

This is the accepted manuscript (postprint version) of the article. Contentwise, this version is identical to the published manuscript, but there may be differences in typography, layout, and editing.

Please cite the final published version:

Liebst, L. S., Philpot, R., Levine, M., and Lindegaard, M. R. (2020). Cross-National CCTV Footage Shows Low Victimization Risk for Bystander Interveners in Public Conflicts, *Psychology of Violence*, <http://dx.doi.org/10.1037/vio0000299>

Cross-National CCTV Footage Shows Low Victimization Risk for Bystander Interveners in Public Conflicts

Lasse Suonperä Liebst^{1*}, Richard Philpot², Mark Levine², and Marie Rosenkrantz Lindegaard^{1,3}

¹ Department of Sociology, University of Copenhagen, Øster Farimagsgade 5, Building 16, 1353 Copenhagen K, Denmark.

³ Department of Psychology, Lancaster University, Lancaster, LA1 4YF, United Kingdom.

³ Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands.

*Corresponding author: Lasse Suonperä Liebst, email: ls1@soc.ku.dk

Author Contributions (CRediT taxonomy): Conceptualization, L.S.L., R.P., and M.R.L.; Methodology, L.S.L., R.P., and M.R.L.; Software, L.S.L.; Validation, R.P.; Formal analysis, L.S.L.; Investigation, L.S.L., R.P., and M.R.L.; Data Curation, L.S.L. and R.P.; Resources, M.L. and M.R.L.; Supervision, M.R.L.; Project Administration, M.R.L.; Funding Acquisition, M.R.L.; Writing—Original Draft, L.S.L.; Writing—Review & Editing, L.S.L., R.P., M.L., and M.R.L.

Abstract

Objective: Accumulating evidence shows that bystanders witnessing public disputes frequently intervene to help. However, little is known regarding the risks entailed for those bystanders who enter the fray to stop conflicts. This study systematically examined the prevalence of bystander victimizations and associated risk factors. Method: Data were a cross-national sample of 93 CCTV video recordings of real-life public disputes, capturing the potential victimizations of 417 intervening and 636 non-intervening bystanders. Results: Data showed that interveners were rarely physically harmed—at a rate of 3.6%—and non-interveners were virtually never victimized. Confirmatory regression results showed that conflict party affiliation was associated with bystander victimization, although only moderately robust. The gender of the intervener was a highly fragile correlate with the outcome. The severity of the conflict at the time of intervention was not found to increase the risk of victimization. Conclusions: Our findings

highlight the ecological value of naturalistic observation for bystander research and emphasize the need for evidence-based bystander intervention recommendations. Data, materials, and postprint are available at osf.io/vyutj.

Keywords: bystander victimization, helping behavior, video observation

Introduction

There are long-standing scholarly assumptions and public concerns that bystanders largely remain passive when witnessing disputes in public settings (Darley & Latané, 1968; Milgram, 1970). However, accumulating evidence questions this notion (Fischer et al., 2011; Stalder, 2008), with recent observational work showing that in as many as nine out of ten real-life public conflicts at least one bystander helps (Philpot, Liebst, Levine, Bernasco, & Lindegaard, 2019). While this high bystander intervention rate is reassuring for potential victims and the public, it raises the question of whether it is safe for bystanders to intervene. Specifically, by entering an ongoing conflict, a bystander may become a target of aggression or be harmed by collateral damage. Despite the relevance for public safety initiatives, the literature examining the victimization risk for helping-giving bystanders is surprisingly limited, and the current paper addresses this research gap.

This lack of bystander victimization research reflects the wider methodological circumstance that the experimental method—the default approach applied in bystander research—cannot, for obvious ethical reasons, expose test subjects to potential harm (Osswald, Greitemeyer, Fischer, & Frey, 2010). Additionally, self-reported accounts commonly applied in victimization research offer unreliable data for studying interaction sequences in violent events (Lindegaard & Bernasco, 2018), including the role played by bystanders (Philpot, Liebst, Møller, Lindegaard, & Levine, 2019).

The few empirical studies on the prevalence of bystander victimizations offer some grounds for optimism, however. Victimization survey data (Planty, 2002) and systematic newspaper assessments (Sherman, Steele, Laufersweiler, Hoffer, & Julian, 1989) suggest that bystanders are unlikely to be harmed during violent crimes, although the risk may vary across types of interpersonal aggression. For example, Hamby, Weber, Grych, and Banyard (2016) note a discrepancy between the percentage of bystanders hurt or threatened when intervening into youth arguments involving relatives (5.8%), and the likelihood of victimization when intervening into sexual assaults (17.8%). Further, a CCTV analysis of conciliatory and aggressive bystander interventions in real-life public assaults in Copenhagen, Denmark, reports a physical victimization rate of 18% (Liebst, Heinskou, & Ejbye-Ernst, 2018). In this recent study, the risk of bystander victimization if entering the conflict to solely de-escalate the fight is lower, at just over one in ten.¹ Reports of higher bystander victimization rates do exist, such as

¹ The reported victimization risk reported in Liebst, Heinskou, et al. (2018) did not distinguish between de-escalatory and escalatory involvement. The cited victimization risk for pure de-escalators is based on a reanalysis of the raw data included in that study.

Moschella, Bennett, and Banyard (2018) who find that approximately one in three interveners in sexual violence experience negative perpetrator responses. Note, however, that this relatively higher prevalence should be considered in light of the very inclusive definition of negative responses applied in this study (e.g., that the perpetrator was ‘upset at’ the intervener, without retaliating physically).

