

Bridging Realities: How Perceptions of Refugees Are Influenced by VR Interactions

Andrey Kirpach

(Alumnus, Faculty of Law, Hitotsubashi University)

November 10, 2023

Abstract

Does virtual reality (VR) possess beneficial qualities as opposed to traditional media? This study examines the effects of viewing a documentary about refugees as an empathy – inducing stimulus and compares the strength of the effects based on the medium used: VR versus a computer screen. An experiment conducted as part of this research could not find evidence that VR is more effective at eliciting empathy for “imagine-other” perspective-taking tasks. On the other hand, the type of perspective-taking itself may be a significant factor. The discussion section synthesizes the experimental results with a critique of VR experiences in international relations contexts from a critical culture studies perspective and highlights how VR experiences are shaped and limited by the power structures within which they are produced.

Conflict of Interest Statement

At the time of writing, the principal investigator, Andrey Kirpach, held a long-term paid internship with Nangok R/Studios Co., Ltd., and had received a job offer from the company. The company specializes in producing augmented reality content and has produced virtual reality content in the past.

In this light, there exists a possible conflict of interest for the principal investigator. However, while the company was aware of this research, it provided neither financial support nor consultation regarding the content of the research. Furthermore, the internship and the job offer are in no way related to the outcome of the research or affected by it.

Financial Disclosures

This research project has received no funding from Hitotsubashi University, nor any other legal body or private individual other than the principal investigator. Expenditures for the equipment, as well as remuneration for study participants, were covered by the principal investigator.

Ethical Considerations

This research has received approval from the Hitotsubashi University Research Ethics Examination Committee, Approval number: 2019C009. The ethics review considered conflict of interest, experiment participant informed consent, experiment participant remuneration, and effects of VR experience on experiment participants, and other issues.

1. Introduction

In recent years, there has been a dramatic growth in virtual reality (VR) experiences that tackle social and international issues, such as life in refugee camps or women's education in Afghanistan (Oculus n.d.), or that provide access to otherwise inaccessible places, such as North Korea (see Geere 2017).

The United Nations has established a VR production program and has used VR films to elicit financial donations from policymakers and influencers (UNVR n.d.b; Gülerk & Kasulke 2018; Rose 2018).

The program description states that VR promotes empathy towards depicted groups in those who experience it. This increase in empathy is believed to improve attitudes towards the group, and stimulate prosocial behavior, for example favorable policy change and financial donations (UNVR n.d.c).

The effects of VR on empathy, as well as the relation between VR experiences, empathy, and altruism, are an area of great academic interest (Ahn et al. 2013; Herrera et al. 2018; Martingano et al. 2021).

This research examines the effects of using virtual reality versus a computer screen when viewing a documentary about refugees and considers the implications of these effects in the context of international relations.

2. Theoretical Framework

What Is Empathy?

Literature reviews reveal the absence of an agreed-upon definition of empathy. Examining empathy from the perspective of various scientific disciplines has produced a number of definitions that, while having some overlaps, nevertheless are not completely interchangeable (Batson & Ahmad 2009; Gerdes et al. 2010; Neumann et al. 2015).

This research draws on a categorization provided by Batson and Ahmad (2009). In an overview of existing empathy-related literature, the authors outline four distinct psychological states that are all referred to as "empathy" in various contexts (Table 2). Batson and Ahmad group these based on the type of psychological state, dividing them into cognitive

states and affective states. The former focuses on the mental processes of imagining oneself or others being placed in various situations. The latter focuses on the emotional response to a situation, or to the feelings of another person.

Another widely used categorization, which is also employed in this research, emphasizes whether empathy is measured and compared between different individuals, or within the same individual at different times. Thus, empathy can be thought of as a trait which varies from person to person (dispositional empathy), or as a momentary affective response to a specific stimulus (situational empathy). An individual’s propensity towards empathy can influence their empathetic response to a given situation. So, dispositional empathy (also known as trait empathy) is a mediator for situational empathy (Davis 1983b).

Table 1: “Four Psychological States Called Empathy in the Intergroup Relations Literature,” from Batson and Ahmad (2009, p. 144)

Type	Psychological state	What the state involves
Cognitive/ perceptual	1. Imagine-self perspective	Imagining how one would think and feel in another’s situation or shoes.
	2. Imagine-other perspective	Imagining how another person thinks or feels given his/her situation.
Affective/ emotional	3. Emotion matching	Feeling as another person feels.
	4. Empathic concern	Feeling for another person who is in need.

Table 2: Situational Empathy vs. Dispositional Empathy

Dispositional (trait) empathy	Situational empathy
A tendency to imagine the feelings of others	An affective response to another’s state

The frameworks of definition covered above are not exhaustive and are provided as examples of how empathy is conceptualized. Synthesizing research from social and natural sciences to produce a more unified view of empathy is an ongoing effort (Gerdes et al. 2010).

Empathy as a Source of Altruism

Prior research primarily led by Daniel Batson and his followers has amassed an extensive body of work conceptualizing empathy and distinguishing it from similar processes (Batson & Ahmad 2009). These works have explored its potential as a source for an individual's altruistic behavior (Batson et al. 1981; Schroeder et al. 1988) and have considered whether eliciting empathy leads to improvement of intergroup relations (Batson et al. 1997; Batson & Ahmad 2009).

3. Prior Research

VR, Empathy, and Prosocial Behavior

Early work by Yee and Bailenson (2006) explored the use of VR to reduce stereotyping of the elderly, suggesting that immersing oneself into a VR avatar of an elderly person had a significant effect on improving negative associations towards the elderly. While the study has its theoretical basis in perspective taking, it did not employ measures of empathy.

