Cooperation versus competition in a mass emergency evacuation: A new laboratory simulation and a new theoretical model

JOHN DRURY
University of Sussex, Brighton, England

CHRIS COCKING
London Metropolitan University, London, England

STEVE REICHER
St. Andrews University, Fife, Scotland

ANDY BURTON
University of Nottingham, Nottingham, England

DAMIAN SCHOFIELD
RMIT University, Melbourne, Victoria, Australia

ANDREW HARDWICK
University of Sussex, Brighton, England

DANIELLE GRAHAM
St. Andrews University, Fife, Scotland

AND

PAUL LANGSTON
University of Nottingham, Nottingham, England

Virtual reality technology is argued to be suitable to the simulation study of mass evacuation behavior, because of the practical and ethical constraints in researching this field. This article describes three studies in which a new virtual reality paradigm was used, in which participants had to escape from a burning underground rail station. Study 1 was carried out in an immersion laboratory and demonstrated that collective identification in the crowd was enhanced by the (shared) threat embodied in emergency itself. In Study 2, high-identification participants were more helpful and pushed less than did low-identification participants. In Study 3, identification and group size were experimentally manipulated, and similar results were obtained. These results support a hypothesis according to which (emergent) collective identity motivates solidarity with strangers. It is concluded that the virtual reality technology developed here represents a promising start, although more can be done to embed it in a traditional psychology laboratory setting.

Virtual reality (VR) technology has been argued to be useful in a number of domains of psychological research. This article reports three studies applying this technology for the first time to the field of emergency evacuation behavior, in order to test a new model of such behavior based on the social identity approach.

Early research on mass evacuation behavior relied on anecdotal evidence and data from the military (e.g., Strauss, 1944). In the 1950s and 1960s, however, the introduction of new laboratory experimental techniques helped in the development of important theoretical advances. Thus, for example, Mintz's (1951) classic simulation study used a bottle containing several corks on strings held by participants, representing the potential for exit jamming that would take place if every individual tried to evacuate simultaneously. Mintz suggested that ineffectual escape in an evacuating crowd is due to individual calculation of costs and benefits, rather than to a contagious outburst of mass irrationality, as assumed by the early *mass panic* models (e.g., McDougall, 1920;

J. Drury, j.drury@sussex.ac.uk

Ross, 1908; see Chertkoff & Kushigian, 1999, for a review of mass panic theories).

Subsequent laboratory simulations were more sophisticated than Mintz's (1951) design and were marked by increased use of digital technology, such as computerized light displays to represent levels of crowding. On the basis of this research, a number of variables have been identified as playing a role in the extent of competition versus cooperation in escape behavior in emergency evacuations. These include group size and perceived threat (Kelley, Condry, Dahlke, & Hill, 1965), exit time (Chertkoff, Kushigian, & McCool, 1996), leadership (Klein, 1976), and "rational" contingencies and emotional arousal (Gross, Kelley, Kruglanski, & Patch, 1972).

However, in the last 3 decades, the field has largely been abandoned by experimental social psychologists. Before describing in detail the VR experimental paradigm employed in the present studies, we will outline how this abandonment came about and, hence, explain why a new experimental paradigm is needed.

Methodological Problems and "Solutions"

The key methodological issue confronting the development of experimental studies of emergency evacuation is that of simulating the impending threat of death. It has been argued that, without some kind of strong motivation provided by the experimenter to approximate this in some way, participants in emergency simulations will not act in the ways observed in real life (Muir, 1996), so the processes behind these behaviors cannot be studied. Classic laboratory studies, such as that by Kelley et al. (1965), induced motivation through a (bogus) threat in their experimental participants. Experimenters might claim that the social and applied importance of such studies outweighs any possible suffering on the part of participants. However, ethical guidelines, from both the American Psychological Association and the British Psychological Society, have become tighter since the 1960s. Classic experimental simulations of mass evacuations would, therefore, now risk falling foul of the stipulation that such studies are acceptable only when there is no stress-free methodological equivalent.

In the absence of a laboratory method that was both psychologically engaging and yet ethically sound, the field of mass evacuation behavior has largely been left to sociologists and others, who have used archival data rather than experiments (e.g., Aguirre, Wenger, & Vigo, 1998). Arguably, therefore, it is outside of experimental social psychology that most of the contributions to the field have been made in recent years. In particular, the normative approach, which stresses the continuity between everyday rule-governed behavior and that in emergency situations (Johnson, 1987a, 1987b, 1988), and affiliation theory (Mawson, 2005; Sime, 1983), which emphasizes the importance of preexisting interpersonal ties in limiting panic responses, represent the current state of theory in this area (Aguirre, 2005). For mainstream social psychology, therefore, the topic of mass evacuation behavior represents a theoretical lacuna that needs to be addressed. This absence is particularly striking given pertinent theoretical advances in the wider field of group processes made in the past 30 years, where the social identity approach has provided explanations across the range of group-based phenomena (Turner, 1999; Turner & Oakes, 1997).

Mathematicians and engineers have also filled the gap left by experimental social psychologists. In both academia and applied settings, increased use has been made of computer simulations of crowd behavior to predict and explain the contours of mass evacuations (e.g., Helbing, Farkas, & Vicsek, 2000). However, it has been argued that, for any sense of realism, such computer models need not only human data (including that from laboratory simulations), but also accurate and up-to-date psychological theory (Langston, Masling, & Asmar, 2006; Still, 2000).

A Social Identity Approach to Mass Evacuation Behavior

Kugihara (2001), an engineer, developed a sophisticated computer-based experimental simulation of a mass evacuation situation. Each participant sat in a booth and watched a display screen of lights indicating whether others were aggressive or made concessions toward each other as they tried to escape from a painful stimulus. The evidence from his studies showed that responses in such contexts could be understood in terms of (aggressive or concessive) group norms (i.e., rules for conduct), which were more salient in larger than smaller groups. These group norms were in turn explained in terms of salient social identities.

As well as the methodological innovation involved, this research was important for its attempt to apply the social identity approach (Tajfel & Turner, 1979) for the first time to the field of evacuation behavior. The social identity approach has become the dominant perspective in group processes and intergroup relations. But the field of evacuations and other mass emergency behavior has remained relatively untouched by it.

The present study followed Kugihara (2001) in applying the social identity approach to mass emergency evacuation behavior. However, we sought to go beyond his work both methodologically and in terms of theory. Rather than being concerned with the contingencies of social identity *contents* (i.e., the different behavioral outcomes of different identity-based norms), the studies described in this article focused on the logic of social identity *processes*. In particular, we sought to draw upon the principles of self-categorization theory (SCT; Turner, 1982, 1985; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) to explain the fact that cooperation, helping, and other expressions of solidarity are common among strangers in mass evacuations (e.g., Blake, Galea, Westeng, & Dixon, 2004; Chertkoff & Kushigian, 1999; Clarke, 2002; Drury, Cocking, & Reicher, in press; Proulx & Fahy, 2003).

What is of theoretical interest in such behavior is the fact that it carries a cost to the helper—in exit delay and even risk to personal safety. The social identity approach and SCT in particular suggests a possible explanation. According to SCT, cognitive representations of the self

take the form of self-categorizations, which may range in inclusivity from the personal (definitions of what makes us unique) to the shared, collective, and even universal (definitions that classify us with others). Collective identity—that is, seeing oneself as interchangeable with other ingroup members on some relevant dimension—means seeing these other ingroup members as, in some sense, part of self. This, in turn, means caring about them and acting in their interests, even where these others are not known or even personally liked.

SCT principles have been applied successfully to explain emergency-related phenomena such as crowd conflict (Reicher, 2001) and mutual support among group members (Haslam, O'Brien, Jetten, Vormedal, & Penna, 2005). Moreover, experimental studies of helping behavior are in line with the hypothesis put forward here that shared ingroup membership enhances cooperation (Levine, Cassidy, Brazier, & Reicher, 2002; Levine, Prosser, Evans, & Reicher, 2005; Levine & Thompson, 2004). Furthermore, a recent comparative interview study of 11 mass emergency events showed that high-identification survivors saw, received, and gave more aid to others than did those low in identification (Drury et al., in press). However, these experimental studies of helping were not simulations of emergency evacuations; and the comparative interview study, by its nature, lacks experimental control. Mass panic theories would suggest that the presence of a crowd and the increased risk of death or injury incurred by helping others would each reduce cooperative behavior and increase competition in an emergency evacuation. Hence, there is a need for an experimental test of the hypothesis that collective identity in a mass evacuation enhances cooperation and reduces competition among crowd members.

