Exploring Human Performance Contributions to Safety in Commercial Aviation

Jon Holbrook, PhD
Crew Systems & Aviation Operations Branch
NASA Langley Research Center

March 12, 2019

Research collaborators

NAME	DOMAIN	ORGANIZATION
Dr. Jon Holbrook	Human Factors	NASA Langley Research Center
Dr. Lawrence Prinzel, III	Human Factors	NASA Langley Research Center
Mr. Brian Smith	Human Factors	NASA Ames Research Center
Ms. Colleen Cardoza	Flight Deck Operations	San Jose State University
Mr. Michael Stewart	Flight Deck Operations	San Jose State University
Mr. Gary Lohr	Air Traffic Operations	NASA Langley Research Center
Mr. Viraj Adduru	Data Science	Universities Space Research Assoc.
Mr. Oliver Ammann	Data Science	San Jose State University
Mr. Ilya Avrekh	Data Science	Stinger Ghaffarian Technologies
Mr. Bryan Matthews	Data Science	Stinger Ghaffarian Technologies

Supported by NASA Engineering and Safety Center; NASA ARMD's System-Wide Safety Project; NASA ARMD's Transformational Tools and Technologies, Autonomous Systems Sub-Project

Aviation is a data-driven industry

- We (rightly) want to make data-driven decisions about safety management and system design.
- The data that are available to us affect how we think about problems and solutions (and vice versa).
- In current-day civil aviation, we collect large volumes of data on the failures and errors that result in incidents and accidents, BUT…

Decision making is biased by the data we consider

- We rarely collect or analyze data on behaviors that result in routine successful outcomes.
- Safety management and system design decisions are based on a small sample of non-representative safety data.

Decision making is biased by the data we consider

- Human error has been implicated in 70% to 80% of accidents in civil and military aviation (Weigmann & Shappell, 2001).

Leads to…

- “To fast-forward to the safest possible operational state for vertical takeoff and landing vehicles, network operators will be interested in the path that realizes full autonomy as quickly as possible.” (Uber, 2016)

- This presupposes that human operators make operations less safe.

A thought experiment

- Human error has been implicated in 70% to 80% of accidents in civil and military aviation (Weigmann & Shappell, 2001).
- Pilots intervene to manage aircraft malfunctions on 20% of normal flights (PARC/CAST, 2013).
- World-wide jet data from 2007-2016 (Boeing, 2016)
  - 244 million departures
  - 388 accidents

A thought experiment

Outcome

Attributed to Human Intervention

Attributed to Human Intervention	Not Accident	Accident	
No	?	?	?
Yes	20%	80%	?
	?	388	244,000,000

- Human error implicated in 80% of accidents.
- Pilots manage malfunctions on 20% of normal flights.
- 388 accidents over 244M departures.

A thought experiment

Outcome

Attributed to Human Intervention

	Not Accident	Accident	
No	?	78	?
Yes	20%	310	?
	?	388	244,000,000

- Human error implicated in 80% of accidents.
- Pilots manage malfunctions on 20% of normal flights.
- 388 accidents over 244M departures.

A thought experiment

Outcome

Attributed to Human Intervention	Not Accident	Accident	
No	?	78	?
Yes	20%	310	?
	243,999,612	388	244,000,000

- Human error implicated in 80% of accidents.
- Pilots manage malfunctions on 20% of normal flights.
- 388 accidents over 244M departures.

A thought experiment

Outcome

Attributed to Human Intervention	Not Accident	Accident	
No	195,199,690	78	?
Yes	48,799,922	310	?
	243,999,612	388	244,000,000

- Human error implicated in 80% of accidents.
- Pilots manage malfunctions on 20% of normal flights.
- 388 accidents over 244M departures.

A thought experiment

Attributed to Human Intervention	Outcome: Not Accident	Outcome: Accident	
No	195,199,690	78	195,199,768
Yes	48,799,922	310	48,800,232
	243,999,612	388	244,000,000

When we characterize safety only in terms of errors and failures, we ignore the vast majority of human impacts on the system.

