Test and Evaluation of Visionic Systems: It's All in the Questions that You Ask

David C. Foyle, PhD [1]
Richard L. Newman, PhD [2]
Becky L. Hooey, MSc [3]

[1] NASA Ames Research Center, Moffett Field, CA
[2] Federal Aviation Administration, Renton, WA
[3] San Jose State University at NASA Ames Research Center

Human Factors research and technology

Outline

- Need for off-nominal testing
- E/SV issues - "insight approach"
- E/SV and HITS scenario recommendations
- Test & evaluation philosophies
- Formal off-nominal method
- Off-nominal event examples

Human Factors
research and technology

Off-Nominal Scenario Testing

CHALLENGE:
Ensure that laboratory studies "scale-up" to fielded system

FAA problem of technical transfer from lab to field

- ASDE-X system, but general (FAA White Paper Parasuraman, Hansman & Bussolari, 2002)
- Advocate early HF input into "very system requirements"

Software testing (esp. NASA space-related)

- Leveson (2001) advocated for off-nominal software testing (testing under unexpected conditions)
- What software should not do (negative requirements)
  - Avoided in software requirements, and
  - Forbidden by some industry standards (not verifiable - infinite testing)

Result:

- Nominal behavior is well-specified
- Off-nominal behavior is incompletely specified

Factor in aviation and space-mission accidents

Human Factors research and technology

E/SV Visionics Issues

Informally-derived Issues - Examples

"Informed, Insightful Researcher Analysis Approach"

Rotorcraft Civil Use of NVGs

Problem: Distance and altitude estimation (safe clearance, landing)
Reason: FOV, resolution, contrast
Assessment: Radio tower -- Height above; distance from

E/SV Usage

Problem: Altitude estimation (approach/landing)
Reason: FOV
Assessment: Objective SA probes, altitude callouts

T-NASA (Taxiway Navigation and Situation Awareness) System

Problem: Compellingness and crew coordination
Reason: Display formats, physical location/availability
Assessment: Induced Captain/First Officer route mismatch

E/SV Visionics Issues

Informally-derived Issues - Examples (cont).

HUD Landing/Approach

Problem: Compellingness and cognitive tunneling
Reason: Display formats, perceptual mechanisms (differential motion)
Assessment: SA probes (incursions), display format research

HUD Minimal Symbology Set

Problem: Recovery from unusual attitude (UA) with full HUD symbology set
Reason: Format configuration clutter
Assessment: Induce UA (via turbulence)

Formal Processes

Functional Hazard Assessment (for Certification)

Problem: Determine minor, major, hazardous, catastrophic hazards
Reason: E.g., Misinterpretation; where in flight envelope
Assessment: (US FAA AC-25.1309-1A) (e.g., absent vs. bad data)

E/SV Visionics Unresolved Issues

Simulator/Flight Test Participants

Experienced vs. Inexperienced (Test Pilots vs. Operational Pilots)

Pathway-in-the-sky displays

- Increased sensitivity as display experience increases

(Wilckens, 1973; Mulder & Mulder, 2004)

E/SV Flight Test Conditions

Problem: Low-visibility *emulation*

Reason: Simulators - are weather conditions realistic, validated?

- Flight test emulated weather - Simulate IMC by VMC with hood, then with step change to VMC at DH
- Flight test actual weather - In actual IMC with go-around; Safety? Scheduling?

Exception: Burgess' 1994 EVS tests

Scenario Recommendations Formal Method

Method: Characterize Study Problem on Four Dimensions (Newman, 2002)

E/SV Systems:

Operational Scenarios

Low-altitude phases of flight --

Terminal navigation, approach/landing, take-off/departure, etc.

Human Error Model

Detection and recognition of external objects/threats

Test Objectives

Target/hazard detection --

Runway incursions, uncharted towers, other objects

E/SV misalignment, Sensor boresight error

(McKay, Guirguis, Zhang & Newman, NATO RTO SCI/SET, 2002)

Test Criteria

Reaction time, hazard/non-hazard assessment accuracy

Scenario Recommendations Formal Method

Advanced Navigation Displays (i.e., Highway-in-the-sky):

Operational Scenarios

Low-altitude phases of flight --

- Terminal area, complex patterns, high-density traffic
- Approach/landing, take-off/departure

Human Error Model

Procedural issues

Situation awareness (detect unsafe situations - navigation blunders, loss of terrain separation)