Experimental control (manipulation and control of relevant variables) and behavioral measures need to be combined with a design that avoids falling foul of current guidelines on the prevention of distress to research participants. We therefore sought to design a computer-based simulation that engaged participants, yet without threatening them with pain. To that end, the Underground Station Evacuation Simulator (USES; Burton & Schofield, 2005) was developed.

Visualization Methodology:
The Underground Station Evacuation Simulator

The USES (Burton & Schofield, 2005) is an interactive computer simulation that visualizes an emergency evacuation at an underground railway station and presents the user with opportunities for cooperating and competing with others. The simulation is modeled using computer game technology (i.e., using similar graphical techniques, physics engines, and user interfaces). The graphical environment within the evacuation simulation was created using 3-D computer-aided design packages similar to those used both in the computer games industry and to create animated models in films such as *Shrek* or *The Incredibles*.[1] The models are created as a 3-D geometric mesh that is later wrapped in (or textured with) photo-

Figure 1. A still from the VR simulation.

graphic images that give the realism achieved in the final models (see, e.g., Figures 1–3).

The simulated evacuation environment is, in fact, more like a 3-D computer game than an animated film. Animated feature films have short sequences that last a few seconds that may take a standard PC hours or even days to render (to create the individual frames of the animation). The evacuation environment, on the other hand, is rendered in real-time as a VR world. It is possible for the user to navigate around the environment freely to view the evacuation from other angles and to interact with virtual characters in the simulation.

In order to achieve the real-time rendering of the world, a technique called *low-polygon modeling* was used.[2] Although, in reality, objects are made up of a lot of very fine detail in terms of their shape, it is often possible to reduce this complexity by ignoring intricate details. Neglecting fine detail can potentially reduce the realism of the world, but this compromise increases the ability for a computer to process the positions of the hundreds of thousands of
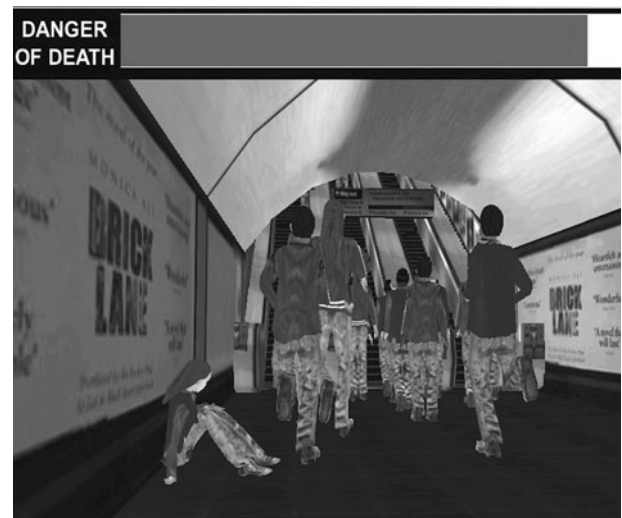

Figure 2. A still from the simulation depicting an evacuee in need of help.

Figure 3. A still from the simulation depicting the user about to push another evacuee aside.

polygons that make up the 3-D virtual environment and render them to the screen. A low-polygon model, textured using digital images (in this case, of an underground rail station) that can be cleverly mapped to the polygonal faces of the object, can replace the detailed information lost by simplifying the mesh geometry. This gives the look of a detailed 3-D model, while retaining relatively low computational requirements.

At the front end, the *task* facing the simulation user is to evacuate the station as soon as possible, while facing bottlenecks caused by the rest of the crowd. Users are able to push characters aside (any number of times) at the press of a key. The amount of pushing is intended as a measure of competition. Users also have to make decisions about whether to stop and help (or to ignore) four separate people (two male, two female) that they encounter during the evacuation, who are apparently injured. The number helped is intended as a measure of cooperation. Faced with one of the four *victims* in need, instructions on the screen invite the user to help or ignore him/her, again at the press of a key. If they choose to ignore the victim, they continue on their way. If they indicate that they will help, some text appears on the screen thanking them for doing so. To represent the personal costs incurred by helping in a real-life mass evacuation, instances of stopping to help a fallen character add to the exit time and, hence, the risk of death for the user. Each user's responses, interactions, and timings are automatically recorded and saved in a data file.

The simulation was designed to be configurable, which allowed a range of experimental conditions to be experienced by different users. A number of key dimensions and system variables can be varied through drop-down menus at the beginning of the program, such as the appearance (clothes) of the characters in the evacuation (red, blue, mixed, or neutral) and the number of other evacuees. To enhance the urgency of exit, a "danger-of-death" bar can be displayed at the top of the screen, the value of which increases over time but reduces as the participant nears the exit and safety. There is also a soundtrack of ambient sounds from an underground railway.

The USES was designed to be embedded within a traditional social psychology laboratory experiment. Thus, the participants' identity can be manipulated by a vignette at the beginning of each trial (verbally, in a written text, or as a PowerPoint presentation) that casts them as either psychological crowd members or individuals in an aggregate crowd.[3] A posttest questionnaire can be used to assess relevant self-report variables—in particular, manipulation checks and feelings toward others.

This VR paradigm avoids use of real or bogus threats of pain to prompt urgency of exit among participants by relying, instead, upon psychological *immersion* or engagement with the interactive visualization program. As such, the methodology is in part based on role-play techniques where participants are asked to cooperate in imagining a scenario (e.g., Ginsburg, 1979). The present methodology goes beyond this, however, through the use of graphics and computer game technologies that many people are familiar with and enjoy using. VR has been argued to be a useful methodological tool in a number of domains—primarily, for its ability to enhance ecological validity while maintaining the control associated with the experimental method (Loomis, Blascovich, & Beall, 1999). More important, it also allows for empirical research in domains in which there might otherwise be ethical restrictions. For example, Slater et al. (2006) used an immersive virtual environment to successfully replicate the classic study of obedience by Milgram (1963); previously, the perceived trauma caused to participants who believed that they had harmed or killed another participant meant that the Milgram paradigm had been impossible to reproduce.

The present study is the first application of VR techniques, so far as we are aware, to the field of mass evacuation or any other form of crowd behavior. Again, part of the rationale for the application of this methodology is the fact that, as has been discussed, real-life emergencies involve stress and suffering. As such, it represents a unique attempt to develop a new method appropriate to the difficult but important theoretical question of explaining cooperation versus competition in a mass evacuation.

Overview

Three studies were carried out using the USES to (1) explore the scope of this methodological technique for investigating aspects of mass evacuation behavior and (2) test the usefulness of the social identity approach to mass emergency evacuation behavior in relation to other theoretical models.

Study 1 was carried out in an immersion laboratory. Conditions of common (vs. personal) fate were compared with a control (1) as antecedents of collective identification and (2) in terms of cooperation (amount of helping offered, time taken to offer help) and competition (i.e., amount of pushing). In Study 2, high and low group identifiers were compared on amount of helping offered, pushing, and concern for others. Possible mediation relationships between group identification, concern, and (1) helping and (2) competition were also assessed, as was the relationship

between the perceived danger of death and the amount of help offered. In Study 3, both collective identity and group size were manipulated to determine their independent and interactive effects on both helping and pushing. Mediation analyses were again carried out, this time using group liking as a possible mediator. The effect of danger of death on both helping and pushing was assessed.

STUDY 1

In a first study, the VR paradigm was used to explore the possible antecedents, as well as the behavioral consequences, of collective identification.

In terms of antecedents, according to SCT, one possible criterion leading otherwise disparate individuals to define themselves in terms of a collective identification is a *common fate*: a perceived relationship to an external force, group, or agent, within which members' fortunes are seen as one (Turner et al., 1987; cf. Campbell, 1958). Put differently, an emergency or disaster can create a sense of "we-ness," creating new social bonds in a crowd of survivors (Clarke, 2002). It is therefore suggested that mass emergency evacuations, like other crowd events (Drury & Reicher, 2000; Reicher, 2001), might broaden people's level of identification, which, as has been shown above, is a possible basis of cooperation.

However, although both field and journalistic evidence support the idea that disasters can both bring people together psychologically and encourage cooperation (e.g., Chertkoff & Kushigian, 1999; Drury et al., in press; Fritz & Williams, 1957; Ripley, 2005), these ideas have not been tested experimentally.

