Informal Social Control: A Randomized Experiment

Christoph Engel & Daniel Nagin

Christoph Engel: Max Planck Institute for Research on Collective Goods, Bonn (Germany)
ORCID 0000-0001-8513-8532

Daniel Nagin: Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA (corresponding author)
ORCID 0000-00030-0075-1851

email: dn03@andrew.cmu.edu

Abstract

Criminological theory is replete with examples of how informal social controls at the community- and micro-levels can deter would-be criminals. Much of the empirical research in this area, however, has focused on community-level pathways rather than testing the micro-level pathways through which informal social control can deter would-be offenders from engaging in criminal activity. We complement this literature by studying micro-level interaction between community members and would-be offenders with the aim of isolating the circumstances under which private citizens can be effective in deterring the would-be offender, and what are the characteristics and motivations of citizens who are willing to play this role? With the help of a lab experiment we can make the socially most important effect visible: the risk that a community member might intervene on behalf of the victim induces a person to refrain from stealing.

1. Introduction

Ideally, crime and disorder are not prosecuted; they are preempted. Yet the police cannot (and should not) be everywhere. Sources of social control, however, extend well beyond the police, including not only police surrogates, such as security guards and surveillance technologies, but also informal sources of social control involving monitoring activities by community members.

The longstanding sociological literature on informal social control takes as its starting point that such monitoring, which can take the form of organized activities, for example block watches, or informal surveillance, is an important supplementary source of crime control that disrupts criminal events or prevents them from occurring altogether (Sampson 2006). This macro-level literature investigates community characteristics that are postulated to facilitate (e.g., homeownership) or detract from (e.g., high rates of residential turnover) a community's capacity to exert effective social control of crime and disorder. Older examples of this research include Shaw and McKay (1942) and Whyte (1947). Two more recent exemplars are the theory of collective efficacy (CE) (Sampson, Raudenbush et al. 1997)and routine activity theory (RAT) (Cohen and Felson 1979).

The concept of guardianship plays a central role in both RAT and CE, but in distinctly different ways. According to Cohen and Felson (1979), capable guardians are individuals or inanimate objects (e.g. surveillance cameras) whose presence may deter a motivated offender from victimizing a vulnerable target by making perpetration of the crime more difficult or risky. The presence of a guardian "serves as a gentle reminder that someone is looking" (Felson and Boba 2010, 28). In CE guardians play a more active role. As described by Sampson (Sampson 2006, 153, emphasis added): "The term collective efficacy is meant to signify an emphasis on shared

beliefs in a neighborhood's capability for action to achieve an intended effect, *coupled with an active sense of engagement on the part of residents*." Such active engagement can come in many forms—some are reactive to a particular incident such as calling out unruly adolescents while others may require coordinated proactive action, for example, to disrupt illegal sex and drug markets.

Our focus here is on the more active form of guardianship envisioned by Sampson and by much of the research on informal social control. Yet much of this research, including that on CE, focuses on the relationship of community-level characteristics to crime and disorder in the community. In so doing it leaves in the background the form of the micro-level interaction between community members and would-be offenders that informal social control theory holds to be instrumental in keeping the risk of crime in check. Specifically, it does not address the circumstances under which private citizens can be effective in deterring would-be offenders, and the characteristics and motivations of citizens who are willing to play this role.

Anderson (1999)'s classic *Code of the Street Decency,* makes clear that the effectiveness of informal social control is highly dependent on both macro-level community characteristics and micro-level factors including the motivations and capacities of community members. Anderson's introductory chapter describes a journey down Germantown Avenue, a major artery in Philadelphia, PA. In Anderson's telling we begin in Chestnut Hill, a prosperous, mostly white but not totally so, community in which crime and violence is infrequent and perpetrated by outsiders. Community members do not fear victimization by other community members - not because of informal control but because the community is entirely populated by what Anderson calls "decent" people who do not have the criminal propensities that effective informal social control aims to keep in check. In Anderson's journey down Germantown Avenue communities become

progressively poorer, with a growing presence of what Anderson calls "street" people, usually young males, who are prone to crime and violence but to varying degrees are kept in check by "decent old heads." The journey ends in Germantown, a desperately poor, drug ridden entirely black ghetto in which the forces of informal social control are overwhelmed by the dominance of street people conducting themselves according to the code of the street.

Drawing on three lines of micro-level research—one on bystander intervention into ongoing criminal events, another on the provision of public goods from the law and economics literature, and third on the willingness of individuals to take personally costly action to punish wrong doing—we constructed an online laboratory experiment designed to help in understanding the conditions under which Anderson's "decent" private citizens acting as guardians taking personally costly action can exert informal social control that is effectively deterring would-be offenders. We do indeed find evidence that private citizen guardians can deter would-be offenders but only if they have capacity to make to make offending unprofitable.

2. The Three Literatures

Bystander intervention research examines the circumstances under which private citizens intervene to disrupt an ongoing criminal event. The main sources of evidence of this research are social psychology experiments investigating how the number of bystanders affects their individual willingness to intervene. For a recent review of this literature see Fischer, Krueger et al. (2011). A second related literature involves post-hoc reviews of CCTV recordings of criminal incidents (cf. Reynald 2009, Lindegaard, Liebst et al. 2017, Liebst, Philpot et al. 2021).This research focuses on when and how bystanders intervene. While both these bystander intervention literatures focus on the decision-making process of the would-be guardians (Nagin, Herman et al. 2023, Barnum, Herman et al. 2024, Jin, Wu et al. 2024), bystander intervention research by its

design has one important limitation—it only examines bystander reactions to an ongoing criminal event. As such, it cannot address the circumstances under which community members are effective in preventing crime altogether.

The two other research lines provide a basis for addressing this limitation. Public goods research from the law and economics literature explores the conditions under which community members are willing to make voluntary contributions to the maintenance of a public good, that is a good that is freely available to all community members. The focus of this literature is overcoming the free-rider problem—community members have an incentive not to make voluntary contributions themselves to the public good's maintenance while still benefitting from the contributions of other community members. For a review see Chaudhuri (2011). Public safety is a textbook public good: every member of the community benefits from vigilance and intervention, but only those community members on the guard bear the safety and time costs of serving in this role. Thus, at the individual-level there is an incentive to free ride off the guardianship contributions of others community members willing to bear the costs of guardianship.

