Social relations and presence of others affect bystander intervention

Liebst, Lasse Suonperä; Philpot, Richard; Bernasco, Wim; Dausel, Kasper Lykke; Ejbye-Ernst, Peter; Nicolaisen, Mathias Holst; Lindegaard, Marie Rosenkrantz

Published in:
Aggressive Behavior

Publication date:
2019

Citation for published version (APA):
Liebst, L. S., Philpot, R., Bernasco, W., Dausel, K. L., Ejbye-Ernst, P., Nicolaisen, M. H., & Lindegaard, M. R. (2019). Social relations and presence of others affect bystander intervention: Evidence from violent incidents captured on CCTV. *Aggressive Behavior*.

Social relations and presence of others affect bystander intervention: Evidence from violent incidents captured on CCTV

Lasse Suonperä **LIEBST** [1]

(ORCID 0000-0003-1062-2447)

Richard **PHILPOT** [1]

(ORCID 0000-0002-0359-2123)

Wim **BERNASCO** [2 3 *]

(ORCID 0000-0002-3385-0883)

Kasper Lykke **DAUSEL** [1]

(ORCID 0000-0001-5279-9000)

Peter **EJBYE-ERNST** [2]

(ORCID 0000-0003-4598-7235)

Mathias Holst **NICOLAISEN** [1]

(ORCID 0000-0002-0283-8144)

Marie Rosenkrantz **LINDEGAARD** [1 2]

(ORCID 0000-0002-1630-774X)

[1] Department of Sociology, University of Copenhagen, Øster Farimagsgade 5A, Bld. 16, Copenhagen K, Denmark

[2] Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), P.O. Box 71304, 1008 BH Amsterdam, The Netherlands

[3] Department of Spatial Economics, School of Business and Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

Acknowledgement: This work was supported by the Danish Council for Independent Research [DFF – 6109-00210] and the Velux Foundation. The funders had no role in the design of the study, data collection and analysis, decision to publish, or preparation of the manuscript. We additionally thank Camilla Bank Friis and Anne Laura Engmann Juul for their contributions to the coding of CCTV footage.

* Corresponding author, e-mail wbernasco@nscr.nl, phone +31-20- 598 5239

Social relations and presence of others affect bystander intervention: Evidence from violent incidents captured on CCTV

Abstract

Are individuals willing to intervene in public violence? Half a century of research on the 'bystander effect' suggests that the more bystanders present at an emergency, the less likely each of them is to provide help. However, recent meta-analytical evidence questions whether this effect generalizes to violent emergencies. Besides the number of bystanders present, an alternative line of research suggests that pre-existing social relations between bystanders and conflict participants are important for explaining whether bystanders provide help. The current paper offers a rare comparison of both factors—social relations and number of bystanders present—as predictors of bystander intervention in real-life violent emergencies. We systematically observed the behavior of 764 bystanders across 81 violent incidents recorded by surveillance cameras in Copenhagen, Denmark. Bystanders were sampled with a case-control design, their behavior was observed and coded, and the probability of intervention was estimated with multilevel regression analyses. The results confirm our hypothesized association between social relations and intervention. However, rather than the expected reversed bystander effect, we found a classical bystander effect, as bystanders were less likely to intervene with increasing bystander presence. We assess these findings in light of recent discussions around the influence of situations versus group-based agency in human helping. Further, we discuss the utility of video data for the assessment of real-life bystander behavior.

Keywords: bystander effect, intervention, social groups, video observation, violence

Introduction

In the presence of others, bystanders are less likely to intervene when they witness someone in need of help (Darley & Latané, 1968). This bystander effect hypothesis is one of the most well-established findings of psychology (Manning, Levine, & Collins, 2007), and is typically interpreted as the product of a diffusion of responsibility, by which the liability to help dilutes across the multiple bystanders present (Latané & Nida, 1981). Paradoxically, although the bystander research field was prompted by the violent 1964-murder of Kitty Genovese, and the inaction of the witnesses present (but see Manning et al., 2007), experimental research has rarely examined bystander behavior in the context of violent attacks (Cherry, 1995; Liebst, Heinskou, & Ejbye-Ernst, 2018). This omission is a result of the practical and ethical infeasibility of exposing participants to dangerous study conditions (Osswald, Greitemeyer, Fischer, & Frey, 2010).

In restricting the analysis of bystander behavior to low-danger laboratory settings, the field risks isolating itself away from the phenomenon it initially set out to explain (Mortensen & Cialdini, 2010; Tinbergen, 1963). Confirming this concern, in the exceptionally few experimental studies that have simulated attacks, it is found that bystanders are equally (Fischer, Greitemeyer, Pollozek, & Frey, 2006), or more (Harari, Harari, & White, 1985), likely to intervene in the presence of others than when alone. Further, a meta-analysis of the experimental literature concludes that the bystander effect attenuates, or even reverses, in high-danger study contexts (Fischer et al., 2011). Taken together, when uncoupling the experimental evidence into the trivial (e.g., a pencil spill, a door that needs to be answered) and the more dangerous emergencies, the classical bystander effect does not seem to generalize across both domains. Rather, in dangerous study contexts, the presence of additional bystanders may provide a welcome physical support that promotes intervention (Fischer & Greitemeyer, 2013). In line with this interpretation, observational evidence from real-life emergencies captured by surveillance cameras shows that bystander presence increases the likelihood of intervention (Levine, Taylor, & Best, 2011). The overall finding that individuals *do* intervene when it is really matters aligns with cross-cultural anthropological accounts suggesting that third-party intervention in everyday conflicts is most likely a human universal (Boehm, 2000; Brown, 1991; Eibl-Eibesfeldt, 1989; Fry, 2000).

Shifting away from a situational emphasis on how additional individuals promote non-intervention, or the potential reversal of such effect, an alternative line of research stresses the importance of group-based agency in bystander helping (Levine & Manning, 2013; Philpot, 2017; Swann & Jetten, 2017). Specifically, those bystanders who are affiliated with an individual involved in the emergency are significantly more likely to intervene than those socially distant. This association is found not only across experimental and observational studies with humans (Levine, Cassidy, Brazier, & Reicher, 2002; Lindegaard et al., 2017; Slater et al., 2013), but also across much non-human primate work (de Waal, 2015). These findings are consistent with an evolutionary theory of cooperation that expects helping behaviors to occur disproportionately between genetically related or reciprocating individuals (de Waal & Preston, 2017; Hamilton & Axelrod, 1981; Vázquez, Gómez, Ordoñana, Swann, & Whitehouse, 2017).

Besides de-escalatory helping, which exists as the main focus of bystander research (Fischer et al., 2011), group membership has also been associated with escalatory interventions by which third-parties fight on behalf of their fellow group members (Black, 1993; Levine, Lowe, Best, & Heim, 2012; Phillips & Cooney, 2005; Swann, Gómez, Huici, Morales, & Hixon, 2010). In these situations, bystanders effectively become partisans in the unfolding conflict. Social relations between bystanders and conflict participants thus seem to foster not only de-escalatory but also escalatory interventions.

