

Bullfighting in the business arena: Evidence-based strategies for detecting and disposing of organizational bullshit

John V. Petrocelli ^{*} 

Wake Forest University, Winston-Salem, NC, USA

ARTICLE INFO

Keywords:

Bullshit
Organizational communication
Evidence-based management
Decision-making
Truth-telling culture
Impression management

ABSTRACT

Organizations frequently tolerate communication delivered without regard for truth, evidence, or established knowledge in decision-making and accountability contexts—what philosophers and psychologists term *bullshit*. Unlike lies, which deliberately subvert truth, bullshit represents an indifference to accuracy, which experimental research suggests may make it both more prevalent and paradoxically more persuasive in domains where evidence-based reasoning is critical. Studies indicate that bullshit-framed arguments may outperform evidence-based reasoning in influencing attitudes, and produce stronger *sleeping effects* over time, with the potential to gradually degrade organizational decision-making quality. This article focuses specifically on bullshit's harmful effects in decision-making, performance evaluation, and accountability reporting—contexts where factual accuracy is essential—while explicitly recognizing that strategic ambiguity, aspirational talk, and selective communication serve legitimate functions in vision-setting, coalition-building, and innovation contexts. The analysis synthesizes experimental findings from social psychology with organizational cases including Alan Mulally's transformation of Ford Motor Company and the Madoff investment fraud to offer practitioners theoretically informed propositions for building more bullshit-resistant decision-making cultures. A managerial playbook offers theoretically derived implementation frameworks with explicit boundary conditions, including proposed reality-testing protocols and meeting reform guidelines. A decision framework helps practitioners distinguish contexts requiring evidence-based clarity from contexts where ambiguity legitimately serves. Drawing on scholarship in psychological safety, employee voice and silence, impression management, and organizational learning, the analysis explores why bullshit may flourish despite its harmful effects and identifies systemic organizational dynamics that warrant attention.

Managerial challenge

Following the 2014 NFL football season, the Seattle Seahawks played the New England Patriots in the Super Bowl. In the closing seconds of the game, the Seahawks found themselves on the one-yard line with a good chance to score a touchdown and win the game. The Seahawks chose to pass the ball. Unfortunately for the Seahawks, the pass was intercepted and the New England Patriots won yet another Super Bowl. Following the game, critics universally declared the Seahawks should have run the ball instead. If the Seahawks had scored the touchdown and won the game, a lot of people would have said they made a good decision—but they didn't and they lost, and that's why a lot of people said the Seahawks should have run the ball instead. If the Seahawks had run the ball and still failed to score, a lot of people would have said they should have passed the ball instead. Psychological research refers to such tendencies

as the *outcome bias*—judging decisions by their outcomes rather than process quality or a priori probabilities—that underlies much of organizational communication, reasoning, and decision-making.

Although the unfortunate case of the Seahawks is drawn from widely reported public events and is used here solely as an accessible illustration of the outcome bias concept—not as organizational data—the relevant empirical evidence for outcome bias comes from controlled experimental research [1,2]. This research shows that people systematically judge decisions as better when outcomes turn out well and worse when outcomes turn out poorly—even when the decision process is identical. Outcome biases may be especially destructive in organizational settings because they align closely with managerial incentive structures. Leaders are often evaluated, promoted, and rewarded largely on visible outcomes rather than the quality of the reasoning that produced them. In organizations, this bias may create powerful pressure to

* Correspondence to: Department of Psychology, Wake Forest University, P.O. Box 7778, Winston-Salem, NC 27109, USA
E-mail address: petrocjv@wfu.edu.

retrospectively justify decisions using confident, professional-sounding narratives that protect reputation and status. Failed initiatives are explained away with jargon about “market headwinds” and “execution challenges.” Successful projects are attributed to “visionary leadership”—regardless of luck’s actual role. Such the communication pattern, delivered with complete indifference to truth or evidence, represents what philosopher Harry Frankfurt termed *bullshit* [3].

Unlike lies, which at least acknowledge truth’s importance by attempting to subvert it, bullshit represents something potentially more dangerous—a total disregard for accuracy. The liar knows the truth and deliberately contradicts it, whereas the bullshitter neither knows nor cares whether their statements are true. While organizations maintain strong sanctions against lying, they typically treat bullshit as a mild social offense [4–6]. However, research suggests this leniency may be misguided.

The far-reaching consequences of workplace bullshit may operate at individual, organizational, and stakeholder levels. At the individual level, employees exposed to organizational bullshit may respond through exit, voice, loyalty, or neglect [7], and empirical evidence suggests that bullshit is associated with a measurable negative effect on job satisfaction—particularly when it originates from a direct supervisor, whose day-to-day communication shapes employee trust, respect, and engagement [8]. Employees most receptive and inspired by corporate bullshit tend to perform more poorly on assessments of work-related decision-making ability [9]. Though bullshit may temporarily bolster an organization’s legitimacy and individual self-confidence, research suggests these short-term gains may be offset by eroded occupational identities, deepened employee cynicism, and increasingly brittle stakeholder relationships [10]. Organizational bullshit may escalate from informal improvisation into routinized procedure and eventually sacred ritual, making it progressively more difficult to identify and address [11]. When unchecked, bullshitting may produce vicious cycles—tolerated bullshitting may beget more bullshitting until empty and misleading communication becomes the dominant language and culture of the organization [12].

Current evidence base

Defining and distinguishing bullshit

Frankfurt’s foundational analysis [3] established that bullshit differs fundamentally from lying. Liars possess beliefs about truth and deliberately contradict them, whereas bullshitters have no concern for truth whatsoever [6,13]. Subsequent empirical work [14,15] operationalized bullshit as communications delivered with indifference to truth, evidence, or established knowledge in the form of rhetorical strategies that create impressions of knowledge without genuine understanding. Bullshit is often characterized by abstract language, missing sources, business jargon, and confident delivery despite absent evidence.

Empirical research on the social perception of bullshitting [16] finds that people tend to evaluate bullshit less negatively than lies, viewing it as expression of opinion or ignorance rather than deliberate deception. Research investigating the detection of bullshit [17] suggests that consistent exposure without appropriate skepticism may erode one’s ability to discriminate substantial from vacuous claims. Study participants regularly exposed to meaningless statements like “Hidden meaning transforms unparalleled abstract beauty.” showed reduced discrimination between genuinely profound statements and nonsense in subsequent testing. Research on the characteristics of *bullibility*—susceptibility to bullshit even in the face of bullshit cues—links individual differences in bullshit receptivity to poorer decision-making across domains including health, financial, and career choices [18,19].

In organizational contexts, bullshit most often appears as strategic ambiguity, jargon-laden explanations, unverifiable claims about “best practices,” and confident assertions unsupported by data [7,10,12]. Importantly, much organizational bullshit is not intentionally deceptive.

Managers frequently produce it under time pressure, uncertainty, or social expectations to appear knowledgeable and decisive.

The persuasive power of bullshit

Experimental research [15] reveals that bullshit can be more influential than evidence-based communication. In studies presenting identical arguments with either bullshit frames (“I’m not really concerned with the evidence...”) or evidence-based frames (“I’m really concerned with the evidence concerning this issue...”), bullshit frames eliminated the traditional argument quality effect. Strong and weak arguments became equally persuasive when framed in bullshit (e.g., “I don’t care about what the data suggest...”), while evidence-based frames (e.g., “I care a great deal about what the data suggest...”) maintained appropriate discrimination between argument quality.

These patterns suggest bullshit may cue what psychologists term “peripheral route processing” [15]—evaluation based on superficial cues like speaker confidence rather than substantive analysis. When confronted with bullshit, the cognitive load required to first decipher claims through layers of jargon, then evaluate their merit, may create vulnerability to influence through style over substance. Further experiments [15] demonstrated that source attractiveness predicted persuasion under bullshit frames but not evidence-based frames, a pattern consistent with the interpretation that bullshit may systematically shift evaluation from content to peripheral cues.

These experimental findings derive primarily from laboratory studies with undergraduate participants. While this research provides valuable controlled tests of causal mechanisms, the generalizability to organizational decision-making contexts—involving experienced professionals, high stakes, complex information environments, and organizational politics—requires careful consideration. The mechanisms identified (peripheral processing, source-based evaluation) have been documented in field settings across multiple domains, lending confidence to their organizational relevance. However, organizational validation through field studies remains an important direction for future research [20,21].

Importantly, the laboratory-to-organization translation question is not merely procedural—it requires substantive engagement with specific moderators where the direction of effect is genuinely uncertain. Four are most consequential for organizational decision-making contexts.

Domain expertise. The intuitive assumption is that domain expertise protects against bullshit susceptibility: an experienced professional should be able to evaluate evidence quality within their field and resist peripheral processing. The empirical picture is more complicated. Research on the antecedents of bullshitting finds that people generally bullshit less as their self-reported knowledge of a topic increases—but this protective effect depends critically on whether knowledge is actual or merely felt [14]. People with *perceived* expertise—a feeling of knowing without the underlying competence—are especially prone to both producing and accepting bullshit in their domain, because they may lack the ability to detect where claims diverge from evidence [18]. Doing so may create a particular risk in organizations, where positional authority is frequently mistaken for domain mastery. Moreover, research on illusory truth suggests that domain knowledge may not reliably protect against fluency-based belief—even knowledgeable participants show the illusory truth effect for statements in their area of expertise [22]. Expertise may therefore confer protection against some bullshit—particularly claims that directly contradict established knowledge—while providing little protection against bullshit that is merely unverifiable or expressed with professional-sounding confidence. In organizational settings, where bullshit typically consists of confident but ambiguous claims rather than outright factual errors, expertise may offer less protection than intuition suggests. Whether the net effect of expertise is protective or amplifying in specific organizational bullshit contexts is an open empirical question yet to be directly tested.

Power differentials. The laboratory experiments [15] presented arguments attributed to peers with no organizational authority over participants. In organizations, bullshit may originate from individuals who hold power over the careers, evaluations, and resources of those receiving the communication. Power differentials may compound bullshit susceptibility through at least two mechanisms. First, organizational hierarchies generate implicit pressure toward deference—research on employee voice and silence indicates that employees are often reticent when interacting with superiors, with evolved and socialized tendencies to avoid challenging authority [23]. These tendencies may operate prior to deliberative processing and may be magnified in high-power-distance organizational cultures [24]. Second, when the bullshitter holds organizational authority, audience members face an additional cognitive task—managing the social consequences of challenge—that consumes capacity that might otherwise be devoted to evidence evaluation. The combined implication is that power differentials likely amplify rather than attenuate the peripheral processing shift that bullshit framing produces—but this specific interaction has yet to be directly examined.

Time pressure and cognitive load. The elaboration likelihood model proposes that the central-versus-peripheral processing distinction is fundamentally a function of motivation and *ability* to carefully scrutinize arguments [15,25,26]. Ability to process centrally is likely reduced under cognitive load and time pressure—conditions that define organizational decision-making contexts far more than they defined the laboratory studies. Senior leaders routinely make consequential decisions after reviewing dense slide decks in compressed timeframes, under conditions of multitasking and competing attentional demands. Research on persuasion and dual-process cognition suggests that even highly capable individuals may default to heuristic processing under cognitive load, shifting evaluation from argument quality to peripheral cues such as source confidence and delivery [25–27]. This suggests that time pressure and cognitive load may specifically amplify the bullshit peripheral processing effect—creating organizational contexts where the laboratory findings likely *underestimate* susceptibility rather than overestimate it. Managers and practitioners should therefore treat time-compressed decision settings as especially high-risk environments for bullshit penetration, precisely because those conditions impair the central route processing that would otherwise allow evidence quality to moderate the bullshit effect.

Decision stakes. High personal consequences might be expected to increase motivation to scrutinize claims carefully, pushing evaluation toward the central route—and thus reducing susceptibility to bullshit. Research on the role of accountability supports this expectation. When people anticipate explaining and justifying their positions to a critical, unlike-minded audience, bullshitting is significantly attenuated [14]. However, accountability effects depend heavily on the direction of the stakes. When career consequences are tied to appearing decisive, confident, and aligned with a powerful communicator—rather than to the accuracy of the evaluation—high stakes may perversely increase peripheral processing by raising the social cost of challenging a confident speaker. The impression management literature suggests that organizational reward systems frequently create precisely these incentive structures [28]. Stakes therefore may have an uncertain and potentially bidirectional moderating effect—they may increase scrutiny when audiences are accountable for evaluation accuracy, while increasing deference when audiences are accountable for relational conformity. Which dynamic dominates in any given organizational decision context is an empirical question that current research cannot definitively answer.