What then might induce community members to take on voluntarily the costs of guardianship? One mechanism for deterring free riding is to allow other community members to punish free riders. Results show that, if the capacity to punish is sufficient, free riding can be effectively deterred (Fehr and Gächter 2000). Remarkably those who engage in costly punishment pay twice: by their own contribution to the first order public good (they behave responsibly themselves), and by containing the risk that others freeride, to the benefit of the entire community (i.e. by a contribution to a second order public good) (Yamagishi 1986, Heckathorn 1989).

However, in the context of our focus here, guardianship activities that deter crime, punishment of free riders by other community members is not feasible. Some other mechanism is required.

Research on the willingness of individuals to take personally costly actions to punish others for their selfish behavior, for instance if they are too miserly in their sharing with others, suggests an intrinsic mechanism. This research finds that sizeable fractions of the population are willing to punish wrongdoing even if punishment is personally costly, and if they derive no personal benefit from intervention whatsoever (Henrich, McElreath et al. 2006). Willingness to exact such punishment increases with would-be punishers' own orientation to share with others, which we henceforth refer to as their social value orientation (SVO).

3. Experimental Design

We build on these three literatures to design an experiment that investigates the circumstances under which the presence of would-be guardians contains the incidence of crime, allowing for heterogeneity of perceptions and motives among both would-be guardians and would-be criminals.

Participants were recruited from the UK-based online platform Prolific which has a large, worldwide participant base. Prolific was used for two reasons. One is that has functionality related to recruiting participants in batches that makes it an attractive and widely used platform for conducting online experiments such as ours. A second relates to data quality. Studies by Peer, Brandimarte et al. (2017) and Chandler, Rosenzweig et al. (2019) conclude that Prolific samples had higher data quality compared to MTurk and online panels such as Qualtrics and Dynata.

Participants are randomly assigned to and informed of one of three roles, designated as A, B, or C. In compensation for their participation, all participants receive a 6£ endowment. Those assigned the A role, would-be thieves, are given the option of taking 3£ of B's endowment. Those

assigned the C role are the would-be agents of informal social control. In their would-be guardian role, they are given the option of punishing A for taking from B, by inflicting a fine on A. The amount contributed by C toward that fine is deducted from their 6£ endowment, thereby making it costly for them to intervene on behalf of B. B participants, the would-be victims, are passive. We manipulate (a) the presence of citizen guardians and their number; (b) the cost of punishment for the punisher and (c) the maximum punishment power.

From the perspective of those in the C role, the experiment is intended to mimic the dilemma confronting real world would-be agents of informal social control. Unlike a police officer who has a duty to intervene if they encounter a problematic event that might escalate into a crime, would-be guardians, who are not so obligated, must balance their sense of duty, if any, to contribute to the safety of their community against the potential cost to themselves (Barnum, Herman et al. 2024, Jin, Wu et al. 2024). In the real world that cost is mainly personal safety, something that cannot be ethically manipulated in an experiment. Instead in the experiment the cost is monetary—a reduced payout for serving as a guardian who punishes A were A to take. In recompense to the external validity loss resulting from this abstraction, the design makes it possible to investigate what is invisible in research on reactive guardianship: the conditions under which knowing of potential guardian intervention deters would-be perpetrators.

Participants were also randomly assigned to one of six experimental conditions summarized in Table 1. Following assignment to role and condition, participants were explained the implications of the role and condition assignments for the payouts for participants, in all roles not just their own. To insure they understood the payout conditions, participants were quizzed on the financial implications of various actions by the A and C participants[1].

---

[1] For details, please see the instructions in the Appendix.

The conditions were designed to vary the maximum punishment that would-be guardians can impose on would-be thieves and the cost to the would-be guardians of imposing a punishment at or less than the maximum allowed through, what we call the fine-to-fee ratio.

Condition	Punishment Maximum (£)	Fine to Fee Ratio	Number of C Role Participants	Sample Size
1	0	0	0	57
2	3	1	1	115
3	6	1	1	110
4	6	2	1	58
5	9	3	1	57
6	9	3	2	50

Table 1
Experimental Conditions

Two conditions served as baselines. In condition 1 there was no C role participant. In this condition, A could take 3£ from B with no risk of punishment. The purpose of this condition was to explore the inhibiting effect of ethical commitments not to harm others for personal gain in the extreme circumstance in which there is no chance of punishment for inflicting harm. Zimring and Hawkins (1973) in their classic book Deterrence: The Legal Threat in Crime Control make the important point that if an individual is committed to refraining from acts prescribed as criminal as a matter personal principle, the threat of punishment or the absence thereof will not affect their behavior because they are already committed to conformance. Pogarsky (2002) uses the term "acutely" conformist to describe such individuals. In condition 2, the C role participant can exact

a maximum punishment of 3£s for taking from B.[2] Thus, in this condition taking is at worst a break-even action—A takes 3£s from B but C punishes her taking with the maximum allowable fine of 3£s. In conditions 3-5 C's capacity to punish is increased by increasing the maximum fine that can be exacted to 6£ in conditions 3 and 4 and to 9£ in condition 5. These conditions are intended to make taking potentially unprofitable with the 9£ maximum potentially resulting in A losing the entire 6£ initial endowment.[3]

The fine to fee ratio determines the cost of punishment to C. In condition 3 each £ of punishment by C translates into only a 1£ fine, whereas in conditions 4 and 5 each £ of punishment by C translates into a 2£ fine and 3£ fine, respectively, were A to take from B. Thus, for example, a 3£ fine will cost C 3£s in condition 3 but only 1.5£s in condition 4, and 1£ in condition 5. Our purpose in varying the fine-to-fee ratio was not only to test its impact on the punishment decisions of C participants but also the taking decisions of A participants. Specifically, we anticipated that increases in the ratio would increase the amount of punishment exacted by C participants and that A participants would anticipate this impact and therefore the taking rate would be a declining function of the ratio.

In the final condition, the number of participants assigned the C role is increased to two. As with condition 5 the punishment maximum is 9£ and the fine-to-fee ratio is 3. The purpose of this condition is to test for the so-called bystander effect (Fischer, Krueger et al. 2011). Each C

---

[2] Originally, sample size was about even across conditions. It turned out that the effect of increasing maximum punishment power from 3£ to 6£ on taking was only marginally significant. To find out whether this seeming effect was spurious, we doubled up the observations on these two conditions. The original design and this extension are preregistered at https://osf.io/gx4sh/?view_only=0cf468a54ccf4f87a288219e06e6b28a. In the Appendix, we report means of all dependent and independent variables, separately for the first and second wave of data collection. It turns out that both waves have been well balanced.