An alternative explanation for enhanced identification and helping in an emergency can be derived from terror management theory (TMT; Greenberg, Solomon, & Pyszczynski, 1997). TMT suggests that one of the qualities that makes us human is the ability to be conscious of our own death and the consequent need to avoid the terror associated with this knowledge. Faced with the thought of personal mortality, people create a sense of psychological endurance that, in turn, provides comfort (Jonas, Schimel, Greenberg, & Pyszczynski, 2002; Osborn, Johnson, & Fisher, 2006). One way to create such psychological endurance is to "cling to the ingroup" (Castano, 2004)—in other words, to increase endorsement of a positively valued social identity and its associated values (which may, e.g., include altruism).

In terms of meta-theory, whereas SCT is a social, *cognitive* theory of the group as a collective self, TMT is a theory of individual motivation. In relation to behavior in life-threatening emergency evacuations, what is needed is a test of whether *common* fate (SCT) or *personal* threat of death (TMT)—or indeed both of them—enhances identification, which, in turn, should prompt more rapid helping and reduce competition. The present experimental paradigm affords precisely a way of doing this. By varying features of the vignette at the beginning of the evacuation, we can manipulate presence and level of threat independently and test their separate effects.

In Study 1, participants in the USES were divided into three conditions: control ("in a hurry to get to the shops"), personal mortality salience (emphasis on personal death), and collective mortality salience (emphasis on common fate through collective death), with self-report measures taken of identification and behavioral measures of cooperation (helping and time taken to help) and competition (pushing). It was expected that there would be more self-reported identification, more cooperation, and less competition in both of the *threat* conditions than in the control.

Method

Design. The design used was a one-way mixed design with three levels. The independent variable was the mortality salience manipulation, with three conditions: control (no mortality salience manipulation); personal mortality salience, with participants instructed to focus on their own death; and collective mortality salience (common fate), with participants instructed to think of the possibility of death that they shared with the others around them.

Participants. Sixty undergraduate students from St. Andrews University took part in this study. Their age ranged from 19 to 25 years. Participation was voluntary.

Dependent measures. Doosje, Ellemers, and Spears's (1995) identification scale was employed, anchored by 1, *strongly agree* (i.e., high identification), to 7, *strongly disagree* (i.e., low identification). Behavioral measures included number of times (out of four) the victims were helped (vs. ignored), the *hesitation* time (in seconds) taken to help the victim, and the number of times people were pushed aside.

Procedure. After being randomly allocated to one of the three conditions, the participants in the two mortality salience conditions were instructed to fill in the Mortality Attitudes Personality Survey (Greenberg et al., 1990). Next, they were told to "relax and clear your mind." They were then instructed to take a few moments to imagine the scene in the 1987 Kings Cross underground rail fire of 1987, to enhance engagement with the visualization.

Next, the individuals were presented with one of three scenarios, depending on their condition, which the experimenter read aloud to ensure that they fully understood. For the control, the scenario was as follows:

1. You have just got off the train and are on a platform in Kings Cross underground station. The station is busy.
2. You and the others around you are hurrying to get out of the station because the sales in Oxford Street have just began and you all want to get some good bargains.
3. You must make your way out of the station as quickly as possible so you do not miss the best deals.

The personal mortality salience condition differed as follows:

3. All of a sudden you see smoke and someone shouts, "There's a fire. You must get out quickly!" The fire is large and your life is seriously in danger. You must get out of the station as quickly as possible.

Finally, the collective mortality salience condition differed as follows:

3. All of a sudden you see smoke and someone shouts, "There's a fire. You must get out quickly!" The fire is large and your life is seriously in danger. You must get out of the station as quickly as possible. While doing this, think of your own death and the deaths of the others around you in the station.

The keyboard facing the participant was connected to three projectors throwing synchronized images of the visualization onto the entire front wall of the lab (4 m wide × 3 m high) and on part of the two side walls. These images filled the participant's entire visual

field. In addition, there was a *surround sound* system of speakers in the room.[4]

Following the visualization, the participants filled in the self-report questionnaire.

Results

Identification. Group identification was found to vary across the conditions [$F(2,57) = 10.41$, $p < .001$, $\acute{\eta} = .27$]. Breaking this down, there was found to be significantly greater identification in the personal mortality condition ($M = 3.35$, $SD = 0.94$) than in the control condition ($M = 5.04$, $SD = 1.25$) [$t(38) = 4.81$, $p < .001$] and significantly greater identification in the collective mortality condition ($M = 3.66$, $SD = 1.48$) than in the control condition [$t(38) = 3.17$, $p = .003$]. However, there was no significant difference between the collective and the personal mortality conditions [$t(38) = 0.79$, $p = .43$]. Together these results showed that threat (whether presented as personal or collective) did indeed create common identification in the emergency itself.

Behavioral measures. Given that personal and collective fate were found to induce common identification to an equal degree, for the behavioral measures, planned comparisons were carried out. It was expected that there would be more help, less delay in offering help, and less pushing in the (combined) mortality conditions than in the control condition. In each case, t tests revealed marginal effects in line with these expectations [help, $t(58) = 1.17$, $p = .13$ (one-tailed); time taken to help, $t(54) = -1.47$, $p = .07$ (one-tailed); and pushing, $t(58) = -1.24$, $p = .11$ (one-tailed); see Table 1 for means and standard deviations].

Discussion

The key finding from this study is the replication in the laboratory of the field observation that features of an emergency evacuation itself can enhance collective identification (Clarke, 2002; Drury et al., in press). This is in line with the prediction based on SCT that common fate can lead individuals to redefine themselves in collective terms in emergencies, just as in other contexts. The comparable results obtained for personal mortality salience are consistent with the role of terror management in endorsing social identification in such circumstances too, however.

The pattern of behavioral evidence, although in the directions expected, was weak. However, the trends that were obtained in this first exploratory study provide at least a prima facie case for considering seriously the hypothesized role of collective identity in enhancing cooperation and reducing competition in a mass emergency evacuation. These results therefore offer us some encouragement for carrying out further studies using this VR paradigm.

On a theoretical level, the next step is to try to improve upon these results. In addition, adding manipulation checks and further measures to a future replication would give us more confidence that the variables specified by SCT are indeed responsible for the pattern of findings obtained here.

Finally, there is also an issue about the practical feasibility of the VR paradigm as implemented in the present study. If we can run the study simply using a PC and yet demonstrate sufficient psychological engagement among participants, the USES can be incorporated into many more research studies, since it would not have to rely on full immersion facilities, which are available in only a minority of psychology department laboratories.

STUDY 2

Study 2 was intended both to measure the level of engagement engendered by the USES methodology (this time run without full immersion facilities) and to compare competing theories of mass evacuation behavior.

Both mass panic theories and modified versions of the normative account suggest that increased perceptions of danger reduce helping behavior in an emergency evacuation (Johnson, Feinberg, & Johnston, 1994; Quarantelli, 1957). Study 2 therefore examined the effects of both danger and identification on helping behavior. We might predict a simple main effect whereby helping would decrease with the increasing rate of danger. However, the effect of danger on helping might be affected by identification in one of two possible ways. The social identity approach would suggest that those in a crowd who define themselves psychologically as part of that crowd would offer more help, even at personal cost. We would therefore expect the absolute amount of cooperation to be greater for high- than for low-identification participants. On top of this, collective identification might moderate the effect of increased danger on cooperation: As danger increases, low-identification participants should help less, whereas this relationship might be attenuated for high-identification participants, who would continue to help to roughly the same degree irrespective of variations in the perceived danger of death.

Any evidence that high identification is associated with general cooperation not only counts against the panic model; it also goes beyond the affiliation approach (Mawson, 2005). The affiliation approach predicts cooperation only with affiliates and, indeed, would suggest that, when the evacuating crowd consists of strangers in an unfamiliar setting, panic (i.e., competitive behaviors) will occur.

