Methodology For Estimation of Benefits Of Human-factors Engineering in SESAR/NextGen Development

Lance Sherry (Ph.D.), Center for Air Transportation Systems Research (CATSR)
George Mason, University
Michael Feary (Ph.D.) NASA Ames Research Center
Jerome Lard (Ph.D.) Thales Research & Technology
Capt. Karl Fenell United Airlines

Eighth USA/Europe Air Traffic Management Research and Development Seminar (ATM2009)

CENTER FOR AIR

Acknowledgements

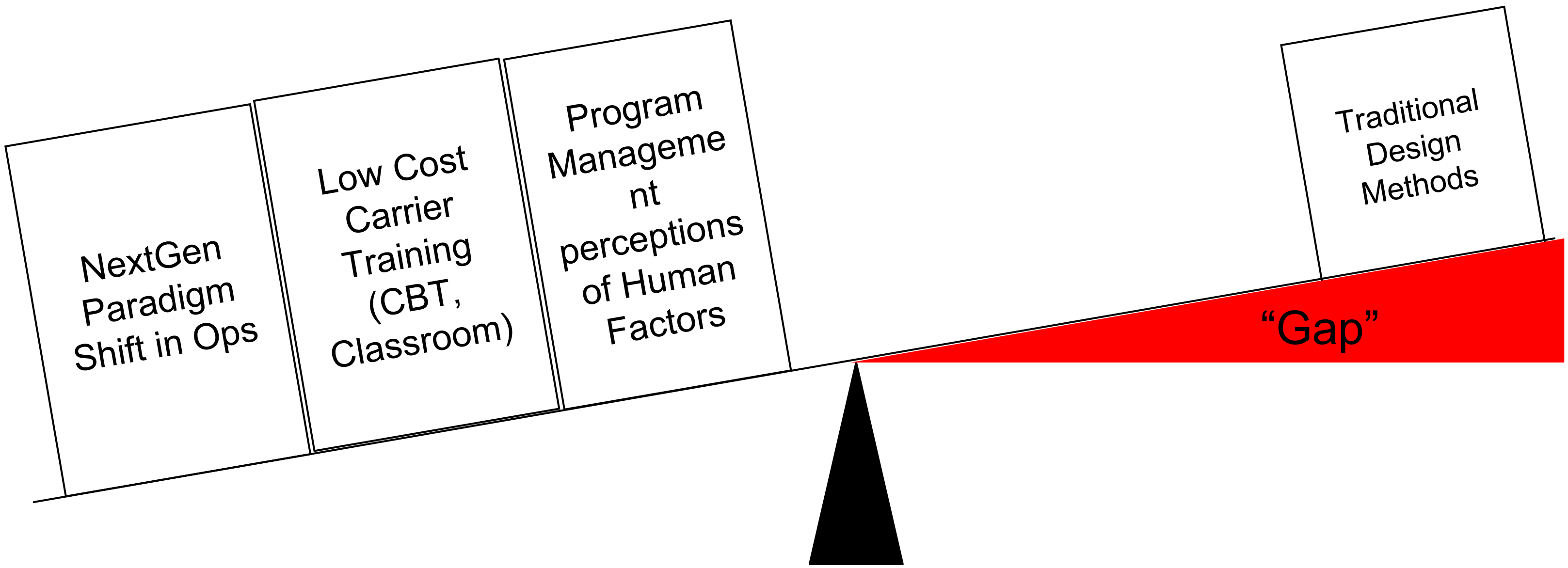
- Project funded by:
  - NASA – Aeronautics – Intelligent Integrated Flightdeck program
  - George Mason University Foundation Funds,
  - THALES IR&D funds
- Steve Young, Amy Pritchett (NASA).
- Dan Boorman, Randall Mumaw (Boeing),
- John Otiker (SWA)
- Peter Polson, Marilyn Blackmon (University of Colorado)
- Andrea Knight, Ricardo Prada (Google)
- Karlin Toner (FAA)
- Robert Mauro (University of Oregon)
- Immanuel Barshi (NASA-ARC),
- Mike Valencia, Midori Tanino (FAA),
- Mike Matessa (Alion Inc.)
- Michael Gordon (CCPace)
- Ken Goldwasser (SEER Tech)
- Tarik Ward (NASA-JSC)
- Terry Thompson (Metron Aviation)
- Debra Shreckenghost (TRACLabs)
- John Shortle, Rajesh Ganesan, John Ferguson,
- Guillermo Calderon-Mesa, Jianfeng Wang, Loni Nath (CATSR/GMU).

