Open access: https://journal.icpa-banyuwangi.ac.id/index.php/skyhawk

SKYHAWK: Jurnal Aviasi Indonesia
Volume. 6, Number. 2, 2026, pp. 787-797

P-ISSN: 2745-6102
E-ISSN: 2776-3854

Cognitive Pathways to Aviation Safety: A Systematic Literature Review of Decision-Making, Situation Awareness, Workload, and Fatigue

Dani Chandra Yudho Pranoto[1], Ahmad Bahrawi[2], Adi Bandono[3], Wawan Gunawan[4], Yew Yun Xin[5]
[12]ATS School, Surabaya Aviation Polytechnic, Surabaya, Indonesia
[34]Post Graduate Educational Technology, PGRI Adi Buana University, Surabaya, Indonesia
[5]International Languages Teacher Training Institute, Malaysia

Email Address: ahmadbahrawi@poltekbangsby.ac.id[2], adibandono@unipasby.ac.id[3], wawan.gunawan@unipasby.ac.id[4], yunxinyew1012@gmail.com[5]
Email Correspondence: dcypranoto@poltekbangsby.ac.id[1]

ARTICLE INFO

Article History

Received: 08 May 2026
Revised: 12 june 2026
Accepted: 23 September 2026

Keywords

Aviation safety, Decision-making, Fatigue, Mental workload, Situation awareness

ABSTRACT

Cognitive determinants are central to aviation safety because they shape how operators perceive information, allocate attention, manage workload, maintain situation awareness, and make decisions under dynamic conditions. This systematic literature review synthesized evidence using a PRISMA-oriented screening and full-text eligibility process. From 5,421 Scopus records, 394 entered title–abstract–keyword screening, and 120 studies were included in the final qualitative synthesis. The evidence was classified into four primary domains: decision-making, risk perception, and cognitive control (n = 29); situation awareness, attention allocation, and visual monitoring (n = 36); mental workload, task complexity, and psychophysiological monitoring (n = 24); and fatigue, sleepiness, stress, and cognitive resilience (n = 31). Findings show that these domains interact dynamically: fatigue and stress reduce cognitive resources, workload constrains attention, degraded attention weakens situation awareness, and impaired situation awareness compromises decision-making and safety performance. The review proposes an integrated cognitive safety perspective linking cognitive state, attentional capacity, situation awareness, decision-making, and safety performance within a socio-technical framework.

1. INTRODUCTION

Aviation is a safety-critical socio-technical domain in which safe performance depends not only on technical reliability but also on how operators perceive, interpret, prioritize, and act on information under changing operational conditions. Pilots, air traffic controllers, maintenance personnel, and other aviation professionals must continually manage limited cognitive resources while monitoring complex systems and responding to uncertainty. Four domains are especially relevant: decision-making, situation awareness, mental workload, and fatigue. Situation awareness describes how operational information is perceived, comprehended, and projected, while Multiple Resource Theory explains how performance can deteriorate when concurrent demands compete for limited perceptual and cognitive resources (Endsley, 1995; Wickens, 2008). Fatigue adds a time-dependent constraint through sleep history, circadian timing, duty structure, workload, and recovery opportunities (Coombes et al., 2020; Hilditch et al., 2023).

Contemporary aviation human-factors research increasingly treats these processes as dynamic and embedded within human-system interaction rather than as isolated sources of error. Operators may adapt through prioritization, monitoring, communication, strategy adjustment, and appropriate use of automation. This is compatible with resilience-oriented safety thinking, which examines how people and systems sustain acceptable performance under variability and uncertainty (Malakis & Kontogiannis, 2023). The same perspective is relevant to Safety Management Systems because cognitive performance is shaped by interfaces, automation, training, staffing, scheduling, reporting conditions, and opportunities for organizational learning (Malakis et al., 2023).

The evidence nevertheless remains fragmented across research traditions. Decision-making research examines risk appraisal, authority, experience, and cognitive control; situation-awareness research emphasizes monitoring and visual attention; workload studies increasingly combine subjective and physiological indicators; and fatigue research addresses sleep, circadian disruption, psychosocial demands, and recovery. Emerging eye tracking, cardiovascular sensing, neurophysiology, and machine-learning methods expand measurement capability but do not eliminate questions of construct validity or operational transferability. A recent SKYHAWK study on stress among fixed-wing cadets further illustrates the continuing relevance of psychological state in aviation training environments (Islam et al., 2024).