The literature also identifies a number of risk factors associated with bystander victimization. In line with the general insight that bystander helping is shaped by group dynamics (Levine & Manning, 2013; Liebst, Philpot, et al., 2019), a pre-existing social tie between bystander and conflict victim is found to be positively associated with bystander victimization (Liebst, Heinskou, et al., 2018). This suggests that bystander intervention on behalf of a victimized group member may leave the intervener vulnerable to the perpetrator’s out-group aggression (see Brewer, 2001). Alternatively, a familiar bystander who intervenes towards an in-group aggressor may be more likely retaliated against than an intervening stranger, suggesting an operating in-group aggression dynamic (see Marques, 1990).

Further, aggressive rather than placatory intervention acts are found to enhance the victimization risk, both among members of the public and law enforcement officers (Liebst, Ejbye-Ernst, et al., 2019; Liebst, Heinskou, et al., 2018; van Reemst, Fischer, & Zwirs, 2015). The nighttime drinking setting is also reported as a situational risk factor for bystander victimizations (Liebst, Heinskou, et al., 2018), corresponding to the finding that these locations are well-known hotspots of violent crime (Hadfield, Lister, & Traynor, 2009). Plausibly, female bystanders are less likely to be victimized than males, because they are perceived as less of a potential physical threat (Christie, 1986; Copes, Hochstetler, & Forsyth, 2013)—note, however, that the empirical evidence supporting a gender-difference is mixed (Belknap & Shelley, 1993; Liebst, Heinskou, et al., 2018; Moschella et al., 2018; Rabe-Hemp & Schuck, 2007). The severity of the conflict is likewise a plausible risk factor of bystander victimization, in line with the argument that violence often gets out of control and, as such, may spillover to bystanders (Collins, 2013). Again, however, the supporting evidence remains inconclusive (Hamby et al., 2016; Liebst, Heinskou, et al., 2018).

Taken together, there is a dearth of knowledge of the prevalence and risk factors of bystander victimization. This is unsatisfactory not only from a scholarly point of view but also from an applied perspective, as it remains uncertain for safety agencies whether it is too dangerous to recommend that bystanders intervene in public conflicts. The current study aims to address this gap by examining the prevalence and associated risk factors of bystander intervention in real-life public conflicts. As such, the current study attempts to validate the findings of Liebst, Heinskou, et al. (2018), which is the only existing direct observational (video-based) study of bystander victimization. As expected from this prior study, and the theoretical view that victimization may be linked to (in or out) group aggression dynamics (Brewer, 2001; Marques, 1990), we hypothesize that intervening bystanders with a conflict party affiliation

have a larger victimization likelihood than intervening strangers (Hypothesis 1). Further, with the aim of clarifying the mixed evidence regarding gender and conflict severity, we hypothesize that female interveners have a lower victimization risk than males (Hypothesis 2), and that the victimization risk is positively associated with conflict severity (Hypothesis 3). Given the confirmatory study ambition (see Cumming, 2014), these hypotheses are tested following a pre-analysis plan (available at osf.io/s4hdx).

Methods

Data comprised of video clips displaying aggressive public incidents captured by actively monitored CCTV cameras in the inner city areas of Amsterdam (the Netherlands), Cape Town (South Africa), and Lancaster (the United Kingdom). The selection of these countries was partially pragmatic, in the sense that access to CCTV data is difficult to acquire and we were able to do so because the researchers involved had local knowledge or prior working relationships with agencies in these cities. One benefit of the current sample is that it is less selected on the dependent variable than, for example, Liebst, Heinskou, et al. (2018) who relied on severe police-reported violent assaults, which most likely over-represent the rate of bystander victimizations. As such, the current data captured incidents ranging from non-violent quarrels to severe violent confrontations of interest to the police. Data access was provided by local police or municipal agencies, and the study was approved by The Danish Data Protection Agency (ref. no 514-0011/18-2000).

The raw sample comprised 1,225 clips.² From this sample, we selected 219 clips conforming to the following criteria: The captured incident occurred in an urbanized, inner-city area. The clips captured a conflict between at least two antagonists, at any level of aggression. The clips had a high quality (e.g., brightness, resolution) that allowed for detailed behavioral coding, with no, or only negligible, breaks in the observed interaction sequences (e.g., a participant moves shortly behind an obstacle, see Nassauer & Legewie, 2018). Given that behavioral video coding is very labor-intensive, we randomly chose a subset of 93 clips, across which all intervening bystanders were coded ($n = 417$). In order to make comparisons between intervening and non-intervening bystanders, we also coded a sample of non-intervening bystanders ($n = 636$). In sparsely populated contexts, we coded all present non-intervening bystanders; in highly populated contexts, we randomly selected a sub-sample of non-interveners.

Several statistical power scenarios were calculated with G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) prior to our logistic regression analysis of victimization risk factors. The sample of 417 intervening bystanders was above the approximate 350 cases found necessary to detect a small effect size (OR ~ 1.60) with a power of 0.80, and $\alpha = .05$. Here, we assumed

² Note that part of the current raw data was analyzed for another study purpose in Philpot et al. (2020). As such, part of the presented sample selection criteria are similar across the two studies.

that any effect size lesser than ‘small’ would be of limited practical significance (Ferguson, 2009; Kirk, 1996).

