The Plausibility Paradox In Small-Scale Virtual Environments

Matti Pouke
University of Oulu, Center for Ubiquitous Computing

Timo Ojala
University of Oulu, Center for Ubiquitous Computing

Katherine J. Mimnaugh
University of Oulu, Center for Ubiquitous Computing

Steven M. LaValle
University of Oulu, Center for Ubiquitous Computing

Abstract
This paper identifies a new phenomenon: when users interact with physically simulated objects in a virtual environment that is much smaller than usual, there is a mismatch between the object physics that they expect and the object physics that is actually correct. We report the findings of our study investigating the relationship between perceived realism and a physically correct approximation of realism in a small-scale virtual environment. We conducted a within-subjects experiment in which 44 subjects performed a simple interaction task with objects, using two physics simulation conditions while scaled down by a factor of ten in a virtual reality application. Although both conditions gave the visual impression of a scaled-down user interacting in a normal-sized environment, the physics conditions affecting the objects were different by simulating either correct behavior at that scale, or incorrect behavior similar to as if a normal sized user was interacting in a world that had been scaled up instead. We found that the significant majority of the users considered the latter condition to be the realistic one. We argue that our findings have implications on many virtual reality and telepresence applications involving operation with physically simulated or physical objects in small scales.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimodal Information Systems—Artificial, augmented, and virtual realities

1 Introduction
The question of whether our body influences how we perceive the world has been pondered in both century-old philosophical texts as well as in modern research. Indeed, many research studies have found a "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 [2, 10, 13, 16, 31]. If one, for example, was somehow shrunk to the size of a doll, that person would be inclined to regard the world as scaled-up and him/herself as normal-sized [31]. Right now, it is less known how such scaling down of oneself would affect a person’s perception of physical phenomena, such as accelerations. Interestingly, if we pay attention to how scaled-down characters interact with their surroundings in many works of fiction, the tendency to represent the world as scaled in comparison to normal-sized protagonists can be observed as well. 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 by rainwater holding onto a pencil, the water and the pencil act more akin to a river and a log. While 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.

Virtual reality (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 scaled-down virtual environments (VEs) are the 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., [9, 22]). In addition, the scaling of users has been utilized in several collaborative mixed reality (MR) systems (e.g., [3, 18]). Teleoperation of robots can require humans to interact with the physical world at micro- and nanoscale. While teleoperation in the physical world can leverage stereoscopic camera systems resembling immersive VR applications [6], purely virtual representations leveraging computer graphics can be used in, for example, educational and training systems for micro- and nanoscale tasks [4, 13]. Robotic surgery systems can perform operations at a microscopic level [7] whereas stereoscopic VR can be utilized in telesurgery [22]. The benefits of VEs have been identified in various design and prototyping processes [15]; these processes can be extended into scaled-down VEs as well. Already two decades ago, both the design [12] and assembly [1] of micro-electrical mechanical systems (MEMS) were prototyped through desktop VEs.

While existing research has addressed many perceptual questions, such as the perception of distance and size after altering one’s virtual size, 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, however, argue that it is not inherently intuitive for humans to perceive physical phenomena, such as rigid body dynamics, in scales that differ greatly from normal human scale. An object dropped from 20cm takes significantly less time
falling than an object dropped from 2m, and their perceived accelerations are also different. Additional physical phenomena, such as fluid dynamics, frictions, and static electricity are similarly becoming more significant when the scale of the operations become smaller. For this reason, additional consideration is required when designing systems in which real or virtual interactions take place in abnormal scales, and it is important to understand human perception regarding physical phenomena while scaled. In this paper, we present our early results investigating the aforementioned phenomenon. More specifically, we will focus on the mismatches between perceived realism and the approximation of physical realism when interacting in a VE while scaled down by a factor of ten. We hypothesize that people are not blind to changes in scale; however, when presented with multiple physical approximations are more likely to consider the one closest to regular human scale to be the most realistic one.

