Virtual Reality In Legal Psychology Research: Advancing Jury Similation Research Methodology

BRIAN MULVANY

SIMON FRASER UNIVERSITY

Jury simulation research methodology is often criticized for lacking sufficient generalizability to impact legal proceedings. The implementation of Virtual Reality (VR) technology would benefit the domain of jury simulation research by addressing some of the persistent areas of concern by courts, such as limited ecological validity and generalizability, and by providing insight into the cognition of simulation participants. Other issues include the lack of decisional consequence experienced by simulation participants, and the contrast between the written transcripts often used as experimental stimuli and what is experienced in a real courtroom. If implemented, VR technology would address these issues by providing more realistic stimuli, conveying consequences for the choices made within-simulation by participants, and using dependent measures such as eye tracking and pupillometry. Furthermore, by developing a software application that facilitates the efficient production of virtual environments by researchers with limited technical knowledge, the costs of integrating VR for jury simulation research can be greatly reduced. Overall, when implemented in conjunction with core methodological advancements in the field of jury simulation research, such as drawing participants from jury pool samples, the use of VR as a tool for jury simulation research would allow for a higher level of ecological validity and generalizability than previously achieved.

Keywords: virtual reality, jury simulation, generalizability, technology

Despite decades of simulation research on jury behaviour, criticism from scholars and those within the legal community remains persistent. Investigators and critics alike have noted that the primary shortcomings of jury simulation research are perpetual methodological practices that decrease the ability to generalize experimental findings to real-word courtrooms (Bornstein, 1999; Diamond, 1997; Krauss & Lieberman, 2017; Vidmar, 2008). Methodology has come under scrutiny regarding this domain of investigation because the intent of jury simulation research is usually applied. While some simulation research is done with the intent of illuminating basic psychological processes, it is evident that much of the research is done with the underlying goal of applying findings from simulations to improve the legal system (Bornstein, 1999). Put differently, jury simulation research tends to pragmatically imply that the way participants behave in a study is the way that real-life jury members will behave in a trial. It is this underlying goal that leaves jury simulation research rightfully open to questions about whether the implemented methods result in adequate generalizability (Bornstein, 1999; Christensen, Johnson, & Turner, 2014). For example, in Lockhart v. McCree (1986), 15 jury studies which alluded to an effect between death-qualified jurors and conviction-proneness were rejected by the U.S Supreme Court on the grounds of the research lacking ecological validity (Krauss & Lieberman, 2017).

It is here that a distinction must be made between two related constructs: ecological validity and generalizability. Generalizability, in this context, is the ability to generalize research findings across case types, legal contexts, jurisdictions, and legal actors, whereas ecological validity is narrower in scope, referring to how closely the research mirrors real-world courtroom practice (Krauss & Lieberman, 2017; Vidmar, 2008). Consequently, when highlighting the need for research to generalize to real-world trials, this discussion is alluding to ecological validity, a validity concern that is a prerequisite to generalizability in the broader sense (Bornstein, 1999).

To allow greater generalizability to jury venires and courtroom settings, current jury research needs to address three key components which past methodologies have had difficulty actualizing: (1) implementing more realistic trial simulations, (2) conveying the weight of consequence to participants' decisions, and (3) using reliable dependent measures. As discussed by

Bornstein (1999) and Diamond (1997), the inability of past jury research methodology to address these issues has played a role in an extensive body of simulation research being criticized by those in the legal community. Considering the current issues with past jury simulation research methodology, virtual reality (VR) technology may be a valuable asset when integrated with jury simulation research because it can increase ecological validity through the presentation of more realistic trial simulations. Additionally, VR can implement more reliable and valid dependent measures such as eye tracking, and it may present a notion of consequence for the choices made by participants.

Overview of Virtual Reality

VR systems are technological tools that are becoming increasingly accessible to researchers. According to Harrison, Haruvy, and Rutström (2011), VR technology refers to "computer-generated 3D real-time environments where users interact with the simulated environment" (p. 87). The virtual environment in turn implements visual and auditory stimulation and invokes a sense of spatial presence which facilitates interactions between people and computer-generated characters in real-time (Harrison et al., 2011).