Coding procedure

A team of six trained student research assistants coded data in accordance with a detailed, interrater agreement tested codebook. This codebook included variable definitions from prior bystander intervention and victimization studies (Levine, Taylor, & Best, 2011; Liebst, Heinskou, et al., 2018), which were further detailed and evaluated through qualitative assessments of a subset of data (see Jones et al., 2016; Lorenz, 1973). This qualitative phase was more interpretative in nature and focused on validating the social meaning of our behavioral codes, considered from the perspective of the participants. Here we followed Valach, von Cranach, and Kalbermatten (1988, p. 252) who “emphasize that it is necessary to use not only physical criteria, but also social meaning, when defining observational units in the construction of observational systems.” By comparison, the subsequent quantitative phase focused on measuring the behavioral units as reliably as possible, ideally with exact agreement across multiple coders (Stemler, 2004).

The first step in the coding procedure involved identifying the, in most cases, two protagonists who initiated or chiefly orchestrated the conflict. All other individuals present in the situation were then defined as bystanders. Those bystanders who performed at least one of the following de-escalatory behaviors were coded as intervening bystanders: pacifying gesturing; pacifying grabbing (e.g., pulling or holding an aggressor back from their target); pacifying pushes; blocking contact; or pacifying touching (e.g., hugging or stroking the body of an antagonist). The coding procedure involved observing each participating bystander, on average around 4 per situation, throughout the entirety of the video clip, lasting on average around 3 minutes. Here, we benefitted from the technical advantage of video data, that each clip may be replayed, zoomed in upon, and slowed down to frame-by-frame instances (Gilmore & Adolph, 2017; Nassauer & Legewie, 2018).

To test the interrater agreement of our variables (Stemler, 2004), we randomly selected 24 videos (25.8% of the sample, containing 38 intervening bystanders) for independent coding. The level of agreement between three coders was evaluated with Krippendorff’s (2004) alpha, following the benchmarks (Landis & Koch, 1977): slight ($< .21$), fair ($\sim .21$), moderate ($\sim .41$), substantial ($\sim .61$), and almost perfect ($\sim .81$). For interrater scores below ‘substantial agreement,’ we also report percentage agreement and Gwet’s (2008) AC_1 agreement statistic, which is more robust to the ‘Kappa’s Paradox’ (Cicchetti & Feinstein, 1990)—the phenomenon in which a high percentage of agreement can result in low interrater scores (though see Krippendorff, 2013).

Measures

Our outcome variable, *bystander victimization*, measures whether the bystander, after the point of first de-escalatory intervention, was victimized by at least one of the following acts: hit; kick; shove; head-butt; grappling (headlock, forceful holding, pullover); attack while on the ground; weapon threat; or weapon use. An AC_1 score of .96 and a percentage agreement of 96.5% suggest high interrater agreement, while a moderate alpha score of .49 appears deflated because the variable is heavily skewed towards non-victimizations (i.e., the raters did not reach near-perfect agreement in the very few victimization cases). We also recorded the victimization of non-intervening bystanders, and for this measure, we reached a fully perfect interrater agreement.

The first explanatory variable, *conflict party affiliation*, measured whether the bystander knew at least one person involved in the conflict at the point of the first intervention. Affiliation was inferred from behavioral displays of personal relationships (i.e., ‘tie-signs,’ Goffman 1971), for example, arriving/leaving the scene together, huddling together, or exchanging greetings (see Afifi & Johnson, 2005; Ge, Collins, & Ruback, 2012). The interrater agreements scores for this variable are moderate, $\alpha = .50$, $AC_1 = .52$, percentage agreement = 75.4%, and this variable should accordingly be interpreted with this in mind. As an alternative operationalization of conflict party affiliation, we also recorded whether the intervening bystander knew a conflict party who had been aggressed against prior to the bystander’s first de-escalatory intervention ($\alpha = .66$). We included, as elaborated further below, this *victim affiliation* measurement to validate how robust our confirmatory analysis is to alternative data specifications.

The second explanatory variable was the bystander’s *gender* ($\alpha = .95$), with female as the reference category. The third explanatory variable was *conflict severity*, which measured the most severe level of aggression on display within a 5-second window prior to the bystander’s first de-escalatory intervention ($\alpha = .90$). This latter measure distinguished between three levels of conflict severity: no observable displays of aggression between the conflict parties; nonviolent displays of aggression (e.g., expressive gestures, poking); violent displays of aggression (e.g., hits, shoves weapon use).³

To control for the possibility that victimization prevalence may vary across geographical contexts with differing levels of violent crimes (Sampson, 2013), we included *urban-national context* as dummy variables: Amsterdam, the Netherlands; Cape Town, South Africa; and Lancaster, the United Kingdom. As a second control, we included the bystander’s *age* ($\alpha = .74$), captured on a continuous scale. To make this variable comparable to binary predictors,

³ Note that we—as described in the codebook and pre-analysis plan analysis—coded violent aggression on two levels, distinguishing between single aggressive acts and highly severe aggressive acts (e.g., violence targeted someone on the ground). We had, however, to collapse these two levels level, because the highly severe aggression level completely separated (perfectly predicted) the victimization outcome, an estimation issue of logistic regression (i.e., no victimization cases were linked with highly severe aggression).

we standardized it prior to analysis by subtracting the mean and dividing by two standard deviations (Gelman, 2008). The rationale for including age as a control is that research documents a robust relationship between young age and the risk of violent victimization (Macmillan, 2001). As a third control variable, we measured the total number of bystanders present at the scene ($\alpha = .87$). This variable, also standardized prior to analysis, was included given research suggesting that bystander presence can play a role in both regulating or amplifying conflict situations (Collins, 2008; Levine et al., 2011). Finally, *aggressive involvement* measured whether the bystander performed any physically escalatory behaviors (e.g., aggressive pointing, hitting) besides de-escalatory acts ($\alpha = .70$). Note that for victimized bystanders, we only included aggressive acts performed prior to their victimization, because we are only interested in how the aggressive acts may be a cause (not a retaliatory effect) of victimization.

