Users are not the Enemy: How Organisations Allocate Blame Through Security Practices

Ivan Flechais

Department of Computer Science,
University of Oxford, OX1 3QD
ivan.flechais@cs.ox.ac.uk

Abstract. Cybersecurity controls are often presented as purely technical measures, yet they often redistribute power: they make activity visible, define permissible conduct, and legitimise sanctions when rules are breached. Through an approach inspired by critical theory, this paper offers an account of organisational endpoint monitoring and the behavioural programmes that typically accompany it. We show how endpoint monitoring produces visibility asymmetries between what employees and organisations can see, embeds routine suspicion through classification and policy triggers, and supports responsibilisation through compliance and phishing metrics that render user failure administratively convenient. These dynamics can shift security labour and blame on to employees and undermine security outcomes by incentivising concealment and workaround practices. We argue that usable security offers the means to challenge such practices, concluding that monitoring must be burden-aware, contestable, and recovery-centric, with explicit constraints on secondary use and punitive governance.

Keywords: Usable Security · Endpoint Monitoring · Organisational Cybersecurity · Security Governance · User Blame.

1 Introduction

Cybersecurity, data security, and privacy all have a common focus on the exercise of power and control: whether by nations, organisations, or individuals. The motivation may stem from a variety of different concerns (e.g. safeguarding sensitive defense, commercial, or personal data), but the means of protection invariably involve solutions that shift the balance of power, e.g. surveillance and endpoint monitoring; setting and revoking access privileges; mandatory access control; strict identification and authentication protocols; compliance with training, standards and policies; controlled information flows; data retention requirements and lawful access; etc.

Within the space of security and privacy, different stakeholders continually engage in the negotiation and contestation of power, which ebbs and flows according to social and technological developments. For example, security dual use explains that controls can be advantageous for defenders but that this can also disadvantage them if these are used by their adversaries. This is particularly discussed in cryptography, where strong cryptography has the potential to be used both by defenders and their adversaries. When technical developments affect the status quo, such as the widespread adoption of convenient communication tools that enable end-to-end encryption, debates arise questioning whether these tools should be redesigned to enable surveillance for national security, policing, or child protection purposes. Rogaway summarises this in his opening statement as "Cryptography rearranges power" [11], but this can be generalised to note that security and privacy are inextricably concerned with the power dynamics of socio-technical systems.

In the previous example, one of the notable aspects of this particular debate is that it recently arose out of a decision to make end-to-end encryption a default option in widely used communication software. The added convenience and ease of use is what ultimately drove the concerns leading to the debate, and helps to highlight the role of usable security in this space.

It is also interesting to note that usable security has deep roots in cryptography – particularly in Kerckhoff's work from 1883 [8], which elegantly frames it as one of six key principles for encryption systems: *"it is necessary, given the circumstances that command its application, that the system be easy to use, requiring neither mental strain nor the knowledge of a long series of rules to observe."* But we largely owe today's recognition of the importance of usable security to Adams & Sasse's work: "Users are not the Enemy" [1]. This paper set the scene for a new paradigm of security centered on the everyday realities of work, showing how security mechanisms and policies frequently fail because they conflict with what users must do to get their jobs done. Since then, the field of usable security has only grown, exploring e.g. user interface design [13], innovative authentication mechanisms [3], secure development lifecycles [5], security awareness campaigns [2], and many more. And yet, underpinning all this is a simple idea: users are not solely responsible for security *system* failures. And so if cryptography rearranges power, then *usable security rearranges blame.*

While end-to-end encryption debates make the politics of security obvious at a societal scale, endpoint monitoring makes the same politics visible inside organisations. In this paper, we draw from critical theory to frame security and privacy as power practices and apply this to endpoint monitoring and the behavioural programmes that commonly accompany it (training, policy attestation, simulated phishing). Inspired by critical theory, we contribute an analysis of endpoint monitoring that identifies recurring power dynamics (visibility asymmetries, classification & suspicion, discipline & responsibilisation, burden distribution, and contestability & recourse), and conclude with a usable security perspective on monitoring regimes that promote legitimacy, contestability, burden fairness, and recovery-centric design.