This paper progresses as follows. Section 2 presents previous research related to this work. Section 3 will outline our research method while in section 4 we will go through our experimental setup. Section 5 will introduce our results. Section 6 discusses our findings while section 7 concludes the paper.

2 Related Research

First-person scalable objective aspects have various subjective effects that have been studied from perceptual and psychological standpoints. One of the most obvious is viewpoint height, as it defines the virtual camera origin in relation to the VE and affects egocentric distance perception [11, 12]. Users’ interaction capabilities such as locomotion speed and interaction distance can be changed according to scale, depending on the purpose of the application [12]. 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 two virtual cameras that are used to render the environment for the user. Changing this distance can affect the users sense of their own size relative to the VE [8, 18].

Humans generally seem to utilize their own body as a primary metric for scale (an effect also referred to as body scaling), and the virtual representation of the user’s body greatly affects their perception of sizes and distances in the virtual environment [10, 11, 17]. Linkenauger et al. [11] studied the role of 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. [17] 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. [31] embodied the entire user in a dolls body as well as in a giants body using a stereoscopic video camera system and a HMD. They found that the embodiment significantly affected the users distance and size perceptions, especially if the user experienced a strong body ownership illusion [25] with the virtual body. Banakos et al. [2] compared the effects of embodying the user as a child vs a scaled-down adult. They found that not only the effect of altered size and distance perceptions was even larger when embodied as a child, but it also made the users associate themselves with childlike personality traits.

The environment, whether real or virtual, affects the perception of scale. Humans generally underestimate egocentric distances in VEs, except when the VE is faithfully modeled to represent a real environment [20]. However, if a familiar room is scaled slightly up or down, underestimations are reintroduced [29]. Langbehn et al. [10] 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, while 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.

Studies in micro- and nanoscale teleoperation have revealed that, due to changes in physics, interactions at these scales can become difficult, but education inside virtual reality environments has been found to alleviate this drawback [14, 23]. Besides this work, there is little research on human perception of physical phenomena in scaled-down VR.

2.1 Presence and Plausibility

The concepts of immersion [28], presence and plausibility [25] are relevant for this study. In Slaters 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) [28]. In addition, the realism of the users response to the VR system depends on two orthogonal components, presence or place illusion (PI) and the plausibility illusion (PSI) [25]. PI refers to the sensation of being in another place, while PSI refers to the perceived believability of the virtual scenario or experience (illusion as being there vs. realness of what is happening) [21]. PSI depends on the extent to which the VE can produce realistic responses for user actions. Rovira et al. [21] 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 (i.e., virtual characters react to the user). In addition, the VE should match the users prior knowledge and expectations [21]. Skarbez et al. [24] 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 while immersion is a technical attribute that affects PI, coherence is a similar technical attribute affecting PSI.

In this study, we used the concept of PSI to study human percep-
We designed an experiment in which the subjects performed a simulation of a doll-sized perspective. We did not use full body tracking or operations took place in one order of magnitude, giving the impression of being virtually shrunk down. More specifically, our hypothesis was that in the true physics condition, the 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 between true physics and movie physics, and we predicted that subjects would be more likely to feel and to expect the movie physics condition to be the more realistic representation. This implies the Plausibility Paradox, a mismatch between perceived realism and the correct approximation of realism.

3.1 Hypotheses

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

H1: In small scale environments, movie physics is more likely to feel realistic to a user than true physics.

H2: In small scale environments, movie physics is more likely to match a user’s expectations than true physics.

3.2 Experimental Apparatus

We designed an experiment in which the subjects performed a simple interaction task in Unreal Engine 4.22 (UE) based VEs using physics conditions described above. 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 [26]. However, we used the default UE VR hand visualization for interaction and to present a medium-fidelity body size cue [17].

To help providing realistic size cues, we modeled the VE acting as the base for the experiment to resemble a location at the main corridor of the campus in which the study took place. The dimensions and materials of the VE were modeled using the real environment as the basis. 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 built-in World to Meters parameter, which automatically scales the player 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 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.

