MazeRunVR: An Open Benchmark for VR Locomotion Performance, Preference and Sickness in the Wild

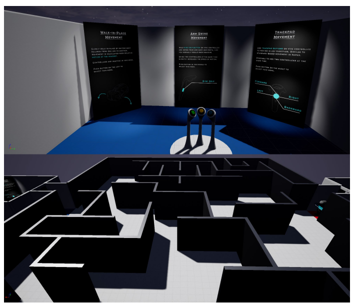

Figure 1: Screenshot of the developed maze navigation game, with the lobby area (top) and the maze itself (bottom)

Kirill Ragozin
Keio University
kirill@kmd.keio.ac.jp

Karola Marky
TU Darmstadt, Keio University
marky@tk.tu-darmstadt.de

Kai Kunze
Keio University
kai@kmd.keio.ac.jp

Yun Suen Pai
University of Auckland
yspai1412@gmail.com

CHI '20 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
DOI: https://doi.org/10.1145/3334480.3383035

Abstract
Locomotion in virtual reality (VR) is one of the biggest problems for large scale adoption of VR applications. Yet, to our knowledge, there are few studies conducted in-the-wild to understand performance metrics and general user preference for different mechanics. In this paper, we present the first steps towards an open framework to create a VR locomotion benchmark. As a viability study, we investigate how well the users move in VR when using three different locomotion mechanics. It was played in over 124 sessions across 10 countries in a period of three weeks. The included prototype locomotion mechanics are arm swing, walk-in-place and trackpad movement. We found that overall, users performed significantly faster using arm swing and trackpad when compared to walk-in-place. For subjective preference, arm swing was significantly more preferred over the other two methods. Finally for induced sickness, walk-in-place was the overall most sickness-inducing locomotion method.

Author Keywords
virtual reality; locomotion; in the wild; framework.

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI);** *Haptic devices;* User studies;

Introduction

VR locomotion is about providing the sense of moving through virtual space, akin to walking or running physically without actually taking up equal amount of physical space. One of the primary issues plaguing VR locomotion is motion sickness, which is caused by our visual system perceiving the optical flow of movement, yet the fluids in our ears do not, creating sensory mismatch. At the same time, VR prioritizes realistic motions and optical flow, which creates a contradiction. To achieve realism, developers need to create a motion input and optical flow that is closest to our physical locomotion in space, but this creates a bigger sensory mismatch; on the other hand, methods like instant teleportation will not be perceived as actual movement and thus reduces motion sickness, yet it is unrealistic which defeats the purpose of using VR. To circumvent this, developers have created a myriad of locomotion methods with mixed results on speed and accuracy, being highly dependant on the content.

Since there are millions of VR users in the current market, we decide to leverage this number to conduct a large-scale study on preference and performance of VR locomotion mechanics by leveraging data from these consumers via an in-the-wild study [12, 9]. In the wild studies have been gaining traction in human-computer interaction (HCI) particularly due to large sample size and a more accurate representation of actual use-case scenarios. To our knowledge, this is also one of the first study to be published for VR researchers in an uncontrolled, in-the-wild scenario [9, 11, 6]. We implemented three common locomotion methods methods; arm swing, walk-in-place, and trackpad movement. These methods will be compared based on three key parameters; speed, simulator sickness and subjective preference. To evaluate speed, we log the completion time of the selected method from each participant. To evaluate simulator sickness, the simulator sickness questionnaire (SSQ) is deployed alongside the game for players to optionally answer before and after any session. Finally, we use the data from play frequency, play time, and subjective feedback to determine the most preferred method. Based on the gathered results, we derived design guidelines for locomotion as well as propose a system to determine user's personal preference and measure their susceptibility to motion sickness to suggest the ideal locomotion mechanic. Our contribution in this paper are the following:

1. We published a VR game that is publicly available under http://mazerun.hcilab.io/ that wraps several established locomotion mechanics into a navigation game to investigate their speed performance, simulator sickness and objective/subjective preference.
2. We found that arm swing performed the fastest and walking-in-place performed the slowest.
3. We found that walk-in-place induced the most motion sickness whereas arms swing was preferable both objectively and subjectively.

