Advances in Human Factors of Transportation, Vol. 224, 2026, 170–177
https://doi.org/10.54941/ahfe1007844

Human Factors Associated With Startle and Surprise Events in Aviation: A Large-Scale Analysis of NASA ASRS Reports

Preeti Yogendra Sharma (Alias Capt Priti Sidharth Singh) and Venkat Ram Reddy Ganuthula

School of Management and Entrepreneurship, Indian Institute of Technology Jodhpur, India

ABSTRACT

This study examined the prevalence and human factors correlates of startle and surprise events in commercial aviation using a large-scale analysis of NASA Aviation Safety Reporting System (ASRS) narratives. A keyword-matching algorithm was applied to 38,655 ASRS reports (2012–2022) to identify 2,642 reports (6.8%) containing startle/surprise-related language. Chi-square tests, odds ratio analyses, and logistic regression compared human factors, flight phase distributions, anomaly types, and outcomes between startle-flagged and non-startle reports. All major human factors—including Fatigue (OR = 1.86, $p < .001$), Physiological factors (OR = 2.11, $p < .001$), Workload (OR = 1.60, $p < .001$), and Confusion (OR = 1.38, $p < .001$)—were significantly over-represented in startle reports. Logistic regression confirmed Physiological factors ($\beta = 0.75$) and Fatigue ($\beta = 0.48$) as the strongest independent predictors. Loss of aircraft control was 2.4 times more prevalent in startle reports. These findings provide large-scale empirical evidence that fatigue, physiological vulnerability, and high workload significantly amplify the risk and severity of startle reactions in operational aviation, supporting the development of targeted Crew Resource Management interventions and evidence-based training protocols.

Keywords: Startle effect, Surprise, Human factors, Aviation safety, ASRS, Pilot performance, Fatigue, Crew resource management

INTRODUCTION

Commercial aviation maintains an exemplary safety record, yet incidents involving inappropriate crew responses to unexpected in-flight situations continue to challenge the industry (Reason, 2000). Among the most consequential human factors phenomena contributing to such incidents are startle and surprise effects—sudden, involuntary physiological and cognitive responses triggered by unanticipated stimuli that can temporarily impair a pilot's capacity for rational decision-making, situational awareness, and psychomotor performance (Martin et al., 2015; Rivera et al., 2014). These phenomena have been implicated in several high-profile accidents, including Air France Flight 447 (BEA, 2012) and West Air Sweden Flight 294 (SHK, 2016), prompting EASA to mandate training provisions addressing startle and surprise management (EASA, 2018).

Received March 20, 2026; Revised April 28, 2026; Accepted May 17, 2026; Available online July 20, 2026

The physiological startle response is a primitive, brainstem-mediated reflex characterized by involuntary muscle contraction, elevated heart rate, and a transient interruption of ongoing cognitive processing (Koch, 1999; Landman et al., 2017). Surprise, by contrast, involves a cognitive mismatch between an individual's mental model and the actual state of affairs, leading to confusion and delayed recognition of the developing situation (Endsley, 1995). While distinct phenomena, they frequently co-occur in operational settings and share the common outcome of degrading human performance at critical moments (Mohrmann and Field, 2017).

Despite growing recognition of startle and surprise in aviation safety, empirical research has been constrained by methodological limitations. Laboratory studies struggle to replicate the ecological validity of real-world operations (Landman et al., 2017). Simulator-based investigations face the inherent limitation that participants anticipate unexpected events (Casner et al., 2013). Survey-based studies, including the authors' prior work involving 44 experienced airline pilots (Sharma and Ganuthula, in preparation), provide valuable self-reported data but are limited by sample size and retrospective bias.

The NASA Aviation Safety Reporting System (ASRS) offers a unique resource for studying startle and surprise at scale. Established in 1976, the ASRS collects voluntary, confidential reports from aviation professionals describing safety-related events (NASA, 2024). Each report contains both structured, pre-coded fields and free-text narrative accounts, enabling a mixed-methods approach combining statistical power with rich contextual detail (Boland and van Rooij, 2017).

