5-7-2019

A Framework for Assessing the Impact of Performance-Based Navigation on Air Traffic Controllers

R. Jordan Hinson

Michael W. Sawyer

Gary L. Serfoss

A FRAMEWORK FOR ASSESSING THE IMPACT OF PERFORMANCE-BASED NAVIGATION ON AIR TRAFFIC CONTROLLERS

R. Jordan Hinson, Michael W. Sawyer, Gary L. Serfoss
Fort Hill Group, LLC
Washington, DC

The air traffic control domain is undergoing significant modernization efforts through technology and procedure enhancements. Understanding the impact of these changes and ensuring these enhancements do not unduly increase controller cognitive workload is essential for success. This research provides a framework for assessing human performance impacts and cognitive workload associated with Performance-Based Navigation for use in an operational air traffic control environment. A panel of human factors and air traffic control subject matter experts assessed a broad set of measures of cognitive workload based on sensitivity, bandwidth, diagnosticity, selectivity, interference, controller acceptance, reliability, and implementation requirements. This resulted in a set of recommended operationally-viable measures of controller cognitive workload. Additionally, a series of potential human performance impacts associated with Performance-Based Navigation were identified. The benefits and limitations of each measure are summarized along with guidance for tailoring the recommended measures based on a research objectives and operational constraints.

The changes introduced by PBN procedures present a wide range of direct and indirect impacts to human performance for both air traffic controllers and flight crews. Achieving the potential benefits associated with PBN procedures requires that controllers and flight crews can effectively assign, execute, modify, and monitor the procedures. Considering the wide range of impacts to controller performance, of particular interest are those that may adversely impact a controller's cognitive workload. Cognitive workload represents just one of the elements of human performance that may be directly or indirectly impacted by PBN procedures. Excessive levels of cognitive workload have been shown to adversely impact human performance in air traffic control and many other similar domains. This paper presents one piece of a larger framework developed to equip the Federal Aviation Administration (FAA) to consistently assess and mitigate the effects of cognitive workload on controller performance. Measuring and managing controller cognitive workload may support the FAA in developing more effective PBN procedures, increasing PBN utilization rates, and ensuring that future technology and procedure changes reduce or do not unduly increase controller cognitive workload.

Methodology

A literature review was first conducted to identify a candidate set of cognitive workload measures for consideration. Each measure identified from the literature was categorized based on five potential measure types: Primary Task (Pri.), Secondary Task (Sec.), Physiological (Phy.), Subjective (Sub.), and Analytical (Ana.) ((Stanton, Salmon, & Rafferty, 2013) (Wilson &

Corlett, 2005). For each measure type, the related measures, source documents, and a brief measure summary were catalogued (Sawyer, Hinson, & Henderson, 2017). To identify which measures would be best given the defined project scope, researchers devised a system for assessing the measures to account for the following criteria developed from the literature: sensitivity (combined with bandwidth), diagnosticity, selectivity, interference, controller acceptance, reliability (combined with transferability), and implementation requirements (Wickens & Hollands, 1999) (Wierwille & Eggemeier, 1993). Scoring criteria definitions were also defined as presented below in Table 1. A workgroup consisting of air traffic control and human factors subject matter experts was then convened to review and rate each measure using the scoring criteria. A consensus approach was taken by the workgroup to assign a value of +1, 0, or -1 for each of the 7 scoring criteria.

Table 1.
ATC Measure Scoring Criteria and Definitions.

Criteria	**1**	**0**	**-1**
Sensitivity *(Sen)*	Measure distinguishes fairly rapid changing levels of cognitive workload, or task load without risk of the measures saturating.	Measure shows moderate variation in task / workload. Scale may become saturated but remains useful to a point.	Measure shows only sensitivity to extreme variations in workload. Measure reaches saturation quickly.
Diagnosticity *(Dia)*	Measure allows the cause of variation in cognitive workload to be identified, or indicates which cognitive resources are most affected.	Measure indicates minimal cause of variation in cognitive workload.	Measure does not indicate cause of variation in workload.
Selectivity *(Sel)*	Measure allows various confounding factors such as noise, physical workload, and emotional stress, to be distinguished from variations in cognitive workload.	Measure accounts for most causes of variation, but may not distinguish some confounding factors or noise.	Measure includes confounding effects which cannot be isolated.
Interference *(Int)*	Measure does not affect primary task performance.	Measure has minimal effect on primary task.	Measure significantly impacts primary task.
Controller Acceptance *(CA)*	Controllers likely accept measure.	Controller is neutral on measure.	Controllers likely reject use of measure.
Reliability *(Rel)*	Measure has documented research of use in ATC.	Measure has documented research with limited use in ATC.	Measure has very little development or validation.
Implementation Requirements *(Imp)*	Neither additional equipment nor specialized personnel are required. Training is minimal.	Minimal equipment or specialized personnel is required.	Significant equipment or specialized personnel are required.