According to SCT, on the other hand, in the same way that high identification increases cooperation and reduces competition, it should also enhance the concern participants express for others, as compared with low-identification participants. Moreover, given that concern

Table 1
Means (and Standard Deviations) for Helping,
Time Taken to Help, and Pushing in Study 1 (in Seconds)

	N Helped		Time Taken to Help		*N* Pushed	
	M	SD	M	SD	M	SD
Control	2.85	1.23	1.52	1.82	3.05	2.63
Personal mortality salience	3.05	1.47	1.18	1.31	1.95	1.50
Collective mortality salience	3.40	0.68	0.71	0.99	2.60	2.56

for others is the motivation for cooperation and is a function of collective identification, we would expect there to be a relationship of mediation between these variables. A mediator is a variable that accounts for the relationship between the predictor and the criterion variable (Baron & Kenny, 1986). That is, strength of identification should indirectly affect helping through its direct effect on concern for others. Strength of identification should predict reduced pushing, as should concern for others, although it is unclear whether the relationship between these three variables should be one of mediation.

Method

Participants. Seventy-two students, 45 of them female, from the University of Sussex took part in this study for course credit or money. There were no differences on any of the relevant variables for either of these subgroups. The age of the participants ranged from 18 to 49 years, the mean age being 22.19.

Dependent measures. As in Study 1, behavioral measures included number of times (out of four) the victims were helped (vs. ignored) and number of times people were pushed aside. Helping was measured in relation to danger, indicated as a percentage on the screen danger bar (see Figure 3), which was recorded in the data file.

Self-report measures were taken of concern for others (five items; e.g., "I wanted to help other people in the evacuation") and identification with the crowd (two items; e.g., "I felt a sense of unity with others evacuating the station"). Manipulation check measures were also taken of self-reported psychological engagement with the scenario depicted in the USES (three items; e.g., "I felt emotionally engaged during the experiment"). For each item, 11-point Likert scales were used (where 1 = *disagree strongly* and 11 = *agree strongly*).

Procedure. The participants were recruited to take part in "a Virtual Reality study of evacuation behavior." On being seated at a PC, they were given headphones to create a sense of immersion in the content of the program.

Unlike in Study 1, the only concern in the present study was that of the behavioral and subjective effects of identification (rather than its possible antecedents). Therefore, we attempted to manipulate high and low identification directly, through inducing the participants to imagine themselves in a psychological or an aggregate crowd, respectively.

The participants allocated to the high-identification condition wore badges and were asked to name three things about themselves as Sussex University students, whereas the low-identification condition participants were each asked to name three things about themselves as individuals. All the participants were then exposed to a brief news video of the Kings Cross fire. In the high-identification condition, the participants read and heard the following via Power-Point presentation immediately before the USES program began:

1. You and other Sussex University students have just been on a march in London about tuition fees. With a large crowd of other Sussex students, you are now going home. You all make your way into an Underground station. Together, you head towards the platform from which you will catch your train . . . .
2. You are just waiting for your train when there is some commotion along the platform. You suddenly hear someone shout "Fire! Get out, get out!" You look behind you and see smoke billowing towards you at one end of the platform. The crowd of people around you looks scared. The air seems to be getting thicker and hotter and you start to choke on the smoke. You realise that you may only have a few minutes to get back up to ground level and survive.
3. But there are other people trying to get out too . . . The station is still packed with other Sussex students from the march . . . .

In the low-identification condition, the wording was identical, except for the opening and third sections:

1. You have just been on a march in London about tuition fees. You have left the rest of the crowd, and are now going home on your own. You make your way into an Underground station. The station is busy with other people—some going home from work, others are tourists, some have been shopping. You head towards the platform from which you will catch your train . . . .
3. But there are other people trying to get out too. The station is packed . . . .

Following the visualization, the participants filled in the self-report questionnaire. The visualization took between 14 and 17 min for each participant.

Results

Manipulation checks. Three *engagement* items (α = .53) were conflated into a composite, the mean score of which (7.34) was significantly above the scale midpoint, indicating that the participants did indeed feel engaged with the visualization [$t(71)$ = 6.92, p < .001]. However, judging from the mean scores of the composite measure of group identification (α = .89), there was no difference in the level of collective identification across the two conditions [low-identification condition, M = 6.12, SD = 2.58; high-identification condition, M = 6.17, SD = 2.59; $t(69)$ = -0.24, p = .81].

Main effects. No gender differences were found for any of the dependent variables, so the analysis excluded this variable. In order to test the effects of identification, a median split was carried out on the self-report measure. Those at the median (6.5, n = 6) were eliminated in order to enhance the difference between those high (n = 34) and low (n = 32) in group identification.

Danger-of-death values at Helping Opportunities 1 and 2 were relatively low (M = 25% and 51%, respectively), whereas at Helping Opportunities 3 and 4, the danger-of-death values were high (M = 88% and 77%, respectively). These were therefore simplified to make just two helping opportunities: low and high danger. The mean amount of help offered at each of these two points was crossed with high and low identification in a 2 × 2 mixed ANOVA to test for a possible moderation effect. There was a main effect for danger of death [$F(1,57)$ = 30.76, p < .001, $\acute{\eta}$ = .39]. Overall, people helped more when danger of death was low (M = 1.74, SD = 0.53) than when it was high (M = 1.08, SD = 0.73). High identifiers (M = 1.59, SD = 0.18) were also found to help more than low identifiers (M = 1.32, SD = 0.21) [$F(1,57)$ = 8.33, p = .005, $\acute{\eta}$ = .17]. However, there was no interaction [$F(1,57)$ = 0.38, p = .54, $\acute{\eta}$ = .02]. Hence, although the high-identification participants offered more help than did the low-identification participants on each occasion, each set of participants were affected in the same way by the level of danger of death; there was no moderation effect.

The behavioral results for identification and helping were echoed in the analysis of conflated self-report measures of concern for others (α = .66): Those high in collective identification were more concerned for others

($M = 8.60$, $SD = 1.13$) than were those low in collective identification ($M = 7.10$, $SD = 1.36$) [$F(1,64) = 23.11$, $p < .001$, $\acute{\eta} = .27$].

Before comparing high- and low-identification participants on pushing, a box-plot was used to identify and exclude outliers (those pushing over 27 times; $n = 8$). (Such high scores may have reflected accidental keypushes, holding the key down continuously, or possible glitches in this early version of the program, since the numbers seemed higher than should be possible in each short trial.[5]) As was expected, those high in collective identification ($M = 3.22$, $SD = 3.58$) pushed significantly fewer times than those low in identification ($M = 8.22$, $SD = 8.03$) [$F(1,59) = 11.46$, $p = .001$, $\acute{\eta} = .17$].

Mediation analysis. To test the proposed indirect relationship between group identification, helping, and concern for others, bootstrapping procedures were employed, using the macro developed by Preacher and Hayes (2004). This method is argued to be more suitable than the Sobel test for smaller samples (Preacher & Hayes, 2004). It also makes fewer assumptions than does the more traditional Baron and Kenny (1986) mediation approach and, so, may be more appropriate for data sets that do not necessarily meet all the parametric criteria (MacKinnon, Fairchild, & Fritz, 2007). First, 5,000 bootstrap samples were generated from the original data set. If the 95% confidence interval for the estimates of mediation effect does not include zero, it suggests that the mediation effect is significant at the .05 level. The bootstrapping results indicated that the mean mediation effects from collective identification through concern for others to amount of help given ($b = .09$ [CI: 0.03, 0.18]) was significant at $p \leq .05$. As the regression equations showed, concern for others fully mediated the effect of collective identification on helping behavior. Thus, collective identification predicted the amount of help given [$\beta = .35$, $t(69) = 3.12$, $p = .003$]. Concern for others also predicted the amount of help given [$\beta = .64$, $t(68) = 6.79$, $p < .001$]. Collective identification predicted concern for others [$\beta = .46$, $t(67) = 4.20$, $p \leq .001$]. Finally, when collective identification and concern for others were put into the equation together, concern for others remained significant, but collective identification was no longer a predictor of helping [$\beta = .07$, $t(66) = 0.67$, $p = .51$].

Using the same procedure, the bootstrapping results indicated that the mean mediation effects from collective identification through concern for others to amount of pushing ($b = -.25$ [CI: -0.69, 0.08]) was not significant at $p \leq .05$. As would be expected on the basis of the social identity approach, high collective identification negatively predicted pushing [$\beta = -0.25$, $t(69) = -2.12$, $p = .04$], and concern for others also predicted pushing, again in a negative direction [$\beta = -0.34$, $t(60) = -2.84$, $p = .006$]. Yet when the two predictors were put into the regression equation together, collective identification remained significant, whereas concern for others was no longer a predictor [$\beta = -0.19$, $t(64) = -1.36$, $p = .18$].