2 Critical theory applied to Endpoint Monitoring

2.1 Critical theory

Critical theory is best treated as a mode of inquiry rather than a single doctrine: it is explicitly normative, reflexive about the social conditions of knowledge, and oriented toward critique that identifies how domination is produced, sustained, and how it might be reduced through plausible alternatives. In its canonical articulation, this contrasts with "traditional" theory that presents itself as neutral description; critical theory instead asks how commonplace arrangements become stabilised as necessary, rational, or inevitable, often in ways that privilege some interests and externalise costs onto others [7,6]. Applied work typically operationalises critical theory through a set of characteristic moves: foregrounding power and interests, analysing legitimation and ideology (the narratives and metrics that justify practices), treating arrangements as historically contingent rather than natural, and maintaining an orientation toward actionable alternatives rather than critique for its own sake.

Critical theory has already been applied to security as an object of study by treating it not as a self-evident technical necessity, but as a social and organisational practice that produces categories, authorities, and exclusions. Within cybersecurity specifically, there have been calls for a "critical cybersecurity" [4] that scrutinises where cybersecurity is enacted, by whom, and with what consequences for those subject to its practices. Related work [10] shows how widely used practitioner methodologies, such as the Cyber Kill Chain, embed predetermined categories and indicators that shape how threats are known and acted upon, demonstrating how cybersecurity "makes politics": by defining what is dangerous and not, thus dominating the discourse around security practice. In information security ethics, critical theory has also been used to surface collective and organisational issues that individualistic ethical framings tend to miss [12], making it a useful lens for examining institutional security controls as governance arrangements rather than merely technical safeguards.

In this paper we examine endpoint monitoring not only as a technical control but as a governance activity that reorganises visibility, classification, accountability, effort, and contestability. Concretely, (i) we examine who gains visibility and control and who bears burdens and risk, (ii) we analyse how compliance framings and behavioural metrics legitimise particular power arrangements, (iii) we examine how classification and enforcement decisions are made contestable (or not), and (iv) we argue that user failure narratives largely serve to displace responsibility and that usable security serves as a clear redress for this.

2.2 Endpoint Monitoring

Endpoint monitoring refers to the collection and use of telemetry from endpoint devices to detect, investigate, and prevent security incidents, e.g. process execution, file activity, network connections, peripheral usage, and data movement.

Typical deployments combine Endpoint Detection and Response (EDR) capabilities to look for evidence of threats, Data Loss Prevention (DLP) controls to prevent or audit sensitive data movement, and centralised logging for forensics and compliance.

Organisations adopt endpoint monitoring to help with security operational needs: faster detection of compromise, ransomware containment, incident reconstruction, and evidence production for compliance and investigations. Monitoring is also attractive because it promises oversight across heterogeneous devices and user behaviours, especially under remote and hybrid work arrangements.

At the same time, endpoint monitoring is a governance infrastructure: it expands organisational visibility into work practices, converts activity into auditable traces, and enables enforcement actions such as quarantining files, isolating devices, throttling access, or escalating investigations. These actions are frequently automatically triggered by policy and classification rather than by context-aware judgement, making the lived experience of security inseparable from institutional oversight. Moreover, telemetry collected for security can be repurposed for performance management or compliance surveillance, intensifying power asymmetries and impacting employee trust.

Endpoint monitoring is commonly paired with security awareness and training programmes that define norms and produce compliance metrics, e.g., training, policy attestations, and simulated phishing. Together, technical telemetry and behavioural measurement create a culture where accountability is often anchored in measurable compliance rather than in systemic conditions (workload, tooling constraints, threats). This socio-technical coupling of monitoring tools and behavioural security programmes is central to how blame is allocated when breakdowns occur.