Compellingness (cognitive capture/tunneling)

Test Objectives

Procedures (HITS reconfiguration due to engine failure/maneuvering)

Detection of off-nominal events (e.g., navigation blunders, diversions, stray aircraft)

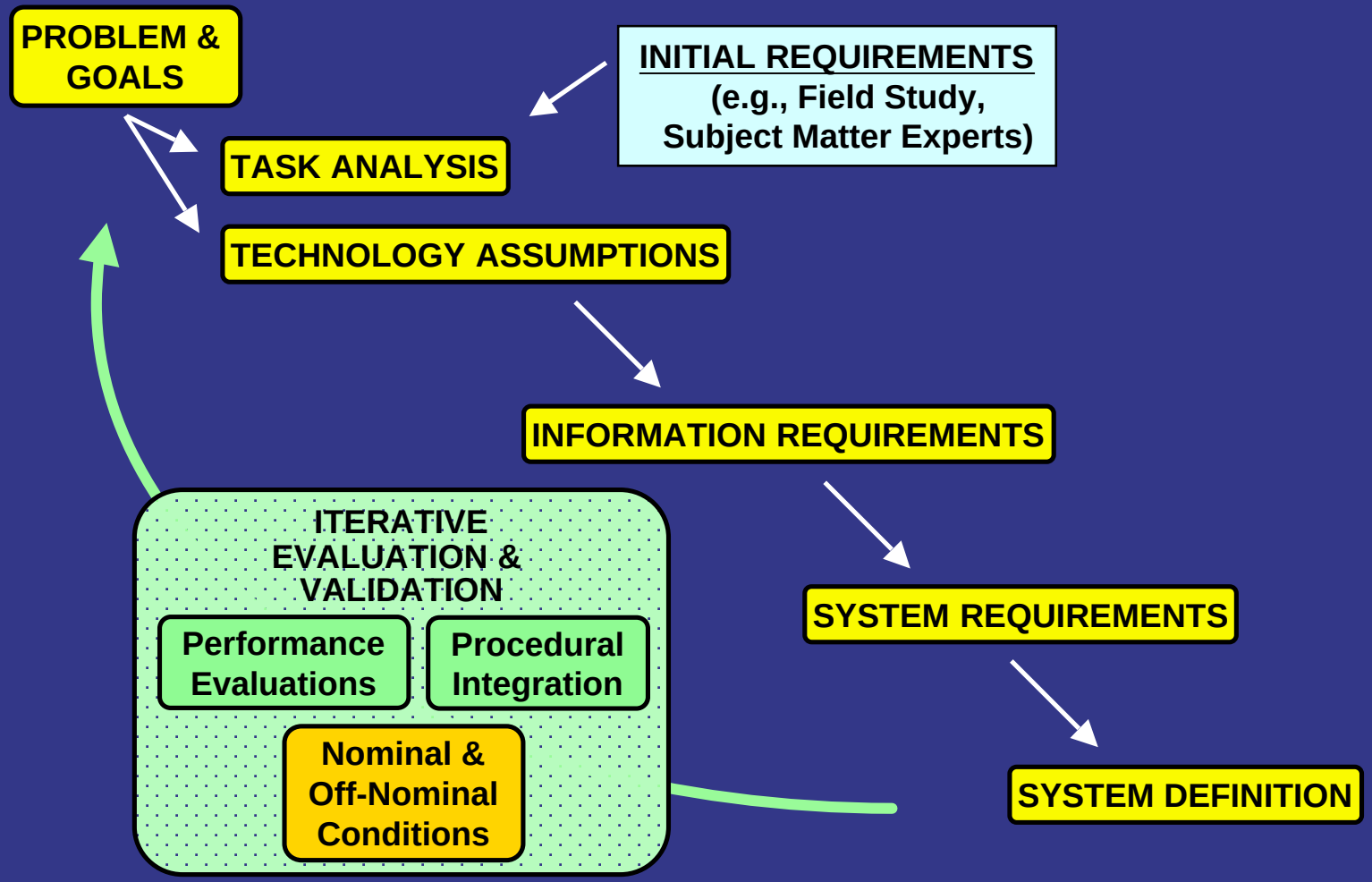
Test Criteria

Reaction time, flight technical errors (esp. turns), SA probes

Human Factors
research and technology

Human-Centered Design Method

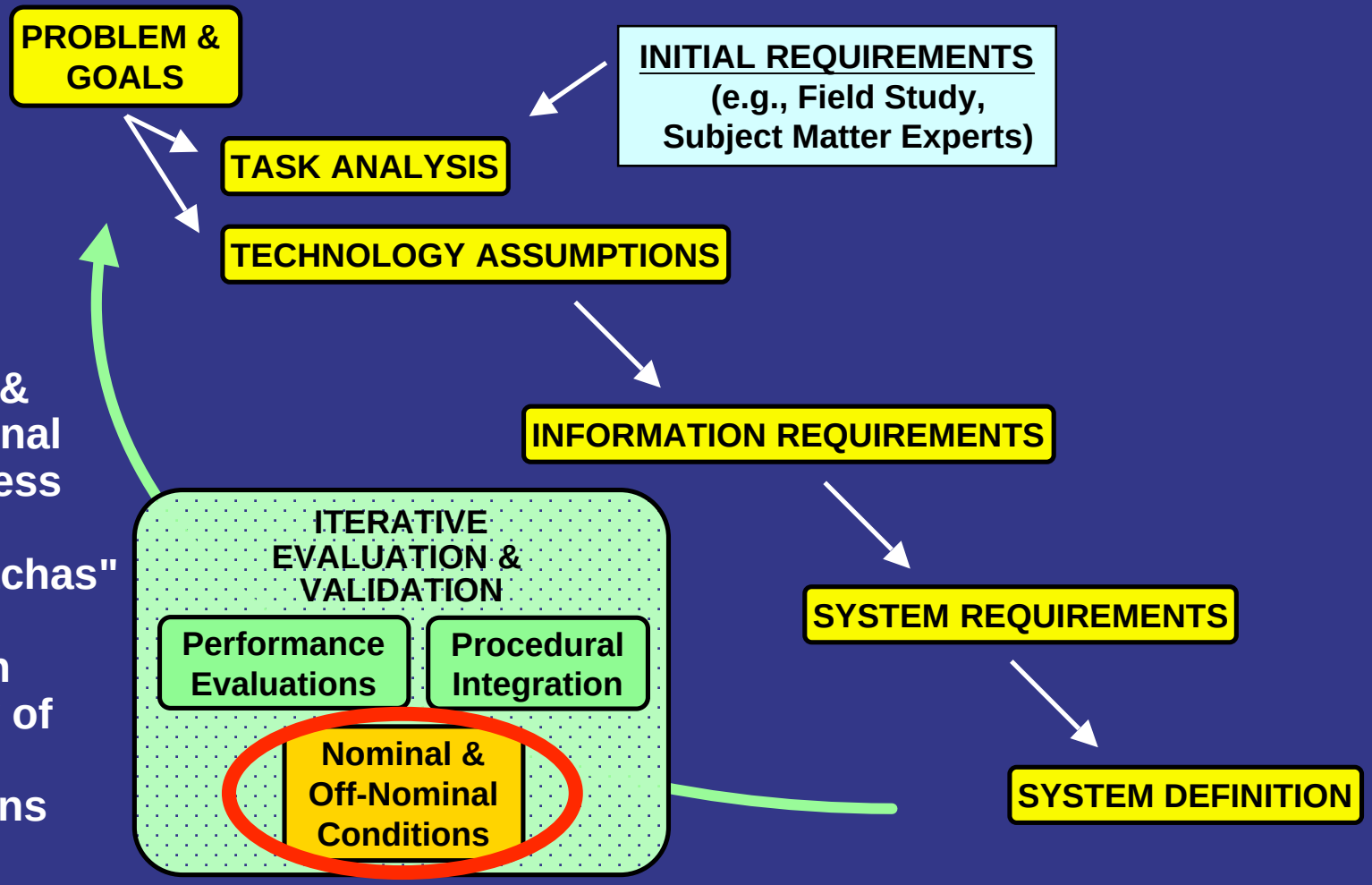
Foyle et al., (1996) - SAE Transactions: Journal of Aerospace
Hooey, Foyle, & Andre (2002) - NATO SET RTO MP-107



Human Factors
research and technology

Human-Centered Design Method

Foyle et al., (1996) - SAE Transactions: Journal of Aerospace
Hooey, Foyle, & Andre (2002) - NATO SET RTO MP-107



- System & operational robustness
- HF "Gotchas"
- Maintain integrity of nominal conditions

Human Factors research and technology

Off-nominal Method

Taxiway Navigation and Situation Awareness (T-NASA) System

Off-Nominal method:

- Medium-fidelity simulation (Foyle, Wilson, Hooey & Johnson, 2002)
- High-fidelity, full-mission simulation (Hooey, Foyle & Andre, 2000)
- Human sequential testing effects (e.g., memory, training, trust)
- Experimental design = "Art"
- Formal method needed

Off-Nominal Method:
Foyle & Hooey (2003)

Human Factors research and technology

Two Philosophies of Scenario Development

Problem observed: Improper balance between nominal and off-nominal scenarios in human-in-the-loop testing

Philosophy #1: Nominal condition emphasis:

- Off-nominal events are very disruptive
- Must protect *nominal condition data*
- Can only be tested on very last trial

Advanced avionics system -- prove it works well ("Engineering approach"):

- Goal of testing is to demonstrate benefits
- Off-nominal testing may contaminate nominal results

Human Factors
research and technology

Two Philosophies of Scenario Development

Problem observed: Improper balance between nominal and off-nominal scenarios in human-in-the-loop testing

Philosophy #1: Nominal condition emphasis:

- Off-nominal events are very disruptive
- Must protect *nominal condition data*
- Can only be tested on very last trial

Philosophy #2: Off-nominal events emphasis:

- *Off-nominal events* are the primary interest
- Tests should not waste time collecting nominal data

Advanced alerting system -- prove it alerts user

- Goal of testing is to verify user response to alerting system
- 90-100% of trials incorporate alert -- so as to not waste sim time

Human Factors
research and technology

Integrating the Two Philosophies

Advantages of integrating two philosophical approaches - testing both nominal and off-nominal events

Nominal Conditions	Both Nominal and Off-nominal Events	Off-nominal Events

Human Factors
research and technology

Integrating the Two Philosophies

Advantages of integrating two philosophical approaches - testing both nominal and off-nominal events

Nominal Conditions

- Normal usage assessment

- Typically encountered conditions - include wide range of routine scenarios
- Usage patterns, workload, efficiency
- Ensure robustness and system success

Both Nominal and Off-nominal Events

Off-nominal Events

Human Factors
research and technology

Integrating the Two Philosophies

Advantages of integrating two philosophical approaches - testing both nominal and off-nominal events

Nominal Conditions

- Normal usage assessment

Both Nominal and Off-nominal Events

Off-nominal Events

- Non-normal usage assessment

- Range from slightly "non-perfect" conditions to partial/full system failures
- Give insight into users' model of system and interactions (failures show user complacency or over-reliance)
- Issues addressed via system design changes, training, procedures

Human Factors
research and technology

Integrating the Two Philosophies

Advantages of integrating two philosophical approaches - testing both nominal and off-nominal events

Nominal Conditions

- Normal usage assessment

- User expectancy manipulation

- Manipulation of relative probabilities
- 80-90% nominal conditions - normal usage
- Caveat: Type and severity of off-nominal event affects probability for “normal usage”

Both Nominal and Off-nominal Events

- Comparative performance measurement

- Assess amount of disruption due to off-nominal event (e.g., turbulence)
- Provides quantitative assessment under worst-case fielded scenarios

Off-nominal Events

- Non-normal usage assessment

- Performance control

Nominal as control condition:

- User "on-task" in nominal
- Off-nominal data not because of "deviant" user

Human Factors research and technology

Development of Off-nominal Events

(Foyle et al, 2002); (Hooey et al, 2000); Foyle & Hooey (2003)

Determined human-system interactions

Identified psychological constructs

Created off-nominal events

Estimated disruptiveness of off-nominal events

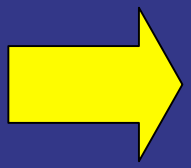
Incorporated off-nominal events into experimental design

Human Factors
research and technology

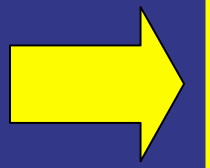
Development of Off-nominal Events

(Foyle et al, 2002); (Hooey et al, 2000); Foyle & Hooey (2003)

Determined human-system interactions



Identified psychological constructs

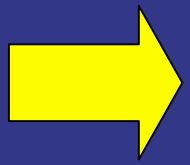


Created off-nominal events

- Unexpected environment or operational changes
- Interactions with other humans, equipment or technologies
- System failures (partial/total)

