Plausibly Deniable Harm: How Everyday Engineering Practices Produce Hidden Risk

Jennifer R. Ayres

Independent Engineering Consultant, Adelaide, South Australia.

ORCID ID 0000-0002-4538-6512.

Abstract

Harm in engineering and infrastructure organisations is commonly attributed to failure, misconduct, or technical breakdown. This paper argues that harm can also arise through ordinary, rule-following practices that appear responsible on paper. Drawing on science and technology studies scholarship, it introduces the concept of plausibly deniable harm to describe situations in which operational and experiential knowledge is systematically excluded through formally legitimate organisational processes, producing foreseeable adverse outcomes that remain administratively unclaimable.

Through conceptual analysis informed by 25+ years of commissioning practice, the paper identifies five mechanisms through which this exclusion occurs: representational abstraction (knowledge translation losses), procedural compliance (substituting process for engagement), temporal deferral (normalising delay), credibility asymmetries (privileging formal over experiential authority), and interactional ambiguity (softening dissent through professional courtesy). These mechanisms interact to produce harm that is formally compliant yet substantively consequential.

By reframing harm as an emergent feature of compliance-oriented governance rather than exceptional failure, the paper extends science and technology studies (STS) debates on responsibility, accountability, and care in sociotechnical systems. It offers an analytical framework for understanding how organisations maintain procedural accountability while remaining substantively unresponsive, with implications for infrastructure governance, organisational design, and future empirical research on epistemic inequality in high-stakes environments.

Keywords: Plausibly deniable harm; Epistemic inequality; Governance and accountability; Operational knowledge; Invisible work; Sociotechnical systems; Science and technology studies

1. Introduction

Across contemporary engineering and infrastructure systems, organisational harm is often attributed to failure, misconduct, or technical breakdowns. Yet science and technology studies (STS) scholarship has long shown that system reliability and safety emerge from sociotechnical arrangements that depend on forms of labour, judgement, and coordination that are routinely undervalued or rendered invisible (Star & Strauss, 1999). This work demonstrates that what keeps systems functioning is not simply technical design, but the ongoing, situated practices through which people interpret, negotiate, and sustain complex infrastructures (Suchman, 1995).

At the same time, STS research has highlighted that not all knowledge within organisations is treated equally. Decisions about what counts as legitimate expertise, whose judgement is trusted, and which concerns are allowed to shape action are structured by institutional hierarchies, representational norms, and governance arrangements (Irwin & Wynne, 1996; Jasanoff, 2004). These dynamics produce systematic asymmetries in whose knowledge is recognised, recorded, or acted upon — what this paper terms epistemic inequality (Bowker & Star, 2008; Cech, 2014; Collins & Evans, 2002).

While existing literatures illuminate how knowledge becomes marginalised, it has been less explicit about how harm is produced and sustained in contexts where exclusion is neither intentional nor overt. In many engineering organisations, harm does not arise from rule-breaking or negligence. Instead, it emerges through ordinary, formally compliant practices that allow organisations to appear diligent while remaining substantively unresponsive to operational realities (Power, 2013). These are environments in which concerns can be acknowledged without being incorporated, and where procedural order can mask the displacement of risk (Mol, 2008; Strathern, 2000).

This paper develops the concept of plausibly deniable harm to describe this patterned organisational outcome. Plausibly deniable harm refers to situations in which foreseeable adverse consequences become administratively unclaimable because decision-making processes adhere to recognised procedures that systematically exclude certain forms of knowledge (Perrow, 2011). The harm is real, but its attribution is diffused through documentation, role separation, and interactional norms that make exclusion appear reasonable, professional, or procedurally justified (Fricker, 2007).

The paper asks a straightforward question that anchors the conceptual contribution:

Why do engineering organisations sometimes arrive at harmful outcomes even when they comply with all required procedures, and how does the marginalisation of operational knowledge contribute to this?

To answer this question, the paper offers a conceptual analysis grounded in STS scholarship on invisible work, expertise, governance, and accountability (Jasanoff, 2004; Suchman, 1995).

The paper proceeds as follows. Section 2 situates plausibly deniable harm within existing STS debates and clarifies its conceptual distinctiveness. Section 3 outlines the analytical approach used to develop the mechanisms. Section 4 elaborates these mechanisms and shows how they interact to produce organisational conditions in which harm becomes both predictable and difficult to contest. Section 5 discusses the implications for STS theories of governance, expertise, and care. Section 6 concludes by reflecting on the conceptual and

practical significance of recognising plausibly deniable harm as a sociotechnical phenomenon. By theorising harm as an emergent property of epistemic exclusion rather than procedural failure, this analysis extends STS scholarship on the politics of knowledge in technoscientific organisations.

2. Theoretical positioning: epistemic inequality and organisational harm

STS scholarship has long challenged the assumption that stability, safety, and system performance naturally follow from technical design alone. Foundational work on invisible labour and maintenance demonstrates that much of what keeps sociotechnical systems functioning is easy to overlook precisely because it works (Star & Strauss, 1999). Tasks that prevent failure, absorb disruption, or compensate for design limitations often disappear from formal accounts of system operation, becoming visible only when something breaks (Orr, 2016; Suchman, 1995). This insight has been especially influential in infrastructure studies, where maintenance and operational care are routinely overshadowed by innovation, optimisation, and expansion (Graham & Thrift, 2007). Operational knowledge in commissioning is often tacit, situational, and difficult to formalise, making it especially vulnerable to exclusions (Denis & Pontille, 2014).

Subsequent scholarship has shown that invisibility is not merely descriptive but political. Labelling certain activities as routine, low-skill, or non-expert reinforces hierarchies of value and authority (Bowker & Star, 2008). In regulated engineering contexts, the work of operators, maintainers, and coordinators is often framed as execution rather than judgement, even when it requires complex situational assessment (Irwin & Wynne, 1996). As a result, knowledge generated through everyday operational engagement is frequently excluded from formal decision-making processes (Bowker & Star, 2008).

Parallel work on expertise examines how organisations distinguish between legitimate and illegitimate knowledge (Collins & Evans, 2002). Expertise is not simply a matter of competence or experience, but of alignment with recognised categories, credentials, and representational forms. Knowledge that can be abstracted, codified, and circulated through documents tends to carry greater authority than experiential or contextual knowledge (Porter, 1995). This privileging of abstraction is especially pronounced in large technical organisations, where governance relies heavily on metrics, plans, and compliance artefacts. (Power, 2013)

The result is epistemic inequality: a systematic imbalance in whose knowledge counts, when it counts, and on what terms (Bowker & Star, 2008). Unlike individual bias or misrecognition, epistemic inequality is embedded in organisational routines, documentation practices, and accountability structures (Hull, 2012). It persists even in the absence of discriminatory intent because institutionalised representational forms already shape what counts as valid evidence.

Work on epistemic injustice provides a complementary but distinct lens (Fricker, 2007). Testimonial and hermeneutical injustice highlight how individuals and groups can be wronged in their capacity as knowers. However, this literature often focuses on interpersonal encounters or cultural marginalisation rather than the organisational mechanisms that normalise exclusion (Medina, 2013). Epistemic injustice concerns interpersonal wrongs; epistemic inequality concerns structural arrangements that determine whose knowledge is institutionally actionable. This distinction is central to the argument developed here (Fricker, 2007). While epistemic injustice captures interpersonal credibility

deficits, commissioning environments also reveal deeper structural patterns of exclusion that exceed interpersonal dynamics.

2.1 From recognition to accountability: epistemic inequality as organisational harm

Frameworks focused on misrecognition or interpersonal injustice help explain how certain knowers are disadvantaged (Fricker, 2007). Still, they do not explain how organisations maintain formal legitimacy while remaining substantively unresponsive to operational knowledge. To address this gap, this section shifts the focus from recognition to accountability, examining how epistemic inequality functions as a mechanism of organisational harm.

In many engineering organisations, decision-making processes are designed to demonstrate that the correct steps have been followed rather than to ensure that all relevant knowledge has been engaged. Documents such as plans, risk registers, and compliance reports serve as proxies for due diligence, allowing organisations to claim responsibility without accounting for operational realities (Power, 2013). In this context, epistemic inequality does not merely disadvantage certain knowers; it shapes how responsibility is distributed and how harm becomes contestable.

2.1.1 Plausibly deniable harm as emergent

Plausibly deniable harm is conceptualised here as an emergent organisational outcome rather than a deliberate strategy. It arises when routine, formally legitimate practices systematically exclude operational knowledge in ways that make harm foreseeable yet administratively unclaimable. Unlike “normal accidents,” which emerge from system complexity, plausibly deniable harm emerges from epistemic inequality embedded in governance arrangements (Perrow, 2011). The term 'deniable' refers not to deliberate denial but to the structural condition whereby harm becomes administratively unclaimable; organisations cannot be held accountable because they followed all procedures correctly.

Plausibly deniable harm describes organisational conditions in which foreseeable adverse outcomes can be disclaimed because decision-making processes have adhered to recognised procedures, even when those procedures exclude critical forms of knowledge. Harm is not denied because it is unforeseeable, but because it falls outside the epistemic boundaries of what the organisation has defined as relevant. The harm is real, but its attribution is diffused through layers of documentation, role separation, and procedural compliance (Strathern, 2000).