This study draws on three complementary theoretical perspectives: the startle reflex model (Koch, 1999), Endsley's (1995) three-level model of Situation Awareness, and the Human Factors Analysis and Classification System (HFACS; Wiegmann and Shappell, 2003). Additionally, the Yerkes-Dodson law (Yerkes and Dodson, 1908) predicts that factors elevating baseline arousal would narrow the margin before a startle stimulus pushes arousal beyond the optimal performance range, motivating this study's focus on fatigue and workload as potential startle amplifiers. The study addresses four research questions: (RQ1) Which human factors are most strongly associated with startle-related language in ASRS narratives? (RQ2) Do fatigue and workload amplify adverse outcomes? (RQ3) How does startle prevalence vary across flight phases and over time? (RQ4) What is the relationship between startle reports and event severity?

METHOD

Data Source and Startle Identification

Data were drawn from the NASA ASRS, accessed through a curated dataset on the Hugging Face platform (Hoole, 2022), containing 38,655 incident reports spanning January 2012 through March 2022 with 111 data fields per report. Startle/surprise reports were identified through systematic keyword-matching applied to concatenated narrative fields. The keyword list was developed through a multi-step process: core terms from the EASA research program (Mohrmann and Field, 2017)—specifically startle, startled, surprise, and surprised—formed the foundation. Additional terms reflecting physiological and cognitive startle manifestations were added, including: froze, frozen, disoriented, heart racing, adrenaline, shocked, panic, tunnel vision, overwhelmed, stunned, jolted, shaken, scared, and frightened (Koch, 1999; Landman et al., 2017). Wildcard pattern matching captured morphological variants. The complete keyword list comprised 21 regular expression patterns, identifying 2,642 reports (6.8%).

Variables and Analytical Approach

The binary startle flag served as the primary dependent variable. Independent variables included 10 ASRS-coded Person 1 human factors: Situational Awareness, Communication Breakdown, Confusion, Distraction, Workload, Time Pressure, Fatigue, Training/Qualification, Human-Machine Interface, and Physiological—Other. Contextual variables included flight phase, time of day, report year, and anomaly type. Event outcomes were coded with attention to loss of aircraft control, emergency landing, diversion, go-around, and near-midair collision.

Analysis proceeded in four stages: (1) descriptive statistics and cross-tabulations; (2) chi-square tests of independence with odds ratios and 95% Wald confidence intervals; (3) multivariate binary logistic regression with all 10 human factors as simultaneous predictors, plus a Fatigue × Workload interaction model; and (4) comparative analyses of event outcomes. All analyses used Python 3.11 with scipy.stats, scikit-learn, and pandas. Statistical significance was set at $\alpha = .05$ with Bonferroni correction for multiple comparisons.

Keyword Validation

A random sample of 150 startle-flagged reports was independently reviewed by two researchers. Inter-rater agreement was substantial ($\kappa = .78$). Among reviewed reports, 82.7% were true positives, 11.3% ambiguous, and 6.0% false positives. A sensitivity analysis using a restricted keyword list (8 most specific terms) identified 1,037 reports (2.7%) with virtually identical human factors associations (Physiological—Other OR = 2.34, Fatigue OR = 2.01), confirming robustness.

RESULTS

Human Factors Associations

Of 38,655 ASRS reports, 2,642 (6.8%) contained startle/surprise keywords. Chi-square tests revealed that 10 of 11 human factors examined were significantly over-represented in startle reports. Table 1 presents the complete results.

Table 1: Chi-square tests and odds ratios for human factors in startle vs. non-startle reports.