This review therefore synthesizes recent evidence across the four domains within a common analytical structure. It asks: (1) what relationships between the four cognitive domains and safety-relevant performance are reported; (2) how the domains are conceptualized and measured and which cross-cutting mechanisms connect them; (3) what methodological patterns and evidential limitations characterize the literature; and (4) how recurring relationships can be integrated into an evidence-informed framework relevant to proactive safety management. The term cognitive pathways is used here to describe recurring configurations of relationships across heterogeneous evidence, not a single causal sequence assumed to operate uniformly across aviation contexts.

2. METHODS

Search Strategy and Scope

The review followed PRISMA 2020 principles for transparent identification, screening, eligibility assessment, and reporting (Page et al., 2021). The literature search was conducted exclusively in Scopus on 7 May 2026 using three conceptual blocks: *("human factors" OR "human* *performance" OR "human error") AND ("aviation safety" OR "flight safety" OR "airline safety") AND ("cognitive factors" OR cognition OR "situational awareness" OR "decision making" OR "mental workload" OR fatigue OR stress OR attention OR "pilot workload")*. The initial search returned 5,421 records. Sequential restrictions produced 2,889 records for 2020 through 7 May 2026, 1,753 articles, 604 Social Sciences records, 399 records after operational aviation actor terms, and 394 English-language journal records. These restrictions define the review as a synthesis of Scopus-indexed literature captured by the predefined search rather than an exhaustive representation of global aviation cognitive-human-factors research.

Eligibility, Extraction, and Coding

Eligible studies addressed an aviation-relevant actor, system, or operational context; examined a focal or closely related cognitive determinant; and reported a cognitive measure or proxy, behavioral or operational performance, or another safety-relevant outcome. Designs could include simulator experiments, surveys, field studies, qualitative research, interventions, mixed methods, modelling, or conceptually oriented studies with interpretable methods. After title, abstract, and keyword screening, 265 records were excluded and 129 reports were sought for retrieval; two could not be retrieved. Formal full-text assessment was completed for 126 reports, with one additional documented abstract-based eligibility exception. Seven full-text reports were excluded, leaving 120 studies: 119 full-text-confirmed studies and one abstract-based exception (Figure 1).

