
Crowdsourcing Virtual Reality Experiments using VRChat

David Saffo

Northeastern University
Boston, MA 02115, USA
saffo.d@husky.neu.edu

Sara Di Bartolomeo

Northeastern University
Boston, MA 02115, USA
dibartolomeo.s@husky.neu.edu

Caglar Yildirim

Northeastern University
Boston, MA 02115, USA
c.yildirim@northeastern.edu

Cody Dunne

Northeastern University
Boston, MA 02115, USA
c.dunne@northeastern.edu

Abstract

Research involving Virtual Reality (VR) headsets is becoming more and more popular. However, scaling VR experiments is challenging as researchers are often limited to using one or a small number of headsets for in-lab studies. One general way to scale experiments is through crowdsourcing so as to have access to a large pool of diverse participants with relatively little expense of time and money. Unfortunately, there is no easy way to crowdsource VR experiments. We demonstrate that it is possible to implement and run crowdsourced VR experiments using a pre-existing massively multiplayer online VR social platform—VRChat. Our small ($n = 10$) demonstration experiment required participants to navigate a maze in VR. Participants searched for two targets then returned to the exit while we captured completion time and position over time. While there are some limitations with using VRChat, overall we have demonstrated a promising approach for running crowdsourced VR experiments.

Author Keywords

Virtual Reality, User Experiments, Crowdsourcing

CCS Concepts

•Human-centered computing → Virtual reality; Usability testing; •Information systems → Crowdsourcing;

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Copyright held by the owner/author(s).

CHI'20 Extended Abstracts, CHI Conference on Human Factors in Computing Systems Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.

[10.1145/3334480.3382829](https://doi.org/10.1145/3334480.3382829)

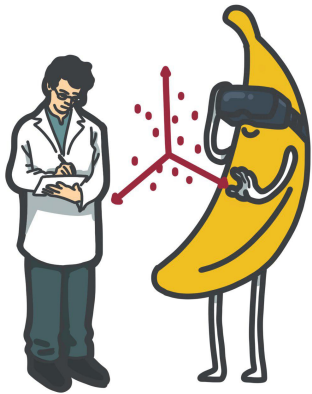


Figure 1: A representation of a VRChat avatar participating in a visualization experiment run by a present experimenter.

Introduction

During the last decade, the use of virtual reality (VR) in human computer interaction (HCI) research has been prolific. The technology of VR head mounted displays (HMDs) has matured significantly since their inception, and they are now capable of providing immersive, comfortable, and convincing experiences. There is now an abundance of VR HMD devices that are widely available to consumers at almost any price point. These factors have driven demand for VR HCI research perhaps higher than it has ever been. However, no research is complete without proper evaluation. User studies are pervasive in HCI and are likewise important for VR research, indeed much of the recently published research relies on user studies [22, 11]. Despite this, VR researchers seeking to evaluate their work face unique challenges that limit the methods available to them for controlled user studies. These challenges include space and/or experimenter availability as well as transport and equipment costs.

The last decade has also seen the rise of the practice of crowdsourcing. Crowdsourcing has been made popular by platforms such as Amazon Mechanical Turk [2, 12, 14, 16] and is widely used for conducting user studies. Researchers can, thanks to these platforms, recruit large and potentially diverse numbers of participants [15] and run them without being bound to a location or a schedule [7]. Unfortunately, though, there currently is no easy way to crowdsource VR user studies, due to the special hardware and computing resources requirement that most participants on crowdsourcing platforms do not meet. Existing platforms also make it difficult to instruct users on how to install the required software in order to participate in a study. Additionally, related sub-domains, such as immersive analytics, have stated the need for evaluating techniques beyond just time and accuracy [6]. Doing so often requires “in

the wild” studies or a greater level of observation and interaction with participants. Traditional crowdsourcing platforms make this difficult, if not impossible, due to limited access to participants. A platform that could provide researchers with easy access to VR equipped users and the flexibility to conduct many styles of controlled user studies would greatly support efforts in evaluating the state of the art in VR HCI, an example of which is described below.

