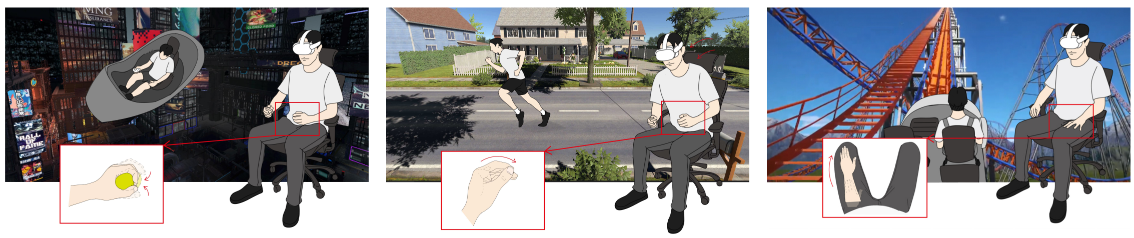
# Continuous Measurement Methods for Transient Physiological Discomfort in VR Locomotion

ANONYMOUS AUTHOR(S)



(a) 'Squeezing Ball' in active locomotion (b) 'Sliding Thumb' in active locomotion (c) 'Rubbing Thigh' in passive locomotion

Fig. 1. The illustrations of 'squeezing ball', 'sliding thumb', and 'rubbing thigh' serve as the most preferred continuous measurement methods for transient physiological discomfort that rapidly changes with VR locomotion variations

Motion sickness, in addition to its persistent long-term effects, also exhibits short-term effects characterized as transient physiological discomfort, which changes rapidly with variations in locomotion. However, such discomforts are challenging to assess using current subjective scales and objective physiological measurements. To tackle this issue, this paper suggests continuous measurement methods designed specifically for evaluating transient physiological discomfort during VR locomotion. Through a user-elicitation study, three preferred measurement methods—'squeezing ball', 'sliding thumb', and 'rubbing thigh'—were identified. These techniques were then evaluated for reliability, validity, attention, presence, and workload, with 'sliding thumb' identified as the most effective option. The paper expands traditional measurement methods to capture users' physiological experiences in VR interactions, offering practical choices for researchers in this field along with an in-depth discussion of design considerations, detailed implementation guidelines, and potential ways to optimize the VR experiences utilizing the measurement data.

CCS Concepts: • **Human-centered computing** → *HCI design and evaluation methods*; **Virtual reality**; User interface design.

Additional Key Words and Phrases: Continuous measurement, Motion sickness, VR locomotion, Sensory conflict

**ACM Reference Format:**

Anonymous Author(s). 2018. Continuous Measurement Methods for Transient Physiological Discomfort in VR Locomotion. In *XXX*. ACM, New York, NY, USA, 26 pages. https://doi.org/XXXXXXX.XXXXXXX

## 1 INTRODUCTION

In real or virtual locomotion environments, humans often experience varying degrees of physiological discomfort due to visual-vestibular conflicts, commonly referred to as motion sickness [51]. The long-term effects of motion sickness are well known due to their severity and significance, and they have been extensively studied [46]. These effects are attributed to the brain's accumulation of sensory conflicts, manifesting as persistent physiological discomfort, such as dizziness, nausea, and eye strain [13]. Motion sickness has also been found to have short-term effects, which reflect



Manuscript submitted to ACM

the instantaneous level of sensory mismatch and manifest as transient physiological discomfort. These symptoms are similar to those of long-term effects but fluctuate at a completely different rate [23]. Unlike the persistent physiological discomfort characterized by long-term effects, transient physiological discomfort occurs immediately when there is an abrupt mismatch between visual and vestibular cues. This often happens to sensitive users in real-world transportation or during VR locomotion events such as the onset of acceleration/deceleration, turning, or sudden stops [7, 49, 52, 70]. However, as speed stabilizes or returns to a stationary state, this discomfort quickly dissipates. In other words, the persistent physiological discomfort induced by the long-term effects continues even after the locomotion stimulus ceases and does not fluctuate immediately with sub-motion changes (e.g., translation, rotation, or stopping). In contrast, the transient physiological discomfort caused by short-term effects rises and falls sharply in response to sub-motion variations [23, 43, 49]. However, research on this phenomenon remains limited due to two major challenges in measuring transient physiological discomfort.

On the one hand, transient physiological discomfort fluctuates rapidly, making traditional assessment methods—such as the Simulator Sickness Questionnaire (SSQ) and the Fast Motion Sickness Scale (FMS) [37, 38]—insufficient in temporal resolution, as they rely on summarizing discomfort experiences over a period of time. On the other hand, this discomfort is highly subjective and involves complex physiological responses related to sensorimotor conflicts (e.g., eye strain, dizziness, nausea). This makes it difficult for objective physiological measurement techniques—such as electroencephalography (EEG), electrocardiography, and eye tracking—to accurately and in real-time associate low-level physiological signals with high-level subjective perceptions. Continuous measurement methods based on users' dynamic self-ratings offer a promising solution to these challenges. The self-report method performed continuously through touchscreens/keyboards/sliders/knobs, has been applied to dynamically assess perceived musical tension, aesthetic responses to music, driver workload, course difficulty, etc [28]. However, applying these to VR locomotion faces two challenges: First, users need to focus on view motion and the discomfort feelings caused by visual-vestibular conflict, providing timely feedback. Therefore, the measurement methods should offer moderate sensory stimulation to avoid increasing the sensory and attentional load. They should also refrain from introducing visual sliders, head actions, or other techniques that might interfere with visual and vestibular perception. Second, changes in discomfort levels could impair the accurate proprioception perception of feedback actions and the timely scoring. Thus, the measurement methods should leverage instinctive actions during discomfort and ensure the timeliness and accuracy of feedback actions.

Measurement methods need to promptly capture transient physiological discomfort, with reliability and validity similar to well-validated scales like FMS [38], while addressing challenges. To achieve this, we initially conducted a user-elicitation study to identify preferred continuous methods for transient discomfort induced by various sub-locomotion. Participants designed gestures/actions that could express discomfort naturally, quickly, and accurately with low load, without affecting visual and vestibular sensations. A total of 121 continuous measurement methods were proposed. Based on feedback, methods combining proprioceptive and tactile feedback, namely 'squeezing ball', 'sliding thumb', and 'rubbing thigh', were found to be the most suitable and preferred.

In the second study, we developed three continuous measurement methods: 'squeezing ball', 'sliding thumb', and 'rubbing thigh', and evaluated their reliability, validity, presence, load, and attention through active and passive locomotion tasks. Results show that 'sliding thumb' outperforms the others, with excellent reliability, strong validity, few distractions, high presence, and low workload. This instinctive self-touch action relies on finger flexibility, small actions, moderate and clear tactile stimulation, and unique tactile changes at the folds of the index finger joints, making it easy to perceive and control with minimal distraction. To our knowledge, this study is the first to design continuous measurement methods for transient physiological discomfort in VR locomotion, focusing on reliability, validity, and natural user experience. Designing such methods offers significant research values: 1) Identifying locomotion situations worthy of attention/avoidance/optimization provides intuitive references for designers to iteratively improve locomotion paths and scenes. 2) It helps understand and reveal the mechanisms and influencing factors of short-term effects of motion sickness, providing attribution references for researchers studying the onset time and intensity of long-term effects of motion sickness. 3) It provides the basis for triggering personalized dynamic interventions and preventive measures, preventing further exacerbation of physiological discomfort. Overall, this paper makes the following three contributions:

i) A set of continuous measurement methods was constructed through a user-elicitation experiment, identifying the characteristics and preferences of users' dynamic ratings on transient physiological discomfort.

ii) The top 3 methods elicited by users, namely 'squeezing ball', 'sliding thumb', and 'rubbing thigh', were comprehensively evaluated, analyzing their performance strengths and weaknesses.

iii) The features of continuous measurement data are analyzed, and design guidelines and uses for continuous measurement methods targeting transient physiological discomfort are proposed, providing valuable references for researchers and designers.

## 2 RELATED WORK

Our work is based on the mechanism of sensory conflicts leading to motion sickness, and is closely linked to input technology for dynamic scoring and VR locomotion.

### 2.1 Sensory Conflict and Motion Sickness

Neuroscience theories suggest that during locomotion, the brain integrates sensory inputs through dynamic multisensory weighting [6, 19]. Sensory conflict, an adaptive mechanism evolved to protect humans, triggers motion sickness when visual and vestibular systems perceive conflicting spatial information, signaling potential danger [56, 57].

Motion sickness typically manifests as a range of physiological discomfort symptoms affecting the head, stomach, and eyes [36]. Based on the rate of change and persistence, these effects can be categorized into long-term and short-term [23]. The long-term effects, which have been extensively studied, involve gradual and persistent discomfort resulting from the accumulation of sensory conflicts over time [62]. These symptoms intensify progressively and can persist even after the sensory conflict has ended, lasting beyond the period of motion exposure [39, 43].

Short-term effects arise from momentary sensory mismatches and manifest as transient discomfort [23]. Some studies used a series sub-locomotion (e.g., speed changes, translation/rotation switches, each lasting a few seconds) to induce sensory conflicts, and measured discomfort during and after the sub-locomotion via self-reports [23, 41, 49, 70], which showed discomfort fluctuating sharply, rising during sub-locomotion changes and quickly reducing after stimulus cessation, and its peak discomfort was stronger than the persistent discomfort after the motion ended. This finer temporal measurement data can capture the short-term effects of motion sickness, high data points represent that the current sub-locomotion significantly disrupts the user experience [23]. Currently, the Fast Motion Sickness Scale (FMS) measures temporal variations by asking users about discomfort at intervals [38]. However, frequent intervals disrupt VR experiences and introduce recall issues and cognitive biases (e.g., the "peak-end rule").

Existing research has largely focused on the long-term effects of motion sickness, while studies on its short-term effects remain scarce. In particular, measurement methods for the transient physiological discomfort associated with short-term effects are still lacking, which has motivated our research.

### 2.2 Continuous Measurement Methods

Users' subjective experiences arise from the brain's sensory-perceptive-cognitive processing mechanisms. Active self-feedback methods are seen to offer direct subjective experiences, especially encompassing various perception and cognitive outcomes [3].

Numerous methods for capturing continuous subjective user experiences have been proposed in fields like psychology, education, and music, enabling moment-to-moment monitoring [32, 59]. Participants typically use digital interfaces (e.g., keyboards, mice, touchscreens, or controllers) to provide continuous and real-time recording of subjective experiences [29], such as changes in emotions, sensations, or cognitive states [25, 67]. Compared to traditional scales, continuous measurement methods offer more detailed and precise data, providing deeper insights into participants' subjective experiences [16]. Recent methods include horizontal sliders, circular dial interfaces, voice expressions, mobile phone controls, face-to-face dials, etc [24, 28]. For example, physical sliders have been used to rate lecture difficulty and TV image quality in real-time [61, 63], knobs to assess motion mismatches in driving simulations [15], and handheld controllers to evaluate passengers' perceived tension, pressure, comfort, and trust in various driving situations [32]. However, they were not specifically designed or evaluated for measuring physiological or psychological states.

Compared to these studies, our study aims to design and evaluate suited continuous measurement methods for novel application scenarios, and solve unique challenges. On the one hand, VR locomotion faces significant sensory conflict challenges, requiring users to naturally, accurately, and quickly provide subjective feedback during physiological discomfort changes. On the other hand, VR locomotion imposes stricter demands on measurement methods, such as not interfering with visual and vestibular sensations, having less physical load and distraction, and well-coordinated with navigation methods in active locomotion.

### 2.3 VR Locomotion

VR locomotion often involves sensory conflicts that disrupt spatial awareness, enjoyment, self-motion perception, and presence, leading to transient and persistent physiological discomfort [9, 47, 58] which is closely tied to the locomotion posture, type, and method. VR can involve various postures—standing, sitting, reclining, or lying—each affecting sensory conflicts during locomotion. Generally, standing locomotion (e.g., walking) tends to create minimal sensory conflicts as it mimics daily movements and offers natural multisensory stimulation. However, maintaining balance and supporting the body can lead to fatigue issues [73]. Sitting compensates for this but leads to stronger sensory conflicts and weaker self-motion perception due to limited body movement [10]. Recently, reclining or lying postures have also been found to induce significant sensory conflicts due to the need for upright redirection [47, 48].