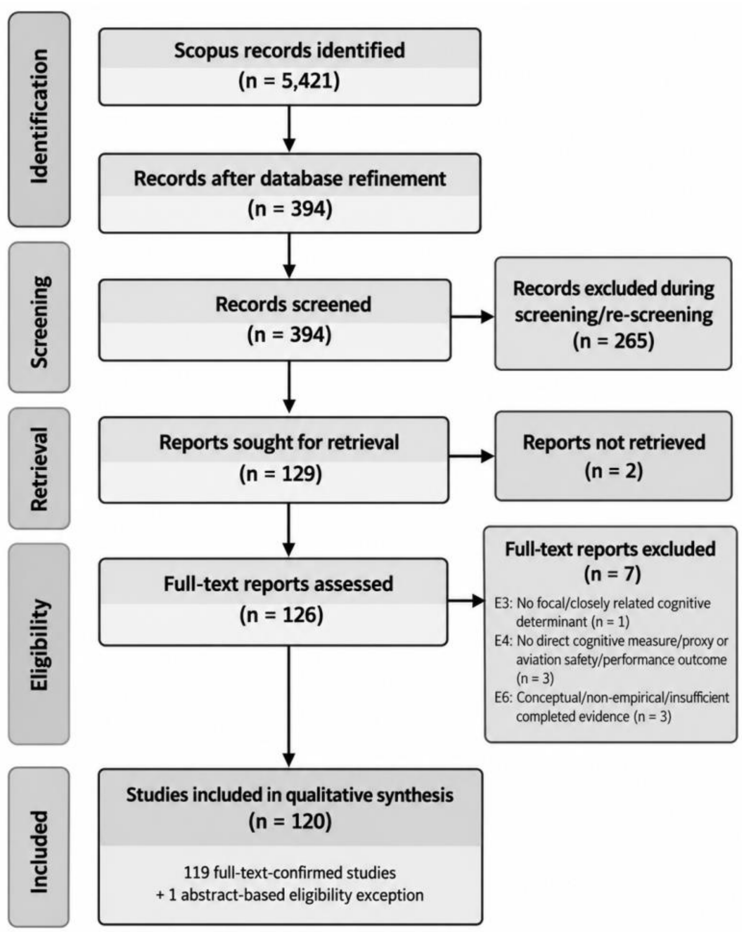

Figure 1. PRISMA-oriented study-selection flow diagram

Data were extracted into a structured evidence matrix and each study was assigned to one mutually exclusive primary domain: decision-making, risk perception, and cognitive control; situation awareness, attention allocation, and visual monitoring; mental workload, task complexity, and psychophysiological monitoring; or fatigue, sleepiness, stress, and cognitive resilience. Cross-cutting mechanisms such as attention, expertise, stress, human-automation interaction, risk perception, resilience, communication, cognitive control, vigilance, and memory were coded using multi-label classification. Measurement approaches were grouped as subjective, behavioral/operational, ocular/visual, physiological/neurophysiological, and computational. Physiological, ocular, and algorithmic outputs were interpreted as indirect and context-sensitive indicators rather than direct measurements of latent cognitive constructs.

Quality Appraisal and Synthesis

Because the evidence base was heterogeneous, methodological quality was appraised using design-sensitive criteria rather than a single numerical score. Criteria covered methodological transparency, design appropriateness, sample or data adequacy, construct validity, measure-construct fit, analytical appropriateness, ecological validity, consistency between evidence and conclusions, and operational transferability. Findings were synthesized narratively rather than statistically pooled because effect estimates, populations, measures, and outcomes were not sufficiently comparable, consistent with synthesis-without-meta-analysis guidance (Campbell et al., 2020). Study frequency was treated as research volume, not evidential strength. Cross-domain relationships were described as associations or evidence-informed conceptual linkages unless the original study explicitly tested a causal or mediational effect. A leave-one-out sensitivity analysis excluding the abstract-based exception reduced the corpus to 119 studies and the fatigue/resilience domain from 31 to 30 studies without materially changing the domain structure or principal synthesis.

3. RESULTS

Overall Evidence Profile

The final corpus comprised 120 studies: situation awareness, attention allocation, and visual monitoring (n = 36, 30.0%); fatigue, sleepiness, stress, and cognitive resilience (n = 31, 25.8%); decision-making, risk perception, and cognitive control (n = 29, 24.2%); and mental workload, task complexity, and psychophysiological monitoring (n = 24, 20.0%). Sixty-three studies (52.5%) were published between 2024 and the search date of 7 May 2026; because 2026 was only a partial publication year, this reflects recent activity rather than a complete annual trend. Simulator or experimental studies accounted for 59 studies (49.2%), followed by survey/cross-sectional designs (n = 22), modelling/computational studies (n = 12), and operational/field studies (n = 9). Behavioral or performance indicators appeared in 112 studies, computational/model-based approaches in 51, self-report in 37, gaze measures in 18, cardiovascular measures in 11, and operational/incident data in 10. These categories were multi-label.

Cross-cutting coding showed that attention or monitoring appeared in 47 studies, experience or expertise in 37, human error or unsafe behavior in 29, stress or emotion in 28, and human-automation interaction in 26. Other recurring mechanisms included risk perception (n = 16), resilience or mindfulness (n = 15), communication or team cognition (n = 12), cognitive control or self-regulation (n = 11), vigilance (n = 8), and working or prospective memory (n = 6). These frequencies describe the distribution of research attention and do not rank causal importance, methodological quality, or strength of evidence. Table 1 integrates the main evidence patterns and interpretive boundaries.