Virtual reality has been used in flight simulators in commercial and military aviation training for decades and is now being utilized in a wide range of research scenarios, from traffic research to attentional research. This increased accessibility of VR technology stems from a rapidly growing consumer interest in VR, and a corresponding increase in commercial production of VR hardware and software. Although still costly to implement in a research scenario, these developments have drastically lowered the cost of such products and increased their availability, making a previously inaccessible technology a viable option for research purposes. For example, early VR systems were marketed at a cost of over $10,000.00 CAD per system (Virtual Reality Society, n.d.). Currently, high-end VR systems are available at a starting point of $1,700.00 CAD for both a headset and computer, although development of software programs currently constitutes the largest expense and is discussed below in more detail (Buzzi, 2018; Greenwald, 2018). While the high accessibility of an experimental material answers some questions regarding its ease of implementation, an in-depth look at the advantages provided by a tool is necessary to answer why it should be chosen over other materials.

Virtual Reality in Jury Simulation Research

Realism

The lack of realistic trial simulations has led to persistent concerns regarding the ecological validity of jury simulation research. In a recent review of the past research, Krauss & Lieberman (2017) con-cluded that among other amendments - such as drawing more realistic participant samples - implementing more realistic stimuli is a key part of increasing the ecological validity and generalizability of jury simulation research. This thought is reflective of the concerns highlighted by Bornstein (1999) who discussed how research settings dissimilar to the courtroom and trial stimuli lacking realism are among the major validity concerns for jury simulation research. The finding that such concerns are still relevant at this point in time reveals that these issues have persisted for several decades. Indeed, the most popular simulation medium, the written transcript, has its greatest limitation in that it has low generalizability to real-world trials (Krauss & Lieberman, 2017). For example, in Free v. Peters (1993), research on jury comprehension that used written stimuli was evaluated as lacking ecological validity by Judge Posner due to it being analogous to a written examination setting, and thus far removed from a trial setting.

Additionally, critics and researchers alike have drawn attention to a feature of written transcripts that further reduces their ecological validity: that they provide participants the opportunity to reread sections of the trial (Pezdek, Avila-Mora, & Sperry, 2010; Rose & Ogloff, 2001). Significantly, this may lead to differences in the comprehension of various trial details between real jurors and mock jurors, further reducing simulation generalizability. In fact, differences have been observed with different trial presentation (e.g., Bornstein et al., 2017; Pezdek et al., 2010). However, it is important to consider that the dependent measures employed have been inconsistent between studies and that the effects of the presentation medium may be different for different dependent measures (Pezdek et al., 2010). Moreover, illuminating the effect of experimental medium on two groups of mock jurors is not equivalent to understanding the effect on comprehension between those exposed to experimental stimuli and those in a real trial. Thus, the issue of low ecological validity for the most commonly used trial stimuli remains unresolved, although the utilization of VR environments may allow for progress in this domain.

By transporting participant jurors to a virtual courtroom where they can see and hear events unfold in real time, VR has the capacity to achieve a level of realism previously unattainable by trial simulations. For example, in a review of several

virtual reality experiments on decision making, a unique feature of VR is described by Harrison et al. (2011) as the ability to present naturalistic cues that allow the participant to become immersed in the task at hand, eliciting the same decision heuristics that would arise in the real-world analog of the virtual environment presented. These naturalistic cues arise from setting up a virtual environment that is familiar or recognizable to participants in the real world, and it is this unique feature of VR that distinguishes it from the laboratory setting (Harrison et al., 2011). This finding may reflect the core strength of utilizing VR for jury simulation research: by immersing participants in a virtual courtroom and eliciting the same psychological processes that are evoked in the analogous real-world environment, the decisions made by mock jurors may be more reflective of those made by a real jury. Furthermore, by bringing a virtual courtroom into the laboratory, realism is gained without having to sacrifice experimental control, since experimenters dictate what stimuli is presented to participants. By contrast, past alternatives to laboratory research, such as field research on jury behaviour, have allowed for increased realism at the cost of experimental control (Bray & Kerr, 1979). In short, VR experiments can facilitate increased ecological validity while retaining the experimental control of the laboratory setting.