Enter VR Chat

Several social platforms gathering users with VR headsets have been growing. Among these, VRChat is currently regarded as the most popular of these platforms [9]. In VRChat, users connect to a virtual hub, and can then visit a multitude of virtual rooms and socialize with other users present in the same room. Each user is represented by an avatar of their choice. Players can also join the game using the desktop version of VRChat, interacting with mouse and keyboard. In 2018, VRChat claimed to have 4 million total users [19], 30% of whom used a VR HMD. It currently has an average of 7000 daily active users [3]. VRChat also has a special feature that sparked our interest: it allows users to upload custom rooms built with Unity by using a proprietary VRChat SDK. The SDK contains special triggers and event handlers that can be triggered by users, in addition to giving the possibility to upload rooms made of and containing any kind of 3D models made by a creator. We started asking ourselves if we could leverage the vast amount of VRChat users who already own VR equipment and use them as experiment participants by building a custom room that contained the implementation of our experiment, in order to run crowdsourced experiments in VRChat.

Research Goal

We identified the potential for VRChat as a platform for VR experiments given its support for user created content as



Figure 2: The original layout of the maze in [1], that we reproduced in our experiment.



Figure 3: The maze used in our experiment.

well as its large existing user base. The goal of this study is to conduct an initial evaluation of the potential of VRChat to be used as a platform for VR user studies. We seek to answer the following questions:

1. What type of studies could currently be implemented?
2. Can users with VR HMDs be easily recruited?
3. What does an experimental procedure look like on this platform?
4. Can experiments on this platform deliver results?

By answering these questions we can assess whether this method of conducting user studies warrants further exploration.

Related Work

The concept of using crowdsourcing methods for VR user studies has been receiving more attention in recent years. The use of platforms such as Mechanical Turk has become common place for HCI user studies as well as other domains. Existing literature supports this method's validity and usefulness [2, 12, 14, 16]. Researchers, such as Heer and Bostock [4] have even been able to successfully reproduce the results of in lab studies using Mechanical Turk. Mottelson et al. [10] explored the feasibility of running VR experiments out of the lab, while Huber et al. [5] examined crowdsourcing VR experiments with a custom platform. Ma et al. [8] explored the use of Mechanical Turk for VR user studies. The authors surveyed 242 participants for various demographic features including which VR HMD devices they have access to. They report 190 of those participants had access to some sort of smartphone-based VR HMD (Samsung Gear VR, Google Cardboard). Only 18 participants had access to more advanced VR equipment, i.e., an HTC Vive. This study showed that Mechanical Turk is a viable platform for conducting crowdsourced VR studies if researchers implement their experiment targeting smartphone-based HMDs. Unfortunately, this is not always

possible or desirable as many researchers primarily use more advanced VR technologies for their high-resolution displays, positional tracking capabilities, and support of motion-control handheld controllers. Mechanical Turk alternatives, such as Prolific¹, have smaller user bases [13] and are likely to also lack users who own VR headsets.

Our Approach

In this paper our goal is to reflect on how we used VRChat as a crowdsourced experiment platform. Thus we are not evaluating user performance, but rather assessing the feasibility of using the VRChat platform on which the usability study is being conducted. To achieve this, we performed a usability study that consisted of a simple experiment that would allow us to explore and answer our research questions.

Choosing the Experiment

We explored a number of VR experiments that we could reproduce in VRChat. Eventually, we decided to use a subset of an experiment conducted by Basu and Johnsen. [1]. The participants in the experiment were asked to run through a VR maze, find two targets inside the maze, and go back to the exit. The experiment was run using two point of views, immersive and non-immersive, and compared the timing between a group of self-declared gamers and non-gamers. Our reasoning for choosing this experiment over others was that it was simple enough to avoid having too many variables influencing the results, and it would give us a quick way to evaluate the process of conducting a user study on the platform.

Implementation

We implemented the maze using Unity and the VRChat SDK [21]. We reproduced the layout of the maze, includ-

¹<https://www.prolific.co/>



Figure 4: The antichamber of the maze in our implementation of the experiment. The planes displaying the over head view of the maze, used for recording purposes, can be seen in the back of the room.



Figure 5: A researcher (the frog, on the right) recruits a VRChat user.