Table 1. Integrated cross-domain synthesis of cognitive human factors in aviation safety (N = 120)

Domain	Evidence profile and mechanisms	Recurring synthesis	Interpretive boundary
Decision-making, risk perception, and cognitive control	n = 29. Simulator/experimental 11; survey 6; modelling 4; mixed 3; intervention 2; other 3. Main mechanisms: error/unsafe behavior 14, risk perception 13, expertise 11, attention 9, cognitive control 8.	Decision performance was repeatedly associated with information acquisition and situation assessment, risk appraisal, reflection and self-regulation, experience, working-memory demand, authority, communication, and the structure of training or decision support (Behrend & Dehais, 2020; Fabre et al., 2022; Xie et al., 2025).	Heterogeneous designs do not support one corpus-wide causal decision pathway. Causal or mediation claims are retained only where directly tested.
Situation awareness, attention allocation, and visual monitoring	n = 36. Simulator/experimental 26 (72.2%); behavioral/performance 35; gaze 15; computational/model-based 13. Main mechanisms: attention 28, expertise 12, human-automation interaction 11, memory 5.	Awareness-related performance depended on cue detection, allocation of attention, information integration, updating, and projection. Distraction and competing demands were associated with missed information, whereas trained scanning and supportive interfaces could improve monitoring (Lefrançois et al., 2021; White & O'Hare, 2022; Xie et al., 2024).	Strong simulator concentration limits direct operational generalization. Gaze indicates visual allocation and monitoring behavior but does not independently establish comprehension, projection, or situation awareness.
Mental workload, task complexity, and psycho-physiological monitoring	n = 24. Simulator/experimental 16 (66.7%); self-report 9; cardiovascular 7; computational 10; EEG 2; fNIRS 1. Main mechanisms: human-automation interaction 10, attention 6, expertise 4.	Workload emerged as a dynamic relationship among external task demand, available resources, strategy, automation, environment, and expertise. Design should calibrate rather than simply minimize workload (Bromfield et al., 2023; Wang et al., 2024; Berthon et al., 2025).	Physiological, neurophysiological, ocular, and computational outputs are indirect and context-sensitive indicators. Classification accuracy does not establish construct validity or operational diagnostic validity.
Fatigue, sleepiness, stress, and cognitive resilience	n = 31. Survey/cross-sectional 14; simulator/experimental 6; modelling 4;	Fatigue and stress were associated with interacting sleep, circadian, schedule, workload, recovery,	Fatigue indicators are heterogeneous and not interchangeable. Individual resilience

	operational/field 4; qualitative 2; intervention 1. Main mechanisms: stress/emotion 16, resilience 12, expertise 10.	psychosocial, and organizational conditions. Recovery, training, reporting, social support, and system controls may support cognitive readiness (Hilditch et al., 2023; Liu et al., 2023; Signal et al., 2023).	strategies should complement rather than replace organizational fatigue controls.
Cross-domain synthesis	N = 120. Simulator/experimental 59 (49.2%); survey 22; modelling 12; operational/field 9. Leading cross-cutting mechanisms: attention 47, expertise 37, error/unsafe behavior 29, stress/emotion 28, human-automation interaction 26.	Cognitive readiness, task demand, attention and monitoring, situation representation, risk appraisal, and decision processes appeared as interdependent components shaped by expertise, automation, communication, training, resilience, system design, and organizational context.	This is an evidence-informed relational synthesis of heterogeneous evidence, not a universal causal sequence. Study frequency indicates research volume rather than evidential strength.

Domain-Specific Findings

Decision-making evidence showed that operational choice was related to the quality and organization of information, social influence, risk appraisal, and cognitive control. Behrend and Dehais (2020) associated cockpit role assignment with different monitoring patterns and decision timing, while Fabre et al. (2022) showed that less-experienced first officers were more likely to follow a risky captain in simulated landing decisions. Xie et al. (2025) reported a study-specific serial pathway linking impulsivity, risk perception, situation awareness, and incident involvement. These findings support decision-making as a convergence point for information state, experience, authority, perceived risk, communication, and self-regulation rather than a purely procedural final choice.

