

# Bystander Affiliation Influences Intervention Behavior: A Virtual Reality Study

SAGE Open  
 July-September 2021: 1–13  
 © The Author(s) 2021  
 DOI: 10.1177/21582440211040076  
[journals.sagepub.com/home/sgo](https://journals.sagepub.com/home/sgo)  


Aitor Rovira<sup>1,2,3</sup> , Richard Southern<sup>4</sup>, David Swapp<sup>1</sup>,  
 Claire Campbell<sup>5</sup>, Jian J. Zhang<sup>4</sup>, Mark Levine<sup>6</sup>, and Mel Slater<sup>1,7</sup>

## Abstract

Traditional work on bystander intervention in violent emergencies has found that the larger the group, the less the chance that any individual will intervene. Here, we tested the impact on helping behavior of the affiliation of the bystanders with respect to the participants. We recruited 40 male supporters of the U.K. Arsenal football club for a two-factor between groups study with 10 participants per group. Each participant spoke with a virtual human Arsenal supporter (V), the scenario displayed in a virtual reality system. During this conversation, another virtual character (P), not an Arsenal fan, verbally abused V for being an Arsenal fan leading eventually to physical pushing. There was a group of three virtual bystanders who were all either Arsenal supporters indicated by their shirts, or football fans wearing unbranded shirts. These bystanders either encouraged the participant to intervene or dissuaded him. We recorded the number of times that participants intervened to help V during the aggression. We found that participants were more likely to intervene when the bystanders were out-group with respect to the participant. By comparing levels of intervention with a “baseline” study (identical except for the presence of bystanders), we conclude that the presence of in-group bystanders decreases helping. We argue therefore that, other things being equal, diffusion of responsibility is more likely to be overcome when participant and victim share group membership, but bystanders do not. Our findings help to develop understanding of how diffusion of responsibility works by combining elements of both the bystander effect and the social identity approach to bystander behavior.

## Keywords

bystander effect, virtual reality, social identity approach, diffusion of responsibility, violence

## Introduction

A person unexpectedly confronted with an emergency, such as seeing a stranger being attacked by another, has a vital decision to make—walk on by, watch but do nothing, or try to intervene in some way to provide help to the victim. Provoked by the need to understand how multiple bystanders apparently allowed the murder of a woman to take place in front of them, Latane and Darley (1968) first established the theory that the more bystanders observing such an event, the less the likelihood that anyone would intervene—see also (Darley & Latané, 1968). This was due to diffusion of responsibility, where if many are present, the responsibility is diminished proportionately. This finding, referred to as the bystander effect, has been repeatedly verified in numerous studies—see, for example, the meta-analysis in Fischer et al.’s study (2011). However, recently the literature has begun to identify additional situational factors that affect the degree of helping behaviors that bystanders exhibit. Key among these is the psychological relationship between bystanders. For example, bystander intervention is increased

when bystanders are friends (Burn, 2009; Latane & Darley, 1969; Latané & Rodin, 1969; Levine & Crowther, 2008; Liebhart, 1972), prior acquaintances (Gottlieb & Carver, 1980), or members of a cohesive group (Rutkowski et al., 1983). Moreover, Levine and colleagues have developed a social identity model of bystander intervention in emergencies, for example (Hopkins et al., 2007; Levine, 1999; Levine & Crowther, 2008; Levine et al., 2002, 2005; Levine & Thompson, 2004; Manning et al., 2007; Reicher et al., 2006).

<sup>1</sup>University College London, UK

<sup>2</sup>Oxford Health NHS Foundation Trust, UK

<sup>3</sup>University of Oxford, UK

<sup>4</sup>Bournemouth University, Poole, UK

<sup>5</sup>Ulster University, Coleraine, UK

<sup>6</sup>Lancaster University, UK

<sup>7</sup>University of Barcelona, Spain

## Corresponding Author:

Aitor Rovira, Department of Psychiatry, University of Oxford, Warneford Hospital, Warneford Lane, Oxford OX3 7JX, UK.

Email: [aitor.rovira@psych.ox.ac.uk](mailto:aitor.rovira@psych.ox.ac.uk)



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of

the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

This work indicates that intervention increases when victims are viewed as common category members and that individuals are more likely to be influenced by fellow bystanders when they are in-group rather than out-group members. Although there is limited research that explores bystander intervention in a context of violence (Borofsky et al., 1971; Fischer et al., 2006; Harari et al., 1985; Schwartz & Gottlieb, 1976, 1980; Shotland & Straw, 1976), Levine and Crowther (2008) have shown that group membership also impacts bystander intervention in the context of violent emergencies. Specifically, they found that when bystanders have a shared group membership, group size can facilitate as well as inhibit helping.

In any actual violent emergency situation, there are therefore many factors that may affect the degree of helping behavior by bystanders—the nature of the emergency itself and its level of potential danger, the perceived or actual relationship in terms of social identity between a bystander and the victim, a bystander and the perpetrator, the group identity between the bystanders themselves, the extent to which bystanders are a coherent group (e.g., of friends) or strangers, as well as their number.

Given the particular challenges of studying violence, these factors are very difficult to study experimentally through the careful control of potentially contributing factors to helping behavior. For ethical and practical reasons, violent emergencies that expose experimental subjects to real-life violence cannot be staged, and even if this were possible, actors would have to tirelessly reproduce identical performances over and over again (and be paid). Alternatively, some studies that focused on observations in real-life situations tend to have low internal validity, as each situation occurs in different contexts and due to different reasons. Situations cannot be compared directly, and observers can easily miss important details of how and why it happened. A popular way to study violent emergencies is with CCTV cameras (Levine et al., 2011). This dramatically increases the number of incidents that can later be analyzed but also has some limitations related to technology—surveillance cameras do not usually record audio, the image resolution is low, and some important events can occur out of the camera angle, and nor is the study under experimental control.

Aitor Rovira et al. (2009) showed how immersive virtual reality (VR) can be used to overcome these problems. There has been 30 years of research on the issue of how people respond to situations and events in VR. This is generally referred to as “presence,” the illusion of being in the place depicted by the VR system (Held & Durlach, 1992; Sanchez-Vives & Slater, 2005; Sheridan, 1992). Presence has been argued to have two dimensions: “Place Illusion” (PI), the illusion of being in the virtual place, and Plausibility (Psi), the illusion that events taking place are really happening—even though in both instances participants know for sure that this is not the case (Slater, 2009). PI is argued to depend on sensorimotor contingencies for

perception that are similar to those in physical reality, essentially using the body to perceive by looking around, looking behind objects, reaching out, bending down, and so on (O’Regan & Noë, 2001). Psi is based on events in the environment responding to actions of the participant, contingent events such as virtual human characters spontaneously addressing the participant, and a match between the environment and expectations where it depicts a situation that is familiar to participants in reality. For example, an early version of the bar scenario used in this paper (Aitor Rovira et al., 2009) failed the Psi test simply because the decor of the bar in which the confrontation took place was not recognized as one that would be frequented by football fans.

