How the general public and security researchers assess moral dilemmas in computer security research

Nicole Dircksen
Paderborn University

Anna Lena Rotthaler
Paderborn University

Tadayoshi Kohno
Georgetown University

Yasemin Acar
Paderborn University

Abstract

Researchers in Computer Security sometimes face ethically unclear situations, in which there are multiple rights and responsibilities to consider which conflict with each other—especially when disclosing vulnerabilities. Public opinion—support or backlash—can be an important decision factor in such moral decision-making. To find out whether the general public and computer security researchers align in their assessment of ethically unclear situations, we plan a survey with general public and security researchers. The participants will be asked to choose one option in moral dilemmas which could arise in vulnerability disclosure.

1 Introduction

Ethical research is a core principle of conducting research and security researchers aim to fulfill this aspiration. Sometimes, there are situations where it is not clear what the ethically correct decision is or where researchers disagree on the ethical decision. The general public might be one factor influencing the decision-making of researchers. While one might argue that "what others think" should not be relevant for ethical decision-making, the opinion of the general public can have valid influence [4, 8]. If there are tensions between ethical and legal behavior or if researchers are not in agreement about the ethically correct behavior, public opinion can be the deciding factor or at least another important aspect of the decision-making. For example, researchers (and the general public) may consider some specific research as the right thing to do, but this research is illegal or in gray areas of the law. The researchers have to consider potential legal consequences in their decision about pursuing the research and the effects on themselves, their future and similar research [12]. If public opinion and public pressure support the researchers, and even might lead to regulatory changes which would further ease this specific research, the ethical decision-making of the researchers could change. There are many examples where researchers acting in good faith were threatened with legal consequences [1].

The public opinion could also influence the ethical decision-making in cases where researchers think that some research is ethical, but the general public or even other researchers assess it differently. Then, still conducting (and publishing) this research could result in backlash within the research community and by the public opinion. This could lead to harm to the researchers themselves (by threats) or their time, energy, and motivation to conduct further research, but could also harm the trust and funding in further research or even the (legal) regulation of it [6, 14].

Therefore, the public opinion and possible consequences can play a role in the decision-making itself and it is important to gain information on the assessment of the public and researchers on ethically unclear situations.

With this study we want to answer the following research questions:

1. **RQ1**: Does the general public, in the EU and US, assess ethically ambiguous situations similarly to security researchers? If any, what are the (cultural) differences?
2. **RQ2**: What are possible reasons for different assessments?

The focus of this work will be on moral dilemmas regarding vulnerability disclosures.

2 Background & Related Work

General Public and Public Opinion. While the general public includes all people (of a society or a country), the term

USENIX Symposium on Usable Privacy and Security (SOUPS) 2026.
August 23–26, 2026, Hannover, Germany.

public opinion describes the aggregated opinion (including view, attitudes, and beliefs) of the public, regarding a specific topic [7]. Public opinion is often measured in polls or surveys but it is more than that. Some include the process of communication and interaction – formed and expressed by people in conversations, papers, books, broadcast media, or nowadays on social media – in the definition of public opinion [7, 11]. What is expressed by people, the "written or voiced public opinion", generally differs from the "average or aggregated public opinion", as some people keep their opinion to themselves whereas others are vocal about it. Public opinion is for example important in the politics of democracies, in commerce (what is "in", modern, or healthy) or in health decisions (what is safe, who is trustworthy) [5, 22, 27]. Sometimes, public opinion erupts in sudden critique of someone or something, increasing the pressure and media attention onto a (perceived) problem or problematic behavior. This can happen as justified, objective criticism, but in some cases such backlash escalates into public shaming and denunciation where the discussion is no longer based on rational arguments but on personal insults or threats [29]. Justified backlash and so-called online firestorms can both have a big influence and lead to serious consequences or changes, where political figures resign, people are fired, or products or advertisements are withdrawn. Researchers are not spared from getting feedback through public opinion, sometimes even in the form of online firestorms [25, 30].