2.2 Differentiating plausibly deniable harm from existing accounts of organisational harm

The concept of plausibly deniable harm builds on, but is distinct from, two influential accounts of organisational harm: Turner’s (1976) analysis of the incubation period preceding disasters, and Vaughan’s (1999) account of normalised deviance. Both have shaped STS work on accountability, risk, and culture, and engaging with them clarifies what is distinctive about the present contribution.

For Turner, harm develops over extended periods as warning signs accumulate across organisational boundaries without being assembled into a coherent picture of risk. The locus of harm production lies in the organisation’s failure to integrate dispersed signals. For Vaughan, anomalies are noticed by operational and engineering staff but are progressively

rationalised and reclassified as acceptable; deviance is produced at the operational level, through the situated meaning-making of those closest to the technology.

Plausibly deniable harm identifies a third pattern that has been less explicitly theorised. It locates the mechanism of harm production neither in failed assembly of warning signs nor in operational rationalisation of anomalies, but in the structural and documentary processes through which situated knowledge is filtered, softened, and rendered administratively inactionable as it moves from operational settings into formal governance arrangements. In the cases that motivate this analysis, operators neither fail to notice anomalies nor rationalise them as normal. They notice them, interpret them accurately, raise them through legitimate organisational channels — and their concerns are received, recorded, and then dissolved through procedural and representational practices that maintain compliance while excluding the substance of what was raised. The harm lies not in unnoticed signals or in normalised deviation, but in the organisational architecture that processes situated knowledge into administrative irrelevance.

This differentiation clarifies the concept of deniability. Deniability here does not refer to deliberate concealment, motivated denial, or the gap between what was knowable and what was known. It describes a structural condition: harm becomes administratively unclaimable because every step in the decision chain produced documentation, followed established procedure, and engaged recognised expertise. There is, in the formal record, no procedural lapse to point to. This structural deniability has a corollary that complicates conventional accounts of accountability: where harm cannot be attributed to procedural failure, responsibility is often unfairly displaced onto individual operators or managers whose situated knowledge was, in fact, the resource the organisation failed to engage. Structural deniability and individual scapegoating are two manifestations of the same epistemic architecture — one that distributes accountability through documents while disclaiming it through the same documentary forms (Bovens, 2007).

The contribution is therefore complementary rather than competing. Turner's incubation period helps explain harm in settings where warning signs disperse across organisational boundaries; Vaughan's normalised deviance helps explain harm where operational interpretation drifts. Plausibly deniable harm helps explain a third pattern, prevalent in highly proceduralised compliance-oriented engineering contexts, where situated knowledge is present, accurate, and raised, but is filtered out of decision-making through formally legitimate organisational practices. Table 1 summarises these distinctions.

Table 1. Differentiating plausibly deniable harm from adjacent accounts of organisational harm

	Turner (1976): Incubation period	Vaughan (1996, 1999): Normalised deviance	This paper: Plausibly deniable harm
Where harm is produced	Failed assembly of dispersed warning signs	Operational rationalisation of anomalies	Structural filtering of situated knowledge through documentary and procedural forms
Status of operational knowledge	Exists in fragments; not aggregated	Present but progressively reinterpreted as acceptable	Present, accurate, and raised — but rendered administratively inactionable
Locus of deviance	Distributed across organisational boundaries	Operational meaning-making	Governance architecture
Role of compliance	Compliance is not the focus	Compliance practices conceal drift	Compliance is the mechanism through which exclusion occurs
Form of accountability after harm	Hindsight reveals missed signals	Hindsight reveals normalised anomalies	Hindsight reveals procedural compliance with substantive exclusion

3. Analytical approach and organisational context

This paper develops a conceptual framework grounded in twenty-five years of commissioning practice across water, wastewater, oil and gas, and infrastructure projects in the United Kingdom, Australia, and Southeast Asia. Commissioning is the phase of engineering work that bridges design, construction, and operations: it involves verifying that built systems function as intended, integrating subsystems, identifying defects, and preparing facilities for handover to operators. It is structurally distinctive within engineering organisations because commissioning practitioners occupy an interface position between multiple actors with different forms of authority — designers, contractors, principal clients, future operators, and regulators — while typically holding little formal decision-making authority of their own (Pollock & Williams, 2008). This position offers a particular vantage point: extended exposure to how organisational processes encounter the material behaviour of complex systems, and to which forms of knowledge are recognised, recorded, or set aside when the two come into tension.

The analysis offered here is best characterised as practice-informed conceptual development rather than systematic empirical research. It is informed throughout by recurring patterns observed across multiple projects and organisational settings, but it does not present a formal empirical dataset and does not claim to measure prevalence, establish causation, or test the proposed mechanisms against a defined sample. The contribution is conceptual: to articulate a pattern that, once named, becomes recognisable in a range of contexts and available for empirical examination.

3.1 Status of the illustrative material

The vignettes presented throughout this paper are composite illustrations drawn from patterns observed across multiple commissioning projects. They are not accounts of single, attributable events. Each scenario assembles elements that recurred across different settings — different facilities, different organisations, different jurisdictions — to render visible an organisational dynamic that operates similarly across them. This compositing serves two purposes. First, it protects the identities of organisations, individuals, and projects, none of which could be identified without breaching the implicit confidentiality of professional commissioning practice. Second, and more importantly for the conceptual argument, it allows the illustrations to foreground the organisational pattern rather than the particularities of any single case. The argument is not that any one project produced any

one outcome through these mechanisms, but that these mechanisms produce a recognisable patterning of outcomes across the kinds of organisational settings in which commissioning takes place.

This approach aligns with established traditions in STS and organisational ethnography that develop conceptual analysis from extended practitioner engagement (Orr, 2016; Suchman, 1995), and with work that uses composite or archetypal cases to surface dynamics that would otherwise remain invisible across distributed organisational settings (Star and Strauss, 1999). The trade-off is acknowledged: composite illustrations cannot be returned to for further empirical examination, and the analysis cannot adjudicate the prevalence of particular mechanisms in particular settings. What composite illustrations can do is make a structural pattern legible, and that is what this paper sets out to do.

3.2 Organisational context: commissioning, contract, and responsibility

The organisational settings on which this analysis draws share several structural features that are relevant to the argument. Commissioning typically occurs within multi-party contractual arrangements involving a principal client (often a utility, asset owner, or public body), a main contractor, specialist subcontractors, an eventual operator (sometimes a separate organisation, sometimes the client), and external regulators. Responsibility for system performance is distributed across these parties through formal contractual mechanisms — design responsibility, build responsibility, commissioning acceptance, performance warranty, operational responsibility — each with its own documentation, milestones, and governance arrangements. This distribution is not incidental to the argument but central to it: the mechanisms developed in Section 4 operate in part because responsibility is fragmented across organisational boundaries in ways that make any single actor's accountability for substantive (as opposed to procedural) outcomes difficult to locate.

This contrasts with what might be called the classical bureaucratic-hierarchical model of organisational responsibility, in which command structures integrate accountability vertically within a single organisation. In commissioning contexts, by contrast, the question of who denies, to whom, and with what consequences rarely has a single answer. A concern raised by a commissioning operator may be received by a project manager, recorded in a register administered by a third party, deferred to a future warranty period managed by a different organisation, and finally absorbed into the operational baseline of yet another. The deniability described in this paper is, in part, a property of this distribution: harm becomes administratively unclaimable not because any single party denies it, but because the architecture across parties provides no location at which substantive accountability can be assembled.

The forms of harm at issue across these settings include environmental discharge (effluent quality failures, contamination events), occupational injury (incidents during operation and maintenance, often arising from design or commissioning decisions made earlier), reliability degradation (premature equipment failure, asset life shortening, eventual public-service disruption), and the cumulative attrition of operational expertise as those whose knowledge is repeatedly filtered out disengage or leave the sector. These harms vary in their visibility, attributability, and consequence; what they share is that each can emerge through the mechanisms described below while every step in the decision chain remains formally defensible.

3.3 Documents as sociotechnical artefacts

Within this context, documents are not neutral records but active sociotechnical artefacts that shape what decisions are possible, whose concerns are considered relevant, and how accountability is judged (Asdal, 2015; Hetherington, 2012; Hull, 2012). Plans, risk registers, acceptance certificates, formal correspondence, and governance records stabilise categories, translate situated activity into abstracted representations, and enable coordination across time and organisational boundaries (Callon and Law, 1982). They also function as governance tools that allow organisations to demonstrate compliance and control in high-uncertainty environments (Power, 2013). This dual role — representational and regulatory — makes documentary practice a primary site at which the mechanisms developed in Section 4 operate. The analysis treats documents not as objects of systematic content analysis but as the medium through which the dynamics under consideration are enacted.

3.4 Scope of the framework and status of the mechanisms

Five mechanisms are developed in Section 4: representational abstraction, procedural compliance, temporal deferral, credibility asymmetries, and interactional ambiguity. These are not claimed to be an exhaustive list. They are the five mechanisms that have been most consistently observed across the commissioning contexts that inform this analysis, and that appear to interact in ways that produce the patterned outcome the paper terms plausibly deniable harm. Other mechanisms — for example, audit cycles, financial-incentive structures, professional indemnity arrangements, or regulatory enforcement gaps — almost certainly contribute to the production of similar outcomes in adjacent settings, and would warrant further analysis. The framework offered here is a starting point for that broader analysis rather than its conclusion.