Related Work

Two strong considerations for implementing a VR locomotion are the speed it can deliver, as well as how realistic it is, i.e how close does it resemble actual human locomotion and motion in three-dimentional space. For applications that prefer realism, the methods use often rely on gesture detection. One of the most well-known VR locomotion mechanic is walking-in-place, where the user can navigate using a stationary walking motion. This is seen in papers like VR-Step [10], which uses the smartphone's accelerometer and gyroscope to detect head bobbing. Such a method is derived from the logic of keeping the motion to be

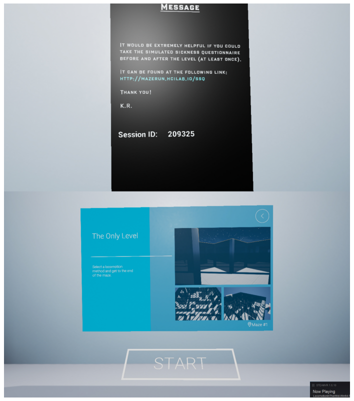

Figure 2: The initial lobby showing basic controls, session ID (top) and start button (bottom)

Figure 3: The difference in motion between (left) trackpad, (middle) armswing, and (right) walk-in-place

as close to actual walking as possible, without actually using the same physical space. This applied to arm swinging as well which is akin to actual walking [8], with the added benefit of consuming less energy than walking-in-place solutions. Methods that prioritize speed on the other hand, compromises on realism. One of the most used method is teleportation, where the user simply points at a position and press a button to teleport there [3]. This method completely eliminates any visual transition in positional change, thus also significantly reduces motion sickness. However, even though teleportation is arguably one of the most used method in deployed applications, we chose not to include it in our game because unlike the other methods, it is an instantaneous transition that breaks the continuity of movement in space.

In-the-wild studies, also known as in-situ, is about conducting a large scale study out of the comfort of a laboratory environment [1]. Such a study can be advantageous as they not only are able to reflect a true use-case scenario of a proposed system, but also allows for gathering of a large sample size. However, ethical concerns need to be addressed, and long term studies require proper planning. Common in-the-wild studies in HCI lies mostly in the realm of public displays and smartphone applications. For example, Claes et al. [2] conducted a study on public display by comparing between in-the-wild, and controlled in-the-wild (CITW), which uses lab-based procedures in an in-the-wild environment. CITW was found to be a viable alternative, though this only applies for hardware research like public displays. Another study placed a physical game machine in the arcade which collected data from 690 participants over a period of a year [5]. The deployment of hardware for in-the-wild studies brings together some novel approaches to conducting studies, such as the participation of the researcher themselves by joining the evaluation process based on different roles [4]. This offers some additional key insight over the proposed system.

Game Development

The entire game was developed using the Unreal Engine. After getting past the main menu and and the starting area of the game shown in Figure 2 and Figure 1 respectively, the procedurally generated maze is loaded. This means that the maze layout is different for each session of the game. The route towards the exit will not take more than 5 minutes, and its procedural nature ensures that the routes cannot be memorized. We choose to keep it short and fun so that users are motivated to perform multiple runs using different locomotion mechanics. The reason we chose to use a maze-like environment for our game is to force the user to move in all four different directions. If they reach a dead end, they will also be forced to make a U-turn. This exposes them to the benefits and limitations of each locomotion method.

At the end of the game, players are presented with a 3-level Likert Scale of their preference for the selected locomotion. We use this as a quick subjective measure of the player's preference.

The three provided locomotion methods are arm swing, walk-in-place, and trackpad movement, with their motions shown in Figure 3.

Arm swing

For arm swing, the user moves depending on the speed of their arm swing. This is achieved by first pressing the grip buttons and observing the relative positions of the VR controllers within a certain time window. The faster the user swings, the faster the locomotion speed until it reaches a maximum threshold.

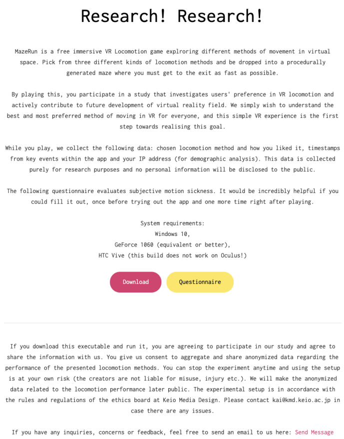

Figure 4: The landing page of the website, detailing the game with the download and questionnaire links

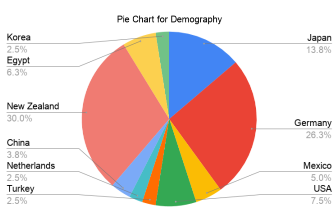

Figure 5: Pie chart for the player demographics

walk-in-place
For walk-in-place, the user is required to walk with both feet while remaining in a fixed position. The movement depends on how fast their head bobs, to simulate walking-in-place. The faster the user walks in place (which in turn, bobs their head), the faster the locomotion speed until it reaches the same maximum threshold as armswing.

