Remarks on Virtual World and Virtual Reality Experiments

Glenn W. Harrison, Ernan Haruvy and E. Elisabet Rutström

December 2010

Forthcoming, *Southern Economic Journal*

Abstract

In recent years there has been increased behavioral research in virtual reality and virtual worlds. While these experiments could offer substantial advantages to researchers, they might also pose risks. We begin by identifying key concepts in virtual experimental research. We next review the critical virtual reality component of virtual worlds. Finally, we offer guidance in conducting virtual world research.

Keywords: Virtual Worlds, Virtual Reality

1. Introduction and Definitions

As humans spend more of their work and leisure time in front of computers, the need for research into the role of *virtual environments* is increasing. The word *virtual* simply means something that is like something else, but without some of the properties of that which it is like. Although that sounds vague, some examples illustrate why one would want a general definition. In the context of computers, the word virtual just means computer simulated. Thus, a virtual environment implies that *something* inherent to the physical environment in which a user is communicating, socializing or gathering information is computer generated or mediated. Virtual worlds are perhaps the most extensive form of virtual environments, involving visual simulations that allow interactions between people in real time. Less extensive forms of virtual environments relax some of these aspects. For example, social networking sites do not include simulations of physical places or require real time interactions, but are still virtual environments where socialization and communications take place.

Virtual reality refers to computer-generated 3D real-time environments where users interact with the simulated environment. These environments can be as visually rich as any movie, but they differ from movies because the actions users take affect the environment. Two important features of these environments are *temporal and spatial realism* (Liang, Shaw and Green (1991), Durlach et al. (2000), Turner and Turner (2006), Mennecke et al. (2008), and Fiore, Harrison, Hughes and Rutstrom (2009)). *Temporal realism* means that events take place in real time and that there are no distracting lags or discrete time lapses. *Spatial realism* means that images are understood by the simulation software as 3D, and that participants can move around and view objects from different perspectives. When the virtual reality is sufficiently "real," users become *immersed* in their virtual experiences: the stimuli from the virtual interactions are dominating their perception and cognition.

What sets virtual *worlds* apart from other virtual *reality* environments is the interaction between human participants. Participants in virtual worlds use an in-world representation, an *avatar*, to engage in any number of activities, including social interactions, commerce, and role play, but also to engage in activities of a more individualistic nature, including artistic creation. As with interactions in the physical world, virtual interactions include production activities by virtual world residents, activities that might alter their environment in a way that affects other residents.

A critical aspect of most virtual worlds is that they are characterized by their persistence: they do not cease to exist when a participant leaves. Thus, events continue to unfold when a user is not present, and the world will have changed during the user's absence.

So far we have described the ways that virtual worlds are similar to the physical world. Virtual worlds differ from physical worlds in several important respects. One is the elimination of real *geographic distance.* Thus, real world residents as far apart as in Melbourne and Uppsala can interact at the same time in the same virtual place through their virtual world resident avatars. From the perspective of an experimental researcher, this allows one to conduct experiments that simultaneously include participants from many different geographic locations.

Virtual worlds also differ from the physical world in that *genetically generated physical distances* are eliminated. Social identity, one's role in society, social norms and opportunities, are largely determined by genetics, such as race, gender, and beauty. Virtual worlds allow individuals to control the appearance of their gender, race, appearance, and other otherwise genetically determined characteristics. The ability to modify appearance can serve to increase experimental control over those features, but may also introduce problems such as misrepresentation of one's true demographic characteristics unless there are other ways to verify these. The anonymity that is possible through virtual interactions may also affect perceptions of social distance, and therefore introduce behavioral effects. Charness, Haruvy, and Sonsino (2007) compare behavior of internet and lab subjects in the "lost-wallet game," and find significantly less trust on the internet.

A final difference between virtual worlds and the physical world is the ability to *temporally coordinate* and to be present at multiple places at the same time. Interactions in the physical world, including those in the experimental laboratory, require participants to be present at the same time. Many virtual interactions, such as email, social networks, blogs, and forums, allow individuals to eliminate the need for coincidence of presence in many cases.

Virtual Reality

The potential benefit to experimental and behavioral research of utilizing virtual reality is that the cues provided are naturalistic, allowing respondents to get immersed in the task in ways that may not be possible using standard text and picture interactions. One area in which task representation is important is the perception of risk and uncertainty. Traditionally, economists have elicited individual risk preferences using choices over lotteries. The common understanding based on research across many years, going as far back as Bernoulli, is that people are most often risk averse. There is, however, a great deal of heterogeneity in risk attitudes across individuals but also across tasks and frames for the same individual.