Ethics in Computer Security Research. Multiple guidelines in security research advise researchers regarding ethical research. The Menlo Report (2012) is a framework for ethical guidelines in computer and information security research, released by the US Department of Homeland Security [3, 19]. There are also the IEEE Code of Ethics and the ACM Code of conduct [2, 17]. Even with guidelines it is not always trivial to do the morally correct thing. For instance, the research community discussed or discusses (in recent publications), among others, if or how it is ethical to use Internet-wide scans to study the Internet [9, 13], to study stolen or publicly available data [10, 28], or to train and use (blackbox) machine learning models [16, 18]. The disclosure of vulnerabilities in software or hardware is another topic with potential ethical questions [26, 31]. The rights and needs of different stakeholders (e. g., users, vendors, researchers) need to be considered. Disclosing a vulnerability (possibly) also allows adversaries to exploit it but makes the vendors and users aware and therefore enables mitigation. To enable a smooth mitigation, there is a process called "Responsible Vulnerability Disclosure", in which the vendor is privately notified about a vulnerability, fixes it, and then the (now fixed) vulnerability can be made public [15]. This ideal process is not always possible, for example when the vendor is unresponsive or the software or hardware with the vulnerability is used in many other systems, where an update or exchange needs to happen.

Moral Dilemmas. A moral dilemma is typically a situation in which one agent has two courses for action, and neither is ideal. Executing one option makes it impossible to do the other, there is a conflict within the options and the agent therefore needs to violate one value or duty to fulfill the other [21, 23]. While dilemmas can be constructed theoretically as a basis for discussion (hypothetical dilemmas), dilemmas in real-life are often more complicated: they can have more than two courses of action and not all consequences might be foreseeable.

3 Methodology

We will present moral dilemmas regarding vulnerability disclosure to both the general public and security researchers to find similarities and differences and ask for reasons for and against the available options.

Survey and Scenario Generation. The scenarios created by Kohno et. al [20] will be used as a starting point. As these scenarios are targeted at security researchers and this survey is also aimed at the general public, the scenarios will be adapted, including simplification to an 8th grade reading level and shortening. While no participant is forced to answer any question in the survey, the scenarios are constructed such that the actor(s) in the dilemma have to choose one option. The scenarios will be clustered in scenario families, where a base scenario is only slightly altered in their setting to create scenario variants. The variations will focus on different elements targeting factors of moral intensity, namely the probable magnitude of consequences, social consensus, and proximity [24]. The modification regarding the consequences can be done in multiple ways (increased or decreased probabilities; higher or lower harm; higher or lower affected parties or groups) and mostly changes the objective pro-contra dynamic of the scenario. It is more complicated to naturally incorporate modifications regarding social consensus or proximity in this within-subject design and will take lower priority.

Free text fields will be used to ask the participants for reasons or explanations of a decision. After constructing a pilot survey, we will obtain approval of our institution's ethical review board and test the survey for comprehensibility and length.

Afterwards, we will run the survey on both the general public—across the EU and US—and security researchers. If this poster is accepted, we would love to use it as an opportunity to invite SOUPS participants to fill in our survey.

Data Analysis. After data cleaning and de-identification, the data analysis will focus on two parts. The selected options per scenario will be analyzed for similarities, differences, or outliers between the groups (security researchers and general public in US and EU). Afterwards, qualitative coding will be applied on the free text answers to get insights into the potential reasons for similar or different answers.

References

[1] Research Threats: Legal Threats Against Security Researchers. https://github.com/disclose/research-threats. Accessed 17/3/2026.

[2] Association for Computing Machinery. ACM Code of Ethics and Professional Conduct. https://www.acm.org/code-of-ethics, 2018. Accessed 19/3/2026.

[3] Michael Bailey, David Dittrich, Erin Kenneally, and Doug Maughan. The menlo report. *IEEE Security & Privacy*, 10(2):71–75, 2012. https://doi.org/10.1109/MSP.2012.52.

[4] Pascal Borry, Paul Schotsmans, and Kris Dierickx. What is the role of empirical research in bioethical reflection and decision-making? an ethical analysis. *Medicine, Health Care and Philosophy*, 7(1):41–53, 2004. https://doi.org/10.1023/B:MHEP.0000021844.57115.9d.

[5] Gracia Brückmann and Thomas Bernauer. What drives public support for policies to enhance electric vehicle adoption? *Environmental Research Letters*, 15(9), 2020. https://doi.org/10.1088/1748-9326/ab90a5.