Taken together, these four moderators suggest a sobering possibility—organizational contexts may amplify rather than attenuate the core laboratory effects, due to the convergence of power deference, cognitive load, and status-based impression management incentives that collectively depress central route processing. Such possibilities do not undermine the theoretical mechanisms identified in the experimental literature—peripheral processing, sleeper effects, and illusory truth

amplification remain theoretically coherent and internally well-replicated. However, it does underscore the urgency of field-based research that tests these mechanisms in real organizational settings, and it reinforces rather than weakens the case for the structural countermeasures introduced in Section 4.

Yet, these dynamics are familiar to most anyone who has observed high-stakes organizational communication. In boardrooms, strategy off-sites, or consultant-led presentations, decision-makers are often confronted with dense slides, abstract language, and confident delivery under severe time constraints. Research on persuasion and dual-process cognition suggests that even experienced leaders may default to heuristic processing under cognitive load, relying on emotional reactions and superficial cues rather than systematic evaluation [27]. In such environments, the cognitive effort required to disentangle substance from rhetoric is substantial. Bullshit thrives precisely under these conditions—high cognitive load, limited time, and social pressure to move forward—where even highly experienced professionals may default to evaluating the communicator rather than the content.

The sleeper effect and illusory truth

Research on delayed persuasion effects [29] indicates that bullshit may produce stronger persuasion effects than lies (i.e., *sleeper effect*—delayed persuasion that results when people initially discount a message due to a noncredible source, but over time forget the source while remembering the message, leading to increased persuasion later than immediately after exposure). When participants learned information came from a bullshitter versus a liar, they initially discounted the bullshit less severely. After 14 days, bullshit-attributed information regained credibility levels, while lie-attributed information remained discounted. This occurs because lies can be definitively flagged as false, while bullshit remains ambiguously or possibly true.

The *illusory truth effect*—where repeated exposure increases perceived truthfulness—further amplifies bullshit’s impact [30]. Information repeatedly flagged as bullshit became more believable over time, while information flagged as lies did not. Repeated organizational exposure to bullshit thus degrades collective capacity to distinguish truth from falsehood, even when bullshit is initially recognized as such.

These mechanisms help explain organizational pathologies like the Bernie Madoff investment scandal. The account of Harry Markopolos—the one analyst who did identify Madoff’s fraud—suggests a pattern consistent with illusory truth dynamics [31]. Investors repeatedly exposed to Madoff’s confident but unsubstantiated claims about proprietary trading strategies initially questioned them, but repeated exposure made them appear more than plausible. Markopolos documented this dynamic—by repeating that story many hundreds of times to the thousands of potential investors he met through the years, Madoff actually came to believe it himself. And because Madoff believed it, it made it even easier for investors to believe [31]. An important interpretive boundary applies here: Madoff was engaged in deliberate, criminal fraud—a category distinct from the bullshit construct, which involves indifference to truth rather than calculated deception. The case is used here not to equate fraud with organizational bullshit, but to illustrate how repeated confident assertion, regardless of its underlying intent, may produce perceptions of credibility over time through a subjective sense of familiarity. The psychological mechanism (illusory truth) is the analytical focus; the case provides a dramatic illustration of its potential scale, not independent evidence of the mechanism itself.

It is worth acknowledging that the empirical findings reviewed in Sections 2.1 through 2.3—concerning bullshit’s persuasive effects, its differential social perception relative to lying, sleeper effects, and illusory truth amplification—derive primarily from a single, developing research program [6,14–16,18,19,29]. This concentration reflects the youth of organizational bullshit as a formal area of scientific inquiry rather than a limitation of the evidence per se. Within this research program, the core findings are well-replicated. Each published study

includes multiple experiments—for example, the bullshit persuasion effects [15] and the bullshit sleeper effect [29] were each replicated across two or more independent experiments within their respective publications, with consistent results across varied attitude objects, participant samples, and procedural variations. Internal replication of this kind provides meaningful confidence in the reliability of the effects under controlled conditions. Independent replication from outside this research program, however, does not yet exist for these specific phenomena. The exceptions [9,17,32] are independent, but address meaningfully different questions (e.g., the reception of pseudo-profound but meaningless statements, rather than the persuasive consequences of bullshit framing in social influence contexts). These are related but distinct phenomena, and these exceptions cannot serve as independent replication of the persuasion, sleeper effect, or illusory truth effects reported here. What this means for practitioners is straightforward: the mechanisms identified—peripheral processing under bullshit framing [15,33], delayed credibility recovery for bullshit versus lies [29], and illusory truth amplification through repetition [30]—are theoretically coherent, internally consistent across multiple experiments, and grounded in well-established persuasion science. They are not, however, yet independently verified. Practitioners should treat these findings as directionally informative and theoretically well-grounded rather than definitively settled, and field-based validation in organizational contexts remains an important priority for this research program.

Boundary conditions: when anti-bullshit guidance applies

If bullshit degrades decision quality, should organizations demand truth, clarity, and evidence in all communication? Research on organizational communication suggests a nuanced answer.

Strategic ambiguity—deliberately vague or multivocal communication—can serve legitimate organizational functions in specific contexts [34,35]. When organizational goals are contested, interests conflict, or creativity is needed, ambiguity can facilitate collective action by permitting multiple groups to read their own meaning into their shared interests and aims [34]. In business schools navigating accreditation while preserving research priorities, managers used “protective ambiguity” to shield specific interests and “invitational ambiguity” to create common ground [34]. In R&D contexts, strategic ambiguity can facilitate dialog across disciplinary boundaries and appreciation of diverse perspectives [35].

Similarly, *aspirational talk*—forward-looking rhetoric about ideals not yet realized—plays a constructive role in organizational change [36]. Vision statements are not meant to be precise operational descriptions but rather motivational pictures of possible futures. Christensen and colleagues argued that corporate social responsibility communications, even when they exceed current practice, can be “performative”—bringing desired realities into being by publicly committing to higher standards [36].

The key distinction lies in purpose and context. Aspirational talk serves legitimate functions when it articulates vision and ideation rather than current operational reality, motivates stakeholders toward higher standards, or creates productive tension between current state and desired future. In contrast, bullshit is most likely to degrade organizational functioning when it substitutes for evidence in decision-making processes, obscures problems requiring operational response, or protects status at the expense of organizational learning.

The recommendations in this article apply primarily to decision-making contexts where evidence quality directly affects outcomes (e.g., strategic choices, performance evaluations, resource allocations, problem diagnosis, operational planning). These contexts require clear distinctions between what is known, what is speculated, and what remains uncertain.

Organizational conditions enabling truth-telling

Understanding why bullshit flourishes requires examining the organizational conditions that suppress candor. Four interconnected literature streams illuminate these dynamics.

Psychological Safety. Amy Edmondson’s research suggests that *psychological safety*—a shared belief that the team is safe for interpersonal risk-taking—is a prerequisite for candor in organizations [24,37–39]. When psychological safety is low, employees withhold concerns, problems, and dissenting views because speaking up feels risky. Edmondson’s seminal study of manufacturing teams found that high psychological safety teams reported more errors—not because they made more mistakes, but because members felt safe admitting them [37]. Edmondson and Bransby [24] and Edmondson and Kerrissey [39] confirm psychological safety predicts performance especially in contexts requiring discretion, creativity, or genuine collaboration. In reviewing evidence across software development contexts, Santana and colleagues [40] find that psychological safety enables voice, reduces silence, and facilitates organizational learning.

Psychological safety does not mean niceness nor comfort; it means candor without fear of punishment [39]. Creating it requires leaders to actively signal that voice matters through three interrelated practices: messaging (explaining why input is needed), modeling (admitting fallibility and asking questions), and mentoring (responding productively when people speak up) [39]. The Ford transformation case (discussed in Section 3.5) exemplifies these dynamics: Mark Fields’ admission of problems was safe only because Alan Mulally (CEO) had explicitly established the conditions for honest communication.

Voice and Silence. Research on employee voice and silence provides complementary insights into upward communication dynamics [23,41]. *Voice*—discretionary upward communication of ideas, concerns, or information about problems—is often withheld even when employees possess valuable information. Multiple barriers include: hierarchical structures that make voice feel inappropriate, cultures that implicitly discourage dissent, and employees’ calculations that remaining silent is safer than speaking up [23,41].

Importantly, silence is not merely absence of voice but an active choice to withhold information that could be relevant. Organizational silence may represent a “barrier to change and development” because problems remain hidden, opportunities are missed, and leaders operate with incomplete information [23]. The decision to voice or remain silent involves complex motivational processes including self-protective motives (e.g., fear of being labeled a complainer), relational concerns (e.g., maintaining relationships with colleagues), and futility beliefs (e.g., “My voice won’t make a difference anyway.”).

Morrison’s framework can help explain why bullshit may proliferate within organizations. When honest communication feels risky, employees may default to safer alternatives—either remaining silent or producing superficially plausible communication that protects their image. Itzhakov and Grau [42] find that high-quality listening—where managers truly attend to employee input—may create conditions more conducive to candid exchange and may reduce defensive communication patterns.

Impression Management. Efforts to create, maintain, or protect desired images—that is, *impression management*—are a ubiquitous part of organizational life [28]. Employees use various tactics, including ingratiation, self-promotion, and exemplification, to shape how others perceive them, particularly in performance-relevant contexts [28]. When accuracy threatens the impression an employee wishes to project, bullshitting may serve as an impression management strategy.

Crucially, impression management motives often conflict with accuracy motives. When asked for honest assessments, people may prioritize “looking good” over “being accurate”—especially when accountability is high but information verifiability is low. Organizational reward systems frequently reinforce impression management over candor: promotions go to those who claim credit convincingly,

performance reviews reward confident self-presentation [28,43] and failures are attributed to external factors while successes are internally attributed [44]. When success depends more on impressive-sounding explanations than verifiable results, and when admitting uncertainty threatens one's image of competence, producing confident but evidence-free communication becomes strategically rational.

Organizational Learning Barriers. Research investigating *organizational learning* identifies defensive routines that inhibit learning from failures [45]. Professionals, particularly successful ones, have rarely experienced failure and thus have limited opportunities to learn from it [45]. When their usual problem-solving approaches prove inadequate, professionals tend to bullshit—deflecting responsibility outward and resisting the kind of critical self-reflection that genuine learning requires (Kiazad, Chen, & Restubog, in press).

Furthermore, one's *theories-in-use*—the actual governing values that shape their behavior—may differ considerably from their more bullshit-based, *espoused theories*. People claim to value openness and inquiry but act to remain in control, maximize winning, suppress negative feelings, and appear rational. Such defensive reasoning may create “organizational defenses”—patterns of interaction that inhibit questioning of underlying assumptions [45]. More recent approaches to defensive behaviors [46] suggest that such patterns may persist even in organizations explicitly committed to psychological safety, requiring sustained attention and deliberate counter-measures.

The organizational learning framework [45] helps to explain why communications like bullshit may persist—they serve defensive functions (Kiazad, Chen, & Restubog, in press). Admitting uncertainty or error threatens feelings of competence, whereas producing confident explanations preserves the appearance of control. Organizations compound this by rewarding people who appear decisive and confident while punishing those who acknowledge mistakes or express doubt. Overcoming these barriers requires making reasoning processes explicit, testing assumptions, and modeling vulnerability at senior levels.

Findings and insights

The four organizational behavior literatures reviewed in Section 2.5—psychological safety, voice and silence, impression management, and organizational learning—each illuminate *why* honest communication is organizationally difficult. They explain why employees stay silent, why leaders reward the appearance of confidence over accuracy, and why organizations resist learning from failure. These are essential contributions. But they do not explain what happens *after* people speak—specifically, what happens when communication that disregards evidence enters a decision-making conversation. For that, a distinct theoretical apparatus is required.

The experimental bullshit literature provides three mechanisms unavailable from any of the adjacent literatures. These mechanisms explain phenomena that psychological safety theory, voice research, and impression management frameworks were not designed to address: why evidence-free communication can sometimes outperform evidence-based communication even when people are listening carefully; why bullshit-based narratives become more credible over time while lie-based narratives do not; and why organizational exposure to bullshit degrades the collective capacity to distinguish truth from falsehood even when individuals initially recognize it as bullshit. Each mechanism has distinct practical implications and demands interventions that cannot be derived from the literature reviewed in Section 2.5 alone.