Situation-awareness evidence was the most experimentally concentrated: 26 of 36 studies used simulator or experimental designs. Distraction, working-memory limits, interface characteristics, expertise, and automation repeatedly modified monitoring performance. Xie et al. (2024) found that greater secondary-task difficulty was associated with poorer perceptual and overall situation-awareness performance, whereas White and O'Hare (2022) demonstrated inattentional blindness during competing simulated-flight tasks. Visual-attention training can also be modified deliberately; Lefrançois et al. (2021) reported improved scanning and subsequent manual-flight performance after training on skilled gaze strategies. Eye tracking, however, indicates where and how operators look rather than whether information has been accurately comprehended or projected.

Workload evidence distinguished external task demand from the operator's workload response. Task phase, time pressure, information structure, environment, and interface design altered subjective and physiological responses, but higher workload was not invariably unsafe. Bromfield et al. (2023) reported improved performance with an angle-of-attack display despite a moderate increase in reported workload, while Wang et al. (2024) used wearable pulse-rate variability to examine workload variation, Mattingsdal et al. (2025) examined helicopter rescue workload, and Berthon et al. (2025) combined subjective, performance, and cardiovascular measures in helicopter maintenance. The stronger inference across studies was workload calibration: demand should remain compatible with monitoring, decision-making, and reserve capacity rather than be minimized indiscriminately.

Fatigue and resilience studies were less dependent on simulators and more often used surveys, field observations, modelling, and sleep-related measures. Fatigue was not reducible to sleep loss alone. Hilditch et al. (2023) identified interacting circadian, workload, rest, schedule, and policy conditions among short-haul pilots; Signal et al. (2023) documented fatigue, sleep opportunity, workload, and performance across ultra-long-range operations; and Liu et al. (2023) connected family-work conflict, fatigue, and safety behavior. Physiological activation also should not be equated automatically with unsafe stress, because arousal can accompany effective engagement under demanding conditions (Vallès-Català et al., 2021). Resilience-related strategies, including psychophysiological self-regulation approaches, may support recovery and regulation (Li et al., 2023), but organizational scheduling, workload design, reporting, and recovery controls remain necessary.

Cross-Domain Synthesis

Across domains, four recurring functional relationships were evident: cognitive readiness was associated with vigilance and resource availability; workload and task complexity were associated with allocation of those resources; attention, memory, monitoring, and human-automation interaction were associated with acquisition and integration of information; and situation representation, risk appraisal, cognitive control, and communication were associated with decision performance. These relationships were reciprocal and context dependent rather than sequential. Decision uncertainty can increase perceived workload, automation can reduce one demand while creating another, and prolonged demand can intensify fatigue. The corpus therefore supports an evidence-informed relational framework in which cognitive processes interact within a broader socio-technical context rather than a validated fatigue-to-workload-to-attention-to-situation-awareness-to-decision causal chain.

4. DISCUSSION

Integrated Interpretation

The principal interpretation is that aviation cognitive safety is better understood as an interacting, context-dependent system than as a collection of independent risk variables. Established theories remain useful within this synthesis: Endsley's model explains perception, comprehension, and projection processes in situation awareness, while Multiple Resource Theory explains interference when concurrent activities compete for limited resources (Endsley, 1995; Wickens, 2008). The review extends these models by placing them in a shared operational context where workload can constrain monitoring, monitoring requirements can generate workload, fatigue can reduce vigilance, demanding operations can intensify fatigue, and uncertainty can redirect attention and risk appraisal. Cognitive vulnerability therefore does not translate mechanically into unsafe performance because operators may compensate through prioritization, communication, procedural support, strategy adjustment, and appropriate automation use.

Experience and technology were similarly conditional modifiers. Experience can support pattern recognition, mental models, scanning, and risk interpretation, yet it does not guarantee safer action under atypical events or poorly calibrated confidence. Automation can lower selected task demands while increasing monitoring, mode-awareness, or trust requirements. The relevant safety question is therefore whether the combination of operator state, task demands, interface characteristics, automation behavior, and available adaptive resources remains compatible with effective performance.

