Immersion at Scale: Researcher's Guide to Ecologically Valid Mobile Experiments

Soo Youn Oh*, Ketaki Shriram, Bireswar Laha, Shawnee Baughman, Elise Ogle, Jeremy Bailenson

Department of Communication, Stanford University

ABSTRACT

While there have been hundreds of psychological studies using virtual reality (VR) over the past few decades, those studies have almost exclusively been conducted in laboratory settings using small samples of college students with little demographic variance. Hence, the generalizability of the results is limited, as not all findings will apply outside the college demographic. In this paper, we present our mobile VR project (*Immersion at Scale*) where we conduct VR experiment sessions in naturalistic settings (e.g., local events, museums, etc.). On average, we were able to collect data from 20-25 people for each 4-hour data collection session of *Immersion at Scale*. We discovered a number of obstacles and opportunities based on bringing VR out into the field. Thus, we do not focus on experimental stimuli and results, but methodological guidelines based on our iterative design improvements from pilot testing.

Keywords: Experimental methodology; Field study; Psychology

Index Terms: HCI design and evaluation methods—user studies; field studies; laboratory experiments; virtual reality

1 INTRODUCTION

The first head-mounted display (HMD), the *Sword of Damocles*, was aptly given its name because it had to be suspended from the ceiling due to its excessive weight [1]. Myron Krueger's *Videoplace* required a minimum of three walls, several projectors, and absolute darkness. In addition to physical constraints, setting up a VR lab was costly. VR was mostly a futuristic topic, tested within the confines of a physical laboratory.

The landscape has changed greatly in the last few years. Major technology companies are currently dedicating a substantial amount of resources to VR. Companies including Facebook (Oculus), Google (Cardboard), HTC (Vive), Samsung (Gear), and Sony (PlayStation VR) showcased their consumer-oriented HMDs and tracking systems at major technology events. These products are a fraction of the cost and weight of their predecessors.

Despite this advancement in technology, much of the research is still being conducted in a select number of VR labs under highly controlled settings with a small number of participants, most of whom are college students. This results in two methodological issues: sample homogeneity and external validity.

To address these limitations, we recently launched *Immersion at Scale (IAS)*, a project with the purpose of gathering a large sample of data from a more diverse population sample in a relatively naturalistic setting.

2 IMMERSION AT SCALE: TEAM & PHYSICAL SETUP

The *IAS* data collection team consists of 5 people: 1 experimenter, 2 research assistants, and 2 participant recruiters. The remainder of this section describes the equipment used in *IAS*.

2.1 Tent

We conducted the experiment sessions in a tent to ensure participant privacy in public areas. There were three factors that were considered in the selection of the tent: size, ease of set-up/take-down, and availability of a partition. We chose a tent that was 13 x 7 x 6.58 feet and included a divider that allowed the partitioning of the tent into a pre-survey and experiment area.

2.2 HMD, Positional Tracker, and Tripod

IAS uses a lightweight HMD with a resolution of 960 x 1080 and a refresh rate of 75 frames per second for each eye. The HMD currently being used is an Oculus Rift DK2. An optical tracking system (Oculus VR™ DK2 Infrared (IR) LED array and IR camera, update rate of 60hz with a 30ms latency rate) and a 3-axis orientation sensor (Oculus VR™ Sensor, update rate of 1000hz with a 30ms latency rate) are used to track the participant's physical head translation (x, y, z) and orientation (pitch, yaw, roll) to appropriately update the rendered viewpoint. A tripod is used to secure the IR camera. The software used to generate and program the world was Worldviz's Vizard.

2.3 Headphones, Headphone Splitter (+ Microphone)

In contrast to laboratories or classrooms, public areas often have high levels of ambient noise. As such, careful considerations should be made if a researcher's VR stimuli include audio instructions or if audio is important for the immersive experience. Even if a noisy environment is not anticipated, it is advisable to use noise-cancelling headphones and a headphone splitter to allow the researcher to provide instructions to the participant via microphone as a precautionary measure.

2.4 Disinfecting Wipes, Furniture, & Participant IDs

Additional components that are crucial—but easy to overlook—are disinfecting wipes and furniture needs (e.g., portable chairs and tables). It is also necessary to prepare a list of randomized participant IDs to ensure participant anonymity.

3 PROCEDURE

While the procedure will differ depending on the nature of the study, this section will provide a general overview of the steps needed to run a VR experiment outside of a laboratory.

3.1 Tent and Equipment Setup

It is recommended that the data collection team arrive at the experiment location at least 45 minutes prior to the experiment start time. After the tent is set up, the experimenter and research assistant(s) can set up the experiment laptop and experiment equipment (HMD, tracking system, headphones, etc.). The participant recruiter(s) can take this time to organize the informed consent forms and participant compensation.