Discussion

The two sorts of manipulation checks taken in this study yielded quite different results. Whereas the participants were not engaged as intended in the identity manipulation contained in the vignette, they indicated that they did feel engaged with the VR.

However, there was sufficient variation in levels of identification among the participants to carry out a comparison of low and high identifiers on the dependent variables, which were found to operate in the directions expected, giving us further reason to regard the USES as a feasible tool for this behavioral domain.

Thus, in line with predictions, those who were found to be high in identification with the rest of the crowd gave more help, expressed more concern, and pushed less than did those low in identification. The relation of mediation between identification, concern, and helping was also exactly as would be predicted by the social identity approach, with the effect of identification on helping being wholly mediated by concern for others.

Concern did not mediate the effect of identification on competition, however, although both identification with the crowd and concern for others separately reduced the amount of pushing displayed. Although identification and concern for others are clearly related, perhaps the pathway from identification to (reduced) pushing is not necessarily the same as that from identification to helping. One might offer help out of concern for others, but one desists from pushing not out of concern, perhaps, but for other identification-based reasons—such as politeness, expectations that ingroup others will themselves be considerate, and lack of impatience. Since these possible factors were not measured in this study, however, these suggestions remain speculations.

Taken together, these results clearly both further confirm the potential of VR methodology for investigating psychological aspects of mass evacuation (even in a setting that is not fully immersive) and give some support to a social identity explanation of such behavior. The results contradict most of the predictions of the mass panic model and add to the dominant affiliation and normative approaches. Increases in the level of danger did reduce helping, but these responses were not disproportionate (exaggerated, irrational), as mass panic approaches would suggest. More important, as we have seen, irrespective of level of danger, absolute levels of helping were greater for high- than for low-identification participants. Neither the mass panic model nor the affiliation approach would predict such high levels of helping and concern and reduced levels of pushing in a crowd of strangers, as occurred for the high-identification participants. At best, the normative approach might suggest that everyday norms of courtesy would continue even in an emergency; but this approach fails to explain why this norm conformity would occur more for high- than for low-identification participants.

As useful as this study has been as a test run of the visualization program and an initial test of the social identity account of mass evacuation behavior, however, there were a number of procedural issues that need to be addressed. In particular, the theoretical claims made here are somewhat weakened by the fact that part of the analysis relied on a post hoc median split on levels of self-reported collective identification.

We can speculate on why the identification measure did not differ between the experimental conditions. One possibility is that the items were created simply for this study, rather than established measures of identification. Hence, they may not have mapped cleanly onto the identification construct. However, the consistency between the general predictions and both the median split comparison analysis and the mediation analysis suggests that the measures were not at fault.

More plausible explanations for the apparent failure of the identification manipulation have to do with features of the manipulation itself. The use of badges and the *name-three-things* technique are tried and tested identity manipulations (Haslam, 2004). However, the manipulation procedure as a whole involved the subsequent provision of additional information that may have served to blur the subjective distinction between conditions.

On the one hand, the shared social category used in the vignette—*demonstrators*—may have little resonance for contemporary U.K. students for whom the campaign over tuition fees is far less prominent now than it was a number of years ago. It may not have been *relevant* enough to student participants to evoke identification with a crowd of other *demonstrators*. On the other hand, and perhaps a more plausible explanation for the failure of the identification manipulation, was that the *personal* and *collective* experimental conditions were not sufficiently distinct. The *demonstration* (collective identification) backstory was held constant, and *personal identification* was manipulated by reference to the participant's being among an aggregate crowd only after being in the *demonstration*. It seems possible, therefore, that this additional *personalizing* information was not prominent enough for those in the low-identification condition.

Through running a third and final version of the study, our intent was to manipulate collective identification by addressing these possible problems of social category relevance and insufficient difference between the vignettes. Successful experimental manipulation would serve to demonstrate more convincingly and cleanly the claimed causal relations between the variables of interest than in the present study.

The third study was also intended to examine the possible mediating role of a further SCT-derived variable different from that employed in Study 2: liking or attraction to others, which should be a function of collective identity. Furthermore, the test of the relationship between perceived danger and behavior was developed by examining whether not only helping, but also pushing is affected by the level of danger. Finally, by varying the size of the crowd in the visualization, this study allowed for a test of SCT not only against panic and affiliation models, but also against social impact theory (Latané & Wolf, 1981), which has previously been applied to predicting (lack of) help in "emergency" situations.

STUDY 3

Social impact theory (e.g., Latané & Wolf, 1981) suggests that group size is a factor determining responses in situations in which helping behavior might be involved.

Specifically, total *social impact* is said to be a function of strength, immediacy, and total number of influences. Hence, according to this theory, the larger the number of individuals over which possible influence (e.g., from a fallen victim) is diffused, the smaller the influence. Thus, the bigger the crowd, the less likely a victim is to receive help from others in the crowd (e.g., Latané & Nida, 1981).

In the present study, therefore, crowd size, as well as collective identity, was manipulated. SCT would suggest that crowd size would make no difference to the amount of helping offered when salience of collective identity was held constant, whereas from the point of view of social impact theory, salience makes no difference to the effect of crowd size on reducing helping behavior. Each theory would lead us to expect main effects and no interaction. However, a third possibility is one of moderation: Crowd size might moderate the effect of collective identity, so that the effect of collective identity on helping will be significantly greater in large than in small crowds.

As before, helping was expected to be reduced and pushing greater when participants did not identify with the group. Although the predictions of social impact theory are clear for helping, they are less clear for pushing, whether with small or large crowds, since the person pushed is not a potential source of influence in the same way as a needy victim. The predictive and possible mediating role of liking was also explored. Finally, we also sought a further test of the predictions of the mass panic and affiliation models by examining the extent to which danger of death predicted not only (lack of) help to victims, but also spontaneous pushing. A version of the mass panic model modified by the tenets of SCT suggests that increased danger would encourage competitive behavior such as pushing, but only for those in the low-identification condition.

Method

Design. This study used a 2 × 2 independent measures design, with low versus high identification as the first level and number of others visibly present in the simulation (32 for high vs. 8 for low) as the second level.

Participants. Forty students from the University of Sussex were recruited to take part in the study. Their age ranged from 20 to 25 years, with the mean age being 22. Seven were male, and the rest were female. All took part without inducements.

Dependent measures. Behavioral measures were the same as those in Studies 1 and 2. Additionally, a measure was taken of the danger-of-death percentage at the time of each spontaneous push by the participant. Two self-report items were used to measure group liking: "How positive did you feel towards the other people in the computer simulation?" and "How much did you like the other people in the computer simulation?" There were manipulation checks for psychological engagement with the VR (three items, based on those in Study 2) and for perceptions of the size of the crowd. Scales were 1–11, where 1 = *agree strongly* and 11 = *disagree strongly*.[6] Since the scale for collective identification was found to be reliable in Study 2, just one item from this scale was used: "How much did you feel part of a group during the simulation?"

Procedure. The participants began the study by reading a news report of the Kings Cross fire. They were then asked to close their eyes and imagine the sights, noises, smells, and other sensations that would be associated with such a scenario. The high-identification participants then were given the following scenario:

1. You have just been to an England football match at Wembley Stadium and are now on your way back to Brighton as you have university in the morning. You and the other England supporters are making your way through the local rail station to the Underground, from where you can get the train back home.

2. You are just about to board the underground train when you hear someone shout "There's a fire, get out, get out!" You look behind you and see large flames at one end of the platform with people running away from the fire. Everybody around you looks scared, and you feel yourself starting to sweat and sense your heart pumping faster. The fire seems to be getting bigger rapidly and you start to choke on the smoke. You realise that you may only have a few minutes to get back up to ground level and away from the fire in order to survive.

The low-identification condition participants were presented with the following:

1. You have spent a long day shopping in central London and are now on the way back to Brighton as you have university in the morning. You are making your way through the local rail station to the Underground, from where you can get the train back home.

The rest of the scenario was otherwise identical to that in the first condition.

To enhance collective identity, in the high-identification condition, the characters in the visualization wore vests that were the same color, whereas in the low-identification condition the characters wore various vests that were different.

Half of the high-identification and half of the low-identification participants were exposed to a condition in which the crowd was small (8 other characters), as opposed to one in which the crowd was large (32 other characters).