Results

Of the 337 bystanders who intervened with only de-escalatory acts during the conflict (i.e., excluding bystanders performing any aggressive acts), 12 (3.6%) were physically victimized ($M = 0.036$, CI 95% [0.019, 0.061]). Across 602 bystanders who were present but remained uninvolved, only one (0.2%) bystander was victimized ($M = 0.002$, CI 95% [0.000, 0.009]). This victimization difference pure de-escalators and non-involved bystanders was statistically significant, $z = -4.18$, $p < .001$, two-tailed. The victimizations of interveners were mainly a single hit ($n = 6$, 50.0%) or several shoves ($n = 5$, 41.7%), while more excessively violent victimizations such as head-butts, violence towards a bystander on the ground, or weapon use were not found in data.

Confirmatory analysis

For our confirmatory hypothesis testing, we applied a logistic regression model, specified with cluster-corrected standard errors to account for our hierarchical data structure, with bystanders nested in conflicts (Huang, 2016). We adjusted the applied alpha level to the size of the analyzed sample (Lakens, 2019), and the number of hypotheses tested (Bland & Altman, 1995).⁴ For our directional hypotheses, we report one-tailed tests and one-sided 95% confidence intervals (Cho & Abe, 2013). Alongside frequentist statistics, we also report Bayes factors (Wagenmakers, 2007), approximated from Bayesian Information Criterion (an approach that does not require specification of a prior, rather a unit-information prior is assumed).

In line with Hypothesis 1, we found that conflict party affiliation was positively associated with bystander victimization (OR = 5.65, 95% CI [1.89, 16.86], $p = .005$, one-sided, $BF_{10} = 3.33$). The statistical significance of this association was suggested by the reported p -value

⁴ Specifically, we calculated a sample standardized alpha, $\alpha_{stan} = \frac{\alpha}{\sqrt{\frac{N}{100}}}$, which we then Bonferroni corrected for the number of main hypotheses (m) tested in the paper $\frac{\alpha_{stan}}{m}$.

below the adjusted alpha threshold of .008, and by the Bayes factor indicating that data was approximately three times more likely under H_1 than H_0 —what should be considered positive evidence evaluated with the thresholds (Raftery, 1995): weak (~1-3), positive (~3-20), strong (~ 20-150), and very strong (150 or larger). In terms of practical significance, conflict party affiliated bystanders had three times the odds of being victimized—a large effect size evaluated with the thresholds (Ferguson, 2009): weak (~ 2.0), moderate (~ 3.0), and strong (~ 4.0). Assessed in absolute terms, however, the predicted probabilities of victimization for conflict party affiliated (5%) and non-affiliated (2%) bystanders, respectively, suggested that the risk was low for both groups.⁵

Next, a statistically significant p -value offered support for Hypothesis 2 that male gender was positively associated with victimization (OR = 3.82, 95% CI [1.68, 8.67], $p = .004$, one-sided, $BF_{10} = 1.54$). However, the related Bayes factor offered only weak evidence in favor of the association, suggesting that data cannot discriminate between H_0 and H_1 . Finally, contrary Hypothesis 3, bystanders were not found to have a higher victimization risk when intervening into nonviolent aggression (OR = 1.93, 95% CI [0.70, 5.29], $p = .143$, one-sided) or violent aggression (OR = 1.22, 95% CI [0.41, 3.65], $p = .383$, one-sided) in comparison to conflicts where there were no immediate observable displays of aggression. This was further stressed by Bayes factor evidence suggesting that it was around five times more likely that the conflict severity was non-associated with bystander victimization than associated ($BF_{01} = 4.79$).

Several of the controls were not found to be associated with the outcome: age (OR = 0.56, 95% CI [0.12, 2.60], $p = .460$, two-sided, $BF_{01} = 16.61$), number of bystanders (OR = 1.39, 95% CI [0.44, 4.47], $p = .576$, two-sided, $BF_{01} = 17.54$), and aggressive involvement (OR = 1.15, 95% CI [0.34, 3.90], $p = .826$, two-sided, $BF_{01} = 19.49$). By comparison, the victimization risk was found to vary across urban-national contexts, as indicated by positive Bayes factor evidence ($BF_{10} = 5.23$). Specifically, we found that the victimization risk was higher in Lancaster (the United Kingdom) as compared to the Amsterdam (the Netherlands) reference (OR = 13.21, $p = .005$, two-sided), while Cape Town (South Africa) was not found to be statistically different from the reference (OR = 2.96, $p = .301$, two-sided).

