

Stress-Testing Dialectical Navigation: A Graded Comparative Analysis of Aviation Navigation Episodes

Abstract

Safety-critical aviation requires crews to navigate mismatches between standing procedures and unfolding conditions under severe time pressure. Yet accident analysis and adaptation research still under-explain how coordination is maintained, how it is substantially strained yet recovered, or how it is lost as episodes unfold. Building on Dialectical Navigation, this paper conducts a graded empirical stress test by comparing three commercial aviation navigation episodes that differ in coherence trajectory. Using official accident and incident investigation reports, each episode is analysed through three coordination mechanisms: guided elasticity, collective calibration, and mutual recovery. The comparison across maintained, strained-but-recovered, and lost episodes shows that outcome variation depends less on the mere presence of procedures, CRM routines, or technical cues than on when mismatch becomes operationally explicit, whether procedural flexing remains bounded, and whether recovery begins while a viable trajectory still remains. The analysis refines Dialectical Navigation by showing that dynamic coherence is best understood as a trajectory rather than a threshold state, identifying collective calibration as the pivotal hinge between bounded adaptation and a recoverable outcome, and clarifying boundary conditions, including cue ambiguity, authority gradients, monitoring-role overload, and shrinking recoverability windows. The paper concludes with implications for procedure design, communication architecture, monitoring and assertion, and feed-forward learning in safety-critical operations and practice.

1. Introduction

Safety-critical aviation episodes are often explained retrospectively in terms of technical malfunction, human error, or failures of crew resource management. These explanations identify important contributing factors, but they do not fully explain how crews move through the episode itself: when procedures cease to fit, how adaptation remains bounded or begins to drift, how coordination becomes explicit, and why recovery is timely in some cases and too late in others. Research on high-risk work shows that uncertainty, tight coupling, and shifting operational margins make reliable performance irreducible to procedural compliance or post hoc reconstruction of failure alone (Reason, 2008; Leveson, 2011; Hollnagel, 2014).

That gap is especially important in aviation because crews rarely confront a simple choice between following procedure and abandoning it. More often, they must preserve the coordinating intent of standard operating procedures while also interpreting ambiguous cues, redistributing workload, and re-anchoring action as the situation changes. The central question is therefore processual: how do teams remain safely coordinated when the standing script is no longer enough but fully improvised action would be equally dangerous?

The companion conceptual paper addressed this gap by proposing Dialectical Navigation as a process theory of safety effectiveness in motion, centred on dynamic coherence and three interdependent coordination mechanisms: guided elasticity, collective calibration, and mutual recovery (Author, under review). Safety is neither guaranteed by standardisation nor rescued by adaptation alone; it depends on how teams navigate the tension between procedural anchoring and adaptive responsiveness while action unfolds (Author, under review; Farjoun, 2010; Smith & Lewis, 2011). This paper extends that conceptual companion by empirically stress-testing Dialectical Navigation across real flight episodes that differ in how coordination is maintained, strained, or lost. For management scholarship, this is also a design problem: how organisations structure procedural governance, communication systems, authority relations, and learning processes so that frontline adaptation remains visible and recoverable under pressure (Bigley & Roberts, 2001; Faraj & Xiao, 2006; Grote, 2012).

A binary contrast between “success” and “failure” obscures important variation. Safety-critical episodes often follow graded trajectories: coherence may be maintained, substantially strained but recovered, or lost before crews can re-anchor safely (Author, under review). A graded comparison therefore better captures coordination under high-reliability conditions. Process research shows that teams under pressure move through shifting forms of interpretation, authority, and coordination within the same episode rather than occupying static states (Weick, 1995; Faraj & Xiao, 2006; Lei et al., 2016). Practically, organisations manage not only routine success and catastrophe, but the large middle ground in which safety depends on strained and often late coordination.

What this misses empirically is mechanism timing. Crews may notice mismatch and voice concern yet still lose coherence if recalibration occurs too late or without sufficient authority to redirect the trajectory. They may recover, but only within a sharply narrowed window that reveals substantial prior degradation rather than robust control. From a managerial standpoint, these distinctions matter because procedure design, briefing architecture, monitoring expectations, and escalation norms are all judged differently when the problem is not whether coordination occurred at all, but whether it remained bounded and recoverable in time. A framework that can distinguish maintained coherence from strained-but-recovered coherence and both from lost coherence therefore offers stronger diagnostic value than one that only separates safe outcomes from accidents.

Accident and incident analyses typically identify what went wrong but under-specify how crews navigate procedural-adaptive tension as events unfold. They rarely distinguish between bounded procedural flexing, late but sufficient recalibration, and recovery initiated only after the viable trajectory has narrowed beyond safe restoration. What remains underspecified is the process by which dynamic coherence is maintained, strained, or lost in motion. The broader literature reinforces this gap. Research on rules and procedural governance clarifies why formalisation is essential yet incomplete; HRO research explains the importance of sensitivity to operations and deference to expertise; CRM research identifies communication, monitoring, and

workload routines; sensemaking research explains interpretation under equivocality; and resilience perspectives foreground adaptation and recovery (Hale & Borys, 2013; Weick & Sutcliffe, 2007; Flin et al., 2008; Weick et al., 2005; Hollnagel, 2014; Woods, 2015). Yet these traditions are often stronger at identifying components, capabilities, or enabling conditions than at explaining the transition work through which teams move between procedural anchoring and adaptive response while action continues (Faraj & Xiao, 2006; Lei et al., 2016). What is therefore missing is not evidence that procedures, adaptation, communication, and recovery all matter, but a process account of how they are coupled within the same navigation episode and how variation in that coupling shapes whether coherence is preserved, strained, or lost.