Despite the co-existence of these partially competing accounts, there have been few attempts to examine the relative contributions of the number of bystanders and social relations in explaining bystander intervention. This may result from the methodological circumstance that “laboratory studies of bystander intervention usually use strangers as research confederates who help to stage the helping dilemma" (Banyard, 2015, p. 30). Fischer and colleagues (2011) included bystander-victim familiarity as a moderator in their meta-analysis and found that the magnitude of the bystander effect was not influenced by whether or not the bystander knew the victim. Similarly, a regression analysis of in-depth interviews reported a significant bystander effect in a model in which social relations were the main predictor of bystander intervention (Phillips & Cooney, 2005). By contrast, Lindegaard and colleagues' (2017) examination of real-life bystander intervention in the aftermath of commercial robberies reported a weak reversed bystander effect in a model where social relations between victims and bystanders, again, dominated the intervention outcome. While these studies assessed the net effects of these two factors, Levine and Crowther (2008), analyzed the interaction between group size and social group identification and found that the inter-relationship between the two factors could both increase or decrease the likelihood of bystander intervention.

These few studies comparing the two factors simultaneously indicate that social relations outperform the number of bystanders as a predictor of intervention, while the evidence regarding the positive, vis-à-vis the negative, direction of the bystander effect remains mixed. However, these studies tend to rely on ecologically limited experimental paradigms and retrospective accounts (Baumeister, Vohs, & Funder, 2007; Swann & Jetten, 2017). An exception is the study of Lindegaard and colleagues (2017), which relied on video-based naturalistic observations of bystanders in the aftermath of non-fatal commercial robberies. However, by analyzing the period *after* the offenders had already left the setting, their study provides limited information on whether bystanders intervene in ongoing violent, dangerous emergencies—i.e., the condition proposed to attenuate or reverse the bystander effect. Overall, there is a dearth of direct comparisons of number of bystanders and social relations as predictors of bystander intervention in violent emergencies. The present study, which utilizes video recordings of public violent assaults, is the first systematic observational study to address this gap.

Given the dangerousness of the violent situations assessed in the present study, we predict a reversed bystander effect, with a positive association between the number of bystanders and the likelihood of bystander intervention (Hypothesis 1). We further predict that bystanders affiliated to a conflict party are more likely to intervene than strangers (Hypothesis 2). As the evidence supporting the reversed bystander effect is less uniform than the evidence in favor of social relations, we predict that the effect of social relations on intervention will be larger in magnitude than the effect of the number of bystanders (Hypothesis 3). These hypotheses align with the majority of bystander research that considers intervention as unambiguously prosocial (i.e., helping behavior), and should therefore apply to de-escalatory interventions. Whether these propositions also fit escalatory interventions, where bystanders become conflict participants, is an open question that we also explore in the empirical analysis.

We control for other factors that may influence the intervention likelihood, including the bystander's gender (Eagly, 2009), whether the bystander is a member of the public or is serving an occupational role (e.g., bouncer) (Hobbs, 2003), whether the event takes place in a nighttime drinking setting or not (Levine et al., 2012; Reynald, 2011), and also for two measures that may affect the bystanders' intervention opportunities: the density of the situation (Macintyre & Homel, 1997) and the spatial proximity of the bystander to the conflict participants (Macintyre & Homel, 1997).

Data and methods

Data

The data consists of 81 surveillance camera recordings of police-reported public violent assaults in central Copenhagen between 2010 and 2012 (replication data and a Stata script are available as

Supporting Information at osf.io).[1] The clips were a subset of a wider sample (N = 164[2]), and were selected if they conformed to the following three criteria. Each clip captured an event of physical violence, with or without intervening bystanders. The clip had a quality (e.g., brightness, resolution) that rendered it possible to conduct a systematic behavioral coding. Each clip captured the duration of the situation, with none, or only negligible, breaks in the coverage (see Nassauer & Legewie, 2018).

Coding procedure

The coding began by identifying the conflicting parties, in most cases, the two individuals between whom the situation initially manifested itself as a conflict. This encounter was identified from displays of direct physical violence or from nonverbal cues of anger and aggression (e.g., emphasizing gestures, forward body inclination, see Dael, Mortillaro, & Scherer, 2012). All individuals entering the ongoing conflict were defined as intervening bystanders.

With the use of a detailed observation codebook, four trained student assistants coded the bystander intervention behaviors (Table A1 in the Appendix) and situational properties (Table A2 in the Appendix) of each clip. This codebook was compiled from existing variable definitions in the literature (e.g., 'de-escalatory' and 'escalatory' intervention types, see Levine et al., 2011) and specified through in-depth qualitative observations of a subsample of videos (see Eibl-Eibesfeldt, 1989; Jones et al., 2016).

In addition to the visual information obtained from the video recordings, each clip also was coupled with a police case file that provided descriptive accounts of the event. Pre-existing social

[1] osf.io is the website of the Open Science Framework, where we deposited our data and Stata script. To secure author anonymity but allow reviewers to access the material, the files are currently stored at an anonymous version at URL `tinyurl.com/moresothan`. The study was approved by the Danish Data Protection Agency (reference 2015-57-0125-0026).

[2] Note that part of this video material is analyzed for another study purpose in ANONIMIZED REFERENCE.

relationships were inferred from nonverbal social behavioral cues (see Murphy, 2016). These cues included interactional displays of collective behavior-in-concert, such as moving in synchrony, shared focus and attention, and bodily proximity (Afifi & Johnson, 2005; Ge, Collins, & Ruback, 2012; Goffman, 1971). In ambiguous cases, coders validated these video-based group assessments against the police case file descriptions.

Interrater reliability

To test the reliability of the variables included in the final analysis, we selected 20 (29%) of the video contexts and 35 (15%) of the intervening bystanders for double coding. All variables included in the analyses reached a Krippendorff's alpha value of $\alpha \geq .80$, recommended by Krippendorff (2004) as the cutoff point for reliable interrater agreement (for the Krippendorff values of all coded variables see Tables A1 and A2 in the Appendix). Disagreements between the coders were resolved through discussion prior to analysis.

Case-control sampling

Because the incidents involved many more non-intervening than intervening bystanders and because the behavioral coding is very time-consuming, we applied a case-control approach (Keogh & Cox, 2014). Here, we randomly selected a sample of non-intervening 'controls,' who were situated in the same time and place as the intervening 'cases,' but without displaying the intervention-outcome of interest (Grimes & Schulz, 2005). For sufficient statistical power, it is recommended to sample at least two, but no more than four, controls per case (Lewallen & Courtright, 1998). With 510 non-intervening bystanders and 215 intervening bystanders included in the study, our control-to-case ratio is 2.4:1 and thus within these recommended thresholds.

Estimation

To account for the hierarchical structure of our data, with bystanders nested into video contexts, data was estimated with 2-level regression models with a random intercept (Hox, Moerbeek, & van de Schoot, 2017). All estimations were calculated with Stata 14’s ‘gllamm’ module using the adaptive quadrature estimation technique (Rabe-Hesketh, Skrondal, & Pickles, 2005). The data showed an average of 9 individuals nested across the 81 contexts, offering a sufficient sample size to obtain unbiased fixed-effect point estimates for most multilevel model specifications (McNeish & Stapleton, 2016).