Ahn et al. (2013) compared traditional verbally induced perspective-taking (PT) with a VR experience designed to simulate color blindness. They investigated differences between the PT and VR groups with regard to 1) sense of "oneness"; 2) attitudes towards color blind people; and 3) helping behavior. They found that these three parameters had larger values for individuals from the VR group. Furthermore, they found that the positive results obtained in the lab were transferable into real-life scenarios.

Schutte and Stilinović (2017) utilized "Clouds Over Sidra," a 360° documentary about refugees, to examine the effects of viewing it on a participant's "engagement" and their empathy directed at the protagonist. They found that both engagement and empathy were significantly higher in the VR condition than the screen condition.

A study by Herrera et al. (2018) investigated both immediate and long-term effects of VR and traditional perspective-taking task interventions. They found that participants who took part in VR perspective-taking tasks

“had more positive attitudes and signed a petition supporting helpful initiatives toward the homeless at significantly higher rates” than participants from traditional PT tasks. They employed an imagine-self perspective task, and found that it induced stronger levels of oneness, empathic concern, and personal distress than traditional PT, which matched the predictions from Batson and colleagues’ theorizing. Their study also investigated effects of perspective-taking on donation levels. However, they did not account for the financial reward to the participants. Therefore, while participants in the VR conditions donated more money than those in the other conditions, it is unclear whether these results can be trusted, as the participants may not have had a sense of attachment to the money.

Gürerk and Kasulke (2018) improved on previous findings by conducting a large-scale investigation specifically targeting the effects of VR and on charitable giving. In a similar vein to Herrera et al. (2018), Gürerk and Kasulke employed an imagine-self perspective-taking task. They measured the difference in attitudes towards refugees in the VR and traditional PT conditions. They found that VR did not lead to higher donations than a control treatment showing the same video on a 2D computer screen.

4. Hypotheses

Drawing from existing literature on the subject, this research puts forth the following hypotheses.

Hypothesis 1 (VR vs. Screen): *Viewing the documentary in VR will elicit stronger empathic responses than viewing via a traditional computer screen.*

Previous research has repeatedly suggested that utilizing media with increasing levels of immersion produces a greater sense of co-presence (i.e., the portrayed characters feel more “real” and “present”). This heightened sense of immersion has been connected to greater empathic concern, personal distress, and sense of “oneness” with the portrayed characters — all proxy measurements for empathy.

Hypothesis 2 (Perspective-taking type & Interactivity): *An “imagine-self” perspective-taking task with interactivity will elicit stronger empathic responses than an “imagine-other” task with little interactivity.*

Herrera et al. have theorized that the type of perspective-taking task may influence the strength of empathic responses produced (Herrera et al. 2018, p. 32).

Prior research has confirmed that “imagine-self” tasks are more effective at eliciting empathic responses than “imagine-other” tasks using an audio medium (Myers et al. 2014). This research verifies those results for perspective-taking tasks in VR environments.

5. Research Methods

5.1 Overview

Following Herrera et al. (2018), in order to verify the hypotheses, an experiment was conducted among university students. Students were asked to view the documentary using either VR or a computer screen. After the intervention, they were asked to answer a survey measuring their empathy.

5.2 Procedure

1. The participant is asked a series of demographic questions.
2. The participant is then assigned into one of two groups (“Screen” and “VR”).
3. As an intervention, the participant is shown “Clouds Over Sidra,” a 360° video documentary depicting the life of a Syrian refugee girl.
 - Participants in the “Screen” group are shown the documentary on a computer screen (as a 360° video). They can change their angle of view using a mouse.
 - Participants in the “VR” group are shown the documentary via a VR device. They are able to change their angle of view by rotating their head.
4. Participants are asked to complete a post-intervention survey measuring outcome variables.

5.3 Participants

Thirty participants were randomly selected from students at Hitotsubashi University. In order to better control for intra-group variance, the selection was limited to Japanese students from the bachelor division.

5.4 Intervention

Stimulus: Clouds Over Sidra

This research used “Clouds Over Sidra” for the following reasons:

1. The film is championed by the UNVR campaign and affiliated VR producers as a pioneer work exemplifying VR’s utility as an “empathy machine.”
2. The UN has cited the film as the cause of the increase in donations at pledging conferences.
3. There is existing scientific and popular literature critically examining the constructed images of refugees in the film, and the relationship between the viewer and the depicted group (Irom 2018; Kool 2016; Nash 2017), which however lacks a quantitative side to how the VR experience influences the attitudes of the participants.
4. There is some prior research into the film’s empathy-inducing effects, which used a different methodology from this study (Schutte & Stilinović 2017).

Equipment

The following equipment was used during the intervention:

1. A 13.9-inch screen laptop was utilized for the viewing of the documentary (Group A, Screen). During the viewing, earphones were utilized for sound output.
2. A 5.1-inch Android-based smartphone was utilized in combination with a VR head-mounted display (HMD) for the viewing of the documentary (Group B, VR). The HMD, when connected to the smartphone’s audio output jack, produced location-aware stereo sound from the

headphones attached to the HMD.

5.5 Post-Intervention Survey

The survey contents were adapted from Herrera et al. (2018) with the following changes:

Firstly, the questionnaire section “n” regarding blatant dehumanization was removed. This was done because 1) the original research failed to find significant results for dehumanization directly after intervention; 2) there was a fear that the question might provoke dishonest or random answers, given that the research participants were university students; and 3) the question has little relevance to the research question and theoretical framework used by this research.

6. Variables

6.1 Population Variables

Table 3: Control Variables

Group	Variable
Demographic Variables	Age Gender
Dispositional Empathy	Beliefs About Empathy Scale Interpersonal Reactivity Index (IRI)

For each participant, the following variables were included as population control.

Age and Gender

Inclusion of these variables was motivated by previous research, which has indicated that these factors may have a significant impact on situational empathy (e.g., Güreker & Kasulke 2018, p. 15).