Protective and Productive Safety*

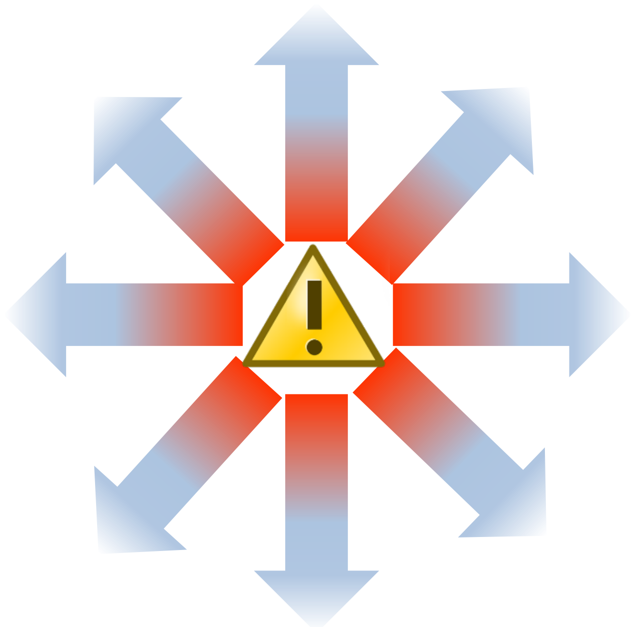
- **Protective Safety** – Prevent or eliminate what can go wrong by analyzing accidents and incidents (Safety-I).
- **Productive Safety** - Support or facilitate what goes well by studying everyday performance (Safety-II).

* Hollnagel, 2016

Why this distinction matters to safety

- Many paths will take you away from what you want to avoid.
- Not every path away from danger is a path toward safety.