To address this gap, the paper conducts a graded comparative analysis of three commercial aviation navigation episodes drawn from the same broad operational domain of multicrew, highly proceduralised, time-pressured flight operations. The cases vary in recovery quality while sharing the core conditions under which Dialectical Navigation should be most visible: consequential procedural-adaptive tension, rich documentary traces, and high-consequence coordination demands. Together, they form a maintained-coherence benchmark, a strained-but-recovered case, and a lower-bound lost-coherence case, enabling a graded empirical stress test rather than a simple success-failure contrast.

The paper asks: How does Dialectical Navigation explain variation in safety effectiveness across graded aviation navigation episodes in which dynamic coherence is maintained, strained but recovered, and lost? Three supporting questions guide the analysis: how guided elasticity, collective calibration, and mutual recovery unfold across episode phases; which conditions amplify or suppress each mechanism; and where the framework explains variation strongly or reveals boundary conditions.

The paper contributes in three ways. First, it provides the first graded empirical stress test of Dialectical Navigation. Second, it refines the framework by specifying how mechanism timing, visibility, and adequacy vary across maintained, recovered, and lost trajectories. Third, it translates those differences into implications for procedure design, communication architecture, authority migration, monitoring and assertion, and learning systems in safety-critical operations. Section 2 presents Dialectical Navigation as an empirical lens, Section 3 explains the method, Section 4 presents the findings, Section 5 discusses theory refinement and practical implications, Section 6 outlines limitations, Section 7 concludes, and the final section lists references.

2. Dialectical Navigation as an Empirical Lens

2.1 Dynamic coherence as a graded trajectory

The companion conceptual paper proposed Dialectical Navigation as a mid-range process framework for explaining how safety effectiveness is achieved when teams navigate the tension between procedural standardisation and adaptive responsiveness in safety-critical work (Author, under review). The framework treats safety effectiveness as dynamic coherence in motion: the minimum alignment in intent, understanding, and coordinated action through which teams continue to act safely as conditions evolve

(Author, under review). This shifts attention from static compliance and post hoc failure attribution to coordination in motion.

In this empirical paper, dynamic coherence is treated as a graded trajectory rather than a binary state. Across navigation episodes, coherence may be maintained, substantially strained but recovered, or lost as mismatch intensifies and recoverability narrows. This matters because the analytic question is not simply why one crew succeeds and another fails, but how the quality of coordination changes across the episode as crews confront shifting cues, time pressure, and procedural inadequacy. The framework is therefore designed to capture not only whether teams adapt, but whether adaptation remains bounded, intelligible, and recoverable while work continues (Author, under review; Faraj & Xiao, 2006; Hollnagel, 2014).

2.2 The three mechanisms

Dialectical Navigation explains variation in dynamic coherence through three interdependent mechanisms: guided elasticity, collective calibration, and mutual recovery (Author, under review). These mechanisms are analytically distinct but can overlap and recur within the same navigation episode.

Guided elasticity refers to procedural flexing that remains bounded, intent-preserving, visible to other team members, and potentially re-enterable. It becomes salient when the standing script no longer fits unfolding demands but crews must still preserve the underlying safety logic of procedure rather than drift into opaque improvisation. Empirically, guided elasticity is visible in bounded resequencing, selective checklist use, substitution of equivalent verification, and explicit efforts to preserve procedural intent while adjusting the immediate course of action (Grote, 2012; Hollnagel, 2014; Woods, 2015).

Collective calibration refers to the interactional work through which teams realign interpretation, pacing, workload, and authority while action proceeds. It becomes salient when local understandings begin to diverge and teams must make that divergence operationally discussable through callouts, cross-checks, updates, challenge-response moves, explicit reprioritisation, or authority migration. Calibration does not require full consensus; it requires sufficient alignment for collective action to remain workable (Faraj & Xiao, 2006; Flin et al., 2008; Weick et al., 2005).

Mutual recovery refers to re-anchoring actions through which teams restore a viable trajectory after strain or disruption and renew the conditions for further coordination. Recovery is mutual because restoring a safe trajectory in safety-critical work rarely depends on one actor alone; it requires reciprocal adjustment of roles, timing, and action. Empirically, mutual recovery is visible in go-around decisions, interruption closure, workload redistribution, and short-loop re-anchoring moves that preserve or restore recoverability (Author, under review; Hollnagel, 2014; Woods, 2015).

Taken together, these mechanisms specify how teams move between procedural anchoring and adaptive response while action is still unfolding. In this paper, they are used as coding lenses for reconstructing each navigation episode rather than as a basis for extended conceptual differentiation.

2.3 Empirical expectations for the graded comparison

The framework generates three empirical expectations that guide the comparison. First, episodes in which dynamic coherence is maintained should show early and bounded guided elasticity, timely recalibration of interpretation and task distribution, and recovery-oriented action before the trajectory becomes precarious. Second, episodes in which coherence is substantially strained but recovered should show delayed or uneven calibration, partial weakening of procedural anchoring, and recovery that occurs late but still within a viable window. Third, episodes in which coherence is lost should show either unbounded continuation, suppressed or insufficient calibration, or recovery attempts that begin only after the recoverability window has effectively closed.

These expectations do not imply a linear stage model. The mechanisms may overlap, intensify, or recur as an episode unfolds. What matters empirically is their timing, visibility, and adequacy relative to the emerging mismatch and the narrowing trajectory for safe re-anchoring. The analytic question is whether the three cases differ not only in final outcome, but in how they enact, delay, suppress, or fail to enact the three mechanisms as the navigation episode develops.

