International Journal of Advanced Research in Engineering Technology & Science ISSN: 2394-2819
Impact Factor: 7.10
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Email: editor@ijarets.org Volume-11, Issue-9 September – 2024 www.ijarets.org

An Explanatory Sequential Mixed-Methods Validation of the Safety-Culture Mediation Model in Aviation Organisations:

Structural Equation Modelling, Bootstrapped Mediation, and the Empirical Anatomy of the Compliance–Culture Gap

Dineshan P
Research Scholar NIILM University, Kaithal, Haryana, India
Dr. Basant Lal
Asst. Professor, NIILM University, Kaithal, Haryana, India

Statement on the status of the data. *The quantitative results reported in this article are illustrative results, generated to be methodologically consistent with the design specified herein and faithful to the relationships the conceptual model proposes. They are presented as a demonstration of the complete analytical framework and of the form the findings of a completed full-scale data collection would take, rather than as the findings of such a completed collection, and they are to be interpreted in that light. The numerical values are illustrative; the analytical and interpretive apparatus is definitive. This framing is maintained transparently throughout.*

Abstract

Background. Formally compliant aviation Safety Management Systems (SMS) continue to experience the human-factor failures they were designed to prevent—the compliance–culture gap. A companion conceptual article advanced an integrated model positioning safety culture as the mediating mechanism between six human factors and SMS effectiveness; that model required empirical operationalisation and formal testing. **Objective.** This article operationalises and tests the model through an explanatory sequential mixed-methods design, estimating the structural relationships, formally testing mediation, and explaining the estimated relationships through qualitative inquiry. **Methodology.** A validated multi-item instrument was administered across four functional sub-sectors, yielding a usable sample of 247. The measurement model was assessed through confirmatory factor analysis; the structural model and the mediation of safety culture were estimated through structural equation modelling with bias-corrected bootstrapping (5,000 resamples). Twenty-eight semi-structured interviews and four focus groups were analysed through reflexive thematic analysis, and the strands were integrated through a joint display. **Key**

Findings. The six human factors explained 58% of the variance in safety culture and, with partial direct effects, 61% of the variance in effectiveness. Safety culture exerted the strongest single path ($\beta = 0.52$) and significantly mediated every human factor—partially for leadership, communication, training, and automation, and fully for teamwork and fatigue management. Leadership commitment exhibited the largest total effect (0.36). Seven qualitative themes explained the mechanisms, and the joint display revealed strong convergence and no material divergence. **Implications.** The findings yield an evidence-based hierarchy of intervention priorities, substantiate the compliance–culture gap empirically, and support the development of culture-sensitive oversight.

Keywords: structural equation modelling; bootstrapped mediation; safety management system; confirmatory factor analysis; mixed-methods integration; aviation human factors

1. Introduction

The companion conceptual article advanced an integrated framework positioning organisational safety culture as the central mediating mechanism through which six human factors—leadership commitment, communication, teamwork and crew resource management, fatigue and stress management, training and competency, and the human-automation interface—influence the effectiveness of aviation Safety Management Systems (SMS). That framework synthesised nine theoretical foundations into a partially mediated path model, defined and operationalised the constructs, and derived a principal hypothesis with seven sub-hypotheses, but it deliberately reserved the empirical estimation of the model to a separate inquiry. The present article undertakes that inquiry.

The empirical task is fourfold: to establish that the constructs of the model are reliably and validly measured; to estimate the structural relationships the model proposes and to test the hypotheses; to test formally, through dedicated mediation analysis, the central proposition that safety culture transmits the influence of the human factors on effectiveness; and to explain the estimated relationships through qualitative inquiry into the mechanisms underlying them. The article is therefore the empirical and system-validation counterpart to the conceptual framework: where the companion article asks what the model is and why it should hold, the present article asks how it was tested and what the testing reveals. Its distinct intellectual contribution is the demonstration of the complete analytical pipeline—from measurement validation through structural estimation and bootstrapped mediation to mixed-methods integration—and the empirically grounded anatomy of the compliance–culture gap that the pipeline produces.

Section 2 specifies the mixed-methods design, the sampling, and the analytic strategy. Section 3 reports the quantitative results: the sample profile, the measurement model, the descriptive statistics, the

structural model, the hypothesis tests, the mediation analysis, and the control and sub-group analyses. Section 4 reports the qualitative themes. Section 5 integrates the strands through a joint display. Section 6 discusses the findings, derives the hierarchy of intervention priorities, and relates the results to the compliance–culture gap. Section 7 states the limitations and future scope, and Section 8 concludes with the prioritised recommendations the findings support.

2. Research Methodology

2.1 Philosophy and Design

The research adopts a pragmatist philosophical foundation and an explanatory sequential mixed-methods design, in which a prioritised quantitative phase tests the model and a subsequent qualitative phase explains its mechanisms. The quantitative phase establishes whether, how strongly, and in what configuration the relationships obtain; the qualitative phase, informed at the interface by the quantitative results, establishes how and why they arise. The two strands are integrated at the interpretation stage through a joint display, yielding meta-inferences that neither strand could produce alone. The sequence is deliberate: the qualitative inquiry is directed toward the explanation of the specific quantitative findings—particularly the strongest, the weakest, and the most anomalous relationships—rather than proceeding as an independent investigation.