Protective
Safety



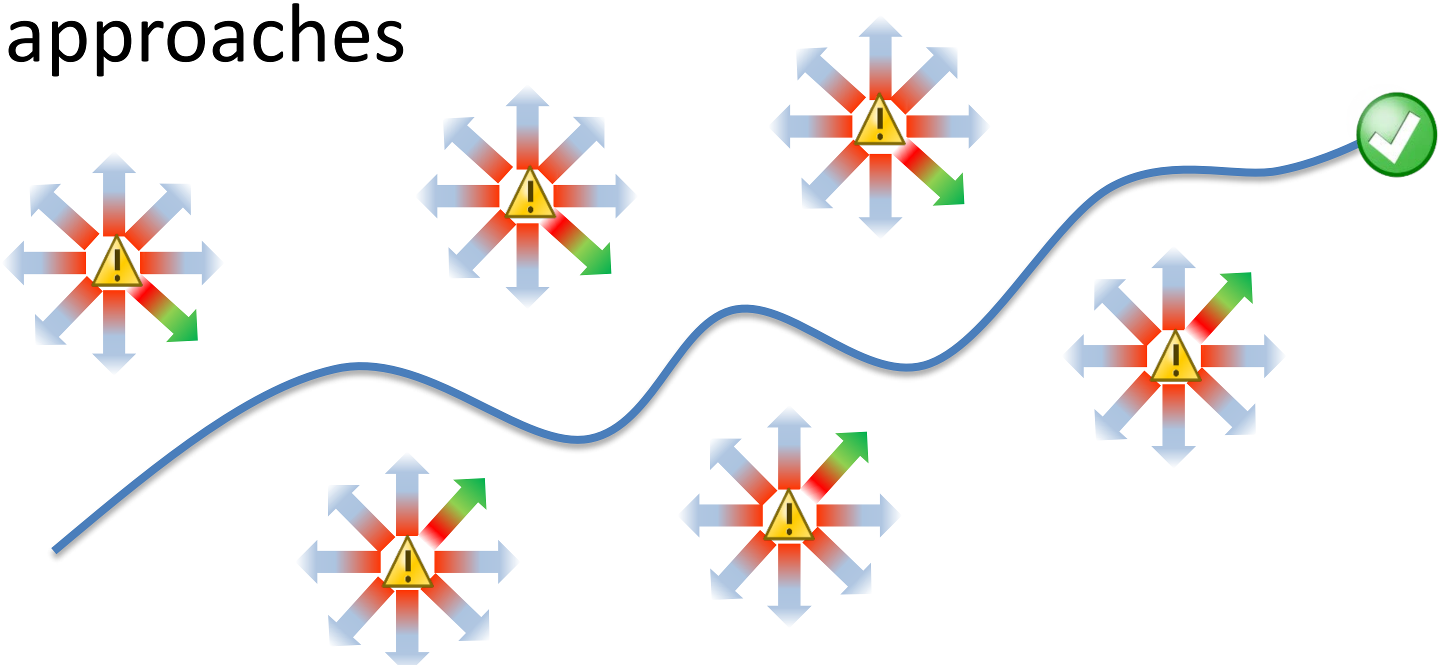
Why this distinction matters to safety

- Only one direction will bring you close to what you want to attain.

Productive
Safety

Why this distinction matters to safety

- Safely and successfully navigating a complex landscape requires both approaches



Why this distinction matters to NASA

- Planning and concepts for future operations in the national airspace system (NAS) include:
  - Decreasing human role in operational/safety decision making
  - Developing in-time safety monitoring, prediction, and mitigation technologies
  - Developing new approach to support verification and validation of new technologies and systems

Why this distinction matters to NASA

- Decreasing human role in operational/safety decision making

- Humans are the primary source of Productive Safety in today's NAS
- The processes by which human operators contribute to safety have been largely unstudied and poorly understood

Why this distinction matters to NASA

- Developing in-time safety monitoring, prediction, and mitigation technologies

- Solutions based on hazards and risks paint an incomplete picture of safety.
- Low frequency of undesired outcomes impact temporal sensitivity of safety assessments

Why this distinction matters to NASA

- Developing new approach to support verification and validation of new technologies and systems

- V&V metrics based on undesired outcomes can be impractical in ultra-safe systems
  - Time necessary to observe effect of a safety intervention in accident statistics is excessive (up to 6 years for a system with a fatal accident rate per operation of 10^{-7})
  - Attributing improvement to a specific intervention becomes intractable due to number of changes over time

Mechanisms of Productive Safety

- **Resilience**: the ability of a system to sustain required operations under both expected and unexpected conditions by adjusting its functioning prior to, during, or following changes, disturbances, and opportunities.*

- Capabilities of resilient systems:
  - Anticipate: “Knowing what to expect” in the future.
  - Monitor: “Knowing what to look for” in the near-term.
  - Respond: “Knowing what to do” in the face of an unexpected disturbance.
  - Learn: “Knowing what has already happened” and learning from that experience.