The use of VR in social psychology has long been argued for (Blascovich, 2002; Blascovich et al., 2002; Loomis et al., 1999) and more recently by Pan and de C Hamilton (2018). Previous work using a VR paradigm showed how group identity played a critical role in fostering or hindering helping behavior in the context of a lone bystander confronted with a fight between two life-sized virtual human characters (Slater et al., 2013). That study used a football team as a marker of group affiliation, where all experimental subjects were strong fans of the English football club Arsenal. Experimental participants were in conversation about football with a (virtual) man (the victim) who was then attacked by a third man (the perpetrator)—verbally but with increasing menace, eventually ending in physical violence. Participants were more likely to intervene to try to stop the argument when the victim was also clearly an Arsenal supporter than when the victim was a general football fan without any particularly obvious affiliation. These results were compatible with earlier findings that in a nonviolent emergency participants (all Manchester United fans) were far more likely to help an injured fallen stranger when he was wearing a Manchester United shirt, compared to the shirt of another team (Levine et al., 2005).

In this article, we turn to the bystander effect itself, but instead of considering whether the number of bystanders influence helping behavior, we concentrate on the group identity of the bystanders in relation to the experimental participants and the victim. Previous findings suggest that fellow bystanders are more influential when they are in-group rather than out-group members. For example, Levine et al. (2002) asked participants to watch a CCTV clip of violence in the presence of two confederates. The confederates were presented as in-group or out-group members and either encouraged or discouraged intervention. Results showed that participants were more influenced under in-group bystander conditions, and were more likely to intervene when encouraged, but less likely to intervene when discouraged.

Here we explore two different possible effects on the level of intervention of the participant. As in the previous experiment reported by Slater et al. (2013), team affiliation was indicated by football shirts, and all participants were Arsenal

supporters. However, in the new experiment, the victim was always an Arsenal supporter. Moreover, a group of three bystanders were present who were either Arsenal supporters as indicated by their shirts, or they were not Arsenal supporters indicated by their unbranded shirts. Therefore, the first factor in the new experiment was the affiliation of the group of three bystanders. The second factor was whether or not these three bystanders encouraged or discouraged the participant to intervene (independently of whether these three were Arsenal supporters or not). Our questions were how much the degree of intervention of the participant to stop the aggression would be influenced by (a) the affiliation of the three bystanders (as indicated by their shirts) and (b) whether the three bystanders encouraged or discouraged such intervention.

Hence, rather than only considering the number of bystanders, as in the classic bystander effect, we consider the effect that the group membership of bystanders (here indicated to the participant by their football shirts) might have on diffusion of responsibility in emergencies. Even though this experiment used the same basic scenario as Slater et al. (2013), it addresses quite different issues. In Slater et al.'s study (2013), there were no bystanders at all (except for the participant). The main question in that case was how the relationship between the affiliation of the victim and participant (both Arsenal supporters or not) would influence helping behavior shown by the participant. Here, in this new experiment, attention shifted to how the affiliation of the three bystanders would influence helping behavior of the participant, given that the participant always shared the same affiliation as the victim.

Classic diffusion of responsibility predicts that diffusion will be distributed across the numbers of others present, irrespective of their psychological relationship to each other. Here, we argue that in a context where the victim is in-group with respect to the participant, diffusion of responsibility will operate when the other bystanders belong to the same group as the participant and victim. More specifically, responsibility will only be diffused across those who are perceived by the participant to have an equal responsibility to help (since their football shirts indicate group affiliation as Arsenal supporters or not). When bystanders are out-group to the participant (in a context where both participant and victim are in-group), they will not be seen to have the same responsibility to act, and thus diffusion of responsibility may not occur. Thus, we expected that the likelihood of intervention will be higher when bystander virtual characters are represented as out-group. This may be independent of whether or not the bystanders encourage or discourage intervention.

## Materials and Methods

### Recruitment

Forty participants were recruited around the university campus. They were aged between 18 and 44, with no significant

differences between groups, all male Arsenal F.C. supporters. The recruitment announcement stated that we were looking for male Arsenal supporters. They had to complete a questionnaire that contained five questions related to football:

- What is your favorite team in the Premier League? (This was asked to make sure that they had read the recruitment announcement).
- How much do you support them?
- How do you usually follow the match when they play?
- How often do you attend football matches?
- How often do you watch football matches on TV?

The question “How much do you support them?” was answered on a scale from 1 (*Not at all*) to 7 (*Very much so*) and only those who answered four or higher were recruited, which was the only filter used. This questionnaire was completed online and included a section to make the appointment to attend the experiment. There was a gap of at least 2 days between completing the online questionnaire and taking part in the experiment. No details about the purpose of the study or the experience in VR were given at this stage.

The number who had applied to be in the study was 153. We filtered out all the participants who did not match the selection criteria (male, 18+ years old, Arsenal supporter, scoring 4+ on the question “how much do you support them?”). The experiment had been designed to recruit 40 participants and recruitment stopped when that number was reached.

### Procedures

Participants arrived at the laboratory and were assigned to a condition (Affiliation and Encouragement) according to the order in which they arrived. Our method of alternate allocation was random for all practical purposes in the sense that there was no deliberate allocation of each participant to an experimental group. Each participant was scheduled solely depending on their availability. On arrival, they were given an information sheet to read, and the procedures of the experiment explained to them. They completed a written informed consent form to obtain agreement for participation (they all agreed). They all filled out a short questionnaire (see Supplemental Text S1). They were explicitly told several times as well as being given in writing that they could withdraw from the experiment at any time without giving reasons. At the end of the experiment, they were debriefed about its purposes.

Upon entering the VR Cave (see below), participants found themselves in a virtual bar that looked like a standard English pub. The decoration included posters and other objects related to football, and a sports magazine program was on TV. There were three people sitting nearby watching TV. In one experimental condition, these three wore Arsenal football shirts and in the other condition unbranded shirts.

The first minute of the scenario was for the participant to acclimatize to the sight and the stereoscopic vision provided by the VR system and to the brightness in the bar, so nothing happened during that time, although we asked them to look for objects related to football. After this, a virtual life-sized character (V), who was wearing an Arsenal shirt walked into the scene and started a conversation with the participant about the Arsenal football team (see Supplemental Text S2 for a full transcript). He asked the participant about the team and the chances they had to win a trophy in the present season, showing a lot of friendship toward the participant and being very optimistic at the same time.

Early in the conversation, a man (P) wearing an unbranded shirt entered the bar and sat down on a stool near where they were having the conversation. After a further 2 minutes approximately, depending on how much time the participant had invested in the conversation, he (P) stood up and approached V accusing him of “staring” at him, and an argument ensued. P’s behavior became increasingly threatening over time. Although V tried to defuse the situation, the argument escalated until the point it reached physical violence (P started to push V toward a wall), when the scenario ended. The confrontation lasted for 2 minutes and 13 seconds.

During the argument, one of the bystanders approached and stood near P and V (next to the participant) and another of the bystanders who was sitting by the table spoke out loud three times. In both conditions, the first utterance was “What is this guy doing?” After this, in the encouraging condition, he tried to encourage the participant toward intervention by saying “Hey, hold on, we should do something about this” and “This guy has lost it, we’ve got do something now.” In the other condition, he tried to dissuade anyone from intervening by saying “There is nothing we can do about it, let’s leave him alone” and “This isn’t our business, let’s leave him alone.” See Figure 1 for an illustration of the environment, and Supplemental S1 Video.

The average overall time for each participant in the VR was 7 minutes. After the experience, they completed a questionnaire regarding the feelings they had during the confrontation. They were also interviewed by the researchers, so they could describe how they felt, as well as giving them the chance to point out anything they would want to mention. Each participant was paid £7, and it took between 30 and 40 minutes for them to complete all the stages.