Human Factor	Startle %	Non-St. %	OR	95% CI	p
Physiological	4.9	2.4	2.11	[1.74, 2.56]	<.001
Fatigue	4.2	2.3	1.86	[1.52, 2.28]	<.001
Workload	20.0	13.5	1.60	[1.45, 1.76]	<.001
Confusion	27.7	21.7	1.38	[1.27, 1.51]	<.001
Time Pressure	13.4	10.3	1.34	[1.19, 1.50]	<.001
Comm. Breakdown	35.2	29.1	1.32	[1.22, 1.43]	<.001
Distraction	22.0	17.9	1.30	[1.19, 1.43]	<.001
Training/Qual.	12.5	10.2	1.26	[1.12, 1.41]	<.001
Sit. Awareness	49.8	44.9	1.21	[1.12, 1.31]	<.001
HMI	10.5	9.2	1.16	[1.02, 1.32]	.027

Logistic Regression

Table 2 presents the multivariate logistic regression results. The model achieved an AUC-ROC of .576, indicating modest but statistically meaningful discriminative ability.

Table 2: Logistic regression predicting startle/surprise report status from human factors.

Predictor	β	SE	Exp(β)	95% CI
Physiological	0.746	0.089	2.108	[1.77, 2.51]
Fatigue	0.478	0.100	1.613	[1.33, 1.96]
Workload	0.328	0.053	1.389	[1.25, 1.54]
Confusion	0.201	0.047	1.223	[1.12, 1.34]
Comm. Breakdown	0.196	0.044	1.217	[1.12, 1.33]
Training/Qual.	0.159	0.059	1.173	[1.05, 1.31]
Sit. Awareness	0.114	0.042	1.121	[1.03, 1.22]
HMI	0.100	0.060	1.105	[0.98, 1.24]
Time Pressure	0.067	0.057	1.070	[0.96, 1.20]
Distraction	0.029	0.050	1.030	[0.93, 1.14]

Physiological—Other emerged as the strongest independent predictor ($\beta = 0.75$, $\text{Exp}(\beta) = 2.11$), followed by Fatigue ($\beta = 0.48$, $\text{Exp}(\beta) = 1.61$) and Workload ($\beta = 0.33$, $\text{Exp}(\beta) = 1.39$). The Fatigue × Workload interaction model yielded a positive interaction coefficient ($\beta = 0.27$, $\text{Exp}(\beta) = 1.31$), suggesting the combined presence of fatigue and workload may amplify startle vulnerability beyond their independent effects.

Flight Phase Distribution and Temporal Trends

Startle reports were distributed across all flight phases. The highest absolute frequencies were observed during cruise (n = 422, 16.0%), initial approach (n = 283, 10.7%), and parked (n = 270, 10.2%). When normalized by total reports per phase, startle prevalence was relatively uniform (5.9–8.2%), contrasting with earlier survey findings in which pilots reported startle predominantly during landings, takeoffs, and turbulence encounters. Annual prevalence ranged from 5.8% to 8.4% with a slight upward trend in later years.

Event Outcomes and Severity

Table 3 presents outcome comparisons. Most critically, loss of aircraft control was reported in 7.1% of startle reports compared to 3.0% of non-startle reports—a 2.4-fold increase.

Table 3: Event outcome prevalence in startle vs. non-startle reports.

Outcome	Startle n (%)	Non-Startle n (%)	Ratio	p
Loss of Aircraft Control	187 (7.1%)	1,075 (3.0%)	2.37	<.001
Evasive Action Taken	482 (18.2%)	5,758 (16.0%)	1.14	.003
NMAC	193 (7.3%)	2,196 (6.1%)	1.20	.012
Go-Around / Missed Appr.	161 (6.1%)	1,577 (4.4%)	1.39	<.001
Diverted	159 (6.0%)	2,435 (6.8%)	0.89	.130
Emergency Landing	176 (6.7%)	3,282 (9.1%)	0.73	<.001

DISCUSSION

This study presents one of the largest empirical analyses of startle and surprise in commercial aviation, examining 2,642 startle-related reports from 38,655 ASRS incidents spanning a decade. The findings provide convergent evidence that specific human factors—particularly fatigue, physiological vulnerability, and workload—are significantly associated with startle events and with operationally significant outcomes, most notably loss of aircraft control.