* Hollnagel, 2016

Work-as-Done vs. Work-as-Imagined

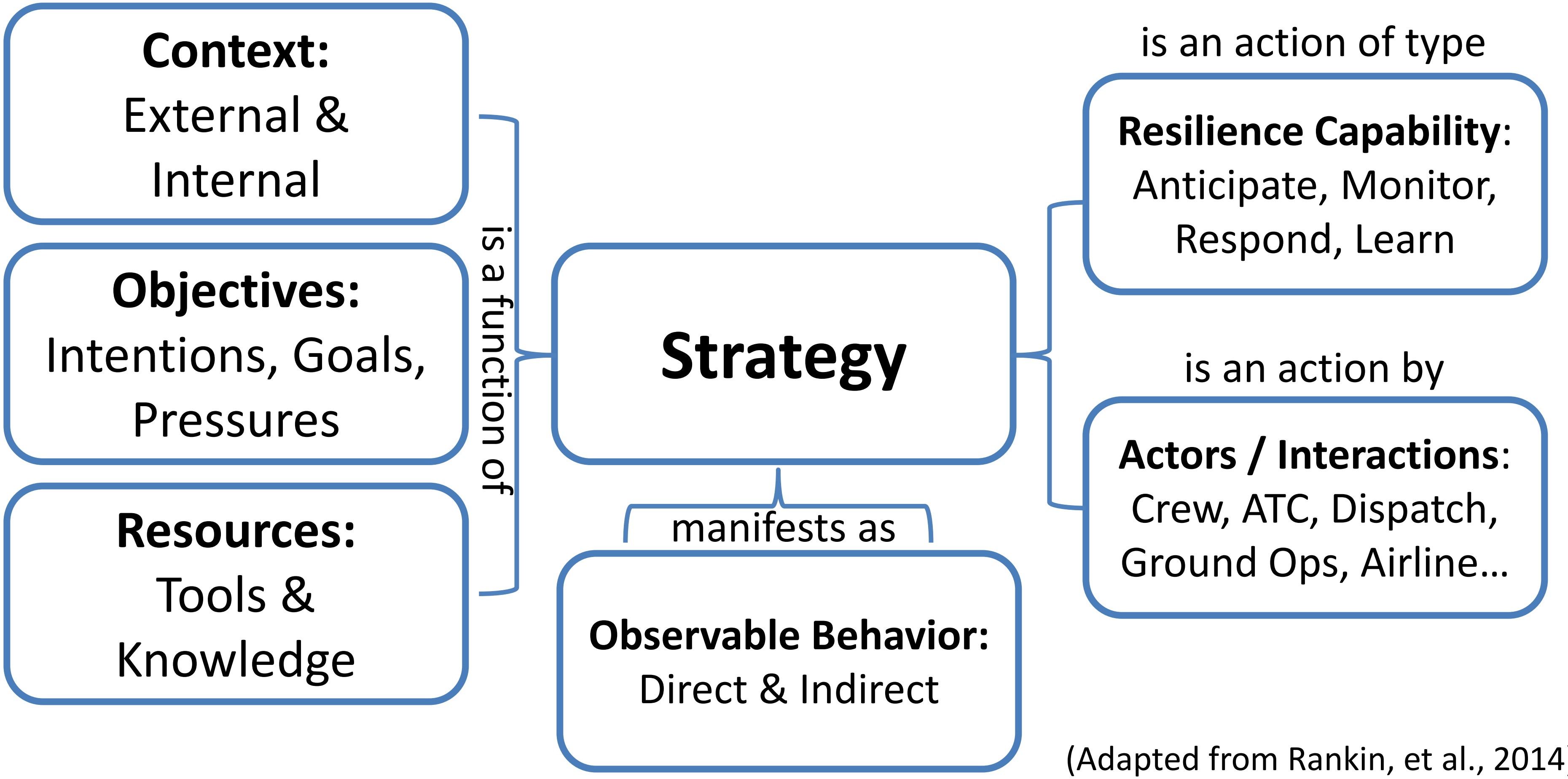
- Work-as-Imagined (Black line) – Procedures, policies, standards, checklists, plans, schedules, regulations
- Work-as-Done (Blue line) – how work actually gets done
  - Sometimes work goes as planned
  - Sometimes work goes better than planned
  - Sometimes work does not go as well as planned, but
  - MOST of the time, actual work is successful!

How can we characterize resilient performance?

- Lots of failure taxonomies, few success taxonomies
  - "Positive" taxonomies largely focused on positive outcomes (e.g., flight canceled/delayed, rejected takeoff, proper following of radio procedures)
- Can we use identify "universally desired" behaviors, regardless of subsequent outcomes?
- Can we identify "language" of resilience?
- Behaviors are complex, and occur within a rich context
  - How can we systematically capture "situated" performance without losing that richness?

Characterizing resilient performance

- No single data source can provide all of this information



(Adapted from Rankin, et al., 2014)

How can we study “work” in aviation?