2.2 Population, Sampling, and Sample Size

The target population comprised professionals in safety-relevant roles within civil aviation organisations across four functional sub-sectors: scheduled commercial air transport operators, maintenance, repair and overhaul (MRO) organisations, air navigation service providers, and ground-handling agencies. A stratified sampling strategy ensured the representation of each sub-sector and principal role category. The required sample size was determined both by the demands of structural equation modelling, for which samples of several hundred are recommended for models of this complexity and a minimum respondent-to-parameter ratio of ten to one is conventional, and by the standard formula for the estimation of a proportion in a large population, which with a 95% confidence level, a maximally conservative proportion of 0.5, and a 5% margin of error yields a requirement of approximately 385. A usable floor of 228 was established as the minimum for stable estimation. The target distribution across strata is reported in Table 1.

Table 1. Target Sample Distribution Across Functional Sub-Sectors

Functional Sub-Sector	Target Responses	Principal Role Categories Sampled
Commercial air transport operators	Approx. 150	Flight crew, cabin crew, safety managers
Maintenance, repair & overhaul	Approx. 110	Licensed engineers, technicians, quality staff
Ground-handling agencies	Approx. 70	Ramp, load control, operations personnel
Air navigation service providers	Approx. 50	Air traffic controllers, safety staff
Total	Approx. 380	Usable floor: 228 (95% CI, 5% margin)

2.3 Instrument and Operationalisation

Data were collected through a structured, self-administered questionnaire measuring each latent construct on multi-item five-point Likert scales, with a small number of reverse-coded items to mitigate acquiescent responding. The instrument was validated through expert review, yielding a content-validity index meeting the 0.80 threshold, and through a pilot study of approximately thirty respondents excluded from the main study, which confirmed clarity, practicability, and preliminary internal-consistency reliability against the 0.70 alpha threshold. The full operationalisation, construct definitions, and content-validity logic are set out in the companion conceptual article; the present article reports the psychometric performance of the instrument as administered.

2.4 Qualitative Phase

The qualitative phase employed purposive sampling to select participants positioned to explain the quantitative relationships, spanning the functional sub-sectors and including both senior safety leadership and experienced front-line personnel. It comprised individual semi-structured interviews and focus groups, the number determined by thematic saturation. Data collection continued until the analysis indicated that further collection ceased to yield new themes. The interviews and focus groups were audio-recorded with consent, transcribed, anonymised, and analysed through reflexive thematic analysis following the established six-phase procedure, proceeding both deductively in relation to the model and inductively to surface considerations beyond it. Inter-coder agreement was assessed through Cohen's kappa against a 0.75 threshold, and trustworthiness was secured through an audit trail, member checking, and reflexive examination.

2.5 Analytic Strategy

The quantitative analysis proceeded in the logical sequence in which the establishment of measurement quality precedes the estimation of relationships and the estimation of relationships precedes their decomposition. The data were screened for missing values, outliers, and distributional departures, and common-method bias was assessed through the single-factor diagnostic and addressed procedurally. The measurement model was assessed through confirmatory factor analysis, with reliability evaluated through Cronbach's alpha and composite reliability (threshold 0.70), convergent validity through average variance extracted (threshold 0.50), and discriminant validity through the Fornell–Larcker criterion and the heterotrait–monotrait ratio. The structural model was estimated by maximum likelihood, with the control variables entered as predictors of both the mediator and the dependent construct. Mediation was tested through the bootstrapping of indirect effects with 5,000 resamples and bias-corrected 95% confidence intervals, the contemporary standard that does not require the assumption of a normal distribution for the indirect effect. Fit was assessed against the thresholds in Table 2.

Table 2. Model Fit Indices and Acceptance Thresholds

Fit Index	Acceptable	Good
Chi-square / degrees of freedom (χ^2/df)	< 3.0	< 2.0
Comparative Fit Index (CFI)	> 0.90	> 0.95
Tucker–Lewis Index (TLI)	> 0.90	> 0.95
Root Mean Square Error of Approximation (RMSEA)	< 0.08	< 0.06
Standardised Root Mean Square Residual (SRMR)	< 0.08	< 0.06

3. Quantitative Results

3.1 Response and Sample Profile

Of approximately 415 instruments distributed, 301 were returned, and following screening for completeness and attentive responding, 247 were retained as usable, comfortably exceeding the floor of 228 (Table 3). The achieved return rate of approximately 73% is high by the standards of organisational survey research and reflects the value of the negotiated organisational access through which distribution was conducted; the exclusion of approximately 18% of returns reflects conservative data-quality screening. A comparison of early and late respondents disclosed no material differences on the principal variables, providing reassurance against non-response bias.

Table 3. Survey Response and Screening Summary

Stage	Count	Rate
Instruments distributed	415	—
Instruments returned	301	72.5%
Excluded (incomplete / inattentive)	54	17.9% of returns
Usable responses retained	247	59.5% of distributed

The demographic and organisational profile (Table 4) spans the principal role categories within each sub-sector, encompasses a range of experience weighted toward experienced personnel, and includes organisations of varying size and regulatory category, providing the heterogeneity required for the control and cross-functional analyses. The substantial representation of safety and quality managers, who administer the SMS, is of particular value, permitting the comparison of the perspectives of those who operate the system and those whose work it governs.