[3] While in principle it would be interesting to test the impact of a fine threat greater than 9£ in which A participants would potentially have to dip into their own resources to pay the fine, such a condition would not be an ethically acceptable manipulation and in any event, would probably be impossible to enforce.

player might want to pass off in whole or part to the cost of disciplining the would-be thief to the other C player, or might expect the other C player to inflict enough punishment, so that further intervention would be overkill. We also use these two conditions to test whether A's taking decision is affected by the number of C players, holding constant punishment potential.

There was no communication among participants. Participants were told ex ante that groups would only be randomly composed after the experiment. At this point taking and punishment decisions would be implemented.[4] With this set up the A participants had to form expectations of the punishment decisions of the C participants in their punishment condition to whom they would be randomly assigned. In so doing it mimics deterrence in the real world where would-be offenders must form expectations of the risk of apprehension and punishment. This expectation was elicited by asking all A participants (as well as the B and C participants) to estimate the average fine exacted by 30 participants in their given condition, should the randomly matched A participant have taken from B. Participants received a monetary reward for the accuracy of their estimate. For detail see the Appendix.

We were not only interested in the effect of the guardian's punishment capacity on taking decisions but also, we want to gain insights into the motives that drive both the decision to take by A participants and the punishment decisions of C participant. To that end, after the main experiment we administered several supplementary tests from all participants for the purpose of eliciting (descriptive) beliefs about the prevalence of stealing, and about punishment, and

---

[4] On Prolific, participants are recruited ad hoc, once the implementation of a study has started. Some participants leave the experiment for technical reasons (e.g. because the connection is unstable), others leave as they are not sufficiently patient while waiting for the next stage of the experiment, or because they find a different online study more profitable, or more interesting. Were one to implement the standard protocol of lab experiments, each group in which a single participant drops out would be lost. We avoid this problem by assembling groups post hoc.

normative beliefs about either choice. The elicitation of both beliefs was incentivized to encourage more careful consideration of responses, using the mechanism introduced by (Krupka and Weber 2013).[5]

We further elicit risk aversion (Holt and Laury 2002) and administer a standard incentivized test of social value orientation (Murphy, Ackermann et al. 2011) in which participants make six different allocations of a monetary reward between themselves and an anonymous other. Most allocations involve their receiving a greater reward at the expense of the anonymous other. A random draw of one of these allocations determines the payout from this test.

Our pre-registered hypotheses are that role A participants will be the more deterred the greater the capacity of role C participants to punish, both in terms of the cost of punishment (the fine to fee ratio), and the maximum amount of punishment that role C participants can inflict on thieves. We expect that increasing the number of bystanders reduces punishment, and that this is anticipated by A participants.[6]

4. Results

3.1 The disciplining effect of a guardian

Does the presence of a guardian and guardian's punishment capacity discipline would-be offenders? As Figure 1 shows, the answer is a qualified yes. Condition 2 in which the guardians punishment capacity at worst makes taking a breakeven proposition is visibly (and statistically) indistinguishable from condition 1 with no guardian and where taking can be done with impunity. Yet in conditions 3 to 5 in which taking can be a losing proposition, the fraction of A participants who take from B is significantly lower than in the absence of a guardian. Descriptively, the taking

---

[5] Participants earn the more money the more accurately they post-predict the responses of the remaining participants in their session.
[6] See the exact wording of the preregistered hypotheses in the Appendix.

rate is lowest in the two conditions (4 and 5) with the highest punishment capacities. However, a two-tailed test does not find a significant difference from condition 3 (while a one-tailed test does): Thus, the results show that if C participants in their guardian role have sufficient punishment capacity to make taking potentially unprofitable (i.e., maximum punishment is power is 6£ or more), A participants are deterred. For a one-tailed test there is also support that a fine-to-fee ratio greater 1 also serves as a deterrent. However, there is no support for an increase in the maximum punishment from 6£ to 9£ or an increase in the fine-to-fee ratio from 2 to 3 acting as a deterrent. As we randomly assigned participants to treatments, differences across treatments permit a causal interpretation.
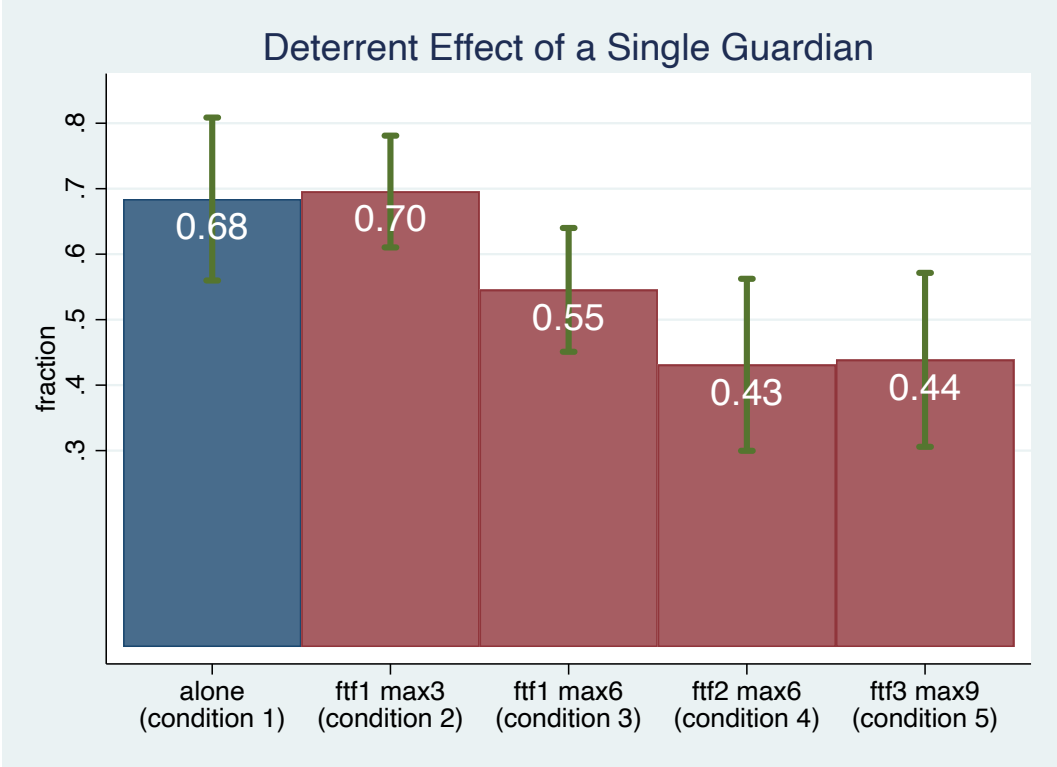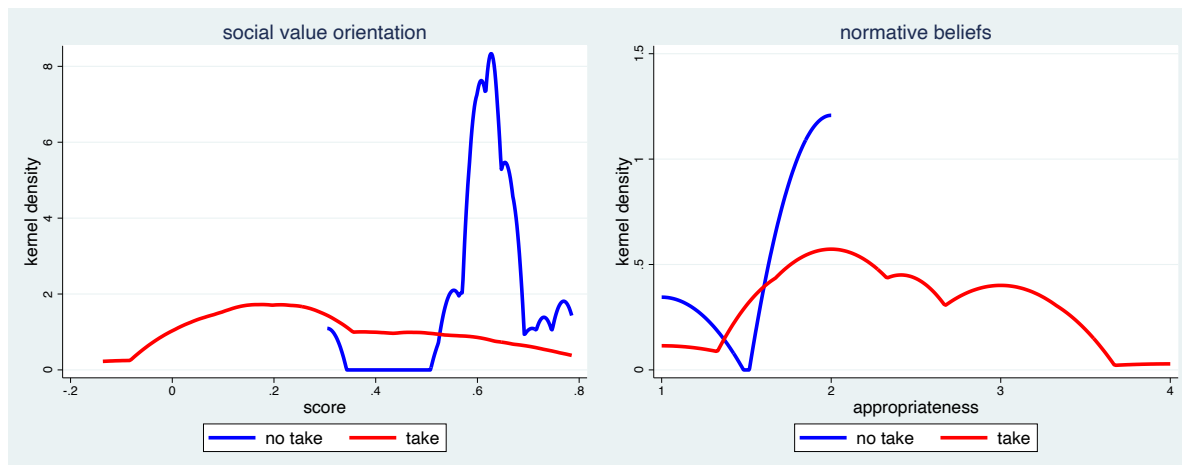