- What data are currently available?
  - Operator-, observer-, and system-generated
  - Access challenge
  - Non-reporting challenge
- How and why are those data collected?
  - Sunk cost challenge
  - Happenstance reporting challenge
- How and why are those data analyzed?
  - Implications for post-hoc coding
  - Big-data challenge, and the need for tools to support analysis of narrative data
- There is no silver bullet
  - Fusing data into a coherent picture
  - De-identification challenge

Research questions

- How to Protective and Productive Safety thinking manifest in current aviation safety data collection and analysis practices?
- Can operators introspect about their own resilient performance?
- Can those introspections support analysis of system-generated data?

Method

- Reviewed state of practice in aviation safety data collection and analysis
- Conducted pilot and air traffic controller interviews to identify examples of resilient behaviors and strategies
- Used those behaviors and strategies to perform targeted analyses of airline FOQA data by asking "how might these strategies manifest in FOQA data?"

Results from Analysis of State of Practice

- Human Factors Analysis and Classification System (HFACS), Line Operational Safety Audits (LOSA), and Aviation Safety Reporting System (ASRS) have detailed coding structures for anomalies and errors, but limited coding for recovery/positive factors.
- Observer-based data collection approaches such as LOSA and Normal Operations Safety Survey (NOSS) code threats, errors, and key problem areas.
  - Focused on *respond* behaviors, but not systematically capturing *anticipate*, *monitor*, or *learn*

Results from Analysis of State of Practice

- Existing operator performance taxonomies (e.g., International Civil Aviation Organization, French Voluntary Reporting System, ASRS) describe specific operator behaviors
  - In the absence of associated situational factors, these behaviors may or may not represent safety-producing performance (e.g., flight canceled/delayed, rejected takeoff, proper following of radio procedures)
  - Current approach does not distinguish between behaviors that support resilient performance (i.e., universally desired behaviors) and behaviors that merely precede desired outcomes (i.e., behaviors which may or may not be desired)