GEORGE MASON UNIVERSITY

Outline

1. Motivation #1: Paradigm Shift in HCI for NextGen Super-Density-Ops
2. Motivation #2: Lessons Learned from Introduction of FMS and Glass Cockpit
3. Problem Statement
4. Mission Scenarios, Tasks and Functions
5. Con-Ops Vulnerability and Design Robustness
6. Conclusions

Motivation



- Shifting paradigms have created "gap"
- Need for:
  - Standards, Methods, Tools

NextGen Con-Ops

Figure 2-6. Elements of a Four-Dimensional Trajectory

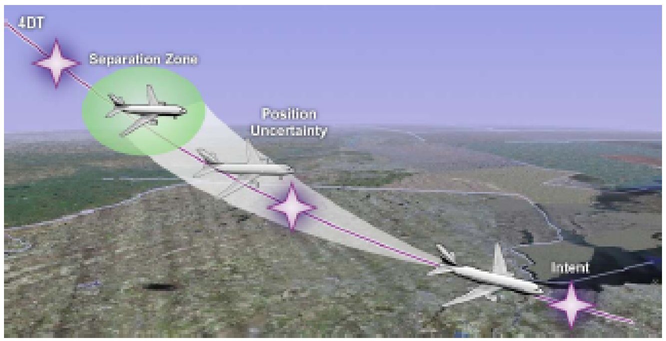


Figure 2-8. Super-Density Arrival/Departure Terminal Operations

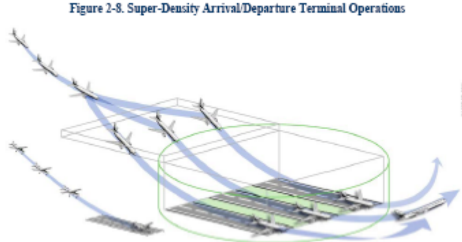
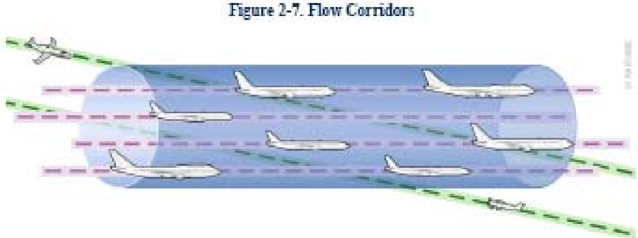


Figure 2-7. Flow Corridors



- **More efficient use of the airspace**
  - **Trajectory-based operations**
  - **Reduced separation in high-density airspace**
- **Performance-based Operations**
  - **increased navigation performance**
  - **Self-separation**
  - ***increased flight-crew proficiency performing required maneuvers***
    - **in a timely and robust manner**

Paradigm Shift in Ops

- Current Air Traffic Control paradigm
  - flights that do <u>not</u> meet required performance do <u>not</u> lose access to preferential airspace
- NextGen
  - flights that <u>cannot</u> respond to off-nominal situations in a timely manner will be shifted to downgraded airspace
    - experience flight delays
    - extended distance routing

Figure 2-7. Flow Corridors

GEORGE MASON UNIVERSITY

Changing Paradigms

- “Low Cost Carrier” Training
  - Traditional Airline Training
    - Student: Teacher = 2:1
    - Emphasis on tutoring
  - LCC Training
    - Computer-based Training
    - Classroom 23:1
- Program Management Perceptions of HF
  - Cold Water and Empty Guns (Rouse, 1987)

Steam-Gauge to Glass

1980's

- 1960 – 1980: Steam-Gauge Cockpit
- Instruments Federated
- Pilot integrates information to:
    - Tune radios, navigate, estimate fuel, compute speeds, select targets, …

- 1985 – Present: "Glass Cockpit"
- Integrated Information
- Automation for: Navigation (Navaid Selection), Flightplan (Nav Data-base, Fuel Calculations

UNIVERSITY

Standards, Design Methods Inadequate

- Transition to glass cockpit was documented by researchers:
  - Wiener, 1988, FAA, 1996, Feary et. al, 1998, BASI, 1999
- **"… like drinking through a firehose…."**
  - Training inadequate
  - Operations not transparent

Current Standards, Design Methods lead to High Probability Failure-to-Complete

1. Boeing 757-223, AAL, Cali, Accident Report - Ladkin (1996)
   1. "He could not be expected to recognize, because it **rarely occurs in regular operations**, …
   2. that the fix he was attempting to locate (Tulua VOR) was by this time behind him,……. and the **FMS generated display did not provide sufficient information** to the flightcrew that the fix was behind the airplane."
   3. "… because of the lack of time, the need to perform multiple tasks in the limited time, … his **attention thus narrowed** to further repeated unsuccessful attempts to locate ULQ [Tulua VOR] through the FMS."
2. B777 FMS Error Messages (65)
   - 57% messages
     - no info underlying causes of the message
     - no guidance to flightcrews on subsequent flightcrew actions to respond to message
   - 36% messages will:
     - occur very infrequently
     - exhibit high mission importance
     - not supported by salient visual cues on the interface

Problem Statement(s)

1. How to design and implement systems to be robust to failures and robust to operator errors?
   – NextGen Research and Policies Issues [JPDO, 2007] - R-46
2. How to measure the "vulnerability" of a Concept-of-Operations to HCI issues
3. How to measure the "robustness of a design" to overcome the "vulnerability" of the Con-ops
4. How to estimate the $ benefits of HCI Engineering in SESAR/NextGen

Tasks In NextGen/SESAR

- Flightcrew use *Automation Functions* to complete *Mission Tasks* in response to all plausible *Environmental Scenarios*
- Mission Tasks are explicit set of tasks required to conduct the mission in presence of all plausible scenarios:
  - ATC instructions
  - Checklist Items
  - Caution/Warning Messages

