Reciprocal Disclosure and the Ethics of Vulnerability Reporting

A Cybersecurity Ethics Case Study of Nightmare Eclipse

James Herrick

Dude Tech IT Out

Author Note

I am the author of the blog post cited in this paper as DudeTechItOut (2025). This paper builds on and provides the philosophical framework for ideas I first explored there.

Correspondence concerning this paper should be addressed to James Herrick.

Email: dudetechitout@gmail.com

Abstract

Bug bounty programs and coordinated vulnerability disclosure (CVD) frameworks are widely regarded as means of aligning the interests of independent security researchers with that of software vendors. This paper argues that the ethical sustainability of these frameworks depends not on their formal structure but on the quality of moral give-and-take (reciprocity) between researcher and organization. Drawing on Cialdini's (2006) principle of reciprocity, Fricker's (2007) concept of epistemic injustice, and the existing ethics of disclosure literature, this paper examines the case of Nightmare Eclipse - a security researcher who shifted from cooperative disclosure to publicly releasing a series of zero-day exploits targeting Microsoft Windows, including core components such as Windows Defender and BitLocker, after experiencing what they describe as systemic dismissal, undervaluation, and legal intimidation.

The case is notable for its scale, as three of the first six disclosed vulnerabilities were exploited in the wild, attackers were able to chain them with ransomware deployments, and the researcher's GitHub and GitLab accounts were swiftly suspended in what can only be described as an institutional effort to suppress further disclosure. These were not vulnerabilities in a niche product either - Windows runs on over 1.4 billion active devices worldwide (Mehdi, 2025), dominates the global desktop market (StatCounter, 2026) and serves as the backbone of most enterprise IT environments - meaning the fallout affected businesses, hospitals, governments and critical infrastructure that depend on it every day.

The argument is that this case exposes a structural failure - and a form of epistemic injustice - in how organizations conceive of their moral obligations toward vulnerability reporters and that the resulting harm to end users, enterprises and the broader international

security environment constitutes a foreseeable consequence of institutional bad faith. The paper proposes a normative framework focused on the concept of *reciprocal disclosure* - a reframing that puts the same expectations on both sides and accounts for the fact that the researcher and the organization do not hold equal power in this relationship.

Keywords: *reciprocal disclosure; epistemic injustice; vulnerability disclosure ethics; retaliation by infrastructure; Nightmare Eclipse; bug bounty programs; coordinated vulnerability disclosure; responsible disclosure; researcher-vendor trust; zero-day disclosure; cybersecurity ethics*

1. Introduction

When a security researcher finds a flaw in software that millions of people rely on, they hold a strange kind of power. And with that power comes a weight. What they do with that knowledge - report it quietly, sell it, publish it - can protect people or expose them. The prevailing institutional response has been to build coordinated vulnerability disclosure (CVD) frameworks and bug bounty programs - structured channels for researchers to report findings to vendors in exchange for recognition, compensation, or both.

At their core, these programs depend on a moral compact - an agreement, if you will. The researcher agrees to disclose privately and responsibly, and the organization agrees to receive, evaluate, and remediate the finding in good faith. When this compact holds, everyone benefits. The researcher gets recognized and paid, the vendor ships a patch before anyone gets hurt, and users stay safer. But none of that works unless both sides actually keep their end of the bargain.

It requires ongoing moral reciprocity - good faith, in plain terms - and when that breaks down, the consequences reach far beyond the two parties involved. The question this paper asks is: when the disclosure relationship devolves, who is responsible for what comes next - and should organizations be held to the same standard they demand of the researchers who protect them?

The case of Nightmare Eclipse answers that question in the most concrete terms possible. Between April and June 2026, this researcher publicly released at least eight zero-day exploits affecting Microsoft Windows, including multiple vulnerabilities in Microsoft Defender that allowed local privilege escalation to SYSTEM-level access. These were not trivial by any means - one exploited a race condition in Defender's own runtime process, meaning the security software designed to protect the system became the vector for compromising it. That is not a low-hanging finding. That is someone who understands Windows internals at a depth most engineers never reach.

Three of these were weaponized in the wild, chained by attackers with ransomware deployments (Barry, 2026). Microsoft responded with legal threats, suspended the researcher's GitHub and GitLab accounts (Montalbano, 2026; Abrams, 2026), and allegedly deleted their Microsoft Security Response Center (MSRC) bug-reporting account (CiphersSecurity, 2026; Okemwa, 2026). The researcher has promised more disclosures and shows no signs of stopping. This breakdown is not abstract - it costs compromised systems, ransomware infections, and organizational damage affecting millions of Windows users worldwide.

The paper examines the case through three philosophical lenses - reciprocity as a moral foundation for disclosure relationships (Cialdini, 2006), epistemic injustice as applied to the

researcher-organization dynamic (Fricker, 2007), and the question of whether going public is ever morally justified - and if so, under what conditions. It draws on consequentialist and deontological reasoning and existing disclosure ethics literature (Pupillo et al., 2018) and critiques of bug bounty platform design (Moussouris, 2021).

The argument here is that while going public with unpatched vulnerabilities is not without ethical cost, the moral responsibility for the resulting harm is shared - and that the current framing of "responsible disclosure" obscures this by putting the weight on only one side. The paper concludes by proposing *reciprocal disclosure* - a framework that holds researchers, organizations, and the platforms between them to the same standard, while acknowledging that they do not enter the relationship with equal power.

2. How Vulnerability Disclosure Is Supposed to Work

Understanding how the Nightmare Eclipse case devolved requires understanding the system it devolved from - what it promises, who it protects, and where the cracks were long before anyone started pulling at them.

2.1 How Coordinated Disclosure Works & Where It Breaks

Coordinated vulnerability disclosure programs formalize the process through which security researchers report vulnerabilities they find to affected vendors, typically through dedicated channels or third-party platforms like HackerOne and Bugcrowd. The agreement is relatively straightforward: the researcher agrees not to disclose publicly for a set period while the vendor works on and implements a fix. In return, the vendor acknowledges the finding,

communicates transparently about remediation timelines, and often pays the researcher based on the severity of the vulnerability.

These are codified in international standards, particularly ISO/IEC 29147 (vulnerability disclosure) (ISO, 2018) and ISO/IEC 30111 (vulnerability handling processes) (ISO, 2019), co-authored by Katie Moussouris, who also launched Microsoft's first bug bounty program (Moussouris, 2021).

These standards exist because managing vulnerabilities is not a courtesy; it's a regulatory and contractual obligation. PCI DSS, HIPAA, SOC 2, and FedRAMP all mandate that organizations identify and remediate vulnerabilities in their systems. The consequences are not theoretical - when CISA added the Nightmare Eclipse vulnerabilities to its Known Exploited Vulnerabilities catalog, federal civilian agencies became legally required to remediate them under CISA's Binding Operational Directive 22-01, since superseded by BOD 26-04, authorized by 44 U.S.C. § 3553(b)(2) (CISA, 2021). Security researchers who report those vulnerabilities are doing work the compliance frameworks assume someone is doing. Silencing them doesn't eliminate the obligation. It just removes one of the most effective ways of meeting it.

This is not some abstract argument - a federal agency has already said the same thing. The US Federal Trade Commission filed a complaint against mobile device manufacturer HTC in which one of the five charges was the failure to implement a process for receiving and addressing security vulnerability reports from third-party researchers (Pupillo et al., 2018). The FTC determined that not having a functional disclosure process is itself a security failure - not an administrative oversight but a violation of reasonable security standards. When Microsoft deleted Nightmare Eclipse's MSRC account, it didn't just silence a researcher. It destroyed the very

channel that federal precedent has already recognized as a baseline security requirement. That is not a vendor exercising discretion - it is a vendor dismantling its own accountability infrastructure while demanding accountability from the person on the other end of it.

The existing philosophical literature on disclosure ethics has primarily framed the question in terms of outcomes and duties. Pupillo et al. (2018) argue that both vendors and third parties have an ethical obligation to disclose vulnerabilities to affected users, and that the timing of that disclosure should be guided by minimizing risk to end users. Their analysis concludes that vendors should release fixes as quickly as possible and that researchers should adopt coordinated disclosure to minimize user risk. That framing is useful - but it misses something fundamental. It assumes both sides are cooperating. And it has no answer for what happens when one side lets it devolve.

Moussouris (2021) highlights a structural problem running underneath much of this debate - the conflation of vulnerability disclosure with bug bounty programs. Vulnerability disclosure is a process for receiving and handling security reports. Bug bounties are an incentive mechanism layered on top. When organizations treat the bounty as the whole program, they create the conditions for the very dysfunction this paper examines. The ethical analysis of disclosure has focused on the researcher's obligations and the vendor's obligations but has given far less attention to the third layer - the platforms, programs, and processes that sit between them and shape the relationship from the inside.

2.2 The Implicit Moral Compact

The moral compact underlying vulnerability disclosure, at a base level, comes down to the researcher agreeing to disclose privately, provide enough technical detail for the vendor to

actually fix the problem - often in the form of a proof-of-concept (PoC) that demonstrates the vulnerability, lays out the steps to replicate it, and sometimes even suggests a path to patching it, though that last part is not typically required - not take advantage of the vulnerability themselves, and give reasonable time for patching before going public. The organization, in return, agrees to acknowledge the report, evaluate it on its technical merits rather than institutional convenience, communicate transparently, compensate fairly where applicable, and - critically - not retaliate against the person who just helped them.

But here is the problem. This agreement is rarely codified in these terms. Bug bounty program terms of service typically spell out the researcher's obligations in considerable detail while the organization's obligations remain ambiguous. The researcher's side of the deal is a contract. The organization's side is a suggestion. That asymmetry in itself is ethically significant. The data shows this: a USENIX study on bug bounty programs has found that poor responsiveness from program managers, downgrades in severity, disputes over validity, and lower than expected payouts are major sources of researcher dissatisfaction (Akgul et al., 2023). The compact's failure is not an anomaly. It is systemic.