3.2.1 Interaction Task

The interaction task consisted of the manipulation of virtual soda cans. The tabs were chosen for the experiment both for their small, consistent mass as well as being a reasonably realistic object to be seen in the simulated VE. We considered a lightweight object to be most realistic for simulating throwing in VR so that one would not need to simulate the decrease in hand acceleration due to increased inertia at the end of the arm or by limited arm strength [5]. 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 while 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 0.175s at 20cm 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 0.6375s at 2m 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 did 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. Figures 2A 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 1g 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. The simulation itself ran at stable 80 FPS which is the maximum frame rate of Oculus Rift S.

3.3 Experimental Procedure

The experiment was carried out as a within-subjects experiment, in which 44 subjects (23 females and 21 males) performed both conditions during one experiment. The order of the conditions was counterbalanced so that there was an equal number of male and female participants starting with each condition. The subjects’ age ranged from 19 to 66, mean and median ages being 30 and 26, respectively. 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.

The experiment was set in a laboratory in which the subjects used the Oculus Rift S system for the experiment. In the beginning of each session, the subjects read through a written Information for Subjects document and signed an informed consent sheet. The subjects were then instructed on using VR hardware, specifically how to use the Rift S motion controller for picking up and throwing objects. Next, they were instructed to stand on a particular starting spot in the laboratory (marked with masking tape), which was 110 cm away from the laptop used for the HMD. When the user was wearing the HMD and the motion controllers comfortably, the following instruction script was read in English or Finnish: “In this experiment, you are in a virtual reality environment, where you are at the university central hallway at night. You have been shrunk down to a size of a barbie doll, approximately 10-times smaller than your current height. You can move around a little bit by taking a few steps (but you don't have to). You will see several pull tabs placed on a book in front of you. We would like you to pick one up and then let it fall to the floor. After that, we would like you to pick one up and throw it across the book in front of you. We would like you to try dropping and throwing the remaining pull tabs as well. After no pull tabs are remaining on the book, we will restart the experiment and ask you to repeat what you just did with the pull tabs. I will now put on the headphones, and then you may begin.”

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 was removed and the subject was asked to respond to a post-experiment questionnaire as well as a background questionnaire on a different laptop (seen left in Fig. [3]). The subjects were asked for any additional comments or questions, and if they can be contacted for future studies, and then given a gift certificate for two euros for participation.

3.3.1 Questionnaires

We collected plausibility related data using two forced choice questions (main questions 1 and 2), 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-Usoh-Steed Presence questionnaire [27][30], as well as a background information questionnaire. The main questions 1 and 2 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?

The main questions 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 or movie physics or vice versa). The first three questions (L1-L3) were bipolar while 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)

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.

4 Results

We used a coding mechanism in Subject IDs to identify the order of conditions for each subject and analyzed the results accordingly. Two subjects were removed from the analysis due to significantly different conditions from issues with the functionality of the software or due to vision impairments.

According to the responses to the main questions, the majority of the subjects considered the movie physics condition as the more realistic one. Out of 44 subjects, 33 participants (73%) responded to the first question that they considered the movie physics condition more realistic, which confirms H1. For the second question, 42 out of 44 (93%) 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 $p = 1.7051^{-8}$ 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 implies that subjects were neither unable to distinguish between the physics conditions, but actually more consistently picked the movie physics response, which was the unrealistic 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 the true physics better matched their expectations.

4.1 Understanding the 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.
Figure 4: A Summary of responses for questions L1-L3 sorted according to differential. Responses closer to 4 are perceived as closer to realism

4.1.1 Open Ended Questions O1 and O2

The purpose of the open-ended questions was to evaluate to what extent the subjects’ responses to the main questions 1 and 2 were related to the perceived realism of the physics. For O1, vast majority of the subjects responded with a reason that can be attributed to the physics conditions (e.g., "Gravity feels more natural," "the tabs were flying in a more natural way" or "Second time they felt too heavy"). Some subjects also gave a secondary reason unrelated to physics (e.g., "movement in space felt more realistic, but the objects lacked 3D, ring pulls are not paper thin" or "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").