Beliefs about Empathy & Interpersonal Reactivity Index (IRI)

These two measures were included in the study following methodology in Herrera et al. (2018). They are employed as measures of dispositional (trait) empathy and used as control variables between the two groups (“Screen” and “VR”). IRI was originally proposed in (Davis 1983a). Note that this research follows Herrera et al. (2018) in omitting the Fantasy (FS) scale from IRI. Only the Empathic Concern (EC), Personal Distress (PD), and Perspective Taking (PT) scales were included.

6.2 Outcome Variables

For each participant, the following variables were included as outcome (dependent) variables:

Table 4: Outcome Variables

Group	Variables
Situational Empathy	Empathic Concern (EC) Personal Distress (PD)
Indirect Empathy-related Indicators	Inclusion of Other in Self (IoS) Social Presence
Prosocial Outcomes	Attitudes Towards Refugees

Empathic Concern and Personal Distress

Originally developed in Batson, Early, and Salvarani (1997), the measure was included in this study for consistency with Herrera et al. (2018). It features questions intended to measure situational Empathic Concern (EC) and situational Personal Distress (PD).

Inclusion of Other in Self (IoS)

Originally developed in Aron et al. (1992), the measure was included in this study for consistency with Herrera et al. (2018). It is included to measure self-other merging in experiment participants with regard to the target out-group (refugees), as a result of watching the documentary. Self-other merging is hypothesized to be an indirect indicator of empathy.

Social Presence

This measure was originally developed in Nowak and Biocca (2003), and was included in this research following work by Herrera et al. (2018). Participants were asked to indicate the extent to which they felt that the people portrayed in the documentary were real and present in the same environment as the participants.

Higher feelings of Social Presence have been linked to higher levels of empathy, so the Social Presence scale was included as an indirect measure of empathy.

Attitudes Towards Refugees

This measure was originally developed in Batson et al. (1997), and then adapted in Herrera et al. (2018) to measure attitudes toward the homeless. It should be noted that Gürerk and Kasulke (2018) also employed an adaptation of the Batson et al. scale to measure Attitudes Towards Refugees, however the adaptation in this paper was developed independently of the work by Gürerk and Kasulke.

6.3 Explanatory Variables

The following measures were chosen as explanatory (independent) variables. Each one of the explanatory variables in this section corresponds to one of the hypotheses (1 or 2).

Media

This measure represents whether the participant was in Group A (Screen) or Group B (VR) and is used to verify Hypothesis 1.

Perspective-taking Type

For each participant who participated in this study, the value of this variable was “imagine-other,” as the participants viewed a documentary describing the lives of refugees and did not themselves have to take on the part of refugees.

Separately from the data collected in this study, the survey results from Herrera et al. (2018) were accessed, and altered in the following way:

1. Responses from individuals who did not participate in perspective-taking tasks at all were stripped.
2. A new column representing the Perspective-taking Type was added, with the value set to “imagine-self.” This was done as the participants in that research were asked to take on the role of homeless people.

Finally, the altered data set from Herrera et al. was then joined with the data set collected as part of this research. The resulting set was then used to verify Hypothesis 2.

Table 5: Hypothesis Operationalization

Hypothesis	Expectation
VR vs. Screen	$EC_{vr} > EC_{screen}$ $PD_{vr} > PD_{screen}$ $IoS_{vr} > IoS_{screen}$ $Presence_{vr} > Presence_{screen}$ $Attitudes_{vr} > Attitudes_{screen}$
Imagine-self vs. Imagine-other	$EC_{self} > EC_{other}$ $PD_{self} > PD_{other}$ $IoS_{self} > IoS_{other}$ $Presence_{self} > Presence_{other}$

>: There is a statistically significant difference in means between the groups.

7. Expected Results

Hypothesis 1: VR vs. Screen

Hypothesis 1 predicts that situational Empathic Concern (EC), as well as indirect indicators of empathy (self-other merging expressed by IoS, and Social Presence) should be higher for participants in Group B (VR) than for those in Group A (Screen).

Furthermore, Personal Distress (PD) is expected to be higher for Group B (VR) than for Group A (Screen). This is attributed to the higher levels of immersion that virtual reality provides.

Finally, Attitudes Towards Refugees are expected to be higher in Group B (VR) than in Group A (Screen), owing to the expected higher levels of empathy.

Hypothesis 2: Imagine-self vs. Imagine-other

Hypothesis 2 predicts that an “imagine-self” perspective-taking task with higher levels of interactivity will elicit stronger levels of empathy than an “imagine-other” task with lower levels of interactivity.

Thus, Empathic Concern (EC), Personal Distress (PD), self-other merging (IoS) and Social Presence are expected to be higher for those who participated in the experiment in Herrera et al. (2018), irrespective of the employed medium, than for those who participated in this study.

The Attitudes Towards Refugees variable was excluded from this analysis, as the data set from Herrera et al. (2018), referred to for “imagine-self”, did not contain it. Theoretically, Attitudes Towards Refugees is expected to be higher for “imagine-self” than for “imagine-other,” but this cannot be verified in this research due to the above data limitations.