Trackpad movement
The trackpad movement is our baseline, where the user simply moves using the trackpad on the Vive controllers, akin to conventional gamepad analog controls. The maximum value for the trackpad, where the finger is furthest from the center dead-zone of the input, corresponds to the maximum speed threshold.

Data Acquisition
After the player completes a session, the logged data is sent to a server that logs them into a comma separated value (CSV) file with one session per line. The logged data includes the unique session identifier, Internet Protocol (IP) address, time the level starts, time the level finishes, selected locomotion method, and selected preference level.

For the SSQ questionnaire, the link is shown on the webpage and also in game, where it brings the player to a Google Form where they fill in the generated session identifier, followed by the SSQ. Even though every player should fill it before and after a session, we made this an optional requirement. We do not control who contributes to the questionnaire as people access the link at their own will.

Publishing the Game
To enable access of the game to a wide audience, we created a website for publicizing the game, where we shared the link across various social media platforms (Facebook, Reddit, Twitter, etc.). On the website, we provide a landing page detailing the gameplay, the purpose of the game, the data being collected, the download link for the game, the SSQ questionnaire, system requirements, consent statement, and comments sections. The description of the game is shown in Figure 4.

Results
MazerunVR was available online starting on the 25th of August 2019 and has an estimated total participants of n = 40 with a total play sessions of n = 80 on the 20th of September 2019. From the logged IP address of the play sessions, 30% of them are from New Zealand, while the second largest player base is from Germany with 26.3%. The remaining countries are Japan (13.8%), USA (7.5%), Egypt (6.3%), Mexico (5%), China (3.8%), Turkey, Netherlands, and Korea(all are 2.5%). This is illustrated in Figure 5.

To evaluate the speed of each locomotion method, we compared between the arm swing (n = 31), walk-in-place (n = 23) and trackpad movement (n = 26). There was a statistically significant difference in the speed between the different locomotion methods, $x^2(2) = 35.985$, $p < 0.001$, with a mean rank score of 53.06 for arm swing, 46.92 for trackpad and 16.3 for walking in place. With Bonferonni correction for adjusted significance, we find that both trackpad ($p < 0.001$) and armswing ($p < 0.001$) was significantly faster than walk-in-place. The plot is shown in Figure 6

For evaluating preference between arm swing (n = 31), walk-in-place (n = 23) and trackpad movement (n = 26), A Kruskal-Wallis H test showed that there was a statistically significant difference in the preference between the different locomotion methods, $x^2(2) = 35.056$, $p < 0.001$, with a mean rank score of 58.06 for arm swing, 33.1 for trackpad and 25.2 for walking in place. Using Bonferonni correction

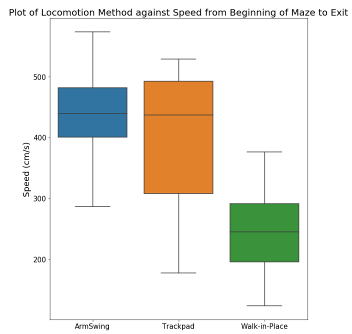

Figure 6: Plot of the speed to navigate from the start to the exit (right)

for adjusted significance, we find that armswing was preferred significantly more than both trackpad ($p < 0.001$) and walk-in-place ($p < 0.001$).

To analyze simulator sickness, we categorize the collected data into induced nausea, oculomotor, disorientation and total score. The sample sizes for the pre-study, armswing, walk-in-place and trackpad are 22, 16, 14 and 18 respectively, aged between 18 to 86 (48 males, Mean: 30.67, STD: 11.457).

