



The Effect of Multimodal Feedback on Perceived Exertion on a VR Exercise Setting

Jon Ram Bruun-Pedersen^(✉), Morten G. Andersen,
Mathias M. Clemmensen, Mads K. Didriksen, Emil J. Wittendorff,
and Stefania Serafin

Aalborg University Copenhagen, A. C. Meyers Vænge 15,
2450 Copenhagen SV, Denmark
{jpe, sts}@create.aau.dk,
{morandl4, mclemm14, mdidril4, ewitttel4}@student.aau.dk

Abstract. This paper seeks to determine if multimodal feedback, from auditory and haptic stimuli, can affect a user's perceived exertion in a virtual reality setting. A simple virtual environment was created in the style of a desert to minimize the amount of visual distractions; a head mounted display was used to display the environment. Users would in the real world drive a Combi Bike and the velocity generated was translated to a vehicle in the virtual environment, moving it forward on a predetermined path. Each user traversed a total of eight hills, two in each of the four conditions. The perceived exertion was measured several times during each condition using the Borg Scale. The results show that there is no significant difference between the four conditions, which had different combinations of auditory and haptic feedback.

Keywords: Virtual reality · Exertion · Multimodal stimuli · Exercise bike

1 Introduction

It is reasonable, that what we have grown to understand about real life settings, directly influences our perception and interpretation of situations or interactions in virtual reality. The study of perceived exertion from physical activity in this relation is interesting for several reasons. It is not highly documented how individual modalities dominate specific parts of the experience of exertion when physical activities are augmented in virtual reality (VR), and its application might be particularly useful in areas such as exercise or rehabilitation using VR. A manipulation of a user's perceived exertion level, could lead to a change of the sensation of fatigue, or the difference between the experience of accomplishment or failure. For instance, for users not otherwise prone to exercise, regulating their sense of exertion could have an effect on their overall exercise experience. If a sense of accomplishment, e.g. for an elderly practitioner, is in part based on a perception of conquering a vast exercise task (such as lifting a series of objects perceived as heavy), an illusion of high exertion could be constructive. If weakened practitioners (for instance, cancer patients) need to exercise but cannot believe themselves capable, perceiving a lower exertion level with that exercise could make the exercise appear easier, and thus be valuable. VR has

previously shown its merits as a motivation tool for VR rehabilitation [1–3], and manipulation of users’ perception of exertion seems possible [4, 5], while uncharted, especially for VR. This paper seeks to explore how to manipulate a user’s perception of exertion in immersive VR. More specifically, the paper presents an experiment where participants are asked to use an immersive VR biking simulation, in which they drive up- or downhill inside the virtual environment, with varying multimodal feedback, and rate the perceived level exertion for each run for each.

2 Related Works

The body of studies on the manipulation of perceived exertion through the use of multimodal feedback are sparse. But while results have not been consistently significant between recent studies on the perceived exertion under certain feedback conditions [4, 5], a shared finding has been that users have indeed perceived a change in perceived exertion, when the activity is augmented with auditory feedback, from variations to pass filters and amplitude.

Looking at studies on behavior related to physical activity based on feedback in VR, Plotnik et al. [6] show how changing the inclination in VR (with walking) altered participants’ gait, even when a physical treadmill was level. Meanwhile, Denton [7] investigated how locomotor respiratory coupling was affected by inclined walking in VR but found no significant changes from level walking and inclined walking. However, both studies used projections on canvas and not head mounted displays (HMD), and no additional multisensory feedback. Even so, Plotnik et al. [6] showed participant reactions to the stimuli provided. Looking specifically at studies using non-visual sensory feedback to affect perceived exertion, Bordegoni et al. [4] showed how altering the frequency and amplitude of the sound, produced by friction from a pulley machine, affected the perception of exertion, using the Borg rating of perceived exertion (RPE) scale [8]. By manipulating sound in real-time, generated from the metal wire connecting the pull handle to the weight bricks, the pulling sound was augmented through headphones, as feedback suggesting increased or decreased weight of the pull. After each pull, participants were asked to rate their level of exertion on the RPE scale. Despite the physical weight on the pulley always remained the same for all conditions, results showed significant differences of exertion for frequency content between low-pass filters at 1000 Hz (more exertion), compared to a highpass filter-split at the same frequency (less exertion) [4].

In subsequent study, Bruun-Pedersen et al. [5] investigated the role of similar auditory feedback, to measure the differences in perceived exertion with a VR based biking simulation. The setup used a DeskCycle couch-bike, a 55’ LED TV for visuals and speakers for audio [5]. When the user pedaled, the TV would provide visual feedback, of moving through a park VE. The base auditory feedback was the sound of wheels, chain and cabin of recumbent bike. The behavior of the feedback (speed, etc.) corresponded with the speed of a user’s pedaling. Similar to Bordegoni et al. [4], this study used 9 conditions; 3 frequency settings and 3 amplitude settings [5]. Participant’s rated their perceived exertion level using the RPE scale [8]. RPE results showed no significant differences between conditions, but qualitative measures suggested that the

majority of participants perceived gradual changes in exertion between conditions. Changes which many participants described as a difference in mechanical resistance from the bike. Perhaps more interesting, was the fact that no participants had been aware of the differences in auditory feedback, while performing experiment trials. When presented to some participants post-trial, however, participant responses were that the individual feedback conditions were considered very different from each other [5] (Fig. 1).