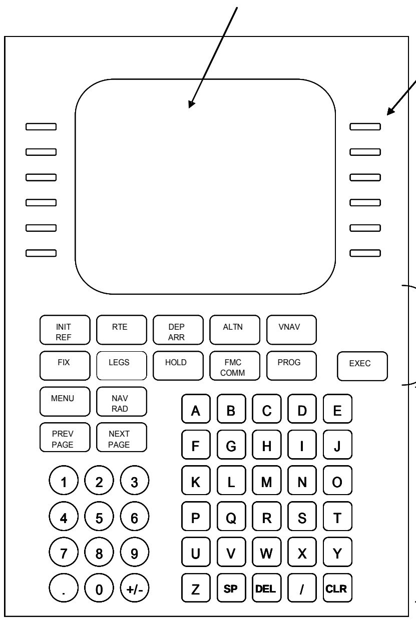
Example Mission Tasks

Mission Tasks triggered by ATC:	Mission Tasks triggered by FMS Error Messages:
• "Climb maintain FL 2-2-0" • "Proceed direct to waypoint XXX" • "Intercept course 0-1-4 degrees from present position" • "For weather, you are cleared to offset your current route 20 nautical miles right. Report when clear of the weather" • "Fly heading 180, join the Gordensville 060 degree radial, track the radial inbound to Gordensville" • "Hold west of Boiler on the 270^{O} radial. Right turns. 10 mile legs. Expect further clearance at 0830"	• Diagnose a mismatch in fuel sensors triggered by message FUEL DISAGREE – PROG 2/2 • Diagnose error in automated procedure tuning triggered by message NAV INVALID – TUNE XXXX – • Diagnose flightcrew data entry problems triggered by messages: INVALID ENTRY, NOT IN DATABASE, and INVALID DELETE • Re-enter data required before takeoff triggered by message: TAKEOFF SPEEDS DELETED –

Example Mission Task

"Hold west of Boiler on the 270 degree radial. Right turns. 10 mile legs. Expect further clearance at 0830"

MCDU page displays data and accepts *inserted* data



Line Select Keys *insert* data typed into Scratchpad into data field

Mode Keys for *accessing* pages

Apha/numeric keys for data *entry*

1) Id Task as Hold task
2) Decide to use the FMS "LNAV Hold: function
3) Press HOLD Function/Mode Key
4) Press LSK 6L, if a holding pattern is already in the route.
5) Line select waypoint identifier for "Boiler" to scratchpad.
6) Press LSK 6L.
7) Enter the quadrant and the radial into the scratchpad, W/270.
8) Press LSK 2L.
9) Enter the turn direction into the scratchpad, R.
10) Press LSK 3L.
11) Enter the leg distance into the scratchpad,10.
12) Press LSK 5L.
13) Enter expect further clearance time into the scratchpad, 0830.
14) Press LSK 3R.
15) Verify the resulting holding pattern on the ND
16) Press EXECUTE
17) Monitor trajectory for entry and following the Holding pattern

Properties of Mission Tasks

1. Mission Impact (or Loss Function)
    - Flight Critical – impacts safe operation
    - Procedure Critical – impacts ability to perform specific procedure
    - Not Flight/Procedure Critical
2. Probability of Failure to Complete
    - Likelihood Pilot/Automation collaborate to complete mission task
3. Frequency of Occurrence
    - Frequent (> once very 5 flights)
    - Infrequent (100 < X < 5 flights)
    - Very Infrequent (> 100 flights)

Con-Ops Scenario Vulnerability

- Con-Ops Scenario Risk =

  Frequency of Occurrence * Mission Impact

- Cumulative Con-ops Scenario Risk =

  $\sum_{\text{All Scenarios}}$ (Frequency of Occurrence * Mission Impact)

George Mason University

Con-Ops Scenario Vulnerability

Boeing 777 FMS Scratchpad Error Messages

	Frequency = FOC(i)			
Mission Impact (estimate additional flight time if fail complete) = ΔFT(i))	Very Infrequent (> 100 flights)	In-frequent (= 20 flights)	Occasional (=5 flights)	All the Time
Flight Critical (+30 mins delay)	**2**	1	-	-
Procedure Critical (10 mins delay)	24	9	5	1
Not Flight or Procedure Critical (< 1 mins delay)	9	1	13	-