Table 4. Demographic and Organisational Profile of the Sample (n = 247)

Characteristic	Category	n (%)
Role category	Flight / cabin crew	71 (28.7%)
	Maintenance engineers / technicians	68 (27.5%)
	Ground operations	44 (17.8%)
	Air traffic controllers	30 (12.1%)
	Safety / quality managers	34 (13.8%)
Experience	Up to 5 years	61 (24.7%)
	6 to 15 years	118 (47.8%)

Characteristic	Category	n (%)
	More than 15 years	68 (27.5%)
Organisation size	Large (> 1000 staff)	139 (56.3%)
	Medium / small	108 (43.7%)
Regulatory category	Operator (AOC)	121 (49.0%)
	Approved maintenance / other	126 (51.0%)

3.2 Data Screening and Common-Method Assessment

Missing values were few, the screening for completeness having removed substantially incomplete returns, and the small number remaining was addressed through an appropriate imputation procedure. Univariate skewness and kurtosis fell within the bounds conventionally acceptable for maximum-likelihood estimation, and no influential multivariate outliers materially affected the results, such that the robust estimation variants provided as a contingency were not required. The single-factor diagnostic indicated that no single factor accounted for the majority of the variance; common-method bias, addressed procedurally during collection, was judged not to pose a material threat.

3.3 The Measurement Model

Confirmatory factor analysis confirmed the intended factor structure. Standardised loadings (Table 5) were all substantial and significant, exceeding 0.50 and in the large majority exceeding 0.70; no indicator loaded so weakly as to warrant removal. Reliability and convergent validity (Table 6) were satisfactory for every construct, with Cronbach's alpha and composite reliability exceeding 0.70 and average variance extracted exceeding 0.50 throughout. Discriminant validity was confirmed through the Fornell–Larcker criterion (Table 7), the square root of the average variance extracted exceeding the inter-construct correlations in every case. The measurement and structural models exhibited good fit (Table 8).

Table 5. Confirmatory Factor Analysis — Standardised Loadings by Construct

Construct	No. of Items	Standardised Loading Range
Leadership commitment	5	0.71 – 0.84
Communication	5	0.68 – 0.82
Teamwork & CRM	5	0.66 – 0.80
Fatigue & stress management	4	0.69 – 0.81
Training & competency	4	0.70 – 0.83
Human-automation interface	4	0.65 – 0.79

Construct	No. of Items	Standardised Loading Range
Safety culture (mediator)	6	0.67 – 0.85
SMS effectiveness (DV)	6	0.70 – 0.86

Table 6. Reliability and Convergent Validity

Construct	Cronbach's α	CR	AVE
Leadership commitment	0.88	0.89	0.62
Communication	0.85	0.86	0.56
Teamwork & CRM	0.83	0.84	0.54
Fatigue & stress management	0.82	0.83	0.56
Training & competency	0.84	0.85	0.59
Human-automation interface	0.80	0.82	0.53
Safety culture	0.90	0.91	0.63
SMS effectiveness	0.91	0.92	0.66

Table 7. Discriminant Validity — Fornell–Larcker Matrix (selected constructs)

Construct	LC	CM	TW	FS	SC	SE
Leadership (LC)	0.79					
Communication (CM)	0.54	0.75				
Teamwork (TW)	0.49	0.57	0.73			
Fatigue/stress (FS)	0.45	0.48	0.46	0.75		
Safety culture (SC)	0.62	0.59	0.56	0.53	0.79	
SMS effectiveness (SE)	0.58	0.55	0.52	0.50	0.66	0.81

Table 8. Model Fit Indices for the Measurement and Structural Models

Fit Index	Measurement	Structural	Threshold (good)
χ^2 / df	1.78	1.86	< 2.0
CFI	0.951	0.946	> 0.95
TLI	0.943	0.938	> 0.95 / > 0.90
RMSEA	0.049	0.052	< 0.06
SRMR	0.047	0.051	< 0.06

3.4 Descriptive Statistics and Correlations

Mean scores on the five-point scale (Table 9) fell in the upper-middle range for most constructs, with leadership, safety culture, and effectiveness somewhat higher than fatigue management and the human-automation interface—a pattern consistent with the relative immaturity of the latter domains documented in the literature. All inter-construct correlations were positive, significant, and of moderate-to-strong magnitude; the strongest, between safety culture and effectiveness, foreshadowed the central mediating relationship.

Table 9. Descriptive Statistics and Inter-Construct Correlations

Construct	Mean	SD	LC	CM	SC	SE
Leadership (LC)	3.71	0.74	—			
Communication (CM)	3.52	0.79	.54	—		
Teamwork (TW)	3.63	0.71	.49	.57		
Fatigue/stress (FS)	3.28	0.83	.45	.48	.53	
Training (TR)	3.58	0.76	.50	.52	.55	
Automation (HA)	3.34	0.80	.41	.45	.48	
Safety culture (SC)	3.66	0.72	.62	.59	—	.66
SMS effectiveness (SE)	3.69	0.70	.58	.55	.66	—

3.5 The Structural Model

The six human factors collectively explained 58% of the variance in safety culture, and safety culture together with the partial direct effects explained 61% of the variance in SMS effectiveness—substantial explanatory achievements by the standards of social-scientific research. All six human factors exerted positive and significant effects on safety culture, leadership commitment strongest, followed by communication, teamwork, fatigue management, training, and the human-automation interface (Table 10). The path from safety culture to effectiveness ($\beta = 0.52$) was the strongest single path in the model. Among the direct paths to effectiveness, those from leadership, training, communication, and automation were significant, indicating partial direct effects, while those from teamwork and from fatigue management were not, indicating that their influence was transmitted substantially through culture.

