Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria
 Published by Research Publishing, Singapore.
doi: 10.3850/ESREL2026061419_esrel26-p26235-cd

The Organizational Black Box: Understanding Systemic Vulnerabilities at the Sharp End / Blunt End Interface

Elliot Quiriconi

Mines Paris, PSL University, Centre for research on risks and crises (CRC), 06904 Sophia Antipolis, France.
E-mail: elliot.quiriconi @etu.minesparis.psl.eu

Didier Delaitre

Mines Paris, PSL University, Centre for research on risks and crises (CRC), 06904 Sophia Antipolis, France.
E-mail: didier.delaitre@minesparis.psl.eu

Eric Rigaud

Mines Paris, PSL University, Centre for research on risks and crises (CRC), 06904 Sophia Antipolis, France.
E-mail: eric.rigaud @minesparis.psl.eu

Safety models for complex sociotechnical systems have progressively shifted from individual error to organizational and systemic perspectives. Reason's defense-in-depth model identifies latent conditions but does not explain how they remain invisible to the actors who produce them. Cook and Woods' sharp end/blunt end distinction reveals informational asymmetries without formalizing the mechanisms that sustain them. This paper proposes the Organizational Black Box model, grounded in Ashby's cybernetic concept of the black box and extended from the technical to the organizational scale, to address this gap. The model articulates three interdependent dimensions: bidirectional opacity between sharp end and blunt end, requisite variety reduction through organizational feedback channels, and delegated trust degradation producing collective automation biases. The framework is applied retrospectively to the BEA final investigation report on flight AF447 Rio-Paris (June 1, 2009). The case study illustrates how these three dimensions interact to produce an emergent systemic vulnerability: opacity sustains silent propagation of erroneous assumptions across organizational boundaries, variety reduction filters corrective signals before they reach regulatory actors, and delegated trust stabilizes the system in its vulnerable configuration. While limited to a single retrospective case and remaining qualitative, the model contributes a complementary analytical lens focused on the generative role of organizational opacity—a dimension acknowledged but insufficiently formalized in existing safety frameworks. Directions for prospective application, formal modeling, and cross-domain validation are discussed.

Keywords: organizational opacity, black box, requisite variety, trust degradation, systemic vulnerability, AF447.

1. Introduction

Cook and Woods (1994) introduced the sharp end/blunt end distinction to characterize the structural division of labor in complex sociotechnical systems. The sharp end refers to frontline practitioners—operators, pilots, clinicians, or technicians—who interact directly and in real time with the hazardous process. They bear the immediate consequences of any failure and must continuously adapt to the unpredictability of the operational environment. The blunt end encompasses the organizational actors who are spatially and temporally removed from operations: designers, managers, regulators, and certification authorities. These actors shape the conditions within which sharp end practitioners operate by allocating resources, defining procedures, certifying equipment, and

structuring training programs, but without being exposed to the operational consequences of their decisions.

This separation generates a critical and bidirectional informational asymmetry. On one side, the sharp end continuously adapts prescribed work to operational realities—improvising, compensating for system deficiencies, and developing local expertise—without these adjustments being visible to the blunt end. On the other, the blunt end designs systems and procedures based on idealized models of how operators should behave, without direct access to the cognitive and adaptive realities of actual practice. The result is a structural gap between work-as-imagined and work-as-done that neither end can fully perceive from its own vantage point. While Reason's (1997) defense-in-depth model acknowledges the role of latent organizational conditions in accident causation, and Cook and Woods' framework highlights the resulting informational asymmetry, neither fully accounts for the mechanisms through which this mutual opacity is actively sustained and why it systematically escapes detection.

It is precisely this gap that the present paper seeks to address. We propose the Organizational Black Box model, grounded in Ashby's (1956) cybernetic concept of the black box and extended from the technical to the organizational scale. The model articulates three interdependent dimensions—bidirectional opacity, requisite variety reduction, and delegated trust degradation—whose circular interaction progressively degrades the system's capacity for self-observation. Its application to the BEA investigation report on flight AF447 Rio-Paris (BEA, 2012) illustrates how organizational structure itself becomes a generative architecture of systemic vulnerability.

