AAM Document Number: AAM-NC-112-001

Working Paper Series
Publications of the National Campaign Subproject

Procedure Automation Rating Matrix

National Campaign Team

Abstract

The National Aeronautics and Space Administration (NASA) Advanced Air Mobility (AAM) National Campaign (NC) is researching the means by which future Urban Air Mobility (UAM) aircraft will operate safely in an integrated and scalable airspace architecture. Consistent with this objective, the NASA NC Airspace Procedures team designed a matrix to evaluate UAM instrument flight procedure design, flyability and interoperability of candidate departure, enroute, and approach architectures in live flight or simulation. The Procedure Automation Rating Matrix (PARM) is a multi-dimensional rating scale designed to provide direct feedback from test pilots and operators to airspace procedure designers developing airspace constructs for the integration and scalability of AAM operations in the National Airspace System (NAS). The PARM is assessed using a hierarchical decision tree that guides the operator through a ten-point alpha-numeric rating scale initiated either with or without the use of automation.

Revision History

Rev	Date	Sections Affected	Description of Change	Who
1	03 OCT 2022	All	Initial feedback and edits	David Zahn
1.1	16 MAY 2023	All	Final version	David Zahn & Sarah Eggum

Table of Contents

Table of Figures

1 Background

Aircraft handling qualities and task workload are currently assessed via the Cooper-Harper rating scale (CHS) or two of its variations: the Bedford Workload Scale (BWS) or the Modified Cooper Harper scale (MCH). The Cooper-Harper rating scale is the current standard for evaluating aircraft handling qualities. It utilizes a decision tree that assesses task completion adequacy, aircraft characteristics, and demands on the pilot to calculate and rate the handling qualities of an aircraft. [1] The Bedford Workload Scale is a modification of the Cooper-Harper rating scale. [2] BWS is a uni-dimensional scale that ranks whether (1) it was possible to complete the task, (2) if workload was tolerable for the task, and (3) if workload was satisfactory without workload reduction. The Modified Cooper Harper scale was developed for complex and automated systems to include task accomplishment, ability, errors, difficulty, performance and mental workload. [3] The MCH scale also focuses on assessing perception, cognition and communication. The PARM is not intended to replace or supplant these existing rating scales which have different purposes than that for which the PARM was designed. In contrast, the purpose of the PARM scale is to qualitatively assess the flyability and complexity of candidate instrument flight procedure constructs.

2 Purpose

The Procedure Automation Rating Matrix (PARM) was created to assess instrument flight procedures by trial and error with the help of NASA and Joby Aviation test pilots at the Joby Aviation high fidelity engineering simulator in Marina, California. The NC airspace procedure research and development required a tool that would provide direct feedback for the following:

- provision of preflight procedural information to the pilot or operator (e.g. airspace management construct)
- execution of flight information with any pilot/operator interface that includes cockpit manual control, remote operation, or automation with any combination thereof (e.g. new and novel approach plate design or multi-function display)
- adequacy of training required to achieve safe and scalable integration of any procedural operational concepts (e.g. commercial, instrument, or fundamentals of instruction (FOI) recommendations)

The NC Procedure Team developed a structure with four main categories, ten rankings and a structure that presents question “gateways” in the binary decision tree format familiar to test pilots from the BWS and MCH. The four main categories rank whether (1) it was possible to validate and accept the proposed procedure, (2) the procedural workload was tolerable, (3) there were no or acceptable depreciation levels of pilot situational awareness, and (4) indicate the projected level of training for median pilot proficiency. Additional criteria for the tool includes complexity and timing for the PARM evaluation.

3 Technical Requirements

The multi-dimensional Procedure Automation Rating Matrix (PARM) rating scale is utilized to quickly provide direct feedback from test pilots and operators to airspace procedure designers developing the airspace constructs for the integration and scalability of AAM operations in the

NAS. The PARM hierarchical decision tree is initiated either with or without the use of automation and guides the operator through the ten-point alpha-numeric rating scale for flight.The goal is to validate a dynamically generated procedure, accept the procedure and then execute the procedure while maintaining situational awareness throughout the duration of the flight. If all aspects of the flight procedure can be safely executed with no appreciable task saturation, the PARM score will elicit the test pilot's or operator's opinion on projected training requirements for the candidate procedure associated with the given research aircraft or aircraft configuration.