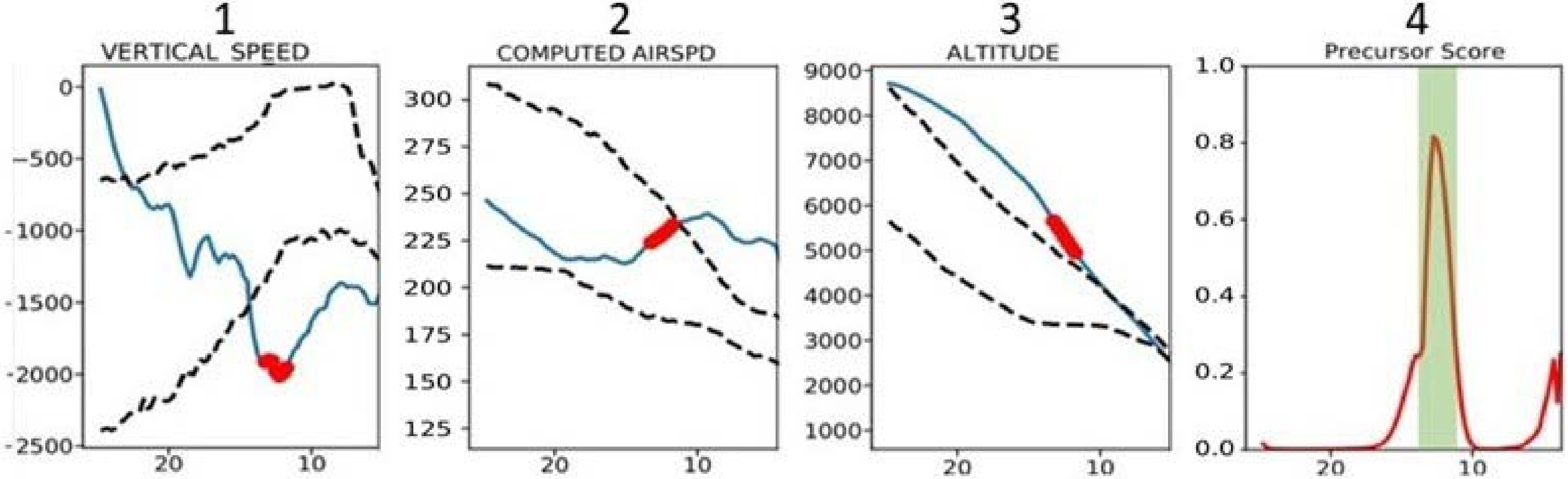
Results from Operator Interviews

Capability	Strategy
Anticipate	Anticipate procedure limits Anticipate knowledge gaps Anticipate resource gaps Prepare alternate plan and identify conditions for triggering
Monitor	Monitor environment for cues that signal a change from normal operations Monitor environment for cues that signal need to adjust/deviate from current plan Monitor own internal state
Respond	Adjust current plan to accommodate others Adjust or deviate from current plan based on risk assessment Negotiate adjustment or deviation from current plan Defer adjusting or deviating from plan to collect more information Manage available resources Recruit additional resources Manage priorities
Learn	Leverage experience and learning to modify or deviate from plan Understand formal expectations Facilitate others' learning

Results from FOQA Analysis: Anticipate Resource Gaps

Example: High-speed exceedance at 1000 ft

- Used sample of 1000 flights, half with adverse event and half without
- Deep Temporal Multiple Instance Learning (DT-MIL) algorithm
  - Detects states ahead of a pre-defined adverse event that have high probability of predicting that event
- Non-event flights examined for high precursor probabilities
- Pilot transferred aircraft energy from altitude to speed, preserving capability to reduce energy further by introducing drag
- More contextual information is needed to fully understand system variability



Findings

- NASA and industry planning and system design in aviation are based on principles and methods focused on predicting and preventing errors.
- Current safety reporting processes are designed to focus on and capture events that degrade safety, but not positive events that bolster safety.
- Defining safety in terms of “things that go right” enabled new methods for exploring existing data.
- Subjective and objective data sources contributed different information toward building an understanding of operators’ resilient performance.

Recommendations

1. Redefine safety in terms of the presence of desired behaviors and the absence of undesired behaviors.
2. Leverage existing data to identify strategies and behaviors that build resource margins and prevent them from degrading.
3. Develop tools to capture new data strategies and behaviors that support resilient performance.
   - From observer-based, operator-based, & system-based data
4. Develop a system-level framework for integrating across data types to facilitate understanding of resilient performance and work-as-done
5. Develop organization-level strategies that promote recognition and reporting of behaviors that support resilient performance.

Concluding thoughts

- Protective Safety thinking is pervasive in system design and safety management cultures in civil aviation
  - Limits the data we collect, the questions we ask, and therefore our understanding of work-as-done
  - Designing systems and making safety management decisions with an inadequate understanding of work-as-done can introduce unrecognized and unknown risks
- Productive Safety thinking represents a complimentary approach to Protective Safety thinking
  - Helps address system design and safety management barriers that arise due to Protective Safety thinking
  - Identifying, collecting, and interpreting data on operator resilient performance is critical for developing integrated, optimized human/technology or autonomous systems

Questions?

jon.holbrook@nasa.gov

757-864-9275

Additional detail is reported in:

NESC Technical Assessment Report, NESC RP-18-01304

Backup Slides

Results from Operator Interviews: "Anticipate" Strategies

Strategy	Behaviors
Anticipate procedure limits	Anticipate when formal procedure (e.g., STAR) won't work
Anticipate knowledge gaps	Anticipate others' intent
Anticipate resource gaps	Anticipate need to "buy time"
	Compare time needed and time available for action
Prepare alternate plan and identify conditions for triggering	Request land at alternate airport (e.g., due to weather) or runway
	Plan for go around (e.g., if preceding aircraft doesn't exit runway)