Cumulative Con-Ops Scenario Risk = 7.7 (mins)
= $232 per flight

GEORGE MASON UNIVERSITY

Con-Ops Design Robustness Risk

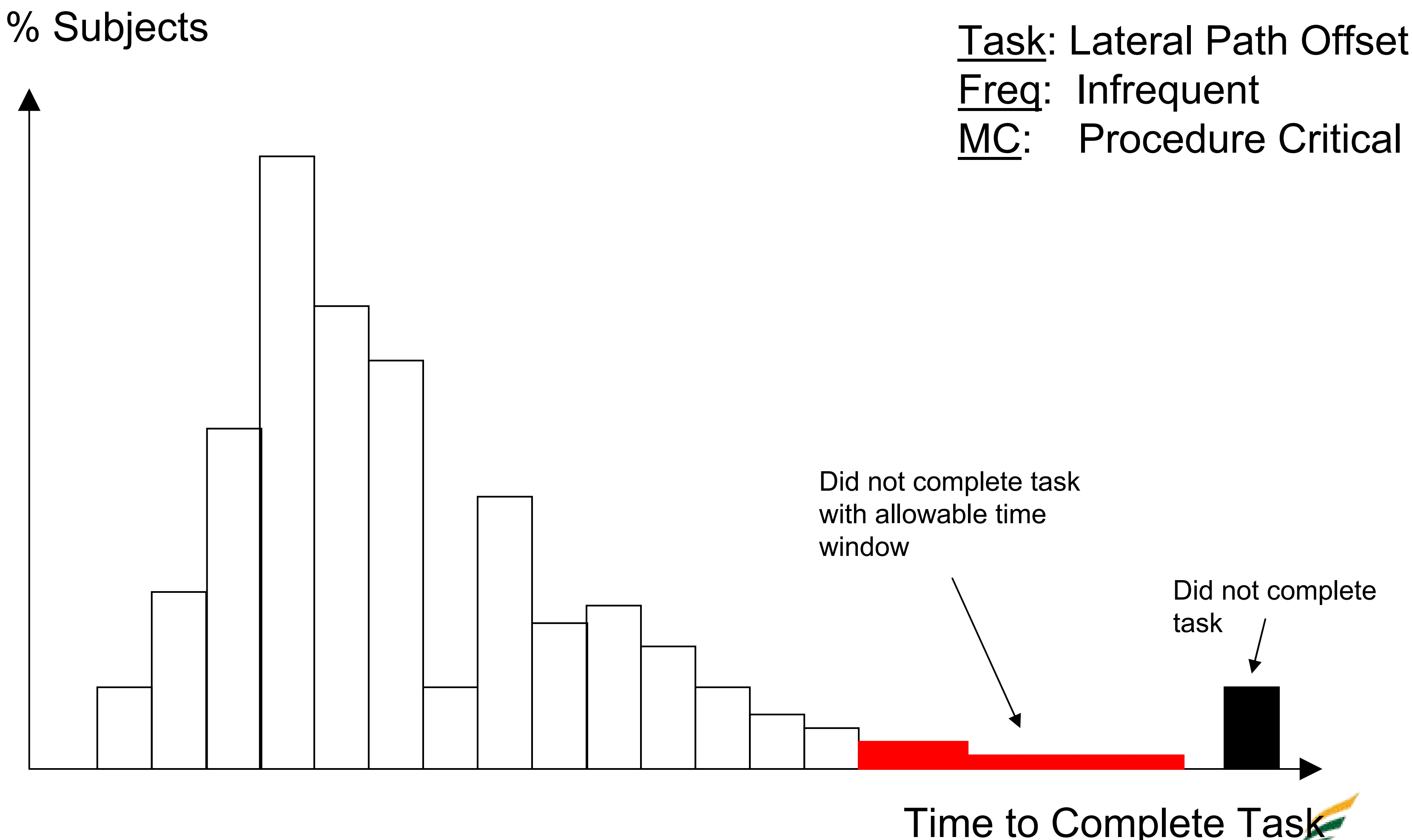
- Con-Ops Scenario Design Robustness Risk =