Consequences of Decisions

Even with more realistic trial stimuli, one of the most potent criticisms of current simulation re-search on jury cognition and behaviour remains: participants have knowledge that their decision will not affect the fate of the person(s) involved (Bornstein, 1999; Diamond, 1997). This represents a key issue because no previous methodology has been able to portray case scenarios in a way where participants feel any sort of consequential weight for the decisions they make as simulated jurors. This has been the basis for court skepticism toward even the most elaborate simulation research, as a court will always point to the fact that participant jurors are aware that the trial is not real, and that they know their decision will not have fateful consequences for the involved parties (Bornstein, 1999; Diamond, 1997). Although empirical evidence is yet to be seen for a relationship between VR immersion and increased perception of responsibility, there are several features of VR that may allow future simulation research to begin bridging this gap. VR blurs the line between what is reality and what is virtual, and immerses participants in a simulated environment where their choices will affect the fate of virtual people. Furthermore, as this technology continues to rapidly advance, the images and audio presented by VR systems may more closely reflect what is experienced in real life.

Regarding VR's capacity to increase participant's sense of importance for decisions affecting a simulated defendant, an area of concern is that the simulated parties - being 3D models of people and not real humans - will not evoke the level of empathy needed to alter participants' perception of the gravity of their choices. However, empirical evidence by Shin (2018) suggests that the real-time interactions of VR promote perception that the narrative is happening in the present, drawing participants into emotional engagement and stimulating empathetic reactions. In addition, the first-person view of VR elicits a much higher engrossment for users than the third-person view utilized by videotaped trial methods, and such engrossment is further instilled by the sense of spatial presence provided by VR (Skulmowski, Bunge, Kaspar, & Pipa, 2014). Significantly, the sense of spatial proximity to a simulated person is enough to evoke an emotional response from research participants, as indicated by moral dilemma studies where participants who perceived closer spatial proximity to virtual people experienced stronger emotional responses (Skulmowski et al., 2014). Therefore, the temporal realism, spatial proximity, and emotional engagement experienced in VR may work to increase participants' perception of decisional consequence, which is an experience unrealized through past experimental materials. Still, experimental jury simulation research that examines the effect of these features on participant's perception of consequence is needed to determine how effective VR environments are in this regard.

Dependent Measures

One final way that VR surmounts past experimental media is its capacity to employ reliable and valid dependent measures. Questionnaire-based dependent measures have faced criticism from members of the judiciary on the grounds of being akin to a written examination (Diamond, 1997). The significance of this lies in the finding that participants who are adequately capable of performing on a real jury may not fully understand the written questions utilized as dependent measures (Diamond, 1997).

Although this issue may be especially pronounced when drawing samples from jury venires where English is an additional language for a large proportion of the population, it can also arise in native English speakers, since reading ability often lags behind verbal comprehension ability (Carrell, 1991). Consequently, questionnaire-based dependent measures may not always be sufficiently reliable or valid. Contrastingly, VR technology can provide dependent measures that are consistent across populations, while giving insight into automatic processes as they happen. One such measure is eye tracking, a method commonly used in a range of research fields such as linguistics, cognitive psychology, and neuroscience. Eye tracking is considered a highly reliable dependent measure because eye movement programming is critically influenced by subcortical structures, resulting in eye movements that precede conscious, deliberate thought about the stimuli present at the location (Goldberg & Wertz, 1972). Thus, eye tracking measures behaviour that is largely automatic, allowing for reliable measurements across diverse populations and samples, and giving researchers insight to where, what, and when participants are directing their attention. Such dependent measures may provide researchers with deeper understanding of participants' responses to the manipulation of independent variables. Moreover, unlike typical infrared eye tracking devices which have eye tracking as their sole function, VR headsets allow for the collection of data while simultaneously projecting the visual component of a simulation to participants.