Multiverse analysis

In addition to the initial confirmatory analysis, we also assessed how robust the hypotheses variables were across a multiverse of all combinations of independent variables (Stegen, Tuerlinckx, Gelman, & Vanpaemel, 2016; Young, 2018). The rationale for this is that statistical analyses often hinge upon arbitrary specification choices (Silberzahn et al., 2017)—but as Leamer (1985, p. 308) highlights, a “fragile inference is not worth taking seriously.” For this

⁵ Note that the predicted probability for nonaffiliated is not significantly different from zero (see osf.io/vyutj). As such, the accuracy of the 2% prediction should be interpreted with caution and effectively indistinguishable from a victimization likelihood of zero.

multiverse analysis procedure, we followed the rule of thumb that sign and significance stability across 50% and 95% of specifications indicate weak and very strong model robustness, respectively (Raftery, 1995).

Across 64 model combinations of independent variables (two-tailed, $\alpha = .05$), conflict party affiliation was statistically significant with a positive sign in 100% of models. Alternatively, we examined the implication of applying an alternative operationalization of affiliation, defined more narrowly as an affiliation to a person who had been aggressed against at the point of intervention—this victim affiliation measurement corresponds to the one applied in Liebst, Heinskou, et al. (2018). This alternative predictor had a positive sign in 100% of models, but only reached statistical significance in 28% of combinations. In sum, across all models specified with one of the two measures of affiliation (i.e., conflict party affiliation or victim affiliation), 64% are positive and significant, suggesting that the hypothesized association of affiliation and victimization hinges in part on how affiliation is measured and should thus be considered only moderately robust. Next, the bystander's gender had a positive sign in 100% of models, while 50% of models reached statistical significance. Finally, the multiverse analysis supports the initial results that severity level is not associated with the victimization outcome. Specifically, bystanders were not found to have a higher victimization risk when intervening into nonviolent aggression or violent aggression (compared to situations with no direct displays of aggression) in any (0%) of the model combinations.

Discussion

The growing evidence suggesting that bystanders witnessing public aggression are often 'active' rather than 'passive' further raises the question of whether it is safe for bystanders to intervene to help. The key finding of the current study is that the prevalence of victimization for both intervening and non-intervening bystanders is small. Specifically, 0.2% of inactive bystanders and 3.6% of actively intervening bystanders were physically victimized, and these victimizations tended to be a single hit, or multiple shoves. This result is overall consistent with prior studies showing that mere presence in conflictual events is rarely dangerous (Planty, 2002; Sherman et al., 1989), and that the victimization risk is limited even for intervening bystanders (Liebst, Heinskou, et al., 2018). More broadly, the low bystander victimization rate agrees with evidence showing that interpersonal violence is relatively rare and only atypically escalates to severe levels—either because violence is hard to perform (Collins, 2019) or due to successful bystander de-escalation (Levine et al., 2011).

It should, however, be noted that the victimization prevalence in the current data is at the lower end when compared to the existing evidence. This between-study variation may be the result of differing victimization operationalizations (e.g., 'physically victimized' versus 'harmed or threatened'), or reflect the diversity of the aggressive situations under examination (e.g., public confrontations as opposed to sexual assaults). The discrepancy in the levels of

victimization between studies may also be an artifact of the applied methods (e.g., real-time observational data instead of retrospective self-reported accounts), or a property of the datasets under examination (Lindgaard & Bernasco, 2018; Philpot, Liebst, Møller, et al., 2019). For example, the higher victimization rate at just over one in ten recorded in Liebst, Heinskou, et al. (2018) is most probably an artifact of relying on serious police-reported violent assaults (as opposed to the range of non-physical and physical confrontations in the current study).

We tested three plausible risk factors of victimization. As expected, we found some, albeit only moderately robust, evidence that conflict party affiliation increases the victimization likelihood (Hypothesis 1). In interpreting this finding, victimization risk could be attributed to out-group or in-group aggression processes (Brewer, 2001; Marques, 1990). Based on a positive association between victim affiliation and bystander victimization, Liebst, Heinskou, et al. (2018) suggested an aggressive out-group dynamic, whereby bystanders socially close to a conflict party are perceived as partisans, and thus eligible targets of aggression. In the current study, however, the association between victimization risk and victim affiliation was highly fragile, and offers thus no clear evidence for out-group aggression. Alternatively, and indicating an aggressive in-group dynamic, a proportion of the bystander victimizations were committed by an apparently familiar person. However, given that these considerations rely on the distribution of very few cases, and that our two measures of affiliation have low interrater agreements, the current study is uncertain about the relative contribution of in- and out-group aggression. A less ambiguous interpretation is that public strangers may intervene with a particularly low victimization risk, likely reflecting their limited social investment in the conflict. This argument resonates with Black's (1993) view that third-parties with a weak partisanship with conflict parties tend to limit conflict outcomes.

Given the rather mixed picture of the potential role being played by group processes in the likelihood of bystander victimization, further exploratory research is required. One of the key questions to be explored is the way bystanders, victims, and perpetrators are making sense of their own psychological relationships to others as the events unfold. There is a body of more qualitative work in the social identity tradition that examines these dynamics in the context of crowd behavior in conflict and emergencies (see e.g., Drury & Reicher, 2000; Stott & Reicher, 1998). Combining analysis of the group dynamics from the perspective of participants with information about behavioral dynamics in real-time will improve our understanding in this area.