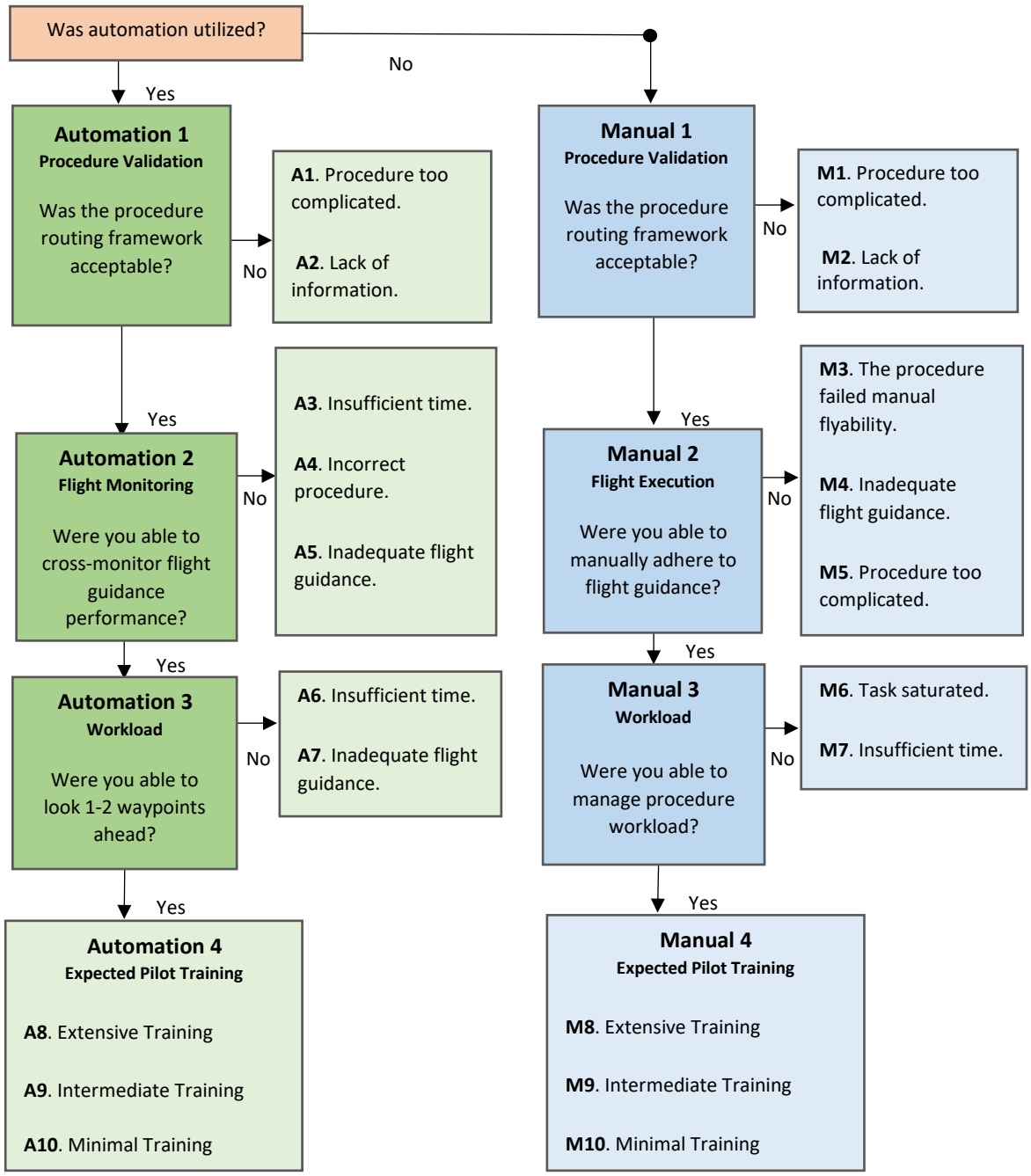

Figure 1: Procedure Automation Rating Matrix (PARM)

4 Evaluation

The pilot or operator rates the ability to fly the instrument procedures on a multi-dimensional alpha-numeric scale from 1 to 10 (best). The pilot/operator enters the rating matrix from the top left corner of the worksheet as opposed to the bottom corner for the BWS or MCH tools. The first top-left question determines which side of the worksheet the evaluation will proceed, either autonomation (green) or manual traditional control (blue). From there the test subject will proceed down the worksheet answering "yes/no" for each information gateway until deriving a rating corresponding to an answer or answers for the gateway. The numeric responses are grouped in four gateway categories:

Procedure Validation
Flight Monitoring
Workload
Pilot Training

The test subject may select one or more answers in each gateway category but then cannot proceed to another category. The rating is scored at the first (lowest) gateway where any type of degradation against procedure execution is noted.