The framework's transferability to other sectors is addressed in Section 5.4. Briefly: the mechanisms appear most likely to operate in organisational contexts that combine documentary accountability, distributed responsibility across multiple parties, asymmetric authority between formal expertise and situated knowledge, and high-stakes outcomes whose attribution is complicated by procedural compliance. Engineering and infrastructure are paradigmatic cases, but not the only ones. The analysis offered here makes no claim about how these mechanisms operate, or whether they operate at all, in classically hierarchical organisational settings or in settings without strong documentary governance.

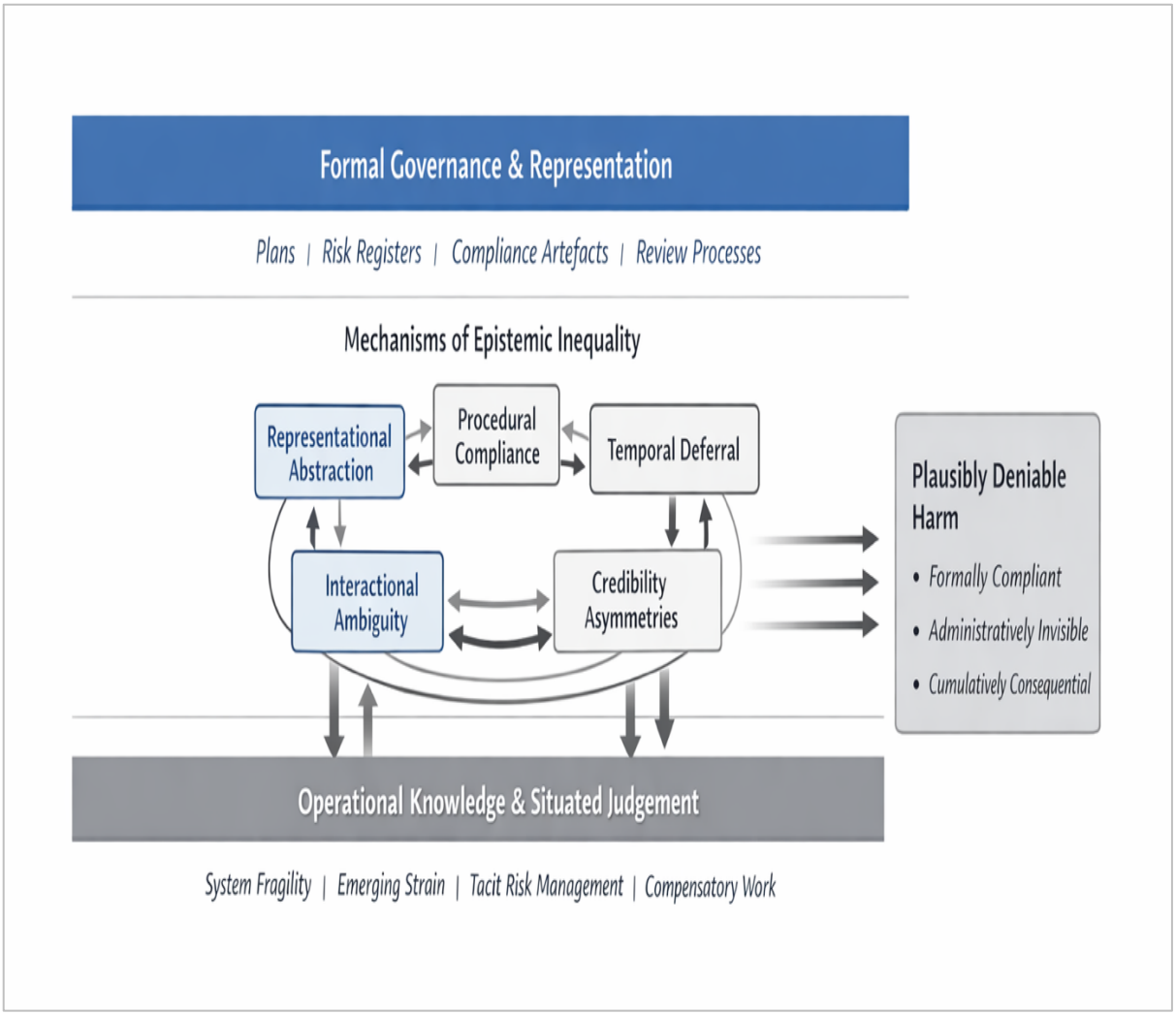

Figure 1. Plausibly deniable harm as an emergent outcome of epistemic inequality

The diagram illustrates how five mechanisms — representational abstraction, procedural compliance, temporal deferral, credibility asymmetries, and interactional ambiguity — operate between formal governance structures and operational knowledge to produce harm that is formally compliant yet substantively consequential.

4. Mechanisms of Epistemic Inequality and Plausibly Deniable Harm

As prior organisational scholarship shows, formal consultation and participation processes can satisfy procedural expectations without meaningfully incorporating the knowledge they solicit, particularly when engagement is bounded by predefined scopes and timelines (Sharma & Bansal, 2020).

Engineering organisations rarely intend to ignore frontline knowledge or make decisions that increase risk. Yet harm can still accumulate through ordinary, compliant, and seemingly reasonable practices. This section explains how that happens. It shows how everyday organisational routines—documentation, planning, consultation, and review—can unintentionally filter out the knowledge needed to prevent harm, while still producing a defensible record of diligence. The mechanisms described here (and summarised in Table 2) are analytically distinct, but in practice they reinforce one another. When they operate together, they create a system where operational insight is softened, delayed, or discounted,

while decisions remain entirely defensible on paper. Plausibly deniable harm does not arise from a single weak point but from the interaction of multiple routine practices that make certain forms of knowledge administratively invisible.

4.1 Representational abstraction and knowledge translation

Engineering organisations rely heavily on documents to move information across teams and decision layers. Plans, schedules, risk registers, and performance reports are treated as the authoritative record of what is known. For operational staff, however, knowledge is often contextual, contingent, and grounded in how systems behave under real-world conditions (Bechky, 2003). When this knowledge is translated into standard templates, much of its nuance is lost. Templates reward clarity, quantification, and alignment with predefined categories (Henderson, 1991; Sykianakis & Bellas, 2011). They leave little room for ambiguity, emerging risk, or the "messy" details that matter most in practice (Porter, 1995).

These patterns take archetypal forms. For instance, in water treatment commissioning, operational observations of pump behaviour under variable load show that this filtering effect is subtle but powerful. A technician's observation that a system "sounds different under load" becomes a generic note about "monitoring performance." A pattern of near-misses becomes a single risk line with a colour code. A complex interaction between ageing infrastructure and seasonal demand is reduced to a simple dependency arrow in a project plan. These transformations make information easier to circulate but harder to interpret. They also create a false sense of completeness: once something is written down, it appears to have been captured, even if the representational form has stripped away the very qualities that made the insight meaningful (Alcadipani et al., 2012).

This is not a failure of individuals but a structural feature of documentation-driven governance. Documents are designed to be stable, portable, and legible to people who were not present when the knowledge was generated. In achieving that stability, they inevitably flatten the situated judgement on which operational work depends. Decisions made based on these sanitised documents appear reasonable and well-informed, even when they miss critical operational realities. Documentation becomes both a communication tool and a shield: it demonstrates diligence while obscuring the limits of understanding.

Not all representational abstraction operates in this way. Engineering work depends on forms of abstraction that make complex systems legible across distance, time, and disciplinary boundaries — drawings, schematics, specifications, calculations — and these abstractions are often enabling rather than smothering (Henderson, 1991). The distinction that matters for the present argument is not between abstraction and non-abstraction but between forms of abstraction that preserve traceability back to situated knowledge and forms that sever it. A schematic that an operator can interrogate, annotate, and contest is a different artefact, in its organisational consequences, from a compliance summary that has aggregated, recategorised, and reformatted a series of observations into a status indicator. The first remains in dialogue with the situated knowledge it represents; the second has been stabilised into a form designed for circulation rather than for return.

A further dynamic deserves explicit acknowledgement. As operational staff become accustomed to working within documentary regimes that recognise only certain forms of expression, their own articulation of concerns can come to be shaped by what those regimes will accept (Vaughan, 1996). A technician who has learned that observations expressed as "the pump sounds different under load" will be dismissed as anecdotal, while observations expressed as "vibration amplitude trending upward against design baseline" will be

recorded, may, over time, come to express concerns in the second form even when the first more accurately captures what is known. The situated knowledge has not disappeared, but its public articulation has been colonised by the documentary forms through which it must travel to be heard. This colonisation is itself a mechanism of plausibly deniable harm: the organisation receives concerns in a form it can process, but the form has already filtered out the qualities of the original observation that made it actionable.