* syoh@stanford.edu

LEAVE 0.5 INCH SPACE AT BOTTOM OF LEFT COLUMN ON FIRST PAGE FOR COPYRIGHT BLOCK

3.2 Participant Recruitment & Informed Consent

After completing the experimental setup, the recruiter(s) invite passers-by to participate in the experiment. Once a person has agreed to participate in the study, the recruiter should obtain informed consent.

3.3 Pre-Survey & Actual Experiment

The pre-survey should be as simple as possible and be designed to take no more than two to three minutes. It is worth noting that field participants typically take significantly more time than lab participants to complete the pre-survey.

Once the participant has completed the pre-survey, the experimenter can administer the actual experiment. Once again, because participants are being recruited on the spot, the researcher should be mindful of minimizing the time needed to complete the study. The experimental stimulus used during the pilot testing of *IAS* lasted for approximately five minutes.

3.4 Debrief and Compensation

After the experiment, the experimenter should thank the participant for his or her time. If there was deception involved, the experimenter should fully debrief the participant regarding the purpose of the study. Once the participant leaves the tent, the recruiter can compensate the participant.

4 Lessons Learned: Points of Concern

4.1 Ambient Noise

One of the first challenges was the unexpected amount of ambient noise. To minimize distractions, we recommend the use of noise-cancelling headphones and a microphone for experimenter-to-participant communication. In addition, it is necessary to carefully assess the noise profiles of candidate venues.

4.2 Different Physical Characteristics

Researchers also need to recognize the different physical capabilities of participants. In the case of *IAS*, for example, there were participants who had trouble standing for an extended period of time without support, people with poor hearing or eyesight, and one person who had mono-vision. In addition, the HMD was sometimes not a good fit for participants who had less prominent supraorbital ridges. This was a truly eye-opening experience for our group as we realized the extent to which studies that focus solely on college students can inadvertently alienate a significant portion of the general population.

4.3 Administrative & Legal Processes

Selecting and securing public venues for data collection involves three key components: finding ideal locations, obtaining official IRB approval, and receiving consent from the authorities at locations of interest. When choosing locations, it is ideal to seek places that are likely to produce a consistent flow of potential participants from diverse backgrounds.

Once locations are approved by the IRB, it is necessary to receive approval from local vendors to set up the experiment. Worth noting is that venues usually charge a fee for space usage, which should be factored early on into the experimental budget.

The final component is compensation. We recommend using gift cards, as they are a fast and reliable method of payment. The guidelines regarding payment of non-citizens vary by institutions, so this should be confirmed with the IRB. In addition, participants should be asked to complete a brief form with their Social Security Number, name and email address for tax audit purposes.

4.4 Range of Movement

As the experiment is being conducted in a public area, the amount of space approved for the study is likely to be limited. The physical movement of the participant, therefore, is constrained. In designing the experiment for *IAS*, we were careful to assess the possible range of movement ('safe zone') within the tent. However, participants (especially children) can make the mistake of attempting to walk outside of this safe zone. The experimenter should thus constantly watch the participant carefully to ensure that he or she does not make sudden movements that could lead to physical injuries or discomfort.

4.5 Hygiene

As a number of participants will be using the HMD, mouse, and headphones in a short span of time, the experimenter should be cognizant of disinfecting them between each participant. If a person is visibly not well, recruiters should abstain from inviting them to participate. In addition, the experimenter should also sanitize his or her hands between participants.

4.6 VR Demos

Finally, we found that providing VR demos for participants facilitated recruitment. In addition to being a good alternative for people who are not eligible for a specific experiment, they can also help friends or family members participants pass the time while they are waiting.

5 Collecting Data with *Immersion at Scale*

One of the greatest benefits of a mobile VR experiment is the outstanding ease with which researchers can collect large samples of data. On average, we were able to collect data from approximately 20-25 participants for each 4-hour data collection session on a single experiment laptop. In addition, most participants had not been exposed to an immersive virtual environment (IVE), and were thus genuinely interested in the experiment. This general enthusiasm regarding VR helps participant recruitment and gives researchers the opportunity to observe how people naïve to and familiar with IVE technology respond to VR.

6 Conclusion

With any new methodology come challenges. When psychologists first started to conduct studies online, a substantial amount of methodology literature was dedicated to best practices in conducting online studies (e.g., [2]). Similarly, as scholars started to explore crowd-sourcing platforms as a venue for research, a number of publications were dedicated to explaining the process and precautions of using this novel research method (e.g., [3]). Similarly, we hope to contribute to the field of VR by offering the guidelines to an impactful methodology that will help researchers conduct studies with a diverse sample population in more naturalistic settings. We believe that this will help our community understand how different people in different environments will respond to VR as the technology begins to make its way to people's everyday lives.

References

[1] J. Cummings and J. Bailenson. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence, *Media Psychology*, 2015.

[2] M. Birnbaum. *Psychological Experiments on the Internet*. Elsevier Press, 2000.

[3] W. Mason and S. Suri. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, 44: 1-23, 2012.