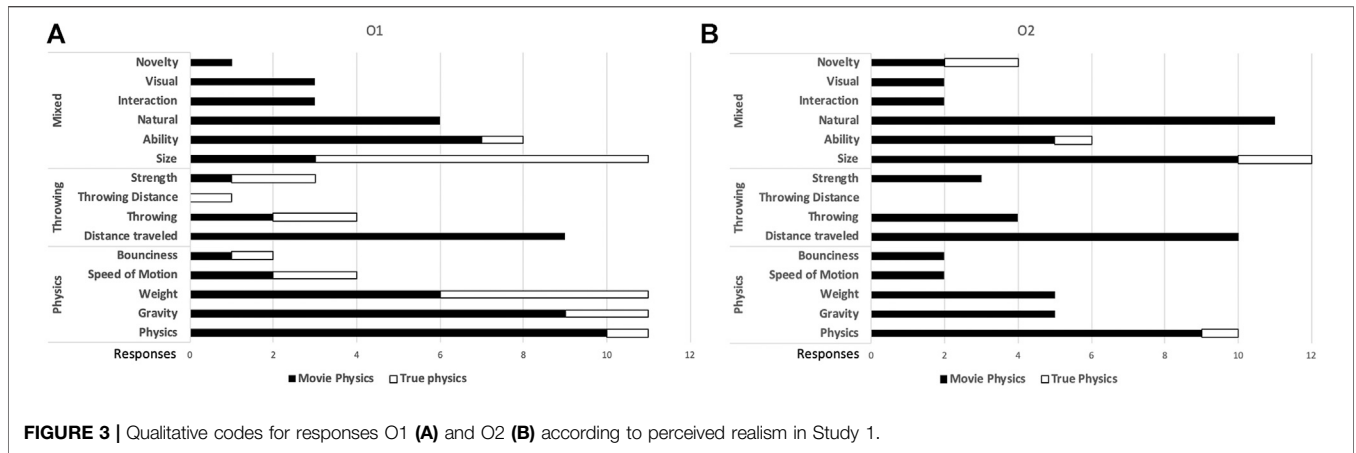


FIGURE 3 | Qualitative codes for responses O1 (A) and O2 (B) according to perceived realism in Study 1.

All questions were presented in either English or Finnish, depending on which was chosen as the preferred language by the subject when signing up for the experiment.

3.2 Results

3.2.1 Hypotheses

According to the responses to *realism* and *expectations*, the majority of the subjects considered the *movie physics* condition as the more realistic one. Out of 44 subjects, 32 participants (73%) responded to the first question that they considered the *movie physics* condition more realistic, which confirms H1. For the second question, 40 out of 44 (91%) subjects responded that the *movie physics* matched their expectations better, which confirms H2. Furthermore, we analyzed the frequencies of responses to questions 1 and 2 with a binomial test and found their corresponding two-tailed *p* values as $p = 0.004$ and $p = 1.7051 \cdot 10^{-8}$, respectively. From this we can conclude that it is unlikely that the responses to questions 1 and 2 were due to chance. In addition, this indicates that subjects were able to distinguish between the two physics conditions and more consistently selected the *movie physics* condition, which was the inaccurate physics condition.

Out of twelve respondents who considered *true physics* more realistic, nine responded that the *movie physics* matched their expectations more. Only one subject considered the *movie physics* more realistic while simultaneously stating that the *true physics* better matched her/his expectations.

3.2.2 Understanding Contributing Factors

We gathered supplementary data to further understand the results. These data include responses to open-ended questions O1 and O2, Likert-scale questions L1–L5, as well as subject background and self-reported level of presence.

The purpose of the open-ended questions was to evaluate to what extent the subjects' responses to *realism* and *expectations* were related to the perceived realism of the physics. The responses consisted of statements typed by the subjects. Thematic analysis with an inductive approach (e.g., Patton, 2005) was carried out independently by two researchers and used to identify codes in the response data. A summary of the codes can be viewed in

Figures 3A,B. ¹Examples of responses in O1 can be seen in **Table 1**, whereas examples of responses in O2 can be seen in **Table 2**.

In short, the responses to questions O1 and O2 indicate that the majority of users (38 out of 44) made their choices primarily according to reasons related to the behavior of the physically simulated tabs. Other primary reasons were related to general interaction and becoming accustomed to controllers. Few references were made to visual details (appearance of tabs and colors) as secondary reasons or general remarks.

3.2.3 Likert Data

Inspecting the Likert responses for questions L1–L5, we found that the *movie physics* condition was closer to perceived realism (median responses closer to 4 in questions 1 and 3 and closer to 7 in questions 4 and 5) in all questions except L2, in which the median response was the same for both conditions. We analyzed the responses to questions L1–L5 with the Wilcoxon Signed Rank test and found that the responses were significantly different ($p < 0.05$) for all questions except L2 ($p = 0.845$). This gives additional confirmation that the subjects perceived the *movie physics* condition more realistic due to differences in the behavior of the physically simulated tabs. A summary of responses including, median, mode and standard deviation for questions L1–L5 can be seen in **Table 3**. In addition, box plots visualizing the medians, interquartile ranges as well as minimum and maximum responses can be seen in **Figures 4A,B**.

3.2.4 Effect of Background and Slater-Usoh-Steed Scores

Furthermore, we used a binary logistic regression to analyze the effects of subject background and presence on their responses to *realism*. We used *educational background*, *gender*, *age*, *VR*

¹*Novelty* refers to cases where an aspect related to the sense of novelty in the VR experience affected the response. *Ability* refers to cases where the subject's perception of his/her own ability to use the VR controllers affected the response. *Size* refers to cases where the perception of subject's virtual size affected the response.

TABLE 1 | Examples of O1 responses (justification to *realism*) in Study 1.

Codes	Preference	Response
Gravity, natural	Movie physics	"Gravity felt more natural"
Gravity, natural	Movie physics	"At the second time, the objects fell to the ground faster, which felt unreal"
Gravity	True physics	"I think when the height of the object is not that high, it should reach the ground faster."
Physics, visual	Movie physics	"movement in space felt more realistic, but the objects lacked 3D, ring pulls are not paper thin"
Ability, distance traveled, physics	Movie physics	"because I was more comfortable with the controllers after using them for some time, and i knew i could do more things now like throwing more far away after some time, and also they were moving more smoothly"
Bounciness, throwing, distance traveled, ability	Movie physics	"I am not sure but I think the second time they still moved a bit after I dropped them to the floor, before being completely still. I think I also managed to throw one of the pull tabs the second time, which felt more realistic than them dropping very quickly just right in front of me after I tried to throw them (but this could also just have been my inability to throw the first time)."
Weight	Movie physics	"Second time they felt too heavy"
Weight, strength, size	True physics	"Pull tabs are not heavy and when I'm small, I probably would not have the strength to throw them afar"
Ability	Movie physics	"I was able to act more normal in the second round. I had worked out the mechanics of the VR better and spent less time attempting to make the task work"

TABLE 2 | Examples of O2 responses (justification to *expectations*) in Study 1.

Codes	Preference	Response
Distance traveled	Movie physics	"As I was taking a swing with my arms I was expecting them to land far away from me which they did only during the first time."
Speed of motion	Movie physics	"In the second time the tabs were falling down surprisingly fast"
Size, weight	Movie physics (different from O1)	"I was not thinking I was shrunk. So it felt estrange to have such heavy pull tabs"
Physics, size	Movie physics (different from O1)	"I didn't think at first (until I saw the previous question) shrinking down would also affect the time it takes for the objects to reach the ground. The physics first time behaved just like in normal life."
Size	Movie physics (different from O1)	"Even though I knew I was shrunk down, I still could not think that way when doing the experiment"
Natural, physics	Movie physics	"The behavior seemed more natural, although probably the laws of the physics tell otherwise"
Ability, throwing	Movie physics (different from O1)	"I thought throwing the pull tabs would be relatively easy, like in the second time"
Weight	Movie physics	"Intuitively I figured things would be light"
Size, novelty	True physics	"I felt that I was really small in that world for the first time."
Natural	Movie physics	"First time. Felt somehow more natural. They didn't have much difference, though"

TABLE 3 | Summary of Likert data in Study 1. In L1–L3, responses closer to 4 are perceived as closer to realism, whereas in L4 and L5, responses closer to 7 are perceived as closer to realism. Responses perceived closer to realism are emphasized in bold (Pouke et al., 2020).