2.3 Critical theory reading of Endpoint Monitoring

Illustrative vignette: *A developer runs a legitimate dependency scanner that triggers an EDR alert resembling credential dumping. The endpoint is automatically isolated from the network. The incident ticket requests a justification and evidence that the tool is approved. Work stops for several hours; the employee worries about reputational impact. The event is later closed as benign, but the employee's subsequent workflow changes: they avoid running similar tools, or they run them on unmanaged devices to prevent future disruption.*

Visibility asymmetry and the production of auditable work. Endpoint monitoring increases the organisation's ability to see and interpret employee activity while employees often have limited visibility into what is collected, how it is interpreted, and what secondary uses may occur. This asymmetry produces a shift in evidentiary power: monitoring data becomes the "true" account of events. When a security incident is investigated, the monitoring data can override contextual explanations, with significant consequences for how responsibility is assigned.

Classification and suspicion as routine governance. Most endpoint monitoring implementations govern through classification: alert thresholds, behavioural baselines, and policy triggers define what counts as suspicious. While classification is an operational necessity, it also introduces a structural risk: errors and false positives can be experienced as accusations. Where the classification process is opaque, employees cannot reliably learn what behaviours are safe, and the burden shifts to them to prove legitimacy in a system that pre-judges them through risk categories.

Discipline and responsibilisation through training and simulation. Training and simulated phishing are often justified as awareness raising, yet they can operate as disciplinary instruments that produce measurable failures and moralised narratives, which can in turn discourage reporting and incentivise concealment when employees anticipate reputational or managerial consequences. These programmes shift the responsibility for systemic vulnerabilities on to employees, for example blaming users for clicking on a phishing link instead of looking at contributory factors such as weak email filtering, deliberate deception, high workload, and workplace urgency. A critical-theory perspective treats security behaviour not as an individual trait but as an outcome of socio-technical conditions. Depending on the design of the monitoring implementation, this can either be improved (by supporting safe work) or worsened (by increasing fear and concealment).

Burden distribution and the emergence of coping strategies. Endpoint monitoring regimes impose ongoing security friction and labour on employees: interruptions, blocked actions, justifications for exceptions, and anxiety about being flagged. When this labour conflicts with organisational priorities (deadlines, service levels, collaboration), predictable coping strategies emerge: using alternative channels, delaying updates, bypassing controls, or avoiding reporting minor mistakes. These are often labelled noncompliance, but a critical view treats them as diagnostic signals of misfit between governance demands and work realities.

Contestability, recourse, and the legitimacy of security decisions. Legitimacy hinges on whether security decisions are contestable. High-impact actions such as device isolation, account suspension, and investigation escalation, should have proportionate recourse: explanations that are actionable, rapid remediation pathways, and human review triggers. Without these, security becomes experienced as arbitrary gatekeeping, which undermines trust and reduces reporting and collaboration with security teams. Contestability is therefore an important consideration for both ethical and instrumental ends.

2.4 Usable Security for Endpoint Monitoring

Endpoint monitoring exemplifies how many contemporary security solutions function by increasing visibility of system and user activity, enabling classification, and creating enforceable governance constructs. Monitoring infrastructures are designed to make actions visible and auditable, which makes it easy to then attribute incidents to individual actions. This can produce a drift from systemic inquiry (identifying what happened) toward person-centred accountability (looking for someone to blame). A critical evaluation of this drift positions this as an organisationally convenient narrative that legitimises expanded monitoring and responsibilisation.

Concretely, usable security contributes methods and evaluation criteria for redesigning such systems, e.g. identifying where friction is unavoidable, where it is wasteful, and how to create recovery paths that preserve security goals under realistic work constraints. Usable security provides a practical approach by treating coping strategies and recurrent "user failures" as evidence of misfit. Under endpoint monitoring, usable security's most consequential intervention is not to make alerts nicer, but reassigning responsibility to designable properties of the regime: burden allocation, exception handling, and contestability. In this sense, usable security "rearranges blame" by turning moralised behavioural accounts into actionable socio-technical redesign tasks, thereby shifting responsibility back toward institutions that have the capacity to change the system.