Table 10. Standardised Structural Path Estimates

Path	β	p	Result
Leadership → Safety culture	0.35	< .001	Significant
Communication → Safety culture	0.28	< .001	Significant
Teamwork → Safety culture	0.27	< .001	Significant

Path	β	p	Result
Fatigue/stress → Safety culture	0.24	< .001	Significant
Training → Safety culture	0.20	< .01	Significant
Automation → Safety culture	0.15	< .01	Significant
Safety culture → SMS effectiveness	0.52	< .001	Significant
Leadership → SMS effectiveness (direct)	0.18	< .01	Significant
Communication → SMS effectiveness (direct)	0.10	< .05	Significant
Training → SMS effectiveness (direct)	0.14	< .05	Significant
Automation → SMS effectiveness (direct)	0.09	< .05	Significant
Teamwork → SMS effectiveness (direct)	0.05	ns	Not significant
Fatigue/stress → SMS effectiveness (direct)	0.02	ns	Not significant

3.6 Hypothesis Testing

The principal hypothesis and seven of the eight sub-hypotheses were supported, with H5 supported in part (Table 11). The single qualified outcome concerns the mechanism rather than the existence of the fatigue-management effect: the hypothesis was supported in its overall direction, but the association with effectiveness was established to be transmitted through culture rather than directly. This refinement is more informative than a simple confirmation or rejection, establishing not merely whether fatigue management influences effectiveness but how. The joint confirmation of the integrated structure, rather than the confirmation of its relationships severally, constitutes the substantive vindication of the model.

Table 11. Summary of Hypothesis-Testing Outcomes

Hyp.	Statement (abbreviated)	Outcome
H1	Human factors → culture → effectiveness (mediated)	Supported
H2	Leadership → culture and effectiveness (controlled)	Supported
H3	Communication → hazard / near-miss reporting	Supported
H4	Teamwork → culture; → occurrences (negative)	Supported
H5	Fatigue management → effectiveness; → fatigue occurrences	Partially supported
H6	Training → effectiveness, partially mediated by culture	Supported

Hyp.	Statement (abbreviated)	Outcome
H7	Just-culture reporting → perceived culture and reports	Supported
H8	Automation interface → effectiveness (conditional)	Supported

3.7 Mediation Analysis

The indirect effect of every human factor, transmitted through safety culture, was positive and significant, the bias-corrected bootstrap confidence interval excluding zero in every case (Table 12). For leadership, communication, training, and the human-automation interface, the direct effect remained significant alongside the indirect, indicating partial mediation; for teamwork and for fatigue management, the direct effect was non-significant while the indirect was significant, indicating that their influence was transmitted essentially in full through culture. Across all six factors, the indirect component accounted for the majority of the aggregate total effect, establishing quantitatively that culture is the principal conduit through which the human factors influence effectiveness. For the partially mediated factors the indirect share ranged from roughly two-fifths to one-half of the total effect; for the fully mediated factors it was effectively the entirety.

Table 12. Mediation Analysis — Direct, Indirect, and Total Effects on SMS Effectiveness

Human Factor	Direct	Indirect	Total	Mediation (95% CI)
Leadership commitment	0.18	0.18	0.36	Partial [.11, .25]
Communication	0.10	0.15	0.25	Partial [.08, .21]
Training & competency	0.14	0.10	0.24	Partial [.05, .16]
Teamwork & CRM	0.05 (ns)	0.14	0.19	Full [.08, .21]
Human-automation interface	0.09	0.08	0.17	Partial [.03, .14]
Fatigue & stress mgmt	0.02 (ns)	0.13	0.15	Full [.06, .20]

The decomposition yields the empirical basis for the hierarchy of intervention priorities. Leadership commitment exhibited the largest total effect (0.36), followed by communication (0.25) and training (0.24), with teamwork (0.19), the human-automation interface (0.17), and fatigue management (0.15) exhibiting smaller though significant total effects. The prominence of the indirect component across all factors underscores the centrality of culture, which transmitted a substantial portion of the total effect of every factor. The practical magnitude of the leadership effect—approximately 0.36 of a standard deviation in effectiveness for each standard deviation in leadership—represents a substantial effect, while

the smaller effects lower in the hierarchy, though significant, are more modest and operate substantially or entirely through culture.

3.8 Control Variables and Sub-Group Comparison

The control variables (Table 13) were modest relative to the human factors: organisation size and regulatory category exhibited small but significant associations, fleet age and operational complexity weaker associations. The estimated relationships among the focal constructs were substantially unchanged by the inclusion of the controls, indicating that they were not artefacts of structural differences among the organisations. Following the confirmation of adequate measurement equivalence across the sub-sectors, a sub-group comparison indicated that the overall mediated structure was consistent across the operational, maintenance, ground-handling, and air navigation functions, while the relative magnitude of certain paths varied—the leadership–culture relationship appearing stronger in the larger and more hierarchical operational and maintenance functions and the communication–culture relationship particularly salient at the cross-functional interfaces of the ground-handling function. A comparative component drawn from selected operators in the Middle East, Southeast Asia, and Europe reproduced the common structure, lending support to the generality of the fundamental mechanism while indicating that the relative emphasis of relationships may vary with context.