The rest of the procedure was the same as that in previous trials. Following the visualization, the participants filled in the self-report questionnaire.

Results

Manipulation checks. The three *engagement* items scaled well ($\alpha = .79$) and were therefore combined in a single item. The mean for these was not different from the midpoint [$M = 6.11$, $SD = 2.16$; $t(39) = 0.32$, $p = .75$], suggesting moderate psychological engagement with the VR. The item measuring the participants' perceptions of the difference in size between the crowds was in the intended direction but did not differ significantly between high- ($M = 4.95$, $SD = 3.23$) and low-density ($M = 5.35$, $SD = 3.23$) conditions [$t(38) = 0.36$, $p = .72$]. Identification differed between the high- ($M = 5.85$, $SD = 2.54$) and low-identification ($M = 8.45$, $SD = 1.64$) conditions in the direction intended [$t(38) = -3.85$, $p = .001$].

Main effects and interactions. Danger-of-death values at Helping Opportunities 1 and 2 were again relatively low (i.e., below the midpoint; 10% and 46%, respectively), as compared with Helping Opportunities 3 and 4 (85% and 56%, respectively). Hence, to look at any possible moderation effects, the data were simplified by conflating Opportunities 1 and 2 and then 3 and 4, respectively, to make just two helping opportunities: low and high danger. These two data points were then crossed with the high- and low-identification conditions in a 2 × 2 mixed ANOVA.

There was a marginal effect for danger of death [$F(1,38) = 3.23$, $p = .08$, $\acute{\eta} = .08$]. Means revealed that slightly more help was offered when danger was low ($M = 1.30$, $SD = $

0.69) than when it was high ($M = 1.05$, $SD = 0.78$). There was a main effect of group identification, with significantly more help offered in high-identification ($M = 0.70$, $SD = 0.29$) than in low-identification ($M = 0.48$, $SD = 0.27$) conditions [$F(1,40) = 6.42$, $p = .02$, $\acute{\eta} = .15$]. However, there was again no interaction between danger of death and identification [$F(1,38) = 0.13$, $p = .72$, $\acute{\eta} < .01$].

For this study, we were able to gather average danger-of-death values for instances of spontaneous pushing, allowing us to analyze the relationship between danger of death, identification, and pushing. There was no significant difference in the mean danger-of-death level for pushing between the high- ($M = 10.21$, $SD = 11.71$) and the low-identification ($M = 8.61$, $SD = 20.51$) conditions [$t(35) = 0.56$, $p = .58$], although the means were in the expected direction.

Looking next at the possible effects of crowd size, there was no difference in the number of victims helped in the small and large crowd conditions [$F(1,40) = 0.00$, $p = 1.00$, $\acute{\eta} = .00$; M for each $= .59$]. Furthermore, there was no interaction of crowd size and identification [$F(1,40) = 1.27$, $p = .27$, $\acute{\eta} = .03$].

The self-report measures of liking ($\alpha = .57$) were conflated into a single measure and were found to echo the behavioral results for helping: Those in the high-identification condition ($M = 6.33$, $SD = 1.21$) liked others in the visualization significantly more than did those in the low-identification condition ($M = 7.95$, $SD = 1.51$) [$t(38) = 3.76$, $p = .001$].

For the analysis of pushing, a box-plot was again used to identify statistical outliers (those over 40 pushes: $n = 3$). Low-identification participants pushed more ($M = 18.39$, $SD = 12.20$) than high-identification participants ($M = 9.26$, $SD = 8.54$), as was predicted [$F(1,37) = 8.27$, $p = .007$, $\acute{\eta} = .20$]. There was more pushing in the large crowd ($M = 17.50$, $SD = 11.58$) than in the smaller one ($M = 10.05$, $SD = 10.08$) [$F(1,37) = 5.37$, $p = .03$, $\acute{\eta} = .14$]. There was a very marginal interaction between crowd size and identification [$F(1,37) = 3.16$, $p = .09$, $\acute{\eta} = .09$]; the greatest amount of pushing took place with the larger crowd and when identification was low ($M = 24.89$), and the least amount of pushing took place when identification was high and the crowd small ($M = 8.40$).

Mediation analysis. To test the possible relationships between collective identification, helping, and liking for others, the bootstrapping procedure was again employed. Five thousand bootstrap samples were generated from the original data set. The analysis indicated that liking for others did not mediate the effect of collective identification on helping ($b < -.001$ [CI: -0.03, 0.25]). The regression equations indicated that identification did indeed predict the amount of help given [$\beta = -0.38$, $t(39) = -2.55$, $p = .02$]. However, liking of others did not predict helping [$\beta = -0.21$, $t(39) = -1.33$, $p = .19$]. Finally, however, in line with SCT, identification predicted liking of others [$\beta = .57$, $t(39) = 4.26$, $p < .001$].

With outliers removed, the same tests were carried out with pushing as the dependent variable. The bootstrap analysis indicated that liking for others did not mediate the effect of identification on pushing ($b = .32$

[CI: −1.04, 1.53]). In the separate regression equations, it was found that there was no relationship between collective identification and pushing [$\beta = .02$, $t(39) = 0.12$, $p = .91$], nor did liking of others predict pushing [$\beta = .07$, $t(39) = 0.41$, $p = .68$].

Discussion

In line with Study 2, it was found that those high in identification gave more help and pushed others less than did those low in identification. Yet, this time, the finding was based on an experimental manipulation of identification, rather than on a post hoc median split of self-reported identification. Again, identification predicted helping. But this time, using a different mediating variable (liking instead of concern), no mediation relation was found between identification and helping. Although identification and liking were correlated, as SCT would suggest (i.e., we like those who we define as in our *ingroup*), there was no relationship between liking and helping. In fact, this too is consistent with SCT, which suggests that we feel desire to help the crowd to the extent that we identify with it, irrespective of whether we like individual members. Indeed, this finding also, at least indirectly, counts against the affiliation approach, where the emphasis is on interpersonal bonds in explaining helping in an emergency.

Crowd size did not affect amount of help given; hence, these findings give no support to social impact theory. (The self-report measures of density might suggest that our manipulation of this variable was not consciously perceived; however, the fact that, as will be discussed below, the participants pushed more in larger groups suggests that there was indeed a difference in the awareness of density across conditions, as was intended.) Neither was there any evidence for a modified version of social impact theory, whereby SCT effects operate only in small crowds. These results are therefore in line with those of Levine and colleagues (e.g., Levine et al., 2002), who, in a variety of (noncrowd) scenarios, have found that collective identification, rather than diffusion of responsibility, explains the extent of helping behavior.

The result for greater pushing in the large crowds, although significant, may not be as psychologically interesting—or indeed as psychological per se—as it might first appear. People may barge and push more in a larger crowd simply because of physical constraints, rather than psychological tendencies toward personal selfishness (Chertkoff & Kushigian, 1999; Cornwell, Harmon, Mason, Merz, & Lampe, 2001).

As with Study 2, there was a main effect of danger of death on helping behavior: the greater the danger, the less helping—even though, once again, those in the high-identification condition still gave more help in absolute terms. This finding is therefore consistent with both SCT and Johnson et al.'s (1994) modified normative theory of emergency evacuation. Moreover, in line with a version of the mass panic model modified by the tenets of SCT, there was a marginal interaction suggesting that increased danger encourages competition only for those low in identification.

GENERAL DISCUSSION

The studies described here are the first to employ VR technology to investigate experimentally one of the most important, yet difficult to study, topics in psychology: mass evacuation behavior. In avoiding both the stress relied upon in many of the classic experimental studies and the intrusion and trauma that might be involved in interviewing survivors from actual emergencies, all three studies had no difficulty in getting ethical clearance from the relevant bodies. They were able to induce a degree of urgency in participants without deception or the use of threats or rewards for escape. Instead, techniques based on role-play, combined with the latest computational graphics technology, were employed to engage participants in the emergency scenario, in a paradigm based on the familiar computer game format. Within this framework, all other relevant variables could be controlled, manipulated, and measured. Thus, we have been able to make a contribution to an area of social psychology that, with a few notable exceptions (Kugihara, 2001), has been largely neglected by experimental social psychology and, hence, by mainstream social psychological theory.

