

## **Bystanders' Perceptions and Reporting Responses in Doxing Cases: The Influence of Online Social Dynamics**

**Yuan Fang**

School of Business, The University of Queensland,  
Brisbane, Queensland, Australia

### **ABSTRACT**

Doxing, the act of publicly revealing personal information online, often with the intent to harm, poses a significant threat to privacy on social media. Despite ongoing efforts to mitigate its impact, effective intervention mechanisms remain inadequate. This research-in-progress explores factors influencing bystanders' willingness to intervene in doxing incidents, specifically, to report doxing. Drawing from the Bystander Intervention Model as a theoretical framework, this study also includes complementary theories: Justice Theory and Social Norms Theory to develop a comprehensive model to explore the dynamics of bystander behavior in the context of doxing. Employing a 4x4 between-subject factorial design, the study examines the effects of others' presence and varying levels of information disclosure on bystanders' intent to take action. By investigating the balance between bystanders' perceived justice and the severity of the doxing incident, the research aims to provide insights that can improve social media content moderation and encourage proactive bystander intervention.

**Keywords:** Doxing, Reporting Behavior, Bystander Intervention, Experiment, Information Disclosure.

### **INTRODUCTION**

Doxing is commonly known as the public release of an individual's personal information online often with the intent to harm the target (Douglas 2016). This phenomenon is recognized as a digital technology-driven issue that can result in severe physical and mental consequences for the targets. Physical consequences include offline harassment and violent attacks, while mental

consequences involve fear and depression (Chen et al. 2018; Eckert and Metzger-Riftkin 2020). In addition to targeting individuals for personal harassment or revenge, doxing is also used as a form of digital vigilantism in collective actions (Davison 2012; Trottier 2019), such as revealing corruption (Demydova), supporting social movements (Bowles 2017; Lau and Lum 2019), and as a tactic in hybrid warfare (Guest 2022; ISD 2022). The increasing number of doxing incidents globally highlights the significant societal impact and complexity of this phenomenon.

In response to the severe consequences of doxing, governments and tech platforms have implemented various countermeasures, including criminalizing doxing (HKSAR 2019; Raftery 2022) and updated policies to regulate doxing behaviors (Bell 2022). However, despite these efforts, doxing continues to persist due to the evolving tactics employed by the doxers and the limitations of content moderation tools (Belanger 2022). The ongoing presence of doxing and the challenges faced by regulators underscore the difficulty of effectively addressing doxing.

Current challenges in mitigating doxing include the overwhelming volume of content (Gillespie 2018), the use of coded language (Citron and Franks 2014), and difficulties in real-time detection algorithms (Crawford and Gillespie 2016). Additionally, platform features like anonymity and rapid information dissemination inadvertently facilitate doxing (Marwick 2017). Given these challenges, algorithmic content moderation alone is insufficient. Instead, effective moderation requires active participation from the online community, specifically through bystanders' intervention. Social media users, as bystanders, can play a crucial role in shaping the outcome of doxing cases by reporting the content or expressing disagreement. Such actions can help mitigate the spread and impact of doxing, complementing the platforms' moderation efforts.

Recognizing the vital role of bystanders in this context, this study aims to explore the factors that influence their willingness to report doxing on social media platforms. This research

will consider both technical and social cognitive factors, as well as unique aspects of doxing identified through a comprehensive literature review (Fang et al. 2023). Therefore, the following research questions are proposed: RQ1: *What factors affect bystanders' willingness to report doxing on social media platforms?* RQ2: *How does the online environment, particularly the presence of others, influence bystanders' willingness to report doxing?* RQ3: *How do different levels of information disclosure affect bystanders' willingness to report doxing?*