Fig. 1. The park VE used by Bruun-Pedersen et al. [5]

The different results between the studies by Bordegoni et al. [4] and Bruun-Pedersen et al. [5] came from similar experiment methodologies. Between the studies, the physical demands by the fitness equipment (i.e. the weight load of a pulley machine [4] and resistance from an DeskCycle [5]) remained constant between conditions, and the auditory feedback (despite representing two different mechanical dynamics) had the very similar condition properties. However, the pulley exercise was centered around a visible, mechanical fitness device, placed in a room with no visual distractions other than the exercise and machine itself [4]. In bike study, the DeskCycle was a placeholder for a digital object (recumbent bike) in a digital environment; a park VE with loads of visuals. The park VE was presented on a TV monitor, while the mechanical pulley was a natural extension of the surrounding environment. In essence, there were several differences between the studies in some key areas. And for the experience of a VE, the (multimodal) display used to mediate the VE plays a decisive role for the effect of the experience, and the relation to the activities performed in that environment.

2.1 The Relevance of Immersion and Presence

Studies on the manipulation of perceived exertion through on multimodal feedback are sparse, especially in relation to immersive VR (IVR), which in this paper is to understood through the definitions of immersion (and presence) offered by Slater [9]. In a VR system, increasing the spectrum of sensory stimuli provided by the system to mediate a virtual environment (VE) to a user, corresponds to increasing the immersive properties of that system [9]. Immersive systems can also be characterized by the collective sensorimotor contingencies (SCs) supported by the system, which describes the number of possible ‘valid actions’ the user is able to perform inside the VE with the system. SCs can be split into two categories; (a) valid sensorimotor actions; actions

which changes the current VE perception (in all sensory modalities), and (b) valid effectual actions; which are user actions which affects changes perception of the VE or changes to the VE [9]. The gain from a highly immersive system, is the potential for a sense of presence inside the VE. Slater defines presence as the combination of a place illusion (PS); users' experience of being transported to another place (even despite being certain that are in fact not), and a plausibility illusion (Psi); which is users' belief that behaviors or occurrences inside that (virtual) place are experienced as occurring (even despite being aware that they are not) [9]. But while system immersion is objectively measurable by its stimuli and SCs, presence is not. Two people can experience different sensations of presence, based on how their actions affect their 'journey' in the VE, and how well it feeds their PS and Psi, and thus their experience in the VE [10].

The plausibility illusion is suggested by Rovira et al. [11] to depend on three conditions; (a) user actions must produce correlated reactions in the VE, (b) the VE must respond directly to the user, also in situations where the user is passive, and (c) the VE and the events occurring in it must be credible, based on the life-time, real-world experience of the user. Designing a VR experience for a specific purpose, in a specific context, should therefore appropriately consider what SCs that users are led to experience, how they are mediated by the system and how consistently users are expected to receive and interact with the mediation. Stimuli given from the system needs to meet the expectations of the user, if intended to reflect the behavior and feedback from a real-world environment.

2.2 Issues to Consider for Exertion Manipulation

It can be argued, that the effect of the auditory feedback from pedaling, in the study by Bruun-Pedersen et al. [5], might have been disrupted by the comprehensive audiovisual stimuli from the VE park (visual flora and fauna, soundscape stimuli from running water, birds, wind, etc.). While the ecology of the park VE might have represented such environment faithfully, it was likely unsupportive of participants' focus on- and perception of the feedback. In hindsight, it can also be argued that the study by Bruun-Pedersen [5] failed to afford the focus and attention on the exercise activity, compared to the disturbance-free experience of the purpose-supportive gym environment of the pulley [4]. In addition, the start position of the VR bike changed between participants, meaning that path characteristics and soundscape features in the VE varied slightly between trials, possibly influencing the consistency of the conditions between trials, and thus the results in the study [5]. Another, more general consideration made obvious from the studies, is a likelihood for the Borg scale to be interpreted depending on the physical attributes of the individual participant, potentially resulting in differences between participants' interpretation of the RPE.

Future studies on using feedback stimuli to manipulate perceived exertion, should thus be able to benefit from system with highly immersive properties, based on the potential to (a) escape possible real-world environment distractions of less immersive systems, (b) focus the participants attention on the activity itself by, (c) offering virtual environment-task disturbance-free and purpose-supportive environment, (d) an expanded usage of SCs, providing, (e) a heightened potential for focus on the task from

a potentially high sense of presence in terms of both place and plausibility, but (f) upholding the guidelines (for Psi) of how actions must produce correlated reactions in the VE, the VE must respond directly to the user, and how the VE and the events occurring in it must be comparable with real-world experience of the user.

2.3 Haptics as Feedback

Taking reference to the studies from Plotnik et al. [6] and Denton [7], revisiting the idea of users' perception when ascending or descending a virtual hill seems interesting. Only this time with biking, and using audition, visuals and haptics for a multimodal IVR mediation, rather than simply visuals from the previous hill-based studies [6, 7].

The relevance of haptics can be seen in various multimodal uses in VR, such as for instance Nordahl et al. [12], where pressure sensors and actuators were combined with motion capture and implemented into a pair of shoes for use in VR. The shoes served as the controller for a surround sound system, as well as the physically based audio-haptic synthesis engine [12]. This footwear-based interaction simulated the feeling of walking on different surfaces. The goal of the study was two-fold; (a) to test whether multimodal feedback in the form of audio and haptics improved the task of walking on a virtual rope while being blindfolded, and (b) to test the importance of auditory and haptic feedback in a VE [12]. Results indicated that participants who started with the audio-haptic feedback had higher mean presence scores. Besides the effect of the haptic feedback, it also showed that the degree of similarity, of locomotive techniques in VR to real life locomotive movement, had a positive influence on presence [12].