[6] danah boyd. Untangling research and practice: What facebook's "emotional contagion" study teaches us. *Research Ethics*, 12(1):4–13, 2016. https://doi.org/10.1177/1747016115583379.

[7] W. Phillips Davidson. public opinion. https://www.britannica.com/topic/public-opinion, 2026. Accessed 17/3/2026.

[8] Eleanor Drinkwater, Elva J. H. Robinson, and Adam G. Hart. Keeping invertebrate research ethical in a landscape of shifting public opinion. *Methods in Ecology and Evolution*, 10(8):1265–1273, 2019. https://doi.org/10.1111/2041-210X.13208.

[9] Saiid El Hajj Chehade, Florian Hantke, and Ben Stock. 403 forbidden? ethically evaluating broken access control in the wild. In *2025 IEEE Symposium on Security and Privacy (SP)*, pages 3218–3235, 2025. https://doi.org/10.1109/SP61157.2025.00252.

[10] Martyna Gliniecka. The ethics of publicly available data research: A situated ethics framework for reddit. *Social Media + Society*, 9(3), 2023. https://doi.org/10.1177/20563051231192021.

[11] Carroll J. Glynn, Susan Herbst, Mark Lindeman, Garret J. O'Keefe, and Robert Y. Shapiro. *Public Opinion*. Routledge, New York London, 2016.

[12] Roger Grimes. What is the best way to report a found vulnerability to a vendor without incurring any risk? https://www.linkedin.com/pulse/what-best-way-report-found-vulnerability-vendor-witho 2022. Accessed 29/5/2026.

[13] Florian Hantke, Sebastian Roth, Rafael Mrowczynski, Christine Utz, and Ben Stock. Where are the red lines? towards ethical server-side scans in security and privacy research. In *2024 IEEE Symposium on Security and Privacy (SP)*, pages 4405–4423, 2024. https://doi.org/10.1109/SP54263.2024.00104.

[14] Jaigris Hodson, Victoria O'Meara, Andrea Galizia, and Chandell Gosse. Quietly coping: The impact of online abuse on research innovation and knowledge workers in research and public education. In *2021 IEEE International Symposium on Technology and Society (ISTAS)*, pages 1–5, 2021. https://doi.org/10.1109/ISTAS52410.2021.9629206.

[15] Allen D Householder, Garret Wassermann, Art Manion, and Chris King. The cert guide to coordinated vulnerability disclosure. Technical report, 2017. https://apps.dtic.mil/sti/tr/pdf/AD1046659.pdf.

[16] Changwu Huang, Zeqi Zhang, Bifei Mao, and Xin Yao. An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence*, 4(4):799–819, 2023. https://doi.org/10.1109/TAI.2022.3194503.

[17] Institute of Electrical & Electronics Engineers. IEEE Code of Ethics. https://www.ieee.org/about/corporate/governance/p7-8. Accessed 19/3/2026.

[18] Haseeb Javed, Hafiz Abdul Muqeet, Tahir Javed, Atiq Ur Rehman, and Rizwan Sadiq. Ethical frameworks for machine learning in sensitive healthcare applications. *IEEE Access*, 12:16233–16254, 2024. https://doi.org/10.1109/ACCESS.2023.3340884.

[19] Erin Kenneally and David Dittrich. The menlo report: Ethical principles guiding information and communication technology research. *SSRN Electronic Journal*, 2012. http://dx.doi.org/10.2139/ssrn.2445102.

[20] Tadayoshi Kohno, Yasemin Acar, and Wulf Loh. Ethical frameworks and computer security trolley problems: Foundations for conversations. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 5145–5162, Anaheim, CA, 2023. USENIX Association. https://www.usenix.org/conference/usenixsecurity23/presentation/kohno.

[21] E. J. Lemmon. Moral dilemmas. *The Philosophical Review*, 71(2):139–158, 1962. http://www.jstor.org/stable/2182983.

[22] Jill J McCluskey, Nicholas Kalaitzandonakes, and Johan Swinnen. Media coverage, public perceptions, and consumer behavior: Insights from new food technologies. *Annual review of resource economics*, 8(1):467–486, 2016. https://doi.org/10.1146/annurev-resource-100913-012630.