Figure 6: The portal used to invite users to the experiment.

ing the position of the props that were used in the original experiment to help players orient themselves. In figures 2 and 3, you can see the comparison of our maze and the original maze. The experiment was uploaded as a private room to VRChat. We also created a control system for the maze only known by the researcher, with a reset button that reset the maze to its original state and three planes whose textures render the maze as seen from top, allowing the researcher to see the choice of path made by the participant (Fig. 4). These planes appear as mirrors to other participants, not allowing them to know the layout of the maze beforehand. Our implementation can be found on [github](https://github.com/VisDunneRight/VRChat_Experiment)².

Running the Experiment

In order to recruit participants, a researcher would log into VRChat and join a public room. The researcher would then find other players and ask them to participate in an experiment. When the players accepted, the researcher would open a portal that allowed them to be teleported into the private experiment room. Once there, participants would be briefed about the experimental procedures. Even though multiple participants could be teleported to the experiment, only one person at a time was allowed inside the maze. Once inside the maze, the entrance would be closed behind the participant and would reopen upon them finding the two targets. While the participant was in the maze, the researcher would look at their choice of path through the special mirrors described beforehand. Due to the limitations of the SDK, we could not set up a fixed camera that recorded the maze from inside VRChat, or have a system to time certain checkpoints inside the maze. Therefore the movements of the player had to be saved by recording the screen of the researcher running the experiment. A possible workaround to some limitations is the use of shaders,

which are the only custom script allowed in custom rooms in VRChat.

Results

Using VRChat we successfully recruited and ran 10 participants. After joining a public world, we began by looking for users using HMDs. We did this by asking users directly if they were using VR, or by observing their in-game movements as VR users have full head and sometimes hand tracking. We found that most users we approached were willing and eager to participate. After users had joined our world, they would spawn in a waiting room where we could give them further instructions. At this stage researchers conducting a user study may also present digital consent forms for participants to read and sign (Fig. 7).

After the participants had been briefed they could begin the maze while we observed their progress from the waiting room. To record user times through the maze we simply recorded our desktop screen while looking at the overhead view of the map. All of our participants were able to navigate and complete the maze relatively quickly with an average completion time of one minute fifty second (0:01:50). These times are faster than those reported in the original study [1], however these results should not be directly compared as movement speed and exact dimensions of the maze are not consistent between the two experiments. Many of the participants commented on the maze being very simple and easy, as well as expressing interest in participating in further studies. VRChat allowed us to quickly recruit VR users, execute our experimental procedure, and to record experimental results.

Discussion

In this section we will revisit and reflect on our research questions with what we have learned from our study.

²https://github.com/VisDunneRight/VRChat_Experiment



Figure 7: A representation of two VRChat avatars accepting a consent form.

Studies That Can Be Implemented

At the time of our study, the biggest limitation of VRChat is the relatively narrow scope of what can be implemented within the SDK. The VRChat SDK [21] uses a basic trigger system where objects can be assigned different types of triggers such as on click, intersect, collide, and timeout. Furthermore, a select number of default Unity functions are also allowed. The SDK is capable of supporting studies that use static stimuli or virtual environments, but it largely rules out implementing stimuli with complex interaction of dynamic functionality. However, VRChat developers have released an alpha version of a new SDK that includes a proprietary visual programming language that they have named Udon [18]. This new SDK should be fully released sometime in 2020 and has the potential to greatly expand the possibility of what can be implemented in VRChat [20].

It is particularly exciting to note that VRChat can also be used to implement collaborative VR studies. Previously, such studies would require custom multiplayer platform development. VRChat not only provides an SDK to create worlds but also all the network capabilities to have several concurrent users all in the same virtual space. This includes the ability to invite many users to the same world, voice chat, and interact with other users, i.e., a user passing a ball to another or using hand gestures to communicate.

It is important to keep in mind that VRChat has policies in place to moderate the content users upload. New accounts have to gain a “trusted” level before being allowed to upload content. This trusted level is gained by spending time in the game, interacting with other players, and performing actions. Researchers that open a new account in order to run experiments will have to go through this process, at least for one uploader account. For us this took a few days but there is no advertised limit. In addition, rooms can be either pub-

lic or private. While public rooms can be accessed by anyone at any time, private rooms are invite-only. A new public room uploaded to VRChat requires approval by moderators which can take up to 48 hours. In our case, we opted for a private room which did not require approval. We recruited users through direct invites instead using “portals”, as described below.