8. Results

8.1 Population Variables Summary

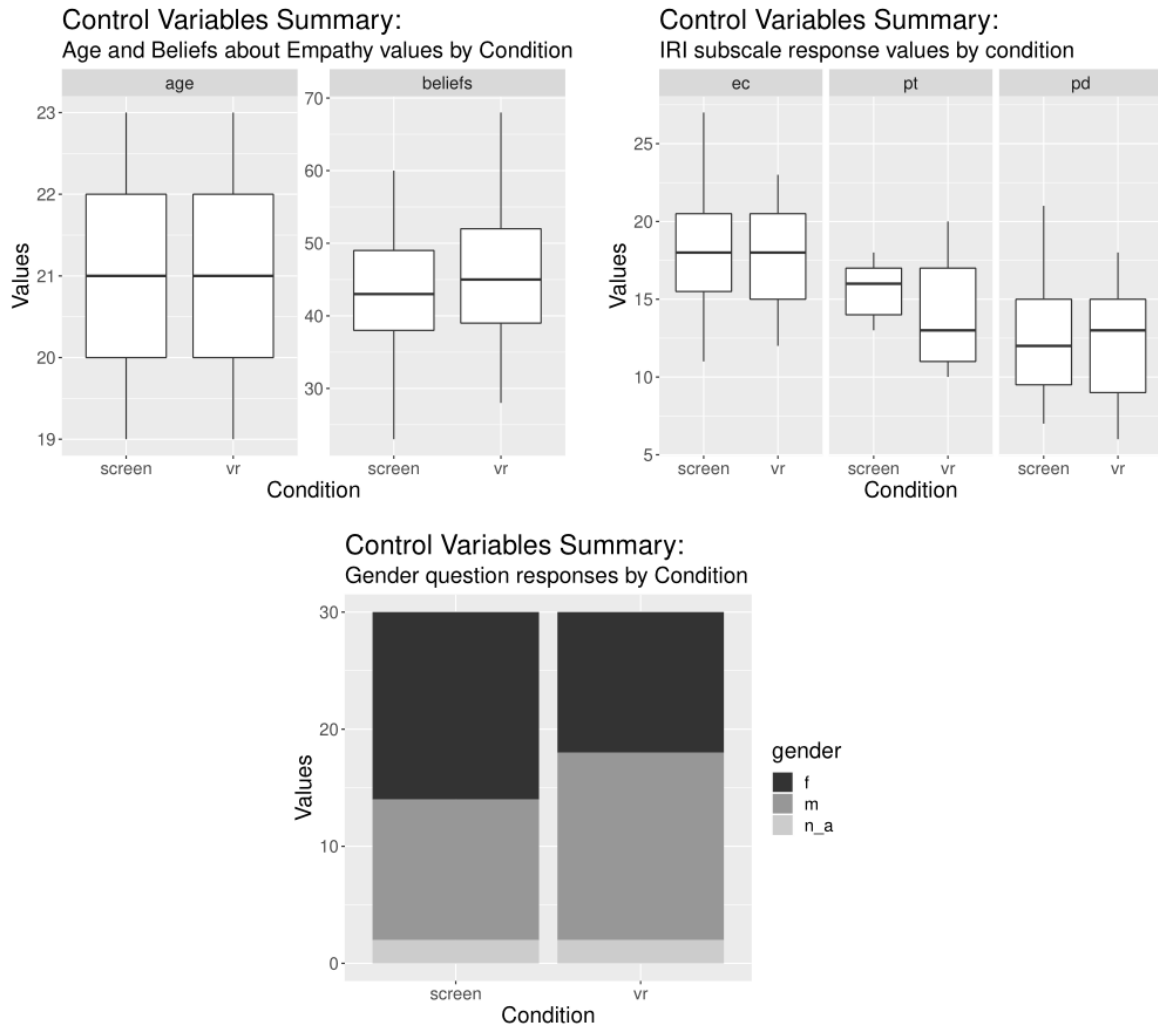
Responses were collected from a total of 30 people.

A multivariate analysis of variance (MANOVA) was performed with each of the continuous population variables as dependent variables, and the Media (Screen vs. VR) variable as the explanatory variable, in order to ascertain that there were no significant differences between Group A (Screen) and Group B (VR) for each of the control variables.

While a visual analysis of the boxplots suggests a slight bias for one of the subscales (PT), the statistical tests nonetheless found no significant differences between the groups in any of the continuous variables, indicating that the random assignment was successful. A full summary of the population variables as well as the results of the tests can be found in Appendix A, Population Variables.

The Gender variable was not included in the MANOVA model, as it is not continuous. The distribution between the “Screen” and “VR” groups was balanced during group assignment.

Figure 1: Control Variables Summary: Screen vs. VR



8.2 Hypothesis Verification

Table 6: Summary of Results per Hypothesis

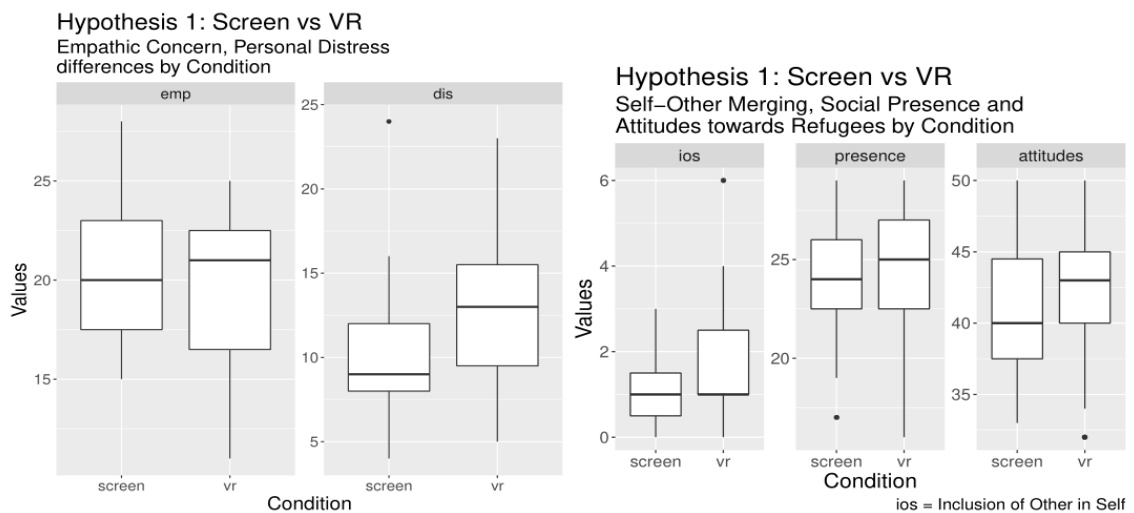
Hypothesis	Expectation	Result (significance code)
VR vs Screen	$EC_{vr} > EC_{screen}$	–
	$PD_{vr} > PD_{screen}$	–
	$IoS_{vr} > IoS_{screen}$	–
	$Presence_{vr} > Presence_{screen}$	–
	$Attitudes_{vr} > Attitudes_{screen}$	–
Imagine-self vs Imagine-other	$EC_{self} > EC_{other}$	–
	$PD_{self} > PD_{other}$	$PD_{self} > PD_{other} ***$
	$IoS_{self} > IoS_{other}$	$IoS_{self} > IoS_{other} ***$
	$Presence_{self} > Presence_{other}$	$Presence_{self} < Presence_{other} ***$