4.1 Flight Procedures with Automation

The PARM is evaluated through the following structure for flight procedures with automation:

Initial Question

Was automation utilized?

For the purpose of this evaluation, automation is defined as the ability of an aircraft to perform control and navigation functions without input by the pilot or remote operator. Automation includes auto-pilot, pilot-assist features, or complete aircraft automation.

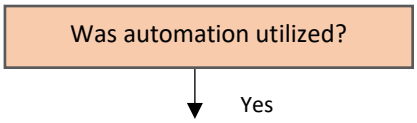

Figure 2: PARM Initial Question

Gateway #1: Procedure Validation

Was the procedure routing framework acceptable?

The second question is designed to evaluate the operator's ability to review and confirm the validity of the proposed procedure. Given dynamic airspace routing, the intent of this question is to assess the ability of the test subject to ingest and understand the proposed operation (expected to be packaged in a different way from traditional or legacy operations).

This question includes the departure, enroute, final approach and missed approach components of the proposed flight.

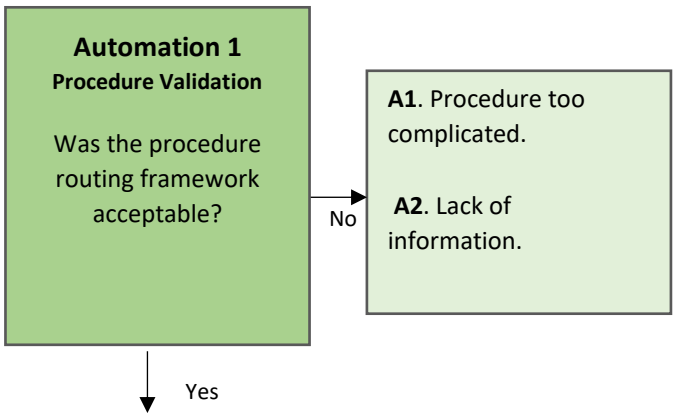

Figure 3: Flight Procedures with Automation Gateway #1: Procedure Validation

A1) Procedure too complicated: Test subject was unable to validate the procedure due to an inability to comprehend or understand the given instructions. This could be from too many routing options (contingencies) closely spaced or overlaid together. Additionally, there could be compound movements in altitude, airspeed, and azimuth without clear guidance delineating which reactive flight path movements need to be made and when. Example comments for A1 include:

1. *"Proposed primary routing was unclear in congested urban environment."*
2. *"Procedure display overview too zoomed out to validate."*
3. *"Proposed procedure executes closely-spaced 90 degree right and left turns at high airspeed which are too difficult to execute."*

(A2) Lack of information: Test subject was unable to validate the procedure due to insufficient information for the proposed departure, enroute, final, or missed approach sequence. This could be resultant from an oversimplification of an autonomous procedure or missing segments of a manual procedure preventing the test subject to "chair fly" or adequately evaluate and accept a procedure. Example comments for A2 include:

1. *"No altitude associated with enroute portion of procedure."*
2. *"Departure climb gradient not available."*

Gateway #2: Flight Monitoring

Were you able to cross-monitor flight guidance performance?

The intent of this question is to measure the ability of the test subject to adequately monitor what the aircraft is doing or is supposed to do. The evaluation includes the timing required to

adequately cross-monitor the aircraft performance between changes in airspeed, altitude, and azimuth.

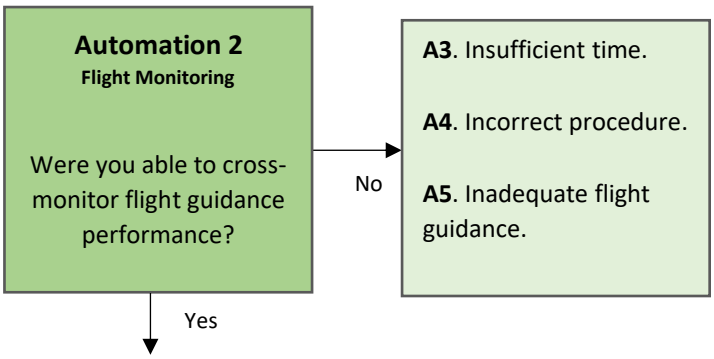

Figure 4: Flight Procedures with Automation Gateway #2: Flight Monitoring

(A3) Insufficient time: The test subject was unable to monitor the aircraft Automation or adhere to the procedure flight guidance due to insufficient time to adequately monitor changes in airspeed, altitude or azimuth. This could result from procedure design error truncating the distance between waypoints that required rapid or simultaneous changes in flight that did not provide enough time for the test subject to safely and efficiently cross monitor the aircraft's performance. Example comments for A3 include:

1. *"Waypoints are too closely spaced together"*
2. *"Assigned airspeed is too fast for the designed procedure."*

(A4) Incorrect procedure: The test subject was unable to monitor the autopilot's performance due to an error in the procedure, coding or flight guidance system. This could occur from a divergent flight guidance in which the procedure calls for an opposite direction than what the aircraft is executing due to an error in coding. This could also be inaccurate spatial data in which the test subject observes the aircraft flying in close proximity to terrain, vertical obstruction or any airspace restrictions. This could also be due to a rounding error in the FMS that the test subject observes the aircraft offset from the intended landing point. Example comments for A4 include:

1. *"Aircraft executed a right turn when procedure called for a left turn."*
2. *"Aircraft flew in close proximity to obstacle (antenna). Required higher climb gradient on departure to maintain adequate clearance."*
3. *"Aircraft was unacceptably offset from landing touchdown point."*

(A5) Inadequate flight guidance: The test subject was unable to effectively monitor the aircraft's performance due to insufficient information. This could be from a lack of information in the pilot's display or the instrument approach plate. The test subject was unable to appropriately cross monitor the Automation due to a lack of primary navigation,

performance, or control indicators available in the cockpit. Example comments for A5 include:

1. *"No time, distance, or heading information available on navigational display."*
2. *"Mismatched information between active waypoint and active leg."*
3. *"No airspeed indicated on flight display."*

Gateway #3: Workload

Were you able to look 1-2 waypoints ahead?

This question measures the ability of the pilot/operator to "stay ahead" of the aircraft while monitoring the aircraft's automation. Specifically designed to assess the pilot workload from cross-monitoring waypoint to waypoint to include the ability to manage the flight by anticipating the next transition or flight sequence.

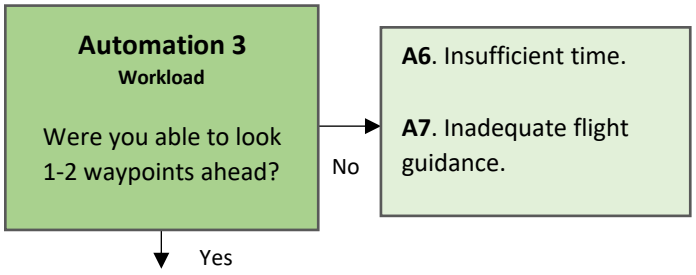

Figure 5: Flight Procedures with Automation Gateway #3: Workload

(A6) Insufficient time: The test subject was unable to maintain situational awareness in the cockpit with an inability to "stay ahead" of the aircraft. This can occur when the pilot is task saturated and is focused on the current operations while unable to evaluate or monitor upcoming events. This answer can be selected if the pilot had appropriate flight guidance ahead of the active leg but felt an inadequacy to ingest additional information beyond the current procedure sequence. Example comments for A6 include:

1. *"Task saturated, complete attention required to keep up with the active leg."*
2. *"Assigned airspeed is too fast or waypoint (deltas) are too closely spaced."*

(A7) Inadequate flight guidance: The test subject was unable to maintain situational awareness ahead of the aircraft due to an insufficient amount of information. The test subject did not have the appropriate waypoint designation, time, distance or heading available to adequately make an informed decision or to accept or reject proposed routing. This could result from the candidate navigational orientation displaying only active leg sequences. Example comments for A7 include:

1. *"No available information beyond active leg."*

2. *"Navigational display set in hemispheric orientation and therefore unable to see next waypoint beyond 180 degree field of view."*

Gateway #4: Pilot Training

Expected pilot training

This question is designed to engage the test subject's opinion of the level of training or experience required to successfully complete the procedure and/or operation. The question measures the extent to which the pilot/operator would need to be familiar with the aircraft, navigation system or additional air service provider.