2.4 Biking in IVR

One of the challenges listed by Nordahl et al., was how virtual egocentric motion should try to correspond to the movement in real life, which however often introduces the issue of the physical limitations of room dimensions [12]. Being a stationary exercise tool, an exercise bike such as the manuped is a useful platform to overcome issues of physical room limitations and should support the real-life experience sensation of egocentric motion, even when not physically moving to a degree. A manuped activates (and occupies) both hands and feet, during exercise sessions. This means that the sensation of egocentric motion inside the VE should be connected to the manuped activity, and possibly considered congruent with a real life sensation of driving a biking device.

Contextualizing the manuped inside the VE is necessary for users' association between the real world and VE. As before, this will be based on the combination of pedal activity on the manuped, and a virtual biking-centric device, which a user controls to move forward inside the VE, based on the pedaling [11].

2.5 The Immersive Properties

Increasing visual immersive properties of the biking experience will come from the multiple modalities (including haptics), as well as a significant increase in the visual domain by introducing an Oculus CV1 VR headset, compared to the previously

mentioned exertion studies [4, 5]. In addition, the study will introduce a new scenario to the perception of exertion, through the task of riding a path which ascends and descends over a hill, taking inspiration from [6, 7]. According to the real-life experience from biking, going downhill should be considered less exertive than going uphill. Meanwhile, creating the sensation with users of experiencing higher exertion going uphill and lower exertion going downhill in VR is not trivial. The next part of the paper will detail how the VR setup has been designed to mediate the drive across different height levels, and how this should incentivize users to experience a different exertion driving uphill and downhill, concurrent with their real-life experiences.

3 Design

The commercial manuped product ‘The Combi Bike’ is a state of the art manuped device from Denmark-based company LEMCO, chosen for its quality construction. Besides solving the problem of limited physical space when moving in VR [3], the manuped is also to be considered a very safe biking device for VR, considering that the user is not placed on a saddle, but in a steady chair. To accommodate the requirement of the least disturbing environmental complexity, the VE created for the study was a desert. The desert VE was designed and developed using the Unity3D engine. The design of the desert VE can be seen in Fig. 2.



Fig. 2. Screenshot of the desert VE

The VE was modelled with the aid of GAIA; a plugin that allows for fast and detailed terrain generation, based on only a few guiding parameters. The terrain was generated based on height maps of sand dunes, which resulted in a natural looking landscape, when compared to sand dune reference imagery, but limited to very few visual elements, to not remove focus on the task of pedaling forward along the path.

The choice of angle for the incline of the hills was determined based on the study by Dentorn [7]. As Denton used a 15° incline but was unable to get any conclusive results. Amplifying the incline by doubling the slope to 30° , was initially chosen and later changed to 20° , based on internal and pilot testing on various angles, finding that 20° was considered the spot between a strong sensation of slope, and an angle which a

device would realistically be able to climb. To increase consistency and reliability between participant trials, as well as control of the sequencing of hills and valleys for participant to cross, the VE road was made as a straight path. This would remove steering from the VE but retain control of the speed of the buggy. While reducing the level of interactivity and freedom in the VE but maintain control and consistency between participant trials. Two hills were implemented as it makes it possible to get respectively an average RPE exertion rating and standard deviations for up- and downhill driving. This should reduce the uncertainty of participant scores [9], while also making it possible to determine if any fatigue would be created for each run through of the VE. Having several hill types should create a more stable average, but the resource demands on participants, related to a longer and more perceptually demanding session with extended periods in VR, kept the hill count to a single entry.

A virtual vehicle for the participants to occupy while in the VE was implemented (also visible in Fig. 2). The vehicle was a 4-wheel sand dune buggy; fitting the theme of the desert VE and the stability of the manuped Combi Bike.

3.1 Multimodal Feedback

Audio and haptic feedback was based on diegetic stimuli from different VE objects and materials. The auditory and haptic feedback was represented by the surface friction between the wheels of the buggy and the gravel road surface. The auditory and haptic feedback could be presented in two different states; static and dynamic (as part of the condition design). Static feedback had constant pitch, independent from the inclination when driving uphill or downhill. Dynamic auditory and haptic feedback was designed similarly between them. Feedback varied by decreasing the pitch when driving uphill and increasing it when driving downhill, inspired by the findings of Bordegoni et al. [4], where altering the frequency of an auditory feedback was found to affect the perceived level of effort. While the auditory and haptic signals were both based on audio signals, they were different from each other, serving purpose in different parts of the frequency spectrum, with haptics focusing on the lower registers (using a Butt-Kicker motor), as Bordegoni et al. [5] showed that as low frequencies were found to affect perceived effort.

3.2 Auditory Feedback and Soundscape Design

An argument brought up in [5] was how the park VE soundscape might have contained too many different noises, which might have cluttered the implemented sounds of the bike. As such keeping to a more consistent and subtle soundscape for this VE, could potentially produce better results. On a similar note; as participants in [5] did not notice of the auditory feedback, there is a possibility that the auditory feedback was *too* subtle. In this study, was therefore designed to be more pronounced.

Despite needing a subtle representation, the inclusion of soundscapes in mediated environment should not be underestimated, as pointed out by Nordahl [13] who shows how soundscapes are able to increase the sense of presence, as well as promote movement inside a VE. The soundscape for this desert VE consisted of simple and subtle wind sounds, to add contextual stimuli in the auditory domain, while remaining

in the background of the multimodal feedback. As previously mentioned, the auditory feedback was the friction sound of the buggy's wheels on the gravel desert road. All sounds were based on recorded samples. The final gravel sound from driving was created by combining several different segments of a gravel friction recording. The combination created variance to avoid an obviously repeating loop of the sample. To control the pitch, the slope of the surface in the VE was remapped, from its minimum negative slope to its maximum slope, for a range of values, which were used as a multiplier for controlling the pitch of the auditory feedback. The pitch varied between a 20% increase and decrease depending on the uphill and downhill slope, which means that when the user is driving up the slope, the pitch would be 80% of its original, and 120% driving downhill. Both static and dynamic conditions saw changes in amplitude to the auditory feedback, controlled from the velocity of the VE vehicle. This meant that when the user was moving fast with high amplitude, low when moving slowly and off when still.