### **LITERATURE REVIEW ON DOXING**

Doxing (or “doxxing”) is a neologism as an altered spelling of “dropping ‘docs’” (short for “documents”) (Garber 2014). It is a term that originated from the Usenet in the 90s (Amanda 2013) and refers to the public release of an individual's personal information on the Internet. In recent years, doxing has emerged as a form of digital vigilantism and gained popularity in the mid-2010s (Tiffany 2022), becoming a significant concern for online privacy and safety. Current research on doxing has been explored across various disciplines, each focusing on distinct aspects of the phenomenon. Legal studies have concentrated on developing and evaluating doxing regulations, examining existing laws through case studies (Crompton 2018; Mery 2020), advocating for more robust social media regulations (Yudiana et al. 2022), and proposing new regulatory frameworks from diverse perspectives (Amiruddin et al. 2021; Bei Li 2018; Calabro 2018; Corbridge 2018; Lindvall 2019; MacAllister 2017; McIntyre 2016; Style 2021). Computer science research has primarily focused on technological solutions, such as methods for detecting doxing (Karimi et al. 2022; Snyder et al. 2017) and strategies for safeguarding organizations from doxing attacks (Khanna et al. 2016). In social sciences, scholars have examined the communication strategies employed in doxing (Lee 2020), its impact on political elections (Hansen and Lim 2019), and the psychological harm inflicted on victims (Chen et al.

2018). Additionally, some studies have explored contexts in which doxing is perceived as acceptable, such as exposing racists or other harmful behaviors (Barry 2021).

Research on addressing the issue of doxing has predominantly focused on proposing legal regulations and developing technological solutions. However, no study has examined the problem from the perspective of involving the audience, specifically through the lens of bystanders. This study, therefore, adopts the bystander intervention framework as the basis for our research to explore the role of bystanders in mitigating the impact of doxing incidents.

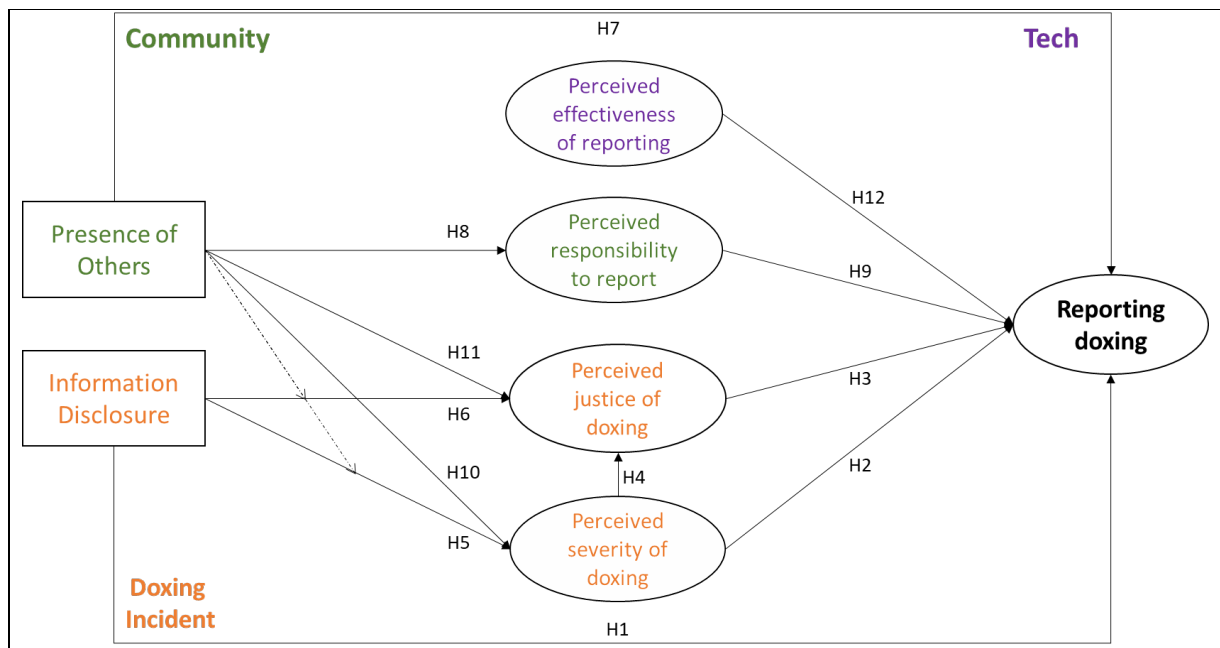
### **THEORETICAL BACKGROUND & HYPOTHESIS DEVELOPMENT**

Figure 1 presents the proposed research model. In the following section, the constructs are categorized into three types: doxing incident-related, online community-related, and technology-related factors. Simultaneously, hypotheses are proposed following the stages of the **Bystander Intervention** Framework (Latané and Darley 1970) while integrating **Justice Theory** (Adams 1965; Bies 1986) and **Social Norms Theory** (Cialdini et al. 1991). The Bystander Intervention model explains that bystander intervention is influenced by four key stages: 1) noticing the event, 2) interpreting it as an emergency, 3) feeling responsible, and 4) knowing how to help. This model serves as the foundation of our research framework.