2. Background

2.1. The AF447 accident

On June 1, 2009, at 2:14 AM, the Airbus A330 operating flight AF 447 crashed into the Atlantic Ocean, killing all 228 people on board. The accident began when the Pitot probes became obstructed by ice crystals, resulting in erroneous speed indications and disconnection of the automation systems. The two co-pilots provided inappropriate control inputs that destabilized the aircraft's trajectory, without following the appropriate procedure for loss of speed indications. The crew did not identify the approach-to-stall nor diagnose the stall situation, which persisted for nearly 4 minutes. The captain, who had left the cockpit, returned 1 minute 30 seconds after the event began, when the aircraft was already in an unrecoverable stall situation. Among the contributing factors, the BEA identified insufficient training for flying in degraded situations, problematic cockpit ergonomics (notably the flight directors issuing misleading commands), and the combined emotional and cognitive workload that prevented the crew from correctly diagnosing the situation.

2.2. From individual error to systemic analysis

By distinguishing active errors from latent errors, Reason (1990, 1997) established an initial conceptual framework for analyzing accidents in complex sociotechnical systems. Active errors correspond to dangerous acts committed by operators in direct contact with the system, whose consequences are immediately perceptible. Latent errors reside in strategic and organizational decisions made upstream, system design, resource allocation and create conditions conducive to failure that can remain dormant before combining with local triggering factors. In his defense-in-depth model, often represented by the "Swiss cheese" metaphor, Reason (1997) demonstrates that organizational accidents result from the alignment of multiple defensive breaches created by successive latent errors. This approach shifts the analytical focus from the individual error toward the organizational conditions that made it possible (Reason, 2000).

Cook and Woods (1994) extend this systemic approach by introducing the sharp end/blunt end distinction. In aviation, the sharp end comprises flight crews while the blunt end encompasses manufacturers, certification authorities, and training organizations, as illustrated by the AF447 case, where insufficient interactions between those who designed the cockpit automation and those who had to operate it in degraded conditions constituted a critical vulnerability. However, while these frameworks illuminate the systemic nature of organizational accidents, they do not provide adequate tools for analyzing the mechanisms through which opacity

itself becomes a generator of systemic vulnerabilities.

It is this gap that the following theoretical framework aims to address.

3. Theoretical Framework: The Organizational Black Box Model

Reason's defense-in-depth model identifies latent conditions but does not explain how they remain invisible to the very actors who produce or suffer from them. Cook and Woods reveal the structural tension between sharp end and blunt end but do not formalize the informational mechanisms that sustain this mutual blindness. To address this gap, we propose a conceptual framework grounded in the cybernetic concept of the "black box" (Ashby, 1956), extended from the technical to the organizational scale. This framework articulates three interdependent dimensions: (1) bidirectional opacity, where each end becomes a black box for the other, (2) requisite variety reduction, where feedback mechanisms impoverish the diversity of operational signals necessary for regulation, and (3) trust degradation, where the delegated confidence system that compensates for opacity simultaneously generates collective automation biases and erodes critical vigilance.

3.1. Foundational concepts

3.1.1. Black box, opacity, and automation

The black box, theorized by Ashby (1956) within cybernetics, is defined as a system that can only be approached through its inputs and outputs, with no knowledge of its internal functioning. Treating a system as a black box is a cognitive choice implying a form of renunciation, voluntary or otherwise, of understanding the interior in order to focus on use.

This concept must be distinguished from opacity, which is a property of the system itself. Opacity describes the degree of non-transparency along a continuum ranging from the "perfectly white box" (fully comprehensible) to the "perfectly black box" (impenetrable), with progressive shades of gray in between. It is because a system exhibits a certain degree of opacity that one may treat it conceptually as a black box. Opacity is partly subjective—dependent on the observer's knowledge and experience—but objective opacity can emerge when a system's complexity exceeds any single individual's cognitive capacities.

The relationship between automation and opacity depends on the complexity involved. Complexity, in the sense of Morin (1990), non-linear interactions, emergence, feedback loops, combined with increasing specialization and finite human cognition, inevitably generates growing opacity. When automated systems simulate complex model interactions in non-linear fashion, even experts cannot always reproduce the computations manually. In modern aviation, the interaction between fly-by-wire control laws, autopilot modes, autothrust logic, and flight envelope protections creates computational complexity that makes the cockpit automation a black box even for experienced crews.

3.1.2. Cognitive consequences: automation biases

Automation exposes users to two interdependent cognitive biases (Goddard et al., 2012): automation bias, an overconfidence in automated outputs leading to commission errors, and complacency bias, an insufficient monitoring of the system generating omission errors. Together, they progressively erode professional intuition and situational awareness.