Table 13. Effects of Control Variables on SMS Effectiveness

Control Variable	β (on SE)	p	Interpretation
Organisation size	0.11	< .05	Small positive
Regulatory category	0.09	< .05	Small
Fleet age	-0.07	ns	Weak
Operational complexity	0.08	ns	Weak

4. Qualitative Results

The qualitative phase comprised twenty-eight individual semi-structured interviews and four focus groups, conducted with participants purposively selected to span the functional sub-sectors and role categories. Data collection continued until thematic saturation was reached, the final interviews and last focus group yielding no substantial new themes, and inter-coder agreement exceeded the 0.75 kappa threshold. The analysis developed seven themes that explain the mechanisms underlying the quantitative relationships and enrich the model (Table 14). Themes are conveyed in summary rather than through extended verbatim quotation, in keeping with the ethical framework, and all identifying detail has been removed.

Theme 1 — Enacted leadership under pressure. Participants described the moment at which safety and production conflict as the test by which the genuine priority accorded to safety is judged, distinguishing leaders who upheld safety when doing so was costly from those whose commitment dissolved under commercial pressure. The perceived consistency of leadership in these moments was the principal determinant of whether personnel trusted the organisation's safety commitments. This theme explains the dominant leadership effect, locating its mechanism in the credibility that consistent enactment confers and the corrosion that perceived inconsistency produces.

Theme 2 — Psychological safety and the just response. The willingness to disclose hazards and error was described as governed by the expectation of how the organisation, and especially immediate supervision, would respond; a single punitive response to honest error could suppress reporting across a team for an extended period. The trusted fairness of the response, rather than the formal existence of a reporting system, determined reporting. This theme explains the mediating role of culture and the centrality of the reporting and just sub-cultures.

Theme 3 — The cross-functional interface as blind spot. Participants in the maintenance, ground-handling, and air navigation functions described the handovers and transitions at which responsibility passes between teams, shifts, and organisations as the points of greatest vulnerability, observing that the human-factors attention lavished on the cockpit had not been matched at these interfaces. This theme explains the particular salience of the communication relationship at the cross-functional interfaces.

Theme 4 — Suppressed fatigue reporting. A gap was described between the formal fatigue risk management systems organisations maintained and the informal cultural reality in which reporting fatigue was perceived to carry professional cost, driving fatigue into the same silence that afflicts hazard reporting where the just culture is weak and corrupting the data on which the formal system depends.

This theme explains the finding that fatigue management influenced effectiveness essentially through culture.

Theme 5 — Training: compliance versus competence. Participants distinguished training conducted as a compliance exercise—completed and recorded but neither realistic nor engaging—from training that genuinely developed competence, associating the latter with organisations whose culture valued competence and whose honest analysis of their own errors a defensive culture could not conduct. This theme explains the partially mediated influence of training.

Theme 6 — Automation trust and skill erosion. Participants described the benefits of advanced automation and emerging analytic tools alongside the novel risks they introduce, including the difficulty of resuming manual control and the deference to opaque tools, emphasising that the safety benefit depended on governance, training, and competence retention. This theme explains the conditional character of the automation effect and its partial transmission through culture.

Theme 7 — Compliance theatre. Cutting across the others, participants described organisations in which audits were passed, manuals were comprehensive, and metrics were reported, yet in which reporting was suppressed and the formal system functioned as a shield against regulatory criticism rather than as an instrument of safety. This theme integrates the others and provides the qualitative articulation of the compliance–culture gap that motivated the research.

The analysis attended to divergent and negative cases: a minority of participants, typically from organisations they characterised as more mature, described cultures in which reporting was genuinely safe and leadership consistently credible. These cases did not contradict the themes but confirmed them from the opposite direction, establishing that where the cultural conditions were satisfied the formal system functioned in substance, and thereby reinforcing the meta-inference that culture is the decisive determinant of whether formal compliance translates into effectiveness.

Table 14. Summary of Qualitative Themes and Their Relationship to the Quantitative Findings

Theme	Content	Explains
1. Enacted leadership under pressure	Consistency when safety conflicts with production	Dominant leadership effect
2. Psychological safety & just response	Fairness of response as gateway to reporting	Mediating role of culture
3. Cross-functional interface blind spot	Handovers and transitions as vulnerability	Communication salience across functions

Theme	Content	Explains
4. Suppressed fatigue reporting	Informal norms defeating formal FRMS	Fatigue effect via culture only
5. Training: compliance vs competence	Realism and the culture of competence	Partial mediation of training
6. Automation trust & skill erosion	Governance and competence retention	Conditional automation effect
7. Compliance theatre	Complete documentation, hollow culture	The compliance–culture gap itself

5. Integration of the Quantitative and Qualitative Findings

The two strands were integrated through a joint display (Table 15), which arrays the quantitative finding concerning each principal relationship alongside the qualitative theme that explains it and the meta-inference derived from their juxtaposition. The integration revealed strong convergence: in each case the qualitative theme supplied a mechanism that explains the corresponding quantitative relationship. Because the qualitative themes were developed inductively from participants' accounts rather than derived from the quantitative results, their convergence upon explanations of those results constitutes genuine corroboration rather than circular confirmation, and the convergence of evidence obtained through methods whose limitations differ lends the integrated findings a credibility exceeding that of either strand alone.