VR locomotion is typically categorized into active and passive types. In active locomotion, users control locomotion through actual/redirected walking, artificial or embodiment methods [8, 45]. Common artificial methods include handheld controllers and teleportation. Controller-based methods are considered effortless and provide a driving metaphor but easily induce strong sensory conflicts, affecting physiological discomfort [2, 14, 42]. Teleportation reduces sensory conflicts by instantly changing the user's position, but sacrifices immersion and self-motion perception [9]. Embodiment methods, such as head joystick [33], swimming [12], and stationary walking [44], offer more natural proprioceptive and vestibular stimulation [20, 21]. Passive locomotion involves less exploration and often includes watching movies/360-degree videos, following predefined routes, or using VR in vehicles [30, 35], which can cause more severe discomfort [17, 46] due to sensory conflicts and unpredictable movement.

These studies guide the selection of postures and methods for VR locomotion in this work. Considering both common usage scenarios and the significant sensory conflict challenges, we selected a seated posture and a handheld controller as the experimental setup. Additionally, while continuous measurement and embodied locomotion methods may involve similar gestures or actions, they are designed for different target contexts, objectives, and challenges, making them unsuitable for direct application to each other's domains.

## 3 STUDY 1: USER-ELICITED CONTINUOUS MEASUREMENT METHODS

This study aims to elicit a set of continuous measurement methods suitable to transient physiological discomfort during VR active and passive locomotion, and explore user preferences.

### 3.1 Ethics and Participants

The experiment obtained ethical approval from our university. To ensure safety and comfort, participants with vestibular disorders or severe motion sickness were excluded. Additionally, users who exhibited insensitivity to vestibular cues and had minimal experience of discomfort induced by sensory conflicts in their daily lives were also excluded. Participants were screened using the Motion Sickness Susceptibility Questionnaire (MSSQ [27]), with those scoring between 5 and 19 considered eligible based on the distribution of scores in the scale paper (25th percentile was 5, 75th percentile was 19). Although the measurement method is designed for all people, participants with professional HCI/design backgrounds are better equipped to generate insights and evaluate the pros and cons of the method to address the design challenges. Therefore, we recruited 24 students from 3 local universities with backgrounds in HCI, interaction design, and related fields (12 male and 12 female, average age = 23.71, SD = 2.44). Among them, 9 participants had high proficiency in VR (weekly usage frequency × number of weeks > 50 times), 7 had some proficiency (50 times > weekly usage frequency × number of weeks > 25 times), and 8 had low proficiency with VR (weekly usage frequency × number of weeks < 25 times).

### 3.2 Configurations

To cover mainstream VR locomotion types, we implemented two conditions: active and passive locomotion, each requiring separate continuous measurement methods. Active locomotion has stricter low-load requirements and limited limb availability, as it involves navigating with one hand while maintaining attention. Exploring methods for both conditions also broadens design possibilities, revealing more practical continuous measurement options.

The main VR interaction postures are standing and sitting. While standing allows natural locomotion with consistent visual and vestibular signals, sitting induces greater sensory conflicts, as the vestibular sense remains stationary while vision perceives motion. Therefore, we chose the sitting posture, where measurement needs are more urgent, as the basic experimental setup. Similarly, since controller-based locomotion presents more sensory conflicts than in-place locomotion by embodied actions while sitting [21, 33], we used a right-handheld controller for active locomotion. Participants utilized the joystick and buttons on the controller to translate, and to rotate while holding the 'grip' button (Figure 2a). Translational acceleration and maximum speed were set to 1.6 m/s and 8 m/s respectively, while rotational acceleration and maximum speed were set to 6 °/s and 30 °/s respectively, according to other studies [34, 69].

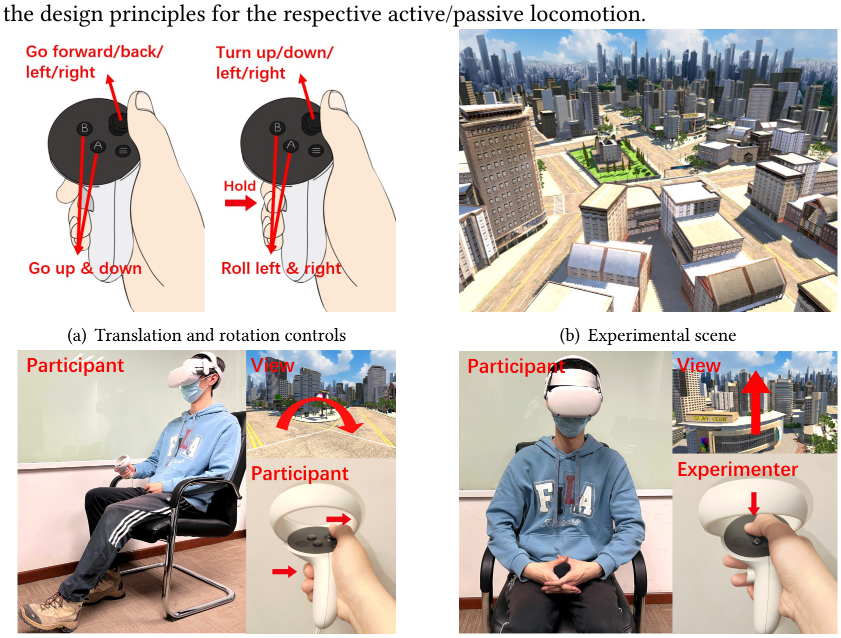
After excluding the visual and vestibular channels that could affect the measurements' validity, few sensory stimulus channels remain, such as proprioception, touch, and sound. So we categorized methods into non-tactile methods (e.g., body actions, voices), self-tactile methods (e.g., touching other body parts during actions), and external-tactile methods (e.g., touching objects outside the body during actions). The latter two provide different sensory experiences: self-tactile method provides dual feelings of the touching and touched parts, while external-tactile method provides more diverse sensations like texture and softness. Notably, methods where tactile feedback remains unchanged during score changes are still considered non-tactile, e.g., doing forearms/hands actions while resting elbows on chair armrests.

To provide rich visual motion stimulations, we utilized a city scene as the experimental setting (Figure 2b). Oculus Quest 2 headsets [65] were employed for the user experiment. The room used for the experiment contained only a chair and the VR device, with no other objects present to avoid providing any unintended cues to the participants.

### 3.3 Procedure

First, participants experienced transient physiological discomfort induced by active locomotion (participant-controlled using a handheld controller, Figure 2c) and passive locomotion (experimenter-controlled using a handheld controller, Figure 2d) in counterbalanced order. These transient discomfort in active and passive locomotion served as references.

Second, participants were tasked with designing 3 non-tactile, 3 self-tactile, and 3 external-tactile methods for each reference. Descriptions of these categories were similar to previous definitions, making them easy for participants with HCI/design backgrounds to understand. If a design didn't align with the needed category, the experimenter would ask for a redesign. If more than 3 methods were suggested within a category, the participant selected the three most suitable. The design process followed a think-aloud protocol [22], where participants elaborated on their designs, including inspirations, score ranges, and score mapping methods, and demonstrated the actions and their score mapping methods to the experimenter. When necessary (e.g., if participants suggest speed mapping), record participants' live actions for subsequent speed analysis. The design principles included no visual displays or actions that affect vestibular perception (e.g., head or trunk actions), and achieved natural, rapid, accurate, and low-load dynamic feedback during physiological discomfort. The experimenter couldn't provide any design-related hints or evaluation but was available to answer other experiment-related questions. The designed methods for active and passive locomotion could be reused, provided that they adhere to the design principles for the respective active/passive locomotion.




(a) Translation and rotation controls

(b) Experimental scene

(c) Elicited process in active locomotion

(d) Elicited process in passive locomotion

Fig. 2. Control methods, scene and elicitation process

Then, participants experienced their designed methods, which involved navigating with a controller or being controlled locomotion by the experimenter, while performing measurement actions based on their discomfort levels. For external-tactile methods, we used daily/3D-printed objects to simulate shape, softness, etc., based on participants' descriptions, which were pre-prepared based on the envisioned design space and placed in a non-experimental room. During this process, participants described their feelings and analyzed the strengths and weaknesses of the methods.

At the end of the study, participants were interviewed to select their most preferred method for each sensory stimulation type (i.e., non-tactile, self-tactile, and external-tactile) and locomotion type (i.e., active and passive), as well as an overall preferred method that could be well-suited for both active and passive locomotion.

The experiment involved two experimenters: one guided participants and recorded their designs, feelings, and preferences, while the other presented references and controlled passive locomotion. Experimenters viewed participants' perspectives in real-time via the Oculus Casting feature [66]. Participants had sufficient rest between locomotion sessions, with the experiment lasting about 1 hour per participant and a $20 cash reward.

### 3.4 Evaluation method

The appearance and preference counts of designed methods reflect their intuitiveness and practicality. We distinguish between user-elicited methods and calculate the appearance and preference count of each gesture under various conditions. Gestures with similar actions and sensory stimulation are categorized as the same method, based on records, descriptions, and discussions. For example, bending the left and right legs is considered the same method, 'bending single leg'.

To assess the agreement among methods from different participants, we calculated the agreement rate of each sensory stimulation type and locomotion type separately. The agreement rate $AR$ for each sensory stimulation type or locomotion type $t$ was calculated using Equation (1) [72]:

$$AR_{(t)} = \sum_{M_i \subseteq M} \frac{M_i(M_i - 1)}{M(M - 1)}, \tag{1}$$

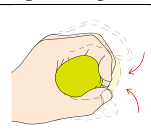
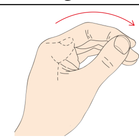
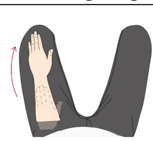
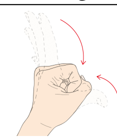
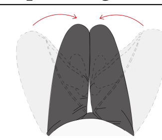
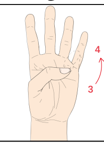
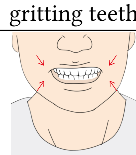
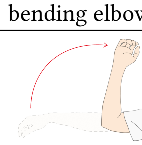
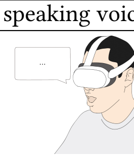
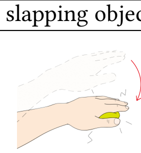
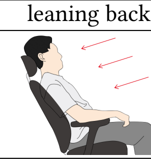
where $M$ represented the total number of elicited methods for a sensory stimulation type or locomotion type, and $M_i$ represented the number of a subset $i$ of identical methods from $M$. Qualitative interpretations for agreement rate were as follows [72]: $low$ ($AR \leq 0.1$), $moderate$ ($0.1 < AR \leq 0.3$), $high$ ($0.3 < AR \leq 0.5$), and $very\ high$ ($AR > 0.5$). The agreement rate was derived using a method set that only considered one method per participant, so we used the per participant's preferred method to calculate $AR$. Finally, we counted overall preferences.

### 3.5 Experimental Results

*3.5.1 User-elicited Continuous Measurement Methods.* We collected 24*3*3=216 continuous measurement methods of active and passive locomotion, respectively. After classification, there were 40 types of non-tactile methods (27 via hands/arms, 11 via legs/feet, 1 via mouth, and 1 via voice), 40 self-tactile methods (16 via touching hands/arms, 17 via legs/feet, 2 via touching torso/stomach, 3 via touching head/face, and 2 via touching teeth), and 41 external-tactile methods (3 via touching the ground, 5 via chairs, 14 via springs, zippers, softballs, etc., and 19 via specially designed hardware like pedals, joysticks, rings, sliders, knobs, buttons, etc.).

*3.5.2 Agreement Rates.* Agreement rates and their levels were as follows: non-tactile methods (0.127, $moderate$), self-tactile methods (0.181, $moderate$), external-tactile methods (0.351, $high$), active locomotion (0.156, $moderate$), passive locomotion (0.091, $low$), and overall preference (0.178, $moderate$).