There was a statistically significant difference in the induced nausea between the different locomotion methods, $x^2(2) = 13.33$, $p < 0.005$, with a mean rank pain score of 27.7 for the baseline, 27.59 for armswing, 42.53 for trackpad and 47.75 for walking in place. Pairwise comparison showed that only walk-in-place induced significant increase in nausea ($p < 0.05$). We also find the baseline level of nausea to already be relatively high. For oculomotor, there was a statistically significant difference between the different locomotion methods, $x^2(2) = 18.036$, $p < 0.001$, with a mean rank pain score of 22.91 for the baseline, 27.53 for armswing, 42 for trackpad and 51.32 for walking in place. We find that walking-place induced significantly higher oculomotor than both the armswing ($p < 0.05$) and the baseline ($p < 0.005$). For disorientation, there was a statistically significant difference between the different locomotion methods, $x^2(2) = 20.784$, $p = 0.001$, with a mean rank pain score of 23.43 for the baseline, 29.34 for armswing, 45.25 for trackpad and 48.96 for walking in place. We find that both trackpad ($p < 0.005$) and walk-in-place ($p < 0.001$) induced significantly higher disorientation over the baseline results. Finally, there was a statistically significant difference in the total sickness score between the different locomotion methods, $x^2(2) = 17.558, p < 0.001$, with a mean rank pain score of 24.8 for the baseline, 28.66 for armswing, 43.81 for trackpad and 49.46 for walking in place. We find that both trackpad ($p < 0.05$) and walk-in-place ($p < 0.005$) induced significantly higher total sickness score over the baseline results.

Discussion

Even though each method was coded to have the same maximum speed, overall it was easier for armswing to achieve that maximum speed compared to trackpad and walk-in-place. Armswing was also deemed overall more enjoyable to use. Even though trackpad only required thumb movements, armswing strikes a careful balance between fun and speed, which contributes to its overall better performance and higher preference. Walking-in-place was the overall least preferable method, as well as being the hardest to achieve the maximum speed. Among the three methods, only walk-in-place was purely gesture-based and required no button input. Therefore, it led to higher false positives, where even though the user has stopped moving, the system occasionally thinks that the user is still moving forward. Drifting also occurs, where after the user stops, the system required a few milliseconds to recognize this before actually stopping. Since walking-in-place follows head movement, this means that the user's view cannot be independent from the body movement. In scenarios where the user needs to go backwards, this required additional time for walk-in-place since they must first stop they're current motion, physically turn, and continue moving. Armswing and trackpad area easier to use in this case, because they can move in all directions at any time, independent of the head direction.

Regarding simulator sickness, walking-in-place induced the most amount of oculomotor. Oculomotor refers to the nerves responsible for eyeball and eyelid movement. Therefore, the nature of walking-in-place that required continuous head movements caused high oculomotor sickness. Even

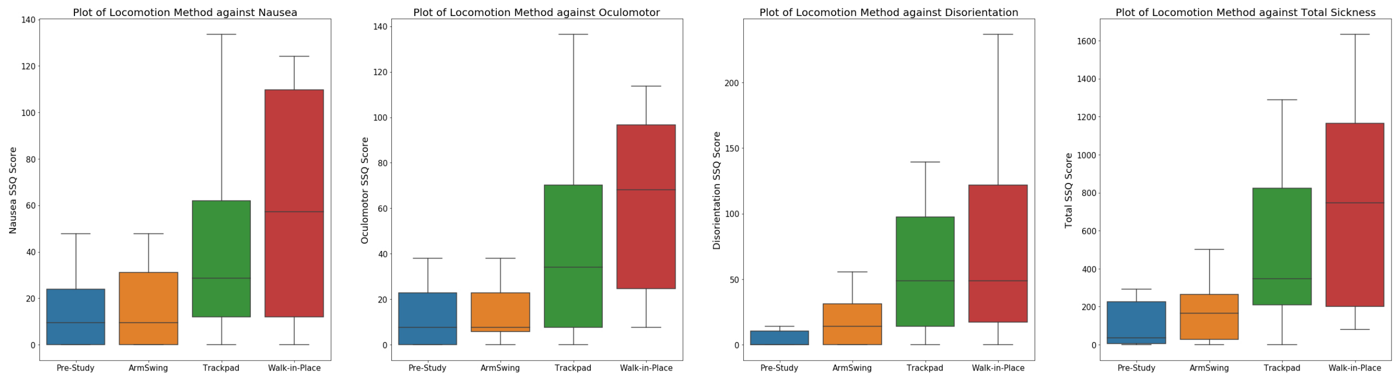

Figure 7: Plot of the nausea (first from left), oculomotor (second from left), disorientation (second from right) and total SSQ score (first from right)

though regular running or jogging also involved periodic head bobbing, other factors such as the weigh of the HMD display could contribute towards higher oculomotor in VR. Nausea, as mentioned previously did not show any significant rise, though we found that this is largely due to the baseline values being already relatively high for the participants. Since this experiment is in-the-wild and we do not control the participants' activities prior to or post experimental session, we cannot determine the cause for sure. One possible reason is that participants may have played the sessions back-to-back without rest, causing a slow rise in nausea before each new session. Regarding disorientation, walking-in-place also induced high disorientation, followed by the trackpad. Disorientation is defined as losing the sense of direction, and walking-in-place's constant head bobbing could potentially cause additional disorientation due to the constant vertical movement. Both the trackpad and armswing methods allow the users to move independently from where they are looking at. However, it is easier to do so with the trackpad which only required thumb movement. Armswing users tend to look towards the direction they are swinging their arms, leading to lesser disorientation.

