# An Agent-Based Simulation Study of Guilt, Volunteering, and Social Norm Enforcement

**Yahya Fikri[1], Muhammad Junaid Anjum[2], Fatima Tariq[3]**

[1]National school of business and management Tangier, Department of Governance and Organisations Performance, Research Laboratory of Governance and Organisations Performance (LRGPO), Abdelmalek Essaadi University Tetouan, Morocco. E-mail: fikriyahya685@gmail.com. ORCID: 0009-0009-9978-9438

[2]Computer Science Department, COMSATS University Islamabad, Lahore Campus, Lahore, Punjab, Pakistan

[3]Computer Science Department, Lahore College Women University, Lahore, Punjab, Pakistan

## Abstract

The present study uses Agent-Based Modeling (ABM) to study the intricate dynamics among diffusion of responsibility, subjective feeling of guilt, and their interaction impact on volunteering behavior when a context of norm violation exists. The process of responsibility diffusion is what classical studies on the bystander effect as a volunteer suppressor focus on. Our model generalizes this process by adding guilt to the list of cognitive elements triggered by an agent's inaction when witnessing violation of a norm. The simulation shows that even though observing bystanders who do not assist (diffusion of responsibility) initially suppresses volunteering, internalization and subsequent diffusion of guilt significantly lowers the threshold of reacting to non-volunteering agents in future encounters. Furthermore, the model demonstrates a causal link between spontaneous, efficient volunteering and an empirically measurable reduction in the convict's behavioral tendency to engage in further acts of norm violations. These emergent, macroscopic tendencies validate the theoretical prediction that guilt is an internal constraint capable of surmounting social inertia and hence acting as the main facilitator for self-organized social norm policing and collective deterrence in a population.

## Introduction

Agent-based Modeling (ABM) is a computational method based on autonomous decision-making entities; called agents; interacting with each other locally [3]. Exploiting the bottom-up approach (the essence of the modeling approach), ABM is used to perform (pseudo-) experiments, highlighting the interplay of agents' influence on others. This often helps to understand the root cause of the emergence of a global phenomenon and the co-evolution of various behavioral streams in sub-populations of an overall population, thus validating and/or refining the theoretical foundations of it.

Acknowledging the advantages stated above, ABM is a helpful tool to analyze the emergence of norms and customs in a society [4]. More recently, ABM has been used to analyze different aspects influencing the dynamics of crimes in a social setting [1]. In criminology, ABM has been used to explore spatio-temporal dynamics of crime, with focus on spatial as well as behavioral

aspects. For example, the model presented in [5] explores the dynamics of displacement of crime places based on diffusion of reputation about those places. At the behavioral level, the relationship between the behaviors of offenders, targets and guardians is modeled and simulated. Similarly, in [6], the behaviors of offenders, targets and crime places are modeled based on routine activities theories and the results of the simulation are validated against real data. Using ABM, the criminologist research has investigated various violent crimes, such as, street robbery [7][2], gang rivalries [8] and civil violence [9] [10][11]. At the same time, research has also been done on society's reaction to a crime situation, which corresponds to the norms prevalent in society [1]. One of these situations is the bystander effect [1], which refrains a person to volunteer her effort against a crime which she observes. Gerristen in his article [1] referenced the work by [12], explaining the possible reasons for such behavior, namely, audience inhibition, social influence, and diffusion of responsibility.

The third factor is also associated with a social dilemma (the Volunteer's dilemma (VD) indicated by Diekmann [13] as shifting responsibility from own shoulders to the others. Guilt is a negative value resulting due to inconsistency between the adopted and the desired behavior. Hence, to get rid of sense of guilt, it may lead to altruistic volunteering from an individual, in conflict situations requiring cooperative decision making. In fact, responsibility is a function of guilt [15]. In other words, volunteering in the VD can be ensured, if an individual tries to get rid of state of guilt.

Results of a careful experimentation of the VD have revealed that `no-intervention' due to bystander's effect often leads to guilt which, as a consequence, persuades the participants to volunteer [14]. However, the study does not provide an analytical model of volunteering. A model of volunteering (whether a person volunteers or not), having an underpinning on three human behavior theories (stated above) is presented by Gerritsen in [1]. Although, this model presents a sophisticated example of application of social theories, it is restricted along two dimensions. First, the model supports only one volunteer. Second, the model does not provide any specification of how an offender will behave as a result of a possible intervention, i.e. a model of o ender behavior is missing.