Sometimes the extent of risk aversion in one settings appears to be different from behavior and revealed preference in other settings, and in other cases inconsistent with expected utility theory. Preference reversals provide one crisp example, but there are others. Moreover, risk aversion is often alleged to be frame-dependent. If true, this is neither appealing nor prescriptive. However, this could be because the choices mathematicians and economists present to individuals are abstract and devoid of context. In naturally occurring settings individuals take calculated risks hundreds of times a day with what seems to be remarkable success. One such example is in driving decisions. Individuals must make hundreds of choices, involving significant risk, including what speed to drive at, how much distance to keep from the vehicle in front, whether to pass a vehicle, whether they can clear a busy intersection in time, and so on. These decisions can reveal a lot about an individual’s risk preference, but they cannot be easily abstracted. Below is a demonstration of a driving simulator used by Dixit, Harrison and Rutstrom (2010) and Andersen et al. (2010) for the purpose of measuring individuals’ risk attitudes through naturalistically generated uncertainty. This driving simulation measures individual choices relating to crash risk or travel time uncertainty. In addition, this simulator environment allows one to study individual route choices in simulated traffic environments. These studies demonstrate that it is possible to build naturalistic virtual reality analogues of the standard, controlled, binary lottery instruments popularly employed in experiments eliciting risk attitudes.

Figure 1. A virtual reality experiment with a driving simulation. Photo courtesy of L-3 MPRI

Another example of virtual reality presentations of uncertain choice scenarios is the wild fire risk management applications in Fiore, Harrison, Hughes and Rutstrom (2009). The controlled burn of parts of forests as a fire management policy avoid future uncontrolled wildfire, and is widely viewed in the fire management field as an effective policy that reduces the risk of large damages. It can thus be viewed as choices with uncertain outcomes, known in economics as ‘‘lotteries’’ or ‘‘prospects’’ that represent a range of final outcomes, each with some probability. Eliciting the true valuations of affected residents’ for controlled burns in their neighborhood may not be realistic in a hypothetical abstract questionnaire, since it does not evoke the same responses as an actual wildfire or controlled burn might. Fiore, Harrison, Hughes and Rutstrom (2009) invested significant attention in the modeling and rendering of virtual trees, forests, and fires, and compared different approaches to visualization in extracting valuations for these probabilistic events through virtual reality simulations of forests and forest fires. Figure 2 illustrates the type of virtual reality environment they developed for this elicitation. They report that risk perceptions are more accurate when using virtual reality than when presenting scenarios using traditional static images and text only.

Figure 2. A virtual reality experiment involving forest fires. Image courtesy of Fiore, Harrison, Hughes and Rutstrom (2009).

Another type of virtual reality research involves online role playing games, particularly focusing on group interactions and group behavior in naturalistic environments. McCabe, Twieg, and Weel (2010) conducted a common pool resource experiment in virtual worlds. They purchased an island on *Second Life* they imaginatively called Hurricane Island. The island is realistic looking, with island-like physical surroundings, and contains eight houses and multiple weather defense stations. These stations defend against hurricanes that frequent the island wreaking havoc on virtual property. Eight islanders live on the island at any one time. Islanders earn money by staying at their virtual homes, but their homes can be damaged by hurricanes. The more damaged their home, the less money they earn. While islanders can repair their homes, this costs them money. Islanders can choose to defend their homes by contributing to a public good in the form of manning one of several weather stations scattered throughout the island. However, this is a costly activity to the attendee with benefits to all islanders. Thus, islanders have strong incentives to free ride, and let someone else defend their homes. Since houses are located at different places on the island, and weather stations are scattered, the equilibria are quite complex. However, the environment here is one of sufficiently high involvement and realism to allow levels of coordination that are not typically observed in the traditional laboratory. Moreover, by observing the communication and interaction between subjects and subsequent actions, McCabe, Twieg, and Weel (2010) are able to decipher the exact process by which coordination comes about. They identify one observed pattern as planning, agreements, information, monitoring and social reward, although this is only one of the observed patterns. This experiment demonstrates the potential of virtual reality, as a component of virtual worlds, to provide evidence of behavior in naturalistic settings, and involving real human social interactions as opposed to abstract laboratory environments.

Guidance in Virtual World Research

The criteria for what constitutes good virtual world research are no different from those applying to laboratory experiments in general: as usual, one needs to worry about internal validity and external validity. Internal validity pertains to the ability of the design to clearly and separately identify the constructs, motives and strategies under investigation. External validity refers to the ability to make general statements about environments and settings outside the laboratory based on laboratory findings.

Internal validity. The first issue for internal validity is making sure the design is incentive compatible. Incentive compatibility means that the incentives in the task presented are aligned with the incentives proposed in the theory being tested. In the presence of unobserved and uncontrolled intrinsic incentives, such as other regarding preferences or risk attitudes, the theory tests are confounded (see e.g. Andersen, Harrison, Lau and Rutstrom (2008)).