Evidence Quality and Ecological Validity

Research volume, methodological quality, consistency, and evidential strength should not be conflated. Study count indicates where research attention has accumulated; methodological quality concerns design, sampling, construct definition, measurement, analysis, and reporting; consistency concerns whether comparable studies show recurring patterns; and evidential strength additionally depends on directness, ecological validity, and transferability. Simulator or experimental studies represented 49.2% of the corpus, including 72.2% of situation-awareness studies and 66.7% of workload studies, whereas only nine studies were operational or field-based. Controlled simulation is not intrinsically weak evidence: it is often the safest and most rigorous way to manipulate distraction, abnormal events, task demand, or automation. Its limitation is direct generalization to line operations, diverse organizational cultures, prolonged exposure, and real operational consequences. Field studies offer greater contextual realism but may provide less experimental control. The two forms of evidence are therefore complementary.

Heterogeneity also limited quantitative synthesis. The reviewed studies differed in aviation populations, simulator or aircraft environments, task scenarios, measurements, and safety outcomes. Situation awareness, workload, and fatigue were each operationalized with multiple non-equivalent measures. Convergence across modalities can increase confidence in a recurring pattern, but it does not make the measures interchangeable. The review therefore did not create a pooled evidence hierarchy or infer strength from frequency. Greater cross-context replication and standardization are required before larger subsets of studies can support meaningful effect-size synthesis.

Measurement and Cognitive-State Monitoring

The growth of eye tracking, cardiovascular sensing, EEG, fNIRS, wearable devices, and machine-learning analysis creates new opportunities for time-resolved assessment, but these technologies do not provide direct readouts of workload, fatigue, stress, attention, or situation awareness. Eye tracking describes visual allocation and can support visual-behavior modelling (Lyu et al., 2024); HRV, PRV, heart rate, and electrodermal activity reflect autonomic activation influenced by multiple processes; EEG represents electrical brain activity; and fNIRS reflects hemodynamic change. Their relationship to latent cognitive constructs is therefore many-to-many. Similarly, high machine-learning classification accuracy within one dataset demonstrates discrimination under that study's labels and conditions, not universal construct validity or operational diagnostic validity. Generalization requires external testing across participants, tasks, aircraft or simulator environments, and operational roles.

Triangulation is consequently central. A physiological or ocular change becomes more interpretable when it converges with task demand, perceived state, behavioral performance, or operational context. Qin et al. (2021), for example, combined cardiovascular and ocular information in fatigue-state classification, while multimodal workload studies increasingly combine wearable and behavioral evidence (Wang et al., 2024). Cognitive-state technology is most defensible as human-in-the-loop decision support for research, training, interface evaluation, or validated fatigue-risk applications rather than as a stand-alone fitness judgment. Implementation also requires governance of privacy, data ownership, consent, explainability, false alarms, and the risk of punitive interpretation.

Safety-II, Resilience, and Safety Management Systems

The relational framework is compatible with Safety-II and resilience perspectives because cognitive human factors are relevant not only when error occurs but also when operators successfully adapt to variability. Adaptation, however, should not be reframed as an individual duty to compensate indefinitely for adverse conditions. Resilience emerges from interactions among individuals, teams, technology, procedures, scheduling, recovery opportunities, training, and organizational support (Malakis & Kontogiannis, 2023). Safety Management Systems provide the organizational mechanism for translating this evidence into hazard identification, occurrence analysis, operational monitoring, learning, training, risk control, and system redesign (Malakis et al., 2023). Fatigue Risk Management Systems apply the same principle specifically to circadian timing, roster design, recovery, reporting, validated predictive tools, and operational judgment. Safety-II and resilience therefore provide an interpretive perspective on adaptive performance, while SMS and FRMS provide mechanisms for action.

Limitations and Research Priorities

Fatigue, sleepiness, stress, and cognitive resilience formed the second-largest primary domain, with 31 studies (25.8% of the final corpus). The evidence profile differed markedly from the other domains. Fourteen studies (45.2%) used survey or cross-sectional designs, six used simulator or experimental designs, four were modelling or computational, four were operational or field studies, two were qualitative, and one evaluated an intervention. Stress or emotion was coded in 16 studies, resilience or mindfulness in 12, experience or expertise in 10, human error or unsafe behavior in five, and vigilance or attention-related constructs in several others. Sixteen studies used self-report measures, 15 used computational or model-based methods, three used sleep or actigraphy data, two used cardiovascular indicators, and a smaller number used eye-tracking, fNIRS, or other physiological measures. Representative evidence patterns are summarized in Table 5.