Figure 1
Deterrent Effect of a Single Guardian
error bars from 95% confidence intervals

Figure 1 holds a further message: Even in the absence of a guardian, and hence without the threat of punishment, less than 70% of A participants take from B. With the help of the battery of supplementary tests, we see why. As Figure 2 shows, those A participants who decide not to take when there is no deterrence radically differ in their motives. Those who do not take have much stronger social preferences for sharing, and a clear normative belief that rules out taking. The population of potential takers is pronouncedly heterogeneous in these regards.

Figure 2
Motives of A participants in the absence of guardians
left panel: social value orientation score, aggregated over 6 allocation choices (see Murphy Ackermann 2011 for calculations)
right panel: scores from Krupka Weber (2013)
kernel density plots

Figure 3 reports the taking rates of A participants by SVO quartile. Taking rates are a pronounced declining function of SVO quartile. Figure 4 shows that the strong association of the taking rate with SVO also has an impact on deterrence. If the taker has a social value orientation score at or above the median, there are no treatment effects. Such participants are intrinsically held back from crime and do not need to be deterred, just as predicted by Zimring and Hawkins (1973) and Pogarsky (2002); it is only those with lesser scruples against taking, namely those with below median SVO scores, who are held back by the risk of punishment. Compared to condition 1 with

no guardian, below median SVO participants are not deterred by condition 1 in which there is no risk of loss from taking, but are deterred by conditions 3-5 with maximum punishment capacities of 6£ or more. There is also some indication that a fine-to-fee ratio greater than 1 may additionally have a deterrent effect.

The Figure 4 result makes clear that the ethical stance on taking has a very pronounced impact on taking decisions. To compare the relative impacts of ethical versus sanctions concerns on taking decision Table 2 reports a standardized regression relating the taking decision of A participants to their perceptions of anticipated fine, their SVO score, and their assessment of the appropriateness of taking on a scale from 1 for very appropriate to 4 for very inappropriate. Each is significantly related to taking at the .05 level or better. Notwithstanding, the magnitudes of the association of SVO and appropriateness with taking are three times larger than that with anticipated punishment. Still another indication of this magnitude difference is that the R^2 of a regression of taking on conditions 1-5 is .05. With the addition of the participant SVO scores and inappropriateness assessments, R^2 increases 6-fold to .30. Both measures of effect magnitude imply that socialization about against harming others is far more effective in discouraging such behavior than the threat of punishment.
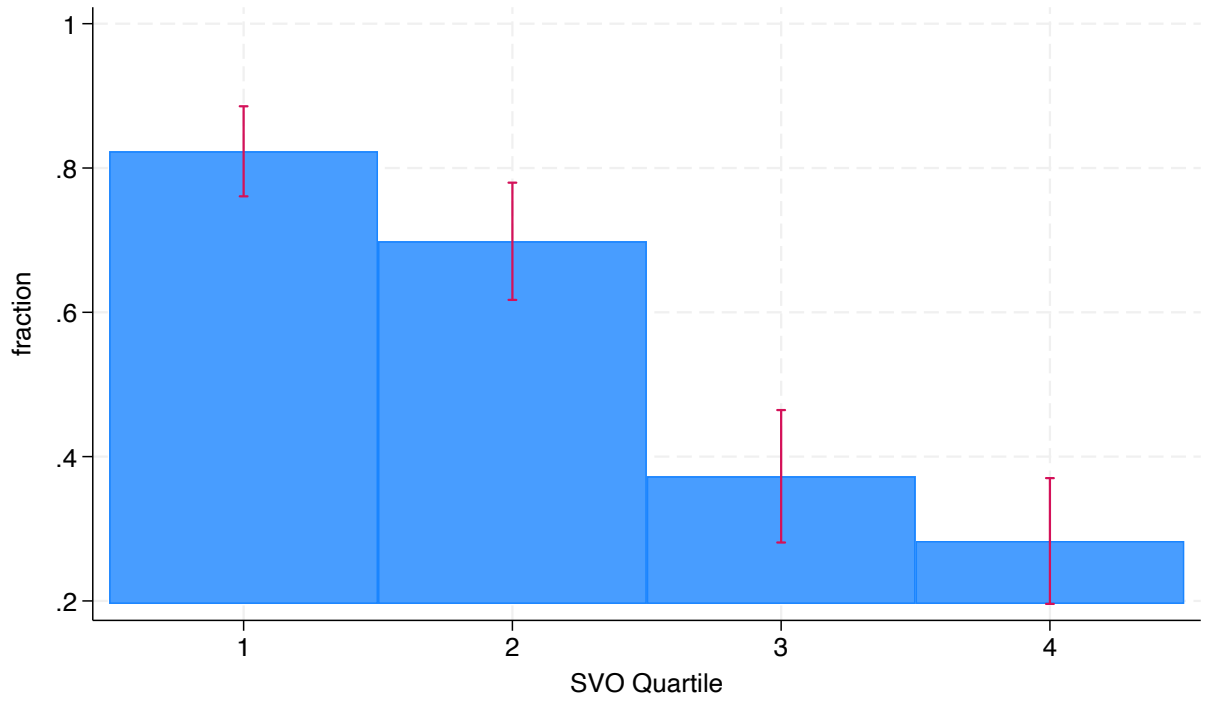
Figure 3
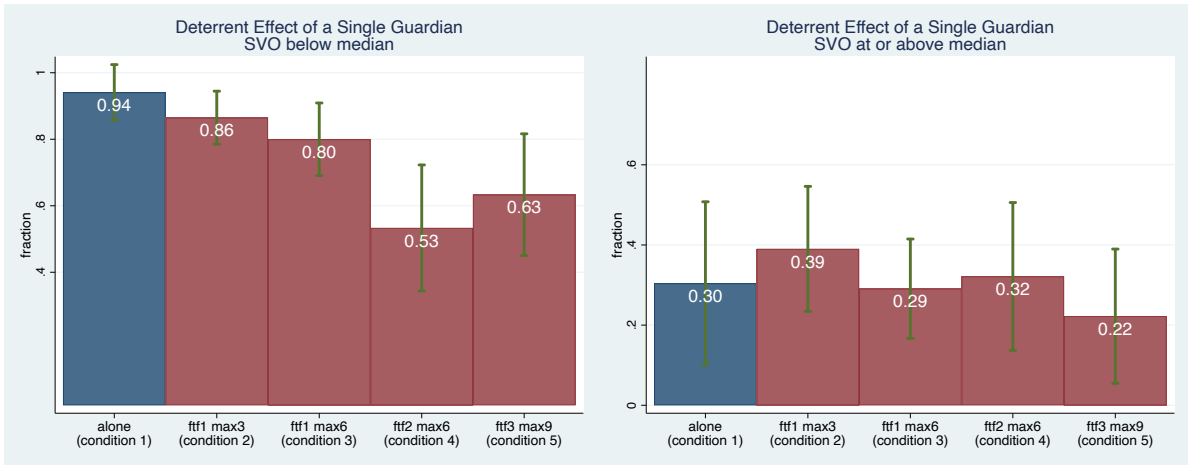*Taking Rates by SVO Quartile*