It is not clear if these problems will be alleviated or exacerbated with virtual worlds. More likely, better controls in the design of experiments are needed, rather than a different platform. Nevertheless, it may be the case that virtual worlds will make it more difficult to reliably assess demographics and other controls that are important for identification (see Duffy, 2010, in this issue). When recruitment takes place entirely within the virtual world this is likely a serious problem, as discussed below.

A second issue for internal validity is controlling for perception confounds. In an abstract lab environment such perception confounds include the background and natural frames subjects might be bringing with them based on their life experiences and that they then impose on the abstract description (Harrison and List (2004), Harrison and Rutstrom (2001)). The dictator game, for example, is one of the least natural environments because it involves giving money to an anonymous stranger for no particularly good reason. Thus, a subject might imagine a charitable situation or think of the experiment as an environment testing his willingness to "look generous" in some sense, or the subject may be influenced to think of the experimenter as part of the game since the fact that the experimenter designed the artefactual environment becomes salient. Virtual worlds have the ability to bring a field dimension to many experiments, as described in the previous section, in a way that could remove or diminish the need for subjects to apply their own (internal) frames.

Crucial to internal validity is the answer to the question: who are your subjects and what do you know about them? One such concern is gender misrepresentation. Gender swapping, the use of an avatar of the opposite gender, is quite common in virtual worlds (Yee, (2001), Huh and Williams (2009), and Hussain and Griffiths (2008)), with males roughly twice as likely to engage in this practice as females (Yee (2001), Huh and Williams (2009)).

Looks can be deceiving. On the virtual world of Second Life, for example, nearly all avatars are young and attractive, and casual inspection informs us that many people interested in virtual worlds are not. An important purpose for virtual worlds is to create an alternate life, so that differences from one's own life, so that one's characteristics in virtual worlds, projected as well as self-reported, might be expected to be false. Depending on the application and research question, this may or may not be a critical issue (see related concerns in Duffy (2010, this issue). It is certainly an important issue for external validity, however, as we discuss next. A more critical concern pertains to cheating. It is certainly possible for subjects to attempt to participate multiple times. This is also possible in the lab, of course, and experimenters try to keep careful databases to prevent such instances. In the virtual world, subjects may do so relatively costlessly by obtaining multiple avatars. It is therefore important to adopt careful recruiting methods to avoid such problems. One commonly used technique is to recruit only subjects who participate on forums and groups, which they presumably do not do under multiple personalities. Another is to recruit passer-bys (Fullbrun, Richwien and Sadrieh (2009)). The tightest control can be obtained by traditional lab recruitment methods where the experimenter has a chance to verify who the subjects are.

External Validity. One of the most common criticisms of lab experiments is that they may not apply to the real world. That is, it is not clear that college students participating in an experiment with a relatively low cash amount at stake behave the same way as business people, firm managers, and governments. The same criticism can be leveled against virtual world experiments. Moreover, a common criticism is that the virtual world results may not generalize back to the lab! Of course, the larger of the virtual worlds have populations comparable in size to medium-sized countries, presumably with a high degree of demographic heterogeneity, so the concern about generalizability to college students may be misplaced.

The real issue, of course, for all experimenters, whether their environment is the lab, the virtual world, or the field, is sample selection. Simple catalogues of demographic variability say nothing about the insidious effects on inference of sample selection on unobservables. That said, it is important to know who your subjects are in order to extrapolate to other populations, and this simple premiss may be widely violated in virtual worlds, as discussed above. Any claim for greater diversity or some superiority of virtual world subjects should probably be replaced by a discussion of selection issues and the differences of these procedures and sample selection biases relative to student populations or populations of subjects for field experiments.

A second aspect of external validity of great concern in the lab and in virtual worlds is that of naturalistic cues. An important use of virtual research is to generate counterfactual dynamic scenarios with naturalistic field cues and scientific realism. What this means is that subjects need to see an environment that is familiar, or at least recognizable, to them in naturally occurring settings, and then this environment can be perturbed in some controlled manner to generate policy implications. This contrasts sharply with the standard presentation frames of the mainstay technology of experiments. At this point in the discussion, it is important to identify the difference between naturalistic cues and framing. Framing implies setting up an environment that would be suggestive to subjects as to how to behave, making certain aspects of the information relatively more salient and triggering specific decision heuristics; naturalistic cues implies setting up an environment as close to the environment that subjects naturally make decisions in, and giving them those same decisions to make. The latter is a key benefit of virtual research that is not available in laboratory settings.