2.3 The Nightmare Eclipse Case

The case under examination involves a security researcher operating under the pseudonym Nightmare Eclipse (also known as Chaotic Eclipse, Dead Eclipse, MSNightmare), who between April and June 2026 publicly released at least eight zero-day exploits targeting Microsoft Windows. The following chronological account is drawn from multiple independent sources, including SecurityWeek (Arghire, 2026), BleepingComputer (Toulas, 2026; Abrams, 2026), Dark Reading (Montalbano, 2026), The Register (Lyons, 2026a; 2026b), CyberNews

(2026), and Barracuda Networks (2026). This paper does not adjudicate all factual claims of either party but instead takes the documented public record as the basis for philosophical analysis.

Phase 1: Prior reporting and the disputed breakdown of channels (pre-April 2026).

Nightmare Eclipse says they first worked within Microsoft's existing vulnerability disclosure process, submitting reports through the Microsoft Security Response Center (MSRC). The researcher's stated grievances, documented across multiple independent news sources, include the following claims. That Microsoft allegedly deleted the MSRC account used to submit bug reports, cutting off access to their own prior submissions (CiphersSecurity, 2026). That Microsoft withheld bounty payments the researcher believed they had earned (Okemwa, 2026). That Microsoft removed their attribution from at least one published advisory (Martin, 2026). Finally, that they were told personally by a Microsoft representative that the company would "ruin my life and they did" (Ferreira, 2026).

Microsoft, for its part, pushes back on these claims. A spokesperson told The Record that the company "does not remove MSRC researcher portal accounts" and "cannot confirm which account this person is claiming was deactivated" (Martin, 2026). Microsoft has also stated that the exploits were "never reported via its official channels prior to being made public" (Lyons, 2026a). This paper does not adjudicate between these accounts.

What is independently documented and not disputed by either party is the following: the researcher published working zero-day exploits; Microsoft patched the vulnerabilities those exploits targeted; Microsoft invoked its Digital Crimes Unit in a public statement widely

interpreted as a threat of legal action; and GitHub and GitLab suspended the researcher's accounts.

For purposes of philosophical analysis, the critical question is not whether Microsoft's or the researcher's factual account is correct, but what the independently verifiable institutional conduct - legal threats, platform bans, and the reintroduction of "responsible disclosure" language - reveals about the ethical dynamics of the disclosure relationship. Separately, William Dormann of Tharros has observed that "MSRC used to be quite excellent to work with. But to save money, Microsoft fired the skilled people, leaving flowchart followers" (Ferreira, 2026), suggesting that the breakdown with Nightmare Eclipse may reflect broader institutional degradation rather than an isolated interpersonal conflict.

Phase 2: Public disclosure begins (April 2026).

On April 2, 2026, Nightmare Eclipse identified BlueHammer (CVE-2026-33825, CVSS 7.8 - Microsoft Important), a Windows Defender signature update workflow time-of-check to time-of-use (TOCTOU) flaw. This flaw allows any authenticated local user to escalate privileges to SYSTEM, the highest privilege level in Windows - above even Administrator. For context, most organizations deem anything above 7.0 a critical or urgent fix and federal agencies under CISA's Binding Operational Directive are legally mandated to patch known exploited vulnerabilities in this severity range within a set time frame. SYSTEM-level access gives the attacker unrestricted control over the machine: credential harvesting, malware installation, lateral movement, data exfiltration, and the ability to persist undetected. The exploit chains five legitimate Windows components - Defender's update mechanism, the Volume Shadow Copy

Service, the Cloud Files API, opportunistic locks, and the offline registry - into a single escalation path.

As Dormann confirmed to BleepingComputer, the flaw is “not easy to exploit” (Toulas, 2026), which makes the fact that it works at all a testament to the researcher’s depth of understanding of Windows internals. The exploit was published on GitHub with a warning: “I was not bluffing Microsoft and I’m doing it again” (Toulas, 2026). BlueHammer was available on GitHub from April 2 until the account was suspended on May 23 (CiphersSecurity, 2026), a 51-day window during which the proof-of-concept code was publicly accessible and downloadable.

Phase 3: Escalation (April - May 2026).

In the coming weeks, Nightmare Eclipse rolled out five more exploits. RedSun (CVE-2026-41091, CVSS 7.8) and UnDefend (CVE-2026-45498, CVSS 4.0) targeted Windows Defender - RedSun giving SYSTEM-level access via a link-following vulnerability, and UnDefend silently disabling Defender’s ability to receive updates and detect new threats while making the system appear healthy (Barry, 2026). While a CVSS score of 4.0 might sound like a minor issue on paper - in reality - disabling an organization’s primary endpoint protection during an active intrusion flips the whole threat landscape on its head. YellowKey (CVE-2026-45585, CVSS 6.8) targeted BitLocker, allowing an attacker with physical access to bypass device encryption and gain a SYSTEM-level shell on the protected volume - the exact outcome the encryption was meant to stop. GreenPlasma (CVE-2026-45586, CVSS 7.8) leveraged the Windows Collaborative Translation Framework, and MiniPlasma (CVE-2020-17103, CVSS 7.0) exploited a flaw in the Windows Cloud Files Mini Filter Driver, originally reported and

supposedly fixed in 2020, both granting SYSTEM-level access through components most security teams would never think to monitor (Barry, 2026). Across all six disclosed vulnerabilities, the average CVSS score is 6.9 - firmly in the High severity range - and five of the six grant SYSTEM-level access, the highest privilege tier Windows offers.

Three of the first six - BlueHammer, RedSun, and UnDefend - were confirmed as being exploited in real-world attacks shortly after publication (CyberNews, 2026). Attackers were chaining these exploits with ransomware deployments (as reported in Barry, 2026), assembling a multi-stage attack chain that could escalate privileges via BlueHammer, RedSun, or MiniPlasma, blind endpoint detection via UnDefend, access encrypted volumes via YellowKey, and persist through GreenPlasma. CISA added the vulnerabilities to its Known Exploited Vulnerabilities catalog.

Phase 4: Institutional retaliation and silencing (May 2026).

On May 23, GitHub suspended Nightmare Eclipse's account, with GitLab doing the same on May 26. On May 28, Microsoft's Digital Crimes Unit issued a public statement suggesting the possibility of criminal prosecution, stating that it would "continue bringing cases against these actors and those that enable their criminal activity" (Lyons, 2026a). This received immediate backlash from the cybersecurity community. Kevin Beaumont, a prominent security researcher and former Microsoft employee, described the situation as "a dumpster fire of [Microsoft's] own making" (Lyons, 2026a).

Dustin Childs, head of threat awareness at Trend Micro's Zero Day Initiative, observed: "CVD is a two-way street. The vendor has some responsibility as well, so to go out publicly stating this person violated CVD without showing any of the correspondence seems bold"

(Lyons, 2026a). Katie Moussouris, who pioneered Microsoft's own bug bounty program, noted that Microsoft's post revived the term "responsible disclosure" - language Microsoft had formally retired in 2010 - and called its reappearance "loaded," writing that "no vendor uses that term unless they want to call someone irresponsible" (Martin, 2026).

Microsoft subsequently clarified it had "no intention to pursue action against individuals conducting or publishing their security research" and, notably, dropped the term "responsible disclosure" from its clarification, reverting to "Coordinated Vulnerability Disclosure" (Martin, 2026). Microsoft also acknowledged that "some interactions have fallen short" and that it was "working to learn" from those incidents (Martin, 2026).

Nightmare Eclipse wrote: "When I actively asked you to communicate with me, you refused, humiliated me and made sure to insult me in front of people. You defame me in public with your CVE-2026-45585 advisory even though you literally deleted the Microsoft account I used to report bugs to you with and I got zero pennies from doing so and I still happily did like an idiot" (Lyons, 2026a). RedSun was exploited in the wild for six weeks before Microsoft issued a formal CVE or patch (Parham, 2026) - a notable gap from a company invoking disclosure transparency as its justification.

Phase 5: Continued adversarial disclosure (June 2026).

Microsoft's June 2026 Patch Tuesday addressed GreenPlasma (CVE-2026-45586) and YellowKey (CVE-2026-45585). Within hours, Nightmare Eclipse released RoguePlanet from a self-hosted code repository, having been banned from both GitHub and GitLab. RoguePlanet exploits a race condition in Microsoft Defender to escalate local privileges to SYSTEM on fully patched Windows 10 and Windows 11 machines.

ThreatLocker independently confirmed the exploit's viability (Abrams, 2026). The following day, the researcher released GreatXML, a BitLocker bypass discovered in four hours, bringing the total exploit count to eight (Lyons, 2026b). The researcher has promised further disclosures and has stated that they possess "a batch of memory corruption vulnerabilities in Defender" alongside vulnerabilities in other components (Montalbano, 2026).

This five-phase trajectory - from cooperative disclosure through institutional rejection and retaliation to sustained adversarial release, with documented real-world exploitation affecting millions of Windows users - provides the empirical foundation for the ethical analysis that follows.

3. Reciprocity as the Moral Foundation

Understanding why the Nightmare Eclipse case escalated the way it did requires understanding what held the disclosure relationship together in the first place - and what happens when that foundation is removed.

3.1 Where Reciprocity Comes From

Robert Cialdini's (2006) principle of reciprocity holds that individuals feel a strong psychological obligation to return favors, concessions, and acts of goodwill. While Cialdini frames this primarily as a social-psychological mechanism, the principle has deep roots in moral philosophy. Reciprocity plays a central role in contractualist ethics (Scanlon, 1998), Axelrod's (1984) evolutionary account of cooperation, and the broader philosophical view that moral relations are underpinned by mutual obligation rather than one-way duty.