Figure 4
Treatment Effect on Taking Conditional on Social Value Orientation
median split by social value orientation score
error bars from 95% confidence intervals

Variable	Coefficient	Standard Error	t
Expected Fine	-.109	.0432	-2.53
SVO	-.362	.0480	-7.54
Inappropriateness of Taking	-.301	.0392	-7.67

Table 2
Standardized Regression of Take on Expected Fine, SVO, and Appropriateness of Taking

Turning now to the effect of number of would-be guardians, we had expected that potential takers would be less deterred if they face more than one potential guardian. As Figure 5 shows, this is not the case. The take rate is virtually identical if a second C participant has power to intervene.
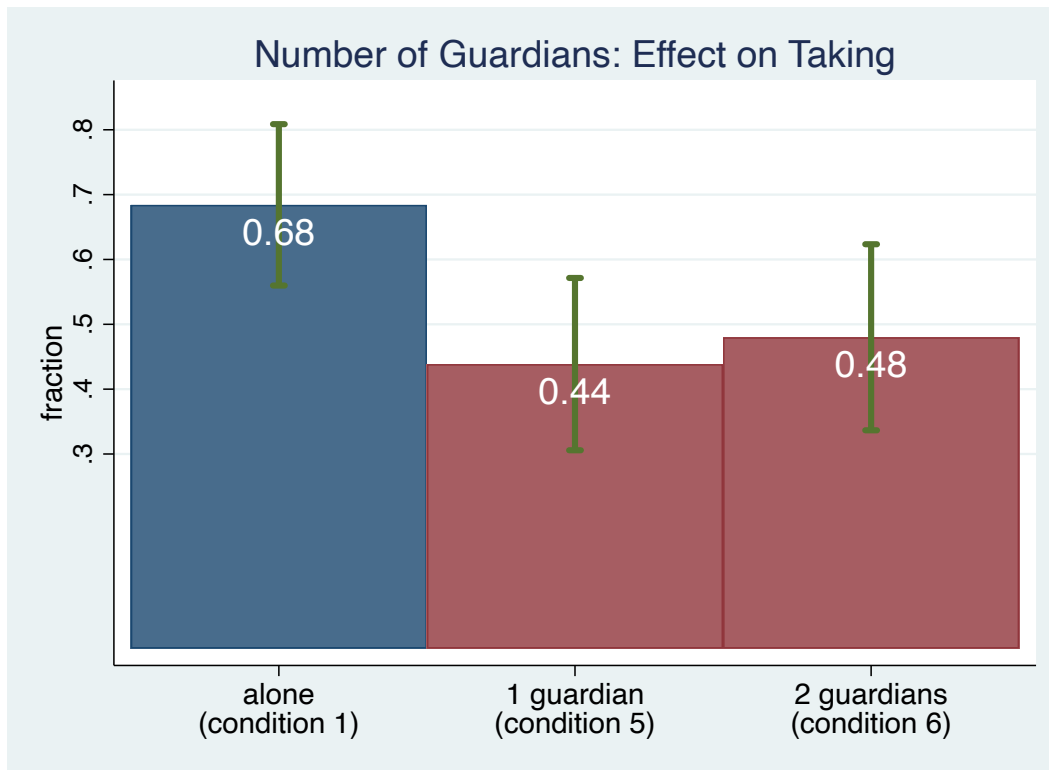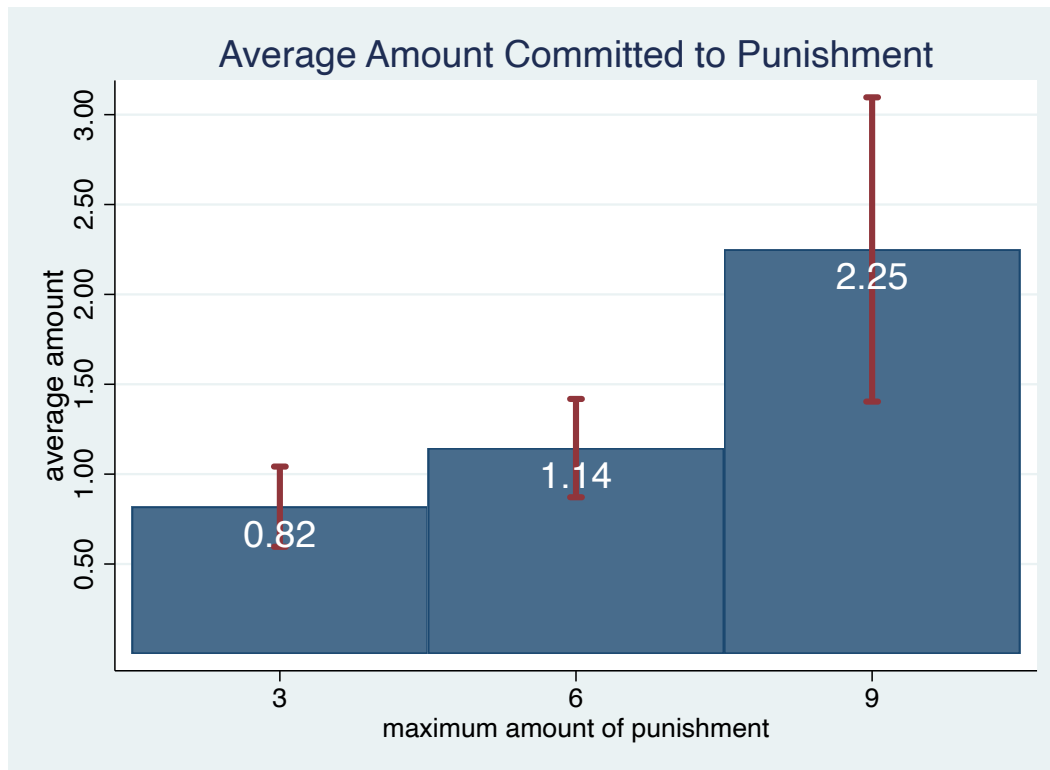
Figure 5
Effect of Number of Guardians on Taking
error bars from 95% confidence intervals