4.2 Procedural compliance as a substitute for engagement

Most engineering organisations have well-defined processes for consultation, review, and sign-off. These processes are designed to ensure that relevant voices are heard. In practice, procedural compliance can become a substitute for genuine engagement with operational knowledge (Power, 2004, 2013). Operational staff may be invited to provide feedback, but the process's structure often limits what can be incorporated. Consultation windows may be too short, feedback may need to fit predefined categories, and issues outside the project scope may be deemed inapplicable. Concerns that require cross-team coordination may be acknowledged but deferred. When harm later occurs, organisations can point to the completed process as evidence that they acted responsibly.

During water treatment facility design reviews, operational staff identify critical sequencing issues—startup procedures that require simultaneous valve operations beyond night-shift staffing, or emergency shutdown sequences that are incompatible with actual access routes. These concerns are raised during formal consultation, recorded, and acknowledged. However, governance structures lack mechanisms to revisit design decisions past procurement milestones. The feedback is categorised as "operational" rather than "design-critical" and is deferred from design. When staffing limitations later become apparent, documented consultation serves as evidence of proper process, even as substantive design limitations persist.

4.3 Temporal deferral and the management of urgency

Time is one of the most powerful and least visible mechanisms shaping how knowledge is valued. Operational concerns often relate to slow degradation, cumulative risk, or capacity limits that do not align with project timelines or reporting cycles (Otsuki, 2024). A typical organisational response is deferral. Issues are framed as not urgent, requiring further evidence, better addressed in the next phase, or outside the current decision window. Deferral allows decisions to proceed without explicitly rejecting operational knowledge. It maintains the appearance of prudence—“we need more data”—while postponing action indefinitely (Asdal & Hobæk, 2020; Barua, 2024).

Temporal deferral is particularly potent because it is rarely experienced as a decision at all. It is experienced as waiting, monitoring, gathering more information, or aligning with the next governance cycle. Yet these delays shape the conditions under which harm becomes possible. By the time a concern becomes urgent enough to act on, the organisation may have lost the opportunity to intervene effectively.

Temporal deferral operates powerfully through defect classification and warranty systems. During pre-operational testing, an operator notices intermittent pump cavitation—not severe enough to fail acceptance criteria but suggesting bearing wear that will worsen under continuous operation. The observation is recorded and categorised as "monitoring required," and the system is accepted because its current performance meets specifications. Six months later, during warranty review, the observation has been superseded by newer issues, the commissioning team has moved on, and context is lost. When the pump fails 18

months after handover, the warranty has expired, failure is classified as normal operational wear, and the early warning lies buried in closed-out commissioning records. A foreseeable reliability concern becomes normalised into the operational baseline, its progression documented but never arrested. Harm emerges gradually; attribution diffuses across multiple reasonable deferrals.

4.4 Asymmetries of credibility and authority

Not all knowledge carries the same weight in organisational decision-making. Formal authority, professional status, and proximity to decision structures often matter more than experiential familiarity with system behaviour. This creates predictable credibility asymmetries. Operational staff may understand how systems behave under strain but lack the authority to have their assessments treated as decisive (Wynne, 1992). Analysts or managers may produce abstract assessments that align with recognised expertise, giving their views greater legitimacy. Concerns raised informally or verbally may be discounted because they do not appear in the official record (Alcadipani et al., 2012).

Decisions can be framed as professional judgement rather than exclusion. This is a key pathway through which plausibly deniable harm is produced: the organisation can claim that it acted on the best available information, even when the information that mattered most was never treated as authoritative.

Credibility asymmetries also shape how risk is narrated. A manager's assurance that "the system is within tolerance" carries more institutional weight than a technician's observation that "it's been behaving strangely for weeks." When harm occurs, authoritative assessments serve as evidence of due diligence, even when operational insight was more accurate.

Credibility asymmetries manifest acutely during commissioning handover. A senior plant operator with 20 years' experience notes that the chemical dosing equipment configuration forces technicians into awkward positions near live electrical panels during routine maintenance, inevitably leading to unsafe workarounds or deferrals. This assessment draws on embodied knowledge of how maintenance is performed under time pressure, using available tools. Meanwhile, a consulting engineer produces a formal "maintainability assessment" concluding access meets design standards based on manufacturer specifications and ergonomic guidelines. The written engineering assessment carries institutional weight because it aligns with recognised expertise categories and produces documentable conclusions. The operator's knowledge is acknowledged but not treated as authoritative enough to delay acceptance or require modification. When an injury later occurs—a technician losing their balance while reaching into a confined space—the organisation demonstrates due diligence through a professional engineering assessment. The operator's warning existed but lacked epistemic authority to shape decisions.

The mechanism described here operates in tension with a body of organisational scholarship on voice, psychological safety, and the empowerment of situated knowledge. Edmondson's (Edmondson, 1999, 2019) work on psychological safety in healthcare, aviation, and engineering teams has shown that organisations can be deliberately structured to encourage operational staff to raise concerns without fear of professional penalty, and that such structuring measurably reduces error and improves system reliability. Comparable work on nursing voice in surgical settings has documented protocols — formalised speak-up procedures, veto rights to halt procedures, structured handover briefings — that confer on operational staff the institutional authority to stop or redirect decisions that more senior or formally credentialled actors would otherwise carry forward (Leape et al., 2012; Okuyama et al., 2014). Where such arrangements exist and operate as intended, the credibility asymmetry described here is blunted or reversed: situated knowledge acquires institutional standing precisely because the organisation has constructed mechanisms to grant it that standing.

This literature strengthens rather than undermines the present argument. It demonstrates that credibility asymmetries are not inevitable features of complex sociotechnical systems but specific properties of how authority and voice are organisationally arranged. The mechanism described here operates in settings where such arrangements are absent or weak — where, for example, an operator's concerns about maintainability are received as input to be weighed against a consulting engineer's formal assessment, with no procedural mechanism to halt acceptance pending resolution. The comparative point is significant for the conceptual contribution: the conditions that produce plausibly deniable harm are organisationally produced and could, in principle, be organisationally altered.

Table 2. Mechanisms of Epistemic Inequality and Their Contribution to Plausibly Deniable Harm

Mechanism	What it does	How it produces epistemic inequality	How it contributes to plausibly deniable harm
Representational abstraction	Translates situated knowledge into standardised forms	Filters out ambiguity and nuance and contextual detail	Decisions appear informed while critical operational realities are obscured
Procedural compliance	Substitutes process completion for substantive engagement	Limits influence of operational knowledge to predefined scopes	Organisations can claim diligence even when concerns are not acted upon
Temporal deferral	Postpones engagement with operational concerns	Reclassifies urgent issues as premature or out of scope	Harm accumulates while appearing prudently managed
Credibility asymmetries	Privileges abstracted expertise over experiential knowledge	Discounts operational judgement through hierarchical norms	Exclusion appears as professional judgement rather than dismissal
Interactional ambiguity	Maintains courteous interactional norms	Softens or reframes concerns reducing their perceived urgency	Harm is enacted through non-confrontational deflection rather than explicit refusal

Table 2 shows that each mechanism is individually defensible but, collectively, produces organisational conditions in which operational knowledge is systematically excluded while decisions remain procedurally accountable.

4.5 Interactional ambiguity and formal politeness

In this paper, "interactional ambiguity" refers to the organisational effects of professional politeness, where concerns are acknowledged in ways that reduce their decisional force rather than their legitimacy.

Operational staff often learn to frame concerns carefully—as questions rather than assertions, as suggestions rather than warnings, as "just wanting to flag" rather than insisting on action. The result is that legitimate concerns reach decision-makers in forms that are easy to acknowledge but difficult to act on.

Conversely, when operational concerns are raised more directly, recipients may reframe them through polite deflection. A statement that "this configuration will create safety risks during maintenance" becomes heard as "some additional training might be helpful." A warning about system fragility becomes "we appreciate your caution and will keep monitoring." The interactional norm—maintain courtesy, avoid conflict, preserve professional relationships—takes precedence over engaging with the substance of the concern.

This mechanism is particularly insidious because it operates through practices that appear respectful and professional. No one is being dismissed or silenced in ways that would be administratively visible. Instead, concerns are acknowledged, thanked, and gently set aside. The harm that later emerges cannot be traced to anyone being rude or unprofessional. The courtesy itself becomes a way of managing dissent without appearing to do so, making exclusion administratively invisible while maintaining an interpersonal sense of respect.

4.6 The fallibility of situated knowledge

The argument developed here should not be read as a claim that situated knowledge is always correct, or that operational judgement should override formal expertise. Operators, technicians, and commissioning staff can be wrong, and the situated character of their knowledge does not exempt it from error. Patterns identified through repeated exposure may reflect coincidence rather than causation; embodied familiarity with system behaviour may carry forward assumptions that no longer hold as systems are reconfigured; the same proximity that grants distinctive observational access can also produce overgeneralisation from local cases. Vaughan's (1999) account of normalised deviance documents precisely this: operational interpretation itself can drift, with consequences that are not corrected by attention to situated knowledge alone.

What the present analysis claims is something narrower and, for that reason, more tractable. It claims that organisational mechanisms systematically filter situated knowledge out of decision-making processes, and that this filtering produces patterned outcomes regardless of whether any particular instance of situated knowledge would have been correct on the merits. The harm is not that operators are right and are not listened to; the harm is that the organisational architecture cannot engage with what operators know at all, whether to validate, contest, integrate, or correct it. A governance arrangement that engages operational knowledge can find that knowledge wrong and proceed accordingly. A governance arrangement that filters operational knowledge before engagement cannot make that finding, because the substance of the knowledge never reaches the surface at which findings are made. The mechanisms described above operate at this prior level: not at the level of who is right, but at the level of what gets to be considered.