In the case of vulnerability disclosure, reciprocity operates as follows: the act of private disclosure by the researcher is a unilateral concession - they relinquish the option of public disclosure, which would bring immediate recognition and pressure for rapid patching, in favor of a cooperative approach that requires trust in the organization's good faith. This concession creates what we can describe as a *reciprocal obligation*: the organization is morally bound to honor the trust placed in it by responding with transparency, fairness and good faith.

3.2 How the Nightmare Eclipse Case Broke the Compact

When an organization fails to meet its reciprocal obligations - dismissing findings, downgrading severity without justification, going silent, retaliating - the moral foundation of the disclosure relationship erodes. The Nightmare Eclipse case is not a subtle erosion. The researcher alleges their MSRC account was deleted, their findings went uncompensated, legal action was threatened, and their code hosting accounts were suspended across two platforms. Every channel for cooperation was closed, and then the organization called the researcher uncooperative. At what point does the obligation to cooperate dissolve? Reciprocity, as a philosophical framework, has a clear answer: obligations sustained by mutual cooperation cannot survive the unilateral abandonment of that cooperation by one party. You cannot demand someone play by the rules of a game you have stopped playing.

This does not mean the researcher's subsequent actions come without ethical cost - the consequences for end users are real and are examined in Section 5. But the moral calculus cannot be conducted without accounting for the organization's role in precipitating the breakdown. You cannot remove the cause and then judge only the effect. There is a pattern worth noting in cases like this - when the disclosure relationship breaks down, the cooperative history tends to

disappear from the narrative. The reports filed in good faith, the unpaid labor, the time put in working within the system - none of it factors into how the researcher is ultimately characterized. The story begins at the moment of public disclosure, as if nothing came before it. Whether that framing is intentional or simply a consequence of how institutions construct their own narratives, it raises a question the ethics of disclosure cannot ignore - can one side's conduct be evaluated in isolation from the relationship that produced it?

It is also worth noting that the adversarial dynamic is not one from which the researcher profits or thrives. Nightmare Eclipse has publicly expressed that the development of RoguePlanet had a great toll on their "mental and physical health" and "drained [their] soul" (Montalbano, 2026). This is ethically significant because it dispels any narrative that depicts the researcher as a gleeful saboteur reveling in the chaos they have created. The pattern is consistent with the psychological literature on workplace ostracism and illegitimate task assignment: sustained rejection and the imposition of tasks that are perceived as unjust do not produce determination but burnout, detachment, and diminished well-being (Williams, 2007). Williams specifically notes that when an individual's sense of control and meaningful existence are sufficiently threatened by exclusion, "a desire to be noticed may supplant a desire to be liked" (2007, p. 444) - a dynamic that describes the Nightmare Eclipse trajectory with uncomfortable precision.

Research on illegitimate tasks - tasks that violate an employee's understanding of what is appropriate for their role - has demonstrated that such tasks function as a distinct category of workplace stressor, threatening professional identity and producing strain, emotional exhaustion, and decreased well-being through what Semmer and colleagues have termed "stress-as-offense-to-self." Semmer et al. specifically identify the mechanism as an "implicit message of

disrespect" that constitutes a direct "threat to the self" - the harm comes not from the difficulty of the work but from what the institutional response communicates about the value of the person doing it (2015, p. 34). When a researcher submits findings through official channels and receives silence, account deletion, and public condemnation in return, the institutional response functions as precisely this kind of identity-threatening stressor - one sustained over months and escalated at every stage.

The parallel to Nightmare Eclipse is direct - a researcher whose professional contributions are dismissed, whose reporting account was allegedly deleted, whose work goes uncompensated, and who is then publicly accused of irresponsibility by the institution that closed every legitimate channel available to them, is experiencing precisely the kind of illegitimate treatment that this literature predicts will produce psychological harm. The researcher's own words confirm the prediction - the development of RoguePlanet "drained [their] soul" (Montalbano, 2026), and they have acknowledged significant degradation of their "mental and physical health." The researcher is not flourishing in the adversarial mode - they have been ground into it by the very dynamic the organization's conduct created. Williams (2007) and Semmer et al. (2015) do not merely explain this outcome, but instead predict it.

3.3 The Litmus Test as Reciprocity Testing

A related phenomenon, described in practitioner literature as the "litmus test" approach (DudeTechItOut, 2025). This involves submitting low-severity or proof-of-concept findings as an initial assessment of a program's good faith. How the organization responds to these early, low-stakes submissions acts as a proxy for how it will react to more serious findings. In Axelrod's (1984) terms, the litmus test represents the first cooperative move in an iterated game,

where the researcher cooperates first and observes whether cooperation is reciprocated. This is a form of reciprocity testing: the researcher extends a small measure of trust and assesses whether it is returned.

The litmus test is ethically significant because it reveals that researchers are not passive participants in disclosure programs but active moral agents evaluating the trustworthiness of their counterparts. Organizations that fail the litmus test - through silence, dismissal, downgraded severity, or withheld credit - send what Semmer et al. (2015) would identify as an "implicit message of disrespect," one that communicates to the researcher that their professional contribution has no value to the institution.

In most cases, this results in quiet disengagement - precisely the pattern Childs described when he noted that researchers "stopped looking at Microsoft altogether because they were too difficult to work with" (Lyons, 2026a). In extreme cases, as with Nightmare Eclipse, it results in adversarial disclosure. Williams (2007) predicts both outcomes - when belonging and self-esteem are the primary needs threatened, individuals withdraw; when control and meaningful existence are threatened, they respond with antisocial behavior designed to force recognition. The litmus test failure determines which path the researcher takes. Organizations do not merely lose a report - they determine the trajectory of the entire relationship.

4. Epistemic Injustice and the Researcher-Organization Dynamic

Reciprocity explains why the relationship broke down. Epistemic injustice explains how the breakdown was made possible - and why the researcher's testimony was never given the weight it deserved.

4.1 What Epistemic Injustice Means

Miranda Fricker's (2007) concept of epistemic injustice gives us the vocabulary for something security researchers experience but rarely name. Fricker identifies two forms: *testimonial injustice*, where a speaker's credibility is unfairly deflated - not because they are wrong, but because of who they are relative to the institution hearing them - and *hermeneutical injustice*, where a speaker lacks the interpretive framework to articulate their own experience.

The researcher-organization relationship is ripe for the first kind. A security researcher, particularly one operating independently without institutional affiliation, submits a vulnerability report that is evaluated not only on its technical merits but through the lens of institutional power. The organization decides what counts as a valid finding, assigns severity, and determines compensation - the researcher brings the technical knowledge, but the organization holds the authority, and those are not the same thing.

Moussouris identified this exact gap when she observed that Microsoft's response to Nightmare Eclipse "confusingly claims their program ensures researchers are compensated and publicly acknowledged in a statement answering a researcher who says he got neither" (Lyons, 2026a). The institution's framework for evaluating the researcher's credibility is built on premises the researcher's own experience contradicts - and the institution's version prevails because the institution controls the platform.

4.2 Application to the Nightmare Eclipse Case

When Microsoft reportedly dismissed Nightmare Eclipse's findings, allegedly deleted their MSRC account, and threatened legal action, this constitutes a form of testimonial injustice:

the researcher's credible technical testimony was deflated not because it lacked evidential support, but because acknowledging it would have imposed costs on the institution. The subsequent suspension of accounts on GitHub, GitLab, and MSRC compounds this injustice by removing the researcher's capacity to testify at all - silencing the speaker rather than engaging with their claims. The irony that GitHub is owned by Microsoft adds a troubling dimension of institutional leverage.

The fact that a single company can simultaneously be the vendor whose software is vulnerable, the platform that hosts the researcher's code, and the entity that decides whether that code stays online raises broader questions about concentration of power in the technology industry - questions that, while beyond the scope of this paper, are difficult to ignore when they converge on a single researcher's ability to be heard.

This framing recontextualizes the researcher's subsequent actions. Public disclosure of zero-day exploits, viewed in isolation, appears irresponsible. Viewed as a response to systematic epistemic injustice - as the only remaining avenue for a silenced speaker to be heard - it takes on a different moral character. As Kevin Beaumont observed, Microsoft had previously reportedly engaged SandboxEscaper after she published similar zero-day exploits, conduct that Microsoft's official blog now characterizes as criminal (Lyons, 2026a). This inconsistency is not merely hypocrisy - it is testimonial injustice driven by institutional convenience.

When SandboxEscaper's exploits could be co-opted by absorbing the researcher into the organization, the same act of public disclosure was reframed as talent worth acquiring. When Nightmare Eclipse's exploits cannot be co-opted, the identical conduct is reframed as criminal. The credibility of the testimony has not changed; what has changed is whether acknowledging it

serves the institution's interests. This is the mechanism Fricker (2007) describes: the speaker's credibility is deflated not by the quality of their evidence but by the hearer's institutional positioning.

4.3 Structural Epistemic Injustice in Bug Bounty Programs

The epistemic injustice observed in the Nightmare Eclipse case is not anomalous but structural. The terms of service of bug bounty programs often grant organizations unilateral authority to decide the severity, pay, and scope of a bug, and researchers who contest the severity assessment have limited recourse, particularly when third-party platforms defer to the organization's judgment. The USENIX study provides empirical evidence of this (Akgul et al., 2023). These are not merely operational inefficiencies but sites of epistemic injustice, where the expert testimony of the researcher is systematically devalued.