3.2 Guardian choices

For conditions 2-5 with a single guardian, Figure 6 reports the average fine imposed by would-be guardian C participants by their maximum punishment amount. While the amount of the commitment from their 6£ endowment does not vary across condition, the amount of punishment exacted does vary significantly depending on how much punishment power C participants have.

Figure 6
Severity of Punishment by Punishment Power
single bystander
error bars from 95% confidence intervals

Also, while takers make no distinction between a single and two potential guardians (Figure 5) against our expectations, if there are two individuals with punishment power, total punishment is substantially and significantly higher as shown in Figure 7.
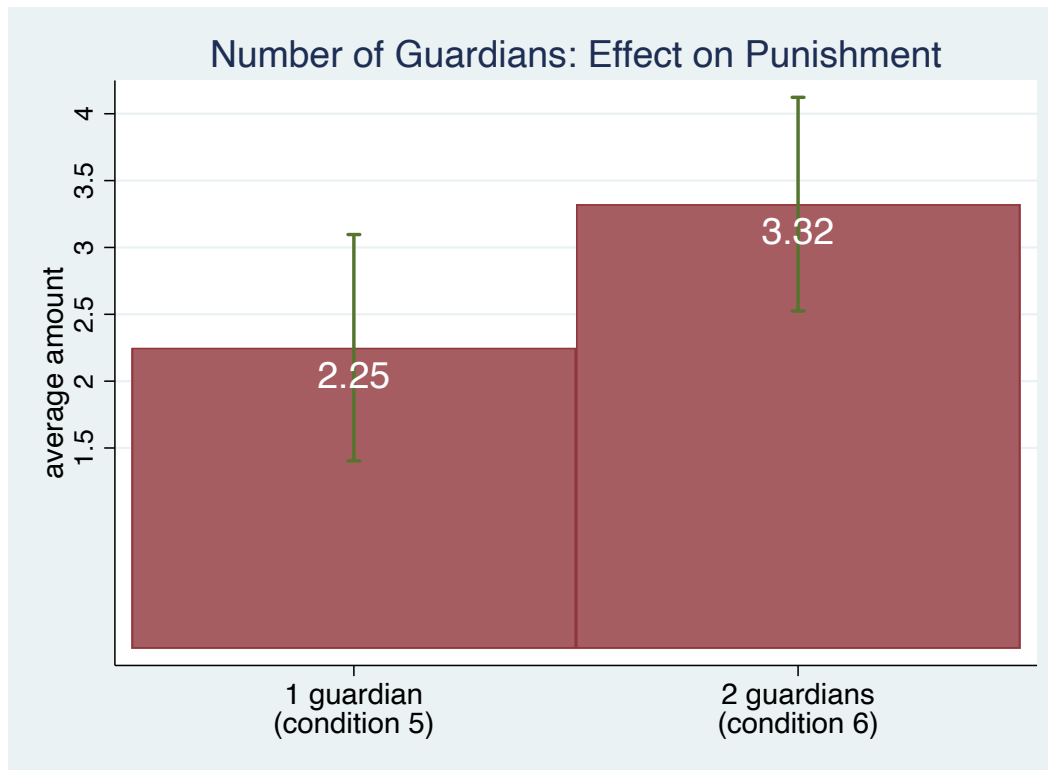
Figure 7
Severity of Punishment by Number of Guardians
maximum punishment 9
error bars from 95% confidence intervals

What then does affect the commitments, if any, of C role participants to the punishment for taking? For single guardian conditions 2 to 5, Table 3 reports a regression of the amount C participants commit to punishment on treatment condition, the participant's SVO and assessment of the appropriateness of punishment on scale from 1 for very inappropriate to 4 for very appropriate. While condition is not significantly related to punishment commitment, it is significantly and positively related to the C participant's SVO and their assessment of the appropriateness of punishment for taking.

Variable	Coefficient	Standard Error	T
Condition 3	.077	.165	.47
Condition 4	.054	.195	.28
Condition 5	-106	.167	-.63
SVO	.536	.247	2.17
Appropriateness of Punishment	.488	.108	4.51
Constant	1.97	.345	5.72

Table 3
Regression of Amount Committed to Punishment on Condition, SVO and Appropriateness of Punishment
single guardian
robust standard errors

5. Discussion and Conclusions

Guardianship is an active area of research in criminology. But neither social psychology experiments nor the ex-post analysis of video footage from crime scenes can make visible what arguably is the greatest social benefit of guardianship: observing bystanders who seem to be vigilant, and might be willing to intervene, keeping potential offenders in check. In the lab, this socially most desirable effect of guardianship can be observed: the pre-emption of criminal acts in the first place.[7]

---

[7] Research on reactive guardianship in which bystanders intervene in ongoing event makes the important point that intervention can come in many forms, from notification of the police to own physical intervention. However, if the threat of guardian intervention in some form deters the criminal event, the specific form of the intervention is secondary to the would-be offender's perception of the guardian capacity to intervene in some fashion that makes undertaking the offense unattractive. In this experiment that capacity is determined by the punishment maximum and the fine-to-fee ratio.