Sampling weights

To make the randomly selected controls representative of the actual number of non-intervening bystanders in each context, data was modelled using sampling weights (Lohr, 2010). All interveners were assigned a weight of 1, and controls were assigned a weight equal to the number of selected controls as a proportion of the total number non-interveners. In the relatively few contexts where the number of selected controls exceeded the number of non-interveners, the controls were assigned a weight of 1. Prior to analysis, the weights were scaled to suit multilevel modelling (Carle, 2009).

Robustness tests

In addition to confirmatory tests of the three hypotheses and an exploratory comparison between escalatory and de-escalatory intervention, we conduct sensitivity analyses to assess the robustness of our results against other reasonable data and model specifications (Steegen, Tuerlinckx, Gelman, & Vanpaemel, 2016). These analyses included estimating combinations of independent variables using two alternative sampling weight scalings (Carle, 2009), and including the number of bystanders as a quadratic term, given that earlier research suggests that the negative association between number of bystanders and intervention diminishes curvilinearly with increasing numbers (Latané, 1981).

Measures

Dependent variables

We defined *bystander intervention* as a binary variable, distinguishing bystanders who intervene into the conflict (with either escalatory or de-escalatory acts) from bystanders that do not intervene. *Decomposed bystander intervention* was measured as a multinomial variable, distinguishing four possible bystanders based on their actions: non-intervention, only de-escalatory acts, only escalatory acts, and a mix of de-escalatory and escalatory acts. De-escalatory acts included making open-handed gestures, non-forceful touching, blocking contact between parties, holding a person back, hauling and pushing the antagonists apart. Escalatory acts included pointing and threatening gestures, throwing a person, pushing, shoving, hitting, kicking, violence against a person on the ground, and weapon use (see Table A1 in the Appendix). Table 1 presents descriptive statistics of the dependent, independent and control variables measured at the individual level. At the context-level, at least one bystander intervened in 85.0 percent of the 81 videos. In total, there were 217 intervening bystanders, with an average of 2.7 interveners per situation.

—— INSERT TABLE 1 HERE ——

Independent variables

The *number of bystanders* was a count of the individuals present in the emergency. This context-level predictor was standardized by subtracting the mean and dividing by two standard deviations as to make it comparable to the effect sizes obtained from the binary predictors (see Gelman, 2008). The bystander's *social relation* was measured with a binary variable, distinguishing bystanders who have a social relationship to an individual involved in the conflict from bystanders who do not know any of the conflict parties.

Control variables

To control for omitted-variable bias and based on findings of prior studies, we included five control variables. The bystander's *gender* was coded as male or female. This variable was included because of evidence showing that men tend to act more 'heroic and chivalrous' in their helping behavior than women (Eagly, 2009; Taylor et al., 2000). *Nighttime drinking settings* were defined as situations occurring in proximity to a bar/nightclub or during the weekend nights. This control variable was included as evidence shows that bystander involvement is a pervasive aspect of these settings (Levine et al., 2012; Parks, Osgood, Felson, Wells, & Graham, 2013).

Further, given that most of our incidents occur in drinking settings, it is plausible that the intervention likelihood is shaped by whether the bystander is performing an occupational role, e.g., as a bar staff or bouncer (Hobbs, 2003; Sampson, Eck, & Dunham, 2010). The *occupational role* of bystanders was captured with a binary variable, distinguishing bystanders who were at work from those who were not. Because physical proximity between individuals may facilitate helping behavior (Fujisawa, Kutsukake, & Hasegawa, 2006), we included a measurement of *spatial proximity*, distinguishing situations where the bystander was within a 2-meters radius from where the conflict initiates from situations in which the bystander was outside of this radius.

Finally, as levels of crowding may be associated with anti-social outcomes at public venues (Macintyre & Homel, 1997), we included *people density* as a control, distinguishing high density and low density situations. Density was assessed by whether it was possible to walk in a straight line across the setting without bumping into others present (low density) or not (high density).

Results

Figure 1 graphically shows the odds ratio estimates and associated confidence intervals of two multilevel binomial logistic regression models comparing bystander intervention with non-intervention. Full details of both models are presented in Table A3 in the Appendix. Both the key variables and control variables are listed on the vertical axis, while the effect sizes (odds ratios) are on the horizontal axis. The estimated odds ratios of the models are printed as dots and diamonds, respectively. The 95% percent confidence intervals are presented as horizontal lines around the estimates. The vertical line indicates an odds ratio of 1, reflecting the absence of any statistical relation.

The first model (estimates indicated in black with dots) includes only the two key variables, i.e., social relations and number of bystanders present. Contrary to the hypothesized reversed bystander effect, but in line with the classical bystander effect, we find that the number of bystanders is negatively associated with the likelihood of intervention. The effect size of this standardized variable (OR = 0.28) is large, as evaluated with Rosenthal's (1996) odds ratio effects size categories. Confirming our expectation, having a social relationship tie to a conflict party is positively associated with intervention. Compared to a stranger, the odds of intervening are more than 20 times larger for a bystander with a social relation to a conflict party, than for an unrelated bystander. Even if assessed conservatively from the lower band of the confidence interval (95% CI = [9.98, 42.17]), the estimated odds ratio is very large.

—— INSERT FIGURE 1 HERE ——

In the second model (estimates shown in grey with diamonds) the five control variables are included to account for confounding influences on the key variables. Confounding is almost negligible, as the estimates of the two key variables are very similar to those in the first model (0.24 and 18.17, respectively). With respect to the control variables, only the bystander's gender is significantly related to intervention, with males' odds of intervention being 3.6 times larger than that of females.

—— INSERT FIGURE 2 HERE ——

To further explore whether the influence of bystander numbers and social relations generalize across de-escalatory and escalatory intervention types, we decomposed the intervening bystanders into three groups: those who displayed only de-escalatory interventions, those who displayed only escalatory interventions and those who displayed both de-escalatory *and* escalatory interventions (the *mixed* group). We estimated two multilevel multinomial logistic regression models to distinguish effects of the key and control variables across these three groups and the non-intervention reference category. Details of both models are presented in Table A4 in the Appendix. To limit the amount of information displayed, Figure 2 includes only the results of the model that includes both the key variables and the controls. Further, the variable that measured whether the bystander was acting in a professional role ('bystander at work') is excluded because it completely separates the escalatory intervention from non-intervention (no bystanders at work intervened with in an escalatory manner), a phenomenon that renders it impossible to estimate the effect of the predictor in a logistic model.

With respect to the effects of the number of bystanders present, Figure 2 supports the following conclusions. Increasing numbers of bystanders are found to be statistically associated with lower odds of de-escalatory, while escalatory, and mixed intervention outcomes are not statistically related to the outcome. Additional tests demonstrate that only the effect size difference between de-escalatory intervention (0.19) and escalatory intervention (0.68) is significant ($\chi^2(1) = 11.63$, $p < .01$) but not those involving the mixed interventions. Social relations do have a positive and statistically significant effect on the odds of all three intervention types. The difference between the estimates of the de-escalatory and the mixed intervention types is significant ($\chi^2(1) = 8.17$, $p < .01$) but not the differences involving the escalating intervention. Similar to the confirmatory analysis, gender is the only control variable significantly related to intervention. Males are more likely than women to display de-escalatory, escalatory, and mixed interventions. These effects sizes do not significantly differ between the three intervention types ($\chi^2(1) = 2.35$, $p = .12$ for escalatory versus de-escalatory intervention, $\chi^2(1) = 1.42$, $p = .23$ for de-escalatory versus mixed intervention, and $\chi^2(1) = .38$, $p = .54$ for escalatory versus mixed intervention).