Table 1. The top 12 user-elicited continuous measurement methods ranked by overall preferences

| Method | squeezing ball | sliding thumb | rubbing thigh | clenching fist | spreading knee | drawing number |
|---|---|---|---|---|---|---|
| Schematic diagram |  |  |  |  |  |  |
| Sensations | external-tactile | self-tactile | self-tactile | non-tactile | non-tactile | non-tactile |
| Mapping method | L-distance(19)<br>NL-force(11)<br>L-force(8) | NL-distance(10)<br>L-distance(8)<br>L-angle(5) | L-distance(19)<br>L-frequency(4) | L-distance(18)<br>L-area(6) | L-distance(10)<br>L-angle(8) | semantic(12) |
| Score range | 0-10(20)<br>0-5(12)<br>0-4(4)<br>0-100(2) | 0-10(13)<br>0-5(7)<br>0-3(2)<br>0-7(1) | 0-10(12)<br>0-5(6)<br>0-100(5) | 0-10(11)<br>0-4(6)<br>0-5(5)<br>0-20(2) | 0-10(8)<br>0-5(8)<br>0-7(2) | 0-5(8)<br>0-10(4) |
| Appearance | 20;18 | 12;11 | 13;10 | 11;13 | 8;10 | 3;9 |
| Preference | 14;8;4;8 | 8;5;5;6 | 7;3;4;5 | 5;1;2;1 | 5;2;0;1 | 4;1;2;1 |
| Method | pedaling | gritting teeth | bending elbow | speaking voice | slapping object | leaning back |
| Schematic diagram |  |  |  |  |  |  |
| Sensations | external-tactile | self-tactile | non-tactile | non-tactile | external-tactile | external-tactile |
| Mapping method | L-angle(12)<br>NL-angle(5)<br>L-distance(9) | NL-pressure(5)<br>L-pressure(3) | L-distance(13)<br>L-angle(13) | semantics(6)<br>L-volume(3)<br>NL-speed(3)<br>NL-pitch(1) | NL-force(6)<br>L-frequency(4)<br>NL-voice(2) | NL-pressure(4)<br>L-pressure(3) |
| Score range | 0-10(13)<br>0-5(11)<br>0-100(2) | 0-5(3)<br>0-2(3)<br>0-10(2) | 0-10(14)<br>0-4(6)<br>0-5(4)<br>0-20(2) | 0-10(6)<br>0-5(5)<br>0-2(2) | 0-10(5)<br>0-5(4)<br>0-3(3) | 0-10(4)<br>0-5(2)<br>0-100(1) |
| Appearance | 16;10 | 5;3 | 14;12 | 8;5 | 7;5 | 5;2 |
| Preference | 3;1;1;1 | 1;0;1;1 | 4;2;1;0 | 3;1;1;0 | 2;1;0;0 | 2;1;0;0 |

*3.5.3 Appearance Count.* For active locomotion, the most frequent method involved single-handed actions (139 times), followed by leg/foot actions (54 times). For passive locomotion, single-handed actions were the most frequent (132 times), followed by double-handed actions (46 times).

In non-tactile methods, finger/wrist/arm actions were most common (81 times), followed by leg/foot actions (48 times), voice commands (13 times), and mouth actions (i.e., 2 times of opening mouth size). In self-tactile methods, the frequencies from most to least were: hand/arm touching hand/arm (64 times), hand/arm touching leg/foot (39 times), leg/foot touching leg/foot (19 times), upper teeth touching lower teeth (8 times), hand touching torso/abdomen (6 times), hand touching head/face (6 times), and tongue touching teeth (2 times). For external-tactile methods, the frequencies from most to least were: hand/arm touching daily/3D-printed objects (58 times), leg/foot touching specially designed hardware (30 times), hand/arm touching specially designed hardware (24 times), hand/arm touching seated chairs (13 times), leg/foot touching the ground (10 times), waist/back touching seated chairs (7 times), thigh touching daily/3D-printed objects (1 time), and thigh touching seated chairs (1 time).

*3.5.4 Score Mapping.* The most common attributes mapped to scores were spatial properties (253 times), such as distance, angle, trajectory, and area formed with other body parts/objects. Next were intensity attributes (108 times), like force, pressure, volume, and pitch, followed by temporal attributes (45 times), including speed, frequency, and speech rate. Semantic attributes (26 times), such as the meaning of actions and verbal expressions, were least frequent.

For score ranges, 0-10 was the most common (216 times), followed by 0-5 (165 times). Regarding mapping types, linear mapping was the most common (247 times), mainly used for spatial properties to score mapping. Non-linear mapping followed (159 times), mainly used for intensity and temporal attributes to score mapping.

*3.5.5 Top 12 Preferred Methods.* We identified the top 12 preferred methods based on overall preference counts, as shown in Table 1. Among them, "Mapping method" column includes mapping attributes (e.g., distance, angle) and mapping types (e.g., linear (**L**) and nonlinear (**NL**)); the numbers before and after the semicolon in "Appearance" column represent the appearance count of the method in active and passive locomotion, respectively; In "Preference" column, the 4 numbers separated by semicolons represent the preference counts for its sensory stimulation type, active locomotion, passive locomotion, and overall preference, respectively.

### 3.6 Discussion

*3.6.1 Reasons for Elicitation and Preference.* Overall, we found that many methods (22 in total) were inspired by instinctive reactions to discomfort and received more preferences (7 entered the top 12). This is understandable, as humans often involuntarily or habitually perform certain actions when feeling uncomfortable, such as clenching fists, sliding thumbs, rubbing thighs, gritting teeth, vocalizing, etc. Participants believed that people are more likely to accept and adapt to such methods with a high level of interactive naturalness because they align with physiological responses and are less likely to be forgotten or ignored during locomotion.

We also found that many methods are inspired by everyday actions like gesturing numbers, drawing on surfaces, pressing foot pedals, turning knobs, sliding sliders, or moving joysticks. While these are commonly used for other contexts like controlling VR locomotion or providing real-time feedback on difficulty, tension, or comfort, they didn't receive high preference in our study. For instance, while five participants elicited 'sliding slider' and six elicited 'turning knob', only one participant favored them in the external-tactile category, and none preferred them overall. Five participants often forgot to provide timely ratings or felt confused when perceiving the current rating. Additionally, three participants noted that perceiving corresponding scores through proprioception while rotating fingers on the knob was more challenging than while moving fingers.

*3.6.2 Locomotion Type's Influence.* In active locomotion, despite the task constraints limiting one hand for controlling movement, most users still tended to design measurement methods using the other hand. Eleven users (45.8%) found it easier to coordinate actions with both hands independently rather than coordinating one hand with other body parts. In passive locomotion, with both hands being free, more actions involving both hands/arms emerged, further reducing the use of leg/foot-based measurement methods. The increased variety (26 types) of actions involving both hands/arms contributed to a lower agreement rate in passive locomotion. However, the prevalence of methods based on single hand/arm actions did not decrease with the increase in methods involving both hands. Twelve users (50%)

considered single hand/arm actions to have a lower burden compared to dual hand/arm actions and sufficient for providing feedback on subjective discomfort.

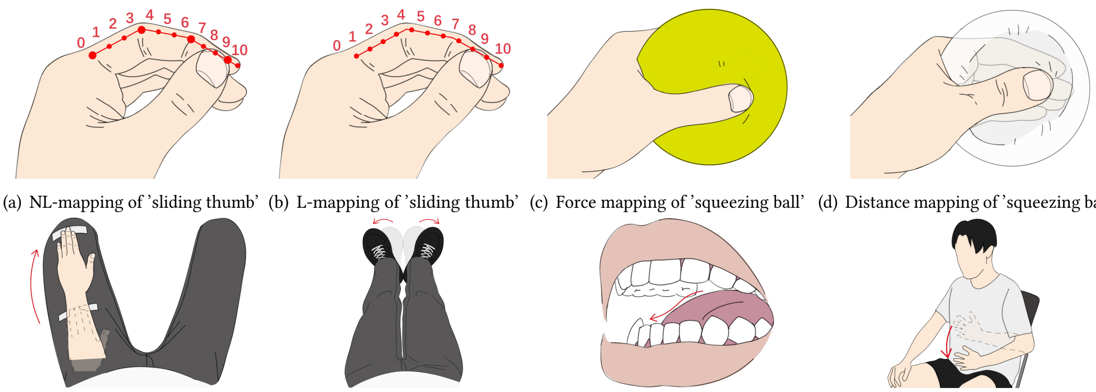
*3.6.3 Sensory stimulation type's Influence.* For the non-tactile methods, most participants (21 people, 87.5%) found it easy to perceive the position of their limbs through proprioception. They also noted that providing feedback through voice allowed them to immediately perceive the attributes of the sound, meeting the design principle of rapid and accurate feedback. Additionally, many participants (16 people, 66.7%) felt that using proprioception for aerial movements was more effortless compared to the tactile feedback methods, as it did not require overcoming friction or resistance from self or external objects. Although voice methods were proposed frequently, they were less preferred due to being strenuous and awkward for some users, especially when expressing discomfort through volume, speech rate, and pitch.

For the self-tactile methods, all participants (24 people) found that adding tactile feedback on the proprioception enhanced the realism of active feedback, particularly when there was dual tactile feedback from both the active and touched body parts, intensifying the perceptual experience. This category had more instinctive approaches (10 types) than the other two types of measurement methods. This is because people typically expect self-tactile feedback to distract and provide self-comfort when feeling uncomfortable, although it may not necessarily alleviate physiological discomfort like dizziness. Additionally, many participants (17 people, 70.8%) believed that performing actions on the body surface could provide better support and stability, resulting in more stable and fluid movement execution.

For external-tactile methods, most participants (22 people, 91.7%) found tactile feedback from external objects more novel and diverse than self-tactile feedback. They appreciated object surfaces as reference points for controlling movements and adjusting feedback scores, with many valuing varied textures, shapes, and temperatures. Despite the variety, this category had the highest agreement rate, mainly because a significant number of participants (21 people, 87.5%) elicited the 'squeezing ball' method, with 14 preferring it. However, some participants noted that strong tactile stimuli, while enhancing accuracy, could distract from visual motion. For instance, three participants felt the slider's friction diverted attention from the locomotion task, suggesting a need for balanced and moderate sensory stimulation.

*3.6.4 Mapping Method.* For the mapping methods, most participants preferred using linear spatial attributes (e.g., distance, angle, etc.) to map scores, considering it the simplest and most intuitive approach. Non-linear intensity attributes were also widely inspired, as participants believed they aligned with instinctual responses to discomfort, although they found them effortful. Time attributes (frequency, speed, etc.) and semantic attributes were generally considered less real-time and more effortful. Regarding score range design, participants favored two score ranges: 0-10 and 0-5, to increase scoring accuracy and reduce load, respectively. Interestingly, participants customized mapping methods and score ranges based on the characteristics of the measurement method. A typical example is 'sliding thumb', where most participants elicited by this method divide the score range based on areas on the index finger that can induce tactile changes. They designate the creases of the three joints on the index finger, along with the junction of the fingertip and the fingerpulp, as 0, 3, 6, and 9 points, and distribute other scores linearly across corresponding segments of each finger (Figure 3a). Participants believe that despite the unequal lengths of each finger segment, this non-linear score mapping based on tactile information is expected to have higher perceptual accuracy and action control than linear mapping (Figure 3b).

*3.6.5 The Top 3 Methods.* The top three preferred methods overall were 'squeezing ball', 'sliding thumb', and 'rubbing thigh', with the majority of preferential votes (19 participants, 79.2%). Eleven participants simultaneously elicited 2 or all of the 3 methods, and they were often undecided on selecting the most preferred one. This suggests these methods




(a) NL-mapping of 'sliding thumb' (b) L-mapping of 'sliding thumb' (c) Force mapping of 'squeezing ball' (d) Distance mapping of 'squeezing ball'

(e) 'Rubbing thigh' with tactile changes (f) 'Opening feet' (g) 'Tongue against teeth' (h) 'Massaging stomach'

Fig. 3. Ingenious design ideas of score mapping methods and other elicited interesting measurement methods

are considered the most suitable and applicable for both active and passive locomotion. These methods involve tactile feedback, indicating that compared to semi-air actions with only proprioceptive stimuli, additional tactile stimuli make participants' active feedback more vivid and engaging, enhancing the perception of scoring action.