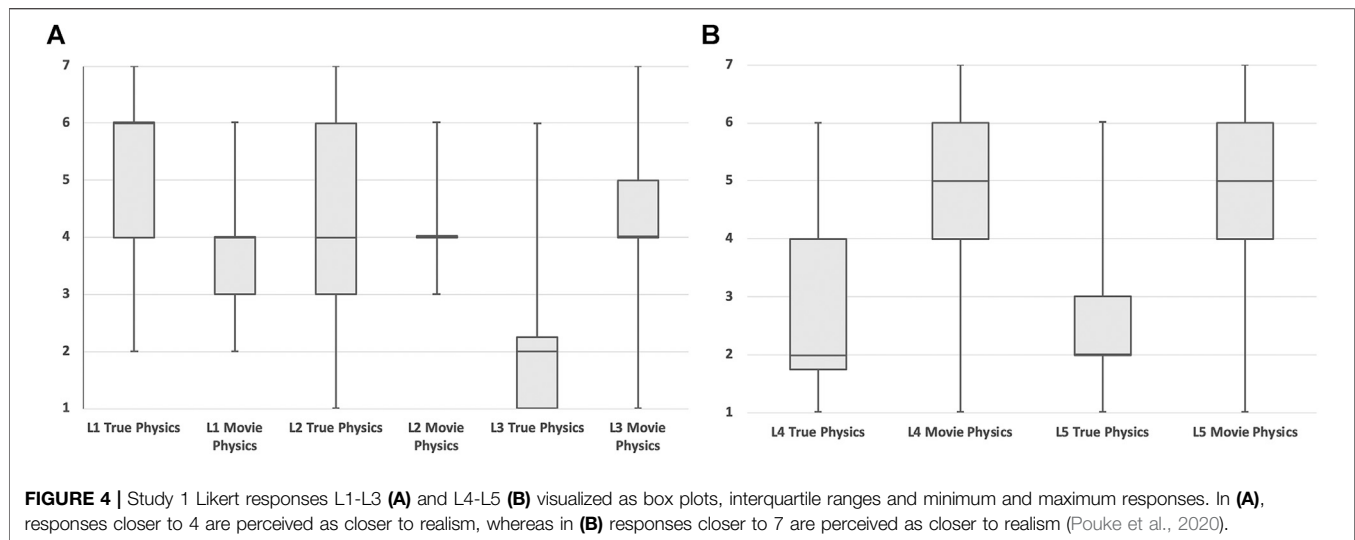
Question	Condition	Median	Mode	STD
L1: Falling speed	true physics	6	6	1.4
	movie physics	4	4	0.8
L2: Speed when thrown	true physics	4	6	1.8
	movie physics	4	4	0.9
L3: Distance when thrown	true physics	2	2	1.3
	movie physics	4	4	1.1
L4: Bounciness	true physics	2	2	1.6
	movie physics	5	6	1.5
L5: Gravity	true physics	2	2	1.4
	movie physics	5	6	1.4

experience, gaming experience, SUS average and SUS score as independent variables and the response to *realism* as the dependent variable.

For analysis purposes, we transformed the Background Questionnaire responses to *educational background* into a binary variable consisting of roughly equal sized groups of *natural sciences and engineering* (25 subjects) and *social*

sciences (19 subjects). In addition, the open responses to *VR experience* and *gaming experience* was transformed into respective ordinal variables ranging from 0 (no experience) to 4 (plenty of experience). When interpreting the *gaming experience* responses, additional emphasis was given to recent experience as well as experience regarding PC and console based 3D gaming (such as first person shooters and simulators) due to the tendency of such games to contain game physics simulations similar to those used in this experiment. The responses to SUS scores were transformed into two ordinal variables consisting of the average of responses as well as the computed SUS score (the number of responses with 6 or 7).

The logistic regression model was unable to predict the response using the independent variables. The model explained 17% of the variance (Nagelkerke's R^2) in perceived realism. Although the overall classification rate was 72.7%, only 16.7% (two responses) of the *true physics* responses were correctly classified. None of the independent variables had a significant effect on the prediction of the response ($p = 0.184-0.858$). According to this analysis, the perception of realism was not significantly affected by the background, education or gaming experience of our subjects. The level of presence according to self-reported SUS score did not have any effect either.



3.2.5 Perception of Mass and Strength

Although we never queried subjects directly regarding the physical properties of the tabs themselves, several subjects commented on the weight of the tabs or their own strength when interacting with the tabs. Five of the subjects who responded in English commented on the feeling of the perceived heaviness of the tabs (see **Table 1**). It is interesting to consider these spontaneous responses regarding differences in the weight of the tabs given that there was no change in the controllers that the subjects used for each condition. This could be an indication of a pseudohaptic effect (Lécuyer, 2009); for example, manipulating the control-to-display ratio of the visual feedback when lifting an object can give the user an illusion of increased weight (Samad et al., 2019). However, it is possible that the subjects were simply referring to the visible trajectories and falling speed of objects (as in the tabs *seemed* heavier instead of the tabs *feeling heavier*). Several of the responses in Finnish specifically contemplated the assumed weight of the tabs in regards to how more much power they would have needed to use to throw the tabs given their reduction in size. To investigate these findings further, we added additional pseudohaptic related questions in Study 2.

4 STUDY 2

4.1 Study Design

In Study 2, we wanted to investigate the perception of rigid body dynamics while the user was enlarged by a factor of ten. We followed a methodology similar to Study 1 so that we could easily compare subjects' perceptions in small and large scales. We introduced minor methodological changes described below.

4.1.1 Hypotheses

Our hypotheses for Study 2 were similar to those of Study 1.

H3: For a scaled-up user, *movie physics* is more likely to feel realistic than *true physics*.

H4: For a scaled-up user, *movie physics* is more likely to match a user's expectations than *true physics*.

4.1.2 Virtual Environment

Study 2 portrayed the subject as a giant, 10 times larger than a regular human. Similarly to Study 1, the VE was also based on a real-world environment we expected to be familiar for most of our subjects. More specifically, the VE depicted a marketplace and its surroundings located in the center of the City of Oulu, Finland. The environment used 3D assets from the "Virtual Oulu" model described in Alatalo et al. (2016). The assets were imported into a UE 4.24 scene. Some of the original materials were remade to follow a contemporary physically based rendering (PBR) workflow for improved esthetics. To enrich the model with additional size cues, the marketplace area of the model was augmented with additional detail such as street furniture, trees, and foliage that were placed using Google Maps photographs and satellite photos as reference. GIS data from the City of Oulu was used to generate non-textured faraway buildings seen in the background of the scene. Also, generic textured building blocks were used in some areas to generate buildings not present in the original Virtual Oulu model but close enough to the viewer so that untextured models were not feasible. Although our aim was to make the scene appear realistic for the subjects, we took minor liberties in the placement of certain scene objects to make the scene more appropriate for the experiment. Namely, the immediate marketplace surroundings were left relatively empty to prevent the subjects from hitting random objects and making unwanted "plausibility noise"; we wanted our subjects to keep their focus on the trajectories of the logs, and not get distracted by the effects of various collisions. In addition, the position of the trees next to the shoreline was adjusted so that the logs had a free passage to water (see **Figure 5A**).

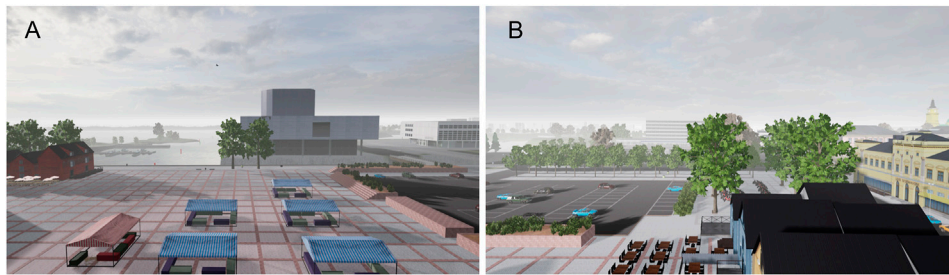


FIGURE 5 | The VE used in Study 2 as seen by subjects in the beginning of the experiment when looking forward (A) and right (B)

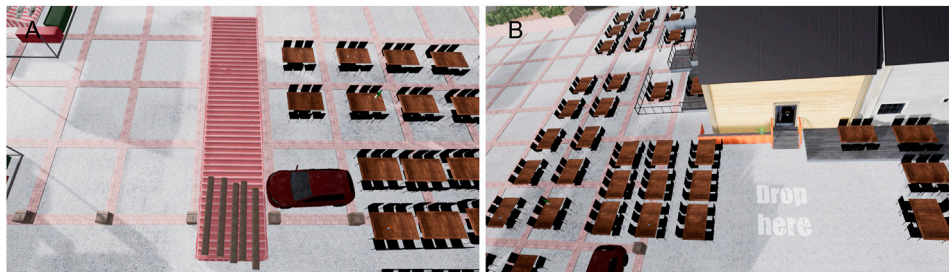


FIGURE 6 | The container with the logs as seen in the beginning of the experiment (A) and the area in which the first three logs were to be dropped (B)

In addition to Virtual Oulu assets, GIS data and self-modeled assets, several commercial packages from the Unreal Marketplace were used in the VE. Animated seagulls and pigeons from the *Birds* package were scaled to correct size (approximated wingspans 70 and 140 cm, respectively) and deployed in the scene to provide animated size cues. The commercial packages “*Nordic Harbor*”, “*Country Side*”, “*Vehicle Variety Pack*”, “*Modern City Downtown*”, “*Sky Pack*” as well as “*Trucks and Trailers*” were also utilized for foliage, vehicles, street furniture, and other minor details, such as traffic signs. Water shader and buoyancy for logs was generated using the *Waterline Pro* package. Screenshots of the scene can be seen in **Figures 5A,B**.