Five subjects out of 44 gave a response related to general interaction or learning effect as the primary reason (e.g., "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" or "I’m kind of feel the same for both times. But maybe the second is more realistic just because I get use to it").

Finally, one subject gave a reason completely unrelated to interaction ("the Coke can made the situation plausible."). Observing recorded video material, it can be seen that this subject stopped to admire the Coca Cola can for a moment after one of the tabs landed close to it in the movie physics condition. However, we do not know the exact reason why the Coca Cola can was chosen as the response. The cans were the same in both conditions so it may have been a more general comment.

Open question O2 asked the subjects to report the reason for their choice of answer for the main question 2, asking which condition matched their expectations more. The subjects who had different responses for main questions 1 and 2 reported their justifications for this (e.g., "I was not thinking I was shrunk. So it felt strange to have such heavy pull tabs," "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," or "Even though I knew I was shrunk down, I somehow could not think about it that way while doing the experiment").

The subjects who chose movie physics confirmed their answers for their first response by simply referring to their earlier response, or providing additional reasoning, such as "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" and "In the second time, the tabs were falling down surprisingly fast."

The subjects who chose movie physics while reporting reasoning unrelated to physics in O1 mostly confirmed this reasoning in O2 as well. For example, "I was paying more attention to the behavior of the pull tabs, while in the first time the environment caught most of my attention" (O1). "For the same reason as the previous response, although for some reason I liked the color of the pull tabs better in the first time, it was somehow more clear" (O2). However, one subject referring to general interaction in O1 ("It felt easy to pick them up") gave another reason in O2 that seems more associated with physics, "First time throwing them felt more natural" (O2). One of the subjects preferring movie physics expressed doubts for his responses to both questions with the statement in O2 as "The be-
We used a binary logistic regression to analyze the effects of subject background and presence on their responses to main question 1. We used Educational Background, Gender, Age, VR Experience, Gaming Experience, SUS Average and SUS Score as independent variables and the response to main question 1 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 were 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 primary reasons related to the behavior of the physically simulated tabs. Other primary reasons were related to general interaction and learning effects. Three references were made to visual details as secondary reasons or general remarks (the thickness of tabs in O1 and two references to colors in O2).

4.1.2 Likert Responses
Inspecting the Likert responses for questions L1-L5, we found that the movie physics condition was closer to perceived realism (average responses closer to 4 in questions 1-3 and closer to 7 in questions 4 and 5) in all questions L1-L5. However, in question L2, the difference is small compared to other questions. We analyzed the responses to questions L1-L5 with the Wilcoxon Signed Rank test and found that the responses were significantly different (p < 0.005) 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, mean, mode and standard deviation of questions L1-L5 can be seen in Table 1.

### Table 1: Likert questionnaire data summarized. Responses perceived closer to realism are emphasized in bold

<table>
<thead>
<tr>
<th>Question</th>
<th>Condition</th>
<th>Mean</th>
<th>Mode</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>True Physics</td>
<td>5.3</td>
<td>6</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>Movie Physics</td>
<td>3.8</td>
<td>4</td>
<td>0.8</td>
</tr>
<tr>
<td>L2</td>
<td>True Physics</td>
<td>4.2</td>
<td>6</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Movie Physics</td>
<td>4.1</td>
<td>4</td>
<td>0.9</td>
</tr>
<tr>
<td>L3</td>
<td>True Physics</td>
<td>2.2</td>
<td>2</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>Movie Physics</td>
<td>4.3</td>
<td>4</td>
<td>1.1</td>
</tr>
<tr>
<td>L4</td>
<td>True Physics</td>
<td>2.7</td>
<td>2</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>Movie Physics</td>
<td>4.9</td>
<td>6</td>
<td>1.5</td>
</tr>
<tr>
<td>L5</td>
<td>True Physics</td>
<td>2.5</td>
<td>2</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>Movie Physics</td>
<td>4.8</td>
<td>6</td>
<td>1.4</td>
</tr>
</tbody>
</table>

4.1.3 Effect of Background and Self-Reported Presence
We used a binary logistic regression to analyze the effects of subject background and presence on their responses to main question 1. We used Educational Background, Gender, Age, VR Experience, Gaming Experience, SUS Average and SUS Score as independent variables and the response to main question 1 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 were 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 average of responses as well as the computed SUS score (the number of responses per subject with a score of 6 or 7) [27]. 36 out of 44 subjects had a SUS Score larger than 0, with the median score being 3.