Dormann, principal vulnerability analyst at Tharros, notes that the institutional degradation of MSRC, where skilled evaluators are replaced by "flowchart followers" (Ferreira, 2026), indicates a structural mechanism through which testimonial injustice becomes institutionally embedded. When an organization replaces personnel capable of exercising technical judgment with procedural gatekeepers, it systematically loses the capacity for the kind of good-faith engagement that reciprocal disclosure requires. The procedural gatekeeper cannot evaluate whether a vulnerability finding has merit; they can only verify whether the submission meets formal requirements. When the submission does not fit the flowchart, it is rejected - not because it lacks technical validity but because the institution has, by design, made itself unable to hear what the researcher is saying. This is testimonial injustice operationalized at the institutional level.

The institutional degradation Dormann describes (Ferreira, 2026) has consequences beyond a single researcher. Childs, who spent seven years working in Microsoft security, told The Register that “while some companies have improved, Microsoft has not. If anything, they are seen as difficult to work with, especially if your bug is Moderate instead of Critical. I’ve had researchers tell me that they stopped looking at Microsoft altogether because they were too difficult to work with” (Lyons, 2026a).

The Nightmare Eclipse case is not an isolated breakdown - it is the most visible symptom of a researcher exodus that has been building for years. Empirical research confirms the pattern. In a large-scale USENIX study of bug bounty participants, one researcher described a dynamic directly analogous to the Nightmare Eclipse grievance: “Sometimes, when the program got everything they needed from you... the bug is getting fixed... severity gets lowered, and you get lower rewards than you expected. And then, they just suddenly stop responding to your comments, and that’s a situation that happens all the time” (Akgul et al., 2023, p. 2284).

It is important to note that not all observers share the framing presented here. Barry (2026), in an analysis of the Nightmare Eclipse exploit chain for Barracuda Networks, explicitly labels the researcher “a malicious actor - not a whistleblower, not a responsible disclosure advocate and not a neutral researcher.” This characterization deserves engagement rather than dismissal. The philosophical analysis presented in this paper does not depend on whether Nightmare Eclipse is characterized as a malicious actor, a frustrated researcher, or something in between. The ethical dynamics of reciprocity, epistemic injustice, and institutional power operate the same way regardless of the label applied to the researcher. What matters for the argument is not the researcher’s character but the institutional conditions that preceded and shaped their conduct.

Even if one accepts the characterization of the researcher as a malicious actor, the question remains: what institutional failures made it possible for a single individual to weaponize this many vulnerabilities in core Windows components, and could a different institutional response at any earlier stage have prevented the escalation? It is also worth noting what that label asks the reader to overlook. This is a researcher who chained five legitimate Windows components into a single escalation path, found race conditions in Defender's runtime process, bypassed BitLocker encryption, and did so repeatedly across multiple core components over the span of weeks.

The technical quality of that work is not in dispute - Microsoft patched every vulnerability the researcher disclosed, which is itself an acknowledgment that the findings were valid. Labeling someone a malicious actor while simultaneously treating their work as credible enough to warrant emergency patches is not a consistent position. To call the findings dangerous enough to patch in days, while simultaneously calling the person who found them too illegitimate to engage with, is not a security assessment - it is institutional convenience dressed up as one. It is a rhetorical move that discounts the expertise in order to avoid reckoning with the grievance behind it - and that is, in itself, a form of the testimonial injustice this section describes.

There is a financial angle here that deserves attention. Microsoft's bug bounty program pays between $30,000 and $100,000 per endpoint zero-day and up to $250,000 for Hyper-V exploits (Okemwa, 2026). Nightmare Eclipse released at least eight working zero-day exploits and, by their own account, received nothing. The researcher themselves acknowledged the financial sacrifice: "I could have made some insane cash selling this, but no amount of money will stand between me and my determination against Microsoft" (Okemwa, 2026). A researcher

sitting on vulnerabilities of this caliber had a clear financial path through official channels - and an even more lucrative one through the zero-day broker market, where exploits of this quality command significantly higher prices. The fact that they chose to release them publicly, for free, with no compensation, raises a question the “malicious actor” framing struggles to answer - why would someone with this level of skill and this many viable exploits walk away from the money unless the channels designed to compensate them had already failed?

There is a further irony worth noting: Microsoft patched every vulnerability Nightmare Eclipse disclosed. Each patch improved the security posture of over a billion Windows devices. The work that Microsoft’s own blog characterized as “never justifiable” is work that Microsoft nonetheless used - and used without compensating the person who produced it. The bounty program exists precisely to pay for this kind of labor. In this case, Microsoft received the security benefit of eight vulnerability discoveries and paid nothing for any of them. Using Microsoft’s own bounty program as a benchmark, eight endpoint zero-day vulnerabilities would warrant between $240,000 and $800,000 in bounty payments - compensation the researcher claims was never received. The incentive structure of a publicly traded company does not naturally produce generosity toward the independent researchers it depends on - and the current disclosure framework has no mechanism to compel it.

The structural nature of this problem is not limited to the Nightmare Eclipse case. Moussouris (2021) has described an incident involving the Capital One data breach in which a security researcher reportedly submitted a report through a bug bounty platform stating that Capital One customer data had been leaked and providing a link to the data dump. The platform closed the report as out of scope because it reported a breach rather than a specific technical vulnerability. As Moussouris recounted: a person literally told the company their customers' data

was exposed, and the system designed to receive that information rejected it on procedural grounds. That is not a process failure - it is the process working exactly as designed, and the design is the problem.

It is also worth noting that the epistemic injustice described here has an international dimension that this paper, focused on a single case study, does not fully explore. The bug bounty researcher population is globally distributed (Akgul et al., 2023), and researchers in the Global South face additional structural disadvantages that compound the dynamics analyzed here: payment processing barriers that make compensation difficult or impossible to collect, visa-related legal exposure that makes even good-faith disclosure risky, and language gaps in report evaluation that can cause technically valid findings to be dismissed on the basis of communication style rather than substance. These barriers represent an additional layer of testimonial injustice that the existing disclosure ethics literature has largely overlooked, and future work should examine how the reciprocal disclosure framework proposed here might be adapted to account for the structural inequities facing researchers outside the Anglophone West.

5. Ethics of Going Public

The previous sections examined why the disclosure relationship collapsed. This section turns to the harder question - what happened after the collapse, who was harmed, and where the moral responsibility for that harm actually sits.

5.1 Consequentialist Analysis

The consequentialist case against public disclosure of unpatched vulnerabilities is straightforward: it exposes end users to exploitation during the window between disclosure and patching. In the Nightmare Eclipse case, this is by no means hypothetical. Three of the first six initially disclosed exploits were confirmed as being used in real-world attacks. Barry (2026) documented attackers chaining these exploits with ransomware deployments. CISA added the vulnerabilities to its Known Exploited Vulnerabilities catalog. The harm is real, measurable, and affecting millions of Windows users worldwide.

However, a complete consequentialist analysis must also account for the counterfactual. Had the researcher continued to engage cooperatively with an organization that was, by their account, refusing to engage in good faith, the speed and quality of remediation would have depended entirely on the organization's internal prioritization - a process that, by multiple accounts, has degraded. Childs has observed that Microsoft is "difficult to work with, especially if your bug is Moderate instead of Critical" (Lyons, 2026a), suggesting that lower-severity findings face significant deprioritization.

RedSun was exploited in the wild for six weeks before Microsoft issued a formal CVE or patch (Parham, 2026). And MiniPlasma (CVE-2020-17103) exploited a vulnerability that Google Project Zero reported in 2020 and that Microsoft claimed to have patched in December of that year - six years later, the original proof-of-concept still achieved SYSTEM-level access on fully patched Windows 11 systems (Bayram, 2026). While, none of this proves that every vulnerability would go unpatched indefinitely, as Microsoft does patch thousands of CVEs each year, it does demonstrate that without external pressure, the verification that remediation was genuine is not guaranteed - and when verification fails, the gap between a claimed fix and an actual fix can persist for years, as the MiniPlasma case illustrates.

This is not a novel philosophical position - it is, in fact, a well-established industry standard. It is the operational premise behind the disclosure policies of the most respected institutions in vulnerability research. Google's Project Zero imposes a 90-day disclosure deadline - if a vendor does not patch within that window, technical details and proof-of-concept code are published regardless, on the explicit rationale that public transparency accelerates remediation (Willis, 2025). The U.S. government's coordination center, CERT/CC, has maintained a 45-day deadline for over a decade (Moussouris, 2021). Open Bug Bounty, a non-profit coordinated disclosure platform that has facilitated the remediation of over one million vulnerabilities, follows the same 90-day model based on ISO 29147 guidelines. The principle across all three is identical: public disclosure protects end users by forcing vendors to act. The only thing that differs in the Nightmare Eclipse case is the institutional authority behind the person invoking it.

Public disclosure, while harmful in the short term, has been shown to create immediate pressure for patching - and in this case, Microsoft patched the first six vulnerabilities disclosed by Nightmare Eclipse promptly once they were public, with three receiving out-of-band emergency patches before the next scheduled Patch Tuesday. The uncomfortable implication is that the researcher's public disclosure, whatever its motivations, accomplished what the official channels had not: it forced the remediation of vulnerabilities that were actively being exploited in the wild, produced patches that now protect over a billion Windows devices, and exposed a six-year-old fix that had never actually worked. The moral calculus is therefore not unambiguous - the act that caused short-term harm to end users may also be the act that ultimately protected them.

5.2 Deontological Considerations

From a deontological perspective, the central question is whether the researcher has a duty to protect end users that persists regardless of the organization's behavior. Pupillo et al. (2018) identify as a central goal of coordinated vulnerability disclosure that users receive enough information to evaluate risks to their own systems - in other words, that the people affected by a vulnerability have a right to know it exists. Critically, the report assigns this responsibility to both vendors and researchers, not to researchers alone. A strict Kantian analysis might suggest that the duty not to cause harm through public disclosure is categorical.