Finding 1: Bullshit triggers peripheral processing, bypassing critical evaluation

Psychological safety research explains why people withhold concerns and stay silent [37,41]. It does not explain—and was not designed to explain—what happens when people do speak, and what they say is delivered with indifference to evidence. The experimental evidence [15]

suggests a mechanism specific to bullshit: when communication is framed with explicit disregard for evidence (“I’m not really concerned with the data here...”), receivers tend to shift from *central route processing* (i.e., careful evaluation of argument quality) to *peripheral route processing* (i.e., superficial cues like speaker confidence, attractiveness, and fluency drive judgment).

Prior persuasion research [47,48] suggests that untrustworthy or biased sources trigger *greater* scrutiny, not less. Bullshit violates this expectation. Because bullshit signals that the communicator is not taking evidence seriously, receivers follow the communicator’s lead—they stop scrutinizing evidence too. The bullshitter and the evidence-based communicator effectively *model* the processing mode they expect from their audience [15]. In controlled experiments, this shift meant that strong and weak arguments became equally persuasive under bullshit frames, while evidence-based frames maintained appropriate discrimination between argument quality. Source attractiveness predicted persuasion only under bullshit frames, not evidence-based ones.

For organizations, this finding points to a phenomenon that psychological safety research was not designed to address: why boardrooms full of psychologically safe, engaged, truth-willing professionals may still produce poor decisions. If a senior executive frames a strategic recommendation with confident jargon and apparent certainty regardless of evidence, even well-intentioned listeners may shift toward peripheral processing. Research suggests susceptibility to bullshit receptivity is widespread [32] and that expertise may not reliably protect against it [22]. The theoretical implication is not only that organizations need to make people safer to speak (the psychological safety solution), but that organizations may also need structural countermeasures that interrupt the peripheral processing shift itself (i.e., encouraging listeners to return to central route evaluation at the moment bullshit cues are detected). Section 4.8 introduces “processing restoration protocols” designed specifically for this purpose.

Finding 2: Bullshit produces stronger and more durable persuasion effects than lies

Voice research and impression management theory explain why people produce deceptive or evasive communication. They do not explain—and were not designed to explain—the differential *downstream consequences* of different types of untrustworthy communication. Bullshit sleeper effect research [29] suggests a finding with potentially important organizational implications: bullshit and lies may follow categorically different temporal trajectories of influence.

In controlled experiments [29], participants who learned they had been exposed to either a liar or a bullshitter initially discounted both sources. But after 14 days, bullshit-attributed information had fully rebounded to near-control levels of believability, while lie-attributed information remained relatively discounted. The mechanism appears to be the ambiguity of bullshit: because bullshit is not necessarily false—the bullshitter simply does not know or care—people cannot decisively flag bullshit-derived beliefs as incorrect the way they can flag known lies. The discounting cue (i.e., “That was bullshit.”) erodes in memory faster than the original message content, while the discounting cue for lies (i.e., “That was false.”) remains anchored to the falsity judgment and continues to suppress belief.

The difference in discounting cues may matter for organizational narrative management. Organizations routinely recirculate strategic narratives—origin stories, acquisition rationales, performance explanations, vision statements. If the sleeper effect dynamics identified in the laboratory extend to organizational settings, bullshit-based narratives may not simply fade or be corrected over time; rather, they may gain credibility through the passage of time. This possibility points to a dynamic that voice research and impression management frameworks were not designed to address: why organizational bullshit may be resistant to correction even in psychologically safe environments. The problem may not only be that people stay silent, but that once bullshit

enters organizational memory, conditions may exist for it to regenerate. Section 4.9 introduces “narrative origin audits” designed to address this possibility.

Finding 3: Repeated exposure to bullshit corrupts collective epistemic capacity

Organizational learning theory [45,49] explains why organizations resist learning from failure through defensive routines. Neither framework addresses what happens to an organization’s collective ability to distinguish true from false claims as a function of repeated exposure to evidence-free communication. Petrocelli and colleagues’ illusory truth effect research [30] suggests that bullshit may produce a stronger illusory truth effect than lies or neutral communication—repeated exposure to bullshit-framed statements appears to increase their perceived truthfulness significantly more than repeated exposure to the same statements framed as lies. The mechanism reflects the same ambiguity asymmetry as the sleeper effect—because bullshit is not definitively false, familiarity with a bullshit claim accumulates as a positive epistemic signal rather than a warning signal. Repetition increases the subjective fluency of processing the claim, and fluency is misattributed to truth. For lies, the false-label remains available as a corrective anchor, but for bullshit, no such anchor exists. Critically, this illusory truth effect emerges even when participants initially recognized the statements as bullshit—the effect persists through correction attempts [30].

If these laboratory findings extend to organizational settings, each repeated status meeting where problems are systematically understated, each quarterly review where optimistic projections go unchallenged, and each strategy session where jargon substitutes for analysis may not simply pass without consequence—each instance may compound. On this account, repeated organizational exposure to bullshit could progressively degrade the collective capacity to identify when claims require scrutiny, raising the threshold of what “sounds plausible.” This would represent a form of epistemic erosion that Argyris’s [45] defensive routines framework does not directly address, because it operates not through individual defensiveness but through the mechanical process of familiarity-induced belief. Section 4.10 introduces “organizational inoculation protocols”—theoretically derived interventions targeting the repetition channels through which bullshit may become embedded in organizational memory.

Finding 4: Bullshit receptivity is universal

Receptivity to bullshit does not appear to be a straightforward function of low intelligence, low expertise, or low motivation [15,22]. Research suggests it may instead reflect a predictable cognitive response to a specific communication pattern—confident delivery combined with explicit or implicit disregard for evidence—that may operate broadly across audience types. Research further suggests that individuals systematically miscalibrate their own bullshit-detection abilities, with high-performers tending to underestimate their susceptibility and low-performers overestimating it [32]. Such miscalibration suggests that training individuals simply to “be more skeptical” or “think more critically” may be insufficient on its own. The problem may not be a cognitive skill deficit but a contextual trigger that can bypass existing critical thinking skills.

The Bernie Madoff investment fraud illustrates bullshit receptivity at scale, though the case requires careful analytic framing. The primary data source is Markopolos’s investigative account [31], supplemented by subsequent regulatory and legal proceedings—all retrospective reconstructions rather than prospective organizational observation. Within those limits, the case is consistent with the theoretical account of universal bullshit receptivity: Madoff’s investment fraud succeeded for 18 years partly because sophisticated investors failed to apply basic plausibility tests to claims that violated elementary arithmetic.

Madoff claimed to achieve consistent returns through a “split-strike

conversion strategy,” using financial jargon that sounded sophisticated enough to justify his seemingly impossible performance. Harry Markopolos, the fraud investigator who tried to expose Madoff, understood what should have been obvious to any careful observer. Madoff’s claimed strategy required far more put options than existed in the entire market—he would have needed \$3 billion to \$65 billion worth of options that simply did not exist. His returns showed impossible consistency, corresponding to only 6% of market fluctuations despite claiming to be market-driven. When plotted on a graph, Madoff’s performance showed a steady 45-degree rise while the S&P 100 displayed its normal volatility. Yet thousands of sophisticated investors, blinded by Madoff’s reputation and status, failed to apply basic plausibility tests to his claims [31]. The combination of impressive credentials, confident presentation, and technical-sounding explanations is consistent with the pattern of bullshit receptivity the experimental literature describes. Several alternative explanations deserve acknowledgment, however. Investor due diligence failures may have reflected regulatory overconfidence, deliberate willful blindness motivated by fee arrangements, or social network dynamics in which peers’ endorsements substituted for independent analysis—none of which require a bullshit-specific mechanism to explain. The case is therefore used here as a contextually plausible illustration of peripheral processing at scale, not as independent organizational evidence that bullshit receptivity was the operative mechanism.

Both empirical research and cases like the Madoff investment fraud suggest receptivity to bullshit is not a failure of intelligence, but rather a product of *contextual cues*. Without such external mechanisms to interrupt peripheral route processing and force attention to evidence quality, effective countermeasures should therefore be *structural*, not dispositional. Organizations may therefore benefit from building structural features (e.g., meeting formats, decision protocols, documentation requirements) designed to elicit central route processing, independent of any individual’s in-the-moment motivation. The Managerial Playbook’s structural interventions (Section 4) are designed as theoretically informed proposals for this problem.

Finding 5: Truth-focused cultures create competitive advantage—the Ford transformation

Alan Mulally’s transformation of Ford Motor Company is used here as an extended illustrative case. The primary data source is a single journalistic account [50]—a richly detailed but non-systematic reconstruction that cannot substitute for primary organizational research. The case is analyzed here not as confirmatory evidence but as a theoretically interpretable instance: one in which the conditions the framework identifies as necessary for bullshit-resistant culture appear, on the available account, to have been present, and in which organizational outcomes consistent with the theoretical predictions appear to have followed. The evidentiary limitations of this reading are discussed explicitly at the end of this section. With those caveats in mind, the case proceeds as follows. Upon becoming CEO in 2006 with Ford hemorrhaging \$17 billion annually [50], Mulally instituted weekly Business Plan Reviews using color-coded status reporting (green/yellow/red). For seven weeks, every executive reported green despite massive losses.

The breakthrough came when Mark Fields displayed the first red light for an unresolvable suspension problem. Instead of punishment, Mulally applauded. Other executives immediately offered solutions; the problem was resolved quickly. Subsequent meetings produced what observers called “a rainbow of colors you would expect to see from the truth.” While competitors required government bailouts, Ford ultimately avoided bankruptcy and returned to profitability.

Mulally’s response is consistent with the psychological safety framework [37], and can be read as an illustration of its three interrelated practices in action. First, he used messaging to explain why voice mattered: Ford faced existential crisis requiring everyone’s input to identify and solve problems. Mulally explicitly declared the weekly

Business Plan Review (BPR) meetings as forums where candor would be rewarded rather than punished.

Second, Mulally modeled the desired behavior. When Mark Fields courageously presented the first “red light”—admitting his division had serious problems with a new model launch—the room went silent, expecting Fields’ punishment. Instead, Mulally led applause, asking “Who can help Mark with this?” This public reinforcement signaled that admitting problems was safe and valued.

Third, Mulally mentored through productive responses to voice. Rather than reacting defensively or punitively when problems emerged, he focused on collaborative problem-solving. Over time, the BPR meetings filled with red lights as leaders competed to identify and address issues rather than hide them. Ford’s turnaround under Mulally is consistent with what research on employee voice and silence predicts: when leaders remove the structural, cultural, and motivational barriers to speaking up, voice replaces silence [23,41]. Ford’s initial silence stemmed from hierarchical culture, futility beliefs (previous CEOs had punished bearers of bad news), and self-protective motives. Mulally systematically addressed each barrier—flattening hierarchy in BPR meetings, demonstrating that voice led to action, and making accurate reporting safer than silence. Ford’s previous culture rewarded looking good over being accurate—a classic manifestation of impression management eclipsing accuracy. Mulally restructured incentives so that career advancement required identifying real problems and solving them, not hiding problems and claiming success [50]—making evidence-based communication strategically rational and impression management counterproductive [28].

Mulally encouraged aspirational communications for vision-setting (e.g., “One Ford, One Team, One Plan”), but demanded evidence-based candor in operational meetings. Extending well beyond Ford’s c-suite, he systematically embedded accurate reporting throughout Ford’s culture. Mulally lined the Thunderbird Room with 30 additional chairs for guests, allowing executives to invite mid-level managers, engineers, or factory workers to observe meetings. At each session’s end, Mulally would ask guests what they thought, signaling throughout the organization that honest communication was not weakness, but a strength. Research on busyness culture [51] finds that organizations rewarding impression management over substance may be associated with more bullshit and reduced productivity, engagement, and decision quality.

Ford avoided bankruptcy without government assistance (unlike GM and Chrysler), returned to profitability faster than competitors, and rebuilt product quality ratings. Some analysts attributed this success significantly to Ford’s cultural transformation: creating an organization where problems surfaced quickly, enabling rapid response—suggesting that candor may function as a competitive advantage under these conditions. Arena and colleagues [52] find that culture clusters in networks, spreads contagiously through relationships, and depends critically on context—dynamics consistent with what Mulally appears to have harnessed at Ford.

Viewed through the lens of the bullshit-specific mechanisms identified in Findings 1–4, Mulally’s Business Plan Review system may have accomplished something more precise than psychological safety research alone would predict. The color-coded, data-anchored reporting format functioned in a manner consistent with a structural peripheral-processing countermeasure: it made evidence quality the primary subject of attention, potentially interrupting the shift toward peripheral cues that bullshit framing might otherwise produce. The weekly cadence may have functioned as a de facto narrative audit that limited any bullshit-based explanation from aging into organizational memory long enough to benefit substantially from the sleeper effect. And the consistent repetition of accurate, color-honest assessments—rather than optimistic narratives—is consistent with what an inoculation effect against illusory truth might look like in practice. This interpretation is theoretically coherent, but remains inferential given the evidentiary limitations of the case discussed above.