3. Method

3.1 Research design

This study uses a graded comparative case design to conduct an empirical stress test of Dialectical Navigation as a process theory of safety effectiveness (Author, under review). Rather than contrasting only one success case with one failure case, the design compares three commercial aviation navigation episodes that vary in coherence trajectory: one in which dynamic coherence was maintained, one in which coherence was substantially strained but recovered, and one in which coherence was lost before safe re-anchoring could occur. The design follows a theoretical replication logic rather than a sampling logic aimed at representativeness (Eisenhardt, 1989; Yin, 2018). The aim is to assess whether the framework discriminates meaningfully across episodes that differ in the timing, visibility, and adequacy of guided elasticity, collective calibration, and mutual recovery under comparable procedural-adaptive tension.

The graded design matters because the framework treats safety effectiveness as dynamic coherence in motion rather than as a binary end state. Comparing maintained, recovered, and lost trajectories makes it possible to examine whether outcome variation is attributable, at least in part, to how coordination unfolds within the episode rather than to static differences in disturbance, environmental severity, or individual competence alone.

3.2 Case selection and data sources

Cases were selected purposively for theoretical leverage. All three come from multicrew commercial aviation conducted in highly proceduralised, time-pressured, high-consequence settings. Selection followed four criteria: (1) occurrence within the same broad domain of proceduralised multicrew flight operations; (2) presence of consequential procedural-adaptive tension generated by mismatch between the standing operational script and unfolding conditions; (3) sufficient documentary

richness to reconstruct interaction timing, role relations, decision points, and recovery attempts; and (4) outcome variation such that the cases together span a graded range of dynamic coherence trajectories (Yin, 2018).

The selected cases satisfy these criteria. US Airways Flight 1549 serves as the maintained-coherence benchmark because the NTSB explicitly links survivability to flight-crew decision-making and CRM during acute disruption (NTSB, 2010). Air Canada Flight 759 provides the strained-but-recovered case, in which coherence degraded substantially during approach but was restored through a late go-around (NTSB, 2018). HL7762 provides the lost-coherence case, in which descent continued below decision altitude without the required visual references, the monitoring role did not assert the expected go-around, and CRM did not function appropriately (JTSB, 2016). Together, the cases provide theoretically contrasting instances within the same operational domain.

The analysis relies exclusively on primary documentary evidence from official accident and incident investigations. Within each report, the analysis draws on event chronologies, cockpit voice recorder and flight data recorder traces where available, role expectations associated with pilot-flying and pilot-monitoring coordination, procedural and checklist references, investigator analyses, probable-cause findings, and post-event safety recommendations (NTSB, 2010, 2018; JTSB, 2016). This documentary strategy provides an auditable chain from reported interactional trace to interpretive coding while reducing dependence on retrospective self-report.

3.3 Unit of analysis and analytical framework

The unit of analysis is the navigation episode. Following Dialectical Navigation, a navigation episode begins when mismatch between standing procedural expectations and unfolding operational reality becomes consequential for trajectory safety or decision timing and ends when the crew either restores a workable alignment in intent, understanding, and action or the episode culminates in material breakdown or safe outcome closure (Author, under review). Using the episode rather than the whole flight keeps attention on coordination in motion rather than on the event as a single undifferentiated outcome.

Episode boundaries were specified explicitly for each case using investigator chronologies and interactional evidence. In US1549, the episode begins with the loss of thrust that renders the normal departure script unusable and ends with the forced water landing judged survivable by investigators (NTSB, 2010). In AC759, the episode begins when visual-procedural mismatch during approach becomes operationally consequential and ends with restoration of a safe climb trajectory through go-around action (NTSB, 2018). In HL7762, the episode begins when approach ambiguity makes continued descent contingent on recalibration or re-anchoring and ends with collision following failed or delayed recovery action (JTSB, 2016).

Coding was guided by the four analytical elements set out in Section 2: dynamic coherence, guided elasticity, collective calibration, and mutual recovery (Author, under review). To operationalise those elements comparably across cases, each episode was

coded through a shared six-phase structure: (A) baseline procedural anchoring; (B) trigger or mismatch emergence; (C) guided elasticity; (D) collective calibration; (E) mutual recovery; and (F) outcome and feed-forward consequences. This phase structure applies Langley's (1999) temporal-bracketing logic and allows comparison to focus on when and how mechanisms become visible, delayed, suppressed, or ineffective across cases.

3.4 Coding procedure, comparative logic, and trustworthiness

Coding proceeded through a structured evidence-to-mechanism mapping process. For each case, reconstruction began with chronological extraction of temporally ordered interactional data from the investigation materials, including recorder-derived traces where available, reported crew exchanges, action sequences, and documented role expectations. These chronologies were then segmented into the six shared phases in order to isolate key transitions associated with mismatch emergence, adaptive deviation, recalibration, and recovery (Langley, 1999). Within each phase, interactions and decision points were first coded inductively for coordination relevance—for example, callouts, trajectory reassessment, checklist adaptation, workload redistribution, go-around initiation, or silence in the face of voiced concern—and only then deductively mapped to guided elasticity, collective calibration, or mutual recovery. This sequencing kept coding anchored in documentary trace before theoretical meaning was assigned.

Particular attention was given to ambiguous or suppressed signals. Instances in which monitoring cues were generated but did not precipitate directive coordination were retained as analytically important because they help distinguish delayed calibration from absent calibration and failed recovery from no recovery attempt. Analytic memos were used throughout to document provisional interpretations, revisit mechanism attributions across phases, and record why particular traces were treated as evidence of bounded adaptation, recalibration, or late re-anchoring (Miles, Huberman, & Saldaña, 2014; Pratt, 2009).