However, the researcher is not the sole moral agent. The organization bears a duty of care to its users - specifically, the duty to take reasonable steps to remediate known vulnerabilities and to maintain functional channels through which those vulnerabilities can be reported. When Microsoft allegedly deleted Nightmare Eclipse's MSRC account, it did not merely fail in its duty to the researcher - it failed in its moral and ethical duty to the users those reports were designed to protect. The deontological burden is therefore shared.

5.3 Virtue Ethics and the Character of Disclosure

A virtue ethics perspective asks not "what are the rules?" but "what would a person of good character do?" A virtuous researcher might exhaust all cooperative avenues before resorting to public disclosure. A virtuous organization would not create the conditions that make such resort necessary. The Nightmare Eclipse case suggests that the researcher did, at least initially, attempt cooperative engagement. The escalation occurred after a series of interactions that demonstrated the organization was not acting virtuously. This does not make the researcher's actions unequivocally virtuous - the harm to end users counts against them - but it puts those actions in a context that virtue ethics takes seriously.

Further complicating the virtue ethics analysis is the researcher's own language, which grew increasingly hostile as the dispute escalated. In a May 2026 post, Nightmare Eclipse wrote: "Mark this date July 14th, I will make sure your bones are shattered that day" (Lyons, 2026a). Dustin Childs called the language "troubling" (Lyons, 2026a), and Moussouris observed the "incendiary language" (Lyons, 2026a). This rhetoric cannot be ignored or sanitized.

A hostile reviewer might argue that such language reveals the researcher's true character - not a frustrated professional, but someone motivated by vengeance. However, the timeline is critical to evaluating this language in context. The threatening rhetoric appeared in late May 2026, after the documented sequence of events: the alleged MSRC account deletion and bounty withholding, the GitHub ban (May 23), the GitLab ban (May 26), and Microsoft's Digital Crimes Unit threat of criminal prosecution (May 28).

Whatever one makes of the researcher's character, the incendiary language is a consequence of the breakdown, not its cause. A virtue ethics framework does not excuse the language, but it does ask whether the institutional conditions that produced it bear moral weight - and whether a person of good character, subjected to the same sequence of institutional actions, might also depart from virtuous conduct. The answer to that question is not comfortable, but it is philosophically relevant.

There is also something revealing about which part of the researcher's conduct draws the most attention. Nightmare Eclipse's findings were impeccable enough to be graded as not only Important but worthy of emergency patches - effectively making Microsoft's entire security operation react to one person, a person they had already decided was not worth listening to. But what drew the attention? Language.

The reaction from institutional observers was not to ask how those vulnerabilities existed in the first place - it was to focus on the language used in a blog post. The technical substance of the work was treated as nothing more than background noise, and the tone did not just become the story - it became the centerfold. That pattern is itself a form of the epistemic dynamic described in Section 4 - the speaker's credibility is not engaged on the merits of what they said but deflected through the way they said it.

The USENIX study documents this trajectory empirically - one participant was banned from a major bug bounty platform over a dispute, and another noted that "exploiting the bugs themselves becomes more attractive when they get frustrated with bug-bounty programs" (Akgul et al., 2023, p. 2284). The institutional response of tone policing a frustrated researcher is not safety - it is the creation of conditions under which the next researcher decides not to disclose at all, and that is a direct detriment to the very end users these institutional responses claim to protect. Does this justify the threatening language? No. But there is something deeply inconsistent about treating someone's work as an emergency and their choice of language as the real problem.

5.4 The Moral Status of Vulnerability Knowledge and Exploit Chaining

Another ethical aspect is the moral status of vulnerability knowledge itself and the reality of exploit chaining. Cases of exploit chaining of low-severity findings like self-XSS, open redirects, or information leaks to achieve full account takeovers are documented in the practitioner literature (DudeTechItOut, 2025). The PACMAN attack on Apple M1 chips demonstrated a related principle at the hardware level: a memory corruption vulnerability, which was considered mitigated by Pointer Authentication, became fully exploitable when combined

with a speculative execution side-channel attack (Ravichandran et al., 2022). In both cases, vulnerabilities dismissed as inconsequential in isolation proved critical in combination.

The Nightmare Eclipse case makes this danger concrete. UnDefend, rated CVSS 4.0 - a medium-severity finding of the kind Childs noted Microsoft treats as not worth engaging with (Lyons, 2026a) - was exploited in the wild and required an emergency out-of-band patch (Parham, 2026). Barry (2026), in an analysis for Barracuda Networks, documented a fully operational attack chain assembled from the disclosed exploits: escalate privileges, blind endpoint detection, bypass disk encryption, and persist through alternative escalation paths. This chain was not theoretical - it was observed in active use.

When organizations systematically dismiss low-severity findings, they leave potential exploit chains intact. This is not merely a technical failure; it is a moral one, with direct implications for the duty of care owed to end users. An organization that dismisses low-severity findings on the assumption that they are inconsequential in isolation is making a bet that no attacker will think to combine them - a bet that the chaining literature, and this case study, shows is consistently lost.

5.5 The Counterargument: Special Obligations to End Users

The strongest objection to the framework presented here holds that vulnerability knowledge affecting millions of users creates a special obligation that persists no matter how badly the organization behaves. On this view, end users are innocent third parties who did not choose to be caught in the researcher-organization dispute. The researcher, by publicly disclosing unpatched vulnerabilities, instrumentalizes these users as collateral in a personal grievance. No amount of institutional bad faith by Microsoft justifies exposing millions of people to

ransomware, data theft, and system compromise. This is the weight of the power described in the introduction - the strange kind of power a researcher holds when they find a flaw in software that a billion people depend on. The objection is that power carries obligation regardless of how the other side behaves.

This objection deserves to be taken seriously, and this paper does not dismiss it. The harm to end users in the Nightmare Eclipse case is not speculative. Ransomware deployments leveraging these exploits are documented. CISA added the vulnerabilities to its Known Exploited Vulnerabilities catalog precisely because real organizations were being compromised. The researcher's stated frustration, however legitimate, does not erase these consequences. A stronger version of this objection holds that public disclosure of zero-day exploits constitutes reckless endangerment analogous to publishing instructions for building weapons - that the act itself is wrongful regardless of the circumstances that motivated it.

This analogy, however, does not hold up. The vulnerability already exists in deployed systems used by millions of people. The researcher did not create the danger; they revealed it. Consider a closer analogy: a structural engineer tells a building owner the foundation is cracking. The owner ignores them, fires them, threatens to sue them, and then tells the tenants everything is fine. The engineer, out of options, warns the tenants directly. The tenants are now alarmed, but they were always in danger - they just did not know it. The building owner's failure to act on the original warning is what left them exposed. Blaming the engineer for the alarm does not fix the foundation - and it does not absolve the owner who knew about the cracks and chose to do nothing.

Apply this to the Nightmare Eclipse case and an uncomfortable question emerges. Once the vulnerabilities were publicly disclosed, Microsoft patched them promptly - some within days. That speed demonstrates the capacity was always there. The question is not whether Microsoft knew about these specific vulnerabilities beforehand - that is disputed - but why an organization with the resources to patch this quickly depends on public embarrassment to move at that pace. If the official disclosure channels had been functioning as designed, would these vulnerabilities have received the same urgency? Or does the speed of the response itself suggest that public pressure produces results that private reporting does not? This does not justify the researcher's method, but it does complicate the attribution of causal responsibility considerably.

The millions of affected users are, in a very real sense, casualties of Microsoft's decision to treat a security researcher as an adversary rather than an ally. Even setting aside the disputed MSRC account deletion, the confirmed actions speak clearly - Microsoft banned the researcher’s GitHub account - a platform Microsoft owns - rather than restricting access to the proof-of-concept code while preserving a channel for communication.

That is not de-escalation - it is the deliberate closing of every door through which the researcher could have been brought back into a cooperative relationship. Let us not forget - Microsoft is not a startup operating in a vacuum. It is a company that holds dominant market share in enterprise IT environments and whose products run hospitals, governments, banks, and critical infrastructure whether or not those institutions had a choice in the matter. That position comes with a moral and ethical duty to the people who depend on it - a duty that does not disappear when a researcher becomes an inconvenience.

Moussouris, who established Microsoft's own bug bounty program, has named this dynamic directly: "The bugs are Microsoft's. They wrote the code and they own the risk to customers. Often researchers who previously work with a vendor respond in the extreme only when they feel there is no other choice. The power they hold is not at all proportionate to the vendor. This is a David and Goliath dynamic we don't like to see play out, especially since it's users who lose when coordination negotiations fail" (Lyons, 2026a). That the person who built the program is now describing its failure in these terms is itself evidence of how far the institutional relationship has degraded.

The special obligation to protect end users does exist - but it binds the organization at least as strongly as it binds the researcher. When the organization abrogates its end of that obligation, the moral weight of the resulting harm shifts accordingly. This is not to excuse the researcher's choices; it is to insist that moral responsibility be distributed in proportion to causal contribution. As Dustin Childs, head of threat awareness at Trend Micro's Zero Day Initiative, told Dark Reading: "customers will be affected by these disclosures, even if it means they have to engage their emergency patch process instead of getting exploited" (Montalbano, 2026). Childs' framing is instructive - acknowledging the harm while implicitly recognizing that the disruption of emergency patching is preferable to the alternative of continued exploitation. The harm is real, but its causes are distributed.

Perhaps most consequentially, the harm is not contained to one researcher. Other security researchers have begun leaking Microsoft exploits in what Ammar Askar described as "one of the few levers I have to try to influence MSRC" (Jones, 2026). Askar published a proof-of-concept exploit for a Visual Studio Code vulnerability within an hour of disclosing it, citing past

experiences in which Microsoft “silently fixed the bug I pointed out without any credit” and “marked it as not having any security impact” (Jones, 2026).