3.2. Dimension 1: Bidirectional opacity

The sharp end/blunt end separation produces a bidirectional opacity: each end becomes a black box for the other.

From the sharp end's perspective, systems designed by the blunt end, cockpit automation, flight control laws, operational procedures, constitute black boxes whose internal logic is only partially accessible. Pilots interact with automated modes through limited interfaces that provide outputs without exposing the underlying reasoning. When the system transitions between modes or degrades its protections, crews must infer what happened from indirect cues, often under time pressure and cognitive load.

From the blunt end's perspective, operational reality is equally opaque. Designers, certification engineers, and training managers conceive systems based on models of how operators should behave, the prescribed work. But

they have limited access to actual work: the real-time adaptations, improvisations, and cognitive strategies that crews deploy facing unexpected situations. The blunt end operates on a simplified representation of operational reality, filtered through incident reports and flight data monitoring that capture only a fraction of actual practice.

This bidirectional opacity is structurally generative of vulnerability. Each end makes decisions based on an impoverished model of the other, creating blind spots precisely at the interface where critical safety information should flow.

3.3. Dimension 2: Requisite variety reduction

Ashby's (1956) Law of Requisite Variety states that a regulatory system must possess at least as much variety as the system it seeks to regulate. Applied to the sharp end/blunt end interface, this principle reveals a fundamental structural deficit. The sharp end operates in an environment of high variety: each flight presents a unique combination of weather, traffic, technical state, crew composition, and operational pressures. However, the channels through which this information reaches the blunt end operate as variety-reducing filters. Incident reporting systems require categorization into predefined taxonomies. Flight data monitoring captures parametric deviations but not the cognitive processes behind them. Safety reports narrativize events into causal chains that necessarily simplify original complexity.

The blunt end therefore regulates based on a low-variety model of a high-variety reality, violating Ashby's Law and creating structural blind spots. Situations that do not fit existing categories, novel failure modes, unexpected mode interactions, may be filtered out before reaching the actors who design systems and procedures. Crucially, this variety reduction is not a dysfunction but a designed feature of organizational information systems, which must compress to function at scale. The vulnerability is therefore inherent to the architecture of the sharp end/blunt end interface.

3.4. Dimension 3: Trust degradation and collective automation biases

3.4.1. The non-linear relationship between opacity and trust

The framework of automation levels (Parasuraman et al., 2000) shows that operators must base decisions on results from processes whose internal logic escapes them. This should theoretically provoke distrust. However, Hoff and Bashir (2015) demonstrate that trust results from three interacting factors: dispositional trust (linked to professional culture, pilots trained in rigorous frameworks are predisposed to trust validated systems), learned trust (linked to perceived reliability, opacity can paradoxically generate higher trust toward systems deemed reliable), and situational trust (linked to context, the organizational ecosystem that certifies, trains, and validates creates psychological safety that may override individual vigilance).

The operator thus does not place trust in the black box itself, but in the organizational ecosystem that validates and supports it. In aviation, the exceptional safety record of automated flight reinforces learned trust to the point where manual reversion in degraded situations becomes psychologically difficult.

3.4.2. Organizational trust and emergent vulnerabilities

When shifting scale from the technical to the organizational level, the same trust mechanism replicates. The organization structures itself as an assemblage of black boxes: the manufacturer's design office is a black box for the certification authority; the certification authority is a black box for the airline's training department; the training department is a black box for line pilots. Each actor takes another domain's conclusions as reliable input without mastering the underlying methods or assumptions.

The biases identified at the technical scale transpose through structural homology. Trust in an automated output becomes trust in another department's production. Task focus becomes strict perimeter delimitation, entailing critical negligence of interfaces. Institutional validation, while necessary, inhibits questioning. Each actor becomes an expert in their task but progressively loses the capacity to understand how their contribution integrates into the overall safety objective. Risk becomes more abstract, perceived as weaker than its actual value. The automation paradox reappears at the organizational level: supposed system reliability decreases active supervision of interfaces.

This environment generates an emergent effect: the silent circulation of potentially erroneous assumptions across organizational boundaries. An anomaly introduced in one black box, a design assumption, a training gap, can propagate to the next without critical examination, creating a gap between perceived component robustness and actual interface vulnerability.