The comparison followed a within-case then cross-case logic. Each case was first reconstructed independently to identify episode boundaries, phase transitions, and salient interactional events. Cross-case comparison then examined patterned variation along five dimensions: mismatch recognition, boundedness of procedural adaptation, shifts between implicit and explicit coordination, authority migration between pilot-flying and pilot-monitoring roles, and timing and adequacy of recovery action. The purpose was to assess whether maintained, strained-but-recovered, and lost trajectories correspond to meaningful differences in mechanism timing and adequacy.

Boundary conditions and rival explanations were considered explicitly. Potential boundary conditions included cue ambiguity, workload compression, degraded visibility, authority gradients, monitoring-role overload, and limited opportunity for procedural re-entry. Rival contextual explanations—including environmental severity, aircraft state, fatigue, training deficits, and experience differences—were considered during episode reconstruction and revisited during cross-case comparison to assess whether the observed trajectories could plausibly be attributed to contextual severity alone (Yin,

2018). These factors were treated as potential amplifiers or suppressors of the three mechanisms rather than as residual background detail.

Trustworthiness was supported through a transparent chain of evidence linking report-based interactional traces to mechanism attribution, use of a common coding template across all three cases, phase-linked evidence matrices, explicit rival-explanation checks, and iterative memoing of coding decisions and refinements (Miles, Huberman, & Saldaña, 2014; Pratt, 2009). These procedures do not remove interpretive judgement, but they make that judgement more transparent, comparable, and open to scrutiny.

4. Findings

This section reconstructs three aviation navigation episodes through the lens of guided elasticity, collective calibration, and mutual recovery. The aim is not to retell each accident or incident in full, but to show how dynamic coherence was maintained, substantially strained but recovered, or lost as crews moved from baseline procedural anchoring through mismatch, adaptation, recalibration, and recovery. To keep the section aligned with the comparative logic set out in Section 3, each case is read against the same five dimensions: mismatch recognition, boundedness of procedural adaptation, shifts from implicit to explicit coordination, authority migration between pilot-flying and pilot-monitoring roles, and the timing and adequacy of recovery action.

4.1 US1549 - maintained coherence

US Airways Flight 1549 serves as the maintained-coherence benchmark because mismatch became immediate and unambiguous. Once the bird strike produced an almost complete loss of thrust in both engines shortly after departure, the standing departure and return logic no longer fit the aircraft state or the time available. That early mismatch recognition mattered because it sharply reduced interpretive lag and forced re-anchoring almost at once (NTSB, 2010).

Procedural adaptation in this case remained clearly bounded. The crew did not attempt literal checklist completion under conditions in which time and altitude made that impossible, but neither did they drift into opaque improvisation. Instead, they used the most relevant checklist logic as an anchor while selectively departing from sequence when necessary to preserve control. The captain's early start of the APU, which the NTSB later judged important for survivability, is the clearest example of guided elasticity: procedure was stretched, but its safety intent and intelligibility were preserved (NTSB, 2010). Coordination also shifted to a sufficiently explicit form very early. The first officer moved into checklist and support work while the captain concentrated on aircraft control, external communication, and landing choice. That shift did not require dramatic authority migration because PF/PM differentiation remained stable and mutually intelligible throughout the episode. Recovery therefore began before the trajectory became precarious in an unrecoverable sense. The ditching decision re-anchored the episode around a new survivable end state rather than an airport return later judged unlikely to succeed under real timing constraints (NTSB, 2010).

US1549 therefore shows maintained coherence not because disruption was minor, but because all five dimensions aligned early: mismatch was unmistakable, adaptation remained bounded, coordination became selectively explicit without role confusion, authority remained clear, and recovery began while a viable trajectory still remained.

4.2 AC759 - strained but recovered coherence

Air Canada Flight 759 provides the intermediate case, in which coherence was substantially strained during approach but ultimately recovered. The episode unfolded more gradually than US1549. The crew had been cleared for a familiar visual approach to Runway 28R, yet Runway 28L was closed, the relevant NOTAM information was reviewed ineffectively, and the first officer did not manually tune the ILS frequency for 28R as backup lateral guidance, nor did the captain verify it. Mismatch recognition was therefore delayed rather than absent: the approach remained formally routine even as the interpretive and procedural anchors that should have stabilized it were weakening (NTSB, 2018).

On the second dimension, procedural adaptation was not fully collapsed, but it was weakly bounded. The crew continued within a visual-approach frame that still seemed plausible to them for too long. In a stronger episode, bounded flexing would have involved earlier re-anchoring to backup guidance, stronger cross-checking, or a more explicit reconsideration of whether the visual script remained valid. Instead, expectation increasingly substituted for verification as the captain interpreted Taxiway C as Runway 28R while the first officer, acting as pilot monitoring, was occupied programming missed-approach information. The episode therefore shows degraded guided elasticity rather than either rigid compliance or outright abandonment of procedure (NTSB, 2018).

The clearest strain appears in the next two dimensions. The shift from implicit to explicit coordination came late. Salient cues were available, including surface lighting and other indicators of misalignment, and the crew did express uncertainty, including querying whether the runway was clear. But the available documentary record suggests that uncertainty did not become operationally decisive soon enough to reorganise the crew's interpretation of the airport surface (NTSB, 2018). At the same time, authority migration from the captain's landing frame to a stronger pilot-monitoring corrective role was muted until the very end. Because the cockpit voice recorder was overwritten, the interactional record is thinner than in the other two cases, so the evidence warrants caution. Still, the NTSB's account supports the narrower claim that monitoring weakened at precisely the point where more directive recalibration was needed and that CRM breakdown formed part of the event sequence (NTSB, 2018).