[23] Terrance McConnell. Moral Dilemmas. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2024 edition, 2024. https://plato.stanford.edu/archives/spr2024/entries/moral-dilemmas/.

[24] Joan M McMahon and Robert J Harvey. An analysis of the factor structure of jones' moral intensity construct. *Journal of Business Ethics*, 64(4):381–404, 2006. https://doi.org/10.1007/s10551-006-0006-5.

[25] Richard Medina, Emily Nicolosi, Collin Riley, and Phoebe McNeally. Reflections on hateful responses to research on hate in the united states. *Public Understanding of Science*, 30(6):797–806, 2021. PMID: 33641493, https://doi.org/10.1177/0963662521994698.

[26] Giovane C. M. Moura and John Heidemann. Vulnerability disclosure considered stressful. *SIGCOMM Comput. Commun. Rev.*, 53(2):2–10, 2023. https://doi.org/10.1145/3610381.3610383.

[27] Geoboo Song. Understanding public perceptions of benefits and risks of childhood vaccinations in the united states. *Risk Analysis*, 34(3):541–555, 2014. https://onlinelibrary.wiley.com/doi/pdf/10.1111/risa.12114.

[28] Daniel R. Thomas, Sergio Pastrana, Alice Hutchings, Richard Clayton, and Alastair R. Beresford. Ethical issues in research using datasets of illicit origin. In *Proceedings of the 2017 Internet Measurement Conference*, IMC '17, page 445–462, New York, NY, USA, 2017. Association for Computing Machinery. https://doi.org/10.1145/3131365.3131389.

[29] Daniel Trottier, Qian Huang, and Rashid Gabdulhakov. *Digital Media, Denunciation and Shaming: The Court of Public Opinion*. Taylor & Francis, 2025. https://doi.org/10.4324/9781003453017.

[30] George Veletsianos, Shandell Houlden, Jaigris Hodson, and Chandell Gosse. Women scholars' experiences with online harassment and abuse: Self-protection, resistance, acceptance, and self-blame. *New Media & Society*, 20(12):4689–4708, 2018. https://doi.org/10.1177/1461444818781324.

[31] Mojtaba Zaheri, Yossi Oren, and Reza Curtmola. Targeted deanonymization via the cache side channel: Attacks and defenses. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 1505–1523, Boston, MA, 2022. USENIX Association. https://www.usenix.org/conference/usenixsecurity22/presentation/zaheri.

4 Appendix

4.1 Example scenario

Vulnerability in a medical device.
Company A produced a lifesaving wireless implantable medical device. They produced too many devices before they were needed. As a result, Company A went bankrupt and closed.

Doctors continued to implant the devices. There are no similar devices which are produced by other companies.

Researchers found a problem, a vulnerability, within the devices. This problem could allow cybercriminals to cause harm to the patients. However, the chance that anyone would exploit this problem is small, but not zero. Because Company A does not exist anymore, the problem cannot be fixed by them.

The choice for the researchers:
Do not tell anyone about the problem:

- Patients do not know that their device is vulnerable.
- Patients keep their device and have better health.
- New patients get the device implanted and have better health.

Publish the findings:

- The patients, the public and therefore adversaries can know about the problem.
- Patients can decide if they want to receive the device or have it removed.
- It can harm the health of the patients if they do not receive the device or have it removed.
- It can hurt patients or their relatives to know that there is a vulnerable device in the body. Even when the chance is very small, that anyone would use the problem to cause harm.

Variations.

- Company A still exists but has abandoned this device and ignores the notifications of the researchers
- Company A still exists and continues to develop the device, but they cannot be contacted (all mails cannot be delivered)

- It is likely that adversaries would exploit the problem to harm or blackmail patients
- It is likely that adversaries are already exploiting the problem
- The vulnerability allows to remotely read out the model's name and type of the device
- The vulnerability allows to remotely read out personal health data (measured data, name of patient or doctor)
- The vulnerability allows to remotely shutdown the device
- The vulnerability allows to remotely control the device (change settings, trigger functions)
- The participant or a persons close to them has the device implanted