>: There is a statistically significant difference in means between the groups.

Significance codes: *** 0.001, ** 0.01, * 0.05, . 0.1

Hypothesis 1: VR vs. Screen

Figure 2: Hypothesis 1: Screen vs. VR

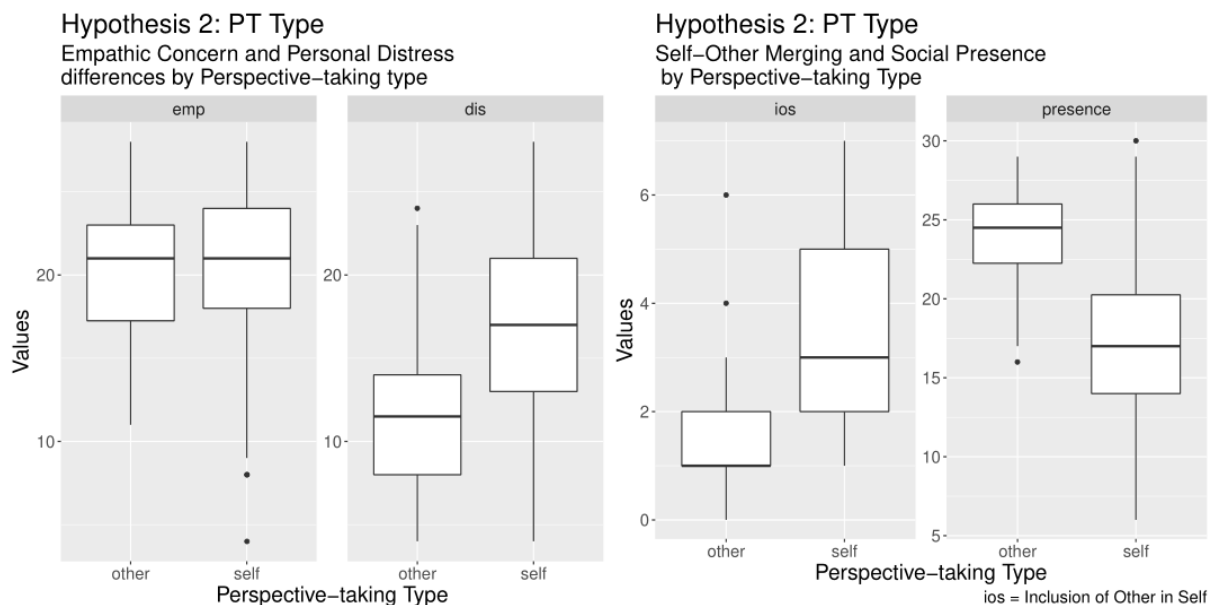


A multivariate analysis of variance (MANOVA) was performed with the Empathic Concern, Personal Distress, Inclusion of Other in Self, Social Presence, and Attitudes Towards Refugees as dependent variables, and with the Media (Screen vs. VR) variable as the explanatory variable. In addition, separate analysis of variance (ANOVA) was performed with each of the dependent variables above.

The results of the overall MANOVA were not statistically significant ($p = 0.553$), nor were the results of the individual ANOVAs ($p > 0.2$ for each).

Hypothesis 2: Perspective-taking Type

Figure 3: Hypothesis 2: Perspective-taking Type & Interactivity



A multivariate analysis of variance (MANOVA) was performed on the combined data set from this experiment and from Herrera et al. (2018), with the dependent variables set to Empathic Concern (emp), Personal Distress (dis), Inclusion of Other in Self (ios), and Social Presence.

The results of the MANOVA showed statistically significant differences between participants in the “imagine-self” group versus the “imagine-other” group with respect to the outcome variables ($p < 0.001$). The means of all continuous dependent variables were found to be significantly different for the two groups (with $p < 0.001$ for all of them), with the exception of Empathic Concern, for which no statistically significant differences were found. The full results can be found in Appendix C.

For the variables which had statistically significant differences, the following relationships were observed:

- Personal Distress was higher for imagine-self tasks than for imagine-other tasks.
- IoS (self-other merging) was higher for imagine-self tasks than for imagine-other tasks.
- Social Presence was lower for imagine-self tasks than for imagine-other tasks.

To further examine the relationship between Empathic Concern, Personal Distress, the type of media, and the perspective-taking task, an additional investigation using a linear regression model was conducted.

Two models were constructed, one for Empathic Concern and one for Personal Distress. In both cases, media (Screen vs. VR) and PT type (Imagine-self vs. Imagine-other) were included as explanatory variables. Furthermore, the relevant IRI subscale (EC for Empathic Concern, PD for Personal Distress respectively), as well as age and gender were included as controls.

Both models failed to find a statistically significant relationship between the type of media and the dependent variable. In both models, the type of perspective-taking task was found to be a significant predictor (negative, $p < 0.05$ for Empathic Concern; positive, $p < 0.001$ for Personal Distress).