Before drawing prescriptive conclusions from the Ford case, its evidentiary limitations deserve direct acknowledgment. The account presented here draws primarily from a single popular-press narrative [50] rather than from primary research, systematic case analysis, or independent corroboration. This creates three interpretive constraints that practitioners should hold in mind. First, the account is retrospective and success-biased—Ford’s turnaround is visible and attributable precisely because it succeeded. Organizations that attempted similar candor-focused reforms without achieving the same outcome are absent from this record not because they do not exist, but because failed transformations rarely generate the documentary trail that successful ones do. Second, the case is highly compressed. Complex multi-year organizational change involving thousands of employees, hundreds of decisions, and significant external factors including the 2008 financial crisis is rendered as a single pivotal moment—Mark Fields’ first red light. Such compression likely understates the duration, difficulty, and contingency of the actual transformation, and risks creating an overly optimistic template for practitioners. Third, the case cannot rule out alternative explanations. Ford’s survival without a government bailout may reflect product decisions made before Mulally’s tenure (Ford had mortgaged its assets to secure credit in 2006, before the financial crisis made credit unavailable), timing advantages, prior financial hedging, or external market factors distinct from its communication culture. The Business Plan Review Meeting may have been a correlate of Ford’s recovery as much as a cause—leaders committed to transparency may also make better product and financial decisions through channels unrelated to communication format. These limitations do not negate the case’s illustrative value. The patterns it displays (public reinforcement of honest reporting, structured status formats, leader modeling) are consistent with the research literature on psychological safety and organizational learning. It should therefore be read as a plausible illustration of those mechanisms in action, not as independent organizational evidence for their effectiveness.

A counterexample: When truth-telling initiatives stall

The Ford case illustrates a best-case trajectory. A contrasting pattern—one that is arguably more common and more instructive—is visible in the documented experience of organizations that implement candor initiatives without adequately addressing the structural and motivational conditions that bullshit-resistant cultures require. Wells Fargo’s repeated internal escalation failures between 2013 and 2016 provide a well-documented and formally investigated instance of this pattern. The primary data sources for this account are a U.S. Senate investigative report [53] and an independent case analysis from Stanford’s Graduate School of Business [54]—both based on documentary evidence and formal testimony rather than ethnographic or organizational research. These sources establish the factual contours of the case and the theoretical interpretation applied here should be read as analytic inference rather than documented organizational fact. Employees across the retail banking division reported awareness of the fraudulent account-opening practices—later confirmed to involve millions of unauthorized accounts—through multiple internal channels, including ethics hotlines and direct manager communications, years before the practices became public [53]. The organization was not, on its surface, lacking in truth-telling infrastructure, as it maintained formal whistleblower channels, ethics training programs, and stated values explicitly committed to integrity. What it appeared to lack were the structural conditions that make truth-telling organizationally rational. Employees who raised concerns were documented to have faced retaliation, reassignment, or termination. The incentive system—aggressive cross-selling targets enforced through high-frequency performance monitoring—made impression management over accuracy the only rational career strategy. The organizational narrative that Wells Fargo was a community-oriented, customer-first institution was repeated with such regularity across internal communications, training materials, and

public statements that it became, in the terms of this paper, an institutionalized bullshit-based narrative. Rather than being established through evidence and tested against reality, it was simply repeated—and the more familiar it became, the harder its disconnection from operational reality was for insiders to articulate without appearing disloyal [53,54].

Alternative explanations for this pattern deserve acknowledgment. The escalation failures may have reflected straightforward managerial suppression of voice without any illusory truth mechanism—employees may have stayed silent simply because retaliation was certain, not because the institutional narrative had acquired credibility through repetition. The misalignment between espoused and operational culture may also reflect deliberate leadership strategy rather than the gradual epistemic drift that the bullshit-specific framework describes. These alternatives cannot be ruled out on the basis of available evidence. What the case illustrates is the structural signature of a bullshit-prone environment—formal truth-telling channels coexisting with incentive systems that make truthful communication career-threatening—rather than confirming the specific psychological mechanisms through which those conditions operated.

Viewed through the theoretical lens developed here, the Wells Fargo pattern is consistent with all three bullshit-specific dynamics identified in Sections 3.1 through 3.3—though this interpretation is necessarily inferential rather than empirically verified in this context. The peripheral processing dynamic: senior executives evaluating risk reports filtered through layers of jargon-saturated metrics may have had their attention redirected from evidence quality to performance narrative, with confident assertions about the organization's customer-first culture functioning as peripheral cues that displaced scrutiny of underlying data. The sleeper effect dynamic: the institutional narrative of integrity, repeatedly asserted without evidentiary grounding, appears to have become more credible over time to both internal and external audiences even as contrary evidence accumulated. The illusory truth dynamic: repeated organizational exposure to stated values and success narratives may have progressively raised the threshold at which dissonance between espoused culture and operational reality felt salient enough to warrant voice. These mechanism-mappings are offered as theoretically grounded interpretation, not as documented organizational fact.

Critically, Wells Fargo had surface-level equivalents of elements proposed in the Managerial Playbook (Section 4). It had formal voice channels. It had stated candor norms. What it appears to have lacked were the structural conditions that make those elements effective: an incentive system that made accuracy strategically rational rather than career-threatening, structural processing countermeasures that interrupt peripheral evaluation of evidence-free performance narratives, and any systematic process for subjecting core organizational narratives to the kind of origin audit that would have revealed their evidentiary status. The failure is therefore instructive.

The contrast between Ford and Wells Fargo should not be read as a clean natural experiment—both cases involve far too many confounding factors for causal attribution, and both accounts rely on after-the-fact reconstruction rather than prospective organizational research. The alternative explanations for each case named above cannot be ruled out, and the mechanisms proposed here remain inferential. Together, however, the two cases usefully illustrate contrasting structural signatures: one in which data-anchored reporting formats, public reinforcement of honesty, and incentive realignment were present and organizational outcomes improved; the other in which formal voice infrastructure coexisted with incentive systems that made accurate communication career-threatening and organizational outcomes deteriorated. The appropriate evidence standard for both cases is illustrative rather than confirmatory, and that distinction should be held clearly in mind when drawing practical lessons.

Managerial playbook: A 90-day plan to reduce organizational bullshit

This playbook is organized around a critical distinction the existing organizational behavior literature does not make. Psychological safety research [37], voice theory [41], and impression management research [28] together explain *why truth-telling is organizationally difficult*—people fear interpersonal risk, silence is safer than voice, and careers reward impression management over accuracy. The foundational actions in Weeks 1–8 address these barriers. They are necessary conditions for what follows. But they are not sufficient.

Even in psychologically safe environments where voice is encouraged and impression management is de-incentivized, theory and experimental research suggest that three bullshit-specific dynamics may continue to operate: evidence-free communication may shift listeners into peripheral processing regardless of their willingness to speak; bullshit-based narratives may regain credibility over time while lying-based narratives do not; and repeated organizational exposure to bullshit may gradually erode collective epistemic standards through illusory truth effects. These theoretically grounded dynamics suggest the value of interventions derived not only from psychological safety or voice theory, but from the experimental bullshit literature itself. Weeks 9–12 introduce those proposed interventions. Neither phase is likely to work without the other: the structural countermeasures in Weeks 9–12 are most likely to require the cultural foundation of Weeks 1–8, and the cultural foundation alone may be insufficient without the structural mechanisms designed to interrupt bullshit's persuasive dynamics.

Organizational bullshit is most likely to thrive when three conditions exist: 1) people fear speaking up [24,38,41]; 2) people are rewarded for confidence over evidence [14]; and 3) careers advance based on looking good rather than being accurate [28,43]. This playbook offers theoretically informed actions designed to address each condition (see Fig. 1).

Pick one regular meeting you lead—your leadership team meeting, weekly staff meeting, or project review. Use this as your testing ground.

Action 1: Change how meetings start (Week 1)

What it can look like in practice: The director opens the quarterly review the way she always does: “Great work everyone, let’s look at where we are.” Three people in the room know the product launch timeline is no longer viable. None of them say so, because he has never asked.

What to do: Begin your next meeting by saying: “I need everyone’s honest input on [the decision/project/situation]. I don’t have all the answers, and I need to hear concerns, not just agreement.” Then share something you’re uncertain about. Keep it brief: “I’m not sure our timeline is realistic” or “I may be wrong about customer priorities here...”.

How it can work: This action targets the psychological safety barrier—a precondition necessary before bullshit-specific structural interventions can take effect. Edmondson [37] and Edmondson and Kerissey [39] find that people are less likely to speak up until leaders demonstrate it is safe to do so, and that leader modeling is among the most reliable mechanisms for shifting this dynamic. Admitting uncertainty may signal that candor is valued over the appearance of certainty. The link between psychological safety and reduced bullshitting specifically—as distinct from reduced silence generally—has not been directly tested; the connection is theoretically inferred from research on the antecedents of bullshitting [14] and the conditions under which impression management substitutes for accuracy.

What success may look like: In the first meeting, people may still hold back. By meeting 2 or 3, someone may test whether you meant it by raising a concern. How you respond will be critical—thank them before addressing the concern. Say “I appreciate you raising this...” then focus on solving the problem, not defending your position [37,39].

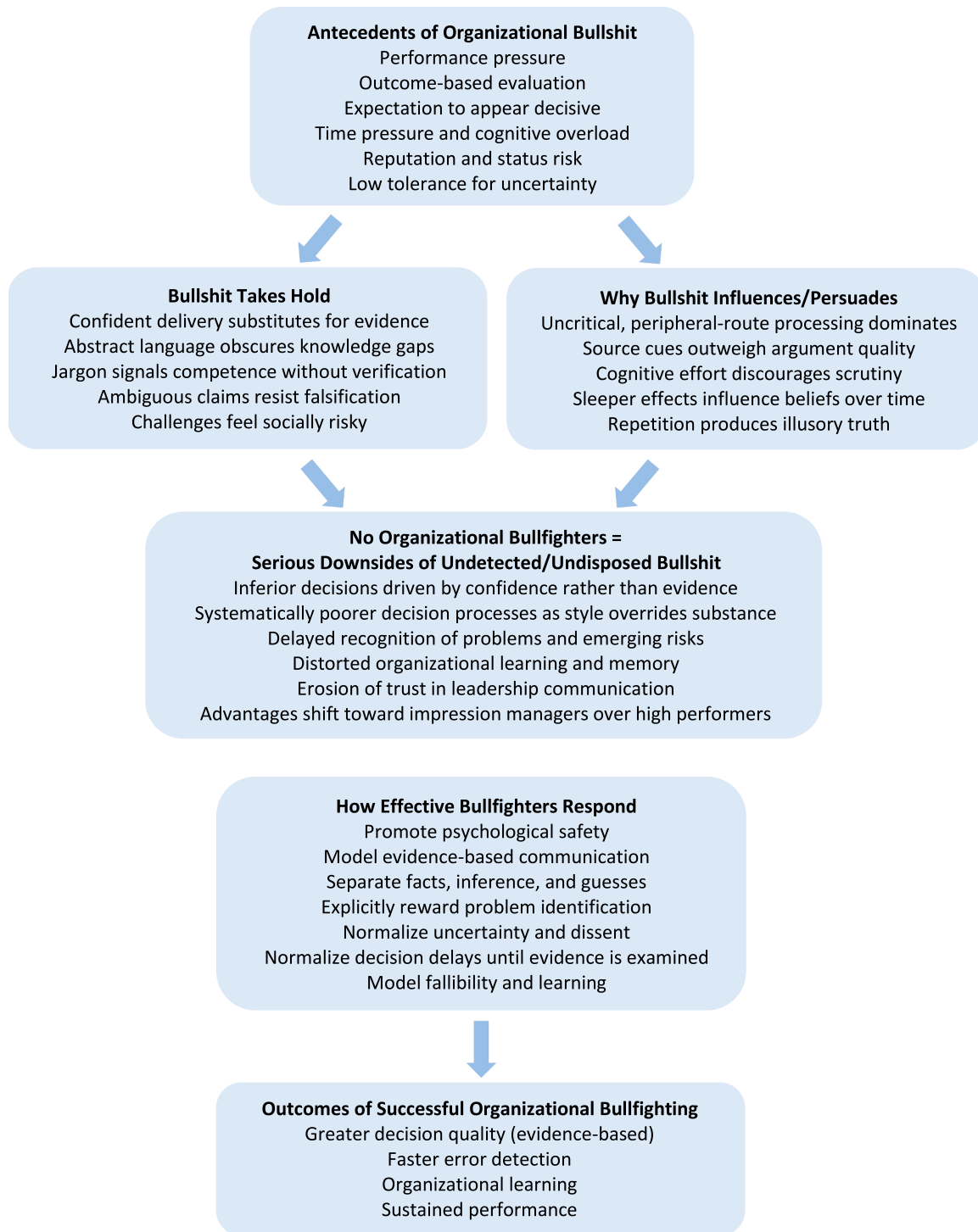


Fig. 1. Framework: From bullshit antecedents to bullfighting strategies of detection and disposal.