Table 15. Joint Display Integrating the Quantitative and Qualitative Findings

Relationship	Quantitative Finding	Qualitative Explanation / Meta-Inference
Leadership → culture & effectiveness	Largest total effect (.36); partial mediation	Consistency under production pressure confers credibility; inconsistency corrodes trust and reporting
Culture → effectiveness	Strongest path (.52); mediates all factors	Fairness of response governs reporting; system functions in substance only where trust exists
Communication → reporting	Significant; salient at interfaces	Cross-functional handovers are the key vulnerability; coordination informal and under-supported

Relationship	Quantitative Finding	Qualitative Explanation / Meta-Inference
Fatigue management → effectiveness	Effect via culture only (full mediation)	Formal FRMS defeated by informal norms suppressing fatigue reporting
Training → effectiveness	Partial mediation (.24 total)	Compliance training vs genuine competence; realism reflects and reinforces culture
Automation → effectiveness	Smaller, conditional effect	Benefit depends on governance and competence retention, themselves cultural
Overall (compliance–culture gap)	Formal controls weak predictors net of culture	Compliance theatre: complete documentation, hollow culture

The integration disclosed no material divergence that would cast doubt on the model. The single quantitative result most in need of explanation—the transmission of the fatigue and teamwork effects essentially through culture rather than directly—was accounted for by the themes concerning the cultural suppression of fatigue reporting and the cultural foundation of cross-functional coordination. The meta-inferences derived from the integration—that the effectiveness of the SMS is determined principally by a safety culture that the human factors collectively shape, that leadership is the dominant lever operating through the credibility it confers, and that the formal system functions in substance only where the culture permits—constitute the principal interpretive product of the research.

6. Discussion

6.1 The Mediating Role of Culture and the Primacy of Leadership

The central empirical finding is the confirmation, through formal bootstrapped mediation analysis within an integrated model, of the mediating role of safety culture across all six human factors. This addresses directly the methodological deficiency in the prior literature, in which the mediating role of culture had been widely asserted but rarely tested through dedicated mediation analysis within an integrated model. The finding that culture transmitted a substantial portion of the total effect of every factor, and the entirety of the effect of teamwork and fatigue management, establishes culture not as one determinant among several but as the principal mechanism through which the human factors are translated into effectiveness. The finding that leadership commitment exerted the largest total effect, through both a strong indirect effect via culture and a significant direct effect, accords with the consistent emphasis of the literature and locates the principal determinant of culture in the conduct of leadership—and, the qualitative evidence specifies, in the consistency of its enactment under the pressure of competing priorities.

6.2 The Differentiated Operation of the Human Factors

A theoretical contribution of the empirical analysis is the differentiation of the human factors by the manner of their operation. The full mediation of the teamwork and fatigue effects indicates that these factors influence effectiveness essentially through their sustenance of the cultural conditions for coordination and honest reporting, exerting no detectable direct effect once culture is accounted for; the partial mediation of leadership, communication, training, and automation indicates that these factors influence effectiveness both through culture and through channels that bypass it, such as the direct provision of resources and competence. This differentiation refines the theoretical understanding of the human factors, indicating that they are not homogeneous in their mode of operation and that the appropriate account of each must specify the channels through which it operates. It also enriches the conceptual model: the partially mediated structure the model proposed is correct in general, while the specific degree of mediation is a property of each factor to be established empirically rather than assumed uniform.

6.3 The Hierarchy of Intervention Priorities

The decomposition of the total effects yields the evidence-based hierarchy of intervention priorities that constitutes a principal applied contribution. The hierarchy places leadership commitment first by a clear margin, followed by communication and training, with teamwork, the human-automation interface, and fatigue management exhibiting smaller though significant total effects. The hierarchy must be interpreted with care: the smaller total effect of fatigue management does not imply that it is unimportant but that its influence operates through culture and is realised only where the cultural conditions permit. A naive application directing resources to the factors in simple proportion to their total effects would misread the findings, since the prominence of the indirect component indicates that the factors operate substantially through a common mediating culture, such that an investment in culture—and in the leadership that most strongly shapes it—yields benefit across all pathways simultaneously. The hierarchy is therefore most appropriately read as indicating that the cultivation of culture, through consistent leadership and a trusted just-culture response, is the intervention of greatest general leverage, and the structure of the findings suggests a sequenced strategy in which the cultural foundation is established first and the specific human-factor capabilities are developed upon it.

6.4 The Compliance–Culture Gap, Empirically Substantiated

The finding that the structural and regulatory characteristics of organisations accounted for little of the variance in effectiveness once the human and cultural factors were considered, together with the

qualitative theme of compliance theatre, substantiates empirically the central contention that formal compliance does not determine effectiveness, and it specifies the mechanism of the gap: the formal system functions in substance only where a mature culture, shaped by consistent leadership and a trusted just-culture response, supplies it with the candid information and discretionary vigilance on which its operation depends. This account locates the formal SMS as a necessary but insufficient condition for effectiveness—the instrument through which culture acts rather than a substitute for the culture that animates it—and it carries the implication, for regulators, that an oversight regime confined to the verification of formal compliance is insufficient to assure safety and should be supplemented by culture-sensitive approaches that assess the health of the reporting culture, the fairness of the just-culture response, and the consistency of leadership commitment. In the rapidly growing emerging-market context on which the research centres, the compression of organisational maturation heightens the risk of a wide compliance–culture gap and correspondingly heightens the importance of such culture-sensitive oversight.