The method we have developed here has a further obvious advantage over most of the techniques used in mainstream experimental social psychology, as well as in sociological survey research. Rather than relying on tick-boxes indicating what participants "intend" to do, the key measures in the present study were behavioral: helping, hesitation, and pushing. The self-report measures we took were additional, not central to the analysis.

The USES has not been produced for commercial availability. It was developed by the researchers for their own research purposes.[7] However, we have done more here than demonstrate the general utility and viability of VR technology for the domain of mass evacuation. We have also specified the computational techniques required. The low-polygon modeling and digital-imaging techniques described here can be employed to re-create any other kind of scenario that researchers wish to use as the backdrop to an emergency evacuation, whether it be an office block, theater, or street. The low computational demands mean that the resultant program will be able to run on most standard PCs and laptops. This is a practical consideration that experimentalists will find attractive, since it makes the paradigm amenable to student projects and other research outside of large-budget research programs.

The one clear weakness in the use of the new technique lay not in the technology itself, but in our attempt consistently to embed it in a traditional psychology laboratory setting. The framing of the VR with an identity manipulation was clearly weak in Study 2. It has been noted previously that identity variables are harder than other group-relevant perceptions (e.g., collective injustice and efficacy) to manipulate (Van Zomeren, Postmes, & Spears, 2008). Indeed, although there are a variety of techniques for manipulating salience of real or laboratory identities, some are more robust than others (Haslam, 2004). Clearly, more consideration needs to be given to choosing the type

of social categories that we think (student) participants might relate to best (e.g., protester vs. football fan) and the best techniques with which to make such social categories salient. Alternatively, we need to consider employing as participants people for whom a given social identity is most clearly relevant. For example, did the Scottish students in Study 1 find the London Underground setting meaningful and engaging (or remote)?

Having said this, taken together the effects and correlations of identification across the three studies allow us to draw a number of provisional conclusions in terms of psychological theory, as well as methodology. First, and most broadly, the results from these three studies suggest the usefulness of collective-level psychological theory in understanding aspects of behavior in mass emergencies. Whereas in classic group psychology, *panic* in an emergency is understood as the dissolution of collectivity (Freud, 1921/1985), our analysis is consistent with the view that collective bonds may be strengthened and even created through the experience of an emergency. Thus, the first study suggested that social identity could operate as a function of the (shared) threat when people experience an emergency.

Studies 2 and 3 showed that the amount of help offered by the participants was affected proportionately by levels of danger, but also that the participants responded differently to that level of danger depending on whether their personal or collective identities were salient. Across all three studies, collective identity was the key predictor accounting for variations not only in cooperation offered, but also the (relative lack of) pushing. Study 2 also showed that the effect of collective identity in helping was mediated by concern for others.

Thus, we would argue that the social identity approach to mass emergency evacuation we have tested and found support for in the studies reported here takes us beyond previous advances in the field in certain important ways. Although the normative approach (e.g., Johnson, 1987a, 1987b) clearly transcends the mass panic model by pointing to the predominance of sociality even within life-threatening mass emergency crowds, it collapses back into the mass panic model when it claims that competitive behavior increases at times of greatest urgency (Johnson et al., 1994; cf. Quarantelli, 1957). The normative approach follows the affiliation model in emphasizing the role of preexisting interpersonal relationships in limiting panic (e.g., Cornwell, 2003; Feinberg & Johnson, 2001; Johnson et al., 1994). But the affiliation model too collapses back into the assumptions of the mass panic approach when it suggests that panic can occur when there are no familiar figures or features present (Mawson, 2005). What is missing from all these models is an explanation for the evidence of positive affective, motivational, and behavioral relations to a crowd of strangers in an emergency evacuation. By contrast, SCT suggests that the shift from personal collective identity leads us to define others' interests in line with our own, instead of understanding them as counterposed. In such contexts, one might even help strangers at a risk to the personal self.

There are valid questions concerning how well the processes analyzed here translate into genuine emergencies, where the threat of collective death is real and not imagined. Questions like this about the validity of simulation studies of course extend beyond the specificities of this particular domain. In our defense, we can perhaps take some encouragement from the high levels of participant engagement and face validity demonstrated in another recent simulation in which, as in the present study, group membership in a high-intensity environment was manipulated. Haslam and Reicher's (2006) BBC *prison* simulation study took psychometric, physiological, behavioral, and observational data to show the dynamic relationship between identification and stress. With so many objective measures, as with the case of the behavioral measures taken here, it is hard to argue that simulation experiments in general are necessarily the poor relations of studies of "real" group contexts. Thus, although it is ultimately necessary to supplement the findings described here with field, archival, and interview data, rather than dwelling on the weaknesses of the methodology, we would seek finally to contextualize our VR paradigm as a positive contribution to group processes research.

Taken in this context, this study is a success, in that it reconnects the discipline with some of its more creative and productive traditions. As we have seen, laboratory research simulating mass emergency situations flourished in the 1950s and 1960s. This was the height of the popularity of the *group dynamics* approach to social psychology, when group processes were studied behaviorally, in the lab, rather than (as now) largely through questionnaires. This period produced some of the most engaging, provocative, imaginative, and readable studies found in social psychology textbooks today. Our aim has been to develop an engaging experimental study that also meets modern standards in research ethics. To the extent that we have achieved this, contemporary theories of group processes, such as SCT, can inform the field of mass emergency behavior, from which mainstream social psychology has been disconnected for too long.

REFERENCES

AGUIRRE, B. E. (2005). Commentary on "Understanding mass panic and other collective responses to threat and disaster": Emergency evacuations, panic, and social psychology. *Psychiatry*, **68**, 121-129.

AGUIRRE, B. E., WENGER, D., & VIGO, G. (1998). A test of the emergent norm theory of collective behavior. *Sociological Forum*, **13**, 301-320.

BARON, R. M., & KENNY, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality & Social Psychology*, **51**, 1173-1182.

BLAKE, S. J., GALEA, E. R., WESTENG, H., & DIXON, A. J. P. (2004). An analysis of human behavior during the WTC disaster of 11 September 2001 based on published survivor accounts. In *Proceedings of Third International Symposium on Human Behavior in Fire* (pp. 181-192). London: InterScience Communications.

BURTON, A., & SCHOFIELD, D. (2005). *Underground station evacuation simulator* (Version 3.1). Nottingham: University of Nottingham.

CAMPBELL, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, **3**, 14-25.

CASTANO, E. (2004). In case of death, cling to the ingroup. *European Journal of Social Psychology*, **34**, 375-384.

CHERTKOFF, J. M., & KUSHIGIAN, R. H. (1999). *Don't panic: The psychology of emergency egress and ingress.* Westport, CT: Praeger.

CHERTKOFF, J. M., KUSHIGIAN, R. H., & McCOOL, M. A. (1996). Interdependent exiting: The effects of group size, time limit, and gender of the coordination of exiting. *Journal of Environmental Psychology*, **16**, 109-121.

CLARKE, L. (2002). Panic: Myth or reality? *Contexts*, **1**, 21-26.

CORNWELL, B. (2003). Bonded fatalities: Relational and ecological dimensions of a fire evacuation. *Sociological Quarterly*, **44**, 617-638.

CORNWELL, B., HARMON, W., MASON, M., MERZ, B., & LAMPE, M. (2001). Panic or situational constraints? The case of the M/V Estonia. *International Journal of Mass Emergencies & Disasters*, **19**, 5-25.

DOOSJE, B., ELLEMERS, N., & SPEARS, R. (1995). Perceived intragroup variability as a function of group status and identification. *Journal of Experimental Social Psychology*, **31**, 410-436.

DRURY, J., COCKING, C., & REICHER, S. (in press). Everyone for themselves? A comparative study of crowd solidarity among emergency survivors. *British Journal of Social Psychology.* doi:10.1348/014466608X357893

DRURY, J., & REICHER, S. (2000). Collective action and psychological change: The emergence of new social identities. *British Journal of Social Psychology*, **39**, 579-604.

FEINBERG, W. E., & JOHNSON, N. R. (2001). Primary group size and fatality risk in a fire disaster. In *Human behavior in fire: Understanding human behavior for better fire safety design* (pp. 11-22). London: InterScience Communications.

FREUD, S. (1985). Group psychology and the analysis of the ego (J. Strachey, Trans.). In A. Dickson (Ed.), *Civilization, society and religion* (pp. 91-178). Harmondsworth, U.K.: Penguin. (Original work published 1921)

FRITZ, C. E., & WILLIAMS, H. B. (1957). The human being in disasters: A research perspective. *Annals of the American Academy of Political & Social Science*, **309**, 42-51.