4.7 Cascading effects: how the mechanisms reinforce one another

While analytically distinct, these five mechanisms rarely operate in isolation. In practice, they interact and reinforce one another, creating organisational conditions in which epistemic

exclusion becomes increasingly difficult to contest. Each mechanism creates conditions that make the others more likely to occur, forming what can be understood as cascading epistemic closure—a self-reinforcing system through which operational knowledge becomes progressively marginalised while formal governance appears increasingly robust.

Representational abstraction enables procedural compliance by producing documents that appear comprehensive while filtering operational nuance. Once concerns are translated into standardised risk registers, they can be addressed through established review processes ;that demonstrate diligence, however, the abstraction strips away contextual detail that would make them actionable. Temporal deferral, then, depends on these abstracted representations to appear reasonable—a concern categorised as "monitoring required" seems prudent in documentary form, even as the urgency communicated through direct operational observation is diluted when translated into risk ratings.

Credibility asymmetries determine which concerns generate temporal urgency and which can be deferred. When engineers with formal authority identify risks, escalation is immediate; when operators raise similar concerns, they may be categorised as requiring further evidence. These asymmetries make exclusion appear as appropriate prioritisation rather than dismissal. Interactional ambiguity operates across all mechanisms, providing social lubrication that allows exclusion without confrontation. Professional politeness enables representational abstraction by framing translation as collaborative refinement, facilitates procedural compliance by maintaining collegial relationships, and makes temporal deferral appear respectful—"we appreciate your input and will consider it carefully"—rather than dismissive. Through courtesy, deflection becomes harder to name and contest.

4.7.1 An extended illustration: mechanisms operating in concert

The interaction of these mechanisms can be illustrated through a composite vignette assembled from patterns observed across several wastewater commissioning projects. The scenario described below is not an account of any single project but renders visible an organisational dynamic that has recurred across different facilities, jurisdictions, and contractual arrangements. Figure 2 summarises the mechanisms' trace through the case.

Setting. The vignette is located in the United Kingdom water industry, where regulated water and wastewater utilities operate within five-year capital investment and price-control cycles set by the economic regulator, known as Asset Management Periods (AMPs); the current cycle, AMP8, runs from 2025 to 2030. Continuous monitoring of discharges is governed by the Environment Agency's Monitoring Certification Scheme (MCERTS), which certifies the performance of online instruments used for compliance monitoring, specifies maximum tolerances for parameters such as drift, and requires that users select, install, and operate certified instruments appropriately for their application (Environment Agency, 2010). Discharge consents — the legally enforceable limits on effluent quality — are monitored through these instruments, with breaches subject to regulatory investigation and financial penalty. The project at the centre of this vignette is the upgrade and recommissioning of a wastewater treatment facility serving a regional population, undertaken under a design-build contract between a principal water utility (the asset owner and eventual operator) and a main engineering contractor. A specialist subcontractor supplies the online process instrumentation — duty and standby MCERTS-certified ammonia analysers on the final effluent, with comparable arrangements for turbidity and pH. Commissioning is conducted by a team drawn from the main contractor, with operational staff from the utility participating in acceptance testing and preparing to assume

operational responsibility at handover. Calibration is verified against certified reference samples at commissioning; the maintenance schedule that travels with the handover documentation prescribes a six-monthly verification cycle, in line with the manufacturer's recommendation and consistent with maintaining performance within the MCERTS drift tolerance of five per cent of reading. Performance is governed by contractual acceptance criteria, a twelve-month defects liability period, and discharge consent monitored by the Environment Agency post-handover.

The observation. During the operational testing phase, the duty/standby ammonia analyser pair pass their acceptance calibration against certified reference samples and are signed off as compliant. The instruments are MCERTS-certified continuous water monitors, with a maximum permitted drift of five per cent of reading under the certification standard. Over the subsequent weeks of integrated operational testing, however, an experienced operator on the utility's commissioning support team notices that the two analysers in each pair are not tracking each other as closely as they should. The readings agree within combined measurement uncertainty — neither instrument is outside its certified drift tolerance, neither is failing acceptance — but the operator recognises from extensive plant experience that consistent duty/standby agreement is the leading indicator of stable calibration, and that early divergence within tolerance is how analyser drift announces itself. The operator's plant-specific interpretation is that the influent matrix at this site, combined with the sample line configuration and ambient temperature range, produces drift accumulation faster than the manufacturer's reference conditions assume. The six-monthly verification interval transferred from the manufacturer's recommendation is built on an assumption about drift rate per unit time that, in the operator's assessment, will not hold under the operating conditions specific to this plant. The disagreement, in other words, is not about whether the instruments meet certified performance. They do. It is about whether the verification regime — the cadence at which the instruments are cross-checked against grab samples and recalibrated against reference standards — has been correctly matched to this site's actual drift behaviour. The operator's interpretation is not that current readings are wrong, but that the divergence pattern predicts that one or both instruments will exceed certified drift tolerance well before the six-monthly verification interval brings the discrepancy to the maintenance team's attention.

The first translation: representational abstraction. The concern is raised verbally during the formal commissioning review meeting and is recorded in the meeting minutes. The standard acceptance template asks whether the instruments meet the calibration specification at the time of testing. They do. The minute reads: "Operations noted some variation between duty/standby pairs on ammonia and turbidity. Both instruments within calibration tolerance. Acceptance confirmed." The situated interpretation — that the variation pattern indicates a calibration trajectory the six-monthly maintenance cycle will not catch — cannot be expressed within the template's categorial structure and does not appear in the record. The observation has been preserved; the interpretation has not.

The second translation: credibility asymmetries. The commissioning manager seeks confirmation from the instrument vendor, who provides written assurance that the observed variation is within combined uncertainty for the certified instrument class and that the manufacturer's recommended six-monthly verification interval, consistent with maintaining performance within the MCERTS drift tolerance, remains appropriate for this application. This assessment is technically correct against the standard and carries institutional weight: produced by recognised technical experts, formally documented, aligned with both the manufacturer's data sheet and the MCERTS performance specification. The operator's

interpretation — that a monthly grab-sample cross-check would more appropriately match the actual drift rate under this site's conditions — is grounded in experiential knowledge of how comparable instruments have behaved across other facilities with similar influent characteristics. It is, importantly, a tightening of the verification regime rather than a loosening: the situated knowledge would have the plant operate to a stricter monitoring standard than either the manufacturer's recommendation or the MCERTS minimum requires. This direction of refinement is not, however, recognised by the procedural architecture, which is configured to test whether vendor and standard recommendations are being followed, not whether they are the appropriate reference for this specific application. The operator's knowledge, lacking institutional markers of formal credentialled expertise, cannot revise the verification schedule even where its proposed revision would yield more conservative monitoring practice. No procedural mechanism exists to escalate the disagreement; the schedule defaults to manufacturer recommendation as transferred at handover.

The third translation: interactional ambiguity. The decision is communicated through carefully professional language. The operator is thanked for their attention and informed that the observation has been noted and reviewed against manufacturer guidance. A senior member of the commissioning team adds, with evident goodwill, that the maintenance team will of course pick up any drift at the scheduled calibration. The operator's concern has been received, acknowledged, and gently set aside. No one has been dismissed; no one has been rude. The interactional norm of professional courtesy has carried the concern out of the active decision space and into the implicit deferral that constitutes the response.

The fourth translation: procedural compliance. All required steps have been completed. The acceptance calibration has been conducted, the criteria met, the vendor consulted, the meeting minuted, and the maintenance schedule transferred to the operator's computerised maintenance management system at the prescribed six-monthly interval. The procedural record is, on its own terms, exemplary. The substantive question — whether the verification cadence is correct for this plant's drift behaviour — has been moved out of scope without being explicitly answered, because the procedural architecture provides no place at which substantive questions of this kind would be answered.

The fifth translation: temporal deferral. The maintenance schedule, once programmed into the CMMS, governs calibration practice from handover onward. The experienced operator who made the original observation moves on as the commissioning support role winds down. New operators trained against the documented regime calibrate at the six-monthly interval, log the results, and close the work order. Duty/standby divergence within tolerance is not a flagged parameter in the operator interface, and the pattern recognition that constituted the original concern is not transmitted: not because anyone declined to transmit it, but because the documentary regime provides no prompt to do so.

The harm. Fourteen months after handover, between scheduled six-monthly verifications, one of the final effluent ammonia analysers drifts low. The duty/standby pair now disagree by an amount that exceeds combined uncertainty, but the alarm logic compares the duty reading against the discharge consent limit, not the duty reading against the standby. The duty reading shows compliant ammonia for a period of approximately five weeks before the next scheduled verification brings the divergence to the maintenance team's attention and triggers recalibration. The standby reading, retrieved retrospectively from the historian, indicates that actual effluent ammonia was above consent for a substantial portion of that period. A self-reported breach is filed with the regulator. The subsequent investigation

concludes that the instruments met all acceptance criteria at commissioning, that verification was performed in accordance with the documented maintenance schedule and manufacturer guidance, consistent with the MCERTS performance specification, and that the breach arose from instrument drift between scheduled verification intervals, which is a recognised characteristic of online ammonia monitoring at this concentration range. The utility absorbs the regulatory penalty as operator. The instrument vendor's commercial relationship is unaffected; the commissioning contractor's obligations have long since expired. The original observation, recorded eighteen months earlier in commissioning minutes that have since been archived, is not retrieved by the investigation.