Fatigue and Physiological Vulnerability

The finding that Fatigue (OR = 1.86) and Physiological—Other (OR = 2.11) were the strongest predictors provides large-scale empirical support for theoretical predictions. The Yerkes-Dodson framework predicts that fatigued operators have a reduced margin before unexpected stimuli push arousal beyond the optimal range (Stokes and Kite, 2017). This was confirmed in the authors' survey research, where pilots consistently reported heightened startle vulnerability during fatigue and circadian disruption (Sharma and Ganuthula, in preparation). The strong physiological association aligns with recognition that startle is fundamentally a brainstem-mediated reflex activating the autonomic nervous system (Koch, 1999), suggesting that interventions focused solely on cognitive strategies may be insufficient without complementary physiological regulation techniques (EASA, 2018).

Workload, Confusion, and the Cascade Model

The significant associations of Workload (OR = 1.60), Confusion (OR = 1.38), and Communication Breakdown (OR = 1.32) are consistent with a cascade model in which the initial startle leads to narrowed attention, which impairs situational awareness, produces confusion and communication failures, and increases workload (Mohrmann and Field, 2017). The positive Fatigue × Workload interaction ($\beta = 0.27$) suggests multiplicative vulnerability—an operationally important finding for fatigue risk management systems.

Flight Phase Distribution

The relatively uniform startle prevalence across flight phases (5.9–8.2%) challenges common assumptions. While pilots perceive startle as most problematic during high-workload phases (Sharma and Ganuthula, in preparation), the ASRS data suggest that the conditional probability of startle given an incident is relatively constant across the flight envelope. This has implications for training design, suggesting startle preparedness should be embedded across all flight phases rather than concentrated in approach-and-landing scenarios (EASA, 2018).

Loss of Aircraft Control

The 2.4-fold increase in loss of aircraft control is perhaps the most operationally significant finding. Loss of control in flight has been identified as the leading cause of fatal accidents in commercial aviation (Belcastro et al., 2017). This association reinforces the case for dedicated startle management training as a component of Upset Prevention and Recovery Training (EASA, 2018; ICAO, 2014).

Integration with Prior Survey Findings

This ASRS analysis was designed to complement the authors' prior survey of 44 experienced airline pilots and a follow-up survey of 26 pilots (Sharma and Ganuthula, in preparation). The survey found 70% prevalence of in-flight startle experiences, with recovery times ranging from seconds to several minutes. The current analysis provides triangulation against a much larger dataset, confirming: (a) the significant role of fatigue as a startle amplifier, (b) diversity of triggers and operational contexts, (c) association between startle and cognitive disruptions, and (d) potential for escalation into safety-critical situations. The survey finding that Communication Breakdown is over-represented in startle reports (OR = 1.32) provides indirect support for the importance of effective crew communication—and by extension, effective instructor communication in training—as a buffer against startle's cognitive effects.

Limitations

Several limitations warrant consideration. The keyword-matching approach may produce false positives and false negatives. The ASRS consists of voluntary, self-selected reports not representative of all aviation events. The cross-sectional design precludes causal inference. The curated dataset represents a subset of the full ASRS database. Finally, human factors codes are assigned by ASRS analysts rather than reporting pilots.

CONCLUSION

This study demonstrates that specific human factors—most notably fatigue, physiological vulnerability, and workload—are significantly and independently associated with startle and surprise events in commercial aviation. Analysis of 38,655 ASRS reports identified 2,642 (6.8%) containing startle-related language, disproportionately associated with loss of aircraft control (2.4-fold increase). Logistic regression confirmed Physiological factors ($\beta = 0.75$, OR = 2.11) and Fatigue ($\beta = 0.48$, OR = 1.61) as the strongest independent predictors.

These findings make three principal contributions: (1) the first large-scale epidemiological evidence linking specific human factors preconditions to startle events; (2) demonstration that startle vulnerability is broadly distributed across flight phases rather than concentrated in high-workload periods; and (3) integration with complementary survey research supporting comprehensive startle management programs addressing both physiological and cognitive dimensions. As the aviation sector evolves with increasing automation and growing traffic density, evidence-based startle and surprise management training represents a critical investment in aviation safety.