9. Results Discussion

Of the hypotheses proposed in this research, only Hypothesis 2 (perspective-taking type) was supported by experimental results. The remainder of this section will discuss the implications of the results for each of the hypotheses and place them in the context of previous research.

Hypothesis 1: Screen vs. VR

The analysis of the survey results could not find evidence to confirm that the choice of VR over a computer screen produces stronger empathic responses. This seemingly contradicts claims by the United Nations, who have used the same documentary as the one employed in this study

(“Clouds Over Sidra”) in their fundraising efforts and had found an increase in donations when compared to previous years, when VR was not employed [UNVR n.d.a).

This contradiction may be explained by a recent meta-analysis (Martingano et al. 2021) which suggests that VR improves emotional empathy but not cognitive empathy. It can be thus hypothesized that the apparent increase in donations was caused by a momentary arousal of compassionate feelings, but it does not reveal an improved ability to imagine the perspectives of refugees.

Hypothesis 2: Perspective-taking Type

Evidence gathered as part of this research suggests that the type of perspective-taking tasks may significantly influence empathy-related responses in the target individual. However, the specific effects seem to vary between the outcome variables.

“Imagine-self” perspective-taking tasks are expected to produce a mixture of other-oriented Empathic Concern and self-oriented Personal Distress, whereas “imagine-other” is not expected to produce such a mixture. With this in mind, experimental findings for Personal Distress, Empathic Concern, and self-other merging seem to be in line with theoretical predictions.

Personal Distress was found to be significantly stronger for people who were tasked with imagining themselves as being part of the out-group and attempting to simulate that experience, as opposed to people who just watched and listened to a member of the out-group talk about their experience. Furthermore, when controlling for the media employed, “imagine-self” tasks produced weaker Empathic Concern in the participants than “imagine-other” tasks. Similarly, self-other merging (as measured by the Inclusion of Other in Self scale) was significantly higher for participants in the “imagine-self” tasks. All three of these findings support previously theorized phenomena.

Feelings of Social Presence describe how “real” the depicted characters feel to the participant, and whether they feel “present in the same space.” These feelings were found to be significantly stronger for participants in the “imagine-other” condition than for those in the “imagine-self” condition. This could be an interesting finding, indicating that participants in the “imagine-self” condition were less focused on the

presence of other characters, due to being absorbed in the simulation and focusing on personal feelings of discomfort.

Limitations

Overall, the most serious limitation of this study is the sample size. Cohen (1992) recommends a sample of at least 38 people to be able to detect a large effect with an ANOVA for two groups at $\alpha = 0.01$. Progressively larger sample sizes are suggested to be able to detect smaller effect sizes. The sample size employed in testing Hypothesis 1 is enough to detect a large effect at $\alpha = 0.10$ (Personal Distress was found to be different between the two media conditions, with $p = 0.07$). It is possible that other effects have not been detected due to the sample size.

Furthermore, the study is limited in the way that the sampling was performed. Constraining the choice of participants to Japanese university students from the same university and within a certain age group has mitigated intra-group variance. However, it also does not allow the results to be generalized to the entire population, including other age and social groups.

Finally, while the analysis results from Hypothesis 2 do not suffer from a low sample size (as the combined data set containing responses from Herrera et al. (2018) was used), they are nevertheless ambiguous. This is due to the fact that, while the methodology in this study was the same as in Herrera et al., both the out-group depicted in the content, and the nature of the content itself were different: this study was concerned with refugees and had little interactivity, whereas the Herrera et al. research focused on homeless people and was more interactive. Thus, as multiple factors differed between this experiment and that of Herrera et al., it is not possible to separate the effects caused by difference in perspective-taking task from the effects caused by difference in content.

Implications for International Relations

While the use of virtual reality has met some criticism from the academic community (Kool 2016), it is also perceived as a new tool in humanitarian development communication, one which promises to bridge the gap between donors and recipients, by providing the donors with an immersive glimpse into the realities faced by the recipients (Irom 2018).

However, while proponents depict a promising future for the use of

VR within international relations, significant issues should be considered.

Firstly, the empathic response to the consuming of content depends not only on the choice of medium, but also other factors. For example, the relationship context between the participant and the depicted group has been raised (Maner & Gailliot 2007). This research additionally indicates that the type of perspective-taking task performed also plays a role. Finally, VR may be limited in the type of effects it produces, as a recent meta-survey suggests that VR improves emotional empathy, but not cognitive empathy (Martingano et al. 2021).

Thus, empirical evidence suggests that for the practical policy goals of building empathy to improve intergroup international relations, the choice of VR as a medium is not the defining factor for success.

Secondly, a limitation of VR experiences is revealed by examining from a Critical Culture Studies perspective. VR experiences, like traditional media, are the creative product of the people, media companies, and international organizations who create them. VR experiences are political representations that inevitably reflect the biases and framings of their creators. On the viewer's side, interacting with VR experiences may reinforce the very power imbalances that the experiences apparently aim to correct. As Bleiker (2001, p. 511) succinctly summarizes, "representing the political is a form of interpretation that is, by its very nature, incomplete and bound up with the values of the perceiver."

Returning to the target of this investigation, *Clouds Over Sidra*, the act of viewing the documentary can now be seen in a new light. From this point of view, whether a diplomat at a fundraiser watches the documentary via a computer screen or via a VR headset bears little significance.

The documentary was made by two Americans (Gabo Arora and Chris Milk), who made the decisions of where to place the camera and when to turn it on. Transferring the 360° documentary from a flat screen to a VR headset may improve immersion and hide the presence of the authors, but it preserves the power imbalances between the viewer, the filmmaker, and the depicted refugees. This imbalance is exacerbated by the language employed: the filmmaker claims increased credibility to their representation by calling it "reality."