Action 2: Ask for problems directly (Weeks 2-4)

What it can look like in practice: The CEO adds the agenda item—“What problems should we discuss?”—and waits. Twelve seconds of silence. He fills it: “I think things are actually going pretty well.” They are not. Two people had planned to speak. They put their notes away.

What to do: Add a standing agenda item: “What problems should we discuss?” or “What’s keeping you up at night?” Sit quietly and wait for responses. Don’t fill the silence [37,41]. When someone identifies a problem, respond with “Who can help solve this?” instead of “Why did this happen?” [24,37].

How it can work: This action targets the voice and silence barrier—specifically, the futility beliefs and self-protective motives that lead employees to withhold concerns even when they possess relevant information. Many meetings focus primarily on successes and plans, while problems remain hidden because research indicates employees routinely withhold concerns from those in positions to act on them [23,41]—until they become crises. Explicitly asking for problems—and responding by problem-solving rather than blaming—may help establish candor as the organizational default. As with Action 1, the specific connection between reducing organizational silence and reducing bullshitting has not been directly tested; the linkage is theoretically derived from the

observation that when honest communication feels risky, employees may default to superficially plausible but evidence-free alternatives [7, 10].

What success may look like: By week 4, people should be bringing up issues without prompting. If they're not, you may need to be more explicit: "I know there are concerns. What am I not seeing?"

30-day checkpoint

After 4 weeks, ask yourself: 1) "Are people raising concerns they would have previously hidden?"; 2) "When problems surface, do we focus on solving them or finding someone to blame?"; and 3) "Have I responded defensively to any bad news?"—If yes, you've set the work back—acknowledge it and recommit. If the answer to the first two questions is yes, proceed to next actions. If no, continue with Actions 1–2 for another 2–4 weeks. Edmondson and Bransby [24] and Edmondson [37] find that the structural interventions introduced later are unlikely to take hold without this cultural foundation.

Action 3: Separate discussion from decision (Weeks 5-8)

What it can look like in practice: The CFO presents a market-entry proposal at 2:45. The decision is due at 3:00. A senior analyst sees three questionable assumptions in the revenue model but starts mentally drafting implementation next steps instead—the room is already nodding, and the clock makes scrutiny feel like obstruction.

What to do: For important decisions, split the meeting into two parts: Part 1 (first half): "We're only identifying concerns and questions. No defending, no deciding yet." Part 2 (second half): "Now let's decide, knowing what we've heard."

How it can work: This action targets the cognitive load and time-pressure pathway identified in Section 2.2 as a likely amplifier of peripheral route processing. When discussion and decision happen simultaneously, people may tend to rush to agreement. The elaboration likelihood model [25,26] predicts that time pressure and competing attentional demands sharply reduce the ability to evaluate argument quality, shifting listeners toward heuristic cues—exactly the conditions under which Petrocelli's [15] experimental findings suggest bullshit is most persuasive. Separating the evaluation and decision phases is designed to reduce this load, giving evidence more opportunity to surface before social momentum forecloses scrutiny. This specific link—between phase-separated meeting formats and reduced peripheral processing in organizational decisions—is theoretically derived and has not been directly tested in field settings.

What success may look like: Decisions may take longer initially but may result in fewer costly mistakes. Track what you decided and check 60 days later whether it was the right call—a practice consistent with evidence-based management principles that emphasize using organizational data to evaluate and redesign decisions over time [21].

Action 4: Change what you reward (Weeks 6-12)

What it can look like in practice: At her annual review, she is told she "demonstrates strong leadership presence." The colleague who spent the year saying "I'm not sure" and flagging problems early is told he "needs to work on executive confidence."

What to do: In your next round of performance conversations, explicitly praise someone who: 1) identified a problem early; 2) admitted they were wrong about something and changed course; or 3) said "I don't know" instead of making up an answer. Be specific: "You flagged the timeline issue in week 2, which saved us from a major customer problem. That's exactly what I need."

How it can work: This action targets the impression management incentive structure identified as a primary antecedent of bullshitting. Lab studies [14] show that people bullshit less when they anticipate accountability to a critical audience—demonstrating that the

calculation to bullshit is sensitive to perceived consequences. McCarthy and colleagues [7] and Spicer [10] report that bullshitting increases when admitting uncertainty threatens a person's career standing. Publicly rewarding honest communication is designed to shift that calculation—making candor the safer career choice. Whether restructuring performance recognition in this way reduces bullshitting specifically, as opposed to merely reducing silence or increasing voice more generally, has not been tested in organizational field studies and represents an important gap in the evidence base.

What success may look like: Within 2–3 months, people may begin volunteering "I'm not sure" or "I was wrong about that" without apparent fear. This can serve as a signal that the culture is beginning to shift [52].

60-day checkpoint

After 8 weeks, assess: 1) Do meetings regularly surface concerns that lead to better decisions?; 2) Have you publicly recognized honest, evidence-based reporting?; and 3) Are people who admit uncertainty being punished or rewarded? If meetings are working well, expand the format to other teams. If not, focus on your own responses—are you truly rewarding candor or still inadvertently punishing it?

Action 5: Require evidence for big claims (Weeks 9-12)

What it can look like in practice: "We're projecting 22% margin improvement in Q3." The slide is confident; the room moves on. No one asks what is driving the number. It is, in fact, an extrapolation of a single month's data from a pilot program that has since been discontinued.

What to do: For major decisions (budget commitments, strategic shifts, new initiatives), require presenters to distinguish: 1) What we know (based on data); 2) What we think (based on experience/judgment); and 3) What we're guessing (genuine uncertainty). Use this simple test: "What would have to be true for this to work?" If the answer is vague or circular, push back: "Can we test that assumption before committing?"

How it can work: This action targets the peripheral route processing dynamic identified in Finding 1: when evidence quality is never made the explicit subject of attention, bullshit-framed claims are evaluated on peripheral cues rather than substantive merit. Petrocelli's [15] experimental work shows that evidence-based framing—where communicators signal that evidence matters—maintains appropriate discrimination between strong and weak arguments, while bullshit framing eliminates it. Requiring presenters to label claims as data, inference, or assumption is designed to create the same functional effect structurally: making evidence quality the mandatory subject of evaluation rather than leaving it to individual motivation. McCarthy and colleagues [7], Spicer [10], and Kiazad and colleagues [13] all report that bullshit is more likely to proliferate when evidence quality goes unexamined. Asking "What exactly is the basis for this claim?" may help redirect evaluation from communicator style toward substantive content. The direct link between claim-labeling requirements and reduced peripheral processing in organizational meetings has yet to be empirically tested.

What success may look like: People start preemptively labeling their claims: "Here's what the data shows... here's where I'm inferring... here's where I'm honestly guessing." This transparency enables better evaluation of risk [55].

Action 6: Implement processing restoration protocols (Weeks 9-10)

What it can look like in practice: The consultant presents the restructuring case with fluency and confidence, citing three Fortune 500 analogs. The room is persuaded—not by the argument, which no one has evaluated, but by the delivery. No one can name what specific evidence supports the core claim. The consultant is not asked to provide any supporting data.

What to do: When a major claim is presented in any decision-making meeting, introduce a mandatory 90-second interruption before discussion proceeds. During this interruption, each participant independently writes answers to two questions: 1) “What specific evidence supports this claim?” and 2) “What would have to be false for this to be wrong?” Participants share answers before any evaluation or advocacy begins. Designate a rotating “evidence anchor”—a specific person in each meeting whose sole role is to interrupt when claims shift from evidence-labeled to assertion-based, using the prompt: “Can you tell us which category that falls in—data, inference, or assumption?”

How it can work: Psychological safety theory predicts that people will be more likely to speak up when it feels safe. The present action targets a separate, theoretically distinct problem: even when people are willing to speak, Petrocelli’s [15] experimental work suggests that how evidence-free communication is framed may shift listeners’ attention to peripheral route processing. The 90-second independent writing task is intended to interrupt this shift by structurally encouraging central route processing—directing attention to evidence quality before social dynamics and speaker confidence can redirect evaluation toward peripheral cues. The evidence anchor role is designed to externalize the scrutiny function so that it does not depend solely on any individual’s in-the-moment motivation to challenge authority. This is proposed as a structural countermeasure against peripheral route processing, distinct from a voice intervention. The mechanism linkage rests on Petrocelli’s [15] laboratory evidence. Whether the 90-second writing format and evidence anchor role produce equivalent effects in organizational meetings with real career stakes has not been directly tested, and the organizational efficacy of this specific protocol remains an open empirical question.

What success may look like: Presenters may begin preemptively structuring their own claims into evidence tiers (e.g., “Here’s the data, here’s my inference, here’s my assumption.”). Challenges to unsupported assertions may become more routine rather than interpersonally threatening. One indicator of progress is the ratio of evidence-labeled to assertion-labeled claims in meeting notes over time—the proportion of evidence-based claims may increase within 6–8 weeks of consistent application.

Context and conditions: This protocol is most important for irreversible, high-stakes decisions. Calibrate its formality to decision stakes—for routine operational choices, a lighter verbal version should suffice. In hierarchical cultures, the evidence anchor role may need to be assigned to someone with sufficient status to interrupt senior presenters credibly.

Action 7: Conduct narrative origin audits (Weeks 10-11)

What it can look like in practice: “We’re the most customer-centric company in our space” has appeared in every strategy deck for four years. When a new hire asks what data it is based on, the room goes quiet. No one can say. The claim predates everyone present. It has simply always been “true” and it has never been examined.

What to do: Identify the five strategic narratives your organization repeats most frequently—in town halls, strategy documents, investor communications, onboarding materials. For each narrative, convene a 60-minute “origin audit” session with a small team that includes at least one organizational newcomer (i.e., someone who joined within the past 12 months). Ask three questions about each narrative: 1) “What was the original evidence basis for this claim when it was first made?”; 2) “Has that evidence been tested against outcomes since?”; and 3) “Does this narrative still accurately describe current organizational reality, or has it drifted?”

Document what the audit reveals. If a narrative cannot be traced to a verifiable evidence origin, classify it as “asserted” rather than “established” in subsequent communications. If a narrative has not been tested against outcomes, schedule a structured reality-test. If a narrative has drifted from reality, create a formal process for updating or retiring it.

How it can work: Bolino and colleagues [28] explain why people may

resist admitting narratives are wrong—doing so threatens their organizational image and status. The narrative origin audit addresses a distinct, bullshit-specific problem that impression management theory was not designed to capture: even after incentives are restructured, narratives introduced before the cultural change may continue to accrue credibility over time through the sleeper effect [29]. Unlike lies, which tend to remain discounted once identified, bullshit-based narratives are only ambiguously wrong—the bullshitter did not know or care whether they were accurate. Such ambiguity may allow the discounting response to erode in organizational memory while the narrative content persists. The origin audit is designed to interrupt this process by reintroducing an evidence anchor to narratives that have been circulating long enough to potentially benefit from sleeper-effect credibility restoration. The sleeper effect mechanism is supported by Petrocelli and colleagues’ [29] controlled experimental findings; whether structured narrative audits interrupt this dynamic in organizational memory at the scale of strategic narratives has yet to be tested and constitutes a theoretically motivated but empirically unverified proposition.

The organizational newcomer inclusion is research-derived—newcomers have not been exposed to repeated iterations of the narrative and are therefore theoretically less susceptible to its illusory truth accumulation [30]. They may serve as relatively more reliable detectors of narratives that have drifted from reality but feel true to those with longer tenure.

What success may look like: A formal “narrative registry”—a living document classifying organizational narratives by evidence status (established, asserted, or outdated)—reviewed quarterly. Leaders begin publicly retiring outdated narratives rather than allowing them to circulate passively. New employees are specifically asked, within their first 90 days, to flag any organizational claims that seem inconsistent with their observations.

Context and conditions: This audit would likely be most valuable after organizational transitions—mergers, leadership changes, strategic pivots—when inherited narratives from the prior era are most likely to have been bullshit-based and are now accruing unwarranted credibility through repetition. In regulated industries, distinguish between legal communications (which may require specific language) and internal strategic narratives (which should be held to full evidence standards internally).