Similarly to Study 1, the scale-changing effect was achieved by scaling the world-to-meters parameter of UE and player character properties, this time making the user to appear 10 times larger instead of smaller. Also as in Study 1, we defined the rigid body dynamics as simulated by the game engine to act as the *true physics* condition. For the *movie physics* condition, we upscaled the *default gravity Z* and *bounce threshold* (as instructed by UE when scaling gravity) properties by 10 to generate conditions similar to human scale. This approach was taken to avoid generating two different-sized versions of the level so that we could eliminate visual difference (in Study 1, we observed minor lighting differences between the two conditions when using different-sized versions of the levels, see **Section 5.2**). These approaches resulted in object free fall times of 1.97 and 0.68 s, when object is dropped from the height of 18 m in *true physics* and *movie physics* respectively.

4.1.3 Interaction Task

The interaction task in Study 2 resembled the task in Study 1, consisting of dropping and throwing objects. We determined wooden logs as suitable objects for interaction, since they are somewhat familiar sized objects for most locals. The logs were approximately 2.9 m in length and 26.7 cm in diameter, matching the dimensions and mass of commercial pine logs. We placed the logs on top of a container. The container with the logs can be seen in **Figure 6A**.

In Study 1, the subjects were allowed to drop and throw the pull tabs in any way they wished. However, in Study 2 we instructed the subjects to drop exactly three logs to their right and throw two logs into the sea visible in front of the subjects (see **Figure 5A**). We also placed a “Drop Here” text on the ground to depict where exactly the logs should be dropped (see **Figure 6B**). There was a particle splashing effect when the logs hit the water surface as well a buoyancy effect. However, these effects were very subtle due to the distance to the water surface.

The specific instructions for interaction were included because in Study 1 we received feedback indicating that more specific instructions would have helped in observing the motions of the pull tabs. In addition, since the subjects were interacting in a large-scale urban environment, there were countless opportunities for “plausibility noise” which we wanted to avoid (such as subjects expecting logs to realistically break windows, dent cars, knock over tables, and so on). By giving specific instructions, we aimed to ensure that the subjects’ responses were based on the motions of the logs only.

4.1.4 Participants

Similarly to Study 1, the experiment was carried out as a gender-counterbalanced within-subject experiment, this time with 40 participants (20 males and 20 females). Three participants were excluded due to failure to follow instructions or not giving their consent for data use. We did not allow people who had already participated in Study 1 to participate again to keep subjects naive for the purpose of the experiment. The subjects' ages ranged from 20 to 57, with mean and median ages being 26 and 25, respectively. The standard deviation of the ages was 6.0. Each participant was remunerated with a movie ticket worth 10 euros.

4.1.5 Experimental Procedure

Apart from COVID-19 related safety guidelines discussed below, the procedure was largely similar to Study 1. In Study 2, however, the subjects did not wear noise-cancelling headphones as there was no sound in the VE or from students in the laboratory hallways, and thus they were unnecessary.

In Study 1, a researcher checked the HMD before each participant to ensure that the Oculus main menu or other anomalies were not present when starting the experimental apparatus. In Study 2, however, we asked the subject to report what he/she saw in the beginning of the experiment since we could not be in close proximity to the subjects to check ourselves. In addition to checking anomalies, this also allowed us to check whether the subject recognized the VE as the Oulu marketplace. After this, an instructions script was read for the subject. The script confirmed that the subjects were at Oulu marketplace, enlarged 10-fold to a size of a giant, and they were to throw and drop the logs placed on a container in front of them. At the end of the experiment, the subject was given her/his movie ticket.

The experiment in Study 2 was conducted during the COVID-19 pandemic, hence additional safety precautions were taken. At the time of the experiment, the regional state of the epidemic was at so-called "baseline level"².

The research space allowed for a maximum of two researchers who kept within safety distance to the participant. The researchers were also separated from the subject with a see-through barrier. The researchers wore safety masks, which were also offered for subjects. The participants were instructed to operate the VR equipment by themselves during the experiment and a researcher intervened only if necessary.

Virtual reality equipment was sanitized between each participant using a "Cleanbox"³ device. In addition, all equipment and surfaces were wiped with alcohol disinfectants. Researchers wore rubber gloves during the cleaning process and the experiments. The default face padding of Oculus Rift S was covered with a silicone hygiene cover for easier cleaning. In addition, the subjects were offered optional disposable paper face hygiene covers. The research space was air-conditioned and ventilated between subjects. Participants were also asked

²Defined by Finnish Institute of Healthcare as follows: "The baseline corresponds to the situation in Finland in the middle of the summer, 2020. The incidence of infections is low, and the proportion of endemic infections is small."

³<https://www.cleanboxtech.com/>

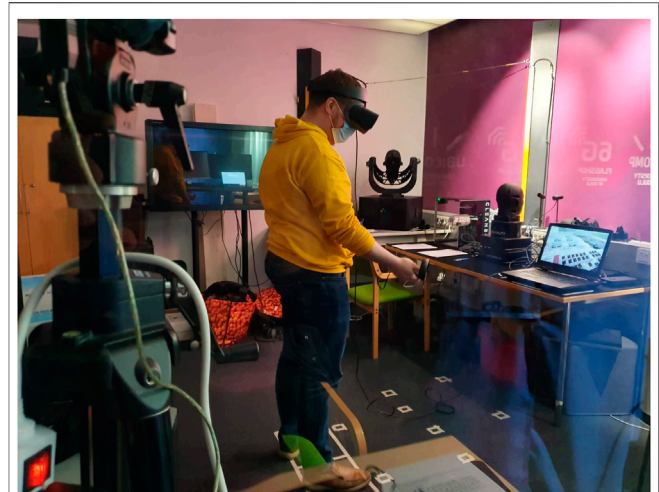


FIGURE 7 | A photo of the research space during Study 2 taken behind a see-through barrier. The same research space was used in both studies.

to use hand disinfectant available in the research space. Participants were asked to join the experiment only when feeling completely healthy. The research space can be seen in **Figure 7**.

4.1.6 Questionnaires

The questionnaires in Study 2 were kept mostly similar to Study 1, consisting of two forced choice questions (*realism* and *expectations*), two open-ended questions (O1 and O2), a 7-point Likert questionnaire concerning log physics (L1–L5), and the extended SUS questionnaire Slater et al. (1994); Usoh et al. (2000). In addition, we added extra 7-point Likert-scale questions L6–L8 concerning the experience of being large and pseudohaptics.

Realism and *expectations* were identical to Study 1, except replacing "pull tabs" with "logs" and "scaled down" with "a giant". The questions were as follows:

1. Thinking back how the logs were behaving in the experiment, which felt more realistic (like what would happen in the real world if you had been a giant), the first or the second time?
2. Thinking back how the logs were behaving in the experiment, which matched your expectations (similar to what would happen in the real world if you had been a giant), the first or the second time?

Similar to Study 1, both *realism* and *expectations* were followed by open-ended questions O1 and O2 stating "Why?".

Questions L1, L3, L4, and L5 were kept similar to Study 1 so that only "pull tabs" were changed into "logs." Since the wording of L2 in Study 1 was found to be problematic, we paraphrased it from "the speed of pull tabs when thrown (slow - fast)" into "time of flight (slow - fast)". The questions L1–L5 were as follows:

- L1 The falling speed of logs (too slow, too fast).

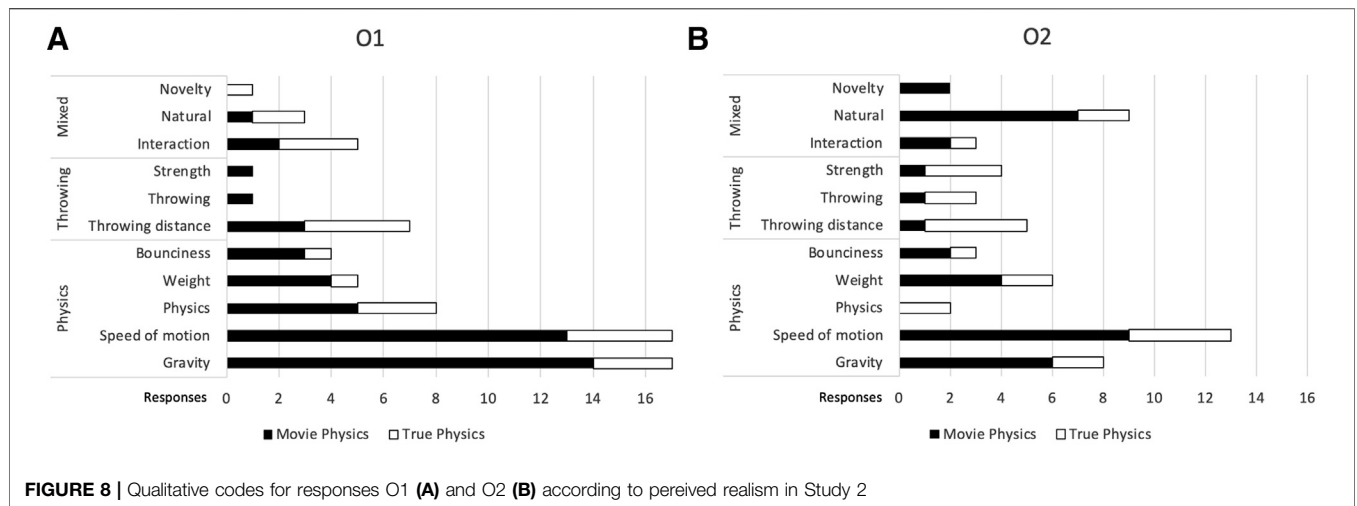


FIGURE 8 | Qualitative codes for responses O1 (A) and O2 (B) according to perceived realism in Study 2

- L2 The time of flight of logs when thrown (too slow, too fast).
- L3 The distance of logs when thrown (too close, too far).
- L4 The way the logs were bouncing when thrown (incorrect, correct).
- L5 The impact of gravity on the logs (incorrect, correct).