Another measurement technique that allows for data collection during a VR simulation is pupillometry, which uses the pre-existing hardware in a VR headset to measure pupil dilation or diameter. As demonstrated in research by Skulmowski et al. (2014), the physiological measure of pupil diameter corresponds to participants' affective arousal or cognitive load. Interestingly, measures of pupil diameter provided indicators of affective responses to experimental stimuli when self-reports purported there were none, demonstrating the utility of this technology to capture data that would not be revealed by means of self-reports (Skulmowski et al., 2014). Thus, with pupillometry capabilities, VR presents yet another dependent measure that is reliable and valid across experimental conditions and participants. Indeed, when used in the context of jury simulation research, measurements that provide data on participants' affective arousal in response to factors such as testimony or prior criminal records will provide a rich psychological tapestry untapped by past methods.

Considerations of Virtual Reality in Research

Ethical

Concerns have been raised regarding the ethical implications of collapsing real and virtual worlds in a research scenario. Madary and Metzinger (2016) described several ethical concerns for the use of VR technology in research, including avoiding harm to participants and describing any risks associated with VR experiences. Therefore, it is the responsibility of researchers to ensure that their stimuli adhere to the principle of non-maleficence and that ongoing, informed consent is obtained from participants. Other such concerns involve the potential mental health risks that have been linked with extensive VR use (Spiegel, 2017). For instance, Aardema, O'Connor, Cote, and Taillon (2010) suggests that with prolonged use, some vulnerable individuals may experience effects that are similar to the symptoms of depersonalization and derealization dissociative disorders (DP/DR), where a sense of detachment to one's thoughts, sensations, actions, and environment are reported. As researchers of psychology, it is imperative that the mental health of participants be protected and prioritized. Therefore, more research is needed to determine the appropriate amount of VR exposure to employ, thereby avoiding prolonged exposure and reducing this risk for individuals who are more susceptible to DP/DR. It is worth noting that it would be a mistake to remove higher-risk individuals from simulations as this would create unequal opportunities to participate in research. Additionally, this practice would bias samples and subsequently reduce generalizability to real juries. Instead, eliminating the risk itself should be a priority for researchers using VR.

Practical

While concerns such as participant behaviour can be addressed during a study's design phase, there remains a logistical concern that must be addressed well before the design phase begins. If researchers desire to create the virtual environments themselves, they will be required to have a knowledge of game engine logic and scripting, 3D modeling and texturing, and possibly animation (Vasser et al., 2017). Learning these skills is an intensive and time-consuming affair, and obtaining proficiency in them often takes several years. The most popular alternative is to hire a professional company to produce the virtual environments necessary for the experiment, which incurs a significant cost. Regarding this, research by Conrad et al. (2015) demonstrated that a lower production quality of dialogue and animation leads to decreased engagement within the virtual environment. Therefore, attempts to decrease the costs associated with production of the environment, characters, and avatars would be detrimental to the immersivity of the experimental paradigm.

However, the cost associated with creating virtual environments can be greatly reduced if an initial investment is made to develop a user-friendly software application that enables researchers with limited technical knowledge to develop and customize their own virtual environments. This initial investment would involve collaboration with those in the field of computer science, game designers, 3D modelers, and animators. The result would be a software toolkit for jury simulation research in VR, similar to that developed by Vasser et al. (2017) for cognitive psychology paradigms. Such an application would include various 3D models that can be manipulated to create unique environments, such as room layouts, the furnishings found in typical courtrooms, and a selection of modelled characters. As opposed to having a company produce the environment from scratch, this type of application would allow for the streamlined production of virtual environments by researchers with limited technical knowledge, thereby greatly reducing the financial expediture associated with the production of environments for VR.

Experimental

Despite technological advances that give VR the capability to present visual stimuli that are increasingly similar to what is seen in real life, participants are still able to discern that the virtual environment is in fact virtual and not reality. This raises the question of whether a participant will behave uncharacteristically due to the perception that they are in a simulated world, resulting in decisions dissimilar to what jurors would make in real-life. Accordingly, research by Yee and Bailenson (2007) indicates that behaviour in a virtual environment which deviates from participant's typical real-life behaviour does indeed occur. Yee and Bailenson (2007) established that atypical behaviour is due to deindividuation of the participant, which is when a participant takes on the visible characteristics of a digital self-avatar, and the corresponding stereotypes associated with those characteristics.