Finally, we conducted a number of sensitivity tests to assess the robustness of our findings against alternative, reasonable data and model specifications. These include an alternative scaling method for our samplings weights, and a curvilinear effect of number of bystanders. In Figures 1 and 2 and the corresponding Tables A3 and A4 in Appendix A, we used scaling method A as described by Carle (2009). Following Carle's recommendation, we also used method B to verify that our findings did not depend on the scaling method. The results of using both scaling methods proved similar, given that all estimates barely differed across the scaling methods. These results are available in the online Supporting Information at osf.io.

Finally, given prior suggestions of a negative curvilinear association between number of bystanders and intervention (Latané, 1981), we estimated the four models shown in Tables A3 and A4 again, but with an added squared number of bystanders term. In support this suggestion, the results demonstrate that for undifferentiated intervention and for de-escalatory intervention, the negative effect of each additional bystander becomes significantly weaker (less negative) as the number of bystanders increases. For example, going from 2 to 3 bystanders reduces the likelihood of intervention more than going from 12 to 13 bystanders. These results are also available in the online Supporting Information at osf.io.

Discussion

Do people help those in need in times of potential danger? Social science has a long tradition of stressing that third-party individuals are indifferent to the plight of others (Cohen, 2001; Manning et al., 2007; Milgram, 1970). A particularly influential account is offered by the bystander field, which stipulates that people rarely intervene to help, because of the collective apathy generated by being together with others. In the present study, relying on naturally occurring data, we contrasted the number of bystanders present against an alternative explanation of bystander involvement that puts social relations between bystanders and conflict participants front stage. Our confirmatory analysis provided no evidence for the reversed bystander effect (Hypothesis 1). Rather, we found that additional bystanders make individual intervention less likely, as expected under the classical bystander effect hypothesis. Further, data offered compelling evidence that the bystanders' social relation with conflict participants are associated with bystander intervention (Hypothesis 2), and that the effect size is larger in magnitude than that of the number of bystanders predictor (Hypothesis 3).

Further, our subsequent exploratory analysis of decomposed bystander intervention suggests that the negative effect of bystander numbers mainly applies to de-escalatory interventions, while social relations with conflict participants are highly predictive of all intervention types—whether de-escalatory, escalatory, or mixed. Finally, the sensitivity analysis indicates that the negative effect of the number of bystanders on de-escalating intervention may diminish with increasing numbers of bystanders (i.e., a decreasing marginal effect), as suggested in earlier bystander research (Latané, 1981).

The bystander effect field has for decades focused on people presence as the chief predictor of intervention behavior—initially as an explanation of non-intervention (Latané & Darley, 1970), and more recently, in dangerous contexts, as a facilitator of intervention (Fischer et al., 2011). Here, with the largest dataset of video captured real-life dangerous conflicts, we do not find evidence of a reversed bystander effect, but instead, a classical bystander effect. This is unexpected, given the recent paradigmatic shift towards an emergent consensus that additional bystanders, in times of danger, offer physical support making intervention more likely (Fischer & Greitemeyer, 2013; Fischer et al., 2011; Levine et al., 2011; Lindegaard et al., 2017).

The reported negative association between bystander numbers and intervention may be received as evidence that bystanders become increasingly apathetic towards the needs of others when situated in more populated contexts (Latané & Darley, 1970). However, we also consider an alternative interpretation, not of collective apathy, but of helping saturation. Unlike the scarcely populated bystander experimental settings, public spaces often contain numerous individuals (with the current study finding an average of 18 bystanders per context), thus offering far more potential help-givers than required to manage a typical conflict. This relatively fixed upper bound of required help-givers has been shown to saturate at around three de-escalatory bystanders (Levine et al., 2011). As such, additional bystanders beyond this point may be surplus to requirements and thus unlikely to intervene (see also Bloch, Liebst, Poder, Christensen, & Heinskou, 2018).

The very strong association between group relations and intervention adds to the accumulating body of evidence showing that group membership is highly predictive of bystander helping (Levine, Cassidy, & Jentzsch, 2010; Lindegaard et al., 2017; Phillips & Cooney, 2005; Slater et al., 2013). Beyond peacekeeping, it is important to recognize that group relationship is also highly predictive of escalatory, aggressive intervention. Here, the intervener acts not as a mediator, but as a partisan who fights on behalf of those in the group (Black, 1993; Phillips & Cooney, 2005; Swann et al., 2010). Given the accumulating evidence supporting group relationship as a key predictor of intervention behavior, it is unfortunate that helping research, and the social sciences more broadly, continue to emphasize the 'power of situation,' at the expense of personal and group-based agency (Lefevor, Fowers, Ahn, Lang, & Cohen, 2017; Smith, 2015; Swann & Jetten, 2017). In the current intervention study, that compares the effect of situational bystander presence to the effect of group dynamics, the latter predictor is many-fold larger in magnitude. As such, people presence matters; in part, as a count in number, but more so as a consideration of the social ties existing between those present.

In addition to these two main predictors, we also included a number of control variables. Male bystanders were found to have a higher likelihood of intervention than females (across all intervention subtypes and model specifications). This is in line with review evidence suggesting a gender-difference in helping behavior, with males being more strength-intensive and risk-averse in their helping strategies than females (Eagly, 2009). Furthermore, occupational role (e.g., as a bouncer) was found to be a perfect predictor of the escalatory outcome category, with zero cases of bystander-workers intervening in a purely escalatory manner (see Table S3 in the Supplemental

Material). This finding suggests that professional ‘place managers’ are less prone to use excessive force than indicated in prior research (Roberts, 2009; Sampson et al., 2010).

In utilizing naturally occurring data, the current work contributes to the scholarly understanding of actual bystander behavior as situated in dangerous emergencies. This was rendered possible by the sampling of police-reported events, all of which contained actual physical assaults. The current high-danger sample satisfies the call for research assessing bystander behavior in violently dangerous emergencies (Fischer et al., 2006), which is difficult to simulate ethically in the lab. The reliance on high-danger police-reported data also incurs several limitations, however. As police-reported data are skewed towards more violently severe conflicts (Lindegaard & Bernasco, 2018; Tarling & Morris, 2010), our data does not capture the more mundane emergencies and non-violent confrontations, commonplace in public settings (Copes, Hochstetler, & Forsyth, 2013). Furthermore, although bystander intervention was predominately de-escalatory in our data, it is likely that the current sample under-represents the proportion of de-escalatory acts, while over-representing the escalatory acts, in the intervention outcome. Specifically, while escalatory bystander interventions may exacerbate the conflict and make it of greater interest to the police, other conflicts successfully de-escalated by bystanders, before they could become severe, are likely absent from our sample (Levine et al., 2012). As such, one should be wary of generalizing the current findings to bystander intervention occurring outside of high-danger, police-reported assaults (see Berk, 1983). Where possible, future research should prioritize random probability sampling of emergency incidents, violent and mundane alike.