  Frequency of Occurrence * Mission Impact * Probability of Failure to Complete

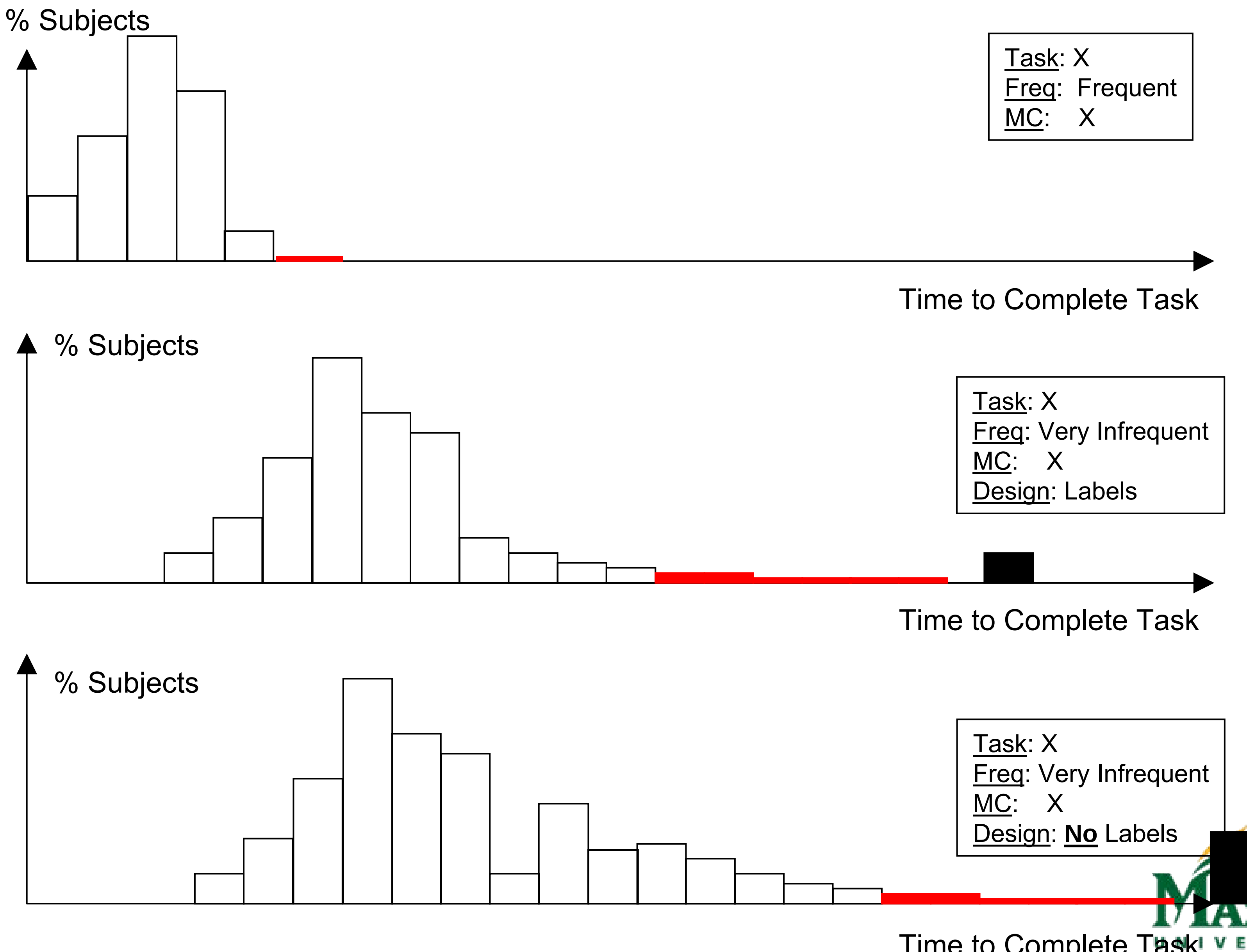
- Cumulative Con-ops Scenario Design Robustness Risk =

  $\Sigma_{\text{All Scenarios}}$ (Con-ops Scenario Design Risk)

Probability of Failure to Complete (PFtC)



PFtC, Frequency and Labels



Phenomenon: Persistent Interaction

- Boeing 757-223, AAL, Cali, Accident Report - Ladkin (1996)

1. "Confusion about the FMS presentation, as is true for use of any computer, is often resolved after ***persistent interaction*** with it.
2. … likely that the Captain believed that the confusion he was encountering was related to *his* use of the FMS…
3. …. and that continued interaction with it would ultimately clarify the situation."
4. "He could not be expected to recognize, because it **rarely occurs in regular operations**, …
5. that the fix he was attempting to locate (Tulua VOR) was by this time behind him,…
6. …. and the **FMS generated display did not provide sufficient information** to the flightcrew that the fix was behind the airplane."
7. "… because of the lack of time, the need to perform multiple tasks in the limited time, …
8. his **attention thus narrowed** to further repeated unsuccessful attempts to locate ULQ [Tulua VOR] through the FMS."

GEORGE MASON UNIVERSITY

Con-Ops Design Robustness

- B777 FMS Error Msgs
  - 57% messages
    - no info underlying causes of the message
    - no guidance to flightcrews on subsequent flightcrew actions to respond to message
  - 36% messages will:
    - occur very infrequently
    - exhibit high mission importance
    - not supported by salient visual cues on the interface

	Frequency = FOC(i)			
Mission Impact (estimate additional flight time if fail complete) = ΔFT(i))	Very Infrequent (> 100 flights)	In-frequent (= 20 flights)	Occasional (=5 flights)	All the Time
Flight Critical (+30 mins delay)	2	1	-	-
Procedure Critical (10 mins delay)	24	9	5	1
Not Flight or Procedure Critical (< 1 mins delay)	9	1	13	-

Con-Ops Scenario Design Robustness Risk =
Frequency of Occurrence * Mission Impact * Probability of Failure to Complete

Cumulative Con-Ops Design Robustness Risk
= 5.3 (minutes)
= $159 per flight

GEORGE MASON UNIVERSITY

Estimated Annual Costs to Airline Z

	Baseline (Traditional Design Methods)
Daily Total Flight Delays	583 hrs
Daily Total Excess Cost	$1.2M
Annual Total Excess Costs	$447M

Solution

1. More reliable human operators
   - Note: Operator performance for infrequent recall has not changed much since Ebbinghaus (1883)
2. Automation designed to support the Mission Task/Scenarios
   - Standards
     - Scenario Vulnerability/Design Robustness, Prob Failure-to-Complete
   - Methodology
     - Task Design
   - Tools (integrated into Design Process)
     - Task Design Document (Sherry, Feary, 2004)
     - HCI Automation (e.g. Cog Tool, John 2004, 2009; ADEPT, Feary 2007)