Nevertheless, since this issue stems from a mismatch of self-representation due to the self-avatar possessing different visible identities than the participant, the solution is rather straightforward: reducing or eliminating this mismatch. One method is to create self-avatars that accurately match the visible identities of participants, such as ethnicity and gender. Another more pragmatically appealing method is to completely eliminate the presence of a self-avatar, thus removing the possibility for participants to observe mismatching cues of self-representation. Yet, research shows that the presence of a visible body which represents the participant in the virtual environment leads to greater immersion in the simulation, thus revealing a considerable downside to this solution (Steed et al., 2016). Ultimately, either of these solutions would prevent participants from unknowingly altering their behaviour to fulfill stereotype expectations of the self-avatar, thus allowing for behaviour and choices that are more congruent with those of real life, and crucially, increasing the generalizability of research findings.

Conclusion

Although VR technology can improve the generalizability of research claims, by itself, it is not enough for jury simulation research to have high external validity. To achieve the highest level of generalizability, researchers must continue to apply and expand on the key progressions that have been made in the domain of jury simulation research, as explained by Diamond (1997) and Krauss and Lieberman (2017). That is, they must continue to draw participants from jury pool samples, implement jury deliberation, use dichotomous dependent measures as response items such as guilty-not-guilty verdicts, and test for consistent results across a variety of simulated cases. These are critical practices because student samples have been shown to differ compared to those taken from jury venires in ways that can affect research results. Similarly, a lack of simulated jury deliberation can introduce questionable results, since research reveals that deliberations can influence outcome. Likewise, a failure to test for similar results across case types raises inevitable questions about generalizability. Hence, maintaining these practices is essential for the generalizability of VR-based research.

In conclusion, for the domain of jury simulation research to benefit from the methodological improvements that advances in technology can bring, VR technology should be implemented as an experimental medium. Due to the capacity of VR to facilitate naturalistic cueing, spatial proximity with simulated people, and immersive experiences through real-time interactions, VR can provide more realistic trial simulations and may be able to invoke the weight of decisional consequence for participants. Additionally, owing to its eye tracking and pupillometry capabilities, VR is an advantageous way to incorporate reliable, valid, and psychologically insightful dependent measures. Utilizing VR in research methodology would allow for the acquisition of more generalizable results while peering into untapped areas of investigation.

References

Aardema, F., O'Connor, K., Cote, S., & Taillon, A. (2010). Virtual reality induces dissociation and lowers sense

of presence in objective reality. Cyberpsychology, Behavior, and Social Networking, 13, 429-435. doi:10.1089/cyber.2009.0164

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? Law and Human Behavior, 23(1), 75-91. doi:10.1023/A:1022326807441

Bornstein, B. H., Golding, J. M., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2017). Mock juror sampling issues in jury simulation research: A meta-analysis. Law and Human Behavior, 41(1), 13-28. doi:10.1037/lhb0000223

Bray, R. M., & Kerr, N. L. (1979). Use of the simulation method in the study of jury behavior: Some methodological considerations. Law and Human Behavior, 3, 107-119. doi:10.1007/BF01039151

Buzzi, M. (2018, June 26). The best computers for the Oculus Rift VR headset. PC Magazine. Retrieved from https://www.pcmag.com/

Carrell, P. L. (1991). Second language reading: Reading ability or language proficiency? Applied Linguistics, 12(2), 159-179. doi:10.1093/applin/12.2.159

Christensen, L. B., Johnson, R. B., & Turner, L. A. (2014). Research methods, design, and analysis (12th ed.). Upper Saddle River, NJ: Pearson Education.