Finally, legitimacy is not separable from effectiveness. Monitoring regimes that are experienced as punitive or inscrutable tend to reduce reporting and encourage concealment, resulting in undermined detection and response. Designing for contestability and humane recovery is therefore part of maintaining operational security, not merely an ethical aspiration.

3 Discussion

Endpoint monitoring is usually justified as an operational necessity (detect compromise, contain ransomware, support forensics), but our analysis suggests it also functions as a governance regime that reshapes visibility, classification, and accountability. This dual role creates a recurring tension: controls that aim to reduce organisational risk can also produce employee insecurity (stress, arbitrariness, fear of sanction, loss of agency), and those experiences can undermine security outcomes by discouraging reporting and incentivising workarounds. This specific issue is emblematic of the tensions that usable security perspectives address: by focussing on the user perspective, usable security looks for *emancipatory* options that challenge entrenched positions. Emancipation is defined as people's ability to "achieve their potential to a greater degree", following Klein and Huynh [9], and is a central characteristic of critical theory. We translate this framing into implications for the design and governance of endpoint monitoring, and outline future research directions.

3.1 Implications for endpoint monitoring

Designing an endpoint monitoring solution requires a careful balancing act between protection, productivity, and usability. We propose the following recommendations to address the issues identified:

Measure burden alongside risk. Monitoring programmes typically optimise detection metrics (alert volume, time-to-respond) while treating employee security labour as external: interruptions, time lost to blocks/quarantines, false-positive handling, and stress about being flagged. When burden is unmanaged, predictable adaptations emerge (avoidance, unmonitored channels, concealment). A practical approach is to track burden metrics (e.g., high-impact automated actions per user, false-positive resolution time, time-to-remediate) and use these as design constraints rather than as post-hoc, external costs.

Contestability-by-design for high-impact actions. Actions such as device isolation, account suspension, and investigation escalation can concentrate power and exacerbate power disparities. And when the basis for action is opaque, this can be experienced as arbitrary and unjustifiable. Designing for contestability does not require revealing detection logic, instead it focuses on proportionate recourse. For high-impact actions, care should be taken to provide an intelligible and justifiable reason, a rapid escalation path to human review, and explicitly reversible interventions where feasible. This improves legitimacy and can increase cooperation with security teams.

Recovery-centric security. Employees encounter monitoring regimes most acutely at failure points: false positives, blocked work, and lockouts. Recovery is therefore a core property of the monitoring regime and not an edge case. Rather than resorting to punitive restriction and indefinite suspicion, greater focus should be placed on recovery flows that restore legitimate work quickly while remaining resilient to genuine attack (e.g., bounded delays, clear next steps, secure "break-the-glass" attention paths).

Decouple learning from punishment. Training and simulated phishing are often coupled to monitoring but can become disciplinary instruments that make user failure administratively convenient. When metrics trigger punishment (explicitly or implicitly), employees rationally conceal mistakes and avoid reporting. Learning-oriented programmes should reward reporting and remediation, separate learning data from HR evaluation, and treat recurrent failure as a signal of system misfit.

Constrain scope creep and secondary use. Telemetry collected for security can easily be repurposed for performance management or general policing, thus intensifying power asymmetries between employees and organisations, and undermining legitimacy. Monitoring programmes therefore need explicit data governance: purpose limitation, retention bounds, privileged access, audit trails, and clear separation between security investigation and managerial surveillance. This is an important condition for maintaining employee cooperation.

3.2 Future research

Addressing these design challenges also opens up possible future research challenges, which we have framed as the following open questions:

- How can monitoring decisions be made meaningfully contestable without disclosing sensitive detection logic?
- Which governance arrangements prevent harmful secondary use of monitoring telemetry while preserving incident response capability?
- What recovery patterns minimise false-positive harm and restore productivity without creating attacker bypasses?
- How do punitive compliance metrics affect reporting, concealment, and organisational learning over time?