GINSBURG, G. P. (1979). The effective use of role playing in social psychological research. In G. P. Ginsburg (Ed.), *Emerging strategies in social psychological research* (pp. 117-155). Chichester, U.K.: Wiley.

GREENBERG, J., PYSZCZYNSKI, T., SOLOMON, S., ROSENBLATT, A., VEEDER, M., KIRKLAND, S., & LYON, D. (1990). Evidence for terror management theory: II. The effects of mortality salience on reactions to those who threaten or bolster the cultural worldview. *Journal of Personality & Social Psychology*, **58**, 308-318.

GREENBERG, J., SOLOMON, S., & PYSZCZYNSKI, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 61-139). San Diego: Academic Press.

GROSS, D. E., KELLEY, H. H., KRUGLANSKI, A. W., & PATCH, M. E. (1972). Contingency of consequences and type of incentive in interdependence escape. *Journal of Experimental Social Psychology*, **8**, 360-377.

HASLAM, S. A. (2004). *Psychology in organizations: The social identity approach* (2nd ed.). London: Sage.

HASLAM, S. A., O'BRIEN, A., JETTEN, J., VORMEDAL, K., & PENNA, S. (2005). Taking the strain: Social identity, social support, and the experience of stress. *British Journal of Social Psychology*, **44**, 355-370.

HASLAM, S. A., & REICHER, S. D. (2006). Stressing the group: Social identity and the unfolding dynamics of responses to stress. *Journal of Applied Psychology*, **91**, 1037-1052.

HELBING, D., FARKAS, I., & VICSEK, T. (2000). Simulating dynamical features of escape panic. *Nature*, **407**, 487-490.

JOHNSON, N. R. (1987a). Panic and the breakdown of social order: Popular myth, social theory, empirical evidence. *Sociological Focus*, **20**, 171-183.

JOHNSON, N. R. (1987b). Panic at "The Who Concert Stampede": An empirical assessment. *Social Problems*, **34**, 362-373.

JOHNSON, N. R. (1988). Fire in a crowded theatre: A descriptive investigation of the emergence of panic. *International Journal of Mass Emergencies & Disasters*, **6**, 7-26.

JOHNSON, N. R., FEINBERG, W. E., & JOHNSTON, D. M. (1994). Microstructure and panic: The impact of social bonds on individual action in collective flight from the Beverly Hills Supper Club Fire. In R. R. Dynes & K. J. Tierney (Eds.), *Disasters, collective behavior, and social organization* (pp. 168-189). Newark: University of Delaware Press.

JONAS, E., SCHIMEL, J., GREENBERG, J., & PYSZCZYNSKI, T. (2002). The Scrooge effect: Evidence that mortality salience increases prosocial attitudes and behavior. *Personality & Social Psychology Bulletin*, **28**, 1342-1353.

KELLEY, H. H., CONDREY, J. C., DAHLKE, A. E., & HILL, A. H. (1965). Collective behavior in a simulated panic situation. *Journal of Experimental Social Psychology*, **1**, 20-54.

KLEIN, A. L. (1976). Changing in leadership appraisal as a function of the stress of a simulated panic situation. *Journal of Personality & Social Psychology*, **34**, 1143-1154.

KUGIHARA, N. (2001). Effects of aggressive behavior and group size on collective escape in an emergency: A test between a social identity model and de-individuation theory. *British Journal of Social Psychology*, **40**, 575-598.

LANGSTON, P. A., MASLING, R., & ASMAR, B. N. (2006). Crowd dynamics discrete element multi-circle model. *Safety Science*, **44**, 395-417.

LATANÉ, B., & NIDA, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin*, **89**, 308-324.

LATANÉ, B., & WOLF, S. (1981). The social impact of majorities and minorities. *Psychological Review*, **88**, 438-453.

LEVINE, M., CASSIDY, C., BRAZIER, G., & REICHER, S. (2002). Self-categorization and bystander intervention: Two experimental studies. *Journal of Applied Social Psychology*, **32**, 1452-1463.

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 & Social Psychology Bulletin*, **31**, 443-453.

LEVINE, M., & THOMPSON, K. (2004). Identity, place, and bystander intervention: Social categories and helping after natural disasters. *Journal of Social Psychology*, **144**, 229-246.

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.

MacKINNON, D. P., FAIRCHILD, A. J., & FRITZ, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, **58**, 593-614.

MAWSON, A. R. (2005). Understanding mass panic and other collective responses to threat and disaster. *Psychiatry*, **68**, 95-113.

McDOUGALL, W. (1920). *The group mind.* New York: Putnam.

MILGRAM, S. (1963). Behavioral study of obedience. *Journal of Abnormal & Social Psychology*, **67**, 371-378.

MINTZ, A. (1951). Non-adaptive group behavior. *Journal of Abnormal & Social Psychology*, **46**, 150-159.

MUIR, H. (1996). Research into the factors influencing survival in aircraft accidents. *Aeronautical Journal*, **100**, 177-181.

OSBORN, C. Y., JOHNSON, B. T., & FISHER, J. D. (2006). After 9/11 at ground zero: The anxiety-buffering effects of worldview support of the first anniversary of 9/11. *Basic & Applied Social Psychology*, **28**, 303-310.

PREACHER, K. J., & HAYES, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, **36**, 717-731.

PROULX, G., & FAHY, R. F. (2003, October). Evacuation of the World Trade Center: What went right? In *Proceedings of the CIB-CTBUH International Conference on Tall Buildings* (pp. 27-34). Ottawa: National Research Council of Canada.

QUARANTELLI, E. (1957). The behavior of panic participants. *Sociology & Social Research*, **41**, 187-194.

REICHER, S. (2001). The psychology of crowd dynamics. In M. A. Hogg

& R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 182-208). Oxford: Blackwell.

Ripley, A. (2005, May 16). How to get out alive. *Time*, **165**, 48-52.

Ross, E. A. (1908). *Social psychology: An outline and source book*. New York: Macmillan.

Sime, J. D. (1983). Affiliative behavior during escape to building exits. *Journal of Environmental Psychology*, **3**, 21-41.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., et al. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PLoS ONE*, **1**, e39. doi:10.1371/journal.pone.0000039

Still, K. (2000). *Crowd dynamics*. Unpublished doctoral dissertation, Department of Mathematics, University of Warwick, U.K. Available from Crowd Dynamics Web site: www.crowddynamics.com/.

Strauss, A. L. (1944). The literature on panic. *Journal of Abnormal & Social Psychology*, **39**, 317-328.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of inter-group conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks/Cole.

Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 15-40). Cambridge: Cambridge University Press.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes: Theory and research* (Vol. 2, pp. 77-122). Greenwich, CT: JAI Press.

Turner, J. C. (1999). Some current issues in research on social identity and self-categorization theories. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 6-34). Oxford: Blackwell.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Blackwell.

Turner, J. C., & Oakes, P. J. (1997). The socially structured mind. In C. McGarty & S. A. Haslam (Eds.), *The message of social psychology* (pp. 355-373). Oxford: Blackwell.

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**, 504-535.

NOTES

1. *Shrek*, copyright Dreamworks (2001); *The Incredibles*, copyright Pixar (2004).

2. A polygon is defined mathematically as a closed plane figure bounded by three or more line segments. In computer graphics terminology, the word polygon has a more specific definition. Here, the term *polygon* is taken to mean a multisided object that can be filled with color or moved around as a single entity. (In most wireframe geometric meshes, computers use triangular polygons.)

3. Here, we follow the social identity model in distinguishing between physical, or aggregate, crowds (numbers of people who are simply physically colocated) and psychological crowds (people who define themselves as having a common identity) (Drury et al., in press; Reicher, 2001).

4. Because the visualization filled the participant's view like a real event, it was thought that the danger bar might draw attention away from the scene and, hence, detract from psychological engagement. It was therefore not employed in this version of the experiment.

5. Since most of the outliers were among the low identifiers, the difference in mean pushing scores between high and low identifiers (5.29 and 35.97, respectively) prior to their exclusion was much greater before than afterward; the exclusion of outliers therefore made for a tougher test of the hypothesis.

6. Note that self-report scales are in the reverse direction to those in Study 2.

7. Other researchers wishing to make use of the USES should contact the authors.