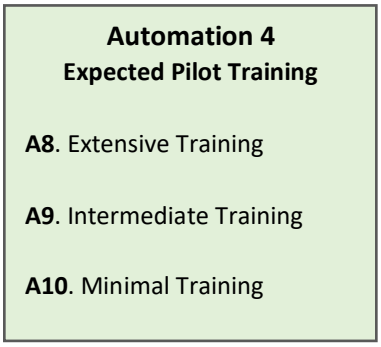

Figure 6: Flight Procedures with Automation Gateway #4: Pilot Training

A8) Extensive training: The test subject felt the tasks could be accomplished but would require a rated pilot with extensive experience and targeted training to achieve a safe level of operation and integration into the National Airspace System. An example comment for A8 includes:

1. *"Extensive training required. Suggested minimum rating be an Airline Transport Pilot (ATP) and update FOI (Fundamentals of Instruction) to include new training methodology."*

(A9) Intermediate training: The test subject felt that the tasks could be accomplished by a rated pilot with additional training to achieve a safe level of operation and integration into the National Airspace System. An example comment for A9 includes:

1. *"Recommend Instrument level of training for adequate pilot proficiency."*

(A10) Minimal training: The test subject felt that the tasks could be accomplished by a student pilot or rated pilot with little to no experience given current levels of training to achieve a safe level of operation and integration into the National Airspace System. An example comment for A10 includes:

1. *"Private pilot equivalent level of training is sufficient for procedure execution."*

4.2 Manual Flight Procedures

The PARM is evaluated through the following structure for manual flight procedures:

Initial Question

Was Automation Utilized?

For the purpose of this evaluation, automation is defined as the ability of an aircraft to perform control and navigation functions without input by the pilot or remote operator. Automation includes auto-pilot, pilot-assist features, or complete aircraft Automation.

Figure 7: PARM Initial Question

Gateway #1: Procedure Validation

Was the procedure validation framework for piloting waypoint to waypoint navigation acceptable?

The second question is designed to evaluate the pilot's ability to review and confirm the validity of the proposed procedure for manual execution. Given dynamic airspace routing, the intent of this question is to assess the ability of the test subject to ingest and understand the proposed operation (expected to be packaged in a different way from traditional or legacy operations). This question includes the departure, enroute, final approach and missed approach components of the proposed flight.

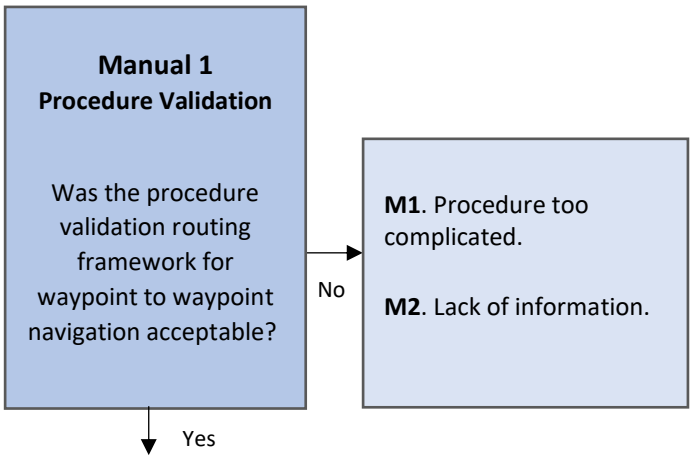

Figure 8: Manual Flight Procedures Gateway #1: Procedure Validation

(M1) Procedure too complicated: Test subject was unable to validate the procedure due to an inability to comprehend or understand the given instructions. This could be from too

many routing options (contingencies) closely spaced or overlaid together. Additionally there could be compound movements in altitude, airspeed, and azimuth without clear guidance delineating which movements need to be made and when. Example comments for M1 include:

1. *"Primary routing was too complex with surrounding terrain and vertical obstructions."*
2. *"Procedcure descent gradient is too steep."*
3. *"Proposed procedure demands airspeed variations that may be unsafe."*

(M2) Lack of information: Test subject was unable to validate the procedure due to insufficient information for the proposed departure, enroute, arrival, or missed sequence. This could be an oversimplification of a procedure or missing segments of a manual procedure preventing the test subject to "chair fly" or adequately evaluate and accept a procedure. Example comments for M2 include:

1. *"No defined missed approach segment."*
2. *"RNP value absent for final approach segment."*

Gateway #2: Flight Execution

Were you able to adhere to flight guidance?