This problem is summarized in Rose (2018, p. 143):

While calling on participants to feel compassion for refugees, and encouraging them to donate, VR documentaries that tell stories of distant suffering can be prey to the same problem that

Godmillow identifies — promoting the status quo, by failing to address the systemic causes of the social problem described, or probing the ways in which the viewer is implicated in those systems.

In conclusion, it is difficult to dismiss the potential for virtual reality to foster empathy and understanding. However, this research along with existing empirical evidence suggests that factors other than the medium strongly influence the degree of empathy produced. Additionally, empathy-inducing VR experiences need to be critically examined within existing power structures and assessed together with biases of the creators. As with other media, the transformative power of VR may be limited if the out-group is not given agency in the production of the media. Ultimately, VR's ability to promote prosocial behavior needs to be further corroborated with empirical evidence as our understanding of empathy advances.

Future work can explore the relationship between VR, empathy, and prosocial behavior. A possible direction is to consider the effects of time: whether longer or more frequent exposure to VR experiences help foster empathy, and whether this leads to a long-lasting improvement of attitudes.

Additionally, future research may consider if VR participation contributes to viewers' developing paternalistic views of refugees, to address existing criticisms of VR documentary use in development aid.

Finally, future work may expand on the attempts made in this study to bridge experimental social psychology and international relations theory, thus providing deeper insights into the processes which govern the minds and hearts of policymakers, and the ways these processes influence politics.

References

- Ahn, S. J., Le, A. M. T., & Bailenson, J. (2013). The effect of embodied experiences on self-other merging, attitude, and helping behavior. *Media Psychology*, 16(1), 7–38.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596. doi: 10.1037/0022-3514.63.4.596
- Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review*, 3(1), 141–177. <https://doi.org/10.1111/j.1751-2409.2009.01013.x>
- Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, 40(2), 290–302. <https://doi.org/10.1037/0022-3514.40.2.290>
- Batson, C. D., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imaging how you would feel. *Personality and Social Psychology Bulletin*, 23(7), 751–758.
- Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., ... & Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, 72(1), 105.
- Bleiker, R. (2001). The aesthetic turn in international political theory. *Millennium*, 30(3), 509–533. DOI: 10.1177/0305829816684262
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.
- Davis, M. H. (1983a). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113.
- Davis, M. H. (1983b). The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality*, 51(2), 167–184.
- Decety, J., & Moriguchi, Y. (2007). The empathic brain and its dysfunction in psychiatric populations: Implications for intervention across different clinical conditions. *BioPsychoSocial Medicine*, 1(1), 22.
- Geere, D. (2017, April 25). Pyongyang VR teleports you inside North Korea, one of the world's most secretive countries. *Wired*. Retrieved from <https://www.wired.co.uk/article/north-korea-vr>

- Gerdes, K. E., Segal, E. A., & Lietz, C. A. (2010). Conceptualising and measuring empathy. *British Journal of Social Work*, 40(7), 2326–2343.
- Gürerk, Ö., & Kasulke, A. (2018). Does virtual reality increase charitable giving? An experimental study. *SSRN Electronic Journal*. 10.2139/ssrn.3072002.
- Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PloS one*, 13(10), e0204494.
- Irom, B. (2018). Virtual reality and the Syrian refugee camps: Humanitarian communication and the politics of empathy. *International Journal of Communication*, 12, 23.
- Kool, H. (2016). The ethics of immersive journalism: A rhetorical analysis of news storytelling with virtual reality technology. *Intersect: The Stanford Journal of Science, Technology, and Society*, 9(3).
- Maner, J. K., & Gailliot, M. T. (2007). Altruism and egoism: Prosocial motivations for helping depend on relationship context. *European Journal of Social Psychology*, 37(2), 347–358.
- Martingano, A. J., Hererra, F., & Konrath, S. (2021). Virtual reality improves emotional but not cognitive empathy: A meta-analysis. *Technology, Mind, and Behavior*, 2(1). doi:10.1037/tmb0000034
- Myers, M. W., Laurent, S. M., & Hodges, S. D. (2014). Perspective taking instructions and self-other overlap: Different motives for helping. *Motivation and Emotion*, 38, 224–234. doi:10.1007/s11031-013-9377-y
- Nash, K. (2017). Virtual reality witness: exploring the ethics of mediated presence. *Studies in Documentary Film*, 12(2), 119–131. doi:10.1080/17503280.2017.1340796
- Neumann, D. L., Chan, R. C., Boyle, G. J., Wang, Y., & Westbury, H. R. (2015). Measures of empathy: Self-report, behavioral, and neuroscientific approaches. In *Measures of personality and social psychological constructs* (pp. 257–289). Academic Press.
- Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 12(5), 481–494.
- Oculus VR for Good. (n.d.). Retrieved June 19, 2019, from <https://www.oculus.com/vr-for-good/>
- Pedwell, C. (2014). *Affective relations: The transnational politics of empathy*. Springer. ISBN 9781137275257

- Rose, M. (2018). The immersive turn: Hype and hope in the emergence of virtual reality as a nonfiction platform. *Studies in Documentary Film*, 12(2), 132–149.
- Schroeder, D. A., Dovidio, J. F., Sibicky, M. E., Matthews, L. L., & Allen, J. L. (1988). Empathic concern and helping behavior: Egoism or altruism? *Journal of Experimental Social Psychology*, 24(4), 333–353.
- Schumann, K., Zaki, J., & Dweck, C. S. (2014). Addressing the empathy deficit: Beliefs about the malleability of empathy predict effortful responses when empathy is challenging. *Journal of Personality and Social Psychology*, 107(3), 475.
- Schutte, N. S., & Stilinović, E. J. (2017). Facilitating empathy through virtual reality. *Motivation and Emotion*, 41(6), 708–712.
- United Nations Virtual Reality (n.d.a). Syrian refugee crisis. Retrieved April 23, 2019, from <http://unvr.sdgactioncampaign.org/cloudoversidra/>
- United Nations Virtual Reality (n.d.b). VR films. Retrieved from <http://unvr.sdgactioncampaign.org/vr-films/>
- United Nations Virtual Reality (n.d.c). About UNVR. Retrieved from <https://unvr.sdgactioncampaign.org/home/about/>
- Yee, N., & Bailenson, J. N. (2006). Walk a mile in digital shoes: The impact of embodied perspective-taking on the reduction of negative stereotyping in immersive virtual environments. *Proceedings of PRESENCE*, 24, 26.