4. Case Study: The AF447 Accident through the Organizational Black Box Model

The BEA final investigation report on the AF447 accident (BEA, 2012) provides a rich empirical basis for applying the Organizational Black Box model. On June 1, 2009, the temporary obstruction of Pitot probes by ice crystals led to automation disconnection and the crew's inability to diagnose and recover from an aerodynamic stall that persisted for nearly four minutes, killing all 228 occupants. Beyond the immediate event sequence, the BEA report reveals structural mechanisms that the three dimensions of our model articulate systematically. Critically, this case study does not seek to re-analyze the accident itself, but to demonstrate how the proposed framework renders visible the organizational opacity mechanisms that existing models, Reason's defense-in-depth and Cook and Woods' sharp end/blunt end distinction, identify without formalizing.

4.1. Dimension 1: Bidirectional opacity

4.1.1. Sharp end opacity: cockpit automation as black box

The transition to Alternate Law 2B removed envelope protections without providing the crew explicit indication of which alternate law was active or what protections remained. The ECAM displayed the generic message "PROT LOST," yet load factor protection remained available, while angle-of-attack protections—the most critical in the developing situation—were not. The BEA notes that "the identification of the precise consequences of a reconfiguration in alternate law is therefore complicated." The crew could not determine whether the stall alarm—impossible in Normal Law—was now a credible signal requiring immediate response. This is a textbook instance of the black box concept applied to the sharp end: the crew interacted with a system whose observable outputs (ECAM messages, alarm activations, flight director bars) provided no reliable access to the underlying state logic.

The flight directors compounded this opacity. Unlike the autopilot and autothrust, which disconnected automatically, the flight directors remained engaged while their tendency bars disappeared and reappeared as input parameters fluctuated, with mode changes detectable only through FMA reading, "probably difficult in a high workload situation." The PF probably followed these bars "without having integrated the change in the engaged longitudinal mode" (BEA, 2012), making them "one of the rare points of coherence in the general incomprehension of the situation." This exemplifies the automation bias predicted by our model: when opacity prevents the operator from evaluating the system's internal reasoning, the automated output becomes a trust anchor followed even when contradicted by other cues, here, the stall alarm and the buffet.

The ECAM itself, designed as a transparency tool, became an opacity generator. Although the flight control computers had detected airspeed inconsistencies, no message communicated this diagnosis to the crew. The ECAM displayed only consequences (autopilot disconnection, autothrust loss, alternate law transition) without exposing their common cause, consuming cognitive resources "at the expense of problem treatment and trajectory monitoring" (BEA, 2012). Meanwhile, the stall alarm's undocumented dependence on angle-of-attack validity created a devastating paradox: when measured speeds fell below 60 knots, angle-of-attack values were invalidated and the alarm stopped. Nose-down inputs reactivated the alarm while nose-up inputs silenced it—the "correct" recovery action appeared to trigger the danger signal, while the "incorrect" action appeared to resolve it.

4.1.2. Blunt end opacity: operational reality as black box

Stall training was conducted exclusively at low altitude (FL100) in a "demonstrative and analytical" manner during initial type qualification, allowing trainees to hear the alarm

"in a situation where it is expected and corrective actions prepared" (BEA, 2012). This approach systematically filtered out the conditions of the actual accident: high-altitude aerodynamics with reduced margins between cruise speed and stall speed, the surprise effect, emotional loading, night flight in turbulence, and degraded flight control law. The BEA observes that "the demonstrative character of the exercises does not allow the crew to apprehend the surprise effect generated by the stall alarm." The operational reality of what happens when pilots face an unexpected stall alarm at FL350 was entirely invisible to those who designed the training curriculum, a direct instantiation of the blunt end's black box problem.

The "Unreliable Airspeed" procedure existed and was technically adequate. However, its design reflected blunt-end assumptions about crew behavior that proved disconnected from operational reality. The procedure was predicated on initial trajectory control and rapid diagnostic capability, a "common mode failure" in the safety model, since "when the capacity for initial trajectory mastery is also lost, the safety model finds itself in common mode default" (BEA, 2012). The study of previous similar events showed this procedure was never applied in cruise, a recurring pattern invisible to its designers, who assumed its existence constituted adequate risk mitigation. Simulator training, following pre-established and known scenarios, could reproduce neither the variability of fault signals nor the surprise effect, leaving crews trained for a simplified version of a complex reality. As the BEA concludes, "the difficulty, even impossibility, of reproducing in the simulator both the complexity and variability of fault signals, combined with the lack of surprise effect linked to a known scenario, did not allow training to be adapted to the situation actually encountered."