#### **Doxing Incident Constructs**

When a bystander encounters a doxing incident on a social media platform, the first step in the intervention process is noticing the event. In the context of doxing this step is heavily influenced by **Information Disclosure**, which refers to the type and amount of personal information revealed about the victim. The more detailed and conspicuous the disclosed information, the more likely a bystander is to recognize it as a situation requiring intervention (Patchin and Hinduja 2006).

**H1:** The increase in information disclosure increases the reporting of doxing.



**Figure 1.** Research Model

Once the incident is noticed, the next step is interpreting it as a problem that requires action. This step involves assessing the **Perceived Severity of Doxing**, which reflects the potential harm or danger posed to the victim. High perceived severity increases the likelihood of bystanders viewing the situation as problematic and in need of intervention (Fischer et al. 2011; Levine 1999).

**H2:** Perceived severity of doxing increases the reporting of doxing.

The dual nature of doxing complicates its evaluation. On one hand, it may be seen as justifiable if doxing is exposing wrongdoing, leading to the perception that it is a fair conduct and thus does not require intervention. While, on the other hand, if doxing escalates or causes severe harm, bystanders may view it as unjust and feel compelled to intervene (Chia et al. 2023). This dual perception underscores that **Perceived Justice of Doxing** plays a nuanced role in determining bystanders' actions. Justice Theory (Adams 1965; Bies 1986) suggests that

individuals are motivated to intervene in situations they perceive as unjust, as this challenges their sense of fairness and moral obligation (Tyler 1990). The perception of injustice increases the likelihood that bystanders will report the post.

**H3:** Perceived justice of doxing decreases the reporting of doxing.

**H4:** Perceived severity of doxing decreases the perceived justice of doxing.

At this stage, the level of information disclosure can significantly affect a bystander's perception of both the severity and justice of doxing. When more detailed information is revealed, such as sensitive personal details, or the involvement of innocent individuals, bystanders are more likely to perceive the situation as more severe. Thus, they may begin to view the doxing as unjust, believing that the consequences are disproportionate to the wrongdoing.

**H5:** The increase in information disclosure increases the perceived severity of doxing.

**H6:** The increase in information disclosure decreases the perceived justice of doxing.

### Online Community Constructs

In the Bystander Intervention model, the construct **Presence of Others** is crucial in shaping bystander behavior. Typically, the presence of others who are concerned about the incident (**Discouraging**) leads to diffusion of responsibility, where bystanders assume that others will take action, thereby reducing their own obligation to intervene (Latane and Darley 1968).

**H7a:** Presence of others – discouraging doxing decreases the reporting of doxing.

**H8:** Presence of others – discouraging doxing decreases the perceived responsibility to report doxing.

**H9:** Perceived responsibility to report increases the reporting of doxing.

However, the presence of others in the context of doxing can also take on a different dimension. Besides those concerned about the incident, there is often a **Presence of Others**—

**Encouraging doxing**, who actively support or escalate the situation by sharing the doxing information further or calling for harmful actions. According to Social Norms Theory (Cialdini et al. 1991), individuals' behaviors are influenced by their perceptions of what is considered normal or acceptable within their social group. In doxing incidents, social norms can either discourage or encourage bystander intervention based on the observed behaviors and opinions of others within the community. The existence of these two distinct types of presence—those concerned versus those escalating—can have different impacts on a bystander's decision to report. While the presence of concerned others might reduce individual responsibility through diffusion, the presence of those escalating the incident could possibly provoke a stronger reaction from bystanders to counteract the harm of the incident (Levine and Crowther 2008). Observing others encourage harmful actions intensifies bystanders' perception of severity, shifting the view of doxing from vigilante justice to personal vendettas, prompting stronger reactions against the behavior.