These are not coordinated attacks - they are individual researchers independently arriving at the same conclusion: that the official channels do not work. Moreover, Nightmare Eclipse themselves reported that other researchers had approached them and in some cases provided vulnerabilities directly (Martin, 2026) - meaning the contagion is not merely independent imitation but active channeling of findings through an adversarial disclosure pathway because researchers have lost faith in the official process.

This is the strongest possible evidence for the framework presented here: Microsoft’s institutional conduct has not deterred adversarial disclosure but normalized it. The consequentialist case for good-faith engagement is not merely that it prevents one researcher from going rogue - it is that bad-faith engagement produces a systemic shift in researcher behavior across the entire ecosystem. The cost is not one disgruntled individual; it is an erosion of the cooperative norms on which coordinated vulnerability disclosure depends.

The question of how this cascading harm might have been prevented leads naturally to the role of intermediary platforms - the institutional structures positioned between researchers and vendors that could, in principle, have intervened before the breakdown became irreversible.

6. The Role of Intermediary Platforms

The existing ethics of disclosure literature focuses on two actors: the researcher and the vendor. Missing from this analysis is a third actor whose role is increasingly consequential: the intermediary platform. Platforms like HackerOne, Bugcrowd, and OpenBugBounty sit between

researchers and vendors - setting program rules, facilitating communication, adjudicating disputes, processing payments. Their design decisions shape the ethical landscape within which disclosure occurs, and those design decisions are not neutral.

HackerOne and Bugcrowd are commercial platforms, and their financial incentives are aligned with the vendor - the vendor pays for the service, the researcher provides the labor. Follow the money and the consequences are predictable: dispute resolution processes that defer to the vendor's severity assessment, non-disclosure clauses that restrict researcher communication, and reputation systems that reward compliance over advocacy. When a researcher submits a finding and the vendor disputes its validity or downgrades its severity, the platform's institutional interest lies in resolving the dispute in the vendor's favor - because the vendor is the paying customer.

The platform does not correct the testimonial injustice described in Section 4 - it participates in it, because the service it offers depends on the continued satisfaction of the paying customer, and the paying customer is never the researcher. The mechanism goes further than financial incentive alignment. Moussouris (2021) has described how bug bounty platforms enforce non-disclosure agreements that bind researchers even when a submitted vulnerability is closed as "not a bug" and never fixed - the researcher cannot discuss the vulnerability, cannot warn users, and in some cases cannot even acknowledge the program's existence.

If a researcher complains publicly, the platform threatens to remove them entirely, cutting off access to every program on the platform and eliminating a source of income. As Moussouris observes, the platforms use the fact that they control the marketplace to coerce compliance from workers they do not employ - a dynamic she compares to gig economy labor exploitation, except

worse, because an Uber driver who accepts a fare at least gets paid for the trip. A bug bounty researcher can invest days of work, submit a valid finding, have it closed as a duplicate or out of scope, receive nothing, and be contractually prohibited from telling anyone about it.

Not all platforms operate in this fashion. OpenBugBounty, for example, is built on a fundamentally different model and occupies a different position. As a non-monetary platform focused on open vulnerability reporting, it provides researchers with public profiles, badges, rankings, and certificates - forms of recognition that function as professional portfolio elements. While it does not solve the financial incentive problem, it demonstrates that non-monetary recognition systems can sustain researcher engagement, particularly for early-career researchers building credibility (DudeTechItOut, 2025). Its transparency model - public reports, public researcher profiles, public organizational responses - provides a degree of accountability that closed commercial platforms lack. When the response is public, dismissing a valid finding carries reputational cost. When the response is private, dismissing a valid finding carries no cost at all.

The Nightmare Eclipse case reveals what happens when no effective intermediary exists. Microsoft's internal MSRC process served as both vulnerability receiver and judge, with no independent mediation. When the researcher-vendor relationship deteriorated, there was no neutral party to arbitrate, no escalation mechanism, and no institutional check on Microsoft's retaliatory actions. The researcher's only remaining option for being heard was public disclosure. A well-designed intermediary platform - one with genuine independence from the vendor, transparent dispute resolution, and meaningful researcher protections - could have prevented this escalation. The absence of such a platform is not ethically neutral; it is a structural gap in the disclosure ecosystem that predictably produces the kind of breakdown observed here.

7. Naming What the Current Model Gets Wrong

This paper does not claim to have the answer. What it offers is a way of naming obligations the current model leaves unnamed and evaluating conduct it currently evaluates in only one direction.

7.1 What the Current Model Fails to Name

The analysis above reveals a set of obligations the current model either ignores or leaves implied. Naming them is not the same as solving them - but you cannot fix what you refuse to name. These fifteen principles are organized in five categories: foundational obligations (1 - 3), epistemic obligations (4 - 6), operational obligations (7 - 9), distributive obligations (10 - 12), and structural obligations (13 - 15).

1. **Reciprocal Obligation:** Organizations and researchers have mutually constitutive moral obligations in vulnerability disclosure. We cannot judge the obligations of one party without considering the other's actions. For a researcher to have a duty to disclose responsibly, an organization must be willing to receive, evaluate, and act on disclosures in good faith. When one party neglects its obligations, the moral weight of the other's obligations is altered accordingly.

2. **Good-Faith Access:** Organizations must maintain functional, accessible reporting channels for researchers - and must not revoke, delete, or restrict access to those channels as a response to disputes. When Microsoft allegedly deleted Nightmare Eclipse's MSRC account, banned their GitHub repository on a platform Microsoft owns, and watched as GitLab followed suit, every legitimate channel through which the researcher could

cooperate was closed. Closing the door and then blaming the researcher for knocking louder is not a security posture. It is *retaliation by infrastructure* - the use of platform control, account management, and institutional access as instruments of punishment rather than engagement.

3. **Non-Retaliation:** Legal threats, platform suppression, account deletion, and other retaliation against good-faith researchers is a categorical violation of the disclosure compact, and it fundamentally changes the moral landscape in which the researcher's subsequent decisions are judged. Microsoft's Digital Crimes Unit statement - threatening to "continue bringing cases against these actors and those that enable their criminal activity" (Lyons, 2026a) - turned the dispute from a disagreement about process into an adversarial confrontation. Once an organization invokes criminal prosecution against a researcher, the relationship cannot return to cooperative disclosure. The threat itself is the damage.

4. **Epistemic Respect:** Organizations have an obligation to evaluate vulnerability reports on their technical merits and to provide transparent, reasoned justifications for severity assessments and remediation decisions. Failure to do so constitutes epistemic injustice as described by Fricker (2007). When Dormann, principal vulnerability analyst at Tharros, observes that MSRC replaced skilled evaluators with "flowchart followers" (Ferreira, 2026), the institutional capacity for epistemic respect has been structurally dismantled. A procedural gatekeeper cannot evaluate whether a finding has merit - they can only verify whether the submission meets formal requirements. When the submission does not fit the flowchart, it is rejected not because it lacks technical validity but because the institution has made itself unable to hear what the researcher is saying.

5. **Substantive Engagement / Tone Displacement:** When a researcher's findings are technically valid, the institutional response must engage with the substance of those findings - not deflect through the tone in which they were delivered. *Tone displacement* - the deflection of attention from the substance of a disclosure to the manner of its delivery - extends beyond the vendor to the industry observers, analysts, and commentators who shape public perception of these disputes. A vulnerability that warrants an emergency patch warrants a substantive response to the person who found it - and when the industry's collective response to eight working zero-day exploits in core Windows components is to debate the tone of a blog post, something has gone wrong that goes beyond one researcher and one vendor. That is a failure of epistemic responsibility across the entire ecosystem. Substantive Engagement is the obligation. *Tone displacement* is what happens when the obligation is abandoned - and in the Nightmare Eclipse case, it happened not just at the vendor level but across the entire institutional response.

6. **Transparent Remediation / Silent Patching:** When a vendor patches a vulnerability, the patch must be accompanied by public acknowledgment - a CVE, an advisory, and attribution where applicable. *Silent patching* - fixing a vulnerability without public disclosure or researcher credit - while publicly condemning the person who reported it is not transparency. It is the appearance of transparency deployed selectively. As Collin Hogue Spears, senior director of solution management at Black Duck, told Dark Reading: “It also means ending what researchers describe here: a flaw patched in silence, and then the finder blamed in public” (Montalbano, 2026). That is precisely the dynamic this principle is designed to name. RedSun was silently patched without a CVE for weeks (Parham, 2026). Askar’s VS Code vulnerability was “silently fixed... without any credit”

(Jones, 2026). When *silent patching* is the norm, researchers lose both the professional recognition and the empirical confirmation that their work mattered.

7. **Time-Bounded Remediation:** Organizations must respond to vulnerability reports within a defined, reasonable timeframe. The industry has already established this norm - Google Project Zero enforces 90 days (Willis, 2025), CERT/CC enforces 45 days (Moussouris, 2021), and Open Bug Bounty follows 90 days under ISO 29147. These deadlines exist because the security community has learned through decades of experience that indefinite silence is not a remediation strategy - it is an invitation to escalation. When an organization receives a vulnerability report and provides no response, no timeline, and no acknowledgment, it is not exercising caution. It is communicating to the researcher that their work has no value. The researcher's subsequent decisions are made in the context of that silence.

8. **Verification of Remediation:** A claimed fix must be an actual fix. When a vendor marks a vulnerability as patched, that patch must demonstrably resolve the issue. MiniPlasma - a vulnerability Microsoft claimed to have patched in December 2020 that remained fully exploitable on patched Windows 11 systems six years later (Bayram, 2026) - demonstrates that without external verification, claimed remediation cannot be taken at face value. This is not an abstract concern. Every user who updated their system in December 2020 believed they were protected. They were not. For six years, a vulnerability that Google Project Zero had reported and Microsoft had claimed to fix remained open. The only reason anyone discovered the failure was because an adversarial researcher tested the patch. The current model has no mechanism to ensure this testing occurs - and when it does not, the word "patched" becomes meaningless.