### The VR System

We used a projection-based VR system known generically as “Cave” (Cruz-Neira et al., 1993). The system consists of a room where three walls and the floor are projection screens. The height is 2.2 m and the floor  $3 \times 3$  m<sup>2</sup>. Each DLP projector delivers an image of  $1440 \times 1050$  pixels with a refresh rate of 100 Hz. Due to the floor having a different aspect ratio than the projected image, the floor image is cropped to  $1100 \times 1050$  pixels. The projectors are controlled by a PC

cluster composed of 4 PCs, each one with an Nvidia Quadro FX 5600 graphics, delivering 3D stereoscopic images synchronized to Crystal Eyes™ shutter glasses worn by the participant. The participant’s head was tracked by an Intersense IS900 system with a refresh rate of 180 Hz that was used to adjust the imagery to his perspective in real time.

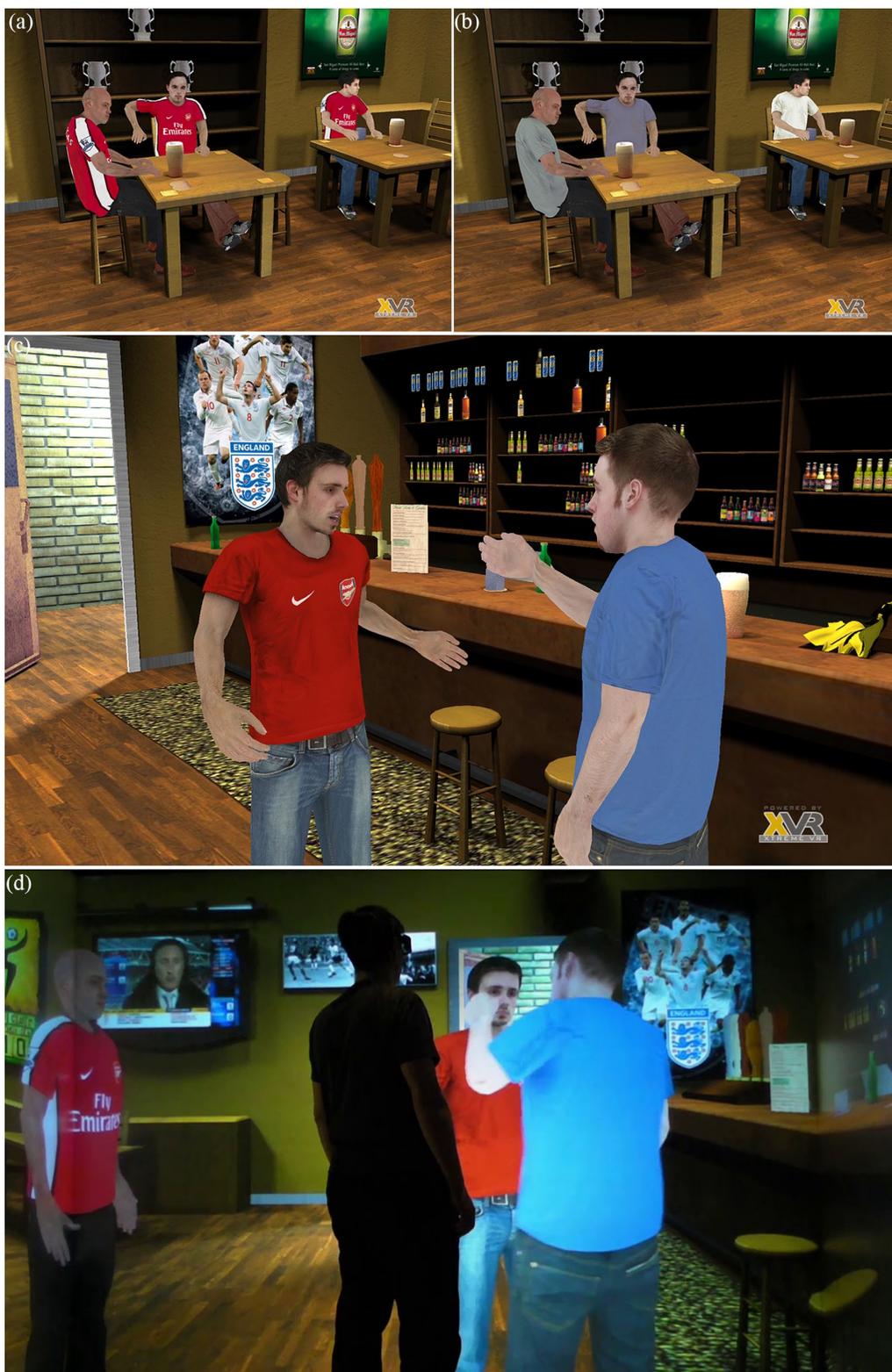
### Counting the Number of Interventions

Any action (verbal or physical) that was executed on purpose to catch the attention of someone else in the scene was considered an intervention. A verbal intervention was considered anything that the participant would say to either the victim, perpetrator, or the other bystanders, excluding interjections that would be more a think-aloud utterance rather than saying something directed to any (or many) of them, if its objective was not to catch their attention, as defined previously (Rovira, 2016). A physical intervention was considered as any attempt to make physical contact with others, such as reaching out to P, and also moving very close to them, even walking in between victim and perpetrator to try to separate them, moving into P’s field of view to catch his attention, waving a hand or any other similar hand gesture.

Two consecutive actions were not considered just one intervention if there was a gap in between of at least 2 seconds. This is the minimum time that we found a person takes to observe the situation waiting for a reaction after an intervention and before intervening again. This is clear in verbal interventions. Some participants did not stop talking for some time, but if they did not stop for a moment to see the consequences of their interventions, then it was considered just one intervention. Our coding did not take into account the length of interventions. A clear example is that when a participant moved in between P and V, and then stayed in that position, this was counted as a single intervention.

The coding was carried out by a researcher uninvolved in the experiment, who was given the instructions above. As a check, one of the experimenters also independently coded the videos. There was high inter-rater reliability. For the number of verbal interventions, the Pearson correlation coefficient  $r = 0.95$ ,  $p < .00005$ . For the number of physical interventions  $r = 0.78$ ,  $p < .00005$ , and for the total number of interventions  $r = 0.92$ ,  $p < .00005$ . The Intraclass Correlation Coefficients (using the Stata 15 function “icc”) are respectively 0.93 (individual), 0.97 (average),  $F(39.0, 40.0) = 29.67$ ,  $p < .0005$  for the number of verbal interventions; 0.65 (individual), 0.79 (average),  $F(39.0, 40.0) = 4.66$ ,  $p < .0005$  for the number of physical interventions; 0.87 (individual), 0.93 (average),  $F(39.0, 40.0) = 13.97$ ,  $p < .0005$  for the total number of interventions. Here the  $F$ -test tests the null hypothesis that the ICC = 0.

Table 1 shows the frequencies of agreement between the two raters. This results in Cohen’s kappa = 0.90. This is high and corresponds to the findings above: there was a high degree of agreement between the two independent raters.



**Figure 1.** The scenario (a) the in-group bystanders sitting around a table, (b) the out-group bystanders sitting around a table, (c) the perpetrator (in blue) picks an argument with the victim, and (d) a photograph of the cave scenario showing a participant, a bystander, and the perpetrator being aggressive against the victim.