At first sight in line with Hypothesis 2, the analysis indicated that male bystanders may have a higher victimization risk than females. However, the very highly nature of the association suggests that the current study adds further to, rather than clarifies, the mixed status of the existing evidence regarding the gender-victimization link. Finally, contrary to Hypothesis 3, we found compelling evidence that conflict severity is not associated with victimization. This finding corresponds with Liebst, Heinskou, et al. (2018) and contradicts Hamby et al. (2016),

and more broadly questions the view that violence often gets out of control and thus may spill over to intervening or even uninvolved bystanders (Collins, 2013). The targeted rather than the uncontrolled display of violence (see Kemper, 2011; Liebst, Lindegaard, & Bernasco, 2019) is further emphasized by the descriptive finding that non-intervening bystanders are virtually never victimized. Taken together, our risk factor analysis only offered unambiguous evidence that conflict severity is non-associated with bystander victimization, while there is ground for caution when interpreting the moderately robust and highly fragile results of affiliation and gender, respectively.

We comment briefly on the included controls. Age was not associated with victimization, indicating that the well-described link between young age and violent victimizations may not generalize to bystander victimizations (Macmillan, 2001). The number of bystanders present was also not found to be associated with victimization and runs counter to the view—both advocated within bystander research (Latané, 1981) and beyond (Collins, 2008)—that mere co-presence of an audience may shape violent outcomes. Contrary to van Reemst et al. (2015) and Liebst, Heinskou, et al. (2018), we did not find that aggressive involvement was a risk factor, indicating that a bystander’s de-escalatory intervention may ‘make up’ for the aggression also performed in the conflict.

Data indicated that interventions in conflicts in Lancaster (the United Kingdom) had a substantially higher victimization rate than the Amsterdam (the Netherlands) reference category. This reflects that the victimizations are unevenly distributed across the subsamples, with 18 out of the 26 (69.2%) victimization events recorded in Lancaster, and only six (23.1%) and two (7.7%) in Amsterdam and Cape Town, respectively. It is perhaps surprising that the Cape Town subsample does not have disproportionately more victimization events, given the high violent crime rates of this urban-national contexts compared to its European counterparts (Global Peace Index, 2018). The victimization skewness could, however, also be attributed to other more specific circumstances, for example, different norms for displaying aggression in night-time drinking settings (Graham & Wells, 2003). Further research should examine how bystander victimization rate varies within and between national, regional, and urban contexts (see Sampson, 2013).

Limitations

The current study has limitations. We acknowledge that our sample is not representative of public place conflicts or interpersonal aggression as such. This is stressed by prior research showing that the victimization risk varies substantially across conflict types (Hamby et al., 2016), as well as studies indicating that conflicts in inner-city public spaces differ from suburban and rural settings (Baumgartner, 1989). Specifically, our sample is constrained by where in public place the cameras are located, which is chiefly in night-time drinking settings. This implies, in turn, that the current sample contains a disproportionately high number of intoxicated

individuals, which may have influenced how the violent events unfolded and the related bystander victimization risk.

It should be acknowledged that CCTV footage does only offer information on a limited aspect of the unfolding conflicts. Video data captures the here-and-now manifestation of conflicts while only very limited information on subjective motivations and chains of interactions that lead up to the situation (Wieviorka, 2014). Moreover, our video clips did not capture sound and thus we could not assess the relative risk associated with the common occurrence of verbal intervention (see Slater et al., 2013). It is likely to have underestimated the victimization prevalence that we could not record verbal victimizations (e.g., threats, scoldings).

We also faced issues with respect to the reliability of our codes. For example, while we reliably captured the non-event of victimization, we were less reliable in capturing the rare events of victimization, and this may have resulted in an underestimation of the number of victimizations across data. Moreover, we only assessed the interrater agreement but not the construct validity of the included variables. As such, in measuring the bystander affiliation to conflict parties, we relied on behavioral cues of affiliation, without definitively knowing who knows whom. This uncertainty, along with the low interrater agreement reached for this variable, may have contributed to the only moderately robust association between affiliation and victimization.

Research implications

Within recent years, the situational approach to the study of violence has regained traction (Bowman, Whitehead, & Raymond, 2018; Collins, 2008; Lindegaard, Bernasco, & Jacques, 2015). Informed by this development, the bystander research field has received criticism for its overreliance on experimental methods with a low ecological validity to assess the causes and consequences of bystander helping behavior (Philpot, Liebst, Levine, et al., 2019). Just as the standard experimental simulations of mundane mishaps (e.g., an ‘accidental’ pencil spill) underestimate the extent to which real-life bystanders intervene when help is required (Fischer et al., 2011; Philpot, Liebst, Levine, et al., 2019), the absence of unstructured violence precludes an assessment of the victimization risk for intervening bystanders (Osswald et al., 2010).

The current paper encapsulates the recommendation that video-based naturalistic observation offers an ecologically valid alternative approach to study real-life bystander behavior (Philpot, Liebst, Møller, et al., 2019). This taps into the wider consideration that social psychology, as stressed by Mortensen and Cialdini (2010), should to a larger extent validate its experimental insights with naturalistic observations. The advantage of conducting video-assisted naturalistic observation is furthermore stressed by the fact that other typical methods—self-reports (interviews, surveys) and on-site observations—are known to provide unreliable information on face-to-face interaction sequences (Jerolmack & Khan, 2014; Morrison, Lee, Gruenewald, & Mair, 2016).