As a final limitation, the very large effect size of group relations may, in part, be inflated because the coders (subconsciously, against their instructions) inferred the bystanders’ relationship ties from whether or not the bystander intervened. In the current study, however, coders had detailed police case files accompanying each video, which were consulted to settle ambiguous video-based assessments of group membership. It is important to note that there were few discrepancies during this qualitative validation. Adding to this, the reported association between group relations and intervention is what one may expect, given that all prior studies (to our best knowledge) testing this association report a positive effect, typically of substantial magnitude. However, future bystander research should, ideally, consider conducting formal interrater *validity* tests (in addition to standard interrater reliability tests) in which video-based assessments are compared against ratings where group membership is definitively known (see Afifi & Johnson, 2005).

Cialdini (1980) describes a 'full cycle' psychology, by which experimentation should be prompted by the naturalistic observation of social phenomena (e.g., the murder of Kitty Genovese), and, in turn, validated through systematic real-world observation. The bystander research field, still largely contained in experimental work, is yet to fully confirm the ecological validity of its setup and findings. A case in point is that bystander studies typically compare rates of intervention when the bystander is alone versus when in the presence of a few others. The prevalence of numerous bystanders in public spaces suggests, however, that solitary conditions—similar to the simulation of non-dangerous emergencies in the presence of strangers only—are over-studied artifacts of the laboratory. With real-life video data, we gain a greater understanding of how bystanders actually behave when together in numbers. This allows a reconsideration of whether non-intervention by individuals in populated settings reflects bystander apathy, or alternatively, bystander surplus. In taking such steps, the field may satisfy the final turn in Cialdini's (1980) cycle, and in doing so, recalibrate the 'external invalidity' (Mook, 1983) of the experimental bystander paradigm towards a higher ecological validity.

Third-party conflict intervention is a probable human universal. Our work evidences that this needs to be understood together with another universal, noted by Brown (1991): in-group favoritism. This bias towards one's own may promote de-escalatory helping towards familiar victims, as shown in the current study. However, the boundaries of 'us' and 'them' may also be an obstacle for the provision of assistance from strangers (Bloom, 2017), and may promote pro-group partisan fighting on behalf of those known (Swann et al., 2010). We suggest that research gravitate away from chiefly using bystander counts to explain non-intervention. Rather, in our view, both the event and the non-event of bystander involvement, as well as its helpful and harmful consequences, calls for an appreciation of the group processes existing between those present.

References

Afifi, W. A., & Johnson, M. L. (2005). The nature and function of tie-signs. In V. L. Manusov (Ed.), *The Sourcebook of Nonverbal Measures: Going Beyond Words* (pp. 189-198). New Jersey: Psychology Press.

Banyard, V. L. (2015). *Toward the next generation of bystander prevention of sexual and relationship violence: Action coils to engage communities*. London: Springer.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*(4), 396-403. doi:10.1111/j.1745-6916.2007.00051.x

Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review, 48*(3), 386-398. doi:10.2307/2095230

Black, D. (1993). *The social structure of right and wrong*. San Diego: Academic Press.

Bloch, C., Liebst, L. S., Poder, P., Christensen, J., & Heinskou, M. B. (2018). Caring collectives and other forms of bystander helping behavior in violent situations. *Current Sociology*. doi:10.1177/0011392118776365

Bloom, P. (2017). Empathy and its discontents. *Trends in cognitive sciences, 21*(1), 24-31. doi:10.1016/j.tics.2016.11.004

Boehm, C. (2000). Conflict and the evolution of social control. *Journal of Consciousness Studies, 7*(1-2), 79-101.

Brown, D. E. (1991). *Human universals*: McGraw-Hill New York.

Carle, A. C. (2009). Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC medical research methodology, 9*(49). doi:10.1186/1471-2288-9-49

Cherry, F. (1995). *The 'stubborn particulars' of social psychology: Essays on the research process*. London: Routledge.

Cialdini, R. B. (1980). Full-cycle social psychology. *Applied social psychology annual, 1*, 21-47.

Cohen, S. (2001). *States of denial: Knowing about atrocities and suffering*. Cambridge: Polity Press.

Copes, H., Hochstetler, A., & Forsyth, C. J. (2013). Peaceful warriors: Codes for violence among adult male bar fighters. *Criminology, 51*(3), 761-794. doi:10.1111/1745-9125.12019

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion, 12*(5), 1085-1101. doi:10.1037/a0025737

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. *Journal of Personality and Social Psychology, 8*(4), 377-383. doi:10.1037/h0025589

de Waal, F. B. M. (2015). Prosocial Primates. In D. A. Schroeder & W. G. Graziano (Eds.), *The Oxford Handbook of Prosocial Behavior* (pp. 61–85). Oxford: Oxford University Press.

de Waal, F. B. M., & Preston, S. D. (2017). Mammalian empathy: behavioural manifestations and neural basis. *Nature Reviews Neuroscience, 18*(8), 498-509. doi:10.1038/nrn.2017.72

Eagly, A. H. (2009). The his and hers of prosocial behavior: an examination of the social psychology of gender. *American Psychologist, 64*(8), 644-658. doi:10.1037/0003-066X.64.8.644

Eibl-Eibesfeldt, I. (1989). *Human ethology*. New York: Aldine de Gruyter.

Fischer, P., & Greitemeyer, T. (2013). The positive bystander effect: Passive bystanders increase helping in situations with high expected negative consequences for the helper. *The Journal of Social Psychology, 153*(1), 1-5. doi:10.1080/00224545.2012.697931

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*(2), 267-278. doi:10.1002/ejsp.297

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., . . . Kainbacher, M. (2011). The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological bulletin, 137*(4), 517-537. doi:10.1080/00224545.2012.697931

Fry, D. P. (2000). Conflict management in cross-cultural perspective. In F. Aureli & F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 334-351). Berkeley: University of California Press.

Fujisawa, K. K., Kutsukake, N., & Hasegawa, T. (2006). Peacemaking and consolation in Japanese preschoolers witnessing peer aggression. *Journal of comparative psychology, 120*(1), 48-57. doi:10.1037/0735-7036.120.1.48

Ge, W., Collins, R. T., & Ruback, R. B. (2012). Vision-based analysis of small groups in pedestrian crowds. *IEEE transactions on pattern analysis and machine intelligence, 34*(5), 1003-1016. doi:10.1109/TPAMI.2011.176

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine, 27*(15), 2865-2873. doi:10.1002/sim.3107

Goffman, E. (1971). *Relations in public. Microstudies of the public order*. New York: Basic Books.

Grimes, D. A., & Schulz, K. F. (2005). Compared to what? Finding controls for case-control studies. *The Lancet, 365*(9468), 1429-1433. doi:10.1016/S0140-6736(05)66379-9

Hamilton, W. D., & Axelrod, R. (1981). The evolution of cooperation. *Science, 211*(27), 1390-1396.

Harari, H., Harari, O., & White, R. V. (1985). The reaction to rape by American male bystanders. *The Journal of social psychology, 125*(5), 653-658. doi:10.1080/00224545.1985.9712039

Hobbs, D. (2003). *Bouncers: Violence and governance in the night-time economy*. Oxford: Oxford University Press.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. London: Routledge.