Proposed Design Standards

1. Probability-of-Failure to Complete < Δ
   – Δ adjusted based on frequency of occurrence
2. Trials-to-Mastery
3. Label-following to meet Semantic Salience standards (see Blackmon, Polson, Sherry, Feary)

Proposed Design Method

- Task Design Document (TDD)
  - Part of SDRL (i.e. required for regulatory Certification)
  - Contents
    - List of plausible scenarios
    - List of automation functions
    - Scenario-to-Function mapping (i.e. Tasks)
    - Probability of Failure-to-Complete for each task

Backup

GEORGE MASON UNIVERSITY

Additional Flight Time Penalty

- Additional Flight Time Incurred by Failure-to-Complete the Task
- Based on Mission Importance
- Flight Critical (+30 mins Δ flight time)
  - must be addressed immediately for continued safe flight
    - error message indicating a mismatch between fuel estimates and actual fuel levels
- Procedure Critical (10 mins Δ flight time)
  - safe flight
  - yield efficiency benefits
    - runway/approach selected in Flightplan is not compatible with the arrival in the flightplan).
- Neither Flight/Procedure Critical (< 1 min Δ flight time)
  - no safety or efficiency implications
  - performed at flightcrew discretion.

Probability of Failure-to-Complete

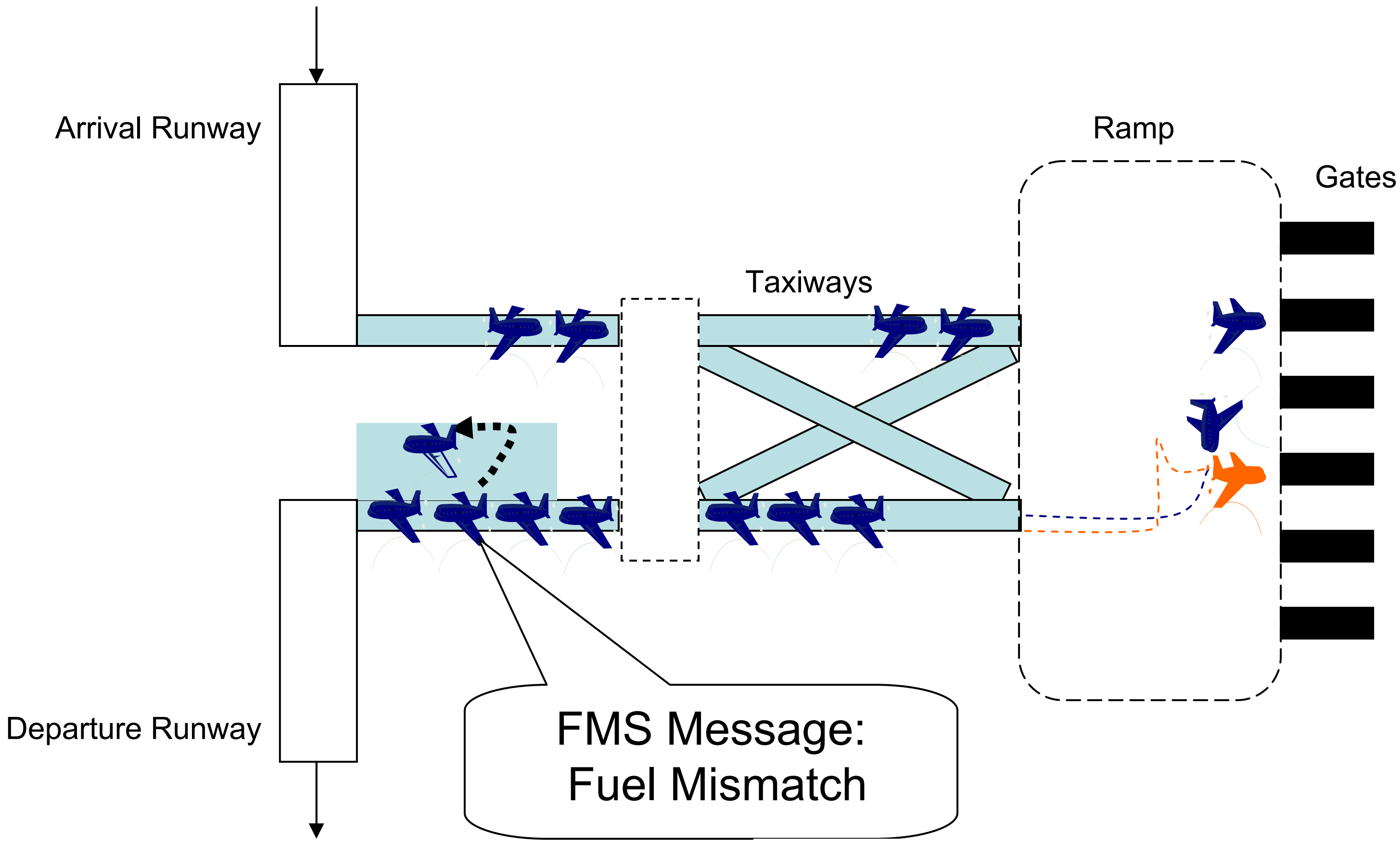
- (PFtC(i))
  - Likelihood of task not being completed (within time window)
  - PFtC < 0.3
    - Frequent Task
    - Infrequent Task WITH visual cues for next pilot action
  - PFtC > 0.3
    - Infrequent Task WITHOUT visual cues for next pilot action