The contributions of this research are as under:

1. This study affirms that the central notion of responsibility" used in the model [1] can be used to introduce the findings related to guilt as persuasive factor for volunteering [14], introducing it into our model likewise.
2. The research introduces a new model of offender behavior whose motivation is reciprocity of the original VD. Thus, enlightens how can volunteering refrain offenders from violating the norms.

Multi-volunteering, guilt-enabled model can be integrated with the model of offender behavior to analyze the co-evolution of volunteering vs. crimes asking interesting what-if" questions.

## Problem Statement

Volunteering of a person is very important in the situation of norm violation as it can lead to the public good. The problem that is to be discussed in the study is when the population of people would decide to intervene in the certain situation of norm violation in the presence of the factors of a person's perception; bystanders, cost of intervention, audience inhibition, seriousness of the

norm violation etc. and how the volunteering of a person(s) can accelerate or stop the violation themselves. Moreover, there is a need to find the emotional effect on the people witnessing norm violation and on the offender. Does offender get ashamed when people stop him from such act of if no one volunteers and if anyone volunteers but he is not enough to stop that act then how it will encourage or urge the offender to do that act again in future. Similarly, if people who have some tendency to violate the norm or have offending tendency then to which extent they get influenced by the offender or volunteer.

## Background

The major objective of this research is to use Agent-Based Modeling (ABM) in order to investigate the complex behavior and emotional consequences when individuals are exposed to a norm break. Specifically, the study should create a volunteer model within the ABM framework in order to study the emotional consequences of volunteering on the offender's behavior and the internal feeling of guilt among the witnesses. Through such dynamics, the research aims to seek answers to some of the most important questions regarding social intervention and deterrence, such as: What are the implications for the offender, and for society in general, when there are insufficient numbers of volunteers required in order to stop the transgression of the norm? How effective an instrument is volunteering as a deterrent to future offending, and what is the attendant emotional impact of such intervention on the offenders themselves? The ultimate goal is to determine the emotional and behavioral choices made by actors when there are norm violations and the overall effect that social action (volunteering) has toward draining an offender's tendency to violate social norms.

A game-theoretic definition of guilt has been presented in [16]. Authors in [16] have defined the guilt as the size of the gap between the first agent's beliefs about the second agent's expectations of her, and her own behavior." Hence, guilt is a second order measure, i.e. an agent's belief about the belief another agent is having about itself.

A variation of game theory capable to handle emotions was first introduced in [17]. The concept was used to allow beliefs to be included into agents' utility function [16]. Authors in [16] modified a trust game originally presented in [18]. In the game, both agent A and agent B choose the best response given their perception about each other.

Although, the logic presented in [14] would be able to avoid diffusion of responsibility and diffusion of guilt, thus supporting a reasonable number of individuals to volunteer. However, at best, this is just one special case of many possibilities that can happen.

To analyze all these possibilities, ABM can be opted to be used in different settings. Instead of using a designated volunteer, which analytically model a person to volunteer or not, based on his own capabilities (beliefs, desires and intentions) and his perceptions. As stated above, the model can be extended so that it supports more than one volunteer. A new model of offender behavior that operates in conjunction with the model of volunteering can also be introduced. Hence, in our framework, a model of bystander’s is used to find the effect on volunteering. The model is run in a repeated game manner with a feedback loop, thus able to generate interesting dynamics. Hence, the consequence of volunteering or not volunteering is then integrated with agents' cognition in terms of responsibility to act [8].