Recovery was late but sufficient. The crew initiated a low-altitude go-around only after the aircraft had descended to extremely small margins above aircraft on Taxiway C, yet that action still restored a viable climb trajectory and prevented catastrophe (NTSB, 2018). AC759 therefore represents a strained-but-recovered trajectory: mismatch became consequential only gradually, procedural anchoring weakened under expectation bias, coordination became explicit too late, authority migration remained

limited until the final moments, and recovery succeeded only within a sharply narrowed window.

4.3 HL7762 - lost coherence

HL7762 represents the lower-bound case, in which dynamic coherence was lost before safe re-anchoring could occur. Here too, mismatch did not begin as a catastrophic technical failure. It emerged through growing visual ambiguity on approach. The JTSCB records the first officer indicating that the runway scene was ambiguous because of cloud conditions, which means that the standing descent script had already become unstable before the formal minimums sequence was exhausted. Mismatch recognition was therefore present, and present early enough to matter, but it remained contestable for too long to interrupt continuation decisively (JTSCB, 2016).

Procedural adaptation in this episode crossed from bounded flexing into drift. In a stronger enactment of guided elasticity, continuation under ambiguity would have remained tightly coupled to explicit verification of whether visual criteria were still satisfied and whether safe re-entry remained possible. Instead, the aircraft continued below decision altitude without the required continuous visual references. The problem is not simply noncompliance; it is that what may have begun as procedural flexing lost its boundedness and outran the conditions that made continuation safe or recoverable (JTSCB, 2016; Dekker, 2011).

The breakdown becomes clearest when the third and fourth dimensions are separated. Coordination did become verbally explicit: the first officer called "Minimum," the captain replied "Continue," and the first officer then stated that the runway was not in sight, followed by the captain's "Wait a second" (JTSCB, 2016). But explicitness alone was insufficient. Authority did not migrate from the captain's continuation logic to the pilot-monitoring role that should have triggered an immediate go-around. In other words, the episode shows suppressed calibration rather than absent calibration. Concern entered the cockpit interactionally, but it did not become directive coordination while a viable correction window still remained. The JTSCB's conclusion that the pilot monitoring did not make the expected go-around call and that CRM did not function appropriately reinforces that reading (JTSCB, 2016).

Recovery was attempted, but too late. The captain eventually initiated a go-around, yet only immediately before collision with approach-lighting structures and terrain in front of the runway threshold (JTSCB, 2016). By that point, earlier drift in procedural adaptation and the failure of authority migration had already narrowed the recoverability window beyond safe restoration. HL7762 therefore marks the lost-coherence condition: mismatch was noticed, concern was voiced, but procedural boundedness, directive recalibration, and timely recovery did not hold together long enough to preserve a safe trajectory.

4.4 Cross-case findings

The cross-case comparison shows that the decisive variation lies less in the abstract presence of procedures, CRM routines, or formal cockpit roles than in how those resources were enacted as mismatch accumulated. All three crews operated in highly

proceduralised multicrew environments. What differed was how quickly mismatch became actionable, whether procedural flexing remained bounded, how coordination shifted from implicit to explicit form, whether the monitoring role could become corrective authority, and whether recovery began while a viable trajectory still remained.

Table 2. Cross-case comparison on the five analytic dimensions of dynamic coherence

Comparative dimension	US1549 (Maintained Coherence)	AC759 (Strained but Recovered)	HL7762 (Lost Coherence)
Mismatch recognition (how early and decisively mismatch became actionable)	Immediate and unambiguous dual-engine thrust loss made the standing flight script unusable at once.	Mismatch emerged gradually under runway-closure conditions, expectation bias, and weakened backup verification; recognition was delayed.	Visual-procedural mismatch was noticed and voiced, but remained contestable too long to interrupt continuation decisively.
Boundedness of procedural adaptation (whether flexing remained intent-preserving and re-enterable)	Strongly bounded guided elasticity: checklist logic was selectively adapted while preserving procedural intent and survivability.	Weakly bounded adaptation: the crew stayed within a visual-approach frame, but backup verification eroded and expectation substituted for cross-checking.	Adaptation drifted into unbounded continuation: descent below decision altitude without the required visual references outran safe procedural re-entry.
Shift from implicit to explicit coordination (whether doubt became shared operational reorientation)	Coordination became selectively explicit early; roles, priorities, and task distribution were clear under severe compression.	Uncertainty surfaced, but explicit recalibration came late and did not reorganise interpretation until margins were extremely narrow.	Concern became verbally explicit, but explicitness did not become directive coordination while recovery remained viable.
Authority migration between PF and PM roles (whether monitoring became corrective authority)	Minimal migration was required; stable PF/PM differentiation and clear authority supported coherence.	Monitoring influence was weakened by heads-down workload and expectation bias; corrective authority remained limited until the final go-around.	Authority migration failed; PM concern was voiced but did not displace the captain's continuation logic in time.
Timing and adequacy of recovery action (relative to the remaining recoverability window)	Early and adequate: the ditching decision re-anchored the episode around a survivable end state while a viable trajectory still remained.	Late but adequate: the low-altitude go-around restored a viable climb trajectory, but only within a sharply narrowed window.	Late and inadequate: the go-around was attempted only after the recoverability window had effectively closed.

Note. Table rows mirror the five cross-case dimensions specified in Section 3.4. PF = pilot flying; PM = pilot monitoring.