Recruiting VR Users

We were able to reliably recruit users with VR HMDs to participate in our study. These users were using a variety of HMDs [17], poses — sitting, standing, or room-scale — and control methods — motion tracking hand-held controllers, keyboard and mouse, or a gamepad. Researchers will need to ask potential participants what they are using to ensure they meet the study’s technology requirements.

Researchers can recruit and run experiments at their own leisure without having to worry too much about the logistics of bringing people into the lab or setting up their own VR equipment. This would be especially helpful for evaluating applications of collaborative VR, as researchers can simply invite multiple users into a virtual room. Currently, recruiting several concurrent in-lab participants and setting up several VR-capable systems is a hindrance to performing collaborative VR evaluations.

We had the most success when approaching small groups of 2–4 people who were not already preoccupied with other activities. VRChat makes inviting users to your world simple and easy. A researcher can drop a “portal” the user can walk into (as we did) or they can invite users directly through the social menu. Most users we spoke to were friendly, respectful, and willing to help us. That said, VRChat is not an inherently professional environment and as a result unprofessional behavior does occur. While this may be an issue for other crowdsourcing platforms as well, it

is more apparent with VRChat as researchers need to be present for recruiting participants and running the study.

Experimental Procedure

We found our experimental procedure on VRChat to be similar to that of a standard in-lab study. With the exception of how participants were recruited in the virtual setting, many of the aspects of the procedure felt familiar. Like an in-lab study, researchers are present to proctor and guide participants through the consent forms, tutorial, experiment, and any exit surveys or debriefings. Even though participants are not physically present in the room with you, the immersive nature of VR makes the experience just as intimate. This immersion also demonstrated to us the value of VRChat as a platform for running collaborative studies.

Producing Results

VRChat does not allow for custom content to communicate with outside services. This means that results cannot be recorded and sent to an external database or downloaded locally. Researchers will need to work around this by implementing into their study a UI that will allow them to externally record results. In our case we set up an in game camera to record the maze from above and we recorded our screen while looking at a plane displaying the overhead view. This allowed us to watch the footage of our participants to record their movement through the maze and completion time. The expanded Udon SDK will help to elevate some of the inconveniences of recording results as it will allow researchers to implement more comprehensive data recording. Researchers will still need UI that will display results of an experiment that they can record externally.

Future Work

Our initial usability study has confirmed that VRChat can be used to run VR user studies. However, there are still

many questions surrounding this method that warrant further study. In this paper, our focus was on reflecting on the process of using VRChat as a crowdsourcing platform for VR experiments. Our experience with VRChat indicates that it is possible to recruit participants from VRChat and collect data. That said, because we did not conduct an actual experiment and compare the results to in lab studies, our findings do not speak to the internal and external validity of results from such crowdsourced studies, which represents a fruitful area of future research. Additionally, given that the new Udon SDK has just been released, we have not had the opportunity to fully test the capabilities of what is experimentally possible on the platform. To address both of these issues, we plan to expand our proposed method to include a more complete experiment and compare the results to in-lab studies. More specifically, using the new Udon SDK, we plan to attempt to reproduce a more complex existing VR user study and examine the validity of the crowdsourced findings compared to this in-lab study. We also plan to implement our own collaborative VR evaluation study to demonstrate the platform's potential for running these logistically complicated studies.

Conclusion

We have shown that it is possible to run a VR usability study using VRChat. We believe that the advantage of aggregating a high number of VR headset owners in a distributed manner is extremely convenient in a research setting. VRChat offers a unique virtual setting and platform that can allow researchers to conduct experiments with methods from both in lab and crowdsourced studies. Although further evaluation of this method is required, our preliminary results and observations give us reason to be optimistic about the viability of this approach. We hope that our experience will help further the research towards crowdsourcing VR studies.