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 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 background, education or gaming in our subjects. The level of presence according to self-reported SUS score did not have any effect either.

4.1.4 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. For example, in responding to why they selected the movie physics as more realistic, one subject commented in English “The pull tabs looked and felt heavier and were easier to throw, as I would expect.” A second subject also commented that the pull tabs in the movie physics condition felt heavier, “They fell in the right place, they had weight and they flew in a realistic projectory [sic].” However, another subject, remarking on why the movie physics was more realistic, said: “Second time they felt too heavy,” referring to the perceived increase in weight of the tabs during the real physics condition. Several of the responses in Finnish related to how the tabs should have felt heavy or about how much more power they would have needed to use to throw the tabs given their reduction in size. It is interesting to consider these spontaneous responses regarding differences in the weight of the tabs given than there was no change in the controllers that the subjects used for each condition. Subjects held the controllers through the whole experiment and while the condition was changed without ever setting them down. 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 tabs felt heavier).

5 Discussion
The results imply that we have identified a strong paradox concerning PSI in small-scale VEs. According to the results, almost a 3/4 majority (73%) of the subjects found the movie physics condition to be more realistic. In addition, a 9/10 majority (91%) of subjects considered the movie physics condition as better matching their expectations. From this, we conclude that even subjects who believed the true physics to be a correct representation of reality still considered it to be surprising. This reasoning was also often present in the responses to open questions O1-O2. According to O1-O2, almost all of the subjects considered the perception of realism to be related to the physics behavior of the tabs, or for general interaction reasons or the learning effect. A few secondary reasons or remarks were made referring to a scene object or other visual details.

We used the Likert scale questionnaires to gather additional insights and confirmation for our results. The questions were focused on various dynamic properties of the tabs so that we could more specifically pinpoint the effects of physics simulations on perceived realism. These responses imply preferences towards movie physics as well, with significant differences regarding the perceived realism of the tab behavior (with the exception of question L2). However, in this question as well, the most popular response indicated a preference for the realism of the movie physics (see Fig. 6B. Throwing speed was neither too fast nor too slow). The Likert responses give us additional confirmation that realistic representations of the physics in scaled-down VEs are not inherently intuitive for users. According to the results, accurate accelerations and falling speeds of objects were perceived as unrealistic. The distance that the subjects
VR is likely to be unintuitive for most. As the scale of operation world or with realistically simulated physics. As can be seen by our were enlarged 10-fold.

were able to throw the tabs was seen mostly as too short (although there were also responses that considered the movie physics enabling too far throwing distances even if true physics was considered short). In addition, responses regarding the bounciness of the of the tabs imply that subjects expected the tabs to behave similarly as if they were enlarged 10-fold.

We inspected the effects of various aspects of the subjects’ background on their responses to O1. It could be that that subjects with a knowledge of physics, for example, might prefer the true physics condition. However, we found no such effects in our subject group. In addition, we did not find self-reported level of presence (27) affecting the response to O1 in our subject group.

5.1 Implications

In this paper, we introduce the Plausibility Paradox in small-scale VEs - when the expectations of a user do not match with reality. We argue that this finding has potential future implications to VR and telepresence applications. Through recent advances in consumer VR hardware as well as sub-microscopic [19] and even atomic [33] level imaging techniques, it is possible that we will witness an increasing exploitation of scaled-down VR in the future. These utilizations could potentially include commercial systems outside of the scientific domain, such as with teleoperated maintenance robots or commercial virtual design solutions at a microscopic scale. However, at this stage, it is not known whether it is 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 by our initial results, the perception of physical phenomena in scaled-down VR is likely to be unintuitive for most. 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 [23]. 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 effectively smooth on a regular scale might become bumpy and full of cracks. Grit and dirt might become actual obstacles for navigation. Vibrations from passersby that would be otherwise indistinguishable might feel like earthquakes.