Practical implications

Bystanders are increasingly appreciated as a crime prevention resource (Eck, 2015), especially as a successful means to target sexual violence on campuses (Kettrey & Marx, 2019). Yet, given the lack of systematic evidence examining the risks posed to intervening bystanders, it is understandable that crime prevention stakeholders may be reluctant to encourage bystanders to take action (Liebst, Philpot, Heinskou, & Lindegaard, 2018). We believe, however, that our results add to the proactive bystander view that the helpful benefits of intervention outweigh the relatively limited costs associated with bystander intervention (Hamby et al., 2016; Hart & Miethe, 2008; Levine, Philpot, & Kovalenko, 2020). However, further evaluations of the benefits against the potential costs should also consider the long-term consequences of intervention for the bystanders. These may include traumatic stress symptoms, but also positive affect related to helping others, particularly in comparison to the negative feelings experienced by those who did not intervene (Witte, Casper, Hackman, & Mulla, 2017).

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.
- Baumgartner, M. P. (1989). *The moral order of a suburb*: Oxford University Press.
- Belknap, J., & Shelley, J. K. (1993). The new lone ranger: Policewomen on patrol. *Am. J. Police*, 12, 47.
- Black, D. (1993). *The social structure of right and wrong*. San Diego: Academic Press.
- Bland, J. M., & Altman, D. G. (1995). Multiple significance tests: the Bonferroni method. *BMJ*, 310(6973), 170.
- Bowman, B., Whitehead, K. A., & Raymond, G. (2018). Situational factors and mechanisms in pathways to violence. *Psychology of violence*, 8(3), 287.
- Brewer, M. B. (2001). Ingroup Identification and Intergroup Conflict: When Does Ingroup Love Become Outgroup Hate? In R. Ashmore, L. Jussim, & D. Wilder (Eds.), *Social Identity, Intergroup Conflict, and Conflict Reduction* (pp. 17-41). Oxford: Oxford University Press.
- Cho, H.-C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate? *Journal of Business Research*, 66(9), 1261-1266.
- Christie, N. (1986). The ideal victim. In *From crime policy to victim policy* (pp. 17-30): Springer.
- Cicchetti, D. V., & Feinstein, A. R. (1990). High agreement but low kappa: II. Resolving the paradoxes. *Journal of clinical epidemiology*, 43(6), 551-558.
- Collins, R. (2008). *Violence: A micro-sociological theory*. Princeton: Princeton University Press.
- Collins, R. (2013). Entering and leaving the tunnel of violence: Micro-sociological dynamics of emotional entrainment in violent interactions. *Current Sociology*, 61(2), 132-151.

- Collins, R. (2019). Preventing Violence: Insights from Micro-Sociology. *Preventing Violence: Insights from Micro-Sociology*, 48(5), 487-494.
- 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
- Cumming, G. (2014). The new statistics: Why and how. *Psychological science*, 25(1), 7-29.
- 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
- Drury, J., & Reicher, S. (2000). Collective action and psychological change: The emergence of new social identities. *British Journal of Social Psychology*, 39(4), 579-604.
- Eck, J. E. (2015). Who should prevent crime at places? The advantages of regulating place managers and challenges to police services. *Policing: A Journal of Policy and Practice*, 9(3), 223-233.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2), 175-191.
- Ferguson, C. J. (2009). An effect size primer: a guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532-538.
- 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
- 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
- Gilmore, R. O., & Adolph, K. E. (2017). Video can make behavioural science more reproducible. *Nature Human Behavior*, 1(7). doi:10.1038/s41562-017-0128
- Goffman, E. (1971). *Relations in public. Microstudies of the public order*. New York: Basic Books.
- Graham, K., & Wells, S. (2003). 'Somebody's gonna get their head kicked in tonight!' Aggression among young males in bars—a question of values? *British Journal of Criminology*, 43(3), 546-566.
- Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61(1), 29-48.
- Hadfield, P., Lister, S., & Traynor, P. (2009). 'This town's a different town today': Policing and regulating the night-time economy. *Criminology & Criminal Justice*, 9(4), 465-485.
- Hamby, S., Weber, M. C., Grych, J., & Banyard, V. (2016). What difference do bystanders make? The association of bystander involvement with victim outcomes in a community sample. *Psychology of violence*, 6(1), 91-102.

- Hart, T. C., & Miethé, T. D. (2008). Exploring bystander presence and intervention in nonfatal violent victimization: When does helping really help? *Violence and Victims, 23*(5), 637-651.
- Huang, F. L. (2016). Alternatives to multilevel modeling for the analysis of clustered data. *The Journal of Experimental Education, 84*(1), 175-196.
- Index, G. P. (2018). *Measuring peace in a complex world*. Retrieved from
- Jerolmack, C., & Khan, S. (2014). Talk is cheap: Ethnography and the attitudinal fallacy. *Sociological methods & research, 43*(2), 178-209.
- 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
- Kemper, T. D. (2011). *Status, power and ritual interaction: A relational reading of Durkheim, Goffman and Collins*. London: Routledge.
- Kettrey, H. H., & Marx, R. A. (2019). The effects of bystander programs on the prevention of sexual assault across the college years: a systematic review and meta-analysis. *Journal of youth and adolescence, 48*(2), 212-227.
- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement, 56*(5), 746-759. doi:doi.org/10.1177/0013164496056005002
- Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research, 30*(3), 411-433.
- Krippendorff, K. (2013). Commentary: A dissenting view on so-called paradoxes of reliability coefficients. *Annals of the International Communication Association, 36*(1), 481-499.
- Lakens, D. (2019). Justify Your Alpha: A Practical Guide. *osf.io/h5e2q*.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics, 159*-174.
- Latané, B. (1981). The psychology of social impact. *American Psychologist, 36*(4), 343-356. doi:10.1037/0003-066X.36.4.343
- Leamer, E. E. (1985). Sensitivity analyses would help. *The American Economic Review, 75*(3), 308-313.
- 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., Philpot, R., & Kovalenko, A. G. (2020). Rethinking the Bystander Effect in Violence Reduction Training Programs. *Social Issues and Policy Review*.
- 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
- Liebst, L. S., Ejbye-Ernst, P., Philpot, R., Bruvik Heinskou, M., Demant, J., Lykke Dausel, K., & Rosenkrantz Lindegaard, M. (2019). Intervention or Involvement: A Video Observational Analysis of Bouncers in Aggressive Encounters. *Deviant Behavior, 1*-10.