The analytic point. The harm in this case was foreseeable to those whose knowledge was filtered out and unforeseeable within the documentary record that constituted the organisation's formal understanding of the system. At no point was a wrong decision made on the terms by which decisions were assessed. Every step — the acceptance calibration, the vendor's assessment, the maintenance schedule, the meeting minutes, the sign-off, the training, the operational practice, the investigation — followed procedure correctly. The harm emerged through the interaction of these procedurally correct steps, each of which translated situated knowledge into a form that could be administratively processed without engaging the substance of what had been observed.

A further dimension of the same architecture deserves attention. The pattern recognition that would have detected the developing drift was held by an experienced operator whose involvement was time-bounded by the commissioning role, and was not transmitted to the operators who would subsequently be responsible for the plant. The situated knowledge required to prevent the harm was present at commissioning, recorded in the minutes, and absent from every artefact that shaped subsequent operational practice.

What gives this case much of its analytical interest is how ordinary it is. There is no exceptional event, no dramatic moment of decision, no point at which any individual could reasonably be said to have erred. The activity is normalised, quiet, and on its surface harmless: calibrations are scheduled and performed, work orders are closed, training is delivered, instruments report values within tolerance, alarms compare readings against consent limits and remain silent. The harm accumulates insidiously, in the gap between what the documentary regime is configured to detect and what an experienced eye would have looked for. The documentary regime is equally indifferent to whether situated knowledge would have tightened or loosened the operating envelope; it processes neither. The harm becomes legible only at the point of breach — by which time it is too late to recover the pattern recognition that would have prevented it. This is what distinguishes plausibly deniable harm from the more familiar cases of accident, error, or rule-breaking around which much organisational scholarship is organised: the harm is not produced by deviation from normal practice but by the routine functioning of normal practice itself.

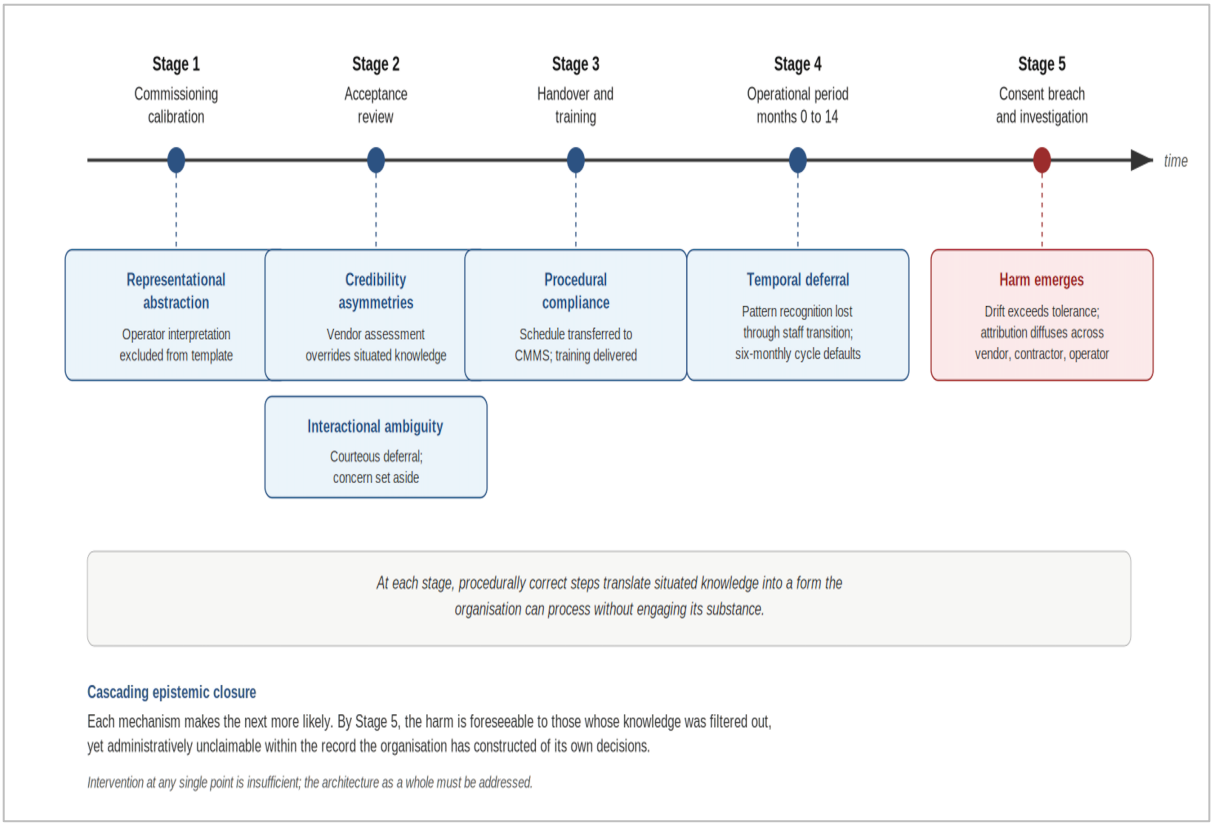

Figure 2. Trace of the wastewater commissioning vignette through five mechanisms over time

5 Discussion: What plausibly deniable harm changes for STS

The concept of plausibly deniable harm offers a way to understand how harm arises from the ordinary functioning of sociotechnical organisations, rather than from dramatic failures. This section reflects on how the analysis extends existing STS conversations about governance, responsibility, expertise, and care. It shows that harm can be produced through systems designed to demonstrate diligence, and that epistemic inequality is a structural feature of how organisations manage knowledge and accountability, providing vocabulary for analysing forms of harm difficult to see, contest, and attribute

5.1 Reframing harm in sociotechnical systems

Much of the STS literature focuses on moments of crisis, breakdown, or exceptional failure (Perrow, 2011). In contrast, plausibly deniable harm shifts attention to the quieter, slower processes through which harm becomes normalised. When organisations rely heavily on documentation and procedural order, they can unintentionally obscure the conditions under which risk is actually increasing (Power, 2013). Decisions appear coherent and defensible because they align with established governance practices, even when those practices have filtered out the knowledge needed to prevent harm (Porter, 1995).

5.2 Rethinking accountability and organisational knowledge

The findings complicate conventional understandings of accountability. Organisations often distribute responsibility across teams, manage uncertainty through documentation, and construct narratives that demonstrate diligence. These practices create a form of procedural accountability that is easy to evidence but difficult to challenge (Power, 2013). When

review processes and documentation practices systematically exclude operational knowledge, accountability becomes a matter of demonstrating that the proper steps were followed rather than engaging with the knowledge needed to make safe, informed decisions (Strathern, 2000).

Procedural accountability operates as a shared narrative about responsible governance, shaping which forms of knowledge are treated as legitimate(Jasanoff & Kim, 2015).

This helps explain why organisations can appear transparent and responsible while remaining unresponsive to the concerns that matter most. It also clarifies why accountability mechanisms often fail to prevent harm: they are designed to verify compliance, not to interrogate the epistemic foundations of decision-making. In many engineering organisations, accountability is retrospective and document-based. It relies on what can be shown, not on what was known (Bowker & Star, 2008). Plausibly deniable harm, therefore, offers a way to analyse how accountability can be maintained procedurally while being hollowed out substantively.

This also has implications for how responsibility is distributed. When harm occurs, organisations often point to completed processes, signed forms, and documented consultations as evidence that they acted responsibly. Yet these artefacts may reflect only a narrow slice of the knowledge available at the time. The concept of plausibly deniable harm helps illuminate how responsibility can be displaced onto processes rather than people, and how this displacement can mask the structural conditions that made harm possible.

5.3 Implications for theories of expertise and care

The analysis also contributes to broader STS debates on expertise and care. Scholars have long argued that system reliability depends on forms of maintenance, coordination, and situated judgement that are undervalued within formal governance structures (Suchman, 1995). Plausibly deniable harm shows that this undervaluation is not simply a matter of recognition. It is a mechanism through which harm is produced. When operational knowledge is abstracted, deferred, or discounted, the labour of care becomes both essential and institutionally unsupported.

This raises questions about what counts as expertise in engineering organisations. Formal authority often carries more weight than experiential familiarity with system behaviour, even when the latter is more accurate (Collins & Evans, 2007). The analysis shows how these credibility asymmetries shape not only whose knowledge is heard but whose knowledge is allowed to matter. It also highlights the emotional and ethical labour carried by operational staff who continue to care for systems even when their concerns are repeatedly softened or deferred. Their expertise is indispensable, yet structurally marginalised.