9. **Fair Compensation:** When a researcher's discovery is covered by an organization's bounty program, they should be paid the bounty. An organization doesn't get the security benefit of being able to find a vulnerability - fix it, protect its users, improve its product - without providing the compensation that its own program promises for that class of work. Using Microsoft's own published rates, eight endpoint zero-days would be worth between $240,000 and $800,000 in bounty payments (Okemwa, 2026). In this case, Microsoft received the security benefit of eight vulnerability discoveries and paid nothing for any of them. The bounty program exists precisely to pay for this kind of labor. When the work is accepted but the worker is not, the program is not functioning as a disclosure incentive - it is functioning as a mechanism for extracting unpaid security research.

10. **User-Centered Harm Accounting:** The moral responsibility for harm to end users resulting from public disclosure must be distributed between the researcher who discloses and the organization whose conduct precipitated the disclosure, in proportion to their respective causal contributions. The current model assigns the entirety of this responsibility to the researcher. This paper argues that when an organization ignores reports, deletes accounts, withholds compensation, and threatens legal action - and the researcher subsequently discloses publicly - the causal chain that produced the harm to end users begins with the organization's conduct, not the researcher's response to it. This does not absolve the researcher. It insists that moral responsibility be distributed rather than dumped on the party with less institutional power.

11. **Proportionate Response:** The researcher's disclosure decisions should be proportionate to the organization's engagement. Good-faith organizational engagement sustains the obligation to cooperate - bad-faith engagement erodes it. This principle operationalizes

Axelrod's (1984) finding that in iterated interactions, cooperation is sustained by reciprocity. The litmus test described in practitioner literature (DudeTechItOut, 2025) is the first cooperative move. When that cooperation is met with silence, dismissal, or hostility, the researcher's subsequent escalation is not a violation of the cooperative norm - it is the predictable consequence of the norm being violated by the other party first. Proportionate response does not mean equivalent retaliation. It means that the researcher's obligations to cooperate diminish as the organization's willingness to reciprocate diminishes.

12. **Duty of Care to Researchers:** Organizations owe a duty of care not only to their users but to the researchers whose labor improves the security of their products. When a researcher reports that their work has "drained [their] soul" (Montalbano, 2026) and acknowledges significant degradation of their "mental and physical health," the institutional conditions that produced that outcome bear moral weight. Semmer et al. (2015) predict exactly this trajectory - illegitimate tasks that communicate an "implicit message of disrespect" constitute a direct "threat to the self," producing strain, exhaustion, and psychological harm. Williams (2007) predicts the behavioral outcome - chronic ostracism depletes coping resources, producing depression and helplessness. The current model treats researcher wellbeing as an externality. It should be treated as an obligation.

13. **Structural Independence / Retaliation by Infrastructure:** When a single entity simultaneously controls the vulnerable software, the platform hosting the researcher's code, and the reporting infrastructure through which vulnerabilities are disclosed, the resulting concentration of power creates conditions under which retaliation can be

exercised across every surface of the researcher's professional life - what this paper terms *retaliation by infrastructure*. In the Nightmare Eclipse case, Microsoft is the vendor whose software is vulnerable, the owner of GitHub where the researcher's code was hosted, and the operator of MSRC through which vulnerabilities are reported. This convergence meant that a single institutional decision could eliminate the researcher's reporting channel, their code repository, and their professional platform simultaneously. The retaliation did not require a lawsuit or a cease-and-desist letter - it required only the exercise of administrative control over infrastructure the researcher depended on. This dynamic must be recognized as a structural conflict of interest - one that the current disclosure model does not acknowledge and has no mechanism to address.

14. **Intermediary Independence:** Third-party disclosure platforms bear an obligation to function as genuine mediators rather than vendor advocates, with transparent dispute resolution mechanisms and meaningful protections for researchers. As Moussouris (2021) has documented, commercial platforms enforce non-disclosure agreements that bind researchers even when findings are closed without remediation, threaten researchers with platform-wide exclusion for non-compliance, and structurally favor the vendor because the vendor is the paying customer. Akgul et al. (2023) confirmed empirically that researchers perceive mediation as biased - "bug-bounty platforms favor the companies that pay them to host" - and that poor platform support ranks among the most significant challenges in the ecosystem. A platform that silences the researcher on behalf of the vendor is not mediating. It is participating in the epistemic injustice the current model was supposed to prevent.

15. **Chaining Awareness:** Organizations need to evaluate vulnerability reports not only for their individual severity but for their potential to be part of exploit chains. A CVSS 4.0 finding dismissed in isolation might be a critical component in a privilege escalation, defense evasion, or encryption bypass chain - as demonstrated in this case study where UnDefend (CVSS 4.0) was a key part of an operational attack chain that included ransomware deployment (Barry, 2026). The PACMAN attack illustrated the same concept at the hardware level: a mitigated vulnerability was rendered fully exploitable when chained with a speculative execution side-channel attack (Ravichandran et al., 2022). An organization that dismisses low-severity findings on the assumption that they are inconsequential in isolation is making a bet that no attacker will think to combine them - a bet that the chaining literature, and this case study, shows is consistently lost.

7.2 From “Responsible Disclosure” to “Reciprocal Disclosure”

The concept of “responsible disclosure,” as currently deployed in industry and academic discourse, places normative weight exclusively on the researcher's conduct. It is the researcher who must be “responsible” - they must disclose privately, wait patiently, accept the vendor’s timeline, and refrain from public communication. The vendor’s corresponding obligations - to respond promptly, evaluate honestly, compensate fairly, and refrain from retaliation - are implied but not captured by the term.

When a vendor fails in these obligations, the language of “responsible disclosure” provides no vocabulary for naming that failure. The rhetorical asymmetry is not accidental. As Moussouris observed when Microsoft revived the term during the Nightmare Eclipse dispute, “no vendor uses that term unless they want to call someone irresponsible” (Martin, 2026).

Microsoft itself had formally retired the phrase in 2010 in favor of "Coordinated Vulnerability Disclosure" - yet reached for it again when it needed to delegitimize a researcher's conduct without examining its own. The term is, in Moussouris' word, "loaded" - it carries a normative judgment that applies in one direction only.

This paper proposes the term *reciprocal disclosure* as a normative reframing. Reciprocal disclosure names a process in which both parties bear articulated, symmetrical obligations and both parties' conduct is subject to ethical evaluation. Under reciprocal disclosure:

The researcher's obligations remain - private disclosure, technical documentation, reasonable patience. These are unchanged from the responsible disclosure model. What changes is that these obligations are explicitly conditioned on the vendor's reciprocal conduct.

The vendor's obligations are elevated from implied expectations to co-equal normative requirements - timely acknowledgment, good-faith evaluation, transparent communication, fair compensation, non-retaliation, and the maintenance of functional reporting channels.

The intermediary platform's obligations are articulated for the first time - independent dispute resolution, researcher protection, and resistance to vendor capture.

Failure by any party is named as failure, using the same evaluative framework. A vendor that deletes a researcher's reporting account has "failed in reciprocal disclosure" just as clearly as a researcher who publishes without notice. The vocabulary of reciprocity makes the bidirectional nature of the obligation explicit and prevents the rhetorical asymmetry embedded in the current "responsible disclosure" framing.

Whether the security industry adopts this language is not for this paper to determine. But if it did, organizations would be accountable for their conduct toward researchers in the same way researchers are currently accountable for their conduct toward organizations. The Nightmare Eclipse case would be understood not as a failure of responsible disclosure by a rogue researcher, but as a bilateral failure of reciprocal disclosure in which both parties - and critically, the millions of affected users - bear the consequences.

7.3 Precedents That Already Exist

This paper is not a policy proposal. It does not prescribe what organizations, platforms, or researchers should do - that work belongs to the people building and operating these systems. What it does is name the obligations the current model leaves unnamed, and provide a framework for evaluating conduct on both sides. That said, it is worth noting that precedents already exist for the kind of bidirectional accountability this framework describes.

Google's Project Zero operates under a 90-day disclosure policy - once a vulnerability is reported to a vendor, the vendor has 90 days to issue a patch before the details are published, regardless of whether a fix is ready. The policy has been controversial but widely respected, and it embodies a form of reciprocal obligation - the researcher commits to a defined waiting period, and the vendor is held to a defined remediation timeline.

If the vendor patches within the window, the disclosure is coordinated. If not, the researcher publishes. It is worth noting that Project Zero actively refuses to submit through bug bounty platform terms - when required to use a platform's ticketing system, researchers include disclaimers in their reports stating that they do not accept the platform's non-disclosure conditions (Moussouris, 2021).

Cisco's Talos research team follows the same practice. The most respected vulnerability research teams in the world have concluded that the current platform model is incompatible with good-faith disclosure. Additionally, CERT/CC - the US government's own coordination center - has maintained a 45-day vulnerability disclosure deadline for over a decade, imposing it on vendors who are non-responsive or evasive. As Moussouris (2021) has noted, anyone who objects to researchers publishing after a deadline should take that up with the United States government, which has been doing exactly that for years.

But even this model has friction points, and acknowledging them strengthens rather than weakens the case for reciprocal disclosure. Microsoft itself has publicly clashed with Project Zero over disclosure timelines, arguing that 90 days is insufficient for complex vulnerabilities in widely deployed systems. Vendors have accused Project Zero of prioritizing rigid deadlines over user safety.

These criticisms are not frivolous - some vulnerabilities genuinely require longer remediation windows, and a mechanical countdown does not account for complexity. What Project Zero's model gets right, however, is the principle - both parties know the rules, both parties bear consequences for their conduct, and the process is transparent. What it gets wrong is the rigidity. Reciprocal disclosure, as described here, improves on the Project Zero model by making the obligations genuinely bidirectional - not just a countdown timer but a framework in which the vendor's conduct during the window (communication, good faith, non-retaliation) is subject to the same ethical evaluation as the researcher's patience. Had the Nightmare Eclipse case occurred under even Project Zero's imperfect framework, the adversarial dynamic would have had no space to develop as it did.