'Squeezing ball' mimics the instinctual clenched fist response to discomfort, with participants finding that adding a soft object enhances the interactive experience. While some users used rubber or elastic balls with limited deformation for grip strength measurement (Figure 3c), most users suggested using spherical airbags, which could be effortlessly fully compressed and quickly restored, mapping scores based on linear deformation or thumb-finger distance (Figure 3d). 'Sliding thumb' mimics the instinctual finger-rubbing response to discomfort and is considered a very rapid and effortless movement due to its small amplitude. With the mapping method and score range mentioned earlier, this method is expected to provide a natural and accurate measurement. 'Rubbing Thigh' mimics the instinctual rubbing of thighs during discomfort, with most participants mapping scores based on the distance their hand moves on the thigh. Although the movement is larger than rubbing fingers, it's a simple, intuitive, and quick self-touch method. Additionally, some users suggested affixing tape or cloth strips at the starting (0) and ending (10) points of the score range to provide tactile feedback different from trousers, helping participants accurately perceive the positions of both score extremes for more precise measurements (Figure 3e).

#### 3.6.6 *Other Interesting Methods.*

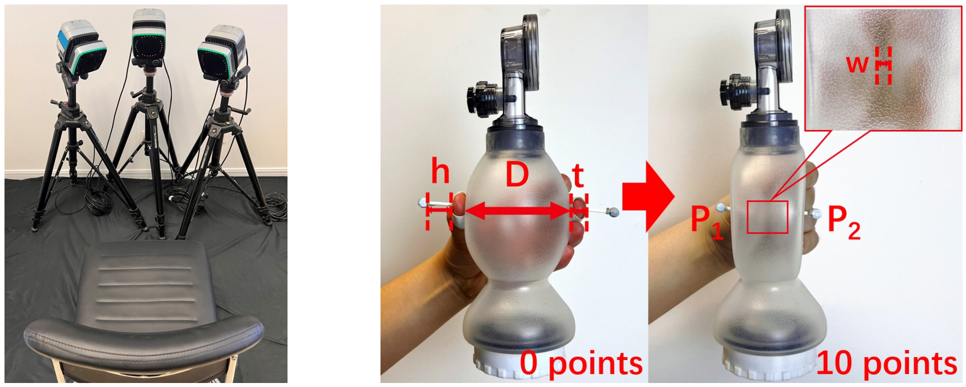
Some less frequently elicited and preferred methods were interesting and creative, suggesting potential. For example, inspired by the small, rapid movement of opening and closing both feet, some participants proposed 'opening feet' (Figure 3f). A few participants drew inspiration from the tongue's endurance during eating to propose 'tongue against teeth' by exerting pressure on teeth with the tongue (Figure 3g). Others, inspired by the nauseous feeling during discomfort, proposed 'massaging stomach' (Figure 3h).

## 4 STUDY 2: EVALUATION OF TOP 3 METHODS

Study 2 aims to evaluate the top 3 methods in Study 1: 'squeezing ball', 'sliding thumb', and 'rubbing thigh', due to such 3 methods having overwhelming preferences and there is no clear 'winner' among them. The study compares their performance in reliability, validity, and user experience to select the most optimal method.

### 4.1 Implements

Based on design preferences in Study 1, all 3 methods map spatial attributes of actions to a 0-10 score range. Thus, we implemented them using computer vision technology. We attached reflective markers to the tracked limbs and used motion capture equipment from Qualisys Arqus [54] (Figure 4a) to bind skeletal joint models and track them in real-time by its built-in algorithm [55].




(a) Tracking environment

(b) 'Squeezing ball' from 0 points to 10

Fig. 4. Tracking and score mapping of 'squeezing ball'

*4.1.1 'Squeezing Ball'.* We customized 3 common soft material spherical airbags: silicone, PVC (thermoplastic), and SEBS (thermoplastic rubber). Participants who designed this method in Study 1 were invited to test these airbags. After discussion, they unanimously agreed that silicone best suited their design ideals due to its softness, superior grip, ease of flattening, and quick rebound. We observed that participants typically placed their thumb and middle finger on opposite sides of the airbag, maintaining at the same height, with their distance equal to the diameter of the airbag. Therefore, we placed a 5mm diameter reflective marker on the nails of the thumb and middle finger, using 3D-printed 2cm rods as bases to prevent the markers from being obscured by the concave spherical airbag when pressed (Figure 4b). We mapped the action to discomfort scores using Equation (2).

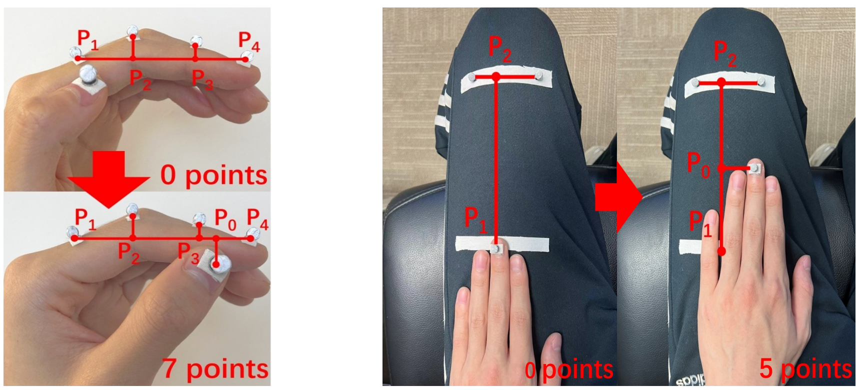
$$score = \lceil 10 \frac{D - d_{P_1 P_2} - 2(h + w + t)}{D} \rceil \tag{2}$$

where $d_{P_i P_j}$ represents the distance from $P_i$ to $P_j$. $P_1$ and $P_2$ represent the two reflective markers. $D$ represents the diameter of the spherical airbag, and $h$ represents the height of the rod base. $t$ represents the thickness of the finger, and $w$ represents the width of the airbag wall (Figure 4b).

*4.1.2 'Sliding Thumb'.* Following the scoring mapping approach in Study 1 (Figure 3a), we placed a 5mm diameter marker at the tip of the thumb. Additionally, we attached 4mm markers at the junction of the index finger's three joints and the boundary between the fingertip and the finger pulp, dividing the range into four segments: 0-3 points, 4-6 points, 7-9 points, and 10 points, as shown in Figure 5a. We mapped the action to discomfort scores using Equation (3).

$$score = \begin{cases} 0, & \text{if } d_{P_0P_4} >= d_{P_1P_4} \\ \lceil 3\frac{d_{P_1P_0}}{d_{P_2P_1}} \rceil, & \text{if } P_0 \in P_1P_2 \\ \lceil 3\frac{d_{P_2P_0}}{d_{P_2P_3}} \rceil + 3, & \text{if } P_0 \in P_2P_3 \\ \lceil 3\frac{d_{P_3P_0}}{d_{P_3P_4}} \rceil + 6, & \text{if } P_0 \in P_3P_4 \\ 10, & \text{if } d_{P_1P_0} >= d_{P_1P_4} \end{cases} \tag{3}$$

where $P_1$ and $P_4$ represent the markers at the first joint of the index finger and the junction of the finger pulp and the fingertip, $P_2$ and $P_3$ represent the projection points of the markers located at the second and third joints of the index finger onto the line segment $P_1P_4$, $P_0$ represents the projection point of the thumb marker to the line segment $P_1P_4$, $d_{P_iP_j}$ represents the distance from $P_i$ to $P_j$ (Figure 5a).




(a) 'Sliding thumb' from 0 points to 7 (b) 'Rubbing thigh' from 0 points to 5

Fig. 5. Score mapping of 'sliding thumb' & 'rubbing thigh'

#### 4.1.3 *'Rubbing Thigh'.*

As the setup in Study 1 (Figure 3e), a 5mm diameter marker was affixed to the tip of the middle finger. When the palm rested on the thigh with its base at the junction of the thigh and torso, the position of the middle fingertip was considered the 0-point position. Tape was affixed to the pants at this location to provide tactile feedback. Additionally, we attached 5mm diameter markers on each side of the knee joint and connected them with tape on the pants to signify the 10-point position (Figure 5b). We mapped the action to discomfort scores using Equation (4).

$$score = \lceil 10\frac{d_{P_1P_0}}{d_{P_1P_2}} \rceil \tag{4}$$

where $d_{P_iP_j}$ represents the distance from $P_i$ to $P_j$. $P_1$ represents the 0-point position. $P_2$ is the projection point of $P_1$ onto the line segment of the two markers at the 10-point position. $P_0$ represents the projection of the marker on the tip of the middle finger onto the line segment $P_1P_2$ during the rating process (Figure 5b).

### 4.2 Participants, Conditions, and Settings

The experiment received ethics approval from our university. Following the criteria outlined according to MSSQ in Section 3.1 for participant selection, 18 participants (9 male, 9 female; mean age = 22.89, SD = 1.91) were recruited from three local universities. The participants were categorized based on their VR experience proficiency: 6 had high proficiency, 5 had some proficiency, and 7 had low proficiency. Participants did not overlap with those in Study 1.

Table 2. Segments in the locomotion route

| Locomotion segments | Included sub-locomotion |
|---|---|
| Part 1 (P1) | go-forward 60m, turn-left 90°, go-forward 55m,turn-right 65° |
| Part 2 (P2) | go-forward 20m,stop 2s,go-forward 15m,stop 2s,go-forward 15m |
| Part 3 (P3) | turn-left 65°,go-forward 70m,turn-right 90°,go-forward 30m,turn-right 55° |
| Part 4 (P4) | roll-right 300°,go-upward 22m,go-left 8m,go-forward 22m |
| Part 5 (P5) | roll-right 60°,turn-right 55°,turn-up 15°,go-forward 70m |
| Part 6 (P6) | turn-down 45°,go-forward 40m,turn-up 30°,turn-right 40° |
| Part 7 (P7) | go-forwards 30m,turn-right 30°,go-left 30m,turn-down 90° |
| Part 8 (P8) | go-backward 12m,go-upward 32m,turn-up 60°,turn-right 80°,go-forward 50m |
| Part 9 (P9) | roll-right 30°,turn-up 10°,go-forward 82m,turn-right 75°,go-forward 30m |
| Part 10 (P10) | roll-left 60°,go-downward 15m, go-forward 60m, roll-left 330° |

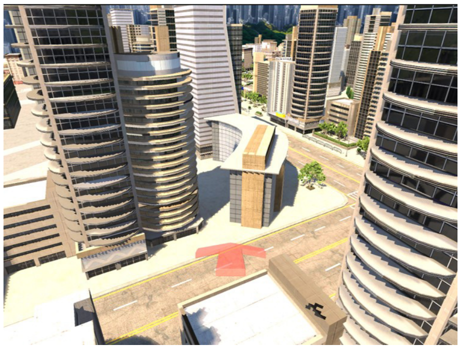
Four experimental conditions were established: three continuous measurement methods (i.e., 'squeezing ball', 'sliding thumb', and 'rubbing thigh'), and one discrete measurement method (i.e., FMS). FMS (Fast Motion Sickness Scale [38]), a widely-used single-item scale to assess physiological discomfort intermittently [47, 68], provides accurate momentary assessments, is less affected by memory decay than multi-item scales[68, 71], and correlates highly with standard scales like SSQ [37]. The study used the same city environment as Study 1, with a flying route of 10 segments and 4-6 sub-locomotions each to cover a wide range of locomotion situations (Table 2). The route design mimics typical VR navigation tasks [11, 14], varying speed, direction, and sub-locomotion to induce visual-vestibular conflicts essential for inducing transient discomfort. The routes were iteratively tested and refined to effectively induce discomfort without being overly intense.

During active locomotion, the method of using a handheld controller for flying remained the same as in Study 1, while semi-transparent arrows were placed as road signs at points of sub-locomotion transitions to help participants accurately and smoothly follow the designated route. These arrows indicate the type (i.e., translation, rotation, pause) and direction of the next sub-locomotion. The arrows were semi-transparent to distinguish them from other scene objects for guidance while not obstructing the visual stimuli provided by the background during the participant's locomotion (Figure 6a and b). Passive locomotion was controlled automatically by the system program along the set route without the arrows (Figure 6c and d). The completion time for the route is 5 minutes.