The new questions L6–L8 assessed the feeling of size and the sensation of weight of the logs in both conditions. The questions were phrased as follows:

- L6 During the experiences, did you feel more like a giant in a normal-sized city, or more like a normal-sized person in a miniature city (normal-sized person, giant).
- L7 When picking up or holding the logs, did you feel a sensation of actual weight (not at all, very much so).
- L8 The logs felt ... (light, heavy).

Similarly to Study 1, all questions were presented either in English or Finnish, depending on the preference of the subject.

4.2 Results

4.2.1 Hypotheses

Again, majority of the subjects considered the *movie physics* condition as the realistic one, but the expectations of the subjects were more mixed, however. For *realism*, 28 out of 40 (70%) subjects chose *movie physics*. As for response to the *expectations*, 25 out of 40 subjects (63%) considered that *movie physics* matched their expectations better. Following the procedure in Study 1, we analyzed the frequencies of responses with a binomial test, and found their corresponding two-tailed *p*-values as $p = 0.017$ and $p = 0.154$. This indicates that the responses to *realism* were significantly biased toward *movie physics*, whereas responses to *expectations* were closer to a random distribution. These results confirm H1, but not H2. Most of the subjects clearly considered *movie physics* as the more realistic condition, but their expectations were more evenly split between *true physics* and *movie physics*. Almost every

subject mentioned recognizing the scene as the Oulu downtown marketplace.

4.2.2 Open-Ended Questions O1 and O2

Thematic analysis using the inductive approach (Patton, 2005) was used to analyze the open-ended questions. The responses were first coded by two independent researchers, after which the final codes were agreed upon. One subject did not respond to the open-ended questions. A summary of codes and their frequencies per each question can be seen in **Figure 8**.

The responses indicate that majority of the subjects considered the motions of the logs as their primary reason of preference; the logs were either moving at a speed they did not feel was realistic, or were under the effect of abnormal gravity. This is especially true for the subjects that perceived *movie physics* as more realistic. For the subjects choosing *true physics*, the ability to throw logs especially far came up relatively more often than for *movie physics* respondents. This could mean that these subjects considered a giant being capable of throwing the logs farther due to increased strength. However, similarly to Study 1, we did not simulate muscle strength per se; the ability to throw the logs far was due to the increased velocity the subjects were able to impart due to being scaled 10-fold. There was only one response to O1, in which *strength* was specifically mentioned, whereas for O2, *strength* came up in four responses. Examples of responses for O1 can be seen in **Table 4**. For O2, examples can be seen in **Table 5**. Distributions of qualitative codes in Study 2 can be seen in **Figure 8**.

4.2.3 Likert Data

Similarly to Study 1, we analyzed questions L1–L5 (*falling speed, time of flight, distance when thrown, bounciness* and *gravity*) to get additional insight into the subjects' perceptions of the motions of the logs. In all questions, the respondents favored *movie physics*, with the median and mode closer to 4 in L1–L3 and closer to 7 in L4 and L5.

TABLE 4 | Examples of O1 responses (justification to *realism*) in Study 2

Codes	Preference	Response
Gravity	Movie physics	"me being big should not affect the gravity of other objects"
Speed of motion	True physics	"Not really sure, but when I picture a giant it feels like that way. Like things going on slow motion."
Speed of motion, gravity	Movie physics	"It happened with normal speed/gravity"
Physics, gravity	Movie physics	"Acceleration felt somewhat realistic, the latter felt like surface of the moon"
Throwing distance	True physics	"when using a strong force, the logs was thrown far away, matching my expectation"
Throwing distance	True physics	"If I were a giant, the logs would fly a little farther, which was highlighted in the second time"
Novelty, physics, bounciness, interaction	True physics	"Everything felt new, not only that you were in VR in the first place, but also the point of view, which was of course higher than normal. It also felt like, in terms of physics, the logs were behaving more realistically in the first time, because I was handling them more carefully. On the second time I just dropped the logs from high up, and they were bouncing any which way"
Gravity, weight, naturalness	Movie physics	"The gravity and motion of the logs felt more natural. In the other one, they were floating like feathers in space and were clearly lighter than real"
Gravity, speed of motion, strength	Movie physics	"According to my own assumptions, objects would feel like they were moving more slowly in relation to myself if I were a giant, but the first time felt more like I was underwater. In my opinion, the second time was more real, even if it was a little fast-ish. I did feel as if I was stronger in the second time, though."

TABLE 5 | Examples of O2 responses (justification to *expectations*) in Study 2

Codes	Preference	Response
Naturalness, speed of motion	True physics (different from O1)	"This bias might be partly because of movies, but also in many real life videos, big objects fall "more slowly" when seen from afar. In the first version the logs were much more slower, which matched my expectations more."
Naturalness, speed of motion	Movie physics (different from O1)	"Somehow faster motions felt more natural"
Naturalness	Movie physics	"Logs felt more credible in the second experiment"
Physics	True physics	"The second time matched my expectations more since the motions of the logs were more realistic"
Speed of motion	Movie physics	"Because the logs acted as they should. A log in the real would not fall slowly."
Interaction, bounciness, throwing	Movie physics	"I find it easy to grab and on throwing it was more realistic. When I drop the log it bounced back as well, making it more realistic. In second, I was also able to see the log clearly when it was in air during the throw."
Speed of motion	Movie physics	"Still the first one. I can not realistically think the world working in slow motion."
Speed of motion	True physics	"because of the speed when I drop the logs"
Strength, throwing distance	True physics	"As a giant I would expect to be stronger, therefore being able to throw the logs further. "

TABLE 6 | Summary of Likert data in Study 2. In L1-L3, responses closer to 4 are perceived as closer to realism, whereas in L4 and L5, responses closer to 7 are perceived as closer to realism. Responses perceived closer to realism are emphasized in bold.

Question	Condition	Median	Mode	STD
L1: Falling speed	true physics	2	2	1.3
	movie physics	5	5	1.0
L2: Time of flight	true physics	2	2	1.7
	movie physics	4	4	1.1
L3: Distance when thrown	true physics	6	6	1.5
	movie physics	3	4	1.0
L4: Bounciness	true physics	3	2	1.7
	movie physics	5	6	1.5
L5: Gravity	true physics	2	2	1.7
	movie physics	5	6	1.4

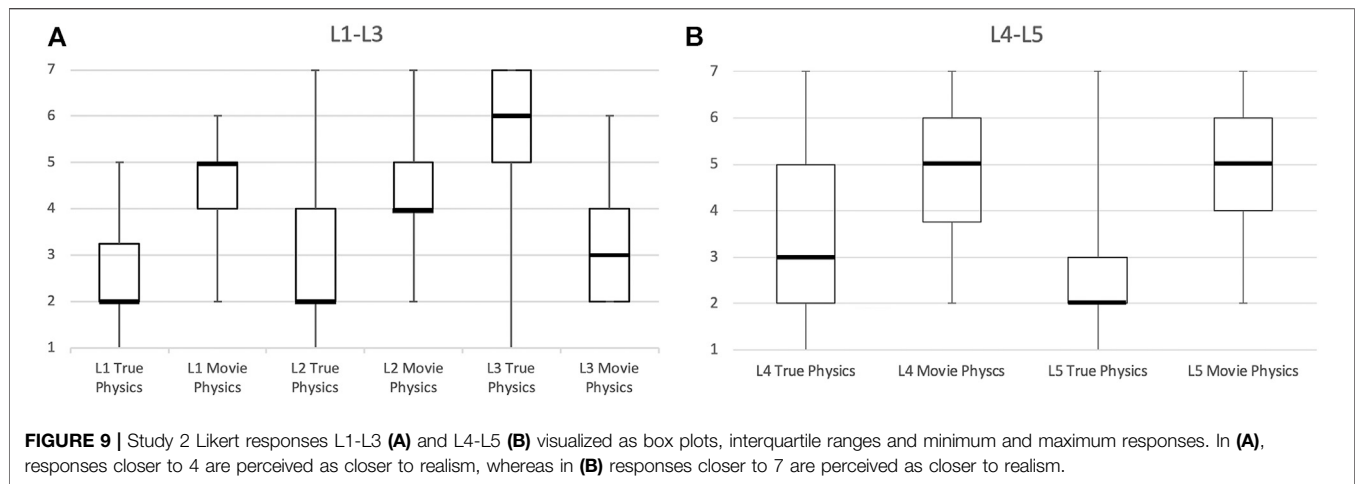
A Wilcoxon Signed Rank test showed that the responses were significantly different ($p < 0.05$) between the conditions for all of the questions except L4 ($p = 0.4$). A summary of the responses can be seen in **Table 6**. Box plots visualizing medians, interquartile ranges, as well as the minimum and maximum responses can be seen in **Figure 9**.