REFERENCES

- [1] Aryabrata Basu and Kyle Johnsen. 2018. Navigating a maze differently - a user study. *CoRR* abs/1805.09454 (2018). <http://arxiv.org/abs/1805.09454>
- [2] Michael Buhrmester, Tracy Kwang, and Samuel D. Gosling. 2011. Amazon's mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science* 6, 1 (23 8 2011), 3–5. DOI: <http://dx.doi.org/10.1177/1745691610393980>
- [3] Steam Charts. 2020. VRChat - Steam Charts. (5 Jan 2020). Retrieved 05 Jan 2020 from <https://steamcharts.com/app/438100>.
- [4] Jeffrey Heer and Michael Bostock. 2010. Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. 203–212. DOI: <http://dx.doi.org/10.1145/1753326.1753357>
- [5] Bernd Huber and Krzysztof Z. Gajos. 2020. Conducting online virtual environment experiments with uncompensated, unsupervised samples. *PLOS ONE* 15, 1 (01 2020), 1–17. DOI: <http://dx.doi.org/10.1371/journal.pone.0227629>
- [6] Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, and Bruce H. Thomas. 2018. *Immersive Analytics*. Springer. Google-Books-ID: vaVyDwAAQBAJ.
- [7] Aniket Kittur, Ed Chi, and Bongwon Suh. 2008. Crwodsourcing User Studies With Mechanical Turk. *Proc. CHI 2008* (01 2008), 453–456. DOI: <http://dx.doi.org/10.1145/1357054.1357127>
- [8] Xiao Ma, Megan Cackett, Leslie Park, Eric Chien, and Mor Naaman. 2018. Web-Based VR Experiments Powered by the Crowd. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. 33–43. DOI: <http://dx.doi.org/10.1145/3178876.3186034>
- [9] Thomas De Moor. 2019. Getting Started with Udon. Medium. (26 Jan 2019). <https://lab.onebonsai.com/social-vr-is-the-weird-future-of-social-media-2fedf4663011>
- [10] Aske Mottelson and Kasper Hornbundefinedk. 2017. Virtual Reality Studies Outside the Laboratory. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST '17)*. Article Article 9, 10 pages. DOI: <http://dx.doi.org/10.1145/3139131.3139141>
- [11] Fares Moustafa and Anthony Steed. 2018. A Longitudinal Study of Small Group Interaction in Social Virtual Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18)*. Article Article 22, 10 pages. DOI: <http://dx.doi.org/10.1145/3281505.3281527>
- [12] G. Paolacci, J. Chandler, and P.G. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5, 5 (2010). <https://ssrn.com/abstract=1626226>
- [13] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology* 70 (2017), 153 – 163. DOI: <http://dx.doi.org/10.1016/j.jesp.2017.01.006>

- [14] David G. Rand. 2012. The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology* 299 (2012), 172 – 179. DOI : <http://dx.doi.org/10.1016/j.jtbi.2011.03.004> Evolution of Cooperation.
- [15] Joel Ross, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. 2010. Who Are the Crowdworkers? Shifting Demographics in Mechanical Turk. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. 2863–2872. DOI : <http://dx.doi.org/10.1145/1753846.1753873>
- [16] Danielle N. Shapiro, Jesse Chandler, and Pam A. Mueller. 2013. Using Mechanical Turk to Study Clinical Populations. *Clinical Psychological Science* 1, 2 (2013), 213–220. DOI : <http://dx.doi.org/10.1177/2167702612469015>
- [17] VRChat. 2019. Control. Docs.VRChat. (8 Nov 2019). Retrieved November 2, 2019 from <https://docs.vrchat.com/docs/controls>.
- [18] @VRChat. 2019a. Getting Started with Udon. Medium. (13 Dec 2019). Retrieved January 5, 2020 from <https://ask.vrchat.com/t/getting-started-with-udon/805>.
- [19] @VRChat. 2019b. Thanks to our community for making 2018 #VRChat's best year yet! Tweet. (18 Jan 2019). Retrieved 02 Nov 2019 from <https://twitter.com/VRChat/status/1086389685268635648>.
- [20] @VRChat. 2019c. VRChat: Unity 2018, Networking, IK, Udon, and More. Medium. (8 Nov 2019). Retrieved November 2, 2019 from <https://medium.com/vrchat/vrchat-unity-2018-networking-ik-udon-and-more-445688972335>.
- [21] VRChat. 2020. VRChat SDK. (2020). <https://docs.vrchat.com/docs/welcome-to-vrchat>
- [22] Yalong Yang, Bernhard Jenny, Tim Dwyer, Kim Marriott, Haohui Chen, and Maxime Cordeil. 2018. Maps and Globes in Virtual Reality. *Computer Graphics Forum* 37, 3 (2018), 427–438. DOI : <http://dx.doi.org/10.1111/cgf.13431>