**Table 1.** Frequencies of Agreement Between the Raters in the Video Coding.

Rater 1	Rater 2	
	Intervention	No intervention
Intervention	23	0
No intervention	2	15

### The Post-VR Experience Questionnaire

The questionnaire administered immediately after the experience contained the following questions, all identical to those by Pomes and Slater (2013) and Rovira (2016), except for the last four. Each answer was scored on a Likert-type scale, from 1 (*completely disagree*) to 7 (*completely agree*).

- After the argument started, I was feeling uncomfortable with the situation.
- After the argument started, I was sometimes concerned for the safety of the man being threatened.
- After the argument started, I was sometimes concerned for my own safety.
- After the argument started, I looked around for help.
- After the argument started, I looked around to check in case the other people might come to make situation worse.
- After the argument started, the victim looked at me wanting help.
- After the argument started, I felt I should move away from those people.
- After the argument started, the aggressor was aware of me looking at him.
- After the argument started, I felt I should do something to stop it.
- After the argument started, I felt I could do something to stop it.
- After the argument started, I felt that I needed to get out.
- My mind started wandering and thinking about other things during the argument.
- The presence of other people in the bar encouraged (or dissuaded) me to intervene.
- The other people in the bar tried to persuade me.
- The other people's utterances had an impact on my behavior.
- Were the other people's utterances encouraging or trying to dissuade you to intervene?

### Statistical Methods

Statistical analysis was carried out using Stata 15 (stata.com). For the ANOVA, the function "anova" was used, with effect sizes produced by the "estat esize" command, and margins analysis with the "margins" command. The bar charts were produced with the "cibar" function (Staudt, 2014).

**Table 2.** Expected Square Root of the Number of Interventions Based on the Earlier Study.

Bystanders affiliation	No encouragement	Encouragement
Out-group	2	3.5
In-group	3.5	5

### Results

This was a two-factor between-group study. The first factor was bystander affiliation (in-group and out-group), and the second factor was encouragement (encourage and discourage). Ten participants were arbitrarily assigned to each cell of the  $2 \times 2$  factorial table (depending solely on the order at which they arrived to the VR laboratory).

A power analysis can be based on the data from a previously published experiment which used the identical scenario (without bystanders) and setup in one of the conditions (Aitor Rovira et al., 2013). The overall mean square root of the number of interventions in that experiment was about 2 (in fact 2.2). We suppose in the current experiment that we had expected the identity of the bystanders and the encouragement to each increase the mean square root of the number of interventions by a modest 1.5. Then our expected table of means of the square root of the number of interventions would be as shown in Table 2.

The overall within cell variance is about 4. This leads to an a priori power of 0.64. If based on expectations from the previous literature that the change might be a little higher, say an increase in the mean square root of the numbers of interventions by only 2, then the four numbers above would become 2, 4, 4, and 6, and the power would be 0.87. Power in this experiment turns out only to be pertinent to encouragement, where the differences were not found to be anywhere near significant.

Age data were recorded in intervals. All participants were in one of the three possible groups, 18 to 25, 26 to 34, or 35 to 44 years old. Considering age as a categorical variable, chi-square tests show that it is not statistically significant for either verbal interventions ( $\chi^2 = 12.675$ ,  $df = 20$ ,  $p = .89$ ) or physical interventions ( $\chi^2 = 4.67$ ,  $df = 12$ ,  $p = .97$ ). Also, the scores provided in the question "How much do you support your team?" were not statistically different for verbal interventions ( $\chi^2 = 36.34$ ,  $df = 30$ ,  $p = .2$ ) or physical ( $\chi^2 = 17$ ,  $df = 18$ ,  $p = .52$ ).

### The Encouragement Factor

Encouragement (encourage and discourage) did not have any effect on the number of interventions. This is probably because the level of encouragement or discouragement offered by the bystanders was quite low in intensity. The bystanders made only three comments, one neutral and two either encouraging intervention (encourage) or discouraging

**Table 3.** Frequencies of Answers to the Question “Were the Other People’s Utterances Encouraging or Trying to Dissuade You to Intervene?”

Answer	Dissuade	Nothing	Persuade	Total
Encouragement:				
Discourage	9	10	1	20
Encourage	0	9	11	20
Total	9	19	12	40

it (discourage) (see Methods). At the end of the VR exposure, participants were asked: “Were the other people’s utterances encouraging or trying to dissuade you to intervene?” with possible answers: dissuade, encourage, or nothing noticed. The responses are given in Table 3. Hence 9 out of 20 of those in the discourage condition and 11 out of 20 of those in the persuade condition noticed the corresponding interventions, whereas 19 out of 40 participants did not notice any intervention.

### Helping Behaviour

Helping behavior can be defined in many different ways, and we consider several different methods of assessing this (all based on the same data). We distinguish between interventions that were verbal (the participant saying something to V or P) and those that were physical, for example, the participant trying to step between them (see Methods for a description of how these interventions were recorded). Then there are the following possibilities by which an intervention can be defined, each one leading to a different type of analysis, but the same results.

**Categorical.** Here each type of intervention is considered as a separate category. From Table 4 we can see, for example, that 15 participants did not intervene at all, 12 made only verbal interventions, all physical interventions were accompanied by verbal ones, and 13 made both physical and verbal interventions.

Considering Table 4(a), it is clear the encouragement factor had no effect. From Table 4(b), we can see that of the 15 participants who made no intervention, 12 were in the condition where the bystanders were in-group. On the contrary, of the 25 who made some intervention (verbal only or verbal and physical) 17 were in the condition where the bystanders were out-group.

Multinomial logistic regression was carried out on the categorical responses using the Stata 15 function “mlogit” on the model affiliation + encouragement + affiliation × encouragement (i.e., the two main effects and the interaction). This shows that taking the base level as “no intervention,” then for the affiliation in-group the significance level is  $p = .019$  of being less than the base level with respect to “verbal only” and  $p = .013$  with respect to being less than

**Table 4.** Frequencies of the Categorical Responses by Affiliation and Encouragement.

(a) Encouragement.

Category	Bystanders out-group	Bystanders in-group	Total
No intervention	8	7	15
Verbal only	6	6	12
Physical only	0	0	0
Verbal and physical	6	7	13
Total	20	20	40