The intent of this question is to measure the ability of the pilot to adequately fly the proposed procedure and measure an ability to adhere to or "keep up" with the low- level truncated transactions in airspeed, altitude and azimuth that are designed to be operated at the desired tempo of the UAM operational use case.

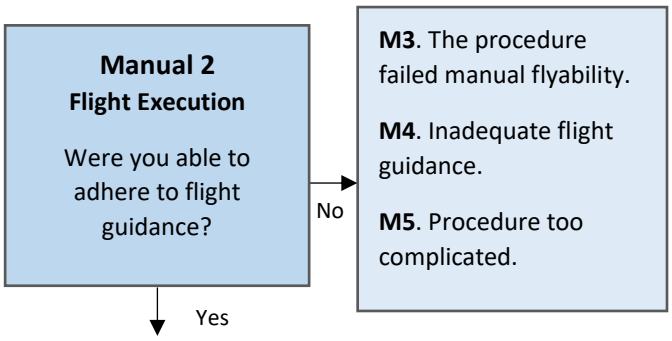

Figure 9: Manual Flight Procedures Gateway #2: Flight Execution

(M3) The procedure failed manual flyability: The test subject was unable to maintain manually control of the aircraft or adhere to the procedure flight guidance. This could result from procedure designer error truncating the distance between waypoints that required rapid or simultaneous changes in flight that did not provide enough time for the test subject to safely or effectively maintain aircraft control. Example comments for M3 include:

1. *"Maneuvers exceed vehicle limitations."*
2. *"Felt like a knife fight in a phone booth making the tight turn to final."*
3. *"Inadequate waypoint spacing which did not allow vehicle to stabilize."*

(M4) Inadequate flight guidance: The test subject was unable to adhere to the flight guidance due to insufficient or unsatisfactory direction provided for the procedure. This could occur when a lack of information is provided by the approach plate or is unavailable to the pilot in the cockpit. The lack of information in either direction, contingency, or latency prevented the test subject from safely executing the procedure. Example comments for M4 include:

1. *"Confused on which direction to turn from final."*
2. *"Lack of information from landing surface to taxiway to parking."*

(M5) Procedure too complicated: The test subject was unable to safely fly the candidate procedure due to its complexity. This could be caused by procedure designer error selecting an incorrect lower altitude causing compound deltas in rapid succession to avoid terrain, vertical obstructions, and/or airspace restrictions. Examples comments for M5 include:

1. *"Maneuver required multiple altitude, airspeed and azimuth deltas."*
2. *"Target altitude unachievable via descent gradient at designated airspeed."*

Gateway #3: Workload

Were you able to manage procedure workload?

This question assumes adequate manual control of the aircraft via inputs by the pilot/operator. It is intended to measure the spare mental capacity of the test subject to maintain peripheral tasks while engaged with the flight procedure. This task defines adequacy as successfully performing radio operations, traffic monitoring, landing check list, etc. in addition to flying duties.

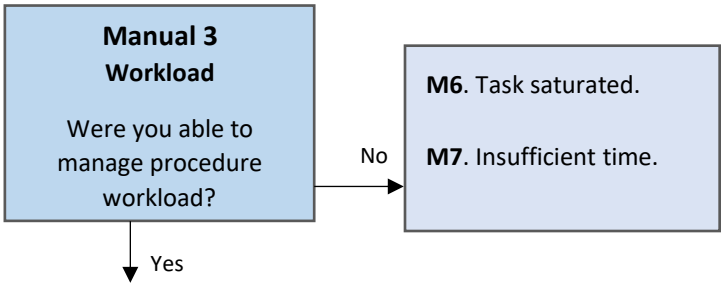

Figure 10: Manual Flight Procedures Gateway #3 Workload

(M6) Task Saturated: The test subject was unable to maintain situational awareness in the cockpit and was not able to "stay ahead" of the aircraft. This could be from pilot task saturation with focus to maintain waypoint to waypoint navigation within acceptable limits. This answer can be selected if the pilot had appropriate flight guidance ahead of the active leg but felt an inability to consume additional information beyond the current procedure sequence. Example comments for M6 include:

1. *"Task saturated; complete attention required to maintain heading, altitude and airspeed."*
2. *"Was able to maintain aircraft control within procedure limits but missed a radio call."*