Appendices

A. Population Variables

Summary Tables

Table 7: Population Variables

Age	Gender	Empathy Beliefs
Min.: 19.00	F: 14	Min.: 23.00
1st Qu.: 20.00	M: 14	1st Qu.: 38.25
Median: 21.00	N/A: 2	Median: 45.00
Mean: 21.07		Mean: 43.90
3rd Qu.: 22.00		3rd Qu.: 50.00
Max.: 23.00		Max.: 68.00

Table 8: Population Variables: IRI Subscales

IRI: EC	IRI: PT	IRI: PD
Min.: 11.00	Min.: 10.00	Min.: 6.00
1st Qu.: 15.00	1st Qu.: 13.00	1st Qu.: 9.00
Median: 18.00	Median: 14.00	Median: 13.00
Mean: 17.73	Mean: 14.77	Mean: 12.67
3rd Qu.: 20.75	3rd Qu.: 17.00	3rd Qu.: 15.00
Max.: 27.00	Max.: 20.00	Max.: 21.00

MANOVA: Population Variable Differences between Screen and VR Groups

Table 9: Multivariate Analysis of Variance: Differences in Age, IRI Subscales (EC, PT, PD) and Beliefs about Empathy between Conditions (Screen vs. VR))

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
Media	1	0.08292	0.434	5	24	0.8203
Residuals	28					

Table 10: Analysis of Variance: Differences in Age between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Media	1	0.000	0.0000	0	1
Residuals	28	43.867	1.5667		

Table 11: Analysis of Variance: Differences in Empathic Concern (IRI Subscale) between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Media	1	0.53	0.5333	0.0309	0.8617
Residuals	28	483.33	17.2619		

Table 12: Analysis of Variance: Differences in Perspective Taking (IRI Subscale) between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Media	1	9.633	9.6333	1.3781	0.2503
Residuals	28	195.733	6.9905		

Table 13: Analysis of Variance: Differences in Personal Distress (IRI Subscale) between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Media	1	0.53	0.53333	0.0371	0.8486
Residuals	28	402.13	14.3619		

Table 14: Analysis of Variance: Differences in Beliefs about Empathy Scale between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Media	1	116.03	116.03	1.0749	0.3087
Residuals	28	3022.67	107.95		

B. Hypothesis 1: Screen vs. VR

Table 15: Multivariate Analysis of Variance: Differences in Empathic Concern, Personal Distress, Inclusion of Other in Self, Social Presence, and Attitudes towards Refugees between the Two Conditions

	Df	Pillai	Approx. F	num Df	den Df	Pr(>F)
Condition	1	0.14463	0.81161	5	24	0.553
Residuals	28					

Table 16: Analysis of Variance: Differences in Empathic Concern between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Condition	1	12.03	12.033	0.6914	0.4127
Residuals	28	487.33	17.405		

Table 17: Analysis of Variance: Differences in Personal Distress between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Condition	1	30.00	30.00	1.2577	0.2716
Residuals	28	667.87	23.852		

Table 18: Analysis of Variance: Differences in Inclusion of Other in Self between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Condition	1	2.133	2.1333	1.1256	0.2978
Residuals	28	53.067	1.8952		

Table 19: Analysis of Variance: Differences in Social Presence between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Condition	1	2.13	2.1333	0.1651	0.6876
Residuals	28	361.87	12.9238		

Table 20: Analysis of Variance: Differences in Attitudes towards Refugees between Conditions

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Condition	1	4.80	4.800	0.1654	0.6873
Residuals	28	812.67	29.024		

C. Hypothesis 2: Perspective-taking Type

Table 21: Multivariate Analysis of Variance: Differences in Empathic Concern, Personal Distress, Inclusion of Other in Self, and Social Presence between Perspective-taking Tasks

	Df	Pillai	approx F	num Df	den Df	Pr(>F)
PT type	1	0.34464	31.685	4	241	<2.2e016***
Residuals	244					

Table 22: Analysis of Variance: Differences in Empathic Concern between Perspective-taking Tasks

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
PT type	1	4	4.0409	0.1783	0.6732
Residuals	244	5530	22.6639		

Table 23: Analysis of Variance: Differences in Personal Distress between Perspective-taking Tasks

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
PT type	1	685.1	685.14	22.721	3.221e-06***
Residuals	244	7357.9	30.16		

Significance codes: ***0.001 **0.01 *0.05 . 0.1

Table 24: Analysis of Variance: Differences in Inclusion of Other in Self between Perspective-taking Tasks

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
PT type	1	84.38	84.383	30.083	1.033e-07***
Residuals	244	684.42	2.805		

Significance codes: ***0.001 **0.01 *0.05 . 0.1

Table 25: Analysis of Variance: Differences in Social Presence between Perspective-taking Tasks

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
PT type	1	1177.2	1177.24	52.24	6.276e-12***
Residuals	244	5498.6	22.54		

Significance codes: ***0.001 **0.01 *0.05 . 0.1

Andrey Kirpach Profile

Born in 1996 in Saint-Petersburg, Russia. Graduated in 2020 from the Faculty of Law at Hitotsubashi University, specializing in International Relations. Currently employed as a software engineer in Tokyo, Japan.