Action 8: Establish organizational inoculation protocols (Weeks 11-12)

What it can look like in practice: The 18% growth projection from last year’s plan was wrong by twelve points. This year’s plan opens with a new projection: 21% growth. No one references last year’s number. The slide moves quickly. It sounds about right—it always does by the third year you’ve heard it.

What to do: Identify the three organizational processes that most frequently repeat the same claims over time (quarterly status reviews, annual planning cycles, and standing committee reports are typical candidates). In each of these processes, introduce a “prior claim checkpoint”—a structured moment at which claims made in the previous cycle are retrieved and evaluated against what has since been learned. In quarterly reviews: begin each review by displaying the three most important claims made in the prior quarter’s review alongside what actually happened. In annual planning: open the planning cycle by reviewing last year’s projections against outcomes, before any new projections are introduced. In standing reports, require all status reports to include a “last period claim, this period evidence” column.

Additionally, introduce “inoculation briefings” before any major presentation of new strategic claims. These 5-minute briefings—delivered by the evidence anchor or a designated team member—explicitly name the type of communication about to occur (e.g., “This presentation will include both evidence-based findings and forward-looking projections that have yet to be tested.”) and prime the audience to maintain distinguishing attention.

How it can work: The illusory truth effect suggests that repeated organizational exposure to bullshit-based claims—even when originally recognized as uncertain—may progressively raise perceived validity, gradually degrading the collective threshold for what counts as “plausible” or “feasible” [30]. The prior claim checkpoint is designed to interrupt this process by reintroducing a reality-test at the moment of highest repetition, before the false familiarity of a claim can be mistaken for evidence. The inoculation briefing is intended to pre-activate critical processing, a technique that research suggests may reduce illusory truth effects by making the repetition-based fluency cue salient rather than unconscious [30].

This action targets the illusory truth mechanism identified in Finding 3: the experimental finding by Petrocelli and colleagues [30] that repeated exposure to bullshit-framed statements increases their perceived validity—even when initially recognized as bullshit—more than equivalent exposure to lie-framed statements. The prior claim checkpoint and inoculation briefing are designed to counteract the specific repetition channel through which this effect accumulates. This is not a generic “critical thinking” intervention but one theoretically derived from the illusory truth mechanism—a mechanism that, if it operates in organizational settings as the experimental literature suggests, would function at least partially independently of individual motivation, organizational culture, or psychological safety levels. The organizational efficacy of the specific prior claim checkpoint and inoculation briefing formats proposed here has not been tested; the intervention design is theoretically motivated by Petrocelli and colleagues’ [30] laboratory findings but requires field validation before its effectiveness in organizational planning and reporting cycles can be confirmed.

What success may look like: Leaders may begin opening planning cycles by routinely naming what prior projections were wrong and what was learned—not defensively, but as standard practice. Standing reports may evolve from performance-advocacy documents toward performance-reality documents. One indicator of progress is the ratio of optimistic projections to outcomes over rolling four-quarter periods—forecast accuracy may improve as inoculation protocols reduce social tolerance for uncalibrated projections.

Context and conditions: This protocol is especially important in fast-moving environments where bullshit is most likely to be produced under time pressure and where its illusory truth effects accumulate fastest. In cultures with high uncertainty avoidance, the prior claim checkpoint may initially feel threatening—frame it not as accountability for failure but as organizational learning from experience.

90-day review

After 12 weeks, evaluate your progress

Culture Check: 1) Can people raise concerns without fear? [24,37]; 2) Do meetings focus on solving problems or hiding them?; and 3) Are honest people advancing or only confident ones?

Decision Quality Check: 1) Review 2–3 major decisions made in the past 90 days; 2) Did they turn out as predicted?; and 3) What did we miss? Could earlier candor or more rigorous evidence scrutiny have surfaced it?

Next Steps: 1) If progress is solid, expand these practices to other teams and meetings’ or 2) If progress is limited, diagnose: Is leadership modeling insufficient? Are incentives still punishing honesty? Return to foundational actions.

What to expect

Weeks 1–4: People may test whether you mean it. Expect initial skepticism. How you respond to the first person who speaks up is likely to influence whether others will follow [37,39].

Weeks 5–8: Meetings may feel slower and more uncomfortable as people voice concerns they previously suppressed. This is progress, not a

problem [38,39,41].

Weeks 9–12: Truth-telling may begin to feel more normal rather than risky, and the structural protocols introduced in Actions 6–8 may begin producing detectable differences in how claims are framed and evaluated. If it doesn’t, assess whether incentives and responses are truly aligned with stated values.

Common mistakes

Asking for honesty but responding defensively when you get it. Research suggests that even a single defensive reaction can significantly set back weeks of trust-building [37,39]. If you catch yourself getting defensive, acknowledge it: “I reacted badly to that feedback yesterday. That’s on me. Please keep telling me what I need to hear.”

Implementing these practices only in your team while the broader organization punishes candor. You can’t control the whole organization, but you can create a pocket of safety in your sphere of influence. Be explicit with your team: “These are our norms, even if they’re different elsewhere.”

Giving up after 30 days because change feels slow. Arena and colleagues [52] find that meaningful culture change often requires 60–90 days or more to begin taking hold. Early discomfort is normal. Persistence matters more than perfection.

Treating the cultural foundation as the complete solution. Organizations frequently implement psychological safety initiatives, voice programs, or candor norms and conclude they have addressed the bullshit problem. Theory and experimental research suggest, however, that these address only its preconditions. Bullshit’s peripheral processing effects, sleeper effect dynamics, and illusory truth accumulation may operate at least partially independently of organizational culture and may continue to function even in high-safety, high-voice environments. The structural interventions in Actions 6–8 are not optional add-ons for mature cultures—they are the elements that make this playbook’s proposed contribution distinct from existing psychological safety or evidence-based management guidance.

Failing to apply the plausibility test. Ask whether claims can be replicated, verified, or logically make sense [55,56]. Before accepting any proposal, require answers to: “Can this be replicated?” “Do the numbers add up?” “Are required resources actually available?” Example: When a consultant promises “30% cost reduction in 90 days,” ask: “Which specific line items will decrease? What is the baseline? Has this worked in organizations of similar size and complexity?” If answers remain vague or redirect to testimonials, the claim fails the plausibility test.

Failing to apply the clarification test. Demand that all proposals be explainable in plain language to someone unfamiliar with the subject [7, 10,45]. Require strategic presentations be understandable to organizational newcomers. Ban jargon-heavy explanations that do not define operational terms. Example: If someone proposes “leveraging synergistic cross-functional integration to optimize stakeholder engagement,” ask them to restate it as: “We want marketing and sales to share customer data weekly so both teams can respond faster to complaints.” If they cannot translate jargon into concrete actions, the proposal lacks substance.

Failing to apply the evidence test. Before implementing any “best practice,” require specific evidence of effectiveness in similar contexts [21]. Require pilot testing rather than organization-wide rollouts. Create space for questioning established practices without being labeled disruptive. Example: Before adopting a new performance management system because “leading companies use it,” require answers to: “Which companies? What were their starting conditions? What metrics improved, by how much, over what timeframe?” Then pilot with one department for six months before broader implementation.

Failing to apply the humility test. Assess whether communicators acknowledge uncertainty, admit knowledge limitations, and demonstrate willingness to change positions based on evidence. Leaders who model intellectual humility create permission for honest assessment

throughout organizations [57]. Example: Compare two responses to a board question about market projections. Bullshit response: “Our analysis definitively shows 20% growth.” Credible response: “Our models suggest 15–25% growth depending on competitor response, and here are the three assumptions most likely to be wrong.” The second acknowledges uncertainty while still providing actionable guidance.

Adapting to your context

If you're not the senior leader: You can still implement these practices with your own team. Model candid communication in interactions with your boss: “I’m uncertain about X.” or “Here’s what could go wrong with Y.” Edmondson and Kerrissey [39] and Morrison [41] find this kind of modeled honesty may create permission for others to follow.

If your culture is highly hierarchical: Direct challenge of authority may remain inappropriate. Create alternative channels: anonymous input mechanisms, one-on-one conversations before group meetings, or asking specific people to voice concerns on behalf of others.

If you work remotely: Effective communication practices call for deliberate attention and active listening [42]. If working virtually, using chat for anonymous questions during presentations and scheduling brief one-on-ones may create space for concerns that might not otherwise surface in large video calls.

The integrated 90-day plan: What changes when

The three phases of this playbook are interdependent but distinct in what they address:

Weeks 1–8 (Foundation): Address why candor is organizationally difficult. Create the psychological safety, voice structures, and incentive realignment that make honest communication feel possible. Source: psychological safety, voice, and impression management literatures. Without this phase, the structural interventions in Weeks 9–12 may be perceived as surveillance rather than support.

Weeks 9–10 (Processing Restoration): Address what happens when bullshit enters the room. Interrupt the peripheral processing shift through structural evidence-anchoring that operates independent of individual motivation. Source: bullshit-persuasion experimental research. Without this, even psychologically safe teams remain vulnerable to well-packaged evidence-free claims.

Weeks 11–12 (Memory Inoculation): Address what happens to bullshit over time. Prevent bullshit-based narratives from regenerating credibility through sleeper effects and illusory truth accumulation. Without this, cultural improvements in Weeks 1–8 may erode as old bullshit narratives continue to circulate.

Boundaries and risks

Extreme power distance and authoritarian leadership

The interventions described here require leadership commitment to psychological safety [24,37]. In organizations where leaders view challenge as insubordination, or in national cultures with extreme power distance, attempting to introduce evidence-accountability norms from middle levels is unlikely to succeed and may place those who try at professional risk. In such contexts, focus on creating protected spaces (e.g., cross-functional working groups, innovation labs) where different norms apply, rather than attempting organization-wide transformation without top leadership support.

Regulatory or legal constraints on communication

Some industries (e.g., financial services, healthcare, defense) operate under regulatory constraints that make external transparency difficult or legally inadvisable [23]. What appears to be “bullshit” may reflect legal advice to avoid admissions that create liability. In these contexts, the

productive distinction is between internal candor (essential for organizational learning) and external communication (subject to legal constraints). Create privileged spaces for internal honesty while maintaining appropriate boundaries on external disclosure.

Contexts genuinely requiring inspirational rather than factual communication

As discussed in Section 2.4, not all organizational communication should be held to strict evidence standards [34–36]. Vision-setting, motivational communication, and aspirational goals serve legitimate functions. Aspirational communication becomes problematic not in its natural habitat of vision-setting, but when it drifts into operational decision-making and displaces the evidence those contexts require. Leaders are therefore encouraged to develop judgment about appropriate communication modes for different contexts.

Weaponization of “truth-telling” norms

Calling something “bullshit” is inherently confrontational. Badly implemented anti-bullshit cultures are likely to suppress several lines of communication and devolve into constant challenges, cynicism, and interpersonal conflict [15,56]. Carucci [46] notes that environments emphasizing candor can inadvertently activate threat responses, turning evidence-accountability norms into instruments of interpersonal conflict rather than organizational improvement. Mitigate this risk by focusing on evidence quality, not personal attacks: “I don’t see evidence for that claim.” rather than “You’re bullshitting.” Establish norms that critique ideas, not people, and that treat uncertainty as natural rather than shameful.

Analysis paralysis

Demanding evidence for every claim—no matter how small—can slow decision-making to a crawl [20,58]. Not every decision warrants extensive evidence documentation. Organizations are therefore encouraged to calibrate evidence standards to decision stakes and urgency. For routine, low-stakes, reversible decisions, reasonable judgment calls may be appropriate. Reserving rigorous evidence protocols for high-stakes, irreversible commitments—and creating explicit categories (reversible vs. irreversible, high-stakes vs. low-stakes) to guide appropriate evidence thresholds—may help avoid this failure mode.

Premature optimization

In genuinely uncertain environments—launching new products, entering new markets, navigating technological disruption—extensive evidence may not be available. Waiting for definitive proof before acting risks missing opportunities. When the potential consequences of missed opportunities appear to exceed the risks of acting without definitive proof, a productive shift may be to move from demanding evidence to demanding explicit acknowledgment of uncertainty and structured learning plans: “We don’t know if this will work, here’s what we need to learn, here’s how we’ll know if we’re wrong.” This approach may acknowledge intellectual humility and preserve intellectual honesty without paralyzing action.