Conclusion

In this work, we present one of the earliest in-the-wild approach towards understanding users' preference, performance and simulator sickness level for different locomotion methods in VR. We published a VR locomotion game, MazeRunVR, on a website and shared the link to the masses which allows us to collect data from a wider demographic. In the future, we wish to have the game available on an online game platform like Steam for a longer period of time to enable us to collect more data. We would also like to include a secondary task alongside locomotion to test the accuracy of movement, i.e the ability to move to specific locations as fast as possible [7]. We would also like to include more locomotion options for a more in-depth comparison data.

REFERENCES

[1] Alan Chamberlain, Andy Crabtree, Tom Rodden, Matt Jones, and Yvonne Rogers. 2012. Research in the Wild: Understanding 'in the Wild' Approaches to Design and Development. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*. ACM, New York, NY, USA, 795–796. DOI: http://dx.doi.org/10.1145/2317956.2318078

[2] Sandy Claes, Niels Wouters, Karin Slegers, and Andrew Vande Moere. 2015. Controlling In-the-Wild Evaluation Studies of Public Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 81–84. DOI: http://dx.doi.org/10.1145/2702123.2702353

[3] Games Cloudhead. 2015. VR Navigation. (2015). http://cloudheadgames.com/cloudhead/vr-navigation/

[4] Rose Johnson, Yvonne Rogers, Janet van der Linden, and Nadia Bianchi-Berthouze. 2012. Being in the Thick of In-the-wild Studies: The Challenges and Insights of Researcher Participation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 1135–1144. DOI: http://dx.doi.org/10.1145/2207676.2208561

[5] Joe Marshall and Paul Tennent. 2017. Touchomatic: Interpersonal Touch Gaming In The Wild. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. ACM, New York, NY, USA, 417–428. DOI: http://dx.doi.org/10.1145/3064663.3064727

[6] Aske Mottelson and Kasper Hornbæk. 2017. Virtual reality studies outside the laboratory. In *Proceedings of the 23rd acm symposium on virtual reality software and technology*. ACM, 9.

[7] Yun Suen Pai, Zikun Chen, Liwei Chan, Megumi Isogai, Hideaki Kimata, and Kai Kunze. 2018. Pinchmove: Improved Accuracy of User Mobility for Near-field Navigation in Virtual Environments. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18)*. ACM, New York, NY, USA, Article 7, 11 pages. DOI: http://dx.doi.org/10.1145/3229434.3229470

[8] Yun Suen Pai and Kai Kunze. 2017. Armswing: Using Arm Swings for Accessible and Immersive Navigation in AR/VR Spaces. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (MUM '17)*. ACM, New York, NY, USA, 189–198. DOI: http://dx.doi.org/10.1145/3152832.3152864

[9] Anthony Steed, Sebastian Frlston, Maria Murcia Lopez, Jason Drummond, Ye Pan, and David Swapp. 2016. An 'in the wild'experiment on presence and embodiment using consumer virtual reality equipment. *IEEE transactions on visualization and computer graphics* 22, 4 (2016), 1406–1414.

[10] Sam Tregillus and Eelke Folmer. 2016. VR-STEP: Walking-in-Place Using Inertial Sensing for Hands Free Navigation in Mobile VR Environments. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1250–1255. DOI: http://dx.doi.org/10.1145/2858036.2858084

[11] Chenglei Wu, Zhihao Tan, Zhi Wang, and Shiqiang Yang. 2017. A dataset for exploring user behaviors in VR spherical video streaming. In *Proceedings of the 8th ACM on Multimedia Systems Conference*. ACM, 193–198.

[12] Jackie Junrui Yang, Christian Holz, Eyal Ofek, and Andrew D Wilson. 2019. DreamWalker: Substituting Real-World Walking Experiences with a Virtual Reality. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. ACM, 1093–1107.