ACKNOWLEDGMENT

The authors acknowledge NASA for maintaining the ASRS database and the aviation professionals who voluntarily submit safety reports.

Author Contributions: Both authors, contributed equally at all stages of research leading to manuscript submission.

Funding Statement: No specific grants received.

Conflict of Interest: No conflicts declared.

Software and AI Usage Statement: The authors used Claude Sonnet 4.6 (www.claude.ai) to assist with structuring the manuscript and shaping its narrative flow. Grammarly (www.grammarly.com) was used for grammar and style improvements. The authors reviewed and edited all content and take full responsibility for the final article.

REFERENCES

BEA. (2012) Final report on the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France Flight AF 447. Bureau d'Enquêtes et d'Analyses.

Belcastro, C. M., Foster, J. V., Shah, G. H. et al. (2017) Aircraft loss of control problem analysis and research toward a holistic solution. Journal of Guidance, Control, and Dynamics, 40(4), 733–775.

Boland, E. J. and van Rooij, J. M. (2017) Review of existing training methods (NLR-CR-2016-621). Netherlands Aerospace Centre.

Casner, S. M., Geven, R. W., Recker, M. P. and Schooler, J. W. (2013) The retention of manual flying skills in the automated cockpit. Human Factors, 56(8), 1506–1516.

EASA. (2018) Startle effect management (Research Report EASA_REP_RESEA_2015_3). European Union Aviation Safety Agency.

Endsley, M. R. (1995) Toward a theory of situation awareness in dynamic systems. Human Factors, 37(1), 32–64.

Hoole, E. (2022) ASRS aviation incident reports dataset. Hugging Face. https://huggingface.co/datasets/elihoole/asrs-aviation-reports

ICAO. (2014) Manual on aeroplane upset prevention and recovery training (Doc 10011). International Civil Aviation Organization.

Koch, M. (1999) The neurobiology of startle. Progress in Neurobiology, 59(2), 107–128.

Landman, A., Groen, E. L., van Paassen, M. M., Bronkhorst, A. W. and Mulder, M. (2017) Dealing with unexpected events on the flight deck: A conceptual model of startle and surprise. Human Factors, 59(8), 1161–1172.

Martin, W. L., Murray, P. S., Bates, P. R. and Lee, P. S. Y. (2015) Fear-potentiated startle: A review from an aviation perspective. International Journal of Aviation Psychology, 25(2), 97–107.

Mohrmann, J. F. W. and Field, J. N. (2017) Review of startle and surprise role in incidents and accidents (NLR-CR-2016-620). Netherlands Aerospace Centre.

NASA. (2024) Aviation Safety Reporting System program overview. https://asrs.arc.nasa.gov/overview/summary.html

Reason, J. (2000) Human error: Models and management. BMJ, 320(7237), 768–770.

Rivera, J., Talone, A. B., Boesser, C. T., Jentsch, F. and Yeh, M. (2014) Startle and surprise on the flight deck. Proceedings of the Human Factors and Ergonomics Society, 58, 1047–1051.

Sharma, P. Y. and Ganuthula, V. R. R. (in preparation) Startle and surprise effects in commercial aviation: A multi-method investigation of pilot experiences and training perceptions.

SHK. (2016) Final report RL 2016:11e regarding the accident on 8 January 2016 with the aeroplane CL-600-2B19 at Örnsköldsvik Airport. Swedish Accident Investigation Authority.

Stokes, A. F. and Kite, K. (2017) Flight stress: Stress, fatigue and performance in aviation. Routledge.

Wiegmann, D. A. and Shappell, S. A. (2003) A human error approach to aviation accident analysis: The Human Factors Analysis and Classification System. Ashgate Publishing.

Yerkes, R. M. and Dodson, J. D. (1908) The relation of strength of stimulus to rapidity of habit-formation. Journal of Comparative Neurology and Psychology, 18(5), 459–482.