3.3 Haptic Feedback

The haptic feedback was controlled through Pure Data, with a patch created and modified using the guide by Farnell [7]. The patch created random pop sounds, using an audio ramp that jumps to one, creates the pop, and randomly fades out to zero. With three of these patches as sub-patches (all linked to the same output), a sound effect was created which was reminiscent to rain. For the haptic stimuli, the 'Buttkicker Advance', a low frequency transducer, was used to translate the audio signal from Pure data to vibrations. Meanwhile, using the patch without modification created too few vibrations to properly simulate the vibration associated with driving on gravel. To solve this, the minimum delay for creating the pop sound was lowered from 30 to 5 and additional six sub-patches was added. With these modifications the impulses were more densely packed, creating vibrations similar to the gravel sensation intended for the road surface.

The haptic feedback was controlled using a separate audio signal created in Pure Data, as Unity cannot split its auditory output between several audio devices. The dynamic values (slope and velocity) obtained from Unity were sent over the Open Sound Control (OSC) protocol to a Pure Data patch.

3.4 VE Movement Speed

The speed of the buggy inside the desert VE, was controlled by the of the Combi Bike pedal arm revolutions, obtained from a custom-built Arduino based GIRO microcontroller [14] strapped to the pedal arm of the Combi Bike (see Fig. 3). The GIRO is made to tracks rotation activity, which it streams to Unity through a UDP network. A Unity script receives the stream and estimates the driving speed inside the VE, based on the values received from the GIRO, and uses the values to drive the user forward at a corresponding speed, with no noticeable delay. For movement speed, gravitational force was deliberately not included in the VE, as it would require the users to work physically harder driving uphill, and less so driving downhill. This would complicate the results, due to its effects on the Borg scale responses from participants [8].



Fig. 3. The GIRO microcontroller [14] strapped to the pedal arm of the Combi Bike

4 Methods

To determine whether dynamic feedback has an effect on perceived exertion three different hypotheses were created, based on our RQ:

- H1: There will be a significant difference in users' perception of exertion between up- and downhill movement by introducing dynamic feedback in a VR exercise setting as opposed to static feedback.
- H2: There will be a significant difference in users' perception of exertion when driving uphill by introducing dynamic feedback in a VR exercise setting as opposed to static feedback.
- H3: There will be a significant difference in users' perception of exertion when driving downhill by introducing dynamic sensory modalities in a VR exercise setting as opposed to static feedback.

4.1 Pilot Test

A pilot test on the VE interaction design was conducted early, with four participants, to determine if there were any adverse health effects. As previously mentioned, results showed that a 30° incline and decline was too steep, combined with forcing the perspective of the camera to follow the rotation of the vehicle, as it caused several instances of strong cybersickness with users. Another aspect effect on cybersickness from the speed of forward movement inside the VE (relative to the pedal speed). The solution was to separate the Unity camera from the vehicle rotation and lower the speed. Combined with 20° hill-angles, cybersickness was notably reduced, and induced a satisfying indication of driving up- or downhill (as it was steep enough to force the user to physically look upwards and downwards to follow the direction of buggy movement).

4.2 Experiment Design

This study used a within-group design, in order to deduce whether different levels of immersion could affect a user's perceived level of exertion in a virtual reality exercise setting. Using a within-group design means that certain factors have to be considered when designing the experiment. By using the participants more than once creates the risk of carry-over effects, such as fatigue and boredom due to the longer test duration for each participant [9]. Additionally, a factor to be aware of when working in VR, is adverse health effects related to cybersickness [11]. Cybersickness has an immense impact on presence [10], which would most likely affect the results of the experiment. In order to minimize the carry-over effects a counterbalanced list of conditions was used. As the experiment had four conditions this created 24 unique combination of conditions [9]. This means that a minimum of 24 participants or a multiple of 24, i.e. 48, 72, were required. The independent variables of the experiment were auditory and haptic feedback. Both had two levels, static and dynamic. The dependent variable was perceived exertion, measured using the Borg Scale [6]. This created four conditions, see Table 1 for an overview of combinations of auditory and haptic levels.

Table 1. Experimental conditions

Condition A	Condition B	Condition C	Condition D
Static Audio & Static Haptic	Dynamic Audio, Static Haptic	Static Audio, Dynamic Haptic	Dynamic Audio & Dynamic Haptic

The four experimental conditions. Containing the different combinations of audio and haptic feedback.

Demographics on test participants (age, gender and previous VR experiences including experiences with adverse health effects) was collected using a questionnaire. In addition, the SUS-questionnaire [15] was used as a post-questionnaire after each condition, to compare the participants' level of presence for each condition. There exists different versions of the SUS-questionnaire. The one used in this study is the 3-item version [15], where ratings can be given on a Likert scale from 1–7, where 6–7 are indicators of presence. The Borg scale of perceived exertion was applied throughout the test. The scale goes from 6–20, where 6 is resting and 20 is the maximal intensity which can be kept for a very short time [8]. The Borg scale can be seen in Appendix 1. As the final part of the experiment, a short semi-structured interview was executed to determine if participants had noticed any differences between the four conditions and if they had experience any adverse health effects.