Cultural mismatch

Truth-telling norms may conflict with national or organizational cultures that prioritize harmony, face-saving, or indirect communication. Implementing Western-style “radical candor” in high-context Asian cultures, for example, may violate deeply held norms about appropriate communication. Adapt interventions to cultural context by creating anonymous feedback channels, using indirect communication protocols, and ensuring that challenges to ideas don’t threaten interpersonal

relationships. The goal is reducing evidence-free decision-making, not importing specific cultural styles of communication.

Rather than organization-wide rollout, piloting anti-bullshit interventions in a single team or division is advisable. Arena and colleagues [52] find that behaviors spread contagiously through networks and that cultural transmission is highly context-dependent—what works in one pocket of an organization may not transmit effectively in another. Piloting before scaling allows practitioners to learn what works in their specific context and identify unexpected failure modes early. Scaling interventions that show concrete benefits in a given organizational context is preferable to broad adoption based on theoretical promise alone.

Conclusion and agenda

Key takeaways

Organizations face a strategic imperative: in an increasingly complex world, clarity may constitute competitive advantage. The research reviewed here suggests that bullshit—communication delivered without regard for truth, evidence, or established knowledge—may pose a more serious threat to organizational effectiveness than commonly recognized. Theory and controlled experimental evidence indicate that bullshit may bypass critical evaluation, produce persistent belief distortions, and gradually degrade collective capacity for sound judgment.

The Ford case illustrates conditions under which systematic reform toward evidence-based communication may produce meaningful competitive benefits, though field-based validation of these propositions remains an important priority. Theory and illustrative evidence suggest that organizations creating environments where candor is rewarded over impression management may gain advantages in decision quality, talent attraction, and crisis resilience. The Managerial Playbook presented here offers theoretically informed implementation frameworks derived from this evidence base. Leaders who invest in systematic processes for identifying and reducing bullshit, while building cultures that reward evidence-based reasoning, may develop clarity advantages that compound over time—though the interventions proposed here should be treated as theoretically grounded starting points subject to organizational learning and adaptation rather than as proven prescriptions.

Research agenda

The lab-to-field gap is the central limitation here, and closing it is the most important research priority going forward. The mechanisms identified here—peripheral route processing under bullshit framing, sleeper effect dynamics, and illusory truth amplification—are theoretically coherent and internally replicated within a single productive research program. What the literature currently lacks is field-based evidence that these mechanisms operate in real organizational settings with experienced professionals, high stakes, complex information environments, and organizational politics. The following research agenda is organized to guide both applied researchers seeking to design validating studies and practitioners seeking to understand what evidence they should look for before investing in these interventions.

Outcome variables: What should researchers measure?

The most important question for field validation is deceptively simple: how should researchers determine whether anti-bullshit interventions actually improve organizational functioning? Decision quality is the obvious candidate, but its operationalization requires careful thought, because organizations naturally evaluate decisions by their outcomes—precisely the outcome bias that Section 1 identifies as a root cause of bullshit in the first place. We do know that employees higher in corporate bullshit receptivity performed more poorly on a validated situational judgment test measuring work-related decision-making, an assessment approach with well-established links to actual

job performance [9]. This matters because organizations rely on sound judgment at every level—from frontline problem-solving to leadership decisions—and a workforce that is more susceptible to corporate bullshit may be systematically disadvantaged when it comes to choosing the best course of action in realistic, high-stakes workplace scenarios. However, measuring decision quality by outcomes alone would reproduce the problem the interventions are designed to solve.

A more defensible operationalization distinguishes *process quality* from *outcome quality* and tracks both over time. Process quality can be assessed through behavioral observation of meeting communication—specifically, the ratio of evidence-labeled claims (data, inference, stated assumption) to bare assertions in decision-relevant conversations. Meeting transcripts or structured observer ratings can provide this measure at low cost, and the processing restoration protocols introduced in Section 4.8 create a natural data collection structure: the evidence anchor role generates a running log of claim types that can be aggregated across meetings and compared pre- versus post-intervention. This operationalization is process-focused, prospective, and independent of outcome luck—the exact properties needed to distinguish genuine decision quality from the outcome bias [1,2] and extended in organizational contexts.

Beyond process quality, researchers should track a portfolio of subsidiary outcomes, each corresponding to a specific mechanism identified in this manuscript. *Speed of problem detection*—the elapsed time between a problem’s emergence and its first documentation in an organizational record—directly tests whether reduction of bullshit translates into earlier surfacing of operationally consequential issues. This can be measured through archival analysis of incident reports, project post-mortems, or escalation logs, comparing detection lag in treated versus control units. *Accuracy of resource allocation*—the degree to which resource commitments made in planning processes correspond to subsequent operational realities—tests whether bullshit-based projections are reduced; it can be operationalized as the mean absolute deviation between planned and realized resource utilization over rolling four-quarter periods. *Post-decision regret and revision rates*—how often organizations reverse or significantly modify consequential decisions within 90 days—provide a proxy for the proportion of decisions made on bullshit-based premises.

Research designs: Units of analysis, time horizons, and implementation fidelity

The appropriate unit of analysis for field studies of anti-bullshit interventions is the *team* or *department*, not the organization. Organizational-level assignment is typically infeasible, and the mechanisms described here—peripheral processing shifts, meeting communication patterns, narrative circulation—operate most proximally at the team level. Team-level assignment also aligns with the playbook’s instruction to pilot in a single team before scaling, creating a natural experimental structure in organizations already implementing the interventions.

The Managerial Playbook’s 90-day timeline represents a realistic minimum for cultural foundation-building, but is likely too short to detect meaningful shifts in the outcome variables described above. Researchers should plan for a minimum 12-month observation window post-intervention, with pre/post measurement at baseline, 90 days, and 12 months. A 6-month measurement point is also warranted to assess whether early gains in meeting communication patterns—which may be detectable within the 90-day window—persist or decay as organizational attention shifts to other priorities.

Implementation fidelity—the degree to which organizations actually execute the prescribed interventions rather than a nominal approximation—is a critical and underappreciated measurement challenge in organizational intervention research [59]. It should be assessed through at least three methods used in combination. *Behavioral observation* of a random sample of meetings in treated units, rated against a fidelity checklist corresponding to Actions 1–8, provides the most direct

measure. *Archival analysis of meeting records*—agenda structures, documented claim categorizations, evidence anchor role assignments—provides scalable fidelity data that does not require observer presence. *Survey measures* of psychological safety [37], voice behavior [41], and impression management orientation [28]—administered to participants in treated and control units—provide individual-level data that can be used both to verify that cultural foundation shifts occurred and to test whether those shifts mediate intervention effects on decision quality outcomes. Without implementation fidelity measurement, null results cannot be interpreted, as they might reflect genuine absence of effect, or they might reflect failure to implement the protocol with sufficient consistency to constitute a valid test.

Boundary-condition tests: Framing section 5 as testable hypotheses

Section 5 identifies several conditions under which the interventions may struggle. Yet, these are not merely caveats—they are testable hypotheses, several of which are both practically important and methodologically accessible. The three most important to test first, in rough order of priority, are as follows.

Power distance as a moderator of intervention effectiveness. The claim in **Section 5.1** that extreme power distance undermines the interventions generates a specific, testable prediction: the effect of processing restoration protocols (Action 6) on the ratio of evidence-labeled to bare-assertion claims should be weaker in organizational units with greater power distance. Power distance can be operationalized at the team level using established survey measures, creating a continuous moderator that can be tested without restricting the sample to extreme cases. This is the highest-priority boundary condition to test because it affects the broadest range of potential implementation contexts and because the psychological safety literature already provides a rich theoretical framework and measurement infrastructure for this moderation [24,38]. Morrison [23] and Edmondson and Bransby [24] find that greater power distance is negatively associated with psychological safety at both individual and group levels, and that this relationship is mediated by employees' reduced willingness to engage in upward communication—exactly the communicative behavior the interventions depend on.

Incentive structure as a moderator of cultural foundation effects. **Section 5** and the impression management literature [28] jointly suggest that psychological safety interventions may be ineffective—or actively counterproductive—when career rewards remain tied to confident performance presentation rather than accuracy. It is reasonable to hypothesize the effect of Weeks 1–8 cultural foundation interventions on voice behavior and problem detection speed will be moderated by the degree to which organizational performance evaluation criteria explicitly reward accuracy over impression management. This can be operationalized through archival analysis of performance review criteria or survey measures of employees' perceptions of what behaviors are actually rewarded in promotion and evaluation decisions, compared to stated organizational values—a gap analogous to what Argyris [45] calls the difference between espoused theories and theories-in-use, extended here to the organizational level.

Regulatory and legal constraint contexts. **Section 5.2** notes that some industries operate under communication constraints that make external transparency legally inadvisable. A comparative study across regulated industries (financial services, healthcare, defense) and less regulated contexts could test whether the interventions produce stronger effects on internal decision-making communication when external communication is legally constrained. The most reasonable prediction is that the gap between internal and external communication candor should be detectably larger in regulated contexts prior to intervention, and that the interventions should narrow this gap in internal settings while leaving external communication patterns unchanged. Such a design would also clarify a conceptual ambiguity in the current framework; that is, whether what appears to be organizational bullshit in regulated industries sometimes reflects legal constraint rather than epistemic indifference, and whether that distinction affects the interventions'

mechanism of action.

Field evidence priorities: The single most important study

The single study that would most advance this research program is a quasi-experimental field study using a *waitlist control design*—in which participating organizations are randomly assigned either to receive the 90-day playbook immediately or to a six-month waiting period before receiving it—tracking a defined set of outcome variables across a minimum 12-month observation window.

The study's target population should be intact work teams (6–15 members) with meaningful decision authority, drawn from organizations in sectors where decision quality can be operationalized with reasonable objectivity (e.g., financial services/resource allocation accuracy, healthcare/preventable incident rates, technology product development/projection accuracy or revision rates, or management consulting/client engagement outcomes). A sample of 40–60 teams would provide adequate statistical power to detect medium-effect moderation while remaining feasible for a multi-site consortium design.

Pre-intervention baseline measurement should include: 1) a structured behavioral observation of two or three representative meetings per team, coded for evidence-labeled versus bare-assertion claims; 2) team-level survey measures of psychological safety [37], voice behavior [41], and impression management orientation [28]; and 3) archival extraction of a 12-month historical baseline on the primary outcome variables (resource allocation accuracy, problem detection lag, decision revision rates). Post-intervention measurement should repeat all three data collection methods at 90 days and 12 months.

The primary test should be whether teams receiving the playbook show greater improvement than waitlist controls on the composite outcome index from pre- to 12-month follow-up. Fidelity should be monitored through monthly check-ins, meeting observation, and archival review, and used as a continuous moderator in secondary analyses to distinguish effective implementation from nominal adoption. The boundary condition tests described in Section 6.2.3 could be embedded within this design at minimal additional cost by including power distance, incentive structure, and regulatory environment as measured moderators at enrollment.

A positive result from this study—demonstrating that teams implementing the playbook show measurable improvements in process-quality metrics and at least one substantive outcome variable relative to waitlist controls—would constitute the first field-based evidence for the organizational relevance of the bullshit persuasion [15], sleeper effect [29], and illusory truth mechanisms [30], and would provide the empirical foundation for the prescriptive claims that currently rest on theoretical inference and illustrative cases. Conveniently, even a null result would be equally informative, as it could identify whether the failure lies at the level of cultural foundation building, structural protocol adoption, fidelity of implementation, or the mechanisms themselves. Either outcome would represent a substantial advance over the current evidence base and would justify the research investment the current framework asks practitioners to make.