4 Conclusion

Endpoint monitoring makes visible a core truth about security and privacy: protective mechanisms are also governance mechanisms that redistribute visibility, agency, and accountability. A critical framing clarifies how monitoring regimes can foster user blame narratives by converting auditable traces into moralised accounts of compliance and failure. Usable security is aligned with this critique because it treats failure and coping strategies as design signals, redirecting responsibility toward institutional choices about burden distribution, contestability, and recovery.

Beyond the practical challenge of building more appropriate monitoring regimes, usable security remains an important perspective that challenges the status quo of security "best practice". It offers solutions that reallocate responsibility from individual users to the socio-technical conditions that make failure predictable, turning "user error" into a redesign problem and not a blame game.

Acknowledgments. This paper is dedicated to the inspirational work and guidance of Prof. Angela Sasse to whom I owe so much.

References

1. Adams, A., Sasse, M.A.: Users are not the enemy. Commun. ACM **42**(12), 40–46 (Dec 1999). https://doi.org/10.1145/322796.322806, https://doi.org/10.1145/322796.322806
2. Bada, M., Sasse, M., Nurse, J.: Cyber security awareness campaigns: Why do they fail to change behavior? In: Proceedings of the International Conference on Cyber Security for Sustainable Society, Coventry, 26-27 February. pp. 118–131 (2015)
3. Brostoff, S., Sasse, M.A.: Are passfaces more usable than passwords? a field trial investigation. In: People and computers XIV—usability or else! Proceedings of HCI 2000. pp. 405–424. Springer (2000)
4. Dwyer, A.C., Stevens, C., Muller, L.P., Cavelty, M.D., Coles-Kemp, L., Thornton, P.: What can a critical cybersecurity do? International Political Sociology **16**(3), olac013 (07 2022). https://doi.org/10.1093/ips/olac013, https://doi.org/10.1093/ips/olac013

5. Flechais, I., Mascolo, C., Sasse, M.A.: Integrating security and usability into the requirements and design process. Int. J. Electron. Secur. Digit. Forensic **1**(1), 12–26 (May 2007). https://doi.org/10.1504/IJESDF.2007.013589, https://doi.org/10.1504/IJESDF.2007.013589
6. Geuss, R.: The Idea of a Critical Theory: Habermas and the Frankfurt School. Cambridge University Press, Cambridge, UK (1981)
7. Horkheimer, M.: Traditionelle und kritische Theorie. Zeitschrift für Sozialforschung **6**(2), 245–294 (1937)
8. Kerckhoffs, A.: La cryptographie militaire. Journal des Sciences Militaires pp. 161–191 (1883)
9. Klein, H.K., Huynh, M.Q.: The critical social theory of jürgen habermas and its implications for IS research. In: Mingers, J., Willcocks, L.P. (eds.) Social theory and philosophy for information systems, pp. 157–237. Wiley, Chichester (2004)
10. Muller, L.P.: Cybersecurity in practice: The vigilant logic of kill chains and threat construction. European Journal of International Security **10**(2), 231–251 (2025). https://doi.org/10.1017/eis.2024.27
11. Rogaway, P.: The moral character of cryptographic work. Cryptology ePrint Archive (2015), https://eprint.iacr.org/2015/1162
12. Stahl, B., Doherty, N., Shaw, M., Janicke, H.: Critical theory as an approach to the ethics of information security. Science and engineering ethics **20** (11 2013). https://doi.org/10.1007/s11948-013-9496-6
13. Whitten, A., Tygar, J.D.: Why johnny can't encrypt: a usability evaluation of pgp 5.0. In: Proceedings of the 8th Conference on USENIX Security Symposium - Volume 8. p. 14. SSYM'99, USENIX Association, USA (1999)