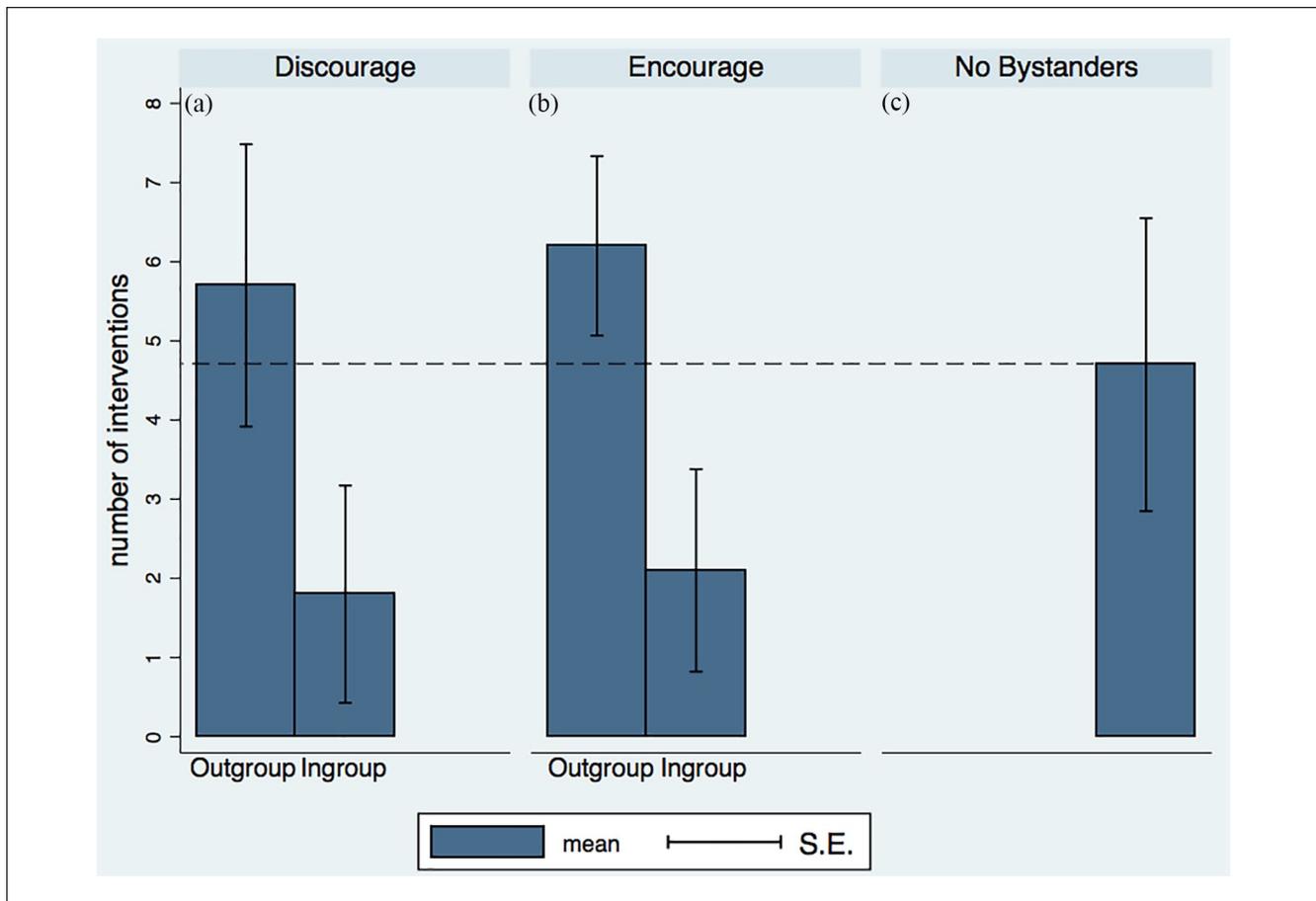
(b) Affiliation.

Category	Bystanders out-group	Bystanders in-group	Total
No intervention	3	12	15
Verbal only	8	4	12
Physical only	0	0	0
Verbal and physical	9	4	13
Total	20	20	40

the base level with respect to “verbal and physical.” There is no interaction effect ( $p = .704$  for the verbal only case and  $p = .613$  for the verbal and physical case), and the main effects of encouragement are  $p = .624$  and  $p = .512$ , respectively. Eliminating encouragement, these significance levels do not change.

**Binary.** Alternatively, an intervention can simply be classified as a binary event—the participant intervened at all at some stage, or never intervened. The frequencies can also be seen in Table 4. A binary logistic regression was carried out using the Stata 15 function logistic on the binary response variable with the same model as above affiliation + encouragement + affiliation × encouragement. The interaction term is not significant ( $p = .613$ ), encouragement is not significant ( $p = .538$ ) and  $p = .079$  for affiliation. Eliminating the encouragement factor  $p = .006$  for affiliation. This does not change using robust standard errors, which allows for possible departures from the model assumptions, and inflates the standard errors of the estimates.

**The number of interventions.** We use the sum of the number of verbal and physical interventions, though the results are the same if each of these is considered separately (Supplemental Tables S1 and S2). Figure 2 shows the means and standard errors of the number of interventions under the various conditions. The evidence suggests that the level of intervention was much less when the bystanders were Arsenal supporters (in-group). These are count variables, and when an ANOVA is fitted for the model affiliation + encouragement + affiliation × encouragement or any subset of this, the residual errors of the fit do not satisfy normality by far (for example, the Shapiro–Wilk test gives  $p = .0001$  for full model and



**Figure 2.** Bar charts showing the means and standard errors of the number of interventions: (a) and (b) by affiliation of the bystanders (out-group, in-group) and encouragement and (c) a condition with no bystanders.

**Table 5.** ANOVA of Square Root of Number of Interventions on Affiliation and Encouragement.

Source	df	F-ratio	p	Partial $\eta^2$
Affiliation	1	11.23	.002	0.24
Encouragement	1	0.83	.368	0.02
Affiliation $\times$ encouragement	1	0.20	.655	0.01

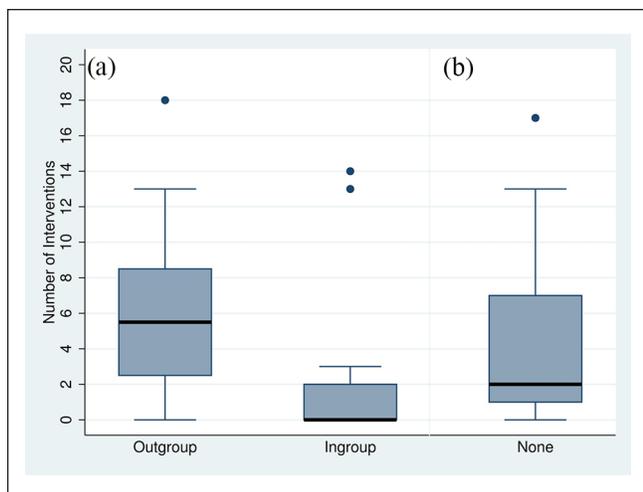
Note. Affiliation out-group = 0, in-group = 1) and encouragement (discourage = 0, encourage = 1). Overall fit:  $F(3, 36) = 4.09$ ,  $p = .013$ ,  $R^2 = .25$ , Shapiro–Wilk’s test for normality of residual errors:  $p = .20$ . ANOVA = analysis of variance.

$p = .00009$  for the model that only includes affiliation). As is common for count data, a square root transformation (Bartlett, 1936) resolves this problem (the same was found by Slater et al. (2013)). Hence, we work with the square root of the number of interventions and the ANOVA results are shown in Table 5. The residual errors of this model satisfy normality [Shapiro–Wilk  $p = .20$ ].

Confidence intervals for all pairwise comparisons of marginal means were computed with an overall 95% confidence level (using Scheffé multiple comparisons). In line with what

is shown in Table 5, the main effect difference between in-group and out-group affiliation did not include 0 ( $-2.42$  to  $-0.60$ ). In addition, the confidence interval for the difference in the means of the conditions (in-group and discourage) and (out-group and encourage) was negative ( $-3.79$  to  $-0.05$ ). All other intervals included 0.

Since neither the interaction term nor encouragement as a main effect contribute to the fit, we can delete these terms from the model, considering only Affiliation. In this case  $F(1,38) = 11.5$ ,  $p = .002$ ,  $R^2 = \eta^2 = 0.23$ . As can be seen from  $R^2$ , the overall goodness of fit remains the same, which additionally shows the noncontribution of encouragement and the interaction term, and the residual errors satisfy normality (Shapiro–Wilk  $p = .35$ ). Figure 3a shows the box plot of the number of interventions by affiliation in order to demonstrate their distribution. It can be seen that apart from few potential outliers (which do not militate against the normality of the model that uses the square root transformation) the vast majority of the distributions conform to the findings of the ANOVA. For example, the whole of the interquartile range of the in-group condition is smaller than the lowest quartile of the out-group condition.