4.3 System Setup

The physical setup can be seen in Fig. 4. A Combi Bike with the GIRO strapped to the handles, was connected to a high-end PC, running the Unity build of the desert VE.



Fig. 4. Schematic of the VR setup

The PC's specifications were as follows:

- NVIDIA GeForce GTX 1080
- Intel Core i7-7700K, 4.2 GHz
- 16 GB VRAM.

The high-level performance ensured the VE running at a stable 90 fps (frames per second), for the best, and least fatiguing and cybersickness-provoking VR experience [16].

Connected to the PC was an Oculus CV1 headset, and a pair of noise reducing headphones, and a ButtKicker attached to a wooden pallet. Through the pallet, the vibrations could reach the user, from the chair, that was placed on the pallet. A carpet was placed under the pallet to reduce noise generated by the ButtKicker and pallet during instances of strong vibrations. In addition to the PC, demographics data, Borg scale ratings from participants, notes from the trials, debriefings, etc. were typed by an observer on a laptop in the background.

4.4 Procedure

24 participants performed the experiment (6 female). All of the participants were selected through non-probability sampling. Age range was 22–31, average age of 24.75 (SD = 2.56). All of the participants had prior experience with VR systems. 11 participants had never become sick from VR before, whereas the remaining 13 had felt varying degrees of sickness in prior VR experiences. Participants were asked to read and sign a consent form, informing of the risk of cybersickness, and seeking permission to record their performance on video. Afterwards, participant would be introduced to the test itself, and how it would be conducted. They were introduced to the Borg Scale, and tested in their understanding of it, to prevent misinterpretations of the values. The participant would go on to fill the demographic items, whether they had prior experience with VR, and any previous occurrences with motion sickness. The session

procedure was explained to the participant, who was first introduced to a training environment (introducing how to use the RPE scale and introduced to the SUS questionnaire). Participants went through the four different conditions after having completed the training environment. Upon main test start, RPE ratings were requested after each incline/decline, with presence questions asked at the end of each condition. The conditions sequence was controlled using a counterbalancing, to minimize carry-over effects. Each participant drove across 2 hills in each condition, measuring their perceived exertion four times, once at the end of each incline and decline. Therefore, the participant would have given 16 Borg measurements and 4 presence measurements by the end of the experiment. As a participant was asked to say the Borg score out loud, the rating was noted by the experimenter. Presence measurements were filled by the participants, which required them to remove the HMD between conditions. After the test there was a debriefing session, asking the participant if they experienced any adverse health effects and if they noticed any changes between conditions.

The data collected was interval data as the Borg Scale uses a numbered scale similar to the Likert scale, which collects interval data. As the experiment conducted was of a 2^2 factorial design, the rule of sphericity is not valid, as it only applies to studies that have factors with more than two levels. The normality of the data must be determined for each condition, which was done using a Shapiro-Wilk test. To determine the existence of statistical differences in perceived exertion, the data was analyzed using a two-way repeated measures ANOVA.

5 Results

The results were analyzed using SPSS 24 and charts were made using Microsoft Excel 2016. Before analyzing the results, the final requirement for parametric data must be evaluated. In terms of parametric statistics, a Shapiro-Wilk test was performed on the data belonging to each of the four conditions. The result for condition A was $D(24) = 0.93$, $p = 0.09$, for condition B the result was $D(24) = 0.93$, $p = 0.11$, for condition C $D(24) = 0.93$, $p = 0.85$ and finally the result for condition D was $D(24) = 0.968$, $p = 0.62$. This means that all of the data is normally distributed, as none of the results were statistically significant and parametric data analysis can be performed on the experiment data.

5.1 Perceived Exertion for Uphill and Downhill Driving

For uphill driving, results of the two-way repeated measures ANOVA, for the data measured in relation to driving uphill in VR, shows that there is no main effect of audio $F(1, 23) = 0.11$, $p = 0.75$, $r = 0.01$. This is also the case with the haptic feedback as there was no difference between the two levels $F(1, 23) = 0.3$, $p = 0.59$, $r = 0.01$. Average scores from the uphill Borg scores can be seen in Fig. 5. There was not found a significant difference for the interaction between audio and haptic feedback when driving uphill $F(1, 23) = 0.0$, $p = 1.0$, $r < 0.01$.

For downhill driving, results, also analyzed using ANOVA, of driving downhill shows that there is no main effect of audio $F(1, 23) = 0.18, p = 0.68, r = 0.01$. Again, there was no main effect of haptic feedback $F(1, 23) = 0.278, p = 0.6, r = 0.01$. The interaction effect between audio and haptic also proved to be non-significant $F(1, 23) = 0.3, p = 0.59, r = 0.01$. The average scores from the downhill Borg scores can be seen in Fig. 6.

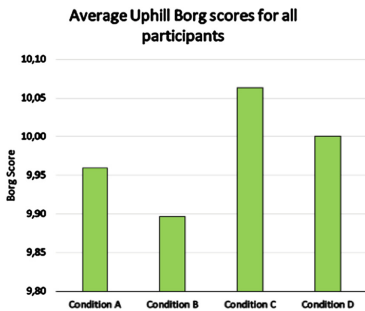


Fig. 5. Averages for each condition, based on participants’ RPE ratings for uphill driving (scale ranges from 6–20). Results are very similar between conditions.