**H7b:** Presence of others – encouraging doxing increases the reporting of doxing.

**H10:** Presence of others – encouraging doxing increases the perceived severity of doxing.

**H11:** Presence of others – encouraging doxing decreases the perceived justice of doxing.

### **Technical Constructs**

Finally, the technological feature of the social media platform influences bystander behavior, particularly through the **Perceived Effectiveness of Reporting** mechanisms. This construct relates to the bystander's beliefs about the efficacy of the platform's reporting system. Even if bystanders notice the incident, interpret it as a problem, and feel responsible, they may still be deterred if they perceive the reporting mechanism as ineffective (Cheung et al. 2020). If

the platform's reporting process is responsive, and reporting is useful, bystanders are more likely to go ahead with reporting the doxing incident (Zuckerman and Reis 1978).

**H12:** Perceived effectiveness of reporting increases the reporting of doxing.

## RESEARCH DESIGN

To empirically test these relationships, we propose a 4 x 4 between-subject factorial design (Table 1) that manipulates the Presence of Others and the level of Information Disclosure.

**Table 1.** Experimental Design

		Information Disclosure			
		Level 1	Level 2	Level 3	Level 4
<b>Presence of Others</b>	None	Group 1	Group 2	Group 3	Group 4
	Discouraging doxing	Group 5	Group 6	Group 7	Group 8
	Encouraging doxing	Group 9	Group 10	Group 11	Group 12
	50% + 50%	Group 13	Group 14	Group 15	Group 16

For Information Disclosure, we created four doxing scenarios on X (formerly known as Twitter), each varying in the amount and type of personal information revealed. For Presence of Others, we manipulated user comments to represent different types of reactions to the doxing post in each scenario. After viewing the doxing scenario, participants will indicate their likelihood of taking various actions, followed by assessments of perceived severity, perceived justice, perceived responsibility to report, and perceived effectiveness of reporting.

## EXPECTED FINDINGS

The expected findings of this study will offer insights into factors influencing bystanders' reporting behaviors in doxing incidents. Inspired by privacy calculus (Laufer and Wolfe 1977), the study posits that bystanders' decisions are influenced by a balance between perceived justice and injustice. The presence of others can further influence this balance, either reinforcing or undermining perceived justice, ultimately shaping bystanders' willingness to report doxing incidents.



## REFERENCES

- Adams, J. S. 1965. "Inequity in Social Exchange," in *Advances in Experimental Social Psychology*. Elsevier, pp. 267-299.
- Amiruddin, N., Manaf, A. A., and Adam, Y. C. 2021. "Trolling and Doxing: Proposal for a Law for Malaysia," *UiTM International Conference on Law & Society 2021 E-Book of Extended Abstract*, pp. 18-20.
- Barry, P. B. 2021. "Doxing Racists," *The Journal of Value Inquiry* (55:3), pp. 457-474.
- Bei Li, L. 2018. "Data Privacy in the Cyber Age: Recommendations for Regulating Doxing and Swatting," *Federal Communications Law Journal* (70:3), pp. 317-328.
- Belanger, A. 2022. "TikTok Can't Stop Users from Doxxing the Supreme Court." from <https://arstechnica.com/tech-policy/2022/06/tiktok-cant-stop-users-from-doxxing-the-supreme-court/>
- Bell, K. 2022. "Meta Will Close a Loophole in Its Doxxing Policy in Response to the Oversight Board." from <https://www.engadget.com/meta-facebook-response-tooversight-board-doxxing-235953823.html>
- Bies, R. J. 1986. "Interactional Justice: Communication Criteria of Fairness," *Research on negotiation in organizations* (1), pp. 43-55.
- Bowles, N. 2017. "How 'Doxing' Became a Mainstream Tool in the Culture Wars." from <https://www.nytimes.com/2017/08/30/technology/doxxing-protests.html>
- Calabro, S. 2018. "From the Message Board to the Front Door: Addressing the Offline Consequences of Race-and Gender-Based Doxxing and Swatting," *Suffolk University Law Review* (51:1), pp. 55-76.
- Chen, Q., Chan, K. L., and Cheung, A. S. Y. 2018. "Doxing Victimization and Emotional Problems among Secondary School Students in Hong Kong," *International Journal of Environmental Research and Public Health* (15:12), pp. 1-8.
- Cheung, M. L., Pires, G., and Rosenberger, P. J. 2020. "The Influence of Perceived Social Media Marketing Elements on Consumer-Brand Engagement and Brand Knowledge," *Asia Pacific Journal of Marketing and Logistics* (32:3), pp. 695-720.
- Chia, S. C., Sun, Y., Lu, F., and Gudmundsdottir, A. 2023. "Doxing, Regulation, and Privacy Protection: Expanding the Behavioral Consequences of the Third-Person Effect," *Asian Journal of Communication* (33:3), pp. 289-307.
- Cialdini, R. B., Kallgren, C. A., and Reno, R. R. 1991. "A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior," in *Advances in Experimental Social Psychology*. Elsevier, pp. 201-234.
- Citron, D. K., and Franks, M. A. 2014. "Criminalizing Revenge Porn," *Wake Forest L. Rev.* (49), p. 345.
- Corbridge, Å. 2018. "Responding to Doxing in Australia: Towards a Right to Informational Self-Determination," *University of South Australia Law Review* (3), pp. 1-28.
- Crawford, K., and Gillespie, T. 2016. "What Is a Flag For? Social Media Reporting Tools and the Vocabulary of Complaint," *New Media & Society* (18:3), pp. 410-428.
- Crompton, M. 2018. "Doxing in Australia: A Practitioner's Perspective," *University of South Australia Law Review* (3), pp. 29-32.
- Davison, J. 2012. "Online Vigilantes: Is 'Doxing' a Neighbourhood Watch or Dangerous Witch Hunt?." *CBC News*, from <https://www.cbc.ca/news/science/online-vigilantes-is-doxing-a-neighbourhood-watch-or-dangerous-witch-hunt-1.1132015>