Focus groups to generate list and rate criticality

of off-nominal events



Incorporated off-nominal events into experimental design

Human Factors
research and technology

Development of Off-nominal Events

(Foyle et al, 2002); (Hooey et al, 2000); Foyle & Hooey (2003)

Determined human-system interactions

Identified psychological constructs

Created off-nominal events

- Analyzed list
- Identified common underlying psychological constructs
- E.g., user complacency, distance estimation

Estimated disruptiveness of off-nominal events

Incorporated off-nominal events into experimental design

Human Factors
research and technology

Development of Off-nominal Events

(Foyle et al, 2002); (Hooey et al, 2000); Foyle & Hooey (2003)

Determined human-system interactions

Identified psychological constructs

Created off-nominal events

- Specific off-nominal events created to assess constructs
- Appropriate dependent measures determined

Estimated disruptiveness of off-nominal events

Incorporated off-nominal events into experimental design

Human Factors
research and technology

Development of Off-nominal Events

(Foyle et al, 2002); (Hooey et al, 2000); Foyle & Hooey (2003)

Determined human-system interac

Identified psychological

Created off-nominal events

- Assign disruptiveness rating
- Disruptiveness:
  - Potential to alter user's system usage, visual scan, procedures
  - Impact on following test trials (i.e., negative system trust, crew interactions)

Estimated disruptiveness of off-nominal events

Incorporated off-nominal events into experimental design

Human Factors
research and technology

Development of Off-nominal Events

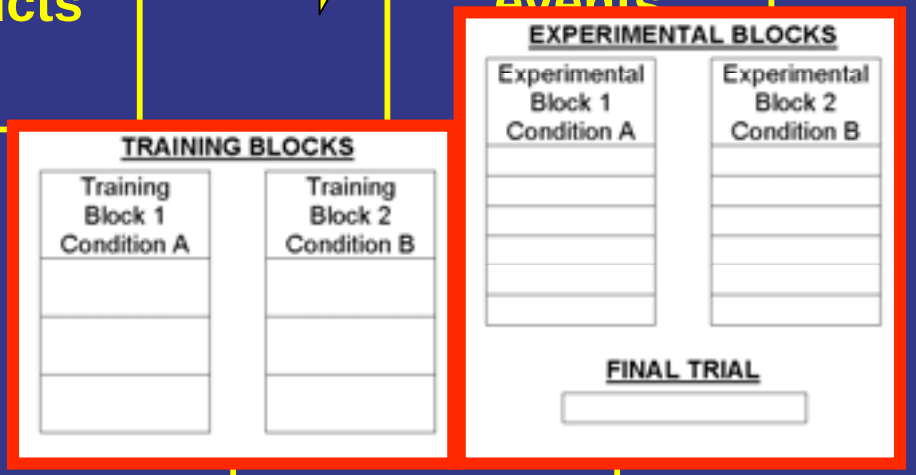
(Foyle et al, 2002); (Hooey et al, 2000); Foyle & Hooey (2003)

Ames Research Center

Determined human-system interactions

Identified psychological constructs

Created off-nominal events



Incorporated off-nominal events into experimental design

Estimated disruptiveness of off-nominal events

Human Factors research and technology

Incorporation into Experimental Design

TRIAL TYPE

TRAINING BLOCKS

	Training Block 1 Condition A	Training Block 2 Condition B
Mixed Nominal & Off-Nominal 1		
Mixed Nominal & Off-Nominal 2		
Mixed Nominal & Off-Nominal 3		

- General simulation familiarization
- Off-nominal stimulus/event familiarization
- Definition of off-nominal response requirement

EXPERIMENTAL BLOCKS

	Experimental Block 1 Condition A	Experimental Block 2 Condition B
RANDOM ORDER: Nominal 1		
Nominal 2		
Nominal 3		
Off-Nominal 1		
Off-Nominal 2		
Off-Nominal 3		

- Separate nominal and off-nominal trials allows for accurate estimation of nominal dependent measures
- Off-nominal trials may contain multiple off-nominal probes

- **Incorporate Low and Medium Disruptiveness off-nominal events**

FINAL TRIAL

Extreme Off-Nominal **"High Disruptiveness"**	

- Single trial at end of experiment for surprising or disruptive event

Incorporation into Experimental Design

TRIAL TYPE

TRAINING BLOCKS

	Training Block 1 Condition A	Training Block 2 Condition B
Mixed Nominal & Off-Nominal 1		
Mixed Nominal & Off-Nominal 2		
Mixed Nominal & Off-Nominal 3		