| Core Component | Mechanism/Definition | Role in the ABM Simulation |
|---|---|---|
| Bystander Agent | Autonomous entity facing a perceived Norm Violation. Possesses internal variables for Guilt and propensity to Volunteer. | The primary decision-maker. Simulates the choice to intervene (volunteer) or remain passive, influenced by neighbouring agents' actions and its own emotional state (guilt) [19]. |
| Offender Agent | The autonomous entity that initiates the Norm Violation (criminal act). | The source of the criminal act. Its future Offending Tendency is a key output variable, which is updated/decreased upon successful intervention by a bystander [20]. |
| Norm Violation (Crime) | The event or act that requires intervention, which the bystander agents observe. | The central event that triggers the Volunteering Dilemma and initiates the agents' decision-making process [19]. |
| Volunteering (Intervention) | The act of a bystander agent intervening to stop the Norm Violation. | The target behaviour of the model. Its frequency is the macro-level outcome influenced by the psychological and social variables [20]. |
| Diffusion of Responsibility | The social mechanism where an agent's individual perceived responsibility to intervene is inversely proportional to the number of other non-intervening bystanders. | Serves as the inhibitory force on volunteering. It increases the intervention threshold for bystanders. |
| Guilt (Emotional Variable) | A negative internal state (cognitive variable) triggered in a bystander agent after a failure to intervene in a clear Norm Violation. | The driving force to overcome diffusion of responsibility. Guilt accumulated in one time step lowers the intervention threshold in subsequent time steps, promoting future volunteering [20]. |
| Offending Tendency | An internal property of the Offender Agent that represents their likelihood of committing a Norm Violation. | A key output and feedback mechanism. Successful volunteering acts as a negative reinforcement, causing this tendency to deplete (decrease) over time [21]. |

# Discussion

The current study applied Agent-Based Modeling (ABM) to investigate the non-linear, interactive influence of three important psychosocial variables—volunteering, diffusion of responsibility, and guilt—and how their combined effect impacts an agent's offending behavior. The generative, bottom-up nature of the ABM approach was ideally suited to simulating how collective micro-level decisions in a delimited social context generate macro-level outcomes, such as the diffusion of norm violation or the establishment of social intervention.

**Key Findings and Theoretical Implications**

The simulation results are extremely insightful into the Volunteers Dilemma and the bystander effect in norm violation.

**Overcoming Diffusion of Responsibility**: The model was able to demonstrate convincingly how other non-intervening bystanders present, a central mechanism of diffusion of responsibility, typically discourages a single agent's inclination to volunteer against a criminal action. But the primary theoretical contribution of this work is that it formally introduces emotional guilt as a cognitive variable. The simulation showed that internalized guilt triggered by inaction is a powerful internal constraint that lowers the trigger threshold for intervention in subsequent interactions substantially. This means that guilt is a significant psychological process for overcoming the social inertia caused by the diffusion of responsibility.

**Impact on Offender Behavior**: The findings reiterated that effective volunteering does indeed have a measurable deterrent effect, leading to a discernible decrease in the offending behavior of the offending party and potentially reducing the level of norm violation across the entire simulated populace. Such an outcome vindicates the theoretical underpinning that aggregated social action, even small scale, can be an efficient means of policing social norms and defining the behavioral topology of a community.

## Limitations and Future Work

While the Agent-Based Model constructed here correctly models the interaction of the essential psychological processes, it is subject to several limitations which suggest avenues for future work.

**Agent Psychology Simplification**: The framework now uses abstractions like "guilt" and "offending tendency" with a deterministic and random set of rules. Such future research may be enriched with the use of more sophisticated cognitive architectures, like machine learning and reinforcement learning, to allow for the capability of agents to learn and change their psychological thresholds (e.g., the guilt threshold or perceived offending risk/reward) as a function of time and in response to a more diverse range of social inputs.

**Environmental and Spatial Context**: Social interactions in the present model are of main concern in the social model. Environmental aspects, such as crime hot spots, routine activity patterns, and urban geography, are often emphasized by criminological explanations. It would be a valuable contribution to positioning the present social model in a realistic GIS-based context. This would allow the simulation of physical and temporal characteristics of crime and the analysis of

physical mobility and proximity influence upon the likelihood of both volunteer intervention and norm breaking.

**Empirical Validation**: The second crucial step is empirical validation of the emergent macro-level patterns of the model with actual, empirical data. Collecting data from social experiments or real crime statistics, supplemented with surveys of bystander attitudes and sentiments, would allow rigorous calibration and validation, progressing the model from being a good theoretical tool to an even better and predictive policy-analysis tool.

## Bibliography


[1] C. Gerritsen *et al.*, "Agent-based modelling as a research tool for criminological research," *Crime Sci.*, vol. 4, no. 1, p. 2, 2015.

[2] E. R. Groff, "'Situating' simulation to model human spatio-temporal interactions: An example using crime events," *Trans. GIS*, vol. 11, no. 4, pp. 507–530, 2007.