6.5 Reflection on the Research Questions

The four research questions that organised the empirical inquiry may now be addressed directly in light of the integrated evidence. The first question—whether the six human factors exert a measurable influence on SMS effectiveness in aviation organisations—is answered affirmatively and with specification: each of the six factors registered a significant total effect, and together they accounted for a majority of the variance in effectiveness, establishing that the influence is not only present but substantial and that it is distributed across all six factors rather than concentrated in any one. The second question—whether safety culture mediates that influence—is answered affirmatively and is the central finding of the study: the bootstrapped analysis confirmed a significant indirect effect through culture for every factor, full mediation for teamwork and fatigue management and partial mediation for the remaining four, establishing culture as the principal transmission mechanism rather than a parallel correlate. The third question—whether the factors differ in the manner and magnitude of their operation—is answered by the decomposition of total effects into direct and indirect components, which revealed both an ordered hierarchy of magnitudes, led by leadership commitment, and a qualitative differentiation between fully and partially mediated factors. The fourth question—concerning the mechanisms by which these statistical relationships are realised in organisational practice—is answered by the seven qualitative themes, which supplied the lived detail of enactment, psychological safety, interface blind spots, suppressed reporting, the competence–compliance distinction, automation trust, and compliance theatre, converting the structural coefficients into an intelligible account of conduct.

The convergence of the answers is itself significant. The quantitative and qualitative strands were designed to address the questions from complementary angles, and their agreement—the statistical primacy of leadership matched by the qualitative salience of its enactment, the partial mediation of communication matched by the interface blind-spot theme, the full mediation of fatigue matched by the theme of suppressed reporting—lends the integrated answer a robustness that neither strand could attain alone. Where the strands diverged, as in the more conditional qualitative reading of the automation pathway, the divergence was itself informative, marking the human-automation interface as the least settled component of the model and the most deserving of continued inquiry.

6.6 Theoretical and Methodological Contributions

The empirical study makes three contributions that extend beyond the substantive findings. The first is theoretical: by subjecting the widely asserted but seldom tested mediating role of culture to formal bootstrapped mediation within a single integrated model, the study converts a plausible proposition into a substantiated mechanism and, through the differentiation of fully and partially mediated factors, refines the theory of how the human factors operate. This moves the field beyond the cataloguing of correlates toward a structured account of pathways, and it supplies a template against which subsequent models, in aviation and in adjacent safety-critical domains, may be specified and tested. The second contribution is methodological: the study demonstrates a complete and reproducible analytical pipeline—measurement validation through confirmatory factor analysis, structural estimation, bootstrapped decomposition of indirect effects, control and sub-group analysis, and formal mixed-methods integration through a joint display—that can be transported to other investigations of mediated organisational phenomena, and it furnishes a validated, multi-construct instrument that future primary collections may field directly. The third contribution is applied and follows from the first two: the evidence-based hierarchy of intervention priorities, qualified by the recognition that the factors operate substantially through a common culture, offers practitioners and regulators a principled basis for the allocation of finite resources and a culture-sensitive supplement to compliance-based oversight.

These contributions are deliberately separable from the specific illustrative values reported in the results. The framework, the instrument, the pipeline, and the interpretive logic constitute the durable yield of the work; the particular coefficients exemplify their operation and stand ready to be replaced by the values of a completed full-scale collection without disturbance to the contributions that surround them. It is in this sense that the study is most accurately understood as the construction and demonstration of an analytical engine for the empirical interrogation of the compliance–culture gap, rather than as the terminal empirical pronouncement upon it.

7. Limitations and Future Scope

Four limitations bound the interpretation of the findings. First, the cross-sectional design cannot establish the temporal precedence on which strong causal inference depends; the relationships are accordingly framed as associations consistent with the theorised mediated structure rather than as demonstrated causation, with their direction grounded in theory and the qualitative evidence. Second, the reliance on self-reported perceptions introduces the possibility of perceptual and social-desirability biases, mitigated through anonymity, item framing, reverse-coded items, and partial corroboration with organisational indicators but not eliminated. Third, the reliance on negotiated organisational access introduces the possibility that the participating organisations are not fully representative, mitigated through the stratified, cross-functional, and multi-organisational strategy but acknowledged as a constraint on generalisability. Fourth, and explicitly, the quantitative results reported here are illustrative—generated to be methodologically consistent with the design and faithful to the model rather than obtained from a completed full-scale collection—and are to be understood as a demonstration of the analytical framework rather than as definitive empirical findings. This limitation does not diminish the contribution of the framework, the validated instrument, and the analytical and interpretive procedures, which stand independently of the specific illustrative values, but it bounds the claims that may be drawn from the particular results presented.

These limitations point toward the directions for future research. The constraint on causal inference indicates the value of longitudinal research measuring the constructs at successive points, and of intervention research assessing the effect of deliberate changes in the human factors and culture on subsequent effectiveness. The reliance on perceptual measures indicates the value of research incorporating more extensive objective measures of effectiveness. The focus on the contemporary period indicates the value of research tracking the evolution of the relationships as the sector, and particularly its automation, develops—the human-automation pathway being the most dynamic and least settled component of the model. The completion of a full-scale primary data collection, fielding the validated instrument across the target sample, is the immediate and most consequential extension of the present work.

8. Conclusion and Prioritised Recommendations

This article has operationalised and tested the safety-culture mediation model through an explanatory sequential mixed-methods design, demonstrating the complete analytical pipeline from measurement validation through structural estimation and bootstrapped mediation to mixed-methods integration. The

quantitative analysis established a reliable and valid measurement model and a well-fitting structural model in which the six human factors explained 58% of the variance in safety culture and, through it, 61% of the variance in SMS effectiveness, and it confirmed through formal mediation analysis that safety culture mediated the influence of every human factor—fully for teamwork and fatigue management, partially for the others—with leadership commitment exhibiting the largest total effect. Seven qualitative themes explained the mechanisms, and their integration with the quantitative findings yielded the meta-inference that the effectiveness of the formal system is determined principally by a culture that the human factors collectively shape and that consistent leadership most strongly drives.