- Demydova, V. 2021. "Doxing as a Form of Online Activism: Case of Alexei Navalny's Film a Palace for Putin," *Electronic Turkish Studies* (16:3), pp. 909-929.
- Douglas, D. M. 2016. "Doxing: A Conceptual Analysis," *Ethics and Information Technology* (18:3), pp. 199-210.
- Eckert, S., and Metzger-Riftkin, J. 2020. "Doxxing, Privacy and Gendered Harassment. The Shock and Normalization of Veillance Cultures," *M&K Medien & Kommunikationswissenschaft* (68:3), pp. 273-287.
- Fang, Y., Risius, M., and Cheung, C. 2023. "Understanding the Current State of Knowledge and Future Directions of Doxing Research: A Social Cognitive Theory Perspective," *Proceedings of the 56th Hawaii International Conference on System Sciences (HICSS-56)*.
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., Heene, M., Wicher, M., and Kainbacher, M. 2011. "The Bystander-Effect: A Meta-Analytic Review on Bystander Intervention in Dangerous and Non-Dangerous Emergencies," *Psychological Bulletin* (137:4), p. 517.
- Garber, M. 2014. "Doxing: An Etymology," in: *The Atlantic*.
- Gillespie, T. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press.
- Guest, P. 2022. "Telegram Has a Serious Doxing Problem." from <https://www.wired.co.uk/article/telegrams-doxing-problem>
- Hansen, I., and Lim, D. J. 2019. "Doxing Democracy: Influencing Elections Via Cyber Voter Interference," *Contemporary Politics* (25:2), pp. 150-171.
- HKSAR. 2019. "Interim Injunction Order Granted by the High Court to Restrain Doxxing and Harassment against Police Officers and Their Families."
- ISD. 2022. "Tales from the Underside: A Kremlin-Approved Hack, Leak & Doxxing Operation." from [https://www.isdglobal.org/digital\\_dispatches/tales-from-the-underside-a-kremlin-approved-hack-leak-doxxing-operation/](https://www.isdglobal.org/digital_dispatches/tales-from-the-underside-a-kremlin-approved-hack-leak-doxxing-operation/)
- Karimi, Y., Squicciarini, A., and Wilson, S. 2022. "Automated Detection of Doxing on Twitter," *Proceedings of the ACM on Human-Computer Interaction*, pp. 1-24.
- Khanna, P., Zavorsky, P., and Lindskog, D. 2016. "Experimental Analysis of Tools Used for Doxing and Proposed New Transforms to Help Organizations Protect against Doxing Attacks," *Procedia Computer Science* (94), pp. 459-464.
- Latane, B., and Darley, J. M. 1968. "Group Inhibition of Bystander Intervention in Emergencies," *Journal of Personality and Social Psychology* (10:3), p. 215.
- Latané, B., and Darley, J. M. 1970. "The Unresponsive Bystander: Why Doesn't He Help?," (*No Title*).
- Lau, C., and Lum, A. 2019. "Telegram Social Media Channel Used by Protesters for Doxxing Hong Kong Police Suspended after Leaking Thousands of Pictures and Videos of Officers and Their Families." from <https://www.scmp.com/news/hong-kong/politics/article/3037001/telegram-social-media-channel-used-protesters-doxxing-hong>
- Laufer, R. S., and Wolfe, M. 1977. "Privacy as a Concept and a Social Issue: A Multidimensional Developmental Theory," *Journal of Social Issues* (33:3), pp. 22-42.
- Lee, C. 2020. "Doxxing as Discursive Action in a Social Movement," *Critical Discourse Studies* (19:3), pp. 326-344.