We argue that these challenges provide 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 [4].

5.2 Challenges and Limitations

An obvious outlier in the responses was the question L2 (see Figures 6 and 43). Whereas in the other questions, the responses seem relatively consistent indicating a stronger preference towards one condition or the other, L2 is an exception. Inspecting the distribution of responses in question L2, it can be seen 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 movie physics. Whereas responses in the L2 movie physics condition was considered realistic (4, neither too fast nor too slow) by vast majority, the real physics condition received almost equal number of responses between 2 (too slow) and 6 (too fast). We suspect that the uncharacteristic distribution of responses might be due to a poor wording in the question L2, The speed of pull tabs when thrown. While we tried to ask how the subjects perceived the time of flight of the tabs, it could be that subjects had other interpretations for the question resulting in inconsistent responses. Similar inconsistency is being present in responses from both Finnish and English speaking subjects.

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 a result of their own inability to use the controllers and not related to aspects of the environment. Although some subjects realized during the subsequent movie physics condition that the behavior of the tabs was an experimental manipulation and not due to their own failure, there were still three subjects that stated as their main reason for preferring the movie physics condition to be the fact that they had learned how to use the controllers. For subjects that had the movie physics first, there did not seem to be any ambiguity that the difference in the behavior of the tabs was related to the environment. While a training session helping to learn the controllers might have been helpful, we believe that it could have introduced unwanted priming for subjects regarding the expected behavior of physics.

Another obvious limitation is the fact that it is currently difficult to realistically simulate object mass in VR. While we chose the soda can pull tabs for the task partly because of their light mass, there was some speculation among responses to O1-O2 on whether the weight of the object and/ or simulated arm strength affected object manipulation.

During a few of experiment sessions, there were occurrences which could have broken presence or caused differences in the experiences of the participants. Two subjects became very active in the virtual environment and accidentally bumped into furniture experiences of the participants. Two subjects became very active in the virtual environment and accidentally bumped into furniture.
the differences were nearly impossible to distinguish, there were two responses to O2 that commented on differences between visual appearance of the conditions.

Finally, 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 reading of the instructions was finished. Not observing the tabs properly might have introduced inaccuracies to their responses. This came up with both verbal comments after the experiment as well as responses to O1-O2.

6 Conclusion and Future Work

In this paper, we present a novel phenomenon regarding the plausibility of physical interactions in scaled-down VE; when users interact with physically simulated objects in a VE that is much smaller from regular human scale, there is a mismatch between what they expect and the object physics that is the correct approximation of reality. We argue that this finding opens many interesting avenues for future research regarding mCVEs, scaled-down VR applications in general, as well as telepresence and teleoperation taking place on a reduced scale. Although the Plausibility Paradox discussed here is specifically related to small-scale environments, there are most likely other situations in VR which similar mismatches can exist.

Regarding small-scale VEs, in the future we intend to focus more on, for example, body scaling effect, and its effects on interaction with physically simulated objects. In addition, we consider scales smaller than 1 order of magnitude interesting since we expect them to provide even greater plausibility mismatches in physical interactions. In addition, the subjective perception of weight that appeared in some of the open ended responses can provide interesting research avenues.

Acknowledgments

The authors wish to thank all the subjects for their participation in this study. This work was supported by the COMBAT project (293389) funded by the Strategic Research Council at the Academy of Finland and the PERCEPT project (322637) funded by the Academy of Finland as well as the HUMORcc (6926/31/2018) with physically simulated objects in a VE that is much smaller from regular human scale, there is a mismatch between what they expect and the object physics that is the correct approximation of reality.


[25] M. Slater. Place illusion and plausibility can lead to realistic behaviour