- General simulation familiarization
- Off-nominal stimulus/event familiarization
- Definition of off-nominal response requirement

EXPERIMENTAL BLOCKS

	Experimental Block 1 Condition A	Experimental Block 2 Condition B
RANDOM ORDER: Nominal 1		
Nominal 2		
Nominal 3		
Off-Nominal 1		
Off-Nominal 2		
Off-Nominal 3		

- Separate nominal and off-nominal trials allows for accurate estimation of nominal dependent measures
- Off-nominal trials may contain multiple off-nominal probes

- **Incorporate Low and Medium Disruptiveness off-nominal events**

FINAL TRIAL

Extreme Off-Nominal **"High Disruptiveness"**	

- Single trial at end of experiment for surprising or disruptive event

Incorporation into Experimental Design

TRIAL TYPE

TRAINING BLOCKS

TRIAL TYPE	Training Block 1 Condition A	Training Block 2 Condition B
Mixed Nominal & Off-Nominal 1		
Mixed Nominal & Off-Nominal 2		
Mixed Nominal & Off-Nominal 3		

- General simulation familiarization
- Off-nominal stimulus/event familiarization
- Definition of off-nominal response requirement

EXPERIMENTAL BLOCKS

RANDOM ORDER	Experimental Block 1 Condition A	Experimental Block 2 Condition B
Nominal 1		
Nominal 2		
Nominal 3		
Off-Nominal 1		
Off-Nominal 2		
Off-Nominal 3		

- Separate nominal and off-nominal trials allows for accurate estimation of nominal dependent measures
- Off-nominal trials may contain multiple off-nominal probes

- **Incorporate Low and Medium Disruptiveness off-nominal events**

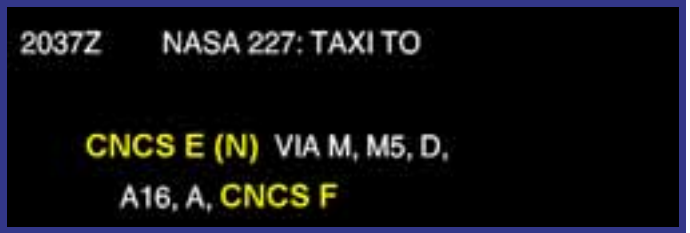
FINAL TRIAL

Extreme Off-Nominal **"High Disruptiveness"**	

- Single trial at end of experiment for surprising or disruptive event

Off-nominal Event Examples

Human-system interaction class: Interactions with other human agents in the system
Constructs: Complacency, levels of processing
Event: ATC issued erroneous taxi clearance
Disruptiveness: ***Low***



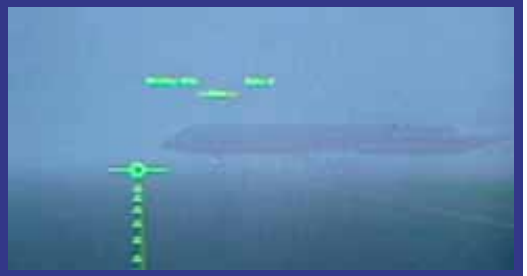
- Clearance always amended (whether or not noticed by pilot)
- Amended clearances typical in actual operation

Human-system interaction class: Failure of the system being tested
Constructs: Crew interaction and display cross-checking
Event: Partial failure of the system - Captain's HUD showed different route than First Officer's taxi map
Disruptiveness: ***High***

- Possible argument over correctness
- Could affect crew communication and teaming
- Could affect system trust; altering usage

Human Factors research and technology

Off-nominal Event Examples (cont.)

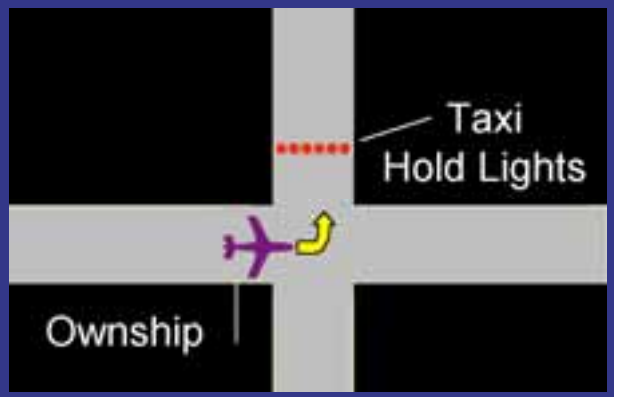


Human-system interaction class: Interactions with other equipment or technologies
Constructs: Complacency, trust, situation awareness
Event: Aircraft taxied in front of ownship - not on taxi traffic display requiring braking (Surveillance system limitation)
Disruptiveness: ***Moderate***