- Levine, M. 1999. "Rethinking Bystander Nonintervention: Social Categorization and the Evidence of Witnesses at the James Bulger Murder Trial," *Human Relations* (52), pp. 1133-1155.
- Levine, M., and Crowther, S. 2008. "The Responsive Bystander: How Social Group Membership and Group Size Can Encourage as Well as Inhibit Bystander Intervention," *Journal of Personality and Social Psychology* (95:6), p. 1429.
- Lindvall, A. J. 2019. "Political Hacktivism: Doxing & the First Amendment," *Creighton Law Review* (53), pp. 1-14.
- MacAllister, J. M. 2017. "The Doxing Dilemma: Seeking a Remedy for the Malicious Publication of Personal Information," *Fordham Law Review* (85:5), pp. 2451-2483.
- Marwick, A. E. 2017. "Scandal or Sex Crime? Gendered Privacy and the Celebrity Nude Photo Leaks," *Ethics and Information Technology* (19), pp. 177-191.
- McIntyre, V. 2016. "Do (X) You Really Want to Hurt Me: Adapting Lied as a Solution to Doxing by Reshaping Intent," *Tulane Journal of Technology and Intellectual Property* (19), pp. 111-134.
- Mery, H. C. 2020. "The Dangers of Doxing and Swatting: Why Texas Should Criminalize These Malicious Forms of Cyberharassment," *St. Mary's Law Journal* (52:3), pp. 905-944.
- Patchin, J. W., and Hinduja, S. 2006. "Bullies Move Beyond the Schoolyard: A Preliminary Look at Cyberbullying," *Youth Violence and Juvenile Justice* (4:2), pp. 148-169.
- Raftery, B. 2022. "Legislation to Protect the Home Addresses of Judges and Other Court Officials." from <https://www.ncsc.org/information-and-resources/trending-topics/trending-topics-landing-pg/legislation-to-protect-the-home-addresses-of-judges-and-other-court-officials2>
- Snyder, P., Doerfler, P., Kanich, C., and McCoy, D. 2017. "Fifteen Minutes of Unwanted Fame: Detecting and Characterizing Doxing," *Proceedings of the 2017 Internet Measurement Conference*, pp. 432-444.
- Styple, K. 2021. "Institutional Doxing and Attribution: Searching for Solutions to a Law-Free Zone," *Georgia Journal of International and Comparative Law* (50:1), pp. 211-239.
- Tiffany, K. 2022. "Doxxing Means Whatever You Want It To." from <https://www.theatlantic.com/technology/archive/2022/04/doxxing-meaning-libs-of-tiktok/629643/>
- Trottier, D. 2019. "Denunciation and Doxing: Towards a Conceptual Model of Digital Vigilantism," *Global Crime* (21:3-4), pp. 196-212.
- Tyler, T. R. 1990. "10 Justice, Self-Interest, and the Legitimacy of Legal and Political Authority," *Beyond self-interest*, p. 171.
- Yudiana, T. C., Rosadi, S. D., and Priowirjanto, E. S. 2022. "The Urgency of Doxing on Social Media Regulation and the Implementation of Right to Be Forgotten on Related Content for the Optimization of Data Privacy Protection in Indonesia," *Padjadjaran Jurnal Ilmu Hukum (Journal of Law)* (9:1), pp. 24-45.
- Zuckerman, M., and Reis, H. T. 1978. "Comparison of Three Models for Predicting Altruistic Behavior," *Journal of Personality and Social Psychology* (36:5), p. 498.