4.2.4 Presence Data

Similarly to Study 1, we acquired self-reported presence data according to the SUS questionnaire. Thirty out of 40 subjects (75%) had an SUS score higher than 0, indicating at least some level of presence. The median SUS score was 1. Again, we divided the subjects into groups of high presence (SUS score >2) and low presence (SUS score <3). The proportion of the high presence group was 16 out of 40 (40%), whereas the low presence group consisted of 24 subjects (60%). Seven subjects out of 16 (44%) from the high presence group chose *true physics*, whereas only five subjects out of 24 (21%) from low presence group did the same. However, according to Fisher’s exact test, this difference was not significant ($p > 0.05$), which means we can assume that belonging to either high or low presence group did not affect the response to *realism*.

4.2.5 Own Size and Pseudohaptics Sensations

We added three new questions L6–L8 to investigate the subjects’ perception of his/her own size, as well as pseudohaptic sensations. It appears that although the subjects generally considered *movie physics* as the more realistic condition, *true physics* was able to more successfully convey the sensation of being large. The median and mode responses to L6, *feeling like a giant*, were 6



and 6, respectively, for the *true physics* conditions. As for *movie physics*, these responses were 4 and 2, respectively. This is somewhat supported by the responses to open-ended questions O1 and O2 (for example, subjects considering true physics more natural, see Table 5). We found these differences to be statistically significant using the Wilcoxon Signed-Rank test ($p = 0.002$).

When inspecting the open-ended data from Study 1, we found a number of subjects mentioning the pull tabs feeling heavier in one condition or another. To investigate this further, we added new questions L7 and L8 to inquire about pseudohaptic sensations. However, these sensations were reported as very low in general. For L7, *sense of actual weight*, the median and mode responses were 1.5 and 1 for *true physics* and 2 and 1 for *movie physics*, respectively. As for L8, *logs felt light/heavy*, median and mode was 2 and 1 for *true physics* and 3 and 1 for *movie physics*. Out of 40 subjects, 4 and 6 subjects reported pseudohaptic sensations stronger than 4 out of 7 in the *true physics* and *movie physics* conditions, respectively. However, Study 2 might have been less appropriate to study the sensations of weight reported by subjects in Study 1, since in Study 2 the physics ranged from normal to perceptually slower, instead of vice versa. Again, using the Wilcoxon Signed-Rank test, we found statistical differences for L7 (*physical sensation of weight*) to be insignificant ($p > 0.05$). However, there was a statistically significant difference ($p = 0.026$) in responses for L8 (the logs were light/heavy). These results could be interpreted so that the subjects generally considered the logs simulated heavier in the *true physics* condition, but failed to notice any differences regarding pseudohaptic sensations, however.

4.2.6 Effect of Background, Presence and Perception of Own Size

Similarly to Study 1, we analyzed the effects of subject background and self-reported presence on their preference on physics. This time around, we also added responses to L6 as variables *feeling like a giant - true physics* and *feeling of own size -*

movie physics to estimate whether subjects' perception of their own size (in essence, the extent of feeling like a giant) affected responses in either condition. Using the same categories and the same coding mechanisms as in Study 1 (this time with 22 subjects categorized having a background in *Natural Sciences and Engineering* and 17 subjects with a *Social Science* background), we performed binary logistic regression analysis. The model explained 38% of the variance (Nagelkerke's R^2) with 79.5% overall accuracy. Although we found, similar to Study 1, that background, gaming or VR experience and self-reported presence did not affect responses, the variable *feeling like a giant - movie physics* had a significant effect ($p = 0.041$). This finding indicates that *true physics* respondents felt smaller specifically during the *movie physics* condition. However, since the distributions of both *true physics* and *movie physics* responses were quite large, but the number of *true physics* respondents was rather small, we would hesitate to put too much confidence in this implication until further evidence is found.

4.3 Comparing Studies 1 and 2

The percentage of subjects that chose *movie physics* for *realism* was 73% in Study 1 and 70% in Study 2. As for *expectations*, these percentages were 91% for Study 1 and 63% for Study 2. We compared the results for *realism* and *expectations* from Studies 1 and 2 with Fisher's exact tests. We found that the difference in number of responses to *realism* was not statistically significant ($p > 0.05$), whereas the number of responses to *expectations* was ($p = 0.003$). This suggests that a majority of similar proportions considered *movie physics* more realistic in both studies. The proportions were largely different for *expectations*. Although almost all subjects considered *movie physics* as better matching their expectations in Study 1, there was not a statistically significant difference in responses for Study 2, even if a slight majority did prefer *movie physics*.

The results for Likert questions L1, L3, and L5 were very similar in Studies 1 and 2, consistently favoring *movie physics*. Responses to L2 were very mixed in Study 1, which we attribute to bad wording of the question. In Study 2, the responses were more consistent and clearly favored *movie physics*. In Study 2, the

responses for L4 were mixed while in Study 1 *movie physics* was preferred.

In both studies, we examined the effect of various contributing factors in an effort to gather additional insights for interpreting the results. In Study 1, we used background data as well as self-reported presence as predictors to *realism*. In Study 2, we also added two new variables *feeling like a giant - true physics* and *feeling like a giant - movie physics*. In Study 1, however, we did not find any significant predictors. In Study 2, a new variable, *feeling like a giant - movie physics* came out as significant.

If we compare the presence scores to those of Study 1, we can see that subjects in Study 2 experienced somewhat less presence. In Study 1, some presence (SUS score >0) was experienced by 82% of the participants with median SUS score being 3. In addition, the proportion of high and low presence groups were almost equal in Study 1 (53% experiencing high sense of presence). In Study 2, 75% responded with SUS score >0 while the median SUS score was 1. The proportions of high and low presence groups were 40 and 60% respectively. However, despite these differences, the SUS scores for Study 1 (44 subjects) and Study 2 (40 subjects) were not statistically different (Mann-Whitney *U* test $p = 0.05$). Also, presence did not have a predictive capability on the preference of realism in either study.

5 DISCUSSION

Our results demonstrate that we have identified a strong paradox concerning PSI in VEs in which the user has been scaled either up or down. Our findings align with the definition of PSI: the plausibility illusion is more dependent on the expectations of the subjects than objective reality (Slater et al., 1994; Skarbez et al., 2017a). We believe, however, that our findings have implications for VR and telepresence applications in which realistic physics are unavoidable, yet users are expected to interact with the environment and respond realistically.

The proportion of the subjects that chose *movie physics* in *realism* was almost identical in Study 1 and Study 2. Close to a 3/4ths majority (73% in Study 1 and 70% in Study 2) chose *movie physics* as the realistic representation. As for *expectations*, the responses were quite different, however. In Study 1, 91% of subjects considered *movie physics* as matching their expectations more, whereas in Study 2 only 63% of the subjects considered the same. It appears that realistically approximated physical phenomenon at a small scale was surprising for almost all subjects. However, many subjects considered *true physics* to better match their expectations at a large scale, even if they actually regarded *movie physics* as the realistic one.

The purpose of open-ended questions regarding the reason why subjects rated one of the physics conditions being more realistic (O1) or matching their expectations better (O2), was first to confirm that the subjects gave their responses according to object motions and not other plausibility related factors, and second to give additional insights, for example regarding different responses to O1 and O2.

In Study 1, according to O1 and O2, almost all of the subjects considered their perception of realism to be related to the physics behavior of the tabs. In addition, a small number of subjects gave responses motivated by general interaction, including learning how to use the controllers correctly. A few secondary reasons or remarks were made referring to a scene object or other visual details. According to the responses to O2, most of the subjects preferring *true physics* as the realistic one stated that during the experiment it was difficult to understand why the physics functioned the way that it did - the behavior of the tabs was still surprising even if they considered it realistic.

As for Study 2, the reasons given by the subjects were also most often related to the behavior of the logs. Some exceptions included interaction (learning to use controllers or other interaction related issues) and novelty (for example, the experience being more overwhelming in the first part of the experiment). In Study 2, no visual aspects came up in the open-ended responses.

Whereas in Study 1, only one subject responded with *movie physics* in O1 and *true physics* in O2, this was the case for five subjects in Study 2. The most popular reason in these cases was the ability to throw the logs farther (3 responses). The other reason was that the slower motions somehow seemed more natural, even if unrealistic, as a giant (2 responses). Another difference to Study 1 was that for Study 2, subjects choosing *true physics* in O1 usually gave the same response also to O2; the behavior of the logs at large scale was not surprising to the same extent as the behavior of the tabs in small scale.

We used Likert scale rating questionnaires to gather additional insight into our findings. The questions focused on various dynamic properties of the objects so that we could more specifically pinpoint the effects of physics simulations on perceived realism. These responses indicated preferences toward *movie physics* as well, with significant differences regarding the perceived realism of the object behavior (with the exceptions of question L2, *speed when thrown*, in Study 1 and L4, *bounciness*, in Study 2). The Likert data further confirms that physically accurate representations of physics in abnormal scales are not inherently intuitive for VR users.

According to our results, accurate accelerations and falling speeds of objects were perceived as unrealistic. The distance that the subjects were able to throw the objects was seen mostly as too short in Study 1 and too long in Study 2. However, there were also responses in both studies that considered *movie physics* to be too extreme.

In Study 1, responses regarding the bounciness of the of the tabs indicated that subjects expected the tabs to behave as if they were enlarged 10-fold. In Study 2, however, the reactions to bounciness were much more mixed; even if median and mode responses preferred *movie physics*, the responses were too mixed to cross the threshold of significance at $p = 0.05$. We believe the main reason for the difference for these responses is the scale. In Study 1, the tabs were not practically bouncing at all in the *true physics* condition. In Study 2, however, the logs were bouncing in both conditions.