In our experiment, we do indeed find such a deterrent effect, provided guardians have sufficient punishment power. The less intervention is costly for them, the more pronounced deterrence. The socially desirable containment of crime obtains whether or not the ability to intervene is held by a single individual, or by two of them. Guardians are willing to punish infractions even if punishing is costly for them. But the lower the cost, the more trenchant the intervention. If punishment power is distributed across two individuals, total interventions are even more severe.

From a policy perspective the results serve as a valuable proof of concept—guardians with sufficient capacity can contain crime. The challenge is to convert this conclusion into actionable policy. How then do the results speak to the reality that the experiment was intended to understand? While there is a very large, decades long literature on informal social control, its focus is on establishing community-level characteristics theorized to affect informal social control either positively or negatively, not on the actions or the effectiveness of community members.

Other literatures, however, do focus on citizen led efforts to prevent crime. Sharkey et al. (Sharkey, Torrats-Espinosa et al. 2017) argue that such groups made an under-appreciated contribution to the sustained drop in the USA's crime rate from the early 1990s to about 2020. They conclude "that every 10 organizations focusing on crime and community life in a city of 100,000 residents lead to a 9 percent reduction in the murder rate, a 6 percent reduction in the violent crime rate, and a 4 percent reduction in the property crime rate." The study, however, does not examine the determinants of the effectiveness of different forms of such organized efforts.

The 2023 National Research Council report "Reducing Racial Inequality in Crime and Justice: Science, Practice, and Policy" commits a chapter to a review of the varying forms of community organized efforts to prevent crime. In forms such as "Operation Ceasefire" community

members work in conjunction with law enforcement to deter violent crime, whereas other forms such as "Cure Violence" work separately from law enforcement and attempt to prevent crime by adopting public health-type epidemic prevention tactics. The details of the philosophies and structures of these citizen led efforts are not relevant for our purposes here. What is relevant is the report's circumspect conclusion: "Given some suggestive but incomplete results from a burgeoning literature on community-based anti-violence and related initiatives, federal and state agencies should explore the significant expansion of community-driven pilot programs that are fielded in combination with strong evaluation strategies" (National Research Council 2023, 225).

Our experiment, we believe, provides a useful framework for designing strong evaluation strategies. The key conclusion of our laboratory experiment is that C role participants can deter A role participants from taking if they have sufficient punishment power. For this discussion of real-world implications of this finding, we restate the conclusion as: private citizen groups can be effective in preventing crime if they have sufficient capacity to deter or by other means discourage would-be offenders from acting out their intentions.

This conclusion restatement implies that the starting point of an evaluation protocol should be an assessment of the group's capacity for crime prevention. This assessment will depend on the group's intended strategy for preventing crime. If it is deterrence via a cooperative working arrangement with law enforcement such as that in Operation Ceasefire, the assessment would involve an evaluation of whether the combined capacities of law enforcement and the community group can deliver a credible deterrence threat. On the other hand, if the group's strategy is to operate independently of law enforcement like, for example, Cure Violence, the assessment would involve an evaluation of whether the intended tactic has a credible potential to prevent crime.

Capacity needs to be assessed relative to the magnitude of the crime problem the group is confronting. In our lab experiment had the A role participants been able to take the entirety of B's 6£ endowment, the 6£ punishment capacity would likely have been ineffective in deterring A, whereas it was effective in deterring A under the condition of A only being able to take 3£s. Thus, we anticipate that citizen vigilance may be ineffective in communities in which public safety is in a very bad state. Returning to Anderson's journey along Germantown Avenue, he concludes that capacity of the "decent" citizens to exert effective informal social control is overwhelmed by the code of streets ethic of the "street" people. Anderson's assessment is shared by Hureau, Braga et al. (2023) who review studies of the effectiveness of Cure Violence-like programs to diffuse gang violence and also conduct an evaluation of a Cure-like program called Street-Safe Boston. Their review concludes that evidence of effectiveness is scant and even worse: programs may be iatrogenic in some instances. The evaluation of Street-Safe Boston finds no evidence of effectiveness. Thus, effective informal social control by private citizens or groups thereof may be limited to the less crime infested communities midway between Chestnut Hill and Germantown.

In our experiment the personal cost of serving as a guardian is monetary. In the real life the dominant cost is safety. Thus, another dimension of the evaluation should be an assessment of the seriousness of the safety risks confronting citizens serving as guardians, particularly from gun violence in places where gun carrying by private citizens, whether legal or illegal, is widespread. Balanced against the safety and also time cost of serving as a guardian may be the social rewards bestowed by community members for working to protect the public safety. These too should be assessed. We anticipate such rewards will be higher in high collective efficacy communities. This hypothesis should be tested both in community-based ecological research and in variations of the lab experiment used here.

At outset we noted that police presence cannot and should not be ubiquitous. Informal social control offers the promise of complementing law enforcement, as in Operation Ceasefire, or substituting for law enforcement as in Cure Violence. Either way, in circumstances such as that in contemporary USA in which large segments of the population are skeptical of law enforcement and the number of law enforcement officers has declined materially, identification of effective complements to and substitutes for conventional law enforcement should be a policy priority. Their identification will require a large investment in rigorous evaluations along the lines we describe above of citizen-based crime prevention programs.

We found that both participants with the ability to steal and participants with the ability to intervene are heterogeneous. The most important dimensions of heterogeneity are participant weight on social well-being, and individual normative convictions. This heterogeneity had a particularly pronounced impact on taking decisions. Those who were high on both these dimensions refrained for taking as a matter of principle, not out of concern for punishment. Supplementary analyses indicate that SVO and normative convictions have a far larger impact on taking decisions than the capacity to punish. Also, would-be guardians who were high on these dimensions were more likely to make the personally costly decision to punish would-be takers for taking. This heterogeneity among both would-be takers and would-be guardians profoundly changes the policy problem. On the one hand, the problem becomes easier: many opportunities for offending will remain unseized in a population that is dominated by individuals socialized to respect the person and property of others, as exemplified by Anderson (1999)'s account of the residents of Chestnut Hill. Public policies such as early childhood interventions that have been shown to advance such socialization should be embraced (Farrington and Welsh 2008). On the other hand, in communities where such dominance is eroded by the presences of large numbers of

individuals lacking such ethical scruples the problem becomes harder: it is not enough to screen out opportunities for offending. Preventing crime is premised on the ability to discriminate between potential perpetrators. Arguably this is where community guardianship shines, as community members can muster local knowledge of individuals who are prone to offend or otherwise engage in disruptive behaviors if not closely surveilled.