Jones, L. K., Jennings, B. M., Goelz, R. M., Haythorn, K. W., Zivot, J. B., & de Waal, F. B. (2016). An ethogram to quantify operating room behavior. *Annals of Behavioral Medicine, 50*(4), 487-496. doi:10.1007/s12160-016-9773-0

Keogh, R. H., & Cox, D. R. (2014). *Case-control studies*. Cambridge: Cambridge University Press.

Krippendorff, K. (2004). Reliability in Content Analysis. *Human Communication Research, 30*(3), 411-433. doi:10.1111/j.1468-2958.2004.tb00738.x

Latané, B. (1981). The psychology of social impact. *American Psychologist, 36*(4), 343-356. doi:10.1037/0003-066X.36.4.343

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.

Latané, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological bulletin, 89*(2), 308-324. doi:10.1037/0033-2909.89.2.308

Lefevor, G. T., Fowers, B. J., Ahn, S., Lang, S. F., & Cohen, L. M. (2017). To what degree do situational influences explain spontaneous helping behaviour? A meta-analysis. *European Review of Social Psychology, 28*(1), 227-256. doi:10.1080/10463283.2017.1367529

Levine, M., Cassidy, C., Brazier, G., & Reicher, S. D. (2002). Self-categorization and bystander non-intervention: Two experimental studies. *Journal of Applied Social Psychology, 32*(7), 1452-1463. doi:10.1111/j.1559-1816.2002.tb01446.x

Levine, M., Cassidy, C., & Jentzsch, I. (2010). The implicit identity effect: Identity primes, group size, and helping. *British Journal of Social Psychology, 49*(4), 785-802. doi:10.1348/014466609X480426

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*(6), 1429-1439. doi:10.1037/a0012634

Levine, M., Lowe, R., Best, R., & Heim, D. (2012). 'We police it ourselves': Group processes in the escalation and regulation of violence in the night-time economy. *European journal of social psychology, 42*(7), 924-932. doi:10.1002/ejsp.1905

Levine, M., & Manning, R. (2013). Social identity, group processes, and helping in emergencies. *European Review of Social Psychology, 24*(1), 225-251. doi:10.1080/10463283.2014.892318

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*(3), 406-412. doi:10.1177/0956797611398495

Lewallen, S., & Courtright, P. (1998). Epidemiology in practice: case-control studies. *Community Eye Health, 11*(28), 57-58.

Liebst, L. S., Heinskou, M. B., & Ejbye-Ernst, P. (2018). On the actual risk of bystander intervention: A statistical study based on naturally occurring violent emergencies. *Journal of Research in Crime and Delinquency, 55*(1), 27-50. doi:10.1177/0022427817710776

Lindegaard, M. R., & Bernasco, W. (2018). Lessons learned from crime caught on camera. *Journal of Research in Crime and Delinquency, 55*(1), 155-186. doi:10.1177/0022427817727830

Lindegaard, M. R., Liebst, L. S., Bernasco, W., Heinskou, M. B., Philpot, R., Levine, M., & Verbeek, P. (2017). Consolation in the aftermath of robberies resembles post-aggression consolation in chimpanzees. *PloS one, 12*(5), e0177725. doi:10.1371/journal.pone.0177725

Lohr, S. (2010). *Sampling: Design and Analysis*. Boston: Brooks/Cole.

Macintyre, S., & Homel, R. (1997). Danger on the dance floor: A study of interior design, crowding and aggression in nightclubs. In R. Homel (Ed.), *Policing for prevention: Reducing crime, public intoxication and injury* (pp. 91-113). Monsey, NY: Criminal Justice Press.

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*(6), 555-562. doi:10.1037/0003-066X.62.6.555

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review, 28*(2), 295-314. doi:10.1007/s10648-014-9287-x

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts and applications*. Los Angeles: Sage.

Milgram, S. (1970). The experience of living in cities. *Science, 167*(3924), 1461-1468.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38*(4), 379-387. doi:10.1037/0003-066X.38.4.379

Mortensen, C. R., & Cialdini, R. B. (2010). Full-cycle social psychology for theory and application. *Social and Personality Psychology Compass, 4*(1), 53-63. doi:10.1111/j.1751-9004.2009.00239.x

Murphy, N. A. (2016). What we know and the future of interpersonal accuracy research. In J. A. Hall, M. S. Mast, & T. V. West (Eds.), *The social psychology of perceiving others accurately* (pp. 404-424). Cambridge: Cambridge University Press.

Nassauer, A., & Legewie, N. (2018). Video Data Analysis: A Methodological Frame for a Novel Research Trend. *Sociological methods & research*, 1-40. doi:10.1177/0049124118769093

Osswald, S., Greitemeyer, T., Fischer, P., & Frey, D. (2010). What is moral courage? Definition, explication, and classification of a complex construct. In C. Pury & S. Lopez (Eds.), *The psychology of courage: Modern research on an ancient virtue* (pp. 149-164). Washington: APA.

Parks, M. J., Osgood, D. W., Felson, R. B., Wells, S., & Graham, K. (2013). Third party involvement in barroom conflicts. *Aggressive behavior, 39*(4), 257-268. doi:10.1002/ab.21475

Phillips, S., & Cooney, M. (2005). Aiding peace, abetting violence: Third parties and the management of conflict. *American Sociological Review, 70*(2), 334-354. doi:10.1177/000312240507000207

Philpot, R. (2017). *Beyond the dyad: The role of groups and third-parties in the trajectory of violence.* (Doctoral dissertation), University of Exeter, Exeter.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2005). Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics, 128*(2), 301-323. doi:10.1016/j.jeconom.2004.08.017

Reynald, D. M. (2011). Factors associated with the guardianship of places: Assessing the relative importance of the spatio-physical and sociodemographic contexts in generating opportunities for capable guardianship. *Journal of Research in Crime and Delinquency, 48*(1), 110-142.

Roberts, J. C. (2009). Bouncers and barroom aggression: A review of the research. *Aggression and Violent Behavior, 14*(1), 59-68. doi:10.1016/j.avb.2008.10.002

Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of social service Research, 21*(4), 37-59. doi:10.1300/J079v21n04_02

Sampson, R., Eck, J. E., & Dunham, J. (2010). Super controllers and crime prevention: A routine activity explanation of crime prevention success and failure. *Security Journal, 23*(1), 37-51. doi:10.1057/sj.2009.17

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, 8*(1), e52766. doi:10.1371/journal.pone.0052766

Smith, C. (2015). *To flourish or destruct: a personalist theory of human goods, motivations, failure, and evil*. Chicago: University of Chicago Press.