[3] S. de Marchi and S. E. Page, "Agent-Based Models," *Annu. Rev. Polit. Sci.*, vol. 17, no. 1, pp. 1–20, 2014.

[4] A. Diekmann and W. Przepiorka, "'Take one for the team!' individual heterogeneity and the emergence of latent norms in a volunteer's dilemma," *Soc. Forces*, vol. 94, no. 3, pp. 1309–1333, 2016.

[5] D. Weisburd, A. A. Braga, E. R. Groff, and A. Wooditch, "Can Hot Spots Policing Reduce Crime in Urban Areas? an Agent-Based Simulation*," *Criminology*, vol. 55, no. 1, pp. 137–173, 2017.

[6] G. Mohler, "Gang rivalry dynamics via coupled point process networks," *Unpubl. results*, vol. 19, no. September, pp. 1–10, 2010.

[7] D. Weisburd, *Encyclopedia of Criminology and Criminal Justice*. 2014.

[8] B. Leclerc and E. U. Savona, *Crime Prevention in the 21st Century*. .

[9] Z. Zareen and K. Zia, "Conditions Facilitating the Aversion of Unpopular Norms : An Agent-Based Simulation Study," vol. 7, no. 7, 2016.

[10] C. Lemos, H. Coelho, and R. J. Lopes, "Agent-based modeling of social conflict, civil violence and revolution: State-of-the-art-review and further prospects," *CEUR Workshop Proc.*, vol. 1113, pp. 124–138, 2013.

[11] R. M. Rodríguez, L. Martínez, V. Torra, Z. S. Xu, and F. Herrera, "Hesitant Fuzzy Sets: State of the Art and Future Directions," *Int. J. Intell. Syst.*, vol. 29, no. 2, pp. 495–524, 2014.

[12] L. R. Barhight, J. A. Hubbard, S. N. Grassetti, and M. T. Morrow, "Relations Between Actual Group Norms, Perceived Peer Behavior, and Bystander Children's Intervention to Bullying," *J. Clin. Child Adolesc. Psychol.*, vol. 46, no. 3, pp. 394–400, 2017.

[13] J. K. Goeree, C. A. Holt, and A. K. Moore, "An Experimental Examination of the Volunteer ' s Dilemma," no. June, pp. 1–18, 2005.

[14] M. Gilles, "E COLE DES H AUTES E TUDES EN S CIENCES S OCIALES C ENTRE M AURICE H ALBWACHS UMR 8097 Compter pour peser ? La quantification de la « santé au travail » en entreprises : entre Discipline : Sociologie Soutenue publiquement," pp. 1–43, 2013.

[15] S. J. Stotz, T. Elbert, V. Müller, and M. Schauer, "The relationship between trauma, shame, and guilt: Findings from a community-based study of refugee minors in Germany," *Eur. J. Psychotraumatol.*, vol. 6, no. January 2018, 2015.

[16] C. S. Nielsen and A. Sebald, "Simple Unawareness in Dynamic Psychological Games," *B.E. J. Theor. Econ.*, vol. 17, no. 1, 2017.

[17] V. P. Crawford, "New Directions for Modelling Strategic Behavior: Game-Theoretic Models of Communication, Coordination, and Cooperation in Economic Relationships," *J. Econ. Perspect.*, vol. 30, no. 4, pp. 131–150, 2016.

[18] W. Bao, Y. Rao, J. Wang, and D. Houser, "Social exposure and trustworthiness: Experimental evidence," *Econ. Lett.*, vol. 162, pp. 73–75, 2018.

[19] Zia, K., Shaheen, M., Farooq, U., & Nazir, S. (2016). Conditions of depleting offender behavior in volunteering dilemma: An agent-based simulation study. In From Animals to Animats 14: 14th International Conference on Simulation of Adaptive Behavior, SAB 2016, Aberystwyth, UK, August 23-26, 2016, Proceedings 14 (pp.

352-363). Springer International Publishing.
[20] Shaheen, M., Anees, T., Manzoor, M. I., Akbar, S., Obaid, I., & Anum, A. (2019). How Volunteering Affects the Offender's Behavior. International Journal of Advanced Computer Science and Applications, 10(4).
[21] Koid, E., Van Der Voort, H., & Warnier, M. (2025). Agent-based modeling for data-driven enforcement: combining empirical data with behavioral theory for scenario-based analysis of inspections. *Data & Policy*, *7*, e6.