Conrad, F. G., Schober, M. F., Jans, M., Orlowski, R. A., Nielsen, D., & Levenstein, R. (2015). Comprehension and engagement in survey interviews with virtual agents. Frontiers in Psychology, 6. doi:10.3389/fpsyg.2015.01578

Diamond, S. S. (1997). Illuminations and shadows from jury simulations. Law and Human Behaviour, 21(5), 561-571. doi:10.1023/A:1024831908377

Free v. Peters, 806 F. Supp. 705 (N. D. Ill. 1992); rev'd, 12 F.3d 700 (7th Cir. 1993), cert, denied, 513 US 967 (1994).

Goldberg, M. E., & Wurtz, R. H. (1972). Activity of superior colliculus in behaving monkey. II. Effect of attention on neuronal responses. Journal of Neurophysiology, 35, 560-574. doi:10.1152/jn.1972.35.4.560

Greenwald, W. (2018, June 22). Oculus Rift vs. HTC Vive: Which virtual reality headset is best? PC Magazine. Retrieved from https://www.pcmag.com/

Harrison, G. W., Haruvy, E., & Rutström, E. E. (2011). Remarks on virtual world and virtual reality experiments. Southern Economic Journal, 78(1), 87-94. doi:10.4284/0038-4038-78.1.87

Krauss, D. A., & Lieberman, J. D. (2017). Managing different aspects of validity in trial simulation research. In M. B. Kovera (Ed.), The psychology of juries (pp. 185-205). Washington, DC, US: American Psychological Association. http://dx.doi.org/10.1037/0000026-009

Lockhart v. McCree, 476 U.S. 162 (1986)

Madary, M., & Metzinger, T. K. (2016). Real virtuality: A code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology. Frontiers in Robotics and AI, 3. doi:10.3389/frobt.2016.00003

Pezdek, K., Avila-Mora, E., & Sperry, K. (2010). Does trial presentation medium matter in jury simulation research? Evaluating the effectiveness of eyewitness expert testimony. Applied Cognitive Psychology, 24(5), 673-690. doi:10.1002/acp.1578

Rose, V. G., & Ogloff, J. R. P. (2001). Evaluating the comprehensibility of jury instructions: A method and an example. Law and Human Behavior, 25(4), 409-431. doi:10.1023/A:1010659703309

Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? Computers in Human Behavior, 78, 64-73. doi:10.1016/j.chb.2017.09.012

Skulmowski, A., Bunge, A., Kaspar, K., & Pipa, G. (2014). Forced-choice decision-making in modified trolley dilemma situations: A virtual reality and eye tracking study. Frontiers in Behavioral Neuroscience, 8. doi:10.3389/fnbeh.2014.00426

Spiegel, J. S. (2017). The ethics of virtual reality technology: Social hazards and public policy recommendations. Science and Engineering Ethics, 24(5), 1537-1550. doi:10.1007/s11948-017-9979-y

Steed, A., Friston, S., Murcia-López, M., Drummond, J., Pan, Y., & Swapp, D. (2016). An “in the wild” experiment on presence and embodiment using consumer virtual reality equipment. IEEE Transactions on Visualization and Computer Graphics, 22(4), 1406-1414. doi:10.1109/TVCG.2016.2518135

Vasser, M., Kängsepp, M., Magomedkerimov, M., Kilvits, K., Stafinjak, V., Kivisik, T., … Aru, J. (2017). VREX: An open-source toolbox for creating 3D virtual reality experiments. BMC Psychology, 5. doi:10.1186/s40359-017-0173-4

Vidmar, N. (2008). Civil juries in ecological context:

Methodological implications for research. In B. H. Bornstein, R. L. Wiener, R. Schopp, & S. Willborn (Eds.), Civil juries and civil justice: Empirical perspectives (pp. 35-65). doi:10.1007/978-0-387-74490-2_3

Virtual Reality Society. (n.d.). History of virtual reality. Retrieved from https://www.vrs.org.uk/virtual-reality/history.html

Yee, N., & Bailenson, J. (2007). The proteus effect: The effect of transformed self-representation on behavior. Human Communication Research, 33(3), 271-290. doi: 10.1111/j.1468-2958.2007.00299.x