5. CONCLUSION

This systematic review synthesized 120 studies across decision-making, situation awareness, mental workload, and fatigue/resilience. The evidence shows that these domains are analytically distinguishable but empirically connected through recurring mechanisms such as attention and monitoring, experience, stress regulation, human-automation interaction, communication, risk appraisal, cognitive control, vigilance, and memory. The principal contribution is therefore an evidence-informed relational framework rather than a fixed causal chain. Cognitive readiness, task demand, information acquisition, situation representation, risk appraisal, and action selection interact dynamically and reciprocally within an operational and organizational context.

The review also shows why multimodal cognitive measurement must remain construct-aware. Behavioral, subjective, ocular, physiological, neurophysiological, and computational approaches provide complementary information, but no single measure directly represents workload, fatigue, stress, or situation awareness. Operational use requires validation, cross-context replication, calibration to individual and task variability, and demonstrated relevance to safety performance. From a management perspective, Safety-II and resilience help interpret adaptive cognitive performance, whereas SMS and FRMS provide organizational mechanisms for learning, hazard control, recovery support, training, and system improvement. Future progress will depend less on identifying isolated cognitive variables than on understanding when and under what operational conditions cognitive, technological, and organizational factors combine to support or constrain safe performance.

6. REFERENCES

Behrend, J., & Dehais, F. (2020). How role assignment impacts decision-making in high-risk environments: Evidence from eye-tracking in aviation. *Safety Science, 129*, 104738. https://doi.org/10.1016/j.ssci.2020.104738

Berthon, L., Bernard, F., Fleury, S., Paquin, R., & Richir, S. (2025). Multi-dimensional measurement of mental workload in industrial context: An experiment in the field of helicopter maintenance. *Applied Ergonomics, 124*, 104599. https://doi.org/10.1016/j.apergo.2025.104599

Bromfield, M. A., Milward, T., Everett, S. B., & Stedmon, A. (2023). Pilot performance and workload whilst using an angle of attack system. *Applied Ergonomics, 112*, 104101. https://doi.org/10.1016/j.apergo.2023.104101

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic review reporting guideline. *BMJ, 368*, l6890. https://doi.org/10.1136/bmj.l6890

Coombes, C., Whale, A., Hunter, R., & Christie, N. (2020). Sleepiness on the flight deck: Reported rates of occurrence and predicted fatigue risk exposure associated with UK airline pilot work schedules. *Safety Science, 129*, 104833. https://doi.org/10.1016/j.ssci.2020.104833

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37(1)*, 32-64. https://doi.org/10.1518/001872095779049543

Fabre, E. F., Matton, N., Beltran, F., Baragona, V., Cuny, C., Imbert, J.-P., Voivret, S., Van Der Henst, J.-B., & Causse, M. (2022). Hierarchy in the cockpit: How captains influence the decision-making of young and inexperienced first officers. *Safety Science, 145*, 105536. https://doi.org/10.1016/j.ssci.2021.105536

Hilditch, C. J., Gregory, K. B., Arsintescu, L., Bathurst, N. G., Nesthus, T. E., Baumgartner, H. M., Lamp, A. C. M., Barger, L. K., & Flynn-Evans, E. E. (2023). Perspectives on fatigue in short-haul flight operations from US pilots: A focus group study. *Transport Policy, 136*, 104-112. https://doi.org/10.1016/j.tranpol.2023.03.004

Islam, F., Sulung, S. D., Wibowo, U. L. N., & Nabila, Y. S. (2024). Pengaruh stres terhadap taruna penerbang sayap tetap dalam menghadapi fase pre-solo flight di API Banyuwangi. *SKYHAWK: Jurnal Aviasi Indonesia, 4(2)*, 366-385. https://doi.org/10.52074/skyhawk.v4i2.258

Lefrançois, O., Matton, N., & Causse, M. (2021). Improving airline pilots' visual scanning and manual flight performance through training on skilled eye gaze strategies. *Safety, 7(4)*, 70. https://doi.org/10.3390/safety7040070