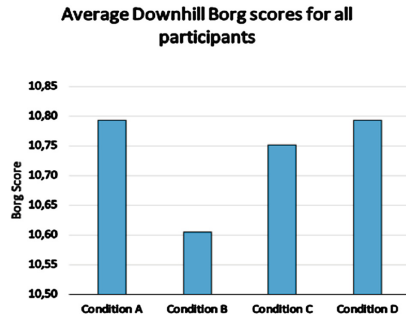


Fig. 6. An average for each condition, based on all participants’ Borg scores for downhill driving (scale ranges from 6 to 20). Results are very similar between conditions.

5.2 Difference Between Uphill and Downhill Driving

To find the difference between the uphill driving data and the downhill data, each participant had their uphill score subtracted from their downhill score, see Fig. 7 for an average score for all participants.

Using these values, the final hypothesis, which sought to determine if there existed a significant difference in perceived exertion between the up- and downhill driving in VR, was analyzed using ANOVA. Just as there were no significant difference for the other two hypotheses no main effect was found between haptic and dynamic audio $F(1, 23) = 0.004, p = 0.95, r < 0.01$. The same applied for the haptic feedback $F(1, 23) = 0.4, p = 0.838, r < 0.01$, this was also the case for the interaction effect $F(1, 23) = 0.624, p = 0.438, r = 0.03$.

5.3 SUS Questionnaire Presence Scores

From the SUS presence questionnaire items, a score of 6 or 7 indicates a sensation of presence. The overall presence scores from participants were very consistent, between respectively the 3 items per condition per user, and also between conditions. 3 of 24 participant scores suggested a sense of presence while driving the buggy in VR (Table 2).

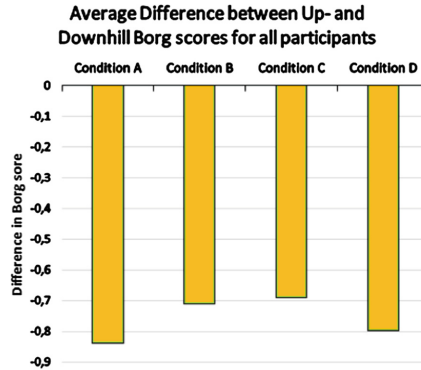


Fig. 7. Average difference between uphill and downhill Borg scores for all participants. Negative scores mean that downhill measurements were higher than uphill measurements.

While only 3 participants felt a strong sense of presence, the score range between participants differs a lot, with most conditions. As such, there are no conditions which induced a higher sense of presence than others, as presence scores for all individual participants scores were very consistent between conditions, and only deviated 0.5 points on average. No one condition was remarkably different from another, in terms of the degree of presence it induced with participants.

Table 2. SUS presence questionnaire results

Condition A	Condition B	Condition C	Condition D
Presence score 4.3	Presence score 4.4	Presence score 4.6	Presence score 4.5
Standard dev. 1.5	Standard dev. 1.4	Standard dev. 2.6	Standard dev. 2.6

6 Discussion

The quantitative results do not illustrate any significant effect of the multimodal stimuli, on participants' perception of exertion. Meanwhile, the somewhat extreme swings in reported presence between participants is curious, despite getting close to *exactly* the same experience, and indicate that there are variables to the IVR biking experiment that cannot be derived simply from the RPE responses. We believe that the large difference in presence scoring can relate to several aspects of the VE design and overall experiment method. As previously stated, the way users relate to VR through their real-life experiences affect their experience in VR, for better or worse. In this case, features from real world physics, as perceived and expected in our VE, did not live up to users' expectations from own experience. Another problem might have been that the experiment asked participants to rate polar opposite mental constructs, in the form of going uphill (more effort) and downhill (less effort), with sounds that have previously suggested a heavier (more effort) and lighter load (less effort). Mixing these in a bag,

combining them, and using the combinations to test an illusion (i.e. something that does not exist to begin with) was most likely not a very good proposition. In some senses, the fact that no real differences were felt, might have been due to confusion between certain mixes of less/more effort effects, that users simply became unsusceptible to the perceptual manipulation.

6.1 Participant Expectations and Polar Opposites

In the debriefing, the majority of participants stated that they experienced the VE differently from their expectations to driving up- and downhill. 14 of the participants directly mentioned that driving downhill felt harder than driving straight or uphill. The rationale was that they expected the buggy to keep rolling, when driving downhill, even if they stopped pedaling. Not meeting this expectation, participants felt they had to work harder to get down the hills. Missing indeed from the VE design is gravity, as the buggy behavior on the hills was not affected by a gravitational pull. Leaving out this element could very well be something which affected the gathered data. In Fig. 7, all the mean values are negative, meaning that downhill driving scored higher on the Borg scale, compared to uphill, which supports participants recollections that driving downhill felt harder, compared to uphill driving. Meanwhile, including gravity in the experiment would change the actual exertion needed to drive up- and downhill. Meaning that measuring perceived exertion would no longer make sense, if used in a similar fashion as in this experiment.

In hindsight, this might generally suggest that there are many complexities to exertion perception studies, which need to be taken into consideration for experiment design. However, perhaps two elements or behaviors of a (real or virtual) world, so fundamentally opposite each other (e.g. forces of pull/push, up/down, etc.) should not be measured against each other, when trying to understand illusions and perceptual manipulations. But simply isolated to gradients within their own domain. At least when experimenting with something that could be described as a perceptual and cognitive construct, more than a real thing. And from a design perspective; placing the same manipulation method on two things act opposite to each other, might also be quite dangerous for the success of the effect.