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science, 11*(5), 702-712. doi:10.1177/1745691616658637

Swann, W. B. J., Gómez, A., Huici, C., Morales, J., & Hixon, J. G. (2010). Identity fusion and self-sacrifice: Arousal as a catalyst of pro-group fighting, dying, and helping behavior. *Journal of personality and social psychology, 99*(5), 824-841. doi:10.1037/a0020014

Swann, W. B. J., & Jetten, J. (2017). Restoring Agency to the Human Actor. *Perspectives on Psychological Science, 12*(3), 382-399. doi:10.1177/1745691616679464

Tarling, R., & Morris, K. (2010). Reporting crime to the police. *The British Journal of Criminology, 50*(3), 474-490. doi:10.1093/bjc/azq011

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychological review, 107*(3), 411-429. doi:10.1037//0033-295X.107.3.411

Tinbergen, N. (1963). On aims and methods of ethology. *Ethology, 20*(4), 410-433. doi:10.1111/j.1439-0310.1963.tb01161.x

Vázquez, A., Gómez, Á., Ordoñana, J. R., Swann, W. B. J., & Whitehouse, H. (2017). Sharing genes fosters identity fusion and altruism. *Self and Identity, 16*(6), 684-702. doi:10.1080/15298868.2017.1296887

Figures

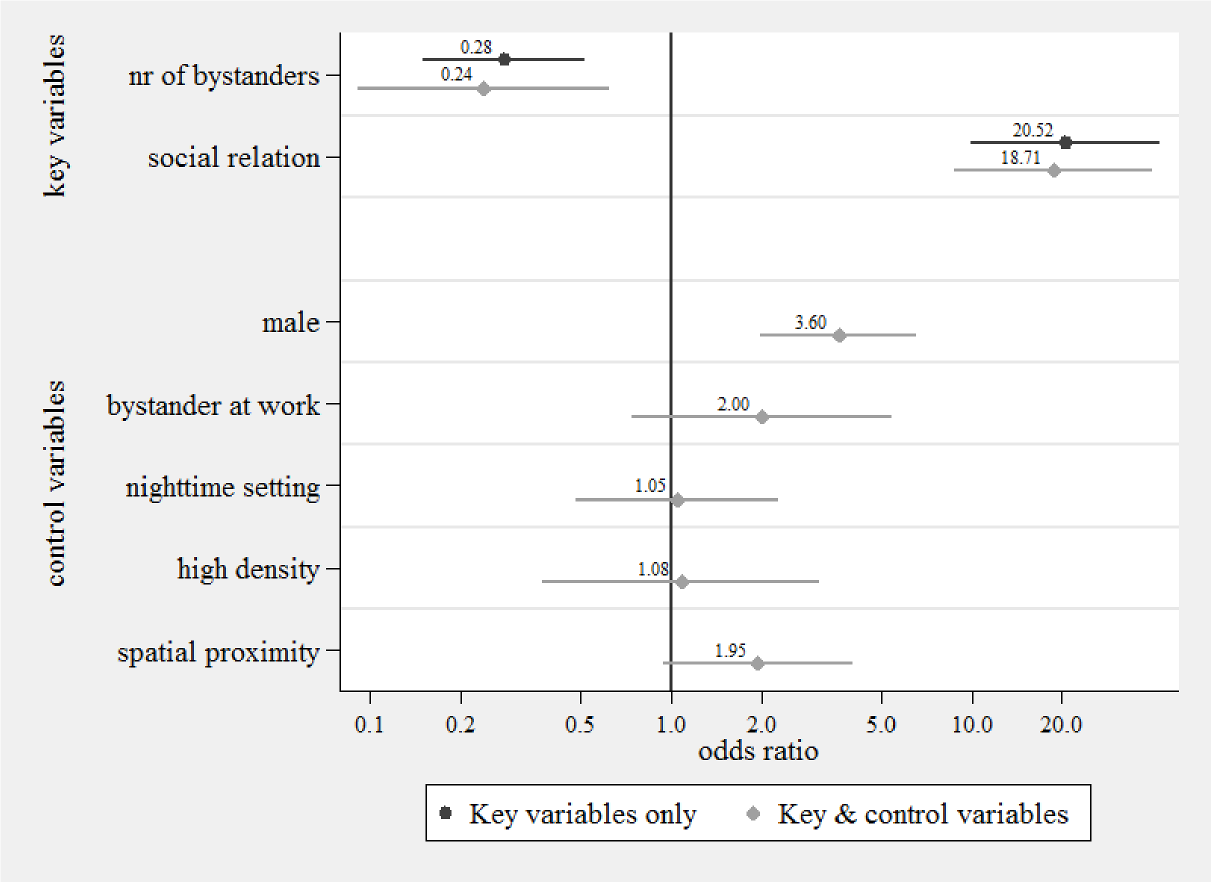

Figure 1: Multilevel binomial logistic regression estimates of bystander intervention.

Complete results reported in Table A3 (Appendix).

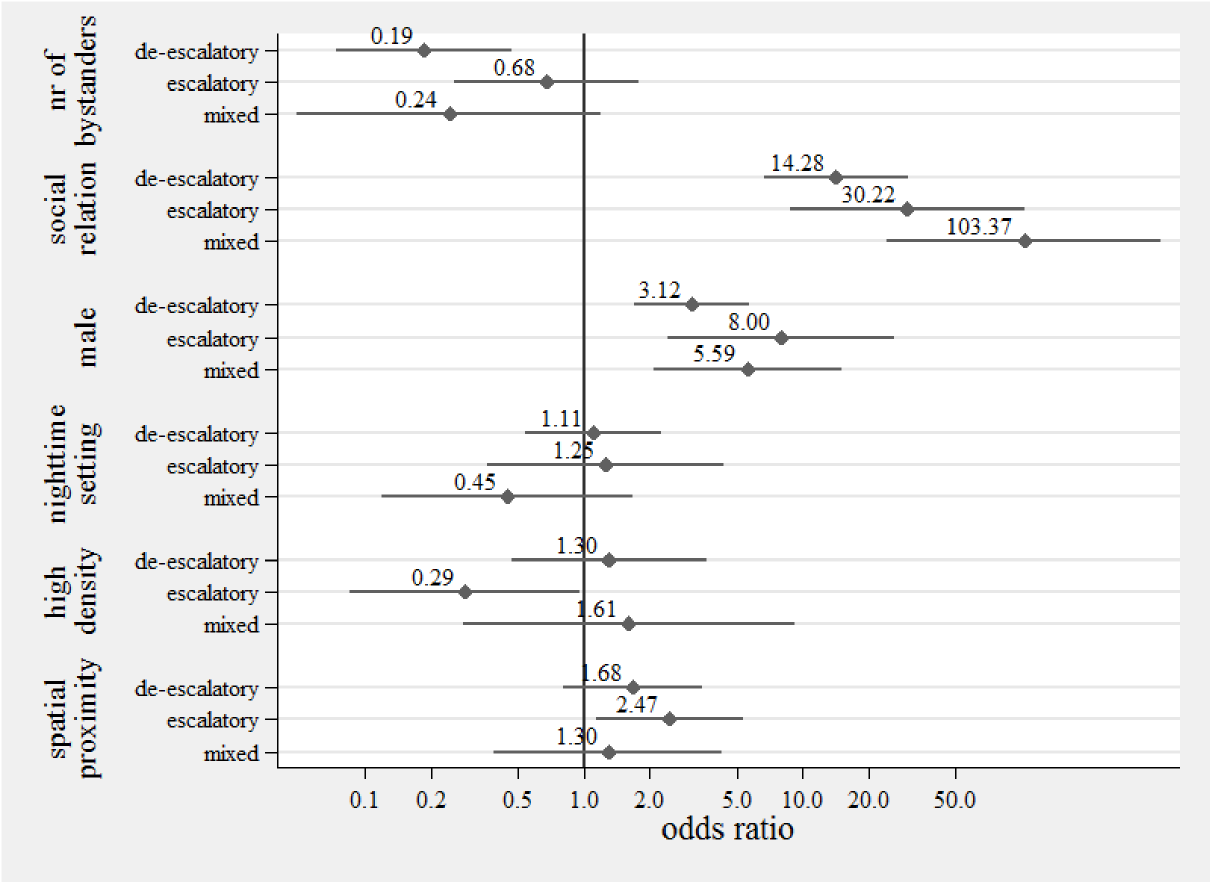

Figure 2: Multilevel multinomial logistic regression estimates of effects of key and control variables on decomposed bystander intervention. No intervention versus de-escalatory, escalatory, and mixed interventions. Complete results reported in Table A4 (Appendix).