The first comparative dimension is mismatch recognition. In US1549, mismatch was immediate and unambiguous: dual-engine thrust loss rendered the standing script unusable at once. That immediacy reduced interpretive lag and forced early re-anchoring. AC759 occupies the middle position. Mismatch emerged through runway-closure conditions, weakened procedural backup, expectation-driven perception, and heads-down monitoring, so the crew remained inside a formally routine visual-approach frame longer than was safe. HL7762 differs again. There, mismatch was recognised and voiced early enough to matter, but because it took the form of contestable visual ambiguity rather than technical rupture, it did not decisively interrupt continuation. Across the three cases, coherence is strongest where mismatch becomes operationally explicit early and weakest where it remains open to continuation.

The second dimension is the boundedness of procedural adaptation. US1549 shows the strongest form of guided elasticity. Procedure was flexed because the available checklist structure could not be followed literally within the time available, yet the crew's actions remained tightly tied to procedural intent and survivability. Procedural anchoring was stretched, not abandoned. AC759 shows a weaker but still recognisable boundedness. The crew remained nominally inside the visual-approach script, yet backup verification eroded and expectation substituted for cross-checking. Adaptation therefore remained possible, but insufficiently bounded. HL7762 marks the lower-bound condition. Continued descent below decision altitude without the required visual references transformed flexing into drift. The contrast is important because it shows that the issue is not simply whether crews depart from procedure, but whether the departure remains visible, justified by procedural intent, and compatible with safe re-entry.

The third dimension concerns shifts from implicit to explicit coordination. Here the cases reveal that explicitness is itself graded. In US1549, coordination became selectively explicit very early. The crew did not need elaborate discussion, but they made roles, priorities, and task distribution clear enough to maintain a common trajectory under compression. In AC759, uncertainty surfaced only gradually and did not become sufficiently explicit soon enough to reorganise the crew's understanding of the airport surface. In HL7762, coordination became verbally explicit in a narrower sense—concern was voiced directly at minimums—but explicitness came without effective reorientation of collective action. The comparison therefore suggests that explicitness matters not as communication volume, but as the timely conversion of local doubt into shared operational significance.

The fourth dimension, authority migration between pilot-flying and pilot-monitoring roles, sharpens that point further. In US1549, authority did not need to migrate dramatically because PF/PM differentiation remained stable and mutually intelligible; clear role execution itself supported coherence. AC759 shows a more fragile pattern. The monitoring function was weakened by workload and expectation bias, and the available evidence does not support a strong claim of early corrective intervention from the pilot-monitoring role. HL7762 displays the clearest failure. Although the pilot monitoring voiced concern, corrective authority did not displace the captain's

continuation frame while recovery remained possible. This distinction matters because shifts to explicit coordination and authority migration are related but not identical. Teams may voice uncertainty without converting that uncertainty into directive action. HL7762 is the clearest example of that decoupling, and AC759 shows a milder version.

The fifth dimension is the timing and adequacy of recovery action. US1549 shows early and adequate recovery: once return to the original plan was no longer viable, the crew re-anchored around a new survivable end state before the episode fragmented into delay or drift. AC759 shows late but still adequate recovery: the go-around was initiated only after margins had narrowed drastically, yet before touchdown commitment made restoration impossible. HL7762 shows late and inadequate recovery: the go-around was initiated, but only after the recoverability window had effectively closed. Recovery is therefore not a simple yes-or-no variable. Its significance depends on whether earlier interactional work has preserved enough boundedness and alignment for re-anchoring to succeed.

Taken together, the five dimensions support a graded reading of dynamic coherence. US1549 represents the maintained condition because the five dimensions aligned early and mutually reinforced one another. AC759 represents the strained-but-recovered condition because all five dimensions weakened, but not beyond restoration. HL7762 represents the lost condition because mismatch recognition, procedural boundedness, explicit coordination, authority migration, and recovery timing came apart before safe re-anchoring could occur. The cases therefore support the core claim of Dialectical Navigation: safety effectiveness varies not simply by whether teams adapt, but by how early and how coherently they convert emerging mismatch into bounded adaptation, shared recalibration, and recoverable action.

5. Discussion

This study set out to empirically stress-test Dialectical Navigation across three aviation navigation episodes that differed in coherence trajectory. The comparison shows that the framework explains more than binary success and failure: the decisive variation lies in how emerging mismatch is translated into bounded adaptation, shared operational significance, and timely re-anchoring. The discussion refines the theory in three respects.

5.1 What the graded comparison adds to Dialectical Navigation

The first refinement is conceptual: dynamic coherence is better understood as a trajectory rather than a threshold condition. A binary success-failure frame obscures the large middle ground in which coordination degrades materially yet remains recoverable. The cases indicate that the decisive variation appears earlier, in whether mismatch becomes operationally consequential soon enough for alignment in intent, understanding, and action to remain repairable. This matters because organisations rarely manage only routine success or catastrophe; they repeatedly confront episodes in which coordination is strained, partially restored, or lost in motion. Dynamic coherence therefore appears less as a state teams either possess or lose and more as

a fluctuating level of recoverable coordination achieved as the episode unfolds (Weick, 1995; Faraj & Xiao, 2006; Author, under review).

The second refinement is processual. Guided elasticity, collective calibration, and mutual recovery vary not simply in presence or absence, but in timing, visibility, and adequacy. All three episodes involved movement beyond the standing script, yet this alone did not explain their divergence. Guided elasticity was safer when procedural flexing remained bounded, visible, and compatible with re-entry; it became dangerous when continuation outran verification. Mutual recovery was likewise graded: a go-around or escape manoeuvre had very different significance depending on whether earlier coordination had preserved enough margin for it to work. This shifts the framework beyond static adaptation-versus-failure labels toward a sharper account of mechanism quality in time (Hollnagel, 2014; Woods, 2015).