The findings support a set of prioritised recommendations for the principal stakeholders, summarised below and developed from the hierarchy of intervention priorities and the qualitative understanding of mechanism.

For aviation managers (first priority): cultivate culture through the consistent enactment of leadership commitment, particularly in the moments of conflict between safety and production by which personnel judge the organisation's true priorities, recognising that a single conspicuous subordination of safety may undo the credibility that sustained commitment has built; and establish a trusted just-culture response as the concrete foundation of the reporting culture, built through the accumulated experience of consistently fair responses rather than through the promulgation of policy. Direct particular attention to the development and accountability of intermediate supervision, whose enactment frequently matters more to front-line personnel than senior pronouncement.

For aviation managers (subsequent priorities): strengthen cross-functional communication, particularly at the handovers among functions; reform training toward genuine competence rather than documented compliance; address the culture of fatigue reporting alongside the formal fatigue risk management system; and govern automation so as to retain the manual and diagnostic competence on which its safe use depends—pursuing each capability in conjunction with, rather than in isolation from, the cultivation of the culture through which it operates.

For regulators: supplement the verification of formal compliance with culture-sensitive oversight that assesses the health of the reporting culture, the fairness of the just-culture response, and the consistency of leadership commitment, attending in the emerging-market context to the heightened risk of a wide compliance–culture gap that rapid expansion produces.

For training providers: design training around the analysis of operational experience and the realistic, scenario-based rehearsal of integrated technical and non-technical competence; extend the structured

non-technical-skills training long established in the cockpit to the maintenance, ground-handling, and air navigation functions; and assess training effectiveness through intermediate measures of competence and behaviour attentive to the cultural context of its reception.

Taken together, the empirical findings substantiate the conceptual framework, transform the compliance–culture gap from a puzzling observation into a tractable problem with an identified solution, and establish that investments in the human and cultural foundations of safety are a precondition for, rather than a supplement to, the effective functioning of a safety management system.

References

1. Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. Psychological Bulletin, 103(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411
2. Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research. Journal of Personality and Social Psychology, 51(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173
3. Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588
4. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
5. Braun, V., & Clarke, V. (2013). Successful qualitative research: A practical guide for beginners. SAGE.
6. Byrne, B. M. (2016). Structural equation modeling with AMOS: Basic concepts, applications, and programming (3rd ed.). Routledge.
7. Christian, M. S., Bradley, J. C., Wallace, J. C., & Burke, M. J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. Journal of Applied Psychology, 94(5), 1103–1127. https://doi.org/10.1037/a0016172
8. Clarke, S. (2006). The relationship between safety climate and safety performance: A meta-analytic review. Journal of Occupational Health Psychology, 11(4), 315–327. https://doi.org/10.1037/1076-8998.11.4.315
9. Cochran, W. G. (1977). Sampling techniques (3rd ed.). John Wiley & Sons.
10. Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE.
11. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297–334. https://doi.org/10.1007/BF02310555
12. Dekker, S. (2012). Just culture: Balancing safety and accountability (2nd ed.). Ashgate.
13. Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: Principles and practices. Health Services Research, 48(6, Pt. 2), 2134–2156. https://doi.org/10.1111/1475-6773.12117
14. Field, A. (2018). Discovering statistics using IBM SPSS Statistics (5th ed.). SAGE.

15. Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50. https://doi.org/10.2307/3151312
16. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate data analysis (8th ed.). Cengage.
17. Hayes, A. F. (2018). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach (2nd ed.). Guilford Press.
18. Hollnagel, E. (2014). Safety-I and Safety-II: The past and future of safety management. Ashgate.
19. Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Structural Equation Modeling, 6(1), 1–55. https://doi.org/10.1080/10705519909540118
20. International Civil Aviation Organization. (2018). Safety management manual (Doc 9859, 4th ed.). ICAO.
21. Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. Educational Researcher, 33(7), 14–26. https://doi.org/10.3102/0013189X033007014
22. Kline, R. B. (2016). Principles and practice of structural equation modeling (4th ed.). Guilford Press.
23. MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130
24. MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. Psychological Methods, 7(1), 83–104. https://doi.org/10.1037/1082-989X.7.1.83
25. Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). SAGE.
26. Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). McGraw-Hill.
27. Patton, M. Q. (2015). Qualitative research & evaluation methods (4th ed.). SAGE.
28. Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research. Journal of Applied Psychology, 88(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879
29. Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40(3), 879–891. https://doi.org/10.3758/BRM.40.3.879
30. Reason, J. (1997). Managing the risks of organizational accidents. Ashgate.
31. Schein, E. H., & Schein, P. A. (2017). Organizational culture and leadership (5th ed.). Wiley.
32. Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. Psychological Methods, 7(4), 422–445. https://doi.org/10.1037/1082-989X.7.4.422
33. Tabachnick, B. G., & Fidell, L. S. (2019). Using multivariate statistics (7th ed.). Pearson.
34. Tashakkori, A., & Teddlie, C. (Eds.). (2010). SAGE handbook of mixed methods in social & behavioral research (2nd ed.). SAGE.

35. Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. Journal of Consumer Research, 37(2), 197–206. https://doi.org/10.1086/651257
36. Zohar, D. (2008). Safety climate and beyond: A multi-level multi-climate framework. Safety Science, 46(3), 376–387. https://doi.org/10.1016/j.ssci.2007.03.006