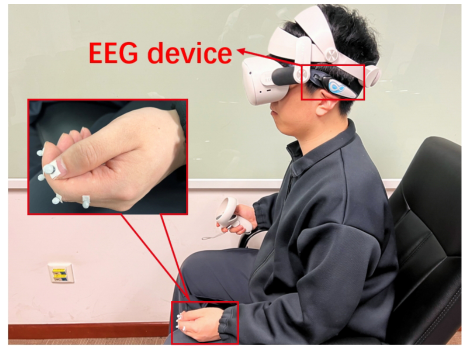
### 4.3 Evaluation Methods

*4.3.1 Reliability.* To validate the novel continuous measurement methods, the results need to demonstrate consistent ratings from participants, both within and between subjects. Within-subject consistency was determined by comparing scores from three repeated continuous measurements of the same participants, while between-subject consistency was determined by comparing scores across all participants. Reliability estimates for each method were calculated using $Cronbach's\,\alpha$, if the $\alpha$ coefficient is below 0.6, consistency reliability is considered insufficient; 0.7-0.8 indicates acceptable reliability, 0.8-0.9 is considerable, and above 0.9 is excellent [64].

*4.3.2 Validity.* Continuous measurement methods should provide measurement validity similar to mainstream discrete measurement methods like FMS. One approach to analyzing this validity is to compare each continuous measurement method with the widely accepted FMS benchmarks periodically. To conduct the validity test, the scores of continuous measurement methods should show significant correlations with the FMS scores for each locomotion segment. For this correlation calculation, multiple continuous feedback scores for each locomotion segment need to be condensed into

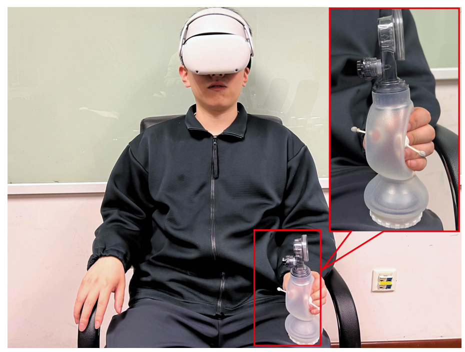


(a) Moving forward in active locomotion




(b) Using 'sliding thumb' while navigating

(c) Rolling rotation in passive locomotion



(d) Using 'squeezing ball' to rating

Fig. 6. The participants completing the locomotion route

one score. We will select a condensation method based on data, such as mean/median/max of continuous scores for each locomotion segment, and then compute using correlation analysis.

The correlation coefficient $r$ ranges between {-1,1}, where $r > 0$ indicates a positive correlation, and vice versa for a negative correlation. $|r|$ between 0.8-1.0 represents a very strong correlation, 0.6-0.8 indicates a strong correlation, 0.4-0.6 suggests a moderate correlation, and $< 0.4$ is considered a weak correlation.

*4.3.3 Attention.* Each measurement method should be evaluated for attention during locomotion to assess the level of distraction caused by active feedback. Methods with higher attention are considered to cause less distraction. In this study, we measured attention using the Brainlink Pro, a commercial portable EEG device, which has been shown to have similar accuracy to traditional multi-electrode EEGs [1, 50]. We utilized its attention values (range: 0-100), calculated based on spectral information such as alpha and beta waves using built-in algorithms. This attention index has been widely used in various studies [5, 26, 53].

*4.3.4 Workload.* Concurrently performing locomotion tasks while using continuous measurement methods should assess cognitive, physical, temporal, and other loads. We utilized the widely-used NASA-TLX for assessing workload across six dimensions, with scores ranging from 0 to 20 for each question [31].

*4.3.5 Presence.* Presence is one of the core experiences in VR. We assessed the impact of each continuous measurement method on presence using the Presence Scale, with scores ranging from 0 to 10 for each question [60]. The total score of the Presence Scale is the cumulative sum of all question scores.

### 4.4 Procedure

This experiment used a repeated measures design, with each participant completing all conditions over 4 days, one per day, in a counterbalanced order. Each day included practice in the experimental scene to anchor discomfort levels for each rating score and familiarize participants with the measurement method. The practice route followed the same sub-locomotion as the formal experiment but in a different sequence. For all 4 measurement methods, participants rated discomfort from 0 (no discomfort), with ratings increasing proportionally with discomfort intensity, to 10 (maximum discomfort). Among the 3 continuous measurement methods, participants adjusted their actions to match the target score in real-time, while the FMS method involved participants rating discomfort after each locomotion segment. Since discomfort levels and 0-10 points were anchored during practice, ratings of 8/9/10 did not necessarily reflect actual severe discomfort. Participants could withdraw at any time if discomfort became unbearable. Practice routes were repeated until participants understood the discomfort range and were proficient in using continuous methods during locomotion.

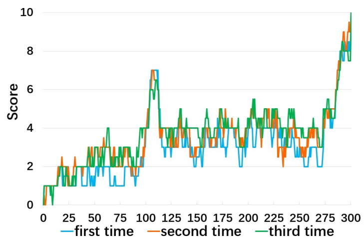
Each day, participants completed one active and one passive locomotion in counterbalanced order, following the formal route. This process was extra repeated 2 times to assess rating consistency. Between locomotion, participants completed the NASA-TLX and Presence Scale, and rested for 10-40 minutes until discomfort subsided. They were repeatedly asked to verbally confirm returning to feelings before the experiment. To monitor attention without adding distractions, lightweight, portable EEG devices were used (Section 4.3.3). Our testing showed that these devices caused minimal interference with task performance. Since EEG signals across the 3 times locomotion processes might be similar, and reattaching the device during rest breaks could be inconvenient, EEG was only worn during one randomly selected locomotion process (both active and passive). At the end of each day, participants were interviewed to describe their experiences with the measurement methods and locomotion. After completing all conditions, they underwent a final interview to analyze and compare the methods.

Participants received a $50 cash reward, even if they withdrew from the experiment due to discomfort. Three participants discontinued the experiment, and we recruited three replacements with similar MSSQ scores.

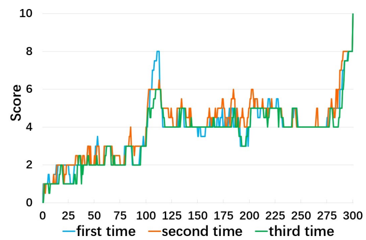
### 4.5 Experimental Results

Since most data did not follow a normal distribution according to a Shapiro-Wilks test, median values were preferred for overall data representation. We utilize line graphs to illustrate the dynamic changes in transient physiological discomfort data obtained through 3 methods, box plots to display the distribution of attention and scale data. Friedman tests, with post-Friedman pairwise comparisons (Dunn's approach and Bonferroni correction [18]) in SPSS, assessed significant differences between various conditions on attention, presence, and workload.

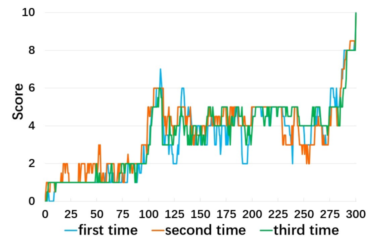
*4.5.1 Reliability.* The dynamic variations of the three repeated locomotion route data from participants using continuous measurement methods are depicted in the line chart (Figure 7), where the x-axis represents time (seconds) and the y-axis represents the median measurement values of all participants at that moment.



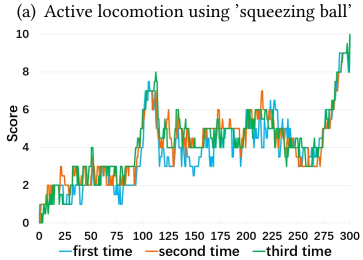

(a) Active locomotion using 'squeezing ball'



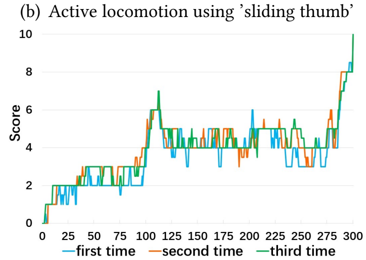

(b) Active locomotion using 'sliding thumb'



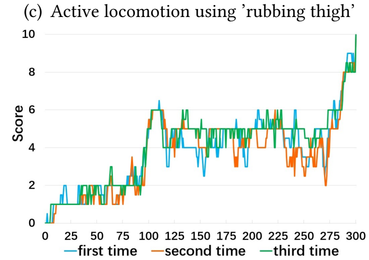

(c) Active locomotion using 'rubbing thigh'




(d) Passive locomotion using 'squeezing ball'




(e) Passive locomotion using 'sliding thumb'




(f) Passive locomotion using 'rubbing thigh'

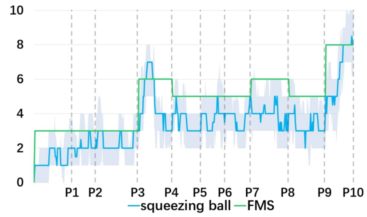
Fig. 7. The discomfort rating of the 3 repeated locomotion route data using 3 continuous measurement methods

We conducted within-subject reliability analysis with 3 items for participants' 3 repeated locomotion. The results show that the $Cronbach's\ \alpha$ for 'squeezing ball' is 0.718 (*acceptable*) in active locomotion and 0.817 in passive locomotion (*considerable*), for 'sliding thumb' is 0.827 in active locomotion (*considerable*) and 0.859 in passive locomotion (*considerable*), and for 'rubbing thigh' is 0.795 (*acceptable*) in active locomotion and 0.774 (*acceptable*) in passive locomotion.

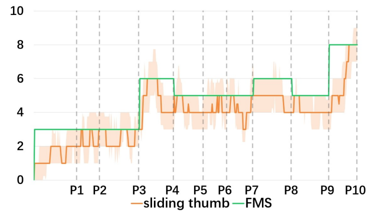
We conducted between-subject reliability analysis with 18 items for each participant. The results show that the $Cronbach's\ \alpha$ for 'squeezing ball' is 0.917 (*excellent*) in active locomotion and 0.928 (*excellent*) in passive locomotion, for 'sliding thumb' is 0.944 (*excellent*) in active locomotion and 0.936 (*excellent*) in passive locomotion, and for 'rubbing thigh' is 0.944 (*excellent*) in active locomotion and 0.946 (*excellent*) in passive locomotion.

*4.5.2 Validity.* To compare each continuous measurement method with the widely used FMS, we simplified the continuous measurement data for each locomotion segment into a single value. This was done by plotting both continuous measurement methods and FMS data on a line chart (Figure 8), where the x-axis represents locomotion segments (Part 1 to Part 10), the y-axis represents the median values of rating scores from all participants at that moment, and the colored area surrounding the continuous measurement curve represents the interquartile range (between the upper and lower quartiles). We observed that the FMS score generally matched the maximum value of the continuous data for the same segment. Thus, for each locomotion segment, we reduced the continuous data to the maximum value, resulting in ten scores per participant. Figure 9 compares the median of these maximum scores with the median FMS scores for each segment.

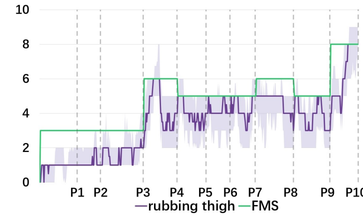
The results of the Spearman correlation analysis between the reduced data for each continuous measurement method and the FMS data showed that the correlation coefficient $r$ for 'squeezing ball' was 0.553 (*moderate*) for active locomotion and 0.640 (*strong*) for passive locomotion, for 'sliding thumb' it was 0.583 (*moderate*) for active locomotion and 0.676 (*strong*) for passive locomotion, and for 'rubbing thigh' it was 0.501 (*moderate*) for active locomotion and 0.597 (*moderate*) for passive locomotion.