In Study 2, we inquired about the extent to which the subjects felt like a giant in a normal-sized city instead of a regular-sized

person in a miniature city. We found a significant difference between the conditions, indicating the subjects in the *true physics* condition felt larger. This may mean that even if the subjects did not generally believe the *true physics* condition to be realistic, it succeeded better in providing the illusion of being large.

We inspected the effects of various aspects of the subjects' background on their responses to O1. It could be that the subjects with knowledge of physics, for example, might prefer the *true physics* condition. However, we found no such effects in either of our subject groups. In addition, we did not find the self-reported level of presence (Slater et al., 1994), either as SUS scores or by dividing subjects into groups of high and low presence, to affect the response to O1 in either study. In Study 2, we found a significant effect for the variable L6 *feeling like a giant - movie physics*. This suggests that the extent to which the subjects experienced the illusion of being a giant in the *movie physics* condition had at least some effect on the subjects' perception of physics. However, the overall performance of the classifier was not very good, and the distributions of the subjects' responses were very large. For this reason, we believe further investigation is necessary before we can claim whether or not the extent of the small-scale or large-scale illusion affects the perception of physics.

In Study 2, we also studied pseudohaptic sensations experienced by the subjects and found that the overall extent of the sensations was very low. A handful of subjects reported strong pseudohaptic sensations. There was, as expected, a perceived difference regarding the overall weight of the logs between conditions.

We found that the level of presence experienced by subjects in Study 2 was somewhat lower. However, as of now, we do not have evidence to claim whether the illusions of being small or large affected self-reported presence or whether, for example, the properties of the VEs used in the studies would explain these differences.

5.1 Implications

Slater, (2009) discussed the role of conformity to expectations, prior beliefs and knowledge for causing and maintaining PSI. Skarbez et al. (2020) conceptualized the former as *coherence*, the reasonable behavior of the VE, which, according to Skarbez, is related to PSI similarly as immersion is related to PI.

Looking at the results against this framework, we can see that in Study 1, *movie physics* was clearly the reasonable behavior for subjects. Even if 27% of subjects considered *true physics* as real, only a handful of subjects considered it matching their expectations. Therefore according to the results of Study 1, realistic object behavior in small scale clearly violated coherence.

According to the results of Study 2, it is somewhat unclear which behavior is the coherent one, even if the results are somewhat pointing toward *movie physics*. Although a significant majority did consider *movie physics* as the realistic behavior, the expectations of subjects were matched almost even. Because of this mixed response to expectations, it is not straightforward to say, which model would yield good coherence in VEs.

Our findings resemble the cases discussed by Nilsson et al. (2017) where a decrease in a particular fidelity component

increased the overall perceived realism. However, we believe that our case differs from the cases discussed by Nilsson et al. (2017) in the sense that the culprit probably lies in the human perception of physics instead of the limitations of the physics engine.

If one was to design a multiscale VR application that would aim at maximizing coherence instead of realism, it would make sense simply to match the physics with the scale of the user, at least in small-scale applications. If the user is allowed to change scale, the physics behavior would follow similarly to Hollywood movies such as *Honey I Shrunk the Kids* where object motions constantly change in speed from scene to scene according to perspective changes. In large scales, however, this type of behavior might lead to bad coherence. In addition, in mCVEs this approach would break since the physics model would not be able to feasibly accommodate all users' perspectives simultaneously during multi-user interaction.

If realistic physics are intended, then users' expectations would have to be modified by some type of training so that realistic behavior does not come up as surprising. According to Skarbez et al. (2020), bad coherence in VEs, especially in relation to unexpectedly behaving environment, can lead to stress and discomfort. In addition, there might be cases where human interaction capabilities are reduced due to unexpected physics. Micro- and nano-level robotics operations are an example of this (Sitti, 2007). For this reason we consider interaction at abnormal scales and perceptual training as important future research directions; even if users would expect realistic physics, their performance might still be affected.

Through recent advances in consumer VR hardware as well as sub-microscopic (Plisson and Zotkina, 2015) and even atomic (Zheng et al., 2017) level imaging techniques, it is possible that we will witness an increasing exploitation of scaled-down VR applications in the future. They could potentially include commercial systems such as teleoperated maintenance robots or commercial virtual design solutions at a microscopic scale. However, at this stage, it is unclear whether it would be intuitive for humans to operate at small scales, especially if it involves operating in the real world or with realistically simulated physics. As can be seen from our results, the perception of physical phenomena as a scaled-down entity is likely to be unintuitive for most. However, it was interesting to note that half of the subjects experienced a strong PI despite the apparent improbability of the experience of being doll-sized. As the scale of operation decreases, perceived frictions and accelerations increase, which has already been found problematic for humans in robotic micro- and nano-level operations (Sitti, 2007). As the scale decreases further, these perceived distortions amplify, and additional phenomena such as fluid dynamics and static electricity, come into play as well. Relative changes in the environment would also provide additional challenges in the physical domain. For example, a floor that is experienced as smooth at a regular scale might become bumpy and full of cracks. Grit and dirt might become actual obstacles for navigation. Vibrations from passersby otherwise indistinguishable might feel like earthquakes. We also investigated the perception of physics at large scale. Study

2 enlarged the subjects 10-fold while giving them a similar interaction task. Although the users believed *movie physics* as realistic similarly to Study 1, the expectations of users was much more mixed. We believe this finding might be useful for designing abnormal-scale VEs where PSI is more important than realism such as games. Realistic physics in small-scale interactions greatly violated the expectations of users while in large-scale slow, realistic motions sometimes seemed natural, even if ultimately unrealistic.

We argue that our study opens up interesting avenues for future VR research. VR education has already been seen as a potential remedy for some issues of small-scale activities in the field of teleoperation (Millet et al., 2008). Further research on the effect of perception-related mismatches on interaction and performance in various applications could yield interesting findings. Also, as of now, we do not know whether the body-scaling effect affects the perception of physics the same way it affects the perception of sizes and distances (e.g., van der Hoort et al., 2011). In both studies, we used virtual hands to provide a body-based size cue, but we did not investigate the effect the absence of these cues would have had. Langbehn et al. (2016) found that groups of human avatars can override the dominant scale otherwise dictated by body-based size cues. Theoretically, this could have implications for perception of physics, as well.

5.2 Challenges and Limitations

Outliers in responses were L2 in Study 1 and L4 in Study 2. Inspecting the distribution of responses in question L2 in Study 1, we see that the *true physics* condition contains responses that are rather uniformly distributed in comparison to the *movie physics* condition; the STD in the *true physics* condition is twice as large as in the *movie physics* condition. Whereas in responses to L2 the *movie physics* condition was considered realistic (4, neither too fast nor too slow) by a vast majority, the *real physics* condition received an almost equal number of responses between 2 (too slow) and 6 (too fast). We suspect that the uncharacteristic distribution of the responses might be due to a poor wording of L2 (*The speed of pull tabs when thrown*). We changed the wording of this question in Study 2 to *“Time of Flight”*.

In Study 2, L4 (the bounciness of the logs) received mixed responses. Although mean and mode responses preferred *movie physics* similarly to other questions, the responses were overall more mixed. We consider bounciness as the most unrealistic aspect of Study 2, since we did not simulate splintering, or otherwise breaking the log due to impact. We considered these aspects as confounding variables in our scope.

In Study 1, according to both verbal comments during the experiment as well as responses to questions O1 and O2, some of the subjects starting with the *true physics* condition thought that the reason for their difficulty in throwing the tabs to a far distance was their own inability to use the controllers and not related to aspects of the VE. Although a training session helping to learn the controllers might have been helpful, it would have introduced unwanted priming. We received these types of responses far less in Study 2, which might be due to opposite behavior of objects when throwing.

Another limitation is the fact that it is currently difficult to realistically simulate object mass in VR since subjects can feel

only the weight of the controllers, even if there are pseudohaptic approaches for simulating weight such as Rietzler et al. (2018) and Samad et al. (2019). There were responses contemplating the role of weight and simulated arm strength on throwing distance, although simulating muscle strength in itself was not in the scope of either study. Human-scale arm motions were simply scaled either down or up by a factor of ten, which resulted in either very small or very large velocity imparted on the thrown object.

During a few experimental sessions, there were occurrences which could have broken presence or caused differences in the experiences of the participants, such as bumping into furniture, stepping onto the HMD cord, or minor physics bugs. Additionally, although we tried to keep the visual appearances of the two conditions as similar as possible in Study 1, the differences in the VE scale led to very subtle differences in their brightness. This deficiency was fixed in Study 2.

There were subjects who were not always paying close attention to the flying or falling characteristics of the tabs, or did not wait until the instructions were read in their entirety. This limitation was somewhat alleviated in Study 2 due to the stricter instructions given to the subjects.

Finally, we used the two distinct questions *realism* and *expectations* to separate what the subjects considered to be the objectively correct physics behavior from what matched their expectations. However, we cannot rule out the possibility that there were subjects who interpreted these questions to mean the same thing. Also, we did not have a control condition, in which the subject would have performed the tasks in a normal scale.