Results

The 33 highest scoring, viable measures are provided in Table 2. Full details on measure identification, assessment, and prioritization are available by technical report (Sawyer et al., 2017).

Table 2.
ATC Measure Assessment Results.

Measure Name	Type	Sen	Dia	Sel	Int	Rel	Imp	CA	**Total**
NASA Task Load Index (TLX)	Sub.	0	1	1	1	1	1	1	**6**
Communications Data	Pri. Ana.	1	1	0	1	0	1	1	**5**
Coordination / Communication Rating	Sub.	1	0	0	1	0	1	1	**4**
Simplified Subjective Workload Assessment Technique (SWAT)	Sub.	0	1	-1	1	1	1	1	**4**
Trajectory-based complexity (TBX)	Ana.	1	0	0	1	0	1	1	**4**
ATC Tape Communication Analysis	Ana.	1	0	-1	1	0	1	1	**3**
Localized traffic density	Ana.	0	0	-1	1	1	1	1	**3**
Number of Handoffs	Ana.	0	0	-1	1	1	1	1	**3**
Simulator Test Score of Performance	Pri.	1	-1	-1	1	1	1	1	**3**
Subjective Workload Assessment Technique	Sub.	0	1	-1	1	1	1	0	**3**
ATC Complexity Measurement	Ana.	0	0	0	1	0	0	1	**2**
Bedford Scale	Sub.	0	-1	0	1	0	1	1	**2**
Checklist to Evaluate Airspace Complexity	Ana.	0	0	-1	1	0	1	1	**2**
Communication time, message length	Ana.	1	0	-1	1	-1	1	1	**2**
Communications Efficiency	Ana.	1	0	-1	1	-1	1	1	**2**
Handoff Acceptance Latency	Ana.	1	-1	-1	1	0	1	1	**2**
Hart & Hauser Rating Scale	Sub.	0	1	1	-1	-1	1	1	**2**
Mental Workload Index (MWLI)	Pri. Ana.	0	0	0	1	-1	1	1	**2**

Measure Name	Type	Sen	Dia	Sel	Int	Rel	Imp	CA	**Total**
Number of control actions	Ana.	0	-1	-1	1	1	1	1	**2**
Performance and Objective Workload Evaluation Research (POWER)	Ana.	1	-1	-1	1	1	0	1	**2**
Projective SWAT	Sub.	0	1	-1	1	-1	1	1	**2**
SME / Over-the-shoulder ratings	Sub.	0	-1	-1	1	1	1	1	**2**
Time required	Ana.	1	-1	0	1	0	0	1	**2**
Air Traffic Workload Input Technique (ATWIT)	Sub. Sec.	0	-1	-1	1	1	1	1	**2**
Behavioral Markers	Pri. Ana.	0	-1	0	0	0	1	1	**1**
Continuous Subjective Assessment of Workload (C-SAW)	Sub.	1	-1	-1	1	0	0	1	**1**
Number of aircraft under control per hour / traffic count	Ana.	0	-1	-1	1	0	1	1	**1**
Recall Ability	Sec.	1	-1	-1	-1	1	1	1	**1**
Respiration	Phy.	0	-1	0	1	0	0	1	**1**
Subjective Workload Dominance (SWORD) Technique	Sub.	0	-1	-1	1	0	1	1	**1**
Task Analysis Workload (TAWL)	Pri. Ana.	-1	1	0	1	-1	0	1	**1**
The Projective SWORD Technique (Pro-SWORD)	Sub.	0	-1	-1	1	0	1	1	**1**
Thermo-vascular activities	Phy.	1	-1	0	1	-1	0	1	**1**

PBN Human Performance Impacts

A list of identified human performance impacts related to cognitive workload impacted by PBN procedures were gathered from a review of operational safety reports from the Aviation Safety Reporting System (ASRS), research studies, industry guidance documents, and interviews with Human Factors and ATC subject matter experts. The resulting list of PBN human performance impacts are grouped into 11 categories listed below in Figure 1.