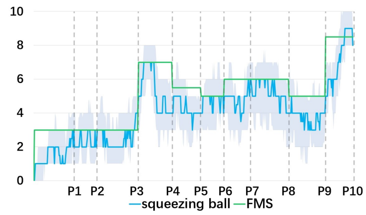

(a) 'Squeezing ball' & FMS in active locomotion



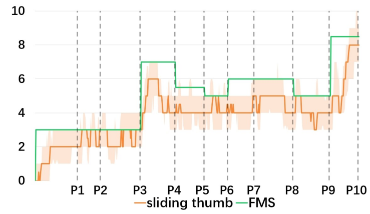

(b) 'Sliding thumb' & FMS in active locomotion



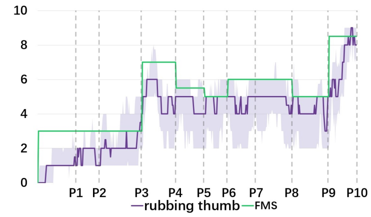

(c) 'Rubbing thigh' & FMS in active locomotion




(d) 'Squeezing ball' & FMS in passive locomotion

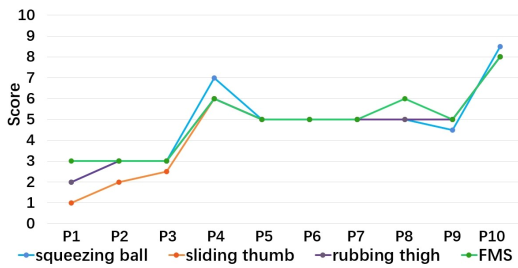



(e) 'Sliding thumb' & FMS in passive locomotion

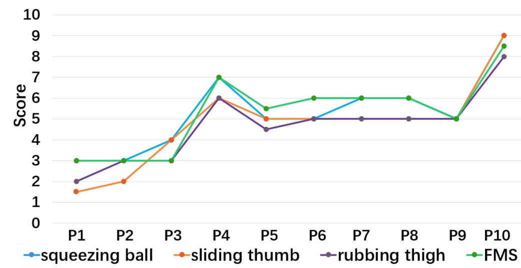



(f) 'Rubbing thigh' & FMS in passive locomotion

Fig. 8. Comparison of the 3 continuous measurement methods and FMS in active and passive locomotion




(a) Comparison of 3 methods after simplified score and FMS in active locomotion



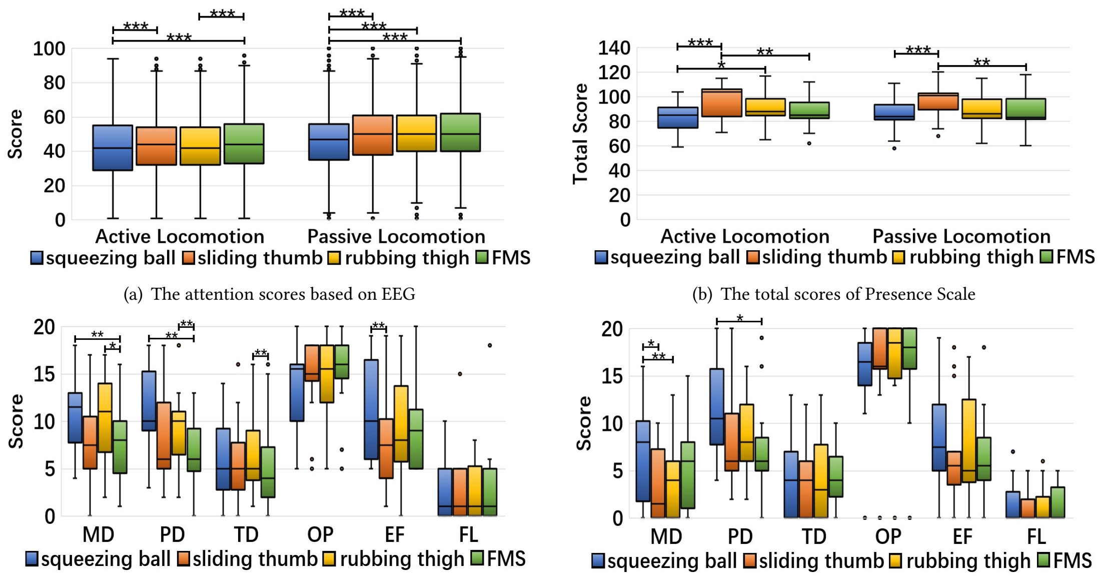

(b) Comparison of 3 methods after simplified score and FMS in passive locomotion

Fig. 9. Comparison of the maximum score for each locomotion segment across 3 continuous measurement methods and FMS

4.5.3 *Attention.* The results of the Friedman test indicated that the attention score was significantly influenced by different measurement methods in both active locomotion ($\chi^2(3) = 54.168, p < 0.001$) and passive locomotion ($\chi^2(3) = 45.609, p < 0.001$). The boxplots illustrating the distribution of attention data and the significant differences between conditions through Pairwise Comparisons are depicted in Figure 10a, where * denotes $p < 0.05$, ** denotes $p < 0.01$, and *** denotes $p < 0.001$.

4.5.4 *Presence.* The results of the Friedman test revealed that the total score of the Presence Scale was significantly influenced by different measurement methods in both active locomotion ($\chi^2(3) = 28.116, p < 0.001$) and passive locomotion ($\chi^2(3) = 22.911, p < 0.001$). The boxplots illustrating the distribution of the total score data of Presence Scale and the significant differences between conditions through Pairwise Comparisons are depicted in Figure 10b.

4.5.5 *Workload.* The results of the Friedman test revealed that the scores of Mental Demand in active locomotion ($\chi^2(3) = 17.089, p = 0.001$) and passive locomotion ($\chi^2(3) = 20.121, p < 0.001$), Physical Demand in active locomotion ($\chi^2(3) = 21.812, p < 0.001$) and passive locomotion ($\chi^2(3) = 12.669, p = 0.005$), Temporal Demand in active locomotion ($\chi^2(3) = 15.154, p = 0.002$), and Effort in active locomotion ($\chi^2(3) = 14.649, p = 0.002$) were significantly influenced by




(a) The attention scores based on EEG

(b) The total scores of Presence Scale

(c) The scores of Nasa-Tlx in active locomotion

(d) The scores of Nasa-Tlx in passive locomotion

Fig. 10. The scores of attention, Presence Scale, Nasa-Tlx of 3 continuous measurement methods and FMS

different measurement methods in NASA-TLX. However, Own Performance and Frustration Level were not significantly affected by the type of measurement method.

The boxplots illustrating the distribution of the data for the six questions of the NASA-TLX and the significant differences between measurement methods through Pairwise Comparisons are presented in Figure 10c and d.

### 4.6 Discussion

*4.6.1 Characteristics of transient physiological discomfort.* We found that all participants' data exhibited clear alternations between peaks and troughs, with high consistency both within and between subjects. This indicates that the transient physiological discomfort induced by locomotion route is rapid and easily perceived. We observed a significant increase in discomfort sensation when there were changes in sub-locomotion, such as transitioning from translation to rotation, from rotation to translation, or from a translation/rotation direction to another direction. This highlights that sudden changes in motion type and direction are easily perceived visually and result in significant momentary visual-vestibular conflicts due to the rich optical flow stimuli they produce.

During the acceleration phase of the same sub-locomotion, transient physiological discomfort also rapidly increased, typically reaching a peak upon reaching maximum speed, especially during rotation. This suggests that visual perception is more sensitive to rotational stimuli compared to translational stimuli, triggering stronger visual-vestibular conflicts. Among rotations, rolling rotation elicited the greatest discomfort sensation. This may be attributed to the infrequency of experiencing rolling rotations in daily life, leading to a lack of brain adaptation to this novel sensory mismatch, thereby exacerbating sensory conflicts. Similar examples include upward/downward translations inducing greater discomfort than translations in other directions, further demonstrating the influence of daily experiences on sensory integration.

Generally, we observed that transient discomfort scores would gradually decline from the peak after maintaining a constant velocity for 1-3 seconds following acceleration to maximum speed. This further illustrated the difficulty of the vestibular system in distinguishing between static and constant linear motion, and the brain to some extent believed that the body was actually moving at a constant velocity, thereby reducing visual-vestibular conflicts [40]. We also found that passive motion typically resulted in higher peaks in physiological discomfort data compared to active motion, reaffirming the role of the brain's motor intention in modulating sensory integration [46].

In addition, we observed that the troughs of physiological discomfort in the later stages of the locomotion path were generally higher than those in the earlier stages. We speculate that this may be due to the gradual onset of the long-term effects of motion sickness, as all participants reported varying degrees of discomfort that persisted well beyond the end of the locomotion path. This suggests that the gradual elevation of the troughs over time is likely caused by the accumulation of frequent instantaneous sensory conflicts, which trigger the persistent discomfort characteristic of long-term effects. This indirectly indicates that continuous measurements of transient physiological discomfort can, to some extent, also reflect the long-term effects of motion sickness. In the middle to later stages of locomotion, the data continued to exhibit noticeable fluctuations in response to sub-locomotion changes, despite the elevated peak and trough levels compared to earlier stages. This suggests that the transient discomfort associated with the short-term effects of motion sickness is further superimposed on the persistent discomfort associated with its long-term effects.

According to participants' feedback, discomfort in the eyes and dizziness were the most commonly reported symptoms of transient physiological discomfort, with almost every sub-locomotion eliciting changes in at least one of these feelings. Nausea was the least frequent feeling, occurring later in the locomotion route, but participants found it to be the least preferable discomfort feeling, often resulting in higher discomfort scores. Participants reported that when experiencing simultaneous changes in multiple discomfort feelings, their scores fluctuated more compared to experiencing only one discomfort sensation. Interestingly, most participants seemed to employ this implicit scoring rule in both continuous measurement methods and FMS, resulting in high reliability within and between subjects. Discrete measurement methods are influenced by the "peak-end rule" when recalling experiences, where subjective ratings tend to reflect the most intense moment (peak) or the final moment (end) of the experience. In the context of measuring transient physiological discomfort, we found a strong correlation between the peak values of the continuous measurement method and the FMS ratings within each locomotion segment.

*4.6.2 Comparison of different measurement methods.* Overall, all three continuous measurement methods demonstrated good within-subject consistency, excellent between-subject consistency, and good correlations with FMS, especially the 'sliding thumb'. This indicates that the reliability and validity of the three continuous measurement methods meet the standards for application in actual locomotion scenarios. Participants did not report forgetting or ignoring continuous ratings when experiencing discomfort, they generally found these methods easy to grasp and could naturally provide feedback actions when experiencing discomfort. This underscores the role of instinctive feedback actions.

Among the three continuous measurement methods, 'sliding thumb' demonstrates outstanding performance, with the highest reliability, highest validity, most focused attention, strongest presence, and lowest workload. Compared to FMS, it also does not significantly disrupt experiences. This method involves small action, is effortless, and utilizes the high flexibility of the fingers, ensuring real-time effectiveness. Participants feedback that this method, compared to others, provides the more moderate tactile stimulation through finger friction, offering sufficient proprioceptive and tactile perception without diverting visual attention, especially with the tactile changes provided by the wrinkles on the index finger at 0, 3, 6, and 9 points, enhancing precise control over scores.

For 'squeezing ball', many participants praised its advantages during interviews, such as the method providing ample and natural tactile feedback during interaction, and the act of squeezing a soft ball providing a certain degree of mental relief during physiological discomfort. Additionally, its reliability and validity in passive motion are obviously higher than in active motion, indicating its susceptibility to the locomotion type. However, the method exhibits the most significant differences in attention and presence, all of which are lower than some other measurement methods, and in mental demand, physical demand, and effort, all of which are higher than some other measurement methods. This suggests that the method still has some notable drawbacks. On one hand, the tactile stimulation produced by this action is more intense than other methods, including the repulsion generated by the spherical airbag's automatic recovery and its cold surface temperature, leading to more distraction, especially during active motion. On the other hand, despite the different tactile stimuli caused by the shape and repulsive force of the airbag at each score, users still frequently reported a vague perception of the current score while squeezing the spherical airbag. We speculate that this is due to the excessively dense and subtle tactile variations, which may increase the sensory perception load, thereby interfering with the integration of proprioceptive and tactile sensations.

For 'rubbing thigh', this action exhibits the highest within-subject consistency and the lowest correlation with FMS. Moreover, compared to FMS, it shows lower attention and higher mental, physical, and temporal demands during active locomotion. We attribute this to two main reasons: firstly, the large amplitude of arm action and their lesser flexibility compared to finger action, which affects the immediacy of this action; secondly, similar to 'squeezing ball,' participants report issues with perceiving the current score clearly during the middle process due to a single friction stimulus, and this perceptual ambiguity is further exacerbated during more distracting active locomotion.