An emerging reality makes this conversation more urgent - AI-accelerated vulnerability discovery. Many observers have pointed out that the traditional 90-day disclosure window was created for a slower era of security research. As one industry analyst pointed out, “the traditional 90-day disclosure embargo was created for a slower world” (PBXScience, 2026), and the availability of AI-powered research tools is dramatically speeding up and reducing the costs of vulnerability discovery. If the volume of vulnerability reports increases by orders of magnitude - as current trends suggest it will - the existing disclosure ecosystem's failures become unsustainable at scale.

The frustration is nothing new either - researchers have been saying this for years, and the industry was not listening. As early as 2009, security researchers were publicly declaring “no more free bugs,” arguing that the legal and professional risks of working with vendors outweighed the benefits and that researchers should sell to any paying customer or disclose publicly (Pupillo et al., 2018). Seventeen years later, the Nightmare Eclipse case demonstrates that the ecosystem has not only failed to resolve that tension - it has intensified it. The framework described here is not only ethically clearer than the current model - it may be practically unavoidable for a world in which the researcher-vendor relationship must function at far greater volume and speed than it does today.

There is a final irony that bears restating. Google’s Project Zero tells vendors - including Microsoft - that they have 90 days to patch or the vulnerability goes public, whether they like it or not, and the industry calls it a disclosure policy. CERT/CC does the same with a 45-day window and the industry calls it coordination. An independent researcher with no institutional backing does something functionally similar and the industry calls it criminal. The principle is identical. The only thing that changed is the power behind the person invoking it.

References

Abrams, L. (2026, June 9). Microsoft Defender 'RoguePlanet' zero-day grants SYSTEM privileges. BleepingComputer. https://www.bleepingcomputer.com/news/microsoft/microsoft-defender-rogueplanet-zero-day-grants-system-privileges/

Akgul, O., Eghtesad, T., Elazari, A., Gnawali, O., Grossklags, J., Mazurek, M. L., Votipka, D., & Laszka, A. (2023). Bug hunters' perspectives on the challenges and benefits of the bug bounty ecosystem. In Proceedings of the 32nd USENIX Security Symposium (pp. 2275–2291). USENIX Association. https://www.usenix.org/system/files/usenixsecurity23-akgul.pdf

Arghire, I. (2026, June 10). New Windows zero-day exploit 'RoguePlanet' released. SecurityWeek. https://www.securityweek.com/new-windows-zero-day-exploit-rogueplanet-released/

Axelrod, R. (1984). The evolution of cooperation. Basic Books.

Barry, C. (2026, May 19). Nightmare-Eclipse: Six zero-days, six weeks and one big grudge. Barracuda Networks Blog. https://blog.barracuda.com/2026/05/19/nightmare-eclipse-zero-days-grudge

Bayram, U. (2026, June 1). The return of a ghost: Unpacking the MiniPlasma zero-day exploit. Picus Security Blog. https://www.picussecurity.com/resource/blog/the-return-of-a-ghost-unpacking-the-miniplasma-zero-day-exploit

Cialdini, R. B. (2006, December 6). The gentle science of persuasion, part two: Reciprocity. W. P. Carey School of Business, Arizona State University. https://news.wpcarey.asu.edu/20061206-gentle-science-persuasion-part-two-reciprocity

CiphersSecurity. (2026, June 2). Nightmare Eclipse Microsoft Windows zero-days 2026. CiphersSecurity. https://cipherssecurity.com/nightmare-eclipse-microsoft-windows-zero-day/

CISA. (2021, November 3). Binding Operational Directive 22-01: Reducing the significant risk of known exploited vulnerabilities. Cybersecurity and Infrastructure Security Agency. https://www.cisa.gov/news-events/directives/bod-22-01-reducing-significant-risk-known-exploited-vulnerabilities

CyberNews. (2026, June 11). Vengeful researcher takes third Microsoft Patch Tuesday sucker punch, posts zero-day exploit on GitHub. CyberNews. https://cybernews.com/security/nightmare-eclipse-rogueplanet-zero-day/

DudeTechItOut. (2025, December 22). Researchers' litmus: Good faith or goodbye. DudeTechItOut. https://dudetechitout.com/posts/researchers-litmus-good-faith-or-goodbye.html

Ferreira, B. (2026, May 27). Microsoft's GitHub bans security researcher who posted zero-day Windows exploits because company ruined their life. Tom's Hardware. https://www.tomshardware.com/tech-industry/cyber-security/microsofts-github-bans-security-researcher-who-posted-zero-day-windows-exploits-because-company-ruined-their-life-expert-claims-action-is-vindictive-and-promises-further-retaliation

Fricker, M. (2007). Epistemic injustice: Power and the ethics of knowing. Oxford University Press.

International Organization for Standardization. (2018). Information technology – Security techniques – Vulnerability disclosure (ISO/IEC 29147:2018). ISO.

International Organization for Standardization. (2019). Information technology – Security techniques – Vulnerability handling processes (ISO/IEC 30111:2019). ISO.

Jones, C. (2026, June 3). Another bug hunter leaks Microsoft exploits in defiance of company's handling of vulnerability disclosures. The Register. https://www.theregister.com/security/2026/06/03/another-bug-hunter-leaks-microsoft-exploits-in-defiance-of-companys-handling-of-vulnerability-disclosures/5250590

Lyons, J. (2026a, May 28). Disgruntled 0-day hunter 'humiliated' by Microsoft pledges 'bone shattering drop' as Redmond calls cops. The Register. https://www.theregister.com/security/2026/05/28/microsoft-0-day-feud-escalates-as-researcher-threatens-another-windows-exploit-dump/5248085

Lyons, J. (2026b, June 11). Microsoft's worst 'Nightmare' unleashes BitLocker bypass 0-day. The Register. https://www.theregister.com/security/2026/06/11/nightmare-eclipse-drops-claimed-bitlocker-bypass-for-microsoft-windows/5254371

Martin, A. (2026, June 1). Microsoft says it will not pursue security researchers after zero-day backlash. The Record from Recorded Future News. https://therecord.media/microsoft-says-it-will-not-pursue-security-researchers-disclosure

Mehdi, Y. (2025, June 24). Stay secure with Windows 11, Copilot+ PCs and Windows 365 before support ends for Windows 10. Windows Experience Blog. https://blogs.windows.com/windowsexperience/2025/06/24/stay-secure-with-windows-11-copilot-pcs-and-windows-365-before-support-ends-for-windows-10/

Montalbano, E. (2026, June 10). Nightmare-Eclipse drops yet another Microsoft exploit, RoguePlanet. Dark Reading. https://www.darkreading.com/vulnerabilities-threats/nightmare-eclipse-microsoft-exploit-rogueplanet

Moussouris, K. (2021). Coordinated vulnerability disclosure and the problem with bug bounty platforms. TechSpective. https://techspective.net/2021/06/21/katie-moussouris-coordinated-vulnerability-disclosure-and-the-problem-with-bug-bounty-platforms/

Okemwa, K. (2026, May 28). "I have proof for every single word": This security researcher's GitHub and Microsoft accounts were deleted. Windows Central. https://www.windowscentral.com/microsoft/security-researcher-github-microsoft-accounts-deleted-windows-11-exploit-bitlocker

Parham, A. (2026, May 21). Microsoft Defender zero-days patched: RedSun, UnDefend exploits already used in live intrusions. TechTimes. https://www.techtimes.com/articles/316957/20260521/microsoft-defender-zero-days-patched-redsun-undefend-exploits-already-used-live-intrusions.htm

PBXScience. (2026, June 25). Microsoft under fire as researcher "Nightmare-Eclipse" drops six Windows zero-days without warning. PBXScience. https://pbxscience.com/microsoft-under-fire-as-researcher-nightmare-eclipse-drops-six-windows-zero-days-without-warning/

Pupillo, L., Ferreira, A., & Varisco, G. (2018). Software vulnerability disclosure in Europe: Technology, policies and legal challenges. CEPS Task Force Reports. Centre for European Policy Studies. https://www.ceps.eu/ceps-publications/software-vulnerability-disclosure-europe-technology-policies-and-legal-challenges/

Ravichandran, J., Na, W. T., Lang, J., & Yan, M. (2022). PACMAN: Attacking ARM pointer authentication with speculative execution. In Proceedings of the 49th Annual International Symposium on Computer Architecture (ISCA '22) (pp. 685–698). ACM. https://doi.org/10.1145/3470496.3527429

Scanlon, T. M. (1998). What we owe to each other. Belknap Press.

Semmer, N. K., Jacobshagen, N., Meier, L. L., Elfering, A., Beehr, T. A., Kälin, W., & Tschan, F. (2015). Illegitimate tasks as a source of work stress. Work & Stress, 29(1), 32–56. https://doi.org/10.1080/02678373.2014.1003996

StatCounter. (2026). Desktop operating system market share worldwide. StatCounter Global Stats. https://gs.statcounter.com/os-market-share/desktop/worldwide/

Toulas, B. (2026, April 6). Disgruntled researcher leaks "BlueHammer" Windows zero-day exploit. BleepingComputer. https://www.bleepingcomputer.com/news/security/disgruntled-researcher-leaks-bluehammer-windows-zero-day-exploit/

Williams, K. D. (2007). Ostracism. Annual Review of Psychology, 58(1), 425–452. https://doi.org/10.1146/annurev.psych.58.110405.085641

Willis, T. (2025, July 29). Policy and disclosure: 2025 edition. Google Project Zero Blog. https://projectzero.google/2025/07/reporting-transparency.html