6 CONCLUSION AND FUTURE WORK

In this paper, we studied a phenomenon regarding the plausibility of physical interactions for scaled-down and scaled-up users in normal-sized VEs; when users interact with physically simulated objects in a VE where the user is scaled 10-fold smaller or larger from a regular human scale, there is a mismatch between expected physics and the accurate approximation of physics at that scale. A similarly sized and significant majority of both scaled-down and scaled-up subjects judged rigid-body dynamics close to human-scale realistic instead of what would be the correct approximation of realism at the resized scale the subjects were on. Almost all subjects at a small-scale considered rigid-body dynamics at that scale to be surprising, while the expectations of large-scale subjects were more mixed.

There are several interesting avenues for future research. It would be intriguing to study whether this phenomenon persists if a normal-sized subject observes, or virtually teleoperates, a giant or a doll performing the tasks presented above. Moreover, we will seek further evidence as to whether the extent of the illusion of being small or being large (modulated by various size cues) affects the perception of physics. In addition, it would be interesting to study if practicing with real-world objects prior to experiment affects the results. In the future, we also intend to study the body scaling effect and its influence on interactions with physically

simulated objects. In addition, we will investigate interaction, performance, and perceptual training at abnormal scales. We will consider scales smaller than 1 order of magnitude since we expect them to provide even greater plausibility mismatches in physical interactions. We will also seek to confirm the existence of our finding outside VR, for example using robotic teleoperation or telepresence at small scale.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MP and SL devised the research scope and the study concepts. MP developed both experimental applications and was responsible

for the overall design and execution of the studies. KM assisted in experimental application design, study design, data collection, and data analysis. AC assisted in data collection. The manuscript was written by MP, KM, AC, TO and SL.

FUNDING

This work was supported by the PIXIE (331822) and PERCEPT (322637) projects funded by the Academy of Finland, the COMBAT project (293389) funded by the Strategic Research Council at the Academy of Finland, as well as the HUMOR (3656/31/2019) project funded by Business Finland.

ACKNOWLEDGMENTS

This paper extends the previous work by Pouke et al. (2020) titled “The Plausibility Paradox For Scaled-Down Users in Virtual Environments” originally published in the IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR 2020). Parts of the original article are reused with permission. The authors also wish to thank all the subjects for their participation in this study. A video displaying the physics conditions in Studies 1 and 2 can be viewed at: <https://youtu.be/Cd0vK7RgZ9w>. A preprint of this article can be found in arXiv (Pouke et al. 2021).

REFERENCES

- Abtahi, P., Gonzalez-Franco, M., Ofek, E., and Steed, A. (2019). “I’m a Giant: Walking in Large Virtual Environments at High Speed Gains,” in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland, UK, May, 2019 (New York, NY: Association for Computing Machinery), 1–13.
- Alatalo, T., Koskela, T., Pouke, M., Alaveses, P., and Ojala, T. (2016). “VirtualOulu: Collaborative, Immersive and Extensible 3d City Model on the Web,” in Proceedings of the 21st International Conference on Web3D Technology, Anaheim, California, July, 2016 (New York, NY: Association for Computing Machinery), 95–103.
- Alex, J., Vikramaditya, B., and Nelson, B. (1998). “A Virtual Reality Teleoperator Interface for Assembly of Hybrid MEMS Prototypes,” in Proceedings of DETC’98 ASME Design Engineering Technical Conference, Atlanta, GA, September 13–16, 1998 (New York, NY: ASME).
- Banakou, D., Groten, R., and Slater, M. (2013). Illusory Ownership of a Virtual Child Body Causes Overestimation of Object Sizes and Implicit Attitude Changes. *Proc. Natl. Acad. Sci.* 110, 12846–12851. doi:10.1073/pnas.1306779110
- Bergström, I., Azevedo, S., Papiotis, P., Saldanha, N., and Slater, M. (2017). The Plausibility of a String Quartet Performance in Virtual Reality. *IEEE Trans. Vis. Comput. Graphics* 23, 1352–1359. doi:10.1109/tvcg.2017.2657138
- Billinghurst, M., Kato, H., and Poupyrev, I. (2001). The MagicBook: A Transitional AR Interface. *Comput. Graphics* 25, 745–753. doi:10.1016/s0097-8493(01)00117-0
- Bolopion, A., and Régnier, S. (2013). A Review of Haptic Feedback Teleoperation Systems for Micromanipulation and Microassembly. *IEEE Trans. Automat. Sci. Eng.* 10, 496–502. doi:10.1109/tase.2013.2245122
- Cross, R. (2004). Physics of Overarm Throwing. *Am. J. Phys.* 72, 305–312. doi:10.1119/1.1634964
- Deng, Z., and Interrante, V. (2019). “Am I Floating or Not?: Sensitivity to Eye Height Manipulations in HMD-Based Immersive Virtual Environments,” in 16th International ACM Symposium on Applied Perception, SAP 2019, Spain, Barcelona, November 19–20, 2016 (New York, NY: Association for Computing Machinery), 1–6.
- Dodgson, N. A. (2004). “Variation and Extrema of Human Interpupillary Distance,” in Stereoscopic Displays and Virtual Reality Systems XI, San Jose, CA, May 21, 2016 (Bellingham, WA: International Society for Optics and Photonics), 36–46.
- Gilbert, S. B. (2016). Perceived Realism of Virtual Environments Depends on Authenticity. *Presence-Teleop. Virt.* 25, 322–324. doi:10.1162/pres_a_00276
- Hatamura, Y., and Morishita, H. (1990). “Direct Coupling System between Nanometer World and Human World,” in IEEE Proceedings on Micro Electro Mechanical Systems, An Investigation of Micro Structures, Sensors, Actuators, Machines and Robots, Napa Valley, CA, February 11–14, 1990 (New York, NY: IEEE).
- Hongo, K., Kobayashi, S., Kakizawa, Y., Koyama, J.-I., Goto, T., Okudera, H., et al. (2002). NeuRobot: Telecontrolled Micromanipulator System for Minimally Invasive Microneurosurgery—Preliminary Results. *Neurosurgery* 51, 985–988. doi:10.1227/00006123-200210000-00024
- Interrante, V., Ries, B., Lindquist, J., Kaeding, M., and Anderson, L. (2008). Elucidating Factors that Can Facilitate Veridical Spatial Perception in Immersive Virtual Environments. *Presence-Teleop. Virt.* 17, 176–198. doi:10.1162/pres.17.2.176
- Izumihara, A., Sasaki, T., Ogino, M., Takamura, R., and Inami, M. (2019). “Transfantom: Transformation into Bodies of Various Scale and Structure in Multiple Spaces,” in ACM SIGGRAPH 2019 Emerging Technologies, Los Angeles, CA, February 2019 (New York, NY: Association for Computing Machinery), 1–2.
- Kim, J., and Interrante, V. (2017). “Dwarf or Giant: The Influence of Interpupillary Distance and Eye Height on Size Perception in Virtual Environments,” in Proceedings of the 27th International Conference on Artificial Reality and Telexistence and 22nd Eurographics Symposium on Virtual Environments, Los Angeles, Australia, November 22–24, 2017 (Geneve, Switzerland: Eurographics Association), 153–160.
- Kopper, R., Ni, Tao., Bowman, D. A., and Pinho, M. (2006). “Design and Evaluation of Navigation Techniques for Multiscale Virtual Environments,”