For FMS, despite its proven advantage in capturing the overall discomfort trend of motion sickness, its low measurement resolution significantly limits its ability to detect transient discomfort changes caused by sub-locomotion. Moreover, 14 users reported FMS frequently interrupted their locomotion experience, with median presence often being the lowest, and 17 users reported that using FMS required some time and effort to recall their experiences.

## 5 GENERAL DISCUSSION

Based on the results and findings from User Studies 1 and 2, this section further proposes a method for distinguishing between the short-term and long-term effects of motion sickness using the collected continuous measurement data; it also offers design guidelines for continuous measurement methods in the unique context of transient physiological discomfort associated with the short-term effects of motion sickness and analyzes the utility and application value of such methods.

### 5.1 Short-Term and Long-Term Effects of Motion Sickness

Building on the study findings, this section further deepens the interpretation of the characteristics of continuous measurement data, with a particular focus on discussing the potential relationship between the short-term and long-term effects of motion sickness—which respectively represent real-time and accumulated sensory conflicts—and how these two effects can be distinguished within the collected continuous measurement data.

All participants reported that in the early phase of locomotion, discomfort quickly diminished to low levels once the motion stabilized (e.g., transitioning/rotating from acceleration to constant velocity). However, in the mid-to-late phase, even after the locomotion stabilized, participants still experienced some level of discomfort, which persisted at varying intensities after the route ended (Section 4.6.1). This suggests that in the early stages of locomotion, the transient discomfort reflecting the short-term effects of motion sickness may serve as an early indicator or precursor to its long-term effects. Additionally, we hypothesize that the valleys in our data reflect persistent discomfort from long-term effects. On the one hand, similar to Fu et al.'s findings [23], the valleys show a gradual upward trend, consistent with the cumulative nature of long-term effects. On the other hand, from a neuroscience perspective, the vestibular signal is sensitive to acceleration and does not cause strong transient conflicts with the uniform motion speed (i.e., valley parts) in the visual signal [4, 40].

Interestingly, even in the mid-to-late phases of locomotion, noticeable peaks and valleys remained, indicating that the temporal characteristics of transient discomfort persisted despite underlying long-term discomfort. This allows for distinguishing short- and long-term effects in mid-to-late phases. Persistent discomfort from long-term effects is reflected by the valleys, while rapid rises and falls (peaks between valleys) indicate transient discomfort. The peak values represent the strongest sensory conflicts during each sub-locomotion and should be given special attention.

### 5.2 Design Guidelines for Continuous Measurement Methods

Unlike other domains involving continuous measurement, assessing transient physiological discomfort caused by the short-term effects of motion sickness presents unique sensory overload challenges. Specifically, participants must simultaneously focus on visual locomotion (and navigational actions), endure and quantify fluctuating discomfort—all while quickly and accurately perceiving the score represented by their current action and adjusting it to match the intended target score. This challenge led to differences in design concepts and user preferences observed in Study 1, as well as performance differences among the three methods evaluated in Study 2. Based on the findings from these two studies, we propose a set of design guidelines to address this challenge and support researchers and designers in developing new or improving existing continuous measurement methods:

- As emphasized in Section 3.3, continuous measurement methods should avoid interfering with visual and vestibular senses to maintain measurement accuracy, as visual-vestibular conflict induces transient discomfort. This means excluding visual displays (e.g., visual sliders or scores) and actions involving head translation/rotation which add vestibular motion information.
- Methods need to perform well in both active and passive locomotion. As discussed in Section 4.6.3, active locomotion introduces additional challenges due to navigation action, making distraction a critical factor.
- As noted in Section 3.6.1, many participants' were inspired by instinctive responses to discomfort, as they are natural, memorable, and easy to learn. So instinctive actions should be prioritized as measurement methods.
- Statistical analysis in Section 3.5.3 and user preferences in Sections 3.6.1 and 3.6.2 show that hand actions, particularly those involving fingers and wrists, outperform other body parts due to their flexibility and lower effort. Designers should prioritize hand-based methods when available.
- Section 3.6.3 suggests that actions combining proprioception and tactile feedback provide better interaction and more accurate score perception than those relying solely on proprioception. Self-tactile methods enhance tactile stimulation, enabling more precise action adjustments during transient discomfort. Designers should integrate methods combining both for optimal effect.
- As discussed in Section 4.3, methods should ensure high intra- and inter-participant consistency (reliability) and strong correlations with validated metrics (validity) to assure repeatable and generalizable data collection.
- Based on Study 2 and the analysis in Section 4.6.2, "squeezing ball" caused sensory overload due to strong tactile stimulation, while "rubbing thigh" lacked clear feedback for intermediate scores. Designers should add tactile feedback points within the score range to provide moderate stimulation (like the wrinkles of "thumb sliding"), reducing perceptual load and improving score accuracy.

## 5.3 The Utility of Continuous Measurement

Given the significant challenges of sensory conflicts faced by VR locomotion scenes, especially active locomotion based on artificial continuous motion methods such as joysticks, steering wheels, and passive locomotion represented by 360-degree videos, this paper's continuous measurement methods can play an important role in improving and optimizing locomotion scenes/routes, dynamic intervention methods, and sensory conflict adaptation training.

Interaction designers can utilize collected data to optimize and improve locomotion scenes and routes to ensure users' physiological comfort. Based on the characteristics of data collected by three continuous measurement methods in Study 2, temporary physiological discomfort usually accompanies significant peaks and troughs in the types, directions, and parameter changes of sub-locomotion. By recognizing the occurrence of large peaks and reducing the time of its corresponding sub-locomotion, substituting other sub-locomotion with smaller peaks for sub-locomotion with large peaks, improving the connection process of sub-locomotion, modifying locomotion parameters to reduce sensory conflict intensity, increasing rest time when peaks or troughs are frequently at higher positions, and adjusting VR scenes at sub-locomotion with large peaks to reduce visual stimuli, measures can be taken to improve and optimize locomotion routes and scenes.

Continuous measurement methods can provide an experimental environment for dynamic intervention methods to help researchers explore new or evaluate existing solutions for reducing sensory conflicts. For example, applying multisensory stimulation methods such as vestibular electrical stimulation/adding moving sound sources to promote multisensory integration when peaks rapidly increase, reducing sensory conflicts through blurring the view edges/adding stable reference objects in the view, and automatically adjusting the types, directions, duration, and speed of sub-locomotion through adaptive algorithms, and even re-plans subsequent locomotion routes when peaks or troughs are maintained at high levels.

Furthermore, data obtained from continuous measurement methods can intuitively reflect individuals' sensitivity to sensory conflicts, allowing for the design of targeted progressive training programs to gradually increase the intensity and complexity of sub-movements, enabling users to gradually adapt to visually induced sensory conflicts and thereby improve the acceptability and effectiveness of locomotion scenes.

## 5.4 Limitation and Future Work

Firstly, as noted in Section 3.2, our measurement method was not designed or evaluated for standing postures or in-place locomotion due to weaker sensory conflicts, limiting its generalizability. Future work will explore transient physiological discomfort in standing postures, such as visual-vestibular conflicts from walking redirection [45], and the compatibility of our methods with in-place locomotion. These scenarios may introduce greater distraction and complicate body-touch interactions, potentially leading to differences in design compared to seated controller use.

Secondly, the proposed measurement method currently applies only to single-hand controller use or natural hand interactions, excluding scenarios where both hands are occupied with controllers, thus limiting its application. Future work will explore situations where both hands are occupied, potentially inspiring alternative methods involving other body parts or innovative controller-based techniques(e.g., finger sliding/tapping/pressing on controllers).

Thirdly, we only explored relatively simple tasks, such as scoring feedback while navigating. Future work will explore more complex scenarios, such as navigating, aiming, and shooting with the dominant-hand controller while providing scoring feedback with the non-dominant hand. These tasks introduce higher distraction and interference and may require special attention to action coordination of two-handed input for accurate measurement and interaction.

Fourth, the uniformity in young age and HCI/design backgrounds may have limited the set of continuous measurement methods in Study 1. For example, older users may prefer easier methods with smaller actions, while non-expert users may focus on a few top-ranked, most intuitive methods.

Fifth, this study focused on the top 3 preferred methods, potentially overlooking other viable methods. Future work will select, develop, and evaluate additional typical continuous measurement methods from user-elicited methods.

Finally, we standardized participants' criteria with absolute ratings to ensure reliable data, calibrating maximum discomfort at 10 points through a practice route (Section 4.4). No participant reported exceeding this in interviews, but it may occur due to factors like body state. We suggest allowing actions to exceed the 10-point marker with higher special scores if discomfort surpasses the pre-anchored maximum.

## 6 CONCLUSION

In conclusion, this study highlights the significance of understanding and effectively measuring transient physiological discomfort in VR locomotion scenarios. Through user-elicitation experiments and comprehensive evaluations, we developed and assessed the top 3 continuous measurement methods ('squeezing ball', 'sliding thumb', and 'rubbing thigh') tailored for capturing transient discomfort experiences. Among them, 'sliding thumb' showed the best performance, providing a natural experience with high reliability, validity, and minimal load. These findings highlight the need for natural continuous measurement methods to improve VR locomotion paths, understand transient discomfort mechanisms, and enable personalized interventions to mitigate discomfort.

## REFERENCES


[1] Ashrant Aryal, Ali Ghahramani, and Burcin Becerik-Gerber. 2017. Monitoring fatigue in construction workers using physiological measurements. *Automation in Construction* 82 (2017), 154–165.

[2] Niels H Bakker, Peter O Passenier, and Peter J Werkhoven. 2003. Effects of head-slaved navigation and the use of teleports on spatial orientation in virtual environments. *Human factors* 45, 1 (2003), 160–169.

[3] Miguel Barreda-Ángeles, Rafael Redondo-Tejedor, and Alexandre Pereda-Baños. 2018. Psychophysiological methods for quality of experience research in virtual reality systems and applications. *MMCT COMSOC Comm Front* 13, 1 (2018), 14–20.

[4] AJ Benson. 1990. Sensory functions and limitations of the vestibular system. *Perception and control of self-motion* (1990), 145–170.

[5] Yulong Bian, Chao Zhou, Yang Zhang, Juan Liu, Jenny Sheng, and Yong-Jin Liu. 2023. Focus on Cooperation: A Face-to-Face VR Serious Game for Relationship Enhancement. *IEEE Transactions on Affective Computing* (2023).

[6] Norbert Bischof and Eckart Scheerer. 1970. Systems analysis of optic-vestibular interaction in the perception of verticality. *Psychologische Forschung* (1970).

[7] Otmar L Bock and Charles M Oman. 1982. Dynamics of subjective discomfort in motion sickness as measured with a magnitude estimation method. *Aviation, Space, and Environmental Medicine* 53, 8 (1982), 773–777.

[8] Costas Boletsis, Jarl Erik Cedergren, et al. 2019. VR locomotion in the new era of virtual reality: an empirical comparison of prevalent techniques. *Advances in Human-Computer Interaction* 2019 (2019), 1–15.

[9] Evren Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Rajiv Dubey. 2016. Point & teleport locomotion technique for virtual reality. In *Proceedings of the 2016 annual symposium on computer-human interaction in play*. ACM, 205–216.

[10] Adolfo M Bronstein, John F Golding, and Michael A Gresty. 2013. Vertigo and dizziness from environmental motion: visual vertigo, motion sickness, and drivers' disorientation. In *Seminars in neurology*, Vol. 33. Thieme Medical Publishers, 219–230.

[11] Fabio Buttussi and Luca Chittaro. 2019. Locomotion in place in virtual reality: A comparative evaluation of joystick, teleport, and leaning. *IEEE transactions on visualization and computer graphics* 27, 1 (2019), 125–136.

[12] Chenyang Cai, Jian He, and Tianren Luo. 2023. Using Redirection to Create a Swimming Experience in VR for the Sitting Position. In *Companion Proceedings of the 28th International Conference on Intelligent User Interfaces*. 68–71.

[13] Umer Asghar Chattha, Uzair Iqbal Janjua, Fozia Anwar, Tahir Mustafa Madni, Muhammad Faisal Cheema, and Sana Iqbal Janjua. 2020. Motion sickness in virtual reality: An empirical evaluation. *IEEE Access* 8 (2020), 130486–130499.