Aircraft Performance	Impacts caused by aircraft deviating from the expected flight path, altitude, and or speed.
ATC Automation	Impacts caused by ATC automation that supports controlling traffic and supporting tasks (map display, flight-plan processing, eligibility, etc.).
Acceptance	Impacts characterized by mistrust in a PBN procedure if it is perceived as less efficient, less safe, flawed, or otherwise inferior to previous or conventional routes.
Communications	Impacts characterized by the effect of PBN procedures on the coordination and communication among air traffic service users including Air Traffic Controllers, Flight Crews, Airport Operators, Traffic Management, etc.
Mixed Equipage	Impacts caused by aircraft using RNAV navigation in the same environment as aircraft using conventional navigational capabilities.
Nominal Operations	
Design of Airspace Procedures	Impacts characterized by the design elements of PBN procedures (speed, course, altitudes, etc.) and interactions with other elements of the airspace (other routes, airspace boundaries, etc.)
Recovery	Impacts characterized by how PBN procedures affect a controller's response to an event that could lead to an adverse outcome.
Monitoring	Impact affecting how a controller monitors the airspace.
Training	Impacts relating to how training is conducted, including when it occurs, how often, what type, and its effectiveness.
Weather & Wake	Impacts to human performance caused by the presence and management of adverse weather conditions and the effects of wake turbulence.

Figure 1. PBN Human Performance Impact Categories.

Conclusion

The PBN cognitive workload assessment framework includes tools for assessing the impact of PBN procedures including recommended measures of cognitive workload. Further selection using the scoring matrix results of each measure resulted in the following recommended PBN Workload Measures for operational use: NASA Task Load Index (TLX), Trajectory Based Complexity Calculation (TBX), Communication Efficiency Rating, and Number of Handoffs. Alternative measures may be selected using the scoring matrix results to match specific research needs and constraints. Additionally, the framework recommends using interviews and impact surveys to assess potential human performance impacts associated with Performance-Based Navigation.

For more information on this framework and the tools supporting this research, see PBN Cognitive Workload Analysis Results Report (Hinson, Serfoss, & Sawyer, 2018b). For a detailed analysis of the science of Cognitive Workload, the pros and cons of the many various evaluation methods, and analysis and discussion of which cognitive workload tools seem most applicable and usable in the PBN air traffic controller environment, see PBN Cognitive Workload Analysis Plan (Sawyer et al., 2017). For a complete guide to the framework providing the appropriate tools and instructions for understanding, analyzing, and beginning to mitigate the impact of PBN procedures on controller performance including a full list of PBN Human Performance Impacts, see the report (Hinson, Serfoss, & Sawyer, 2018a).

Acknowledgements

Fort Hill Group would like to thank the FAA NextGen Human Factors Division (ANG-C1) for supporting and funding this research and Bill Kaliardos for providing technical guidance throughout the project. Additionally, we would like to thank the aviation safety subject matter experts who provided the valuable insight necessary to developing these results. The results presented herein represent the methodology and results of this research project and do not necessarily represent the view of the Federal Aviation Administration.

References

Hinson, R. J., Serfoss, G., Sawyer, M. W. (2018a). *Framework for Collecting and Utilizing PBN Human Performance Metrics.* Washington, DC: delivered to the Federal Aviation Administration with the intent of future publication.

Hinson, R. J., Serfoss, G., Sawyer, M. W. (2018b). *PBN Cognitive Workload Analysis Results Report.* Washington, DC: delivered to the Federal Aviation Administration with the intent of future publication.

Sawyer, M. W., Hinson, R. J., Henderson, A. S. (2017). *PBN Cognitive Workload Analysis Plan.* Washington, DC: Fort Hill Group, Delivered to the Federal Aviation Administration with the intent of future publication.

Stanton, N., Salmon, P. M., & Rafferty, L. A. (2013). Human factors methods: a practical guide for engineeringand design. *Ashgate Publishing, Ltd.*

Wickens, C. D., & Hollands, J. G. (1999). *Engineering psychology and human performance (3rd ed.).* New Jersey: Prentice Hall.

Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors, 35(2)*, 263-281.

Wilson, J. R., & Corlett, N. (Eds.) (2005). *Evaluation of Human Work.* CRC Press.