The strongest refinement concerns collective calibration. Across the comparison, the pivotal difference was not whether crews sensed difficulty, but whether local doubt became shared operational significance early enough to reorganise interpretation, workload, and authority while action continued. Where calibration kept pace with mismatch, guided elasticity remained bounded and recovery could build on it. Where calibration lagged, adaptation weakened into expectation-driven continuation and recovery became compressed or ineffective. The empirical paper thus extends the conceptual companion by showing not only that the three mechanisms matter, but that their temporal coupling determines whether dynamic coherence remains recoverable in motion (Author, under review).

5.2 Boundary conditions and theory refinement

The graded comparison also sharpens the framework's boundary conditions. One is cue ambiguity. When mismatch is immediate and unambiguous, the standing script is displaced quickly and crews are forced into explicit re-anchoring. When mismatch emerges through contestable perceptual or procedural signals, however, teams can remain inside a plausible continuation frame longer than is safe. Under those conditions, the explanatory burden shifts toward collective calibration, because the challenge is not cue absence but the need to make weak or ambiguous cues operationally decisive in time (Weick et al., 2005; Maitlis & Christianson, 2014). A second boundary condition concerns authority migration between monitoring and control roles. Hierarchy matters less as a static feature than as an interactional question: can monitoring become corrective authority while the trajectory is still recoverable? When voiced concern does not become directive coordination, calibration remains verbal rather than operational (Edmondson, 1999; Faraj & Xiao, 2006).

A third boundary condition is the recoverability window itself. Recovery is not solely a matter of technical possibility or crew intention. It depends on whether earlier interactional work has preserved enough boundedness and alignment for re-anchoring to succeed. The same visible act can therefore have different theoretical meaning across episodes. A late go-around may represent effective recovery in one case and failed recovery in another, depending on whether guided elasticity and collective calibration have preserved a viable trajectory up to that point. This refines Dialectical

Navigation by suggesting that the mechanisms are asymmetrically time-sensitive: delay in calibration disproportionately undermines the adequacy of subsequent recovery (Hollnagel, 2014; Woods, 2015; Author, under review).

This positioning also clarifies the framework's relation to adjacent literatures. CRM research identifies routines of monitoring, cross-checking, and communication; HRO and sensemaking work explain why sensitivity to operations, equivocal cues, and deference to expertise matter; resilience perspectives foreground adaptation and recovery. Dialectical Navigation adds an episode-level process account of how these resources are coupled in motion as teams move between procedural anchoring, adaptive response, recalibration, and re-anchoring (Weick et al., 2005; Weick & Sutcliffe, 2007; Flin et al., 2008; Hollnagel, 2014). The empirical paper therefore extends the conceptual companion not by replacing those literatures, but by showing where the framework explains strongly, where its mechanisms are suppressed, and where its theoretical boundaries become visible in practice.

5.3 Practical implications

The managerial implication is not that safety-critical organisations need “more flexibility” in the abstract. They need systems in which adaptation remains bounded, discussable, and recoverable under pressure. First, procedures should function as intent-bearing anchors rather than exhaustive scripts. They should clarify not only normal execution, but also when bounded departure is legitimate, how deviations remain visible to other team members, and what counts as safe procedural re-entry or explicit closure. Second, organisations should distinguish between anchoring communication and calibration communication. Routine phases can rely on confirmation and stable role execution; non-routine phases require interaction that surfaces mismatch, converts local doubt into shared operational significance, and supports temporary shifts in workload and authority.

Third, monitoring and assertion should be treated as designed coordination capabilities rather than as individual interpersonal traits. One of the most consequential differences across the cases lay in whether monitoring signals altered team trajectory while recovery remained viable. Training, simulation, line checks, and debriefs should therefore focus not only on SOP compliance and generic CRM behaviours, but on whether voiced concern actually triggers recalibration and action under ambiguity, workload compression, and time pressure. The practical issue is not simply speaking up; it is whether organisations make corrective escalation usable in real time.

Finally, recovery should be treated as feed-forward re-anchoring, not merely as technical escape or post-event reporting. Learning systems should use incidents, near misses, and strained recoveries to strengthen escalation cues, backup verification, monitoring norms, and re-entry conditions for future episodes. From an engaged management perspective, this shifts attention from abstract compliance-versus-flexibility debates to concrete design choices that preserve recoverability in live operations. That is the paper's managerial contribution: a way to diagnose not merely whether safety held or failed, but where recoverability was preserved, narrowed, or

exhausted, and which organisational design choices are most likely to strengthen future coordination in safety-critical work.

6. Limitations

This study is bounded in three ways. First, it relies on official investigation reports as primary evidence. These reports provide auditable chronologies, procedural references, and selected interactional traces, but they do not offer identical visibility into cockpit coordination across cases. This matters especially in AC759, where interactional evidence is thinner than in the other two episodes, limiting fine-grained reconstruction of how uncertainty was surfaced, resisted, or converted into action. The analysis therefore privileges what can be supported reliably - timing, procedural anchors, documented cues, and investigator findings - rather than inferring interaction that cannot be observed.

Second, the design is intentionally bounded to commercial multicrew aviation in highly proceduralised, time-pressured, high-consequence settings. This strengthens cross-case comparability but limits direct transfer to settings in which role structure, procedural governance, and recoverability windows differ substantially.