The concept of plausibly deniable harm also raises fundamental questions about what counts as care work in engineering organisations and whose care work is valued. Care, as theorised by Mol (2008) and extended by Puig de la Bellacasa (2017) involves attentiveness to vulnerability, maintenance of relational connections, and ongoing practices of noticing and responding. In commissioning contexts, operators and technicians continuously engage in precisely these practices—monitoring system behaviour, noticing anomalies, interpreting performance changes, and anticipating failure modes—that constitute essential care work for sociotechnical systems. Yet this care remains largely invisible within governance structures that privilege abstracted assessment over situated attention (Star and Strauss, 1999).

When credibility asymmetries discount operational judgment, they simultaneously devalue the care work on which system reliability actually depends. This creates a distinctive double bind for operational staff. They must continue caring for systems—attending to fragility, compensating for design limitations, preventing failures through vigilant interpretation—even as organisational governance structures fail to recognise these activities as requiring their specific form of situated knowledge. The emotional and ethical labour involved in maintaining attention to system vulnerability, particularly when concerns are repeatedly deflected through the mechanisms described here, represents an unacknowledged form of organisational burden that extends beyond conventional accounts of invisible work (Strauss, 1985).

This burden is compounded by the epistemic violence enacted when care work is systematically discounted. To have one's professional judgment about system vulnerability repeatedly abstracted into compliance categories, deferred to later phases, or overridden by formal assessments that lack operational grounding is not merely a matter of misrecognition—it is an erosion of the epistemic foundations on which careful practice depends. Operators who observe their warnings documented but not acted upon, acknowledged but not engaged, must decide whether to continue raising concerns or withdraw into compliance with the very documentary systems that filter their knowledge. The former requires sustained emotional labour in the face of recurring deflection; the latter requires accepting complicity in harm that operational knowledge suggests could be prevented.

Plausibly deniable harm, therefore, is not only a matter of epistemic inequality but also a question of whose care work is valued and supported within sociotechnical systems. Organisations that rely on operational care while systematically excluding operational knowledge from governance create conditions in which care becomes both essential and unsustainable. This dynamic helps explain patterns of quiet professional attrition, burnout, and disengagement among operational staff who continue to carry responsibility for system safety without the epistemic authority to shape decisions that determine safety outcomes. It also illuminates why post-incident investigations often find that "warning signs were present but not acted upon"—the care work of noticing was occurring, but the organisational structures for acting on operational knowledge were absent or non-functional.

Recognising care as an epistemic practice, not merely an affective orientation, shifts how we understand what constitutes epistemically just governance. It suggests that valuing operational knowledge requires more than including operators in consultation processes; it requires treating operational care work as a legitimate basis for shaping organisational decisions. This, in turn, challenges conventional divisions between "operational" and "strategic" knowledge, revealing how such categories themselves function as mechanisms of epistemic inequality that enable plausibly deniable harm.

5.4 Conceptual portability beyond engineering

Although this study is grounded in engineering and infrastructure governance, the concept of plausibly deniable harm has broader relevance. Similar patterns can be seen in healthcare, public administration, and algorithmic decision-making, where harm emerges not from overt failure but from governance arrangements that render certain forms of knowledge administratively irrelevant (Jasanoff, 2004). The concept, therefore, offers a portable lens for examining how organisations across sectors manage uncertainty, distribute responsibility, and justify decisions that have harmful consequences.

This portability also suggests that plausibly deniable harm may help bridge conversations across subfields of STS. It connects work on organisational knowledge, maintenance and care, accountability, and sociotechnical imaginaries by showing how these dynamics converge in the production of harm (Jasanoff & Kim, 2015). It also provides a way to analyse how governance systems create the conditions under which harm becomes both possible and deniable.

In healthcare, similar patterns emerge in electronic health record systems, which translate nursing observations of patient deterioration—subtle behavioural changes that experienced nurses recognise as early warnings—into standardised data fields that may not capture the situated assessment that gives observations their significance (Mol, 2008). Once abstracted, concerns can be reviewed through clinical governance processes, demonstrating documentation compliance without engaging the interpretive nursing knowledge that generated concern. Credibility asymmetries privilege physicians' diagnostic conclusions over nurses' judgment (Elhihi et al., 2025), even when nurses' continuous patient contact provides distinctive observational access. When adverse outcomes occur, healthcare organisations can demonstrate proper protocols were followed, yet the harm may have been foreseeable to nursing staff whose knowledge was filtered through these governance mechanisms.

In policy contexts, community consultation processes demonstrate procedural inclusion while filtering experiential knowledge into policy-legible categories. Indigenous environmental knowledge, local residents' understanding of neighbourhood dynamics, or community health workers' insights must fit predetermined frameworks that may render certain concerns out of scope (Li, 2005; Whyte, 2018). Credibility asymmetries privilege technical expertise or academic research over community knowledge, even when lived engagement provides access to dynamics that formal methods miss.

In algorithmic systems, representational abstraction operates through training data and feature selection, where design choices are embedded, and the systems subsequently appear technically neutral (Benjamin, 2019). Procedural compliance is demonstrated through fairness metrics and algorithmic audits (Raji et al., 2020), while credibility asymmetries privilege data science expertise over affected communities' accounts of discriminatory outcomes.

These parallels suggest that plausibly deniable harm may be pervasive in governance contexts that combine documentary accountability, uneven distributions of expertise, organisational complexity, and high-stakes harm attribution. The concept offers an analytical purchase for understanding how harm emerges from formally legitimate arrangements across diverse sociotechnical systems.

5.5 Normative implications: towards epistemically inclusive governance

Finally, the analysis raises normative questions about how sociotechnical governance might be reconfigured to reduce the conditions under which plausibly deniable harm emerges. More epistemically inclusive organisations would recognise operational judgement as a legitimate form of expertise, create pathways for situated knowledge to shape decisions, and treat documentation as a complement rather than a substitute for engagement (Bowker & Star, 2008). Acknowledging the structural production of deniability is a necessary first step toward designing systems that are substantively, rather than merely procedurally, accountable.

5.6 Future empirical directions

This framework suggests several directions for empirical research. First, detailed ethnographic studies of how documentary practices actually unfold in commissioning environments could test whether the mechanisms operate as theorised. Second, a comparative analysis across infrastructure sectors could reveal contextual variations in how epistemic inequality harms. Third, intervention studies examining what happens when organisations adopt more epistemically inclusive practices could demonstrate practical applications.

6. Conclusion: rethinking harm, knowledge, and responsibility

This paper has shown how harm can emerge through the ordinary functioning of engineering and infrastructure organisations. Plausibly deniable harm describes a pattern in which operational knowledge is softened, delayed, or discounted through routine practices that appear reasonable, compliant, and defensible. The five mechanisms developed here — representational abstraction, procedural compliance, temporal deferral, credibility asymmetries, and interactional ambiguity — do not operate in isolation. They reinforce one another, creating a governance environment in which harmful decisions can be justified even when they rest on incomplete or distorted understandings of system behaviour.

The contribution to STS is to offer a way of analysing how harm is produced not through breakdown, misconduct, or normalised deviance, but through the structural and documentary features of sociotechnical governance itself. The framework differentiates from Turner's (1976) account of incubation, which locates harm in the failed assembly of dispersed warning signs, and from Vaughan's (1996, 1999) account of normalised deviance, which locates harm in the operational rationalisation of anomalies. Plausibly deniable harm locates harm in a third pattern: situated knowledge that is present, accurate, and raised is filtered out of decision-making through formally legitimate organisational processes, and the resulting harm becomes administratively unclaimable because every step in the decision chain remained procedurally compliant. The deniability is structural, not deliberate; it is a property of the epistemic architecture rather than of the intentions of any individual or organisation.

The framework has implications across sectors that combine documentary accountability, distributed responsibility, asymmetric authority between formal expertise and situated knowledge, and high-stakes outcomes whose attribution is complicated by procedural compliance. Engineering and infrastructure are paradigmatic cases but not the only ones; the parallels in healthcare, public administration, and algorithmic decision-making suggest that the pattern is more widespread than any single sectoral analysis can capture.

For organisational practice, the analysis offers a practical insight: organisations can meet all procedural requirements and still arrive at decisions that increase risk. Recognising this opens the possibility of designing more epistemically inclusive arrangements — arrangements that value operational judgement, create space for uncertainty, treat documentation as a complement rather than a substitute for engagement, and provide procedural pathways through which situated knowledge can revise decisions in either direction, toward more or less conservative practice. Preventing plausibly deniable harm requires more than compliance with established processes. It requires attention to the epistemic conditions under which decisions are made, the forms of knowledge that are

allowed to shape those decisions, and the structural mechanisms that make some concerns visible while rendering others administratively irrelevant.

Naming this pattern is a precondition for addressing it. The vocabulary of plausibly deniable harm makes visible an organisational dynamic that, once named, becomes available both for empirical investigation and for governance redesign. Future work should test the mechanisms ethnographically across sectors, compare organisational arrangements that successfully engage operational knowledge with those that filter it, and examine the conditions under which the cascading interaction of these mechanisms can be interrupted. The challenge is not to eliminate documentary governance, procedural compliance, or formal expertise — each of which serves indispensable functions in complex sociotechnical systems — but to ensure that the documentary regime remains in dialogue with the situated knowledge on which its own legitimacy ultimately depends.