This in part, could be due to the previously mentioned perspective of expectations within certain environmental or physical conditions, which have been confirmed in this study to apply in VR and in relation to perceived exertion, which we will discuss in a moment. But also, from a test design method perspective, it also has to do with changing too many feedback-oriented aspects, when operating in a within-participant design. When testing multimodal feedback to certain conditions, especially a delicate manipulation as the illusion of exertion has shown to be, having two conditions which could be considered polar opposites to each other (going uphill and going downhill) could be a bad choice for the methodology. And as mentioned previously, coupling it with feedback stimuli which also has suggestive properties to opposite exertion effects, the complication might be too high for the delicacy of the desired effect.

6.2 Make It Plausible, Somehow

When suggesting the existence of non-existent physical changes to the real world, the argument needs to be persuasive. In some sense, this might also have to do with what you are implicitly telling the participants to be a plausible part of the scenario. In the previous (bike) study from Bruun-Pedersen et al. [5], the physical attributes of the path (e.g. height) did not change. Meanwhile, a part of the test procedure was a little mind-game, where the researcher turned the bike-resistance knob between each trial, to *suggest* that a perceived change in resistance was a hypothetical possibility, and thereby suggesting that changes in perceived exertion *could* in fact be plausible in the real world. Meanwhile, doing so makes the study manipulate not only the feedback, but prime the user to change his/her perspective and expectations to the world they are entering. This perspective can be both constructive or destructive to the validity of future studies, but is a point that it seems to be hard to neglect.

6.3 Speed of the Virtual Buggy

The velocity of the virtual buggy also mentioned by participants during debriefing. Some participants felt that the speed of the virtual buggy did not correlate well with the force put into the Combi Bike, and felt that the buggy went slower than expected. Meanwhile, high velocity was responsible for some of the cybersickness that pilot testers felt, even if most was related to camera movement. In the final experiment, 7 participants reported affects from cybersickness, meanwhile just slightly (reasons being mixed, for instance whenever a participant were forced to break to give a Borg scale rating, stated to feel unnatural with no force transfer).

6.4 Various Risks from the Within-Group Design

One of the general risks to always consider is the carry-over effect of VR is currently impossible to neglect when working a within-participant design, as it can bias results between conditions. This study used counterbalancing was used to try and counteract or reduce the potential bias created from such an effect. Our participants had to go through several conditions to reach the end of the experiment. Another is the similarity between conditions, as using of the same VE have likely bored some the participants, making them less engaged in the experiment, even if the desert VE environment was selected to avoid other issues. Fatigue played a larger role than anticipated. Since there were two hills for each condition each participant has eight Borg scores for driving up-and downhill. Comparing the difference in Borg score between driving up the first and second hill, in each participant's first condition, shows that the second hill ($M = 10.17$) scores higher on average compared to the first hill ($M = 8.8$), see Fig. 8.

Since participants are assigned to conditions using counterbalancing the difference in scores for the two hills indicates that the difference is not created by using dynamic and static feedback. The difference most likely stems from fatigue.

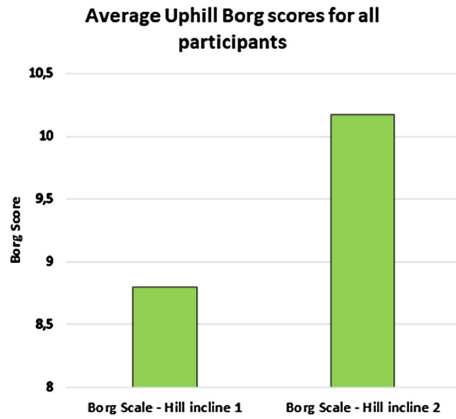


Fig. 8. Average score of the 1 and 2 encountered hill, based on all participants

6.5 Presence

The range of presence scores given to a single item, for a single condition was very different between individual participants, as seen in both Table 3 and Fig. 9.

Table 3. SUS presence questionnaire – score range

Condition A	Condition B	Condition C	Condition D
Lowest average 2.0	Lowest average 1.7	Lowest average 2.0	Lowest average 2.0
Highest average 6.7	Highest average 6.0	Highest average 6.3	Highest average 6.7

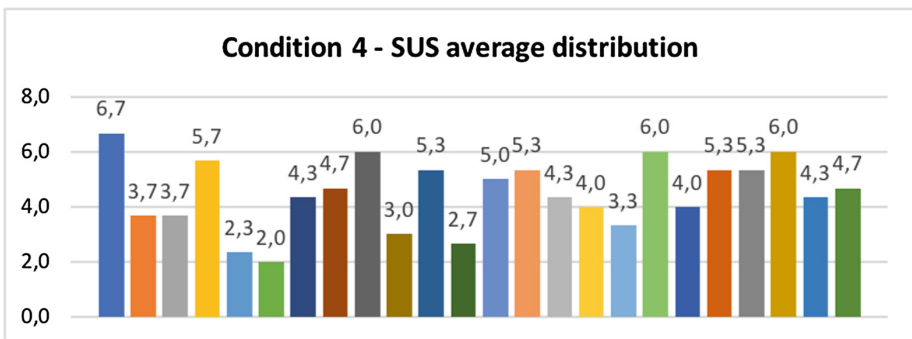


Fig. 9. Averages of the 3 SUS items for each 24 participants, for condition 4. The range is considerable, from almost no sensation - to a powerful sensation of presence.

Over the course of the experiment, we knew there would be several instances where the Oculus headset would have to be removed, resulting in an immersive change, and also risk of this affecting the sensation of presence. For every RPE rating, the participant would be made aware of the test conductor. Between every condition, the participant would also have to remove the Oculus headset to respond to questions. As such, participants were both reminded of the non-virtual world while in VR, and fully switching back and forth. Going forward, all user input should be incorporated into the virtual reality, so that participants would not need interaction with the real world at any point during the experiment (so to speak). Another presence related issue was how in the VE, the buggy had a steering wheel, while in the real world the handles on the Combi Bike were moving hand pedals. As a result, the relationship between VR and real-world interaction was missing, causing a disturbance for the participants, based on the lack of correlation.