**Figure 3.** Box plots showing for the number of interventions (a) by affiliation of the bystanders (out-group and in-group) and encouragement. (b) A condition with no bystanders. The thick horizontal lines are the medians, the boxes are the interquartile ranges (IQR). The whiskers extend from max (min value, 25th quartile –  $1.5 \times \text{IQR}$ ) to min (max value, 75th quartile +  $1.5 \times \text{IQR}$ ).

In order to remove the potential effect of the outliers we used also the nonparametric Wilcoxon rank-sum test for the null hypothesis of equal medians between out-group and in-group. Since this depends only on ranks here, it makes no difference whether the number of interventions or their square root is used. The Wilcoxon test rejects the null hypothesis ( $z = 3.085$ ,  $p = .002$ ). Hence overall, the estimated model fits well what can be seen in Figures 2a and 3a, supporting the idea that when the bystanders are in-group the amount of helping behavior is less than when they are out-group.

### Comparison With Previous Results

In order to directly address the bystander effect itself, we can compare these data with those from the experiment reported by Rovira et al. (2013). In that experiment, carried out about 2 weeks prior to the one described in this paper, there were 10 other participants, again all Arsenal supporters, who experienced the identical scenario using the same equipment, except for the fact that there were no bystanders and as in the current experiment the victim was depicted as an Arsenal supporter. The mean and *SE* of the number of interventions is shown in Figure 2b, and median and interquartile range in Figure 3b. It is clear that the result is close to the out-group affiliation condition. A one-way ANOVA for the square root of the number of interventions has  $F(2,47) = 5.91$ ,  $p = .005$ ,  $R^2 = 0.20$ , Shapiro–Wilk  $p = .23$ . An overall 95% confidence interval for all mean differences between the conditions (Scheffé) is  $-2.63$  to  $-0.39$  for in-group–out-group, but  $-1.92$  to  $.83$  for no

bystanders–out-group and  $-0.41$  to  $2.34$  for no bystanders–in-group.

Finally, in this section, we note that the only nonsignificant results are with respect to the encouragement factor and its interaction with affiliation. The sample size may be considered to be adequate to detect a difference for encouragement because with quite modest assumptions of how the factors might have influenced the results in comparison to the earlier study (Rovira et al., 2013), the a priori power with respect to the analysis of the number of interventions can be computed to be between 0.64 and 0.87 as explained previously in this section.

### Questionnaire Responses

The conditions (affiliation and encouragement) had no noticeable effect on any of the questionnaire responses. However, the question “ShouldStopit” (the feeling that the fight should be stopped) is positively correlated with the number of interventions (independent of condition) (Spearman’s  $\rho = 0.44$ ,  $p = .001$ ,  $n = 50$ ). If we add this as a covariate in the ANCOVA of the square root of the number of interventions on affiliation ( $n = 50$ ), its coefficient is  $0.35 \times 0.10$  (*SE*), with 95% confidence interval 0.12 to 0.56, partial  $\eta^2 = 0.17$ , with overall  $R^2 = 0.34$  (Shapiro–Wilk  $p = .76$ ).

### Discussion

Other things being equal, the social identity of bystanders has an important effect: their shared group affiliation (in-group) with the participant is associated with less helping behavior compared to when the bystanders are out-group. This finding extends our understanding of the way social identity can impact on bystander behavior. For example, the review by Levine and Manning (2013) suggested that the presence of in-group bystanders increases the capacity of the group to influence any particular member—in line with the norms and the values of the group. When the group favors intervention, then in-group bystanders should enhance this tendency. When the group favors inaction, then individuals in the group should be less likely to intervene. More specifically, previous experimental work (Levine et al., 2002) demonstrated the potential for in-group members to enhance, as well as inhibit, the likelihood of helping in an emergency. In Levine et al.’s (2002) Study 1, they showed that bystanders to violence (viewed as a CCTV clip) who encourage or discourage intervention are only influential to the extent that they are viewed as in-group members. In this study, we see that despite the communication manipulation being unsuccessful, there is clear evidence for the inhibitory effect of the presence of in-group bystanders irrespective of the attempts to encourage or discourage.

There are a number of important differences between the design of the study (Levine et al., 2002) and this study (see

Supplemental Table S3). However, the key difference between that study and this is the inclusion of the group membership of the victim. In this study, participants had interacted with the victim prior to the onset of the violent emergency. They do not do so in the experiment of Levine et al. (2002). During the interaction in this study, participants establish common group identity with the victim (they are both Arsenal fans). When the attack happens in front of the other bystanders, the participants need to consider not only their relationship to the victim but also to the other bystanders. This more dynamic and complex set of identity relationships results in the clear effect of group membership on their likelihood of intervention. When those bystanders are in-group, the likelihood of intervention is lower compared to previously found helping levels of in-group victims in the absence of bystanders. When bystanders are out-group—then helping remains at similar levels to previous studies of helping of in-group victims in the absence of bystanders. This suggests a diffusion of responsibility effect in the presence of other in-group members who might also be expected to help.

Our findings therefore help to develop understanding of how diffusion of responsibility works by combining elements of both the classic bystander effect and social identity theory (Tajfel 1974). Classic diffusion of responsibility predicts that diffusion will be distributed across the numbers of others present, irrespective of their psychological relationship to each other. Based on the current findings, we argue that the social identity of the bystanders changes the participant's perception of responsibility. More specifically, we argue that responsibility will mainly be diffused across those who are perceived by the participant to have an equal responsibility to help. When bystanders are out-group to the participant (in a context where both participant and victim are in-group), they will not be seen to have the same responsibility to act, and thus diffusion of responsibility is less likely to occur. In the context of this study, when participants face a clear violent emergency, with the knowledge that the out-group members are unlikely to help, it falls squarely and only on the shoulders of the participant to help the victim.

A second important aspect of this paper is in its use of VR to study bystander behavior in violent emergencies. The meta-analysis of Fischer et al. (2011) argues that the bystander effect does not hold in violent or dangerous emergencies. However, because of both ethical and practical limitations, there has been very little work that has studied the actual behavior of participants during these events (as opposed to collecting self-report or retrospective data). As people tend to respond realistically to virtual events and situations, VR is useful for studies in social psychology—as was pointed out in (Blascovich et al., 2002; Pan & de C Hamilton, 2018). Rather than use human actors, virtual characters perform identically in each condition of the experiment, the environment is completely

under the control of the computer program, written once and for all for a particular study, and it does not require physical setups such as particular spaces. As in this experiment a virtual bar study can take place in a small office and does not require a visit to an actual bar. Moreover, today the cost of good quality VR equipment is less than the cost of many smartphones.

One key aspect of this study was the unsuccessful attempt to manipulate norms of encouragement or discouragement by in-group virtual bystanders. Supplemental S1 Video shows that the encouraging/discouraging statements were clear and should have been heard by the participants. We can consider three reasons why the bystander encouragement statements had no effect. First, the salience of the bystander interventions was low—because the bystanders only made two comments that would encourage or discourage intervention. Given the emotionally charged situation of the attack on the victim by the perpetrator, it is possible that in spite of the presence of the bystanders, a great deal of attention was paid to the actual confrontation, and while the comments of the bystanders should have been heard they were not processed. As reported above, approximately half of the participants did not notice whether the bystanders tried to encourage or dissuade against intervention.