- Emergency braking and higher physiological arousal
- Cause attributed to normal surveillance system limit
- Not attributed to system under test; Trust unaffected

Human-system interaction class: Unexpected changes in the environment or operations
Constructs: Situation awareness, display capture
Event: Unexpected taxiway stoplights requiring quick reaction/near-emergency stop
Disruptiveness: ***Moderate***

- Possibly high physiological arousal
- But low consequence of miss (go unnoticed)



Human Factors
research and technology

Summary

Developed experimental method for off-nominal testing in human-in-the-loop evaluations

Off-nominal testing allows for:

- Understanding of the human-machine system under evaluation
- Uncover design issues that can be addressed
- Determination of training issues and procedures

The method involves:

- Developing issues to be tested
- Define off-nominal events addressing those issues
- Estimating disruptiveness of events
- Incorporate into experimental design
  - Low and moderately disruptive off-nominal events incorporated (Minimal disruption of nominal trial dependent measures)
  - Highly disruptive, "truly surprising" event - Single final trial

Ames Research Center

Human Factors research and technology

Summary (cont.)

Off-nominal testing

- Allows for more *robust* tests and evaluations
- May improve *technical transfer success rate* of systems and concepts from the laboratory to the field

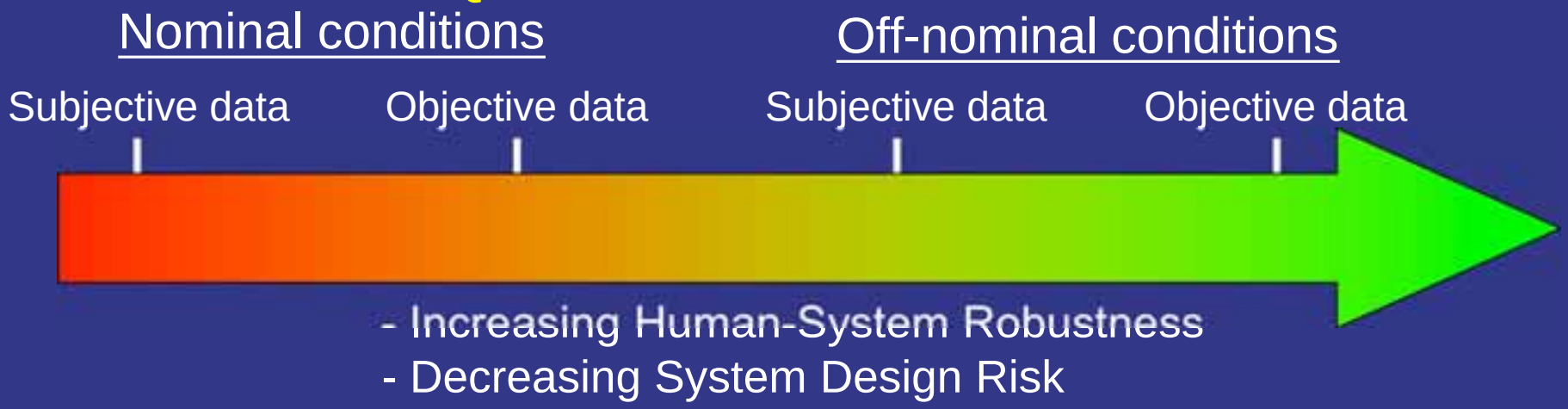
Human Factors
research and technology

Summary (cont.)

Off-nominal testing

- Allows for more *robust* tests and evaluations
- May improve *technical transfer success rate* of systems and concepts from the laboratory to the field

IT'S ALL IN THE QUESTIONS THAT YOU ASK



Human Factors research and technology

Test and Evaluation of Visionic Systems: It's All in the Questions that You Ask

David C. Foyle, PhD [1]
Richard L. Newman, PhD [2]
Becky L. Hooey, MSc [3]

[1] NASA Ames Research Center, Moffett Field, CA
[2] Federal Aviation Administration, Renton, WA
[3] San Jose State University at NASA Ames Research Center

Human Factors research and technology