A qualification is necessary. The argument that situated knowledge is filtered out of organisational decision-making should not be read as a claim that situated knowledge is always correct. Operators can be wrong; experiential familiarity can carry forward assumptions that no longer hold; the same proximity that grants distinctive observational access can produce overgeneralisation. The claim is narrower and, for that reason, more tractable: the mechanisms described produce patterned outcomes regardless of whether any particular instance of situated knowledge would have proven correct on the merits. A governance arrangement that engages operational knowledge can find that knowledge wrong and proceed accordingly. A governance arrangement that filters operational knowledge before engagement cannot make that finding, because the substance of the knowledge never reaches the surface at which findings are made. As the vignette in Section 4.7.1 illustrates, this filtering operates symmetrically: situated knowledge proposing more conservative practice is excluded as efficiently as situated knowledge proposing laxer practice.

AI use statement

The author used generative artificial intelligence (Grammarly) to assist with language editing and structural clarity during manuscript preparation. All conceptual framing, theoretical development, interpretation, and substantive content are the author's own. The author takes full responsibility for the integrity, accuracy, and originality of the work.

References

Alcadipani, R., Khan, F. R., Gantman, E., & Nkomo, S. (2012). Southern voices in management and organization knowledge. *Organization*, *19*(2), 131–143. https://doi.org/10.1177/1350508411431910

Asdal, K., & Hobæk, B. (2020). The modified issue: Turning around parliaments, politics as usual and how to extend issue-politics with a little help from Max Weber. *Social Studies of Science*, *50*, 030631272090284. https://doi.org/10.1177/0306312720902847

Barua, M. (2024). *Plantation Worlds*. Duke University Press. https://doi.org/10.1215/9781478027744

Bechky, B. A. (2003). Sharing Meaning Across Occupational Communities: The Transformation of Understanding on a Production Floor. *Organization Science*, *14*(3), 312–330. https://doi.org/10.1287/orsc.14.3.312.15162

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.

Bovens, M. (2007). Analysing and Assessing Accountability: A Conceptual Framework[1]. *European Law Journal*, *13*(4), 447–468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

Bowker, G. C., & Star, S. L. (2008). *Sorting things out: Classification and its consequences* (1. paperback ed., 8. print). MIT Press.

Cech, E. A. (2014). Culture of Disengagement in Engineering Education? *Science, Technology, & Human Values*, *39*(1), 42–72. https://doi.org/10.1177/0162243913504305

Collins, H., & Evans, R. (2002). The Third Wave of Science Studies: Studies of Expertise and Experience. *Social Studies of Science*, *32*(2), 235–296. https://doi.org/10.1177/0306312702032002003

Collins, H., & Evans, R. (2007). *Rethinking Expertise*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226113623.001.0001

Denis, J., & Pontille, D. (2014). Material Ordering and the Care of Things. *Science, Technology & Human Values*, *40*, 338–367. https://doi.org/10.1177/0162243914553129

Edmondson, A. (1999). Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly*, *44*(2), 350–383. https://doi.org/10.2307/2666999

Edmondson, A. (2019). *The fearless organization: Creating psychological safety in the workplace for learning, innovation, and growth*. John Wiley & Sons, Inc.

Elhihi, E. A., Aljarary, K. L., Alahmadi, M., Adam, J. B., Almwualllad, O. A., Hawsawei, M. S., Hamza, A. A., & Abdullatif Ibrahim, I. A. (2025). The mediating role of moral courage in the relationship between ethical leadership and error reporting behavior among nurses in Saudi Arabia: A structural equation modeling approach. *BMC Nursing*, *24*(1). Scopus. https://doi.org/10.1186/s12912-025-03043-7

Environment Agency. (2010). *MCERTS performance standards for continuous water monitoring equipment, Part 2: Performance standards and test procedures for on-line monitors (Version 3.1). Environment Agency, Bristol, UK.* https://assets.publishing.service.gov.uk/media/5a74fd1ded915d502d6cc9f7/performance_standards_and_test_procedures_for_continuous_water_monitoring_equipment_part_2.pdf

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford university press.

Graham, S., & Thrift, N. (2007). Out of Order: Understanding Repair and Maintenance. *Theory, Culture and Society*, *24*. https://doi.org/10.1177/0263276407075954

Henderson, K. (1991). Flexible Sketches and Inflexible Data Bases: Visual Communication, Conscription Devices, and Boundary Objects in Design Engineering. *Science,*

Technology, & Human Values, *16*(4), 448–473. https://doi.org/10.1177/016224399101600402

Hull, M. S. (2012). *Government of paper: The materiality of bureaucracy in urban Pakistan*. University of California Press.

Irwin, A., & Wynne, B. (Eds.). (1996). *Misunderstanding Science?: The Public Reconstruction of Science and Technology* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511563737

Jasanoff, S. (Ed.). (2004). *States of Knowledge* (0 ed.). Routledge. https://doi.org/10.4324/9780203413845

Jasanoff, S., & Kim, S.-H. (Eds.). (2015). *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. The University of Chicago Press.

Leape, L. L., Shore, M. F., Dienstag, J. L., Mayer, R. J., Edgman-Levitan, S., Meyer, G. S., & Healy, G. B. (2012). Perspective: A Culture of Respect, Part 1. *Academic Medicine*, *87*(7), 845–852. https://doi.org/10.1097/ACM.0b013e318258338d

Li, T. M. (2005). Beyond "the State" and Failed Schemes. *American Anthropologist*, *107*(3), 383–394. https://doi.org/10.1525/aa.2005.107.3.383

Medina, J. (2013). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination* (1st ed.). Oxford University PressNew York. https://doi.org/10.1093/acprof:oso/9780199929023.001.0001

Mol, A. (2008). *The Logic of Care* (0 ed.). Routledge. https://doi.org/10.4324/9780203927076

Okuyama, A., Wagner, C., & Bijnen, B. (2014). Speaking up for patient safety by hospital-based health care professionals: A literature review. *BMC Health Services Research*, *14*(1), 61. https://doi.org/10.1186/1472-6963-14-61

Orr, J. E. (2016). *Talking about Machines: An Ethnography of a Modern Job*. Cornell University Press.

Otsuki, K. (2024). Infrastructural violence and its temporalities. In O. Coutard & D. Florentin (Eds.), *Handbook of Infrastructures and Cities* (pp. 240–254). Edward Elgar Publishing. https://doi.org/10.4337/9781800889156.00026

Perrow, C. (2011). *Normal Accidents: Living with High Risk Technologies - Updated Edition*. Princeton University Press. https://doi.org/10.1515/9781400828494

Pollock, N., & Williams, R. (2008). *Software and Organisations* (0 ed.). Routledge. https://doi.org/10.4324/9780203891940

Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton university press.

Power, M. (2004). *The risk management of everything: Rethinking the politics of uncertainty* (1. publ). Demos.

Power, M. (2013). *The audit society: Rituals of verification* (Repr). Oxford University Press.

Puig de la Bellacasa, M. (2017). *Matters of Care: Speculative Ethics in More Than Human Worlds. University of Minnesota Press. 265 pages.*

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-

to-end framework for internal algorithmic auditing. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33–44. https://doi.org/10.1145/3351095.3372873

Sharma, G., & Bansal, P. (Tima). (2020). Cocreating Rigorous and Relevant Knowledge. *Academy of Management Journal*, *63*(2), 386–410. https://doi.org/10.5465/amj.2016.0487

Star, S. L., & Strauss, A. (1999). Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work (CSCW)*, *8*(1–2), 9–30. https://doi.org/10.1023/A:1008651105359

Strathern, M. (2000). The Tyranny of Transparency. *British Educational Research Journal*, *26*(3), 309–321. https://doi.org/10.1080/713651562

Strauss, A. (1985). *Social organization of medical work*. University of Chicago press.

Suchman, L. (1995). Making work visible. *Communications of the ACM*, *38*(9), 56–64. https://doi.org/10.1145/223248.223263

Sykianakis, N., & Bellas, A. (2011). Organization Politics and the Role of Accounting. *EUROPEAN RESEARCH STUDIES JOURNAL*, *XIV*(Issue 3), 119–134. https://doi.org/10.35808/ersj/330

Turner, B. A. (1976). The Organizational and Interorganizational Development of Disasters. *Administrative Science Quarterly*, *21*(3), 378. https://doi.org/10.2307/2391850

Vaughan, D. (1996). *The Challenger launch decision: Risky technology, culture, and deviance at NASA* (1st ed.). Univ. of Chicago Press.

Vaughan, D. (1999). THE DARK SIDE OF ORGANIZATIONS: Mistake, Misconduct, and Disaster. *Annual Review of Sociology*, *25*(1), 271–305. https://doi.org/10.1146/annurev.soc.25.1.271

Whyte, K. P. (2018). Indigenous science (fiction) for the Anthropocene: Ancestral dystopias and fantasies of climate change crises. *Environment and Planning E: Nature and Space*, *1*(1–2), 224–242. https://doi.org/10.1177/2514848618777621

Wynne, B. (1992). Misunderstood misunderstanding: Social identities and public uptake of science. *Public Understanding of Science*, *1*(3), 281–304. https://doi.org/10.1088/0963-6625/1/3/004