[14] Weiya Chen, Anthony Plancoulaine, Nicolas Férey, Damien Touraine, Julien Nelson, and Patrick Bourdot. 2013. 6DoF navigation in virtual worlds: comparison of joystick-based and head-controlled paradigms. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*. 111–114.

[15] Diane Cleij, Joost Venrooij, Paolo Pretto, Daan M Pool, Max Mulder, and Heinrich H Bülthoff. 2017. Continuous subjective rating of perceived motion incongruence during driving simulation. *IEEE Transactions on Human-Machine Systems* 48, 1 (2017), 17–29.

[16] Guy Cohen-Lazry, Ariel Telpaz, Asaf Degani, and Tal Oron-Gilad. 2020. Identifying sources of discomfort in various road events while riding automated vehicles. In *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*. IEEE, 1–5.

[17] Xiao Dong, Ken Yoshida, and Thomas A Stoffregen. 2011. Control of a virtual vehicle influences postural activity and motion sickness. *Journal of Experimental Psychology: Applied* 17, 2 (2011), 128.

[18] Olive Jean Dunn. 1964. Multiple comparisons using rank sums. *Technometrics* 6, 3 (1964), 241–252.

[19] Marc O Ernst and Martin S Banks. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 6870 (2002), 429–433.

[20] Kim M Fairchild, Beng Hai Lee, Joel Loo, Hern Ng, and Luis Serra. 1993. The heaven and earth virtual reality: Designing applications for novice users. In *Proceedings of IEEE virtual reality annual international symposium*. IEEE, 47–53.

[21] Carlo Flemming, Benjamin Weyers, and Daniel Zielasko. 2022. How to Take a Brake from Embodied Locomotion–Seamless Status Control Methods for Seated Leaning Interfaces. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 728–736.

[22] Marsha E Fonteyn, Benjamin Kuipers, and Susan J Grobe. 1993. A description of think aloud method and protocol analysis. *Qualitative health research* 3, 4 (1993), 430–441.

[23] Rui Fu, Li Ma, Yingshi Guo, Qinyu Sun, Chang Wang, Wei Yuan, and Tingting Lan. 2024. Long and short-term characteristics of motion sickness: a test track investigation in a passenger car. *Cognition, Technology & Work* 26, 2 (2024), 313–324.

[24] John M Geringer and Clifford K Madsen. 2003. Gradual tempo change and aesthetic responses of music majors. *International Journal of Music Education* 1 (2003), 3–15.

[25] John M Geringer, Clifford K Madsen, and Dianne Gregory. 2004. A fifteen-year history of the Continuous Response Digital Interface: Issues relating to validity and reliability. *Bulletin of the Council for Research in Music Education* (2004), 1–15.

[26] Daniela Girardi, Nicole Novielli, Davide Fucci, and Filippo Lanubile. 2020. Recognizing developers' emotions while programming. In *Proceedings of the ACM/IEEE 42nd international conference on software engineering*. 666–677.

[27] John F Golding. 2006. Predicting individual differences in motion sickness susceptibility by questionnaire. *Personality and Individual differences* 41, 2 (2006), 237–248.

[28] Dianne Gregory. 1989. Using computers to measure continuous music responses. *Psychomusicology: A Journal of Research in Music Cognition* 8, 2 (1989), 127.

[29] Dianne Gregory. 1995. Research note: The Continuous Response Digital Interface: An analysis of reliability measures. *Psychomusicology: A Journal of Research in Music Cognition* 14, 1-2 (1995), 197.

[30] Xin Guo, Xin Pu, Youai Xia, Haopeng Guo, Yuyang Wang, and Lili Wang. 2022. Virtual tourism experience of Changbai Mountain scenic spot. In *Second International Symposium on Computer Technology and Information Science (ISCTIS 2022)*, Vol. 12474. SPIE, 467–471.

[31] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[32] Franziska Hartwich, Matthias Beggiato, and Josef F Krems. 2018. Driving comfort, enjoyment and acceptance of automated driving–effects of drivers' age and driving style familiarity. *Ergonomics* 61, 8 (2018), 1017–1032.

[33] Abraham M Hashemian, Matin Lotfaliei, Ashu Adhikari, Ernst Kruijff, and Bernhard E Riecke. 2020. Headjoystick: Improving flying in vr using a novel leaning-based interface. *IEEE Transactions on Visualization and Computer Graphics* 28, 4 (2020), 1792–1809.

[34] Abraham M Hashemian and Bernhard E Riecke. 2017. Leaning-based 360 interfaces: investigating virtual reality navigation interfaces with leaning-based-translation and full-rotation. In *Virtual, Augmented and Mixed Reality: 9th International Conference, VAMR 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings 9*. Springer, 15–32.

[35] Philipp Hock, Sebastian Benedikter, Jan Gugenheimer, and Enrico Rukzio. 2017. Carvr: Enabling in-car virtual reality entertainment. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. ACM, 4034–4044.

[36] Kazuhito Kato and Satoshi Kitazaki. 2006. *A study for understanding carsickness based on the sensory conflict theory*. Technical Report. SAE Technical Paper.

[37] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3, 3 (1993), 203–220.

[38] Behrang Keshavarz and Heiko Hecht. 2011. Validating an efficient method to quantify motion sickness. *Human factors* 53, 4 (2011), 415–426.

[39] James R Lackner. 2014. Motion sickness: more than nausea and vomiting. *Experimental brain research* 232 (2014), 2493–2510.

[40] Francesco Lacquaniti, Gianfranco Bosco, Silvio Gravano, Iole Indovina, Barbara La Scaleia, Vincenzo Maffei, and Myrka Zago. 2014. Multisensory integration and internal models for sensing gravity effects in primates. *BioMed research international* 2, 3 (2014), 124–132.

[41] Marc TM Lambooij, Wijnand A IJsselsteijn, and Ingrid Heynderickx. 2007. Visual discomfort in stereoscopic displays: a review. *Stereoscopic Displays and Virtual Reality Systems XIV* 6490 (2007), 183–195.

[42] William B Lathrop and Mary K Kaiser. 2002. Perceived orientation in physical and virtual environments: Changes in perceived orientation as a function of idiothetic information available. *Presence* 11, 1 (2002), 19–32.

[43] Joseph J LaViola Jr. 2000. A discussion of cybersickness in virtual environments. *ACM Sigchi Bulletin* 32, 1 (2000), 47–56.

[44] Juyoung Lee, Sang Chul Ahn, and Jae-In Hwang. 2018. A walking-in-place method for virtual reality using position and orientation tracking. *Sensors* 18, 9 (2018), 1–19.

[45] Yi-Jun Li, Frank Steinicke, and Miao Wang. 2022. A Comprehensive Review of Redirected Walking Techniques: Taxonomy, Methods, and Future Directions. *Journal of Computer Science and Technology* 37, 3 (2022), 561–583.

[46] Roman Luks and Fotis Liarokapis. 2019. Investigating motion sickness techniques for immersive virtual environments. In *Proceedings of the 12th acm international conference on pervasive technologies related to assistive environments*. ACM, 280–288.

[47] Tianren Luo, Chenyang Cai, Yiwen Zhao, Yachun Fan, Zhigeng Pan, Teng Han, and Feng Tian. 2023. Exploring Locomotion Methods with Upright Redirected Views for VR Users in Reclining & Lying Positions. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, 1–16.

[48] Tianren Luo, Zhenxuan He, Chenyang Cai, Teng Han, Zhigeng Pan, and Feng Tian. 2022. Exploring Sensory Conflict Effect Due to Upright Redirection While Using VR in Reclining & Lying Positions. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–13.

[49] Tianren Luo, Fenglin Lu, Jiafu Lv, Xiaohui Tan, Chang Liu, Fangzhi Yan, Jin Huang, Chun Yu, Teng Han, and Feng Tian. 2024. Exploring Experience Gaps Between Active and Passive Users During Multi-user Locomotion in VR. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.

[50] Rytis Maskeliunas, Robertas Damasevicius, Ignas Martisius, and Mindaugas Vasiljevas. 2016. Consumer-grade EEG devices: are they usable for control tasks? *PeerJ* 4 (2016), e1746.

[51] Kenneth E Money. 1970. Motion sickness. *Physiological reviews* 50, 1 (1970), 1–39.

[52] Dominik Mühlbacher, Markus Tomzig, Katharina Reinmüller, and Lena Rittger. 2020. Methodological considerations concerning motion sickness investigations during automated driving. *Information* 11, 5 (2020), 265.

[53] Geeta U Navalyal and Rahul D Gavas. 2014. A dynamic attention assessment and enhancement tool using computer graphics. *Human-centric Computing and Information Sciences* 4 (2014), 1–7.

[54] Qualisys. 2024. Qualisys Arqus. https://www.qualisys.com/cameras/arqus/. (2024).

[55] Qualisys. 2024. Tracking algorithm of Qualisys. https://www.qualisys.com/integrations/qtm-connect/unity/. (2024).

[56] James T Reason. 1978. Motion sickness adaptation: a neural mismatch model. *Journal of the royal society of medicine* 71, 11 (1978), 819–829.

[57] James T Reason and Joseph John Brand. 1975. *Motion sickness.* Academic press.

[58] Roy A Ruddle and Simon Lessels. 2006. For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological science* 17, 6 (2006), 460–465.

[59] Charles P Schmidt. 1996. Research with the continuous response digital interface: A review with implications for future research. *Philosophy of Music Education Review* (1996), 20–32.

[60] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments* 10, 3 (2001), 266–281.

[61] B Series. 2012. Methodology for the subjective assessment of the quality of television pictures. *Recommendation ITU-R BT* 500, 13 (2012).

[62] Avi Shupak and Carlos R Gordon. 2006. Motion sickness: advances in pathogenesis, prediction, prevention, and treatment. *Aviation, space, and environmental medicine* 77, 12 (2006), 1213–1223.

[63] Namrata Srivastava, Eduardo Velloso, Jason M Lodge, Sarah Erfani, and James Bailey. 2019. Continuous evaluation of video lectures from real-time difficulty self-report. In *Proceedings of the 2019 CHI conference on Human factors in computing systems*. 1–12.

[64] Mohsen Tavakol and Reg Dennick. 2011. Making sense of Cronbach's alpha. *International journal of medical education* 2 (2011), 53.

[65] Meta Technologies. 2022. Oculus Quest 2. https://www.oculus.com/quest-2/. (2022).

[66] Meta technology. 2022. Oculus casting to a screen. https://www.oculus.com/casting. (2022).

[67] Ariel Telpaz, Michael Baltaxe, Ron M Hecht, Guy Cohen-Lazry, Asaf Degani, and Gila Kamhi. 2018. An approach for measurement of passenger comfort: real-time classification based on in-cabin and exterior data. In *2018 21st international conference on intelligent transportation systems (ITSC)*. IEEE, 223–229.

[68] Sebastian Thorp, Alexander Sævild Ree, and Simone Grassini. 2022. Temporal Development of Sense of Presence and Cybersickness during an Immersive VR Experience. *Multimodal Technologies and Interaction* 6, 5 (2022), 31.

[69] Sam Tregillus, Majed Al Zayer, and Eelke Folmer. 2017. Handsfree omnidirectional VR navigation using head tilt. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 4063–4068.

[70] Matthieu Urvoy, Marcus Barkowsky, and Patrick Le Callet. 2013. How visual fatigue and discomfort impact 3D-TV quality of experience: a comprehensive review of technological, psychophysical, and psychological factors. *annals of telecommunications-annales des télécommunications* 68, 11 (2013), 641–655.

[71] Joy Van Baren. 2004. Measuring presence: A guide to current measurement approaches. *Deliverable of the OmniPres project IST-2001-39237* (2004).

[72] Radu-Daniel Vatavu and Jacob O Wobbrock. 2015. Formalizing agreement analysis for elicitation studies: new measures, significance test, and toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1325–1334.

[73] Daniel Zielasko and Bernhard E Riecke. 2021. To sit or not to sit in vr: Analyzing influences and (dis) advantages of posture and embodied interaction. *Computers* 10, 6 (2021), 73.