Data

- Average Flights per Day – 3745
  - 535 aircraft
  - average 7 flights per day
- Average hours utilization per aircraft – 13 hours
- Average flight distance – 630 nautical miles
- Average Direct Operating Cost - $35 per minute
  - Biased by ratio of gate-taxi airborne operations (Bureau of Transportation Statistics)

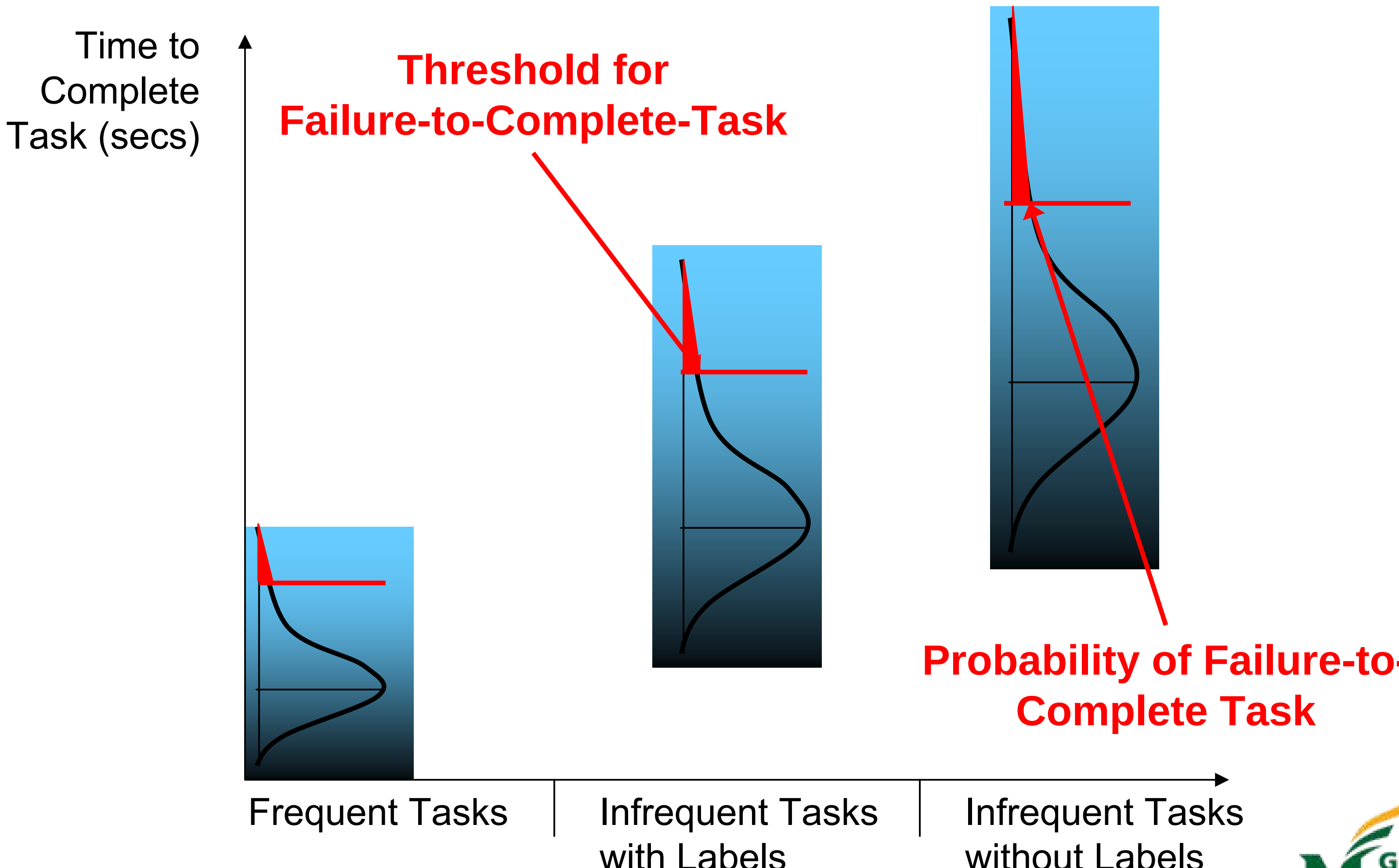
GEORGE MASON UNIVERSITY

Consequence of Paradigm Shift Lose Slot

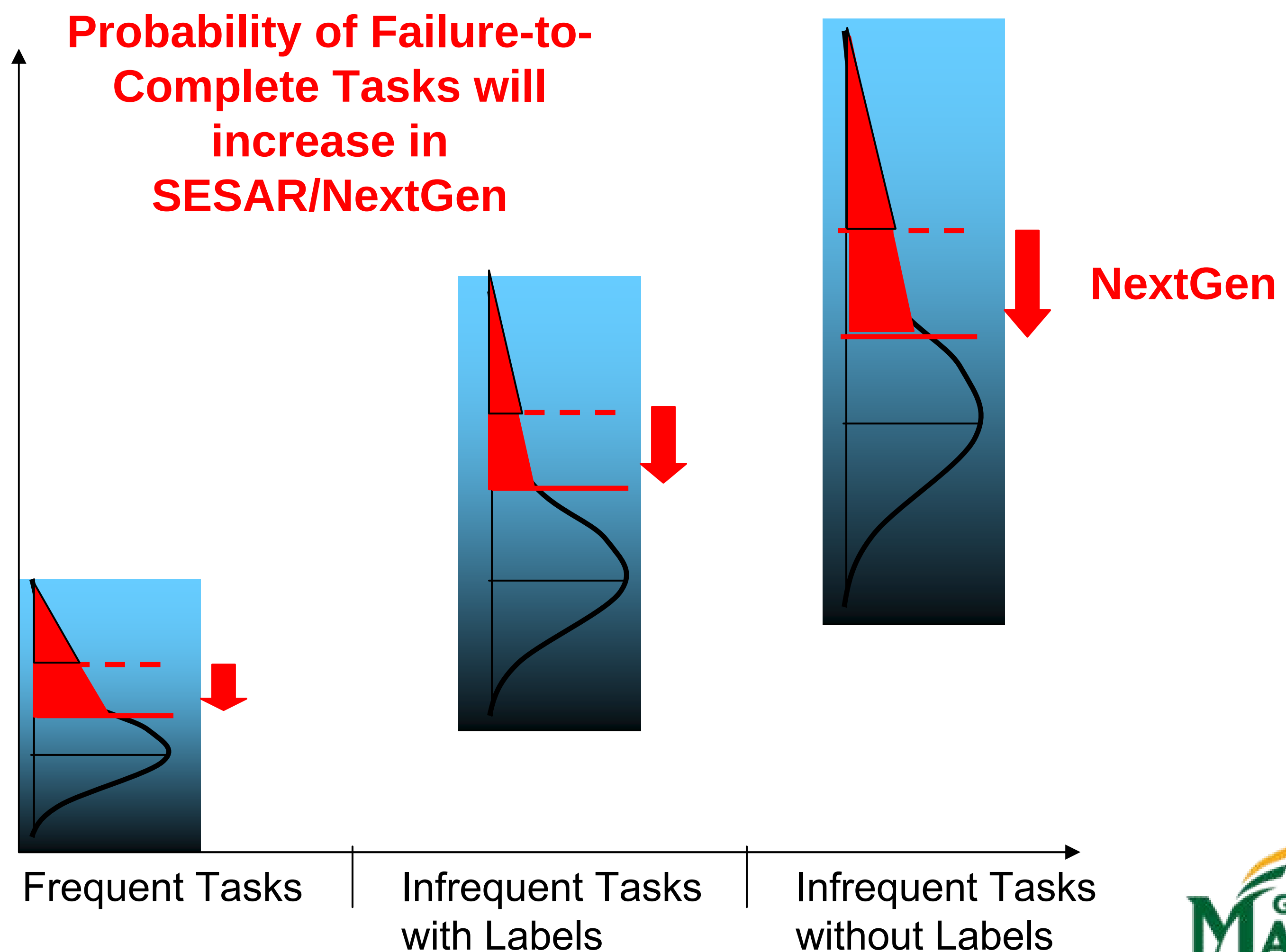


- Flight loses slot in departure queue
- All other flights “bump-up”
- Flight assigned slot when queue over

Result: Prolonged Interaction



Problem: NextGen Perf Reqs



Frequency of Occurrence

- Frequent (> once very 5 flights)
- Infrequent (100 < X < 5 flights)
- Very Infrequent (> 100 flights)

Motivation

- NextGen Research and Policies Issues [JPDO, 2007]:
- R-46:
  - At times of peak demand, major airports conduct Super-Density-Operations
  - Capacity-enhancing arrival and surface procedures are implemented to maximize runway throughput
- Issue:
  - How to design and implement systems to be robust to failures and robust to operator errors?

GEORGE MASON UNIVERSITY

Con-Ops Design Robustness

- Excess Costs per Flight =
  FOC(i) * PFtC(i) * ΔFT(i) * ADOC
  - Task i where i =1 .. n
  - FOC(i) = Frequency of Occurrence
  - PFtC(i) = Probability of Failure to Complete Task i in time ti
  - ΔFT(i) = Additional Flight Time Incurred by Failure to Complete the Task(i)
  - ADOC = Airline Direct Operating Costs

Certification, Qualification, and Proficiency

- Super Density Operations (SDO) dependent on Flight Technical Error (FTE)
  - Equipment *certified* to meet accuracy requirements
  - Pilot *qualified* on aircraft for operations
    - Critical elements of flight only (approach, landing, engine-out)
    - Emphasis is on flying qualities
  -

GEORGE MASON UNIVERSITY