Li, W.-C., Zhang, J., Braithwaite, G., & Kearney, P. (2022). The influence of augmented reality interaction design on pilot's perceived workload and situation awareness. *International Journal of Industrial Ergonomics, 90*, 103382. https://doi.org/10.1016/j.ergon.2022.103382

Li, W.-C., Zhang, J., Braithwaite, G., & Kearney, P. (2023). Quick coherence technique facilitating commercial pilots' psychophysiological resilience to the impact of COVID-19. *Ergonomics, 66(8)*, 1181-1195. https://doi.org/10.1080/00140139.2022.2139416

Liu, B., Xin, X., Ji, M., Yang, S., You, X., & Li, Y. (2023). Can family-work conflict influence safety behavior in airline pilots? The mediating role of fatigue and the moderating role of extraversion. *Safety Science, 163*, 106061. https://doi.org/10.1016/j.ssci.2023.106061

Lyu, M., Li, F., Qu, X., & Li, Q. (2024). Flashlight model: Integrating attention distribution and attention resources for pilots' visual behaviour analysis and performance prediction. *International Journal of Industrial Ergonomics, 103*, 103630. https://doi.org/10.1016/j.ergon.2024.103630

Malakis, S., & Kontogiannis, T. (2023). Team adaptation and safety in aviation. *Safety Science, 158*, 105985. https://doi.org/10.1016/j.ssci.2022.105985

Malakis, S., Kontogiannis, T., & Smoker, A. (2023). A pragmatic approach to the limitations of safety management systems in aviation. *Safety Science, 166*, 106215. https://doi.org/10.1016/j.ssci.2023.106215

Mattingsdal, H., Dahle, T. N., Abrahamsen, E. B., Søvik, S., & Ottestad, W. (2025). Workload in helicopter rescue operations - A comparison of two different rescue methods in a randomized cross-over design. *Safety Science, 181*, 106996. https://doi.org/10.1016/j.ssci.2025.106996

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Qin, H., Zhou, X., Ou, X., Liu, Y., & Xue, C. (2021). Detection of mental fatigue state using heart rate variability and eye metrics during simulated flight. *Human Factors and Ergonomics in Manufacturing & Service Industries, 31(6)*, 680-692. https://doi.org/10.1002/hfm.20927

Signal, T. L., van den Berg, M. J., Zaslona, J. L., Wu, L., Hughes, M., Johnston, B., Dyer, C., Drane, M., & Glover, M. (2023). Managing the challenge of fatigue for pilots operating ultra-long range flights. *Frontiers in Environmental Health, 1*, 1329203. https://doi.org/10.3389/fenvh.2023.1329203

Vallès-Català, T., Pedret, A., Ribes, D., Medina, D., & Traveria, M. (2021). Effects of stress on performance during highly demanding tasks in student pilots. *The International Journal of Aerospace Psychology, 31(1)*, 43-55. https://doi.org/10.1080/24721840.2020.1841564

Wang, Y., Zhang, C., Liu, C., Liu, K., Xu, F., Yuan, J., Jiang, C., Liu, C., & Cao, W. (2024). Analysis on pulse rate variability for pilot workload assessment based on wearable sensor. *Human Factors and Ergonomics in Manufacturing & Service Industries, 34(5)*, 492-504. https://doi.org/10.1002/hfm.21053

White, A., & O'Hare, D. (2022). In plane sight: Inattentional blindness affects visual detection of external targets in simulated flight. *Applied Ergonomics, 99*, 103578. https://doi.org/10.1016/j.apergo.2021.103578

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50(3)*, 449-455. https://doi.org/10.1518/001872008X288394

Xie, X., Yang, J., Li, Y., & Ji, M. (2024). Working memory capacity prevents pilots' loss of situation awareness in distraction scenarios: An eye-movement study. *International Journal of Industrial Ergonomics, 103*, 103667. https://doi.org/10.1016/j.ergon.2024.103667

Xie, X., Zhou, Y., Jiang, X., Xu, S., Zhang, J., Yang, C., & Ji, M. (2025). The impact of impulsivity on incident involvement in Chinese airline pilots: The serial mediating roles of risk perception and situation awareness. *Safety Science, 181*, 106833. https://doi.org/10.1016/j.ssci.2025.106833