Second, the nature of the scenario was one that bordered on violence. In this situation, and part of the advantage of using VR in these types of studies, is that when participants are faced with life-sized human characters in a surrounding 3D environment, this may produce an overwhelming need to decrease arousal discomfort—for example, as illustrated by the stress exhibited by participants in the virtual reprise of one of the conditions of Stanley Milgram's obedience studies (Slater et al., 2006) and more recently (Gonzalez-Franco et al., 2018; Neyret et al., 2020). Thus, concern with arousal reduction might produce different behavior to situations where participants are just required to express an opinion about intervention. Hence the use of VR is advantageous for controlling conditions in complex social encounters and also to depict scenarios that are not possible to study experimentally in physical reality. However, this is also a disadvantage because it means that there is no experimental ground truth against which to compare results from VR experiments. Nor can there ever be such ground truth precisely because those experiments cannot be carried out in physical reality. An alternative would be to radically changing the frame so that it is no longer really about bystander responses to violent behavior but about something else, for example, whether people attempt to help an injured person (Levine et al., 2002). In these circumstances, the results from VR studies can be used to build predictive theory, which can be tested against in real-life observational studies.

A third way in which responses to bystanders in VR might be different is in terms of perceived efficacy of the other bystanders. Work on collective action (Van Zomeren et al., 2008) shows that judgments about the efficacy of

others play an important part in individuals' decisions to act. While virtual humans may be able to signal group membership, or to create emotionally charged environments, they were not programmed to actually intervene in this scenario. Thus, participants could not have expected any support from fellow bystanders should they have chosen to intervene. In our study, the bystander characters are possible objects on which to diffuse responsibility—particularly when trying to reduce anxiety. However, it remains to be seen whether the possibility of actual physical support from bystanders could create conditions that enhance intervention in violence.

In generalizing the findings of this study, a further issue should be taken into consideration—that all participants were male. This was for practical reasons in recruiting a sufficient number of supporters of the Arsenal football club. A meta-analysis conducted by Eagly and Crowley (1986) indicated gender differences in the extent to which men and women help in emergencies and the nature of help that they provide. The bar context in which this study is set, and the violent nature of the emergency may therefore not be generalizable to both genders. Specifically, it is likely that there are different social norms for the conduct of men and women in such an environment and different expectations regarding intervention in a violent altercation. Future studies will need to include female and transgender participants, and consider ethnicity, cultural background, and factors related to personality to observe any potential effect on the results. In addition, the results will need to be tested with supporters from other football teams and from other sports as well. The only reason why we chose a football-related experience was because the strong sense of identity that supporters have with their team, and as a continuation of previous work. Other scenarios could provide a wider variety of data from different cases of social identity. In future studies, the method of counting the number of interventions could be more sophisticated. Here we counted discrete interventions allowing 2 seconds between each one. Hence, one intervention that lasted for example 5 seconds would be counted as equivalent to another that lasted only 1 second. Moreover, interventions can vary in many ways—such as the loudness of voice in a verbal intervention, or the velocity and position of movement in a physical intervention. Hence more robust methods are needed to assess not just the frequency but also the type, duration, and quality of the interventions.

A critical missing element in VR studies is the lack of the possibility of physical consequences to an action—so that the participant can have no rational fear of being physically harmed by the perpetrator on intervention. This is not to say that there may not still be some fear simply based on the perceived situation, or fear of a verbal attack. However, it is possible for there to be an interactive element whereby when the participant intervenes the

perpetrator responds aggressively to the participant, and even some level of haptic feedback where the participant can feel friendly or aggressive touch from the victim and perpetrator. Adding this element of greater physicality is an important way forward in this methodology. Finally, the results of this experiment should be treated as having generated a new empirically grounded hypothesis that would need further studies for verification: diffusion of responsibility in the bystander effect is modulated by salient bystander group identity, other things being equal. More specifically, in the situation where this salient group identity is shared between the victim and a specific bystander, that bystander is more likely to intervene when other bystanders are out-group, since then the only one with the responsibility to intervene is that individual. When the other bystanders are in-group then responsibility is equally shared, and thus the individual is less likely to intervene. The “other things being equal” is important here. For example, for this hypothesis to be valid, the other bystanders, whether in-group or out-group, must be equal in status to the individual, apart from the issue of the salient factor through which group identity is determined. In particular, other bystanders when out-group should not be perceived as posing a threat to the individual. There are still many other considerations here: what proportion of bystanders in a crowd need to be perceived as out-group (or in-group) for the individual to intervene (or not)? Further work is also required on the issue of encouragement. We suggest that VR provides a powerful tool to answer such questions because these are practically and ethically impossible to address with human actors.

## Conclusions

This article demonstrates the significant contribution that VR can make to the study of human behavior in ethically challenging circumstances. It contributes to the literature demonstrating how VR can facilitate the use of the experimental method to study controversial topics with experimental rigor. The development of VR scenarios and the behavior of participants in these VR environments, allows for experimental work with high internal and ecological validity. It is through the strengths of this approach that the paper makes a significant contribution to theory in the social psychology of bystander behavior. By being able to study the interaction of social identity processes and bystander behavior, we are able to develop our understanding of the concept of diffusion of responsibility. This study shows how the presence of in-group bystanders can reduce helping of in-group victims (compared to the helping of in-group victims in the presence of out-group bystanders). Taken together, the paper points to the continued importance, both empirically and theoretically, of being able to study challenging real-world social psychological questions in VR.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This research was supported by the UK EPSRC Project EP/F032420/1 “*Visual and Behavioural Fidelity of Virtual Humans with Applications to Bystander Intervention in Violent Emergencies.*” A.R. is supported by the NIHR Oxford Health Biomedical Research Center BRC-1215-20005. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. M.S. is supported by the European Research Council Advanced Grant MoTIVE #742989.

## Ethical Statement

This study was carried out in accordance with the recommendations and approval of the University Research Ethics Committee with written informed consent from all subjects.

## ORCID iD

Aitor Rovira  <https://orcid.org/0000-0001-9308-5784>

## Data Availability

The raw data are available as Supplemental Material.

## Supplemental Material

Supplemental material for this article is available online.