Results from Operator Interviews: "Monitor" Strategies

Strategy	Behaviors
Monitor environment for cues that signal change from normal ops	Monitor for "non-standard" signals/cues
	Monitor for deviations from normal pace of operations
	Monitor for deviations from normal control "feel" (e.g., weight on controls might indicate fuel imbalance)
Monitor environment for cues that signal need to adjust or deviate from current plan	Monitor party-line radio comms
	Monitor locations of aircraft in the area
	Monitor others' workload
	Monitor for cues (e.g., voice) of crew- or team-member's state (e.g., stress, uncertainty)
Monitor own internal state	Monitor own workload
	Monitor own limits and capabilities

Results from Operator Interviews: “Respond” Strategies

Strategy	Behaviors
Adjust current plan to accommodate others	Change speed to accommodate other aircraft
Adjust or deviate from current plan based on risk assessment	Deviate from procedure based on risk assessment
Negotiate adjustment or deviation from current plan	Negotiate route change
Defer adjusting or deviating from plan to collect more information	Defer action until more information available
Manage available resources	Divide/take/give tasks to balance workload
	Outsource tasks to automation (e.g., use autopilot to fly when handling other tasks)
Recruit additional resources	Ask others (e.g., ATC/dispatch) for assistance/resources
	Ask others (e.g., crewmember, ATC) for information/clarification
Manage priorities	Adjust timing or speed of tasks based on operation pace & workload
	Balance competing goals of formal expectations (e.g., follow procedures, maintain margins, smooth ride, reduce workload)
	Shed/abbreviate tasks to fit timeline/pace of operations

Results from Operator Interviews: "Learn" Strategies

Strategy	Behaviors
Leverage experience & learning to modify or deviate from plan	Predict likelihood of events based on past experience
	Consider historical occurrences with similar contexts
	Mentally simulate procedure
	Use heuristics/rules of thumb (e.g., fly upwind of a thunderstorm)
Understand formal expectations	Know and apply formal expectations (e.g., procedures, regulations, company policies, wx forecasting)
Facilitate others' learning	Teach other crew- or team-member
	Share actionable info with other aircraft/ATC

Results from Operator Surveys: Expanding List of "Positive" Behaviors

- Positive "Event Assessment" codes for ATC in NASA's ASRS database:
  1. Issued advisory/alert
  2. Issued new clearance
  3. Provided assistance
  4. Separated traffic

Resilient performance by operators is common and necessary:

- 92% of tower controllers indicated that they exhibited resilience on the job "at least once per day".
- 75% of tower controllers indicated making traffic management decisions NOT procedurally specified by JO 7110.65 or LOA "at least once per week".

- Potential additional codes identified through ATCT controller survey:
  5. Corrected read-back
  6. Provided weather information
  7. Intervened to prevent unsafe situation
  8. Anticipated potential problem
  9. Developed strategic plan to avoid a problem
  10. Adjusted traffic flow
  11. Cancelled clearance (T/O or Landing)
  12. Coordinated support
  13. Anticipated needs of pilot
  14. Anticipated flow issues
  15. Verified pilot intentions
  16. Repeated transmission for emphasis
  17. Communicated with professionalism/clarity
  18. Offered options/alternatives
  19. Monitored for changes
  20. Anticipated and adjusted for unexpected event

Results from FOQA Analyses: Managing Priorities

- Example: Timing of pre-takeoff control-surface check
  - Performing these checks prior to take-off is a procedural requirement, but the specific timing and spatial location is discretionary
  - Are there patterns to where and when pilots performed the check?
  - Mined from FOQA data for departures at Bercelona-El Prat airport by looking for consecutive full-range motion in rudder angle, aileron angle, and elevator angle during taxi-out.
  - Findings
    - No pilots performed checks before starting to taxi
    - Majority of checks initiated during 90-degree turn onto the taxiway parallel to the departure runway or during the 90-degree turn onto the runway itself
- Existence of discernible patterns suggests that performance variance occurs for strategic reasons, which can be explored in follow-up analyses.

Numbered regions indicate regions where control surface check were most commonly performed.