Executive summary

Organizations face an underappreciated threat in the realm of strategic decision-making and accountability—*bullshit*—communication delivered with indifference to truth, evidence, or established knowledge. Unlike lies, bullshit is judged leniently despite research showing it can be more persuasive and persistent than outright deception in contexts where factual accuracy and evidence-based reasoning are critical. Experimental research in controlled settings suggests mechanisms through which bullshit may produce insidious effects on organizational decision-making and informs a practical playbook derived from theoretically consistent organizational cases. Theoretically grounded propositions include: (1) bullshit may trigger peripheral processing that bypasses critical evaluation; (2) bullshit may produce stronger sleeper

effects than lies; (3) systematic, truth-focused cultures may create competitive advantages. The managerial playbook provides a framework for implementing reality-testing protocols and reforming meeting practices—with a 30–60–90 day implementation timeline for building bullshit-resistant organizations.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author used Grammarly and Claude in order to check and polish grammar, as well as to convert all APA-style in-text citations and references to the sequential numbering-style. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Declaration of Competing Interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

- J. Baron, J.C. Hershey, Outcome bias in decision evaluation, *J. Pers. Soc. Psychol.* 54 (1988) 569–579, <https://doi.org/10.1037/0022-3514.54.4.569>.
- C.E. Seta, J.J. Seta, J.V. Petrocelli, M. McCormick, Even better than the real thing: alternative outcome bias affects decision judgments and decision regret, *Think. Reason 21* (4) (2015) 446–472, <https://doi.org/10.1080/13546783.2015.1034779>.
- H.G. Frankfurt, On bullshit, *Raritan Q. Rev.* 6 (1986) 81–100, <https://doi.org/10.1017/CBO9780511818172.011>.
- C. Ferreira, D. Hannah, I. McCarthy, L. Pitt, S.L. Ferguson, This place is full of it: towards an organizational bullshit perception scale, *Psychol. Rep.* 125 (2022) 448–463, <https://doi.org/10.1177/0033294120978162>.
- W.G. Obenauer, A commentary on Kiazad et al. (2026): Calling bullshit on bullshitting, *Group Organ. Manag.* <https://doi.org/10.1177/10596011251403031>.
- J.V. Petrocelli, K. Kiazad, S.L.D. Restubog, F.X. Chen, A reply to Obenauer (in press): On the empirical distinctions between bullshitting and lying, *Group Organ. Manag.* (in press). <https://doi.org/10.1177/10596011251415209>.
- I.P. McCarthy, D. Hannah, L.F. Pitt, J.M. McCarthy, Confronting indifference toward truth: dealing with workplace bullshit, *Bus. Horiz.* 63 (2020) 253–263, <https://doi.org/10.1016/j.bushor.2020.01.001>.
- M. Fallatah, This is bullshit: The relationship between organizational bullshitting and employee job satisfaction, *Soc. Sci.* 12 (2023) 636, <https://doi.org/10.3390/socsci12110636>.
- S. Littrell, The corporate bullshit receptivity scale: development, validation, and associations with workplace outcomes, *Pers. Individ. Differ.* 255 (2026) 113699, <https://doi.org/10.1016/j.paid.2026.113699>.
- A. Spicer, Shooting the shit: the role of bullshit in organisations, *M@N@gement* 16 (2013) 653–666, <https://doi.org/10.3917/mana.165.0653>.
- A. Spicer, Playing the bullshit game: How empty and misleading communication takes over organizations, *Organ. Theory* 1 (2020) 1–26, <https://doi.org/10.1177/2631787720929704>.
- B. Fugere, C. Hardaway, J. Warshawsky, *Why Business People Speak Like Idiots: A Bullfighter's Guide*, Free Press, New York, 2005.
- K. Kiazad, F.X. Chen, S.L.D. Restubog, Workplace bullshitting: a multidimensional analysis of motives, *Group Organ. Manag.* (in press). <https://doi.org/10.1177/10596011251348757>.
- J.V. Petrocelli, Antecedents of bullshitting, *J. Exp. Soc. Psychol.* 76 (2018) 249–258, <https://doi.org/10.1016/j.jesp.2018.03.004>.
- J.V. Petrocelli, Bullshitting and persuasion: the persuasiveness of a disregard for the truth, *Br. J. Soc. Psychol.* 60 (4) (2021) 1464–1483, <https://doi.org/10.1111/bjso.12453>.
- J.V. Petrocelli, H.E. Silverman, S.X. Shang, Social perception and influence of lies vs. bullshit: a test of the insidious bullshit hypothesis, *Curr. Psychol.* 42 (2023) 9609–9617, <https://doi.org/10.1007/s12144-021-02243-z>.
- G. Pennycook, J.A. Cheyne, N. Barr, D.J. Koehler, J.A. Fugelsang, On the reception and detection of pseudo-profound bullshit, *Judgm. Decis. Mak.* 10 (6) (2015) 549–563, <https://doi.org/10.1017/S1930297500006999>.
- J.V. Petrocelli, J.M. Curran, L.M. Stall, Bullshit can be harmful to your health: bullbility as a precursor to poor decision making (Article), *Curr. Opin. Psychol.* 55 (2024) 101769, <https://doi.org/10.1016/j.copsyc.2023.101769>.
- J.V. Petrocelli, Y. Li, E. Wang, J.M. Curran, Bullshitting and bullbility: conditions and consequences, *Soc. Psychol.* 55 (5) (2024) 262–279, <https://doi.org/10.1027/1864-9335/a000561>.
- E.A. Locke, D.M. Schweiger, G.P. Latham, Participation in decision making: when should it be used? *Organ. Dyn.* 14 (3) (1986) 65–79, [https://doi.org/10.1016/0090-2616\(86\)90032-X](https://doi.org/10.1016/0090-2616(86)90032-X).
- D.M. Rousseau, S. ten Have, Evidence-based change management, *Organ. Dyn.* 51 (3) (2022) 100899, <https://doi.org/10.1016/j.orgdyn.2022.100899>.
- L.K. Fazio, N.M. Brashier, B.K. Payne, E.J. Marsh, Knowledge does not protect against illusory truth, *J. Exp. Psychol. Gen.* 144 (5) (2015) 993–1002, <https://doi.org/10.1037/xge0000098>.
- E.W. Morrison, Employee voice and silence, *Annu. Rev. Organ. Psychol. Organ. Behav.* 1 (2014) 173–197, <https://doi.org/10.1146/annurev-orgpsych-031413-091328>.
- A.C. Edmondson, J.M. Bransby, Psychological safety comes of age: observed themes in an established literature, *Annu. Rev. Organ. Psychol. Organ. Behav.* 10 (2023) 55–78, <https://doi.org/10.1146/annurev-orgpsych-120920-055217>.
- R.E. Petty, J.T. Cacioppo, The effects of involvement on responses to argument quantity and quality: central and peripheral routes to persuasion, *J. Pers. Soc. Psychol.* 46 (1984) 69–81, <https://doi.org/10.1037/0022-3514.46.1.69>.
- R.E. Petty, J.T. Cacioppo, *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*, Springer-Verlag, New York, 1986.
- D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York, 2011.
- M.C. Bolino, K.M. Kacmar, W.H. Turnley, J.B. Gilstrap, A multi-level review of impression management motives and behaviors, *J. Manag.* 34 (6) (2008) 1080–1109, <https://doi.org/10.1177/0149206308324325>.
- J.V. Petrocelli, C.E. Seta, J.J. Seta, Lies and bullshit: the negative effects of misinformation grow stronger over time, *Appl. Cogn. Psychol.* 37 (2023) 409–418, <https://doi.org/10.1002/acp.4043>.
- J.V. Petrocelli, E.N. Rice, J.M. Curran, When fools have greater influence than liars: Testing the insidious bullshit hypothesis with the illusory truth effect, *Curr. Psychol.* (in press).
- H. Markopolos, *No One Would Listen: A True Financial Thriller*, Wiley, Hoboken, 2010.
- S. Littrell, J.A. Fugelsang, Bullshit blind spots: the roles of miscalibration and information processing in bullshit detection, *Think. Reason* 30 (1) (2024) 49–78, <https://doi.org/10.1080/13546783.2023.2189163>.
- R.E. Petty, P. Briñol, Persuasion: from single to multiple to metacognitive processes, *Perspect. Psychol. Sci.* 3 (2) (2008) 137–147, <https://doi.org/10.1111/j.1745-6916.2008.00071.x>.
- J.A. Sillince, P. Jarzabkowski, D. Shaw, Shaping strategic action through the rhetorical construction and exploitation of ambiguity, *Organ. Sci.* 23 (3) (2012) 630–650, <https://doi.org/10.1287/orsc.1110.0670>.
- L. Simeone, Using strategic ambiguity as management practice in academic R&D: an ethnographic study of MIT SENSEable City Lab, *R. & D. Manag.* 47 (2) (2017) 288–298, <https://doi.org/10.1111/radm.12229>.
- L.T. Christensen, M. Morsing, O. Thyssen, CSR as aspirational talk, *Organization* 20 (3) (2013) 372–393, <https://doi.org/10.1177/1350508413478310>.
- A. Edmondson, Psychological safety and learning behavior in work teams, *Adm. Sci. Q.* 44 (2) (1999) 350–383, <https://doi.org/10.2307/2666999>.
- A.C. Edmondson, Z. Lei, Psychological safety: The history, renaissance, and future of an interpersonal construct, *Annu. Rev. Organ. Psychol. Organ. Behav.* 1 (2014) 23–43, <https://doi.org/10.1146/annurev-orgpsych-031413-091305>.
- A.C. Edmondson, M. Kerrissey, What people get wrong about psychological safety, *Harv. Bus. Rev.* 103 (3) (2025) 52–59.
- C. Santana, L. Monte, B. Santana de Araújo Silva, G. Carneiro, S. Freire, J. Santos, M. Mendonça, Psychological safety in software teams: a systematic review, *Inf. Softw. Technol.* 187 (2025) 107838, <https://doi.org/10.1016/j.infsof.2025.107838>.
- E.W. Morrison, Employee voice behavior: Integration and directions for future research, *Acad. Manag. Ann.* 5 (1) (2011) 373–412, <https://doi.org/10.5465/19416520.2011.574506>.
- A.N. Kluger, G. Itzhakov, The power of listening at work, *Annu. Rev. Organ. Psychol. Organ. Behav.* 9 (2022) 121–146.
- H.P. Sims, D.A. Gioia, C.O. Longenecker, Behind the mask: the politics of employee appraisal, *Acad. Manag. Exec.* 1 (3) (1987) 183–193, <https://doi.org/10.5465/AME.1987.4275731>.
- D.T. Miller, M. Ross, Self-serving biases in the attribution of causality: fact or fiction? *Psychol. Bull.* 82 (2) (1975) 213–225.
- C. Argyris, Teaching smart people how to learn, *Harv. Bus. Rev.* 69 (3) (1991) 99–109.
- R. Carucci, 6 defensive behaviors that show up at work—and how psychological safety can help, *Harv. Bus. Rev. Digit. Artic.* (2025) 1–9. (<https://hbr.org/2025/10/6-defensive-behaviors-that-show-up-at-work-and-how-psychological-safety-can-help>).
- J.M. Priester, R.E. Petty, Source attributions and persuasion: perceived honesty as a determinant of message scrutiny, *Pers. Soc. Psychol. Bull.* 21 (1995) 637–654, <https://doi.org/10.1177/0146167295216010>.
- J.R. Priester, R.E. Petty, The influence of spokesperson trustworthiness on message elaboration, attitude strength, and advertising effectiveness, *J. Consum. Psychol.* 13 (2003) 408–421, https://doi.org/10.1207/S15327663JCP1304_08.
- G.P. Huber, Organizational learning: The contributing processes and the literatures, *Organ. Sci.* 2 (1) (1991) 88–115, <https://doi.org/10.1287/orsc.2.1.88>.
- B.G. Hoffman, *American Icon: Alan Mulally and the Fight to Save Ford Motor Company*, Crown Business, New York, 2012.
- A. Waytz, Beware a culture of busyness, *Harv. Bus. Rev.* 101 (2) (2023) 58–67.

- [52] M. Arena, S. Hines, J. Golden III, The three Cs for cultivating organizational culture in a hybrid world, *Organ. Dyn.* 52 (1) (2023) 100958, <https://doi.org/10.1016/j.orgdyn.2023.100958>.
- [53] U.S. Senate Committee on Banking, Housing, and Urban Affairs, An examination of Wells Fargo's unauthorized accounts and the regulatory response, 114th Congress, 2016. (<https://www.govinfo.gov/content/pkg/CHRG-114shrg23001/html/CHRG-114shrg23001.htm>).
- [54] Tayan, B., 2019. The Wells Fargo cross-selling scandal, Stanford Closer Look Series, Stanford Graduate School of Business, 2019. <https://www.gsb.stanford.edu/faculty-research/publications/wells-fargo-cross-selling-scandal>.
- [55] D. Kahneman, D. Lovallo, O. Sibony, Before you make that big decision, *Harv. Bus. Rev.* 89 (6) (2011) 50–60.
- [56] J.V. Petrocelli, *The Life-Changing Science of Detecting Bullshit*, St. Martin's Press, New York, 2021.
- [57] A. Grant, *Think Again: The Power of Knowing What You Don't Know*, Viking, New York, 2021.
- [58] J. Pfeffer, R.I. Sutton, Evidence-based management, *Harv. Bus. Rev.* 84 (1) (2006) 62–74.
- [59] C. Carroll, M. Patterson, S. Wood, A. Booth, J. Rick, S. Balain, A conceptual framework for implementation fidelity, *Implement. Sci.* 2 (2007) 40, <https://doi.org/10.1186/1748-5908-2-40>.