Table

Table 1: Descriptive statistics of unweighted variables

Variable	*M*	*SD*	Min	Max	N
bystander intervention	0.29	0.45	0	1	747
decomposed bystander intervention					
de-escalatory	0.20	0.40	0	1	747
escalatory	0.05	0.21	0	1	747
mixed	0.04	0.20	0	1	747
number of bystanders (unstandardized)	18.28	13.73	1	76	747
number of bystanders (rescaled [1])	0.16	0.52	-0.50	2.36	747
social relation	0.29	0.45	0	1	747
male	0.69	0.46	0	1	747
nighttime drinking setting	0.71	0.45	0	1	747
bystander at work	0.11	0.32	0	1	747
spatial proximity	0.44	0.50	0	0	741
people density	0.38	0.49	0	1	747

[1] Rescaled as $x' = x - \mu_x/2\sigma_x$, i.e. subtract the mean and divide by twice the standard deviation (see Gelman, 2008)

Appendix

Table A1: Summary of bystander intervention codes used to construct the outcome variables.

Behaviors	Qualitative definition	Type
Open hand gestures	The bystander displays a calming hand movement with open hands.	De-escalatory
Non-forceful touching	The bystander touches a person in a non-forceful manner.	De-escalatory
Blocking contact between conflict parties	The bystander blocks a person from reaching a conflict party (i.e., acting as a barrier).	De-escalatory
Holding a person back	The bystander holds a person back from moving further towards the conflict or conflict partner.	De-escalatory
Hauling a person off	The bystander holds a person and pulls/carries that individual away from the conflict or conflict partner.	De-escalatory
Pushing	The bystander pushes a person away from the conflict or conflict partner in a non-aggressive manner.	De-escalatory
Pointing and threatening gestures	The bystander displays an aggressive hand movement, typically pointing at someone in a threating manner.	Escalatory
Throw a person	The bystander firmly grips a person and then throws that person in an aggressive manner.	Escalatory
Shoving	The bystander shoves a person in a forceful and aggressive manner.	Escalatory
Hit	The bystander hits a person with either an open or closed hand.	Escalatory
Several hits	The bystander hits several times with either an open or closed hand.	Escalatory
Kick	The bystander kicks a person.	Escalatory
Several Kicks	The bystander kicks a person several times.	Escalatory
Kick to the head	The bystander kicks a person to the head or stomps on a person's head.	Escalatory
Violence against a person on the ground	The bystander physically attacks a person on the ground.	Escalatory
Weapon use	The bystander physically attacks a person with an object (e.g., billiard ball, bottle, knife).	Escalatory

Note. The above codes were used to construct the binary intervention outcome (i.e., any intervention or none), as well as the bystander intervention outcome de-composed into four outcomes (i.e., de-escalatory, escalatory, mixed, none). The Krippendorff's alphas of the de-escalatory and escalatory intervention codes are .92 and .82, respectively. A mixed outcome is coded for bystanders displaying both escalatory and de-escalatory interventions.

Table A2: Summary of independent variable definitions and related Krippendorff's alphas.

Variable	Description	Krippen-dorff's α
Number of bystanders	The number of bystanders present in the situation at the point when the conflict initiates.	.85
Social relation	The bystander knows at least one person (victim and/or perpetrator) who is physically involved in the conflict. We apply a minimal definition of relationship ties, which include everything from ties established the same day to family ties.	1.0
Male	Gender based on the bystander's visual appearance.	1.0
Bystander at work	The bystander is performing an occupational role (e.g., as a bouncer or bar staff). Excludes emergency services (e.g., medics or police officers).	1.0
Nighttime drinking setting	The incident took place 10PM–7AM during the weekend, or if inside/in front of a drinking establishment.	1.00
High density	The density of everyone present in the situation at the point when the conflict initiates. High density is assessed from whether it is possible to walk across the setting (i.e. dance floor, street) in a straight line, without bumping into someone present.	.83
Spatial proximity	The bystander is within a 2-meters radius from where the conflict initiates.	.81

Table A3: Multilevel binomial logistic regression estimates of bystander intervention.

	Key variables only			Key and control variables		
	OR	95% CI	p	OR	95% CI	p
number of bystanders	0.28***	0.15–0.52	0.00	0.24**	0.09–0.62	0.00
social relation	20.52***	9.98–42.17	0.00	18.71***	8.75–40.03	0.00
male				3.60***	1.98–6.55	0.00
bystander at work				2.00	0.74–5.42	0.17
nighttime setting				1.05	0.48–2.29	0.90
high density				1.08	0.37–3.12	0.89
spatial proximity				1.95	0.94–4.03	0.07
N1 (individuals)	751			741		
N2 (incidents)	81			80		

OR = odds ratio, CI = confidence interval, *** $p < .001$ ** $p < .01$ * $p < .05$

Table A4: Multi-level multinomial logistic regression estimates of decomposed bystander intervention.

	Key variables only			Key and control variables		
	OR	95% CI	p	OR	95% CI	p
de-escalatory						
number of bystanders	0.26***	0.14–0.48	0.00	0.19***	0.07–0.47	0.00
social relation	14.53***	7.06–29.91	0.00	14.28***	6.75–30.22	0.00
male				3.12***	1.70–5.74	0.00
nighttime setting				1.11	0.54–2.28	0.77
high density				1.30	0.47–3.64	0.61
spatial proximity				1.68	0.81–3.50	0.16
escalatory						
number of bystanders	0.43	0.17–1.08	0.07	0.68	0.26–1.79	0.43
social relation	35.70***	9.66–131.85	0.00	30.22***	8.84–103.33	0.00
male				8.00***	2.42–26.50	0.00
nighttime setting				1.25	0.36–4.34	0.72
high density				0.29*	0.09–0.96	0.04
spatial proximity				2.47*	1.14–5.38	0.02
mixed						
number of bystanders	0.24**	0.09–0.66	0.01	0.24	0.05–1.20	0.08
social relation	93.52***	26.50–330.06	0.00	103.37***	24.54–435.40	0.00
male				5.59***	2.08–15.02	0.00
nighttime setting				0.45	0.12–1.67	0.23
high density				1.61	0.28–9.20	0.59
spatial proximity				1.30	0.39–4.32	0.67
N1 (individuals)	751			744		
N2 (incidents)	81			80		

OR = odds ratio, CI = confidence interval, *** p < .001 ** p < .01 * p < .05