Third, the analysis is theory-led. Coding through guided elasticity, collective calibration, and mutual recovery sharpens mechanism timing and comparative clarity, but it also foregrounds coordination process over other explanatory levels. Rival explanations, including environmental severity, aircraft state, fatigue, training, and experience differences, were considered explicitly, but the paper should be read as a comparative stress test of one process framework rather than as an exhaustive causal account of each event.

7. Conclusion

This paper conducted a graded empirical stress test of Dialectical Navigation by comparing three aviation navigation episodes in which dynamic coherence was maintained, substantially strained but recovered, and lost. The analysis shows that the framework explains more than binary success and failure. The decisive variation lies in when mismatch becomes operationally consequential, whether procedural adaptation remains bounded, how quickly coordination is recalibrated, and whether recovery begins while a viable trajectory still remains.

The comparison refines the framework in two core ways. First, dynamic coherence is best understood as a trajectory rather than a threshold state. Second, the mechanisms matter not only by their presence, but by their timing, visibility, and adequacy, with collective calibration emerging as the pivotal hinge between bounded adaptation and recoverable outcome. Safety effectiveness, on this view, is achieved not by procedures alone or by flexibility alone, but by how teams move between anchoring and adaptation without exhausting recoverability in motion.

For management practice, the key consequence is that recoverability must be designed, not assumed. Procedures, communication routines, monitoring roles, and learning systems are most valuable when they help teams surface mismatch early,

convert doubt into coordinated action, and re-anchor before the viable trajectory closes. The paper's empirical message is that safe navigation depends on preserving dynamic coherence while there is still time to recover.

References

- Author. under review. Dialectical navigation: Dynamic coherence in safety-critical work. Manuscript under review.
- Bigley, G. A., & Roberts, K. H. 2001. The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal*, 44: 1281-1299.
- Dekker, S. 2011. *Drift into failure: From hunting broken components to understanding complex systems*. Farnham, UK: Ashgate.
- Edmondson, A. C. 1999. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44: 350-383.
- Eisenhardt, K. M. 1989. Building theories from case study research. *Academy of Management Review*, 14: 532-550.
- Faraj, S., & Xiao, Y. 2006. Coordination in fast-response organizations. *Management Science*, 52: 1155-1169.
- Farjoun, M. 2010. Beyond dualism: Stability and change as a duality. *Academy of Management Review*, 35: 202-225.
- Flin, R., O'Connor, P., & Crichton, M. 2008. *Safety at the sharp end: A guide to non-technical skills*. Aldershot, UK: Ashgate.
- Grote, G. 2012. Safety management in different high-risk domains--all the same? *Safety Science*, 50: 1983-1992.
- Hale, A., & Borys, D. 2013. Working to rule, or working safely? Part 1: A state of the art review. *Safety Science*, 55: 207-221.
- Hollnagel, E. 2014. *Safety-I and Safety-II: The past and future of safety management*. Farnham, UK: Ashgate.
- Japan Transport Safety Board (JTSB). 2016. Aircraft accident investigation report: Asiana Airlines, Inc., Airbus A320-200, HL7762, Hiroshima Airport, April 14, 2015. Tokyo: JTSB.
- Langley, A. 1999. Strategies for theorizing from process data. *Academy of Management Review*, 24: 691-710.
- Lei, Z., Waller, M. J., Hagen, J. U., & Kaplan, S. 2016. Team adaptiveness in dynamic contexts: Contextualizing the roles of interaction patterns and in-process planning. *Group & Organization Management*, 41: 491-525.
- Leveson, N. 2011. *Engineering a safer world: Systems thinking applied to safety*. Cambridge, MA: MIT Press.
- Maitlis, S., & Christianson, M. 2014. Sensemaking in organizations: Taking stock and moving forward. *Academy of Management Annals*, 8: 57-125.
- Miles, M. B., Huberman, A. M., & Saldaña, J. 2014. *Qualitative data analysis: A methods sourcebook* (3rd ed.). Thousand Oaks, CA: Sage.
- National Transportation Safety Board (NTSB). 2010. Loss of thrust in both engines after encountering a flock of birds and subsequent ditching on the Hudson River, US Airways Flight 1549, Airbus A320-214, N106US, Weehawken, New Jersey, January 15, 2009. Aircraft accident report NTSB/AAR-10/03. Washington, DC: NTSB.

- National Transportation Safety Board (NTSB). 2018. Taxiway overflight, Air Canada Flight 759, Airbus A320-211, C-FKCK, San Francisco, California, July 7, 2017. Aircraft incident report NTSB/AIR-18/01. Washington, DC: NTSB.
- Pratt, M. G. 2009. For the lack of a boilerplate: Tips on writing up and reviewing qualitative research. *Academy of Management Journal*, 52: 856-862.
- Reason, J. 2008. *The human contribution: Unsafe acts, accidents and heroic recoveries*. Farnham, UK: Ashgate.
- Smith, W. K., & Lewis, M. W. 2011. Toward a theory of paradox: A dynamic equilibrium model of organizing. *Academy of Management Review*, 36: 381-403.
- Weick, K. E. 1995. *Sensemaking in organizations*. Thousand Oaks, CA: Sage.
- Weick, K. E., & Sutcliffe, K. M. 2007. *Managing the unexpected: Resilient performance in an age of uncertainty* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. 2005. Organizing and the process of sensemaking. *Organization Science*, 16: 409-421.
- Woods, D. D. 2015. Four concepts for resilience and the implications for the future of resilience engineering. *Reliability Engineering & System Safety*, 141: 5-9.
- Yin, R. K. 2018. *Case study research and applications: Design and methods* (6th ed.). Thousand Oaks, CA: Sage.