- in IEEE Virtual Reality Conference (VR 2006), Los Alexandria, VA, March 25–29, 2006 (New York, NY: IEEE), 153–182.
- Krekhov, A., Cmentowski, S., Emmerich, K., Masuch, M., and Krüger, J. (2018). “Gullivr: A Walking-Oriented Technique for Navigation in Virtual Reality Games Based on Virtual Body Resizing,” in Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play, Melbourne, VIC, October 2019 (New York, NY: Association for Computing Machinery), 243–256.
- Langbehn, E., Bruder, G., and Steinicke, F. (2016). “Scale Matters! Analysis of Dominant Scale Estimation in the Presence of Conflicting Cues in Multi-Scale Collaborative Virtual Environments,” in IEEE Symposium on 3D User Interfaces, Greenville, SC, March 19–20, 2016 (New York, NY: IEEE), 211–220.
- Lécuyer, A. (2009). Simulating Haptic Feedback Using Vision: A Survey of Research and Applications of Pseudo-haptic Feedback. *Presence: Teleoperators and Virtual Environments* 18, 39–53. doi:10.1162/pres.18.1.39
- Leyrer, M., Linkenauger, S. A., Bühlhoff, H. H., Kloos, U., and Mohler, B. (2011). “The Influence of Eye Height and Avatars on Egocentric Distance Estimates in Immersive Virtual Environments,” in Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization, Toulouse France, August 2011 (New York, NY: Association for Computing Machinery), 67–74.
- Li, Z., and Sitte, R. (2001). “Virtual Reality Modeling Aid in MEMS Design,” in Electronics and Energetics, Adelaide, Australia, November 21, 2011 (Bellingham, WA: SPIE), 4591. 153–162.
- Linkenauger, S. A., Leyrer, M., Bühlhoff, H. H., and Mohler, B. J. (2013). Welcome to Wonderland: The Influence of the Size and Shape of a Virtual Hand on the Perceived Size and Shape of Virtual Objects. *PLoS one* 8, e68594. doi:10.1371/journal.pone.0068594
- McCoy, J., and Ullman, T. (2019). Judgments of Effort for Magical Violations of Intuitive Physics. *PLoS one* 14, e0217513. doi:10.1371/journal.pone.0217513
- McIntyre, J., Zago, M., Berthoz, A., and Lacquaniti, F. (2001). Does the Brain Model Newton’s Laws? *Nat. Neurosci.* 4, 693–694. doi:10.1038/89477
- Millet, G., Lécuyer, A., Burkhardt, J. M., Haliyo, D. S., and Régner, S. (2008). “Improving Perception and Understanding of Nanoscale Phenomena Using Haptics and Visual Analogy,” in International Conference on Human Haptic Sensing and Touch Enabled Computer Applications, Madrid, Spain, June 10–13, 2017 (Berlin, Germany: Springer), 847–856.
- Mori, M., MacDorman, K., and Kageki, N. (2012). The Uncanny Valley [from the Field]. *IEEE Robot. Automat. Mag.* 19, 98–100. doi:10.1109/mra.2012.2192811
- Mujber, T. S., Szecsi, T., and Hashmi, M. S. J. (2004). Virtual Reality Applications in Manufacturing Process Simulation. *J. Mater. Process. Technol.* 155–156, 1834–1838. doi:10.1016/j.jmatprot.2004.04.401
- Nilsson, N. C., Nordahl, R., and Serafin, S. (2017). “Waiting for the Ultimate Display: Can Decreased Fidelity Positively Influence Perceived Realism?,” in IEEE 3rd Workshop On Everyday Virtual Reality (WEVR), Los Angeles, CA, March 19, 2017 (New York, NY: IEEE), 1–5.
- Ogawa, N., Narumi, T., and Hirose, M. (2017). “Distortion in Perceived Size and Body-Based Scaling in Virtual Environments,” in Proceedings of the 8th Augmented Human International Conference, Osaka, Japan, March 2017 (New York, NY: Association for Computing Machinery), 1–5.
- Ogawa, N., Narumi, T., and Hirose, M. (2019). “Virtual Hand Realism Affects Object Size Perception in Body-Based Scaling,” in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, March 23–27, 2019 (New York, NY: IEEE), 519–528.
- Patton, M. Q. (2005). *Qualitative Research: Encyclopedia Of Statistics in Behavioral Science*. New York, NY: John Wiley & Sons, Inc..
- Piumsomboon, T., Lee, G. A., Ens, B., Thomas, B. H., and Billingham, M. (2018a). Superman vs Giant: A Study on Spatial Perception for a Multi-Scale Mixed Reality Flying Telepresence Interface. *IEEE Trans. Vis. Comput. Graph.* 24, 2974–2982. doi:10.1109/tvcg.2018.2868594
- Piumsomboon, T., Lee, G. A., Hart, J. D., Ens, B., Lindeman, R. W., Thomas, B. H., et al. (2018b). “Mini-me: An Adaptive Avatar for Mixed Reality Remote Collaboration,” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal QC, Canada, April 2018 (New York, NY: Association for Computing Machinery), 1–13.
- Plisson, H., and Zotkina, L. V. (2015). From 2d to 3d at Macro- and Microscopic Scale in Rock Art Studies. *Digital Appl. Archaeology Cult. Heritage* 2, 102–119. doi:10.1016/j.daach.2015.06.002
- Pouke, M., Mimnaugh, K. J., Ojala, T., and LaValle, S. M. (2020). “The Plausibility Paradox for Scaled-Down Users in Virtual Environments,” in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, Atlanta, GA, March 22–26, 2020 (New York, NY: IEEE), 913–921.
- Pouke, M., Mimnaugh, K. J., Chambers, A., Ojala, T., and LaValle, S. M. (2021). The Plausibility Paradox for Resized Users in Virtual Environments. *Preprint: arXiv:2102.03179*.
- Renner, R. S., Velichkovsky, B. M., and Helmer, J. R. (2013). The Perception of Egocentric Distances in Virtual Environments-A Review. *ACM Comput. Surv.* 46, 1–23. doi:10.1145/2543581.2543590
- Rietzler, M., Geiselhart, F., Gugenheimer, J., and Rukzio, E. (2018). “Breaking the Tracking: Enabling Weight Perception Using Perceivable Tracking Offsets,” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland, United Kingdom, April 2020 (New York, NY: Association for Computing Machinery), 1–12.
- Rovira, A., Swapp, D., Spanlang, B., and Slater, M. (2009). The Use of Virtual Reality in the Study of People’s Responses to Violent Incidents. *Front. Behav. Neurosci.* 3, 59. doi:10.3389/neuro.08.059.2009
- Samad, M., Gatti, E., Hermes, A., Benko, H., and Parise, C. (2019). “Pseudo-haptic Weight: Changing the Perceived Weight of Virtual Objects by Manipulating Control-Display Ratio,” in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland, United Kingdom, May, 2020 (New York, NY: Association for Computing Machinery), 1–13.
- Senot, P., Zago, M., Lacquaniti, F., and McIntyre, J. (2005). Anticipating the Effects of Gravity when Intercepting Moving Objects: Differentiating up and Down Based on Nonvisual Cues. *J. Neurophysiol.* 94, 4471–4480. doi:10.1152/jn.00527.2005
- Shenai, M. B., Tubbs, R. S., Guthrie, B. L., and Cohen-Gadol, A. A. (2014). Virtual Interactive Presence for Real-Time, Long-Distance Surgical Collaboration during Complex Microsurgical Procedures. *J Neurosurg.* 121, 277–284. doi:10.3171/2014.4.jns131805
- Sitti, M. (2007). Microscale and Nanoscale Robotics Systems [grand Challenges of Robotics]. *IEEE Robot. Automat. Mag.* 14, 53–60. doi:10.1109/mra.2007.339606
- Skarbez, R., Brooks, F. P., Jr, and Whitton, M. C. (2017a). A Survey of Presence and Related Concepts. *ACM Comput. Surv.* 50, 1–39. doi:10.1145/3134301
- Skarbez, R., Brooks, F., and Whitton, M. (2020). Immersion and Coherence: Research Agenda and Early Results. *IEEE Trans. Vis. Comput. Graph.* 99, 1. doi:10.1109/TVCG.2020.2983701
- Skarbez, R., Neyret, S., Brooks, F. P., Slater, M., and Whitton, M. C. (2017b). A Psychophysical Experiment Regarding Components of the Plausibility Illusion. *IEEE Trans. Vis. Comput. Graphics* 23, 1369–1378. doi:10.1109/tvcg.2017.2657158
- Slater, M., Pérez Marcos, D., Ehrsson, H., and Sanchez-Vives, M. V. (2009). Inducing Illusory Ownership of a Virtual Body. *Front. Neurosci.* 3, 29. doi:10.3389/neuro.01.029.2009
- Slater, M. (2009). Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Phil. Trans. R. Soc. B* 364, 3549–3557. doi:10.1098/rstb.2009.0138
- Slater, M., Usoh, M., and Steed, A. (1994). Depth of Presence in Virtual Environments. *Presence-Teleop. Virt.* 3, 130–144. doi:10.1162/pres.1994.3.2.130
- Slater, M., and Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence-Teleop. Virt.* 6, 603–616. doi:10.1162/pres.1997.6.6.603
- Ullman, T. D., Spelke, E., Battaglia, P., and Tenenbaum, J. B. (2017). Mind Games: Game Engines as an Architecture for Intuitive Physics. *Trends Cogn. Sci.* 21, 649–665. doi:10.1016/j.tics.2017.05.012
- Usoh, M., Catena, E., Arman, S., and Slater, M. (2000). Using Presence Questionnaires in Reality. *Presence-Teleop. Virt.* 9, 497–503. doi:10.1162/105474600566989
- van der Hoort, B., Guterstam, A., and Ehrsson, H. H. (2011). Being Barbie: The Size of One’s Own Body Determines the Perceived Size of the World. *PLoS One* 6, e20195. doi:10.1371/journal.pone.0020195
- Yao, H.-Y., and Hayward, V. (2006). An Experiment on Length Perception with a Virtual Rolling Stone. *Proc. Eurohaptics* 2006, 325–330.
- Zhang, J., Dong, Z., Lindeman, R., and Piumsomboon, T. (2020a). “Spatial Scale Perception for Design Tasks in Virtual reality,” in Symposium on Spatial User Interaction, Canada, October, 2020 (New York, NY: Association for Computing Machinery), 1–3.

- Zhang, J., Piumsomboon, T., Dong, Z., Bai, X., Hoermann, S., and Lindeman, R. (2020b). "Exploring Spatial Scale Perception in Immersive Virtual Reality for Risk Assessment in Interior Design," in Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, April, 2020 (New York, NY: Association for Computing Machinery), 1–8.
- Zhang, X., and Furnas, G. W. (2005). mCVEs: Using Cross-Scale Collaboration to Support User Interaction with Multiscale Structures. *Presence: Teleoperators and Virtual Environments* 14, 31–46. doi:10.1162/1054746053890288
- Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y., and Agard, D. A. (2017). MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy. *Nat. Methods* 14, 331–332. doi:10.1038/nmeth.4193

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

Copyright © 2021 Pouke, Mimnaugh, Chambers, Ojala and LaValle. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.