Table 1. Design Issues in Virtual World Experiments

Internal Validity
Is the design incentive compatible?
Controlling for confounds.
Who are your subjects and what do you know about them?
External Validity
Are the cues natural?
Naturalistic Cues versus Framing.
Are the subjects representative?
Diversity vs. different sample selection process

Conclusion

We love doing experiments in all environments: the lab, the field, and in virtual worlds. But we urge some caution in the excitement over virtual reality and virtual worlds, so that novelty does not mask methodological insights that could not come from simpler, more familiar environments. In this paper we first precisely defined what virtual worlds are, how they are similar to and different from physical worlds and lab environments. Second, we explained and illustrated key aspects of virtual reality, a critical component of virtual worlds, and virtual reality research. Internal validity requires one needs to pay attention to the "R" in virtual reality more than a casual observer might think. Third, we identified the key criteria of establishing quality virtual world research. While these are same criteria as in any experimental research, they translate a bit differently in virtual worlds, and we gave particular attention to issues that concern us. In particular, sample selection issues are ones that affect all experimental research, and just happen to be more severe in virtual worlds. We hope this serves as a valuable roadmap for future research in virtual reality and virtual worlds, as well as the contributions in this issue.

References

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau and E. Elisabet Rutström (2008), "Eliciting Risk and Time Preferences," *Econometrica*, 76(3), 583-618

Andersen, Steffen, Vinayak Dixit, Rami Harb, Glenn W. Harrison, Morten I. Lau, Essam H. Radwan and E. Elisabet Rutström (2010), "Experiments on Driving Under Uncertain Congestion Conditions ", *Unpublished Manuscript*, Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University.

Bernoulli, Daniel (1738), "Specimen theoriae novae de mensura sortis" *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, Tomus V, 175-192, translated as "Exposition of a new theory on the measurement of risk," *Econometrica* 22(1), 1954, 23-26.

Dixit, Vinayak, Glenn W. Harrison and E. Elisabet Rutström (2010), "Risk Attitudes and Risk Perceptions in the Driving Simulator", *Unpublished Manuscript*, Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University.

Durlach, N., G. Allen, R. Darken, R. L. Garnett, J. Loomis, J. Templeman and T. E. von Wiegand (2000), "Virtual Environments and the Enhancement of Spatial Behavior: Towards a Comprehensive Research Agenda," *Presence: Teleoperators and Virtual Environments 9(*6), 593–615.

Fiore, Stephen M., Glenn W. Harrison, Charles E. Hughes and E. Elisabet Rutström (2009), "Virtual Experiments and Environmental Policy," *Journal of Environmental Economics and Management* 57, 65–86

Fullbrun, Sascha, Katharina Richwien and Abdolkarim Sadrieh (2009), "Trust and Trustworthiness in Anonymous Virtual Worlds," *FEMM Working Paper 9033*, Faculty of Economics and Management, Otto-von-Guericke University Magdeburg, 2009

Harrison, Glenn W. (2008), "Neuroeconomics: A Critical Reconsideration," *Economics & Philosophy*, 24(3), 303-344.

Harrison, Glenn W. and John A. List (2004), "Field Experiments," *Journal of Economic Literature*, 42(4), 1013-1059.

Harrison, Glenn W. and E. Elisabet Rutström (2001), "Doing It Both Ways -- Experimental Practice and Heuristic Context," *Behavioral and Brain Sciences,* 24(3), 413-414.

Hussain, Z. and Griffiths, M.D. (2008), "Gender Swapping and Socializing in Cyberspace: An Exploratory Study," *Cyberpsychology & Behavior*, 11 (1), 47-53

Huh, S. and D. Williams (2009), “Dude looks like a lady: Gender swapping in an online game,” In *Online worlds: Convergence of the real and the virtual*, Bainbridge, W. (Ed.). New York: Springer.

Liang, J., C. Shaw and M. Green (1991), “On temporal-spatial realism in the virtual reality environment,” *UIST '91 Proceedings of the 4th annual ACM symposium on User interface software and technology*, 19-25.

MacInnes, I. (2006), “Property Rights, Legal Issues, and Business Models in Virtual World Communities,” *Electronic Commerce Research* (6), 39-56.

Malaby, T. (2006), “Parlaying Value,” *Games and Culture* 1(2), 141-162.

McCabe, Kevin, Peter Twieg and Jaap Weel (2010), “Heterogeneous Dispositions to Cooperate and the Effects of Communication in a Second Life Collective Action Dilemma,” *George Mason University Working Paper*.

Mennecke, Brian E., David McNeill, Matthew Ganis, Edward M. Roche, David A. Bray, Benn Konsynski, Anthony M. Townsend and John Lester (2008) “*Second Life* and Other Virtual Worlds: A Roadmap for Research,” *Communications of the Association for Information Systems*, Volume 22, Article 20, 371-388.

Turner, P. and S. Turner (2006), “Place, Sense of Place, and Presence,” *Presence: Teleoperators & Virtual Environments*, 15(2), 204–217.

Yee, Nicholas (2001), “The Norrathian Scrolls: A Study of EverQuest,” *Unpublished Manuscript*, http://www.nickyee.com/eqt/report.html.