Criminological and sociological research on the roots of crime and its prevention have historically made sparing use of laboratory experiments to test theories and prevention strategies, probably because of external validity concerns. While we acknowledge that our laboratory experiment is far removed from the real world of crime, as it demonstrates, laboratory experiments can nonetheless provide valuable insights into the roots and prevention of crime, the more so the more one has reason to believe that basic behavioral mechanisms are critical.

References

Anderson, Elijah (1999). Code of the Street. New York, WW Norton & Company.

Barnum, Timothy C, Shahina Herman, Jean-Louis van Gelder, Denis Ribeaud, Manuel Eisner and Daniel Nagin (2024). Guardianship: Who Intervenes? How? And Why?

Chandler, Jesse, Cheskie Rosenzweig, Aaron J Moss, Jonathan Robinson and Leib Litman (2019). "Online Panels in Social Science Research: Expanding Sampling Methods Beyond Mechanical Turk." Behavior Research Methods 51: 2022-2038.

Chaudhuri, Ananish (2011). "Sustaining Cooperation in Laboratory Public Goods Experiments. A Selective Survey of the Literature." Experimental Economics 14(1): 47-83.

Cohen, Lawrence E and Marcus Felson (1979). "Social Change and Crime Rate Trends. A Routine Activity Approach." American Sociological Review 44: 588-608.

Farrington, David P and Brandon C Welsh (2008). Saving Children from a Life of Crime: Early Risk Factors and Effective Interventions. New York, Oxford University Press.

Fehr, Ernst and Simon Gächter (2000). "Fairness and Retaliation. The Economics of Reciprocity." Journal of Economic Perspectives 14: 159-181.

Felson, Marcus and Rachel L Boba (2010). Crime and Everyday Life, Sage.

Fischer, Peter, Joachim I Krueger, Tobias Greitemeyer, Claudia Vogrincic, Andreas Kastenmüller, Dieter Frey, Moritz Heene, Magdalena Wicher and Martina Kainbacher (2011). "The Bystander-Effect. A Meta-Analytic Review on Bystander Intervention in Dangerous and Non-Dangerous Emergencies." Psychological Bulletin 137(4): 517-537.

Heckathorn, Douglas D. (1989). "Collective Action and the Second-Order Free-Rider Problem." Rationality and Society **1**: 78-100.

Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, Frank Marlowe, David Tracer and John Ziker (2006). "Costly Punishment across Human Societies." Science **312**(5781): 1767-1770.

Holt, Charles A. and Susan K. Laury (2002). "Risk Aversion and Incentive Effects." American Economic Review **92**: 1644-1655.

Hureau, David M, Anthony A Braga, Tracey Lloyd and Christopher Winship (2023). "Streetwork at the Crossroads: An Evaluation of a Street Gang Outreach Intervention and Holistic Appraisal of the Research Evidence." Criminology **61**(4): 758-794.

Jin, Kelou, Jie Wu, Ran Zhang, Shen Zhang, Xiaoyan Wu, Tingting Wu, Ruolei Gu and Chao Liu (2024). "Observing Heroic Behavior and Its Influencing Factors in Immersive Virtual Environments." Proceedings of the National Academy of Sciences **121**(17): e2314590121.

Krupka, Erin L. and Roberto A. Weber (2013). "Identifying Social Norms Using Cordination Games. Why Does Dictator Game Sharing Vary?" Journal of the European Economic Association **11**(3): 495-524.

Liebst, Lasse Suonperä, Richard Philpot, Mark Levine and Marie Rosenkrantz Lindegaard (2021). "Cross-National Cctv Footage Shows Low Victimization Risk for Bystander Interveners in Public Conflicts." Psychology of Violence **11**(1): vio0000299.

Lindegaard, Marie Rosenkrantz, Lasse Suonperä Liebst, Wim Bernasco, Marie Bruvik Heinskou, Richard Philpot, Mark Levine and Peter Verbeek (2017). "Consolation in the Aftermath of Robberies Resembles Post-Aggression Consolation in Chimpanzees." PloS one **12**(5): e0177725.

Murphy, Ryan O, Kurt A Ackermann and Michel Handgraaf (2011). "Measuring Social Value Orientation." Judgment and Decision Making **6**(8): 771-781.

Nagin, Daniel S, Shaina Herman and Timothy C Barnum (2023). "Collective Guardianship, Reactive and Proactive." Crime and Justice **52**(1): 487-523.

National Research Council (2023). Reducing Racial Inequalities in Crime and Justice: Practice and Policy.

Peer, Eyal, Laura Brandimarte, Sonam Samat and Alessandro Acquisti (2017). "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research." Journal of Experimental Social Psychology **70**: 153-163.

Pogarsky, Greg (2002). "Identifying 'Deterrable' Offenders. Implications for Research on Deterrence." Justice Quarterly **19**: 431-452.

Reynald, Danielle M (2009). "Guardianship in Action. Developing a New Tool for Measurement." Crime Prevention and Community Safety **11**: 1-20.

Sampson, Robert J (2006). Collective Efficacy Theory. Lessons Learned and Directions for Future Inquiry. Taking Stock. F. T. Cullen, J. P. Wright and K. R. Blevins. New Brunswick, Transaction**:** 149-167.

Sampson, Robert J, Stephen W Raudenbush and Felton Earls (1997). "Neighborhoods and Violent Crime. A Multilevel Study of Collective Efficacy." science **277**(5328): 918-924.

Sharkey, Patrick, Gerard Torrats-Espinosa and Delaram Takyar (2017). "Community and the Crime Decline. The Causal Effect of Local Nonprofits on Violent Crime." American Sociological Review **82**(6): 1214-1240.

Shaw, Clifford Robe and Henry Donald McKay (1942). Juvenile Delinquency and Urban Areas. Chicago, University of Chicago Press.

Whyte, William Foote (1947). Street Corner Society. The Social Structure of an Italian Slum. Chicago, University of Chicago Press.

Yamagishi, Toshio (1986). "The Provision of a Sanctioning System as a Public Good." Journal of Personality and Social Psychology **51**: 110-116.

Zimring, Franklin E. and Gordon Hawkins (1973). Deterrence. The Legal Threat in Crime Control. Chicago, Chicago University Press.