(M7) Insufficient time: The test subject was unable to maintain situational awareness ahead of the aircraft due to an insufficient amount of time allocated between pilot required inputs. The test subject did not have enough time to review succeeding waypoints due to an inappropriately selected airspeed for the given procedure. Or an adequate airspeed was selected but an inadequate distance was used to separate waypoints. Example comments for M7 include:

1. *"Assigned airspeed too fast or waypoint (deltas) too closely spaced."*
2. *"Was unable to review or select optimal alternate routing due to the time allotted to review and accept upcoming route changes."*

Gateway #4: Pilot Training

Expected pilot training

This question is designed to engage the test subject's opinion of the level of training or experience required to successfully complete the procedure and/or operation. The question measures the extent to which the pilot/operator would need to be familiar with the aircraft, navigation system or additional air service provider.

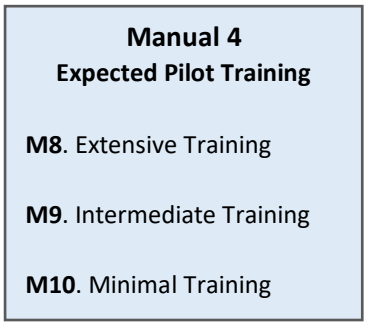

Figure 11: Manual Flight Procedures Gateway #4: Pilot Training

(M8) Extensive training: The test subject felt the tasks could be accomplished but would require a rated pilot with extensive experience and targeted training to achieve a safe level of operation and integration into the National Airspace System. An example comment for M8 includes:

1. ***"****Extensive training required. Suggest minimum rating be an Airline Transport Pilot (ATP) and update FOI (Fundamentals of Instruction) to include new training methodology."*

(M9) Intermediate training: The test subject felt that the tasks could be accomplished by a rated pilot with additional training to achieve a safe level of operation and integration into the National Airspace System. An example comment for M9 includes:

1. *"Recommend Instrument level of training for adequate pilot proficiency."*

(M10) Minimal training: The test subject felt that the tasks could be accomplished by a student pilot or rated pilot with little to no experience given current levels of training to achieve a safe level of operation and integration into the National Airspace System. An example comment for M10 includes:

1. *"Private pilot equivalent level of training is sufficient for procedure execution."*

5 Conclusion

As with any rating matrix, there will be influence and variation from the evaluator, as the scoring and test subject's experience with the scale will vary from test to test. Dynamic procedure automation is uncharted territory, and, therefore, no "norms" for data interpretation exist. However, building upon the MCH and BWS, live flight and simulation constructs will need to be explored to support operational integration, derive averages, and apply detailed evaluator comments for workload, training, and flyability.

Keeping to NASA's objectives to research safe UAM operations in an integrated and scalable airspace architecture, the NC subproject successfully developed a matrix to evaluate candidate instrument approach procedures. The Procedure Automation Rating Matrix (PARM) could be an effective tool to validate candidate procedure applicability. Resultant scoring will elicit the test pilot's or operator's opinion on flyability and projected training requirements tailored to a given research aircraft or aircraft configuration. Direct feedback from test pilots and operators will be crucial for airspace procedure designers and the development of new and novel AAM airspace constructs.

References

[1] R. P. Harper, Jr., and G. E. Cooper, "Handling qualities and pilot evaluation" Journal of Guidance, Control, and Dynamics, vol. 9, pp. 515-529, 1986.

[2] P. F. Borchers, J. A. Franklin, and J. W. Fletcher, Flight Research at Ames, 1940-1997: Fifty-even Years of Development and Validation of Aeronautical Technology (NASA SP-3300) Moffett Field, CA: NASA, Ames Research Center, 1998.

[3] M.L. Cummings, Kevin Myers, Stacey D. Scott, “Modified Cooper Harper Evaluation Tool for Unmanned Vehicle Displays” Moffett Field, CA: NASA, Ames Research Center, 2008.

Appendix A: Acronyms and Abbreviations

This appendix contains acronyms that are used repeatedly throughout this document.

Acronym	Term
AAM	Advanced Air Mobility
BWS	Bedford Workload Scale
eVTOL	electric Vertical Takeoff and Landing
FOI	Fundamentals of Instruction
MCH	Modified Cooper-Harper rating scale
NAS	National Airpsace System
NASA	National Aeronautics and Space Administration
NC	National Campaign
PARM	Procedure Automation Rating Matrix
UAM	Advanced Air Mobility