6.6 The Individual Interpretation of the Borg Scale

The interpretations of the Borg scale have like varied between participants, which could be a reason for how and why the ratings between participants varied in some circumstances. Meanwhile, the advantage of the within-group design is that it is each participant's own variance that is compared between participants. This eliminates the bias of participants starting with different scores. There is no way to guarantee that the same increase in exertion is scored with the same variance between participants. To give the participants the same interpretation of the scale it would require some method for determining how much more exertion is needed based on an initial score, e.g. this is twice as hard as a score of 5 therefore the new score is 10. Participants would then have a starting point from where they would be able to give more similar changes in score, for similar changes in perceived exertion. The Borg scale does contain definitions for the individual ranges, see appendix 1, but these are rather ambiguous and can be interpreted. Meanwhile, adding some concrete examples could aid to limit the uncertainty.

7 Conclusion

This paper presented an experiment, aiming to investigate the possibility of altering the perceived exertion, using auditory and haptic feedback, during a biking session in VR. The study found no significant difference in any of the presented hypotheses. An interesting discovery was made from the post experiment feedback, as 14 participants mentioned, that the missing effect of gravity made driving downhill feel more difficult compared to driving uphill. This and the almost extremely differentiating presence score, has led to reflections on the experiment methodology, which should lead to improved study designs in the future. This relates to understanding the approach necessary when operating a delicate effect, such as the manipulation of perceived exertion, despite the fact that it is trying to invoke a perceptual construct in the real world, of something that (at best) only exists in the virtual. VR is still very young as a technology, and despite not producing significant differences between conditions, it is important to take notice of how to connect the dots properly to achieve that difference in future studies.

References

1. Bruun-Pedersen, J.R., Serafin, S., Kofoed, L.B.: Motivating elderly to exercise – recreational virtual environment for indoor biking. In: 4th International Conference on Serious Games and Applications for Health, Orlando (2016)
2. Bruun-Pedersen, J.R., Pedersen, K.S., Serafin, S., Kofoed, L.B.: Augmented exercise biking with virtual environments for elderly users - a preliminary study for retirement home physical therapy. In: Proceedings of IEEE VR 2014 - Workshop on Virtual and Augmented Assistive Technology, VAAT, Minneapolis (2014)
3. Bruun-Pedersen, J.R., Serafin, S., Kofoed, L.B.: Restorative virtual environment design for augmenting nursing home rehabilitation. *J. Virtual Worlds Res.* **9**(3), 1–24 (2016)
4. Bordegoni, M., Ferrise, F., Grani, F., Bruun-Pedersen, J.R., Serafin, S.: Auditory feedback affects perception of effort when exercising with a Pulley machine. In: Online Proceedings from the International Conference on Multisensory Motor Behavior: Impact of Sound (2013)
5. Bruun-Pedersen, J.R., Grani, F., Serafin, S.: Investigating the role of auditory feedback in a multimodal biking experience. In: 13th International Symposium on Computer Music Multidisciplinary Research, CMMR (2017)
6. Plotnik, M., Bahat, Y., Akiva, Y.A., Gimmon, Y., Inzelberg, R., Kizony, R.: Can we climb with our eyes? Preliminary report on the effect of conflicting virtual scenery on leveled and inclined gait. In: Virtual Rehabilitation, ICVR (2013)
7. Denton, W.C.: Comparison of uphill walking and visual uphill environments on energy expenditure and locomotor-respiratory coupling. Doctoral dissertation, University of Nebraska at Omaha (2017)
8. Borg, G., Linderholm, H.: Perceived exertion and pulse rate during graded exercise in various age groups. *J. Intern. Med.* **181**(S472), 194–206 (1967)
9. Slater, M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. Roy. Soc. Lond. B: Biol. Sci.* **364**, 3549–3557 (2009)
10. Nilsson, N.C., Nordahl, R., Serafin, S.: Immersion revisited: a review of existing definitions of immersion and their relation to different theories of presence. *Hum. Technol.* **12** (2016)
11. Rovira, A., Swapp, D., Spanlang, B., Slater, M.: The use of virtual reality in the study of people’s responses to violent incidents. *Front. Behav. Neurosci.* **3**, 59 (2009)
12. Nordahl, R., Serafin, S., Turchet, L., Nilsson, N.C.: A multimodal architecture for simulating natural interactive walking in virtual environments. *PsychNol. J.* **9**(3) (2011)
13. Nordahl, R.: Sonic interaction design to enhance presence and motion in virtual environments. In: Proceedings of CHI 2008 Workshop on Sonic Interaction Design (2008)
14. Grani, F., Bruun-Pedersen, J.R.: Giro: better biking in virtual reality. In: IEEE 3rd Workshop on Everyday Virtual Reality, WEVR (2017)
15. Usoh, M., Catena, E., Arman, S., Slater, M.: Using presence questionnaires in reality. *Presence: Teleoper. Virtual Environ.* **9**(5), 497–503 (2000)
16. Bruun-Pedersen, J.R., Serafin, S., Kofoed, L.: Going outside while staying inside - exercise motivation with immersive vs. non-immersive recreational virtual environment augmentation for older adult nursing home residents. In: IEEE International Conference on Healthcare Informatics, ICHI 2016 (2016)
17. Bruun-Pedersen, J.R., Serafin, S., Kofoed, L.B.: Simulating nature for elderly users - a design approach for recreational virtual environments. In: Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (2015)