4.2. Dimension 2: Requisite variety reduction

The AF447 case reveals multiple mechanisms through which operational variety was systematically reduced before reaching actors capable of regulatory action, in direct violation of Ashby's Law of Requisite Variety.

Prior to the accident, numerous Pitot icing events had occurred across the A330/A340 fleet. However, crews experiencing brief airspeed losses in cruise did not always file reports capturing the full cognitive and procedural dimension of the event—particularly their non-application of the "Unreliable Airspeed" procedure. The BEA reveals that "a certain number of incidents are not reported in a directly exploitable manner by crews; only a posteriori analysis of recorded data has enabled detection of their safety aspect." The absence of hot testimony or CVR recordings, resulting from "imprecise notification," meant that the sharp end's actual experience (surprise, confusion, procedural failure) was filtered before reaching the blunt end. Operational variety was thus compressed into technical parameters (probe model, icing conditions, duration of speed loss) while the human factors dimension, the signal most relevant for anticipating the AF447 scenario, was lost.

This categorical compression channeled the regulatory response toward component reliability rather than human-system interaction. The BEA notes that "actions undertaken aimed at reducing the risk of probe icing through technical modifications" while "operational experience did not lead to analyzing the operational aspects linked to this fault." The feedback mechanisms treated a systemic interface problem as a component problem—a direct manifestation of variety reduction where the regulatory response addressed the wrong dimension because information channels had filtered out the relevant variety.

The certification process reinforced this variety reduction. The "Major" classification of airspeed loss rested on behavioral assumptions about crew response that the BEA describes as "not verified in the context of the accident." The (J)OEB evaluation during A330 certification produced no mandatory training program for this specific fault condition. More fundamentally, the acceptance of neutral static longitudinal stability in alternate law, based on the presence of Normal Law protections, created a hidden conditional dependency. The certification framework, by filtering scenarios through predefined severity categories and assumed crew capabilities, systematically excluded the possibility that a trained crew might simultaneously lose trajectory control and diagnostic capability.

Regulatory surveillance exhibited a structural variety deficit: Air France underwent approximately 80 in-flight checks for 350,000 annual flights (1:4,000 ratio), and during checks "crews know what is expected and generally

manage to avoid showing" deviations (BEA, 2012). An internal Air France safety report had identified "sometimes weak piloting capabilities," "a loss of common sense," and "declining general aeronautical knowledge", yet these internal signals, generated within the airline's organizational boundary, did not reach regulatory actors with sufficient fidelity to trigger corrective action. The surveillance system was structurally incapable of capturing the variety of actual line operations.

4.3. Dimension 3: Trust degradation and collective automation biases

4.3.1. Technical-level trust

The three trust components identified by Hoff and Bashir (2015) interacted to produce a pathological configuration at the moment of crisis. Dispositional trust operated through professional culture: the PF "may have remained on the habitual schema that the aircraft could not stall and in which a STALL alarm was incoherent" (BEA, 2012). This schema, deeply embedded through years of operating within Normal Law protections, made the stall alarm cognitively inadmissible even as it sounded continuously for 54 seconds. Learned trust, built through uneventful automated flight, had created "confident surveillance of automation due to their performance level and reliability," making manual reversion psychologically difficult and generating a surprise effect disproportionate to the objective technical severity of the initial fault. Situational trust attached to the flight directors, whose green bars became a cognitive anchor despite presenting potentially erroneous commands based on opaque mode logic. The conjunction of the memorized pitch value from stall recovery training (12.5°) with the flight director's orders constituted, for the PF, the sole coherent reference in total confusion—a trust attachment that directly instantiates the model's prediction that opacity paradoxically reinforces, rather than undermines, trust in validated systems.

4.3.2. Organizational-level trust chains

The model predicts that technical-level automation biases replicate at the organizational scale through structural homology, and the AF447 case provides compelling evidence. The absence of positive static longitudinal stability in alternate law had been accepted by certification authorities through "special conditions and particular interpretations" predicated on Normal Law protections. This conditional acceptance was not propagated through the organizational chain: pilots were never informed that their aircraft lacked a fundamental aerodynamic characteristic their basic training taught them to expect. The certification authority treated the manufacturer's design rationale as a validated black box output—organizational automation bias in its purest form.