## References

- Bartlett, M. (1936). The square root transformation in analysis of variance. *Supplement to the Journal of the Royal Statistical Society*, 3(1), 68–78. <https://doi.org/10.2307/2983678>
- Blascovich, J. (2002). Social influence within immersive virtual environments. In R. Schroeder (Ed.), *The social life of avatars* (pp. 127–145). Springer.
- Blascovich, J., Loomis, J., Beall, A. C., Swinth, K., Hoyt, C., & Bailenson, J. N. (2002). Immersive Virtual Environment Technology as a methodological tool for Social Psychology. *Psychology Inquiry*, 13, 103–124. [https://doi.org/10.1207/S15327965PLI1302\\_01](https://doi.org/10.1207/S15327965PLI1302_01)
- Borofsky, G. L., Stollak, G. E., & Messe, L. A. (1971). Sex differences in bystander reactions to physical assault. *Journal of Experimental Social Psychology*, 7, 313–318. [https://doi.org/10.1016/0022-1031\(71\)90031-X](https://doi.org/10.1016/0022-1031(71)90031-X)
- Burn, S. M. (2009). A situational model of sexual assault prevention through bystander intervention. *Sex Roles*, 60(11–12), 779–792. <http://dx.doi.org/10.1007/s11199-008-9581-5>
- Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *SIGGRAPH '93 Proceedings of the 20th annual conference on Computer graphics and interactive techniques* (pp. 135–142). <http://dl.acm.org/citation.cfm?id=16613>
- Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies—Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8, 377–383.
- Eagly, A. H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, 100(3), 283–308.
- Fischer, P., Greitemeyer, T., Pollozek, F., & Frey, D. (2006). The unresponsive bystander: Are bystanders more responsive in dangerous emergencies? *European Journal of Social Psychology*, 36, 267–278.
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., & Frey, D. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, 137, 517–537. <http://psycnet.apa.org/journals/bul/137/4/517/>
- Gonzalez-Franco, M., Slater, M., Birney, M., Swapp, D., Haslam, S. A., & Reicher, S. D. (2018). Participant concerns for the Learner in a Virtual Reality Replication of a Milgram Obedience Study. *PLOS ONE*, 13(12), Article e0209704. <https://doi.org/doi.org/10.1371/journal.pone.0209704>
- Gottlieb, J., & Carver, C. S. (1980). Anticipation of future interaction and the bystander effect. *Journal of Experimental Social Psychology*, 16(3), 253–260.
- Harari, H., Harari, O., & White, R. V. (1985). The reaction to rape by American male bystanders. *The Journal of Social Psychology*, 125, 633–658.
- Held, R. M., & Durlach, N. I. (1992). Telepresence. *PRESENCE: Teleoperators and Virtual Environments*, 1, 109–112.
- Hopkins, N., Reicher, S., Harrison, K., Cassidy, C., Bull, R., & Levine, M. (2007). Helping to improve the group stereotype: On the strategic dimension of prosocial behavior. *Personality and Social Psychology Bulletin*, 33(6), 776–788.
- Latane, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, 10(3), 215–221. [doi:http://dx.doi.org/10.1037/h0026570](http://dx.doi.org/10.1037/h0026570)
- Latane, B., & Darley, J. M. (1969). Bystander apathy. *American Scientist*, 57, 244–268.
- Latané, B., & Rodin, J. (1969). A lady in distress: Inhibiting effects of friends and strangers on bystander intervention. *Journal of Experimental Social Psychology*, 5, 189–202.
- Levine, M. (1999). Rethinking bystander nonintervention: Social categorization and the evidence of witnesses at the James Bulger murder trial. *Human Relations*, 52, 1133–1155.
- Levine, M., Cassidy, C., Brazier, G., & Reicher, S. (2002). Self-categorization and bystander non-intervention: Two experimental studies. *Journal of Applied Social Psychology*, 32(7), 1452–1463.
- Levine, M., & Crowther, S. (2008). The responsive bystander: How social group membership and group size can encourage as well as inhibit bystander intervention. *Journal of Personality and Social Psychology*, 95, 1429–1439.
- Levine, M., & Manning, R. (2013). Social identity, group processes, and helping in emergencies. *European Review of Social Psychology*, 24(1), 225–251.
- Levine, M., Prosser, A., Evans, D., & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin*, 31, 443–453.

- Levine, M., Taylor, P. J., & Best, R. (2011). Third parties, violence, and conflict resolution: The role of group size and collective action in the microregulation of violence. *Psychological Science, 22*, 406–412.
- Levine, M., & Thompson, K. (2004). Identity, place, and bystander intervention: Social categories and helping after natural disasters. *Journal of Social Psychology, 144*, 229–245.
- Liebhart, E. H. (1972). Empathy and emergency helping: The effects of personality, self-concern, and acquaintance. *Journal of Experimental Social Psychology, 8*(5), 404–411.
- Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods Instruments & Computers, 31*, 557–564. <https://doi.org/10.3758/BF03200735>
- Manning, R., Levine, M., & Collins, A. (2007). The Kitty Genovese murder and the social psychology of helping: The parable of the 38 witnesses. *American Psychologist, 62*, 555–562.
- Neyret, S., Oliva, R., Beacco, A., Navarro, X., Valenzuela, J., & Slater, M. (2020). *An embodied perspective as a victim of sexual harassment in virtual reality reduces action conformity in a later milgram obedience scenario* [Scientific reports]. <https://doi.org/10.1038/s41598-020-62932-w>
- O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav Brain Sci, 24*, 939–1031.
- Pan, X., & de C Hamilton, A. F. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology, 109*, 395–417. <https://doi.org/10.1111/bjop.12290>
- Pomes, A., & Slater, M. (2013). Drift and ownership towards a distant virtual body. *Frontiers in Human Neuroscience, 7*. <https://doi.org/10.3389/fnhum.2013.00908>
- Reicher, S., Cassidy, C., Wolpert, I., Hopkins, N., & Levine, M. (2006). Saving Bulgaria's Jews: An analysis of social identity and the mobilisation of social solidarity. *European Journal of Social Psychology, 36*(1), 49–72.
- Rovira, A. (2016). *Simulating social situations in immersive virtual reality—a study of bystander responses to violent emergencies*. University College London.
- Rovira, A., Swapp, D., Southern, R., Zhang, J. J., & Slater, M. (2013, September 9). *The impact of enhanced projector display on the responses of people to a violent scenario in immersive virtual reality* [Paper presentation]. The Virtual Reality (VR), IEEE. <https://www.computer.org/csdl/proceedings-article/vr/2013/06549350/12OmNzb7Znq>
- Rovira, A., Swapp, D., Spanlang, B., & Slater, M. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Frontiers in Behavioral Neuroscience, 3*, 1–10. <https://doi.org/10.3389/neuro.08.059.2009>
- Rutkowski, G. K., Gruder, C. L., & Romer, D. (1983). Group cohesiveness, social norms, and bystander intervention. *Journal of Personality and Social Psychology, 44*(3), 545–552.
- Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience, 6*, 332–339.
- Schwartz, S. H., & Gottlieb, A. (1976). Bystander reactions to a violent theft: Crime in Jerusalem. *Journal of Personality and Social Psychology, 34*, 1188–1199.
- Schwartz, S. H., & Gottlieb, A. (1980). Bystander anonymity and reactions to emergencies. *Journal of Personality and Social Psychology, 39*, 418–430.
- Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *PRESENCE: Teleoperators and Virtual Environments, 1*, 120–126.
- Shotland, R. L., & Straw, M. K. (1976). Bystander response to an assault: When a man attacks a woman. *Journal of Personality and Social Psychology, 34*, 990–999.
- Slater, M. (2009). Place Illusion and Plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B, 364*, 3549–3557. <https://doi.org/10.3389/frobt.2016.00074>
- Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N., & Sanchez-Vives, M. V. (2006). A virtual reprise of the Stanley milgram obedience experiments. *PLOS ONE, 1*, Article e39. <https://doi.org/10.1371/journal.pone.0000039>
- Slater, M., Rovira, A., Southern, R., Swapp, D., Zhang, J. J., Campbell, C., & Levine, M. (2013). Bystander responses to a violent incident in an immersive virtual environment. *PLOS ONE, Article e52766*. <https://doi.org/10.1371/journal.pone.0052766>
- Staudt, A. (2014). *CIBAR: Stata module to plot bar graphs and confidence intervals over groups*. <https://ideas.repec.org/c/boc/bocode/s457805.html>
- Tajfel H. (1974). Social identity and intergroup behaviour. *Social Science Information, 13*(2), 65–93. <https://doi.org/10.1177/053901847401300204>
- Van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin, 134*(4), 504–535.