- 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
- Liebst, L. S., Lindegaard, M. R., & Bernasco, W. (2019). Dissecting the Role of Dominance in Robberies : An Analysis and Implications for Micro-Sociology of Violence. *Journal of interpersonal violence*.
- 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 predict bystander intervention: Evidence from violent incidents captured on CCTV. *Aggressive Behavior*.
- Liebst, L. S., Philpot, R., Heinskou, M. B., & Lindegaard, M. R. (2018). Bystander Intervention in Street Violence: Current Evidence and Implications for Practice. *Samfundsøkonomen*, 4, 11-15.
- 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., Bernasco, W., & Jacques, S. (2015). Consequences of expected and observed victim resistance for offender violence during robbery events. *Journal of Research in Crime and Delinquency*, 52(1), 32-61.
- Lorenz, K. Z. (1973). The fashionable fallacy of dispensing with description. *Naturwissenschaften*, 60(1), 1-9.
- Macmillan, R. (2001). Violence and the life course: The consequences of victimization for personal and social development. *Annual Review of Sociology*, 27(1), 1-22.
- Marques, J. M. (1990). The black sheep effect: Outgroup homogeneity in social comparison settings. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances*. London: Harvester Wheatsheaf.
- Milgram, S. (1970). The experience of living in cities. *Science*, 167(3924), 1461-1468.
- Morrison, C., Lee, J. P., Gruenewald, P. J., & Mair, C. (2016). The reliability of naturalistic observations of social, physical and economic environments of bars. *Addiction Research & Theory*, 24(4), 330-340. doi:10.3109/16066359.2016.1145674
- 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
- Moschella, E. A., Bennett, S., & Banyard, V. L. (2018). Beyond the situational model: Bystander action consequences to intervening in situations involving sexual violence. *Journal of interpersonal violence*, 33(20), 3211-3231.
- 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.

- Philpot, R., Liebst, L. S., Levine, M., Bernasco, W., & Lindegaard, M. R. (2019). Would I be Helped? Cross-National CCTV Footage Shows That Intervention Is the Norm in Public Conflicts. *American Psychologist*, 75(1), 66–75.
- Philpot, R., Liebst, L. S., Møller, K. K., Lindegaard, M. R., & Levine, M. (2019). Capturing violence in the night-time economy: A review of established and emerging methodologies. *Aggression and Violent Behavior*, 46, 56-65.
- Planty, M. (2002). *Third-party involvement in violent crime, 1993-1999 (Bureau of Justice Statistics Special Report)*. Washington, DC: Department of Justice.
- Rabe-Hemp, C. E., & Schuck, A. M. (2007). Violence against police officers: Are female officers at greater risk? *Police Quarterly*, 10(4), 411-428.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological methodology*, 25, 111-164.
- Sampson, R. J. (2013). The place of context: a theory and strategy for criminology's hard problems. *Criminology*, 51(1), 1-31.
- Sherman, L. W., Steele, L., Laufersweiler, D., Hoffer, N., & Julian, S. A. (1989). Stray bullets and “mushrooms”: Random shootings of bystanders in four cities, 1977–1988. *Journal of Quantitative Criminology*, 5(4), 297-316.
- Silberzahn, R., Uhlmann, E. L., Martin, D., Anselmi, P., Aust, F., Awtrey, E. C., . . . Bonnier, E. (2017). Many analysts, one dataset: Making transparent how variations in analytical choices affect results.
- 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
- Stalder, D. R. (2008). Revisiting the issue of safety in numbers: The likelihood of receiving help from a group. *Social Influence*, 3(1), 24-33. doi:10.1080/15534510701766181
- Stegen, 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
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation*, 9(1), 4.
- Stott, C., & Reicher, S. (1998). How Conflict Escalates: The Inter-Group Dynamics of Collective Football CrowdViolence'. *Sociology*, 32(2), 353-377.
- Valach, L., von Cranach, M., & Kalbermatten, U. (1988). Social meaning in the observation of goal directed action. *Semiotica*, 71(3-4), 243-260.
- van Reemst, L., Fischer, T. F., & Zwirs, B. W. (2015). Response decision, emotions, and victimization of police officers. *European Journal of Criminology*, 12(6), 635-657.
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic bulletin & review*, 14(5), 779-804. doi:10.3758/BF03194105

- Wieviorka, M. (2014). The sociological analysis of violence: new perspectives. *The Sociological Review*, 62(2_suppl), 50-64.
- Witte, T. H., Casper, D. M., Hackman, C. L., & Mulla, M. M. (2017). Bystander interventions for sexual assault and dating violence on college campuses: Are we putting bystanders in harm's way? *Journal of American college health*, 65(3), 149-157.
- Young, C. (2018). Model uncertainty and the crisis in science. *Socius*, 4, 2378023117737206.