EASA's analysis of Pitot icing events "confirmed the criticality of the failure" yet concluded not to mandate probe replacement. This risk assessment became a black box input for Air France's safety management. The BEA reveals circular logic: "the existence of an operational procedure associated with airspeed loss led the operator, manufacturer, and authorities to consider that the risk was controlled, in the absence of significant trajectory control loss during known events." Each actor trusted the other's conclusions without mastering the underlying methods or assumptions—the organizational analogue of the individual pilot trusting the flight director without understanding its mode logic. Task focus became strict perimeter delimitation: the manufacturer addressed probe design, the authority assessed certification compliance, the airline managed training schedules, but no actor examined the interfaces where cumulative vulnerability accumulated.

4.4. Synthesis: the organization as black box

The application of the three-dimensional model to AF447 reveals how the dimensions interact to produce an emergent systemic vulnerability irreducible to individual component failures. Bidirectional opacity created the conditions for failure: pilots could not understand the automation's logic, while designers could not perceive the gap between prescribed and actual crew behavior. Requisite variety reduction ensured that corrective signals—non-application of procedures, cognitive inadequacy of altitude-specific training, conditional aircraft stability—were systematically filtered before reaching actors with regulatory power. Trust degradation stabilized the system in its vulnerable state: each

organizational black box trusted the others' outputs, generating collective complacency that inhibited the critical questioning necessary to detect risk accumulation at the interfaces.

The BEA's conclusion captures this emergent property: "the double failure of the planned procedural responses shows the limits of the current safety model." Reason (1997, 2000) recognized that latent conditions are embedded in organizational decisions and can interact. However, the defensive barrier metaphor—widely adopted in safety practice—implicitly frames breaches as independent failures that happen to align. The AF447 case challenges this reading: the three dimensions of the Organizational Black Box model reveal mutually reinforcing mechanisms rather than coincidentally co-occurring deficiencies. The accident is the predictable product of an organizational architecture where opacity, variety reduction, and trust degradation co-produce a structurally blind system—an organizational black box whose vulnerability remains invisible to all constituent actors. Erroneous assumptions—that low-altitude stall training prepares for high-altitude events, that a procedure's existence guarantees its application, that certification-validated characteristics hold across all operational modes, that each actor's risk assessment integrates dependencies embedded in others' assessments—propagated across organizational boundaries without critical examination, carried by the delegated trust that compensates for, but simultaneously perpetuates, opacity at each interface.

5. Conclusion

This paper proposed the Organizational Black Box model to formalize how opacity, variety reduction, and trust degradation interact as a self-reinforcing dynamic at the sharp end/blunt end interface. The AF447 case demonstrated that these three dimensions are not independent contributing factors but a circular architecture that explains why latent vulnerabilities persist—and remain invisible to all constituent actors—despite competent actors at every level.

The model offers three action levers: targeted transparency on critical interface dependencies, institutionalization of calibrated distrust through cross-functional audits, and amplification of organizational variety beyond technical parameters. Several limitations should be acknowledged: the framework was applied retrospectively to a single case, the analysis remains qualitative, and articulation with existing systemic safety models remains to be developed. Future research should test the model's transferability across high-risk domains and explore formal modeling of critical opacity thresholds.

References

Ashby, W. R. (1956). *An Introduction to Cybernetics*. Chapman & Hall.

BEA (2012). *Rapport final sur l'accident survenu le 1er juin 2009 à l'Airbus A330-203 immatriculé F-GZCP exploité par Air France, vol AF 447 Rio de Janeiro – Paris*. Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile.

Cook, R. I., & Woods, D. D. (1994). Operating at the sharp end: The complexity of human error. In M. S. Bogner (Ed.), *Human Error in Medicine* (pp. 255–310). Lawrence Erlbaum Associates.

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1), 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434. https://doi.org/10.1177/0018720814547570

Morin, E. (1990). *Introduction à la pensée complexe*. ESF Éditeur.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30(3), 286–297. https://doi.org/10.1109/3468.844354

Reason, J. (1990). *Human Error*. Cambridge University Press. https://doi.org/10.1017/CBO9781139062367

Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Ashgate.

Reason, J. (2000). Human error: Models and management. *BMJ*, 320(7237), 768–770. https://doi.org/10.1136/bmj.320.7237.768