Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
April 29, 2021
A Regional Conference of the Society for Industrial and Systems Engineering

Leveraging Aviator Training Next to Select Higher Quality Aviators

Marie Docken, Jennifer Gervais, Blake Watson, Vince Wilkinson, Collin Winstead, and Clay Woody

Department of Systems Engineering
United States Military Academy
West Point, NY

Corresponding Author's Email: marie.docken@westpoint.edu

Author Note: The authors would like to thank Department of Systems Engineering at West Point and the United States Army Aviation Center of Excellence for supporting this research. Specifically, the authors would like to give a special thank you to Lieutenant Colonel Erika Salerno, Deputy Director of Organization and Personnel Force Development (OPFD) that oversees the Aviator selection process, Major Chris McFarland, Project Manager for Aviator Training Next, Chief Warrant Officer Five Joe Roland, 2nd Aviation Flight Detachment Commander, and Chief Warrant Officer Five Chadwick Ford, former US Army Chief Tactical Operations Officer.

Abstract: The United States Army Aviation Center of Excellence (USAACE) is committed to developing leaders through the production of aviation professionals. In 2018, USAACE established an innovative research and training program known as Aviator Training Next (ATN). ATN's purpose is to research and implement the capabilities of commercial off-the-shelf virtual reality technology, to enhance the capacity and quality of USAACE's production pipeline. Last year, ATN validated the use of its system as a viable means of training aviators on entry-level flight tasks. These research findings prompted USAACE leadership to explore ATN's capacity as an assessment tool. This paper provides a use case of the ATN system as an augmentative means to enhance the aviation branch selection process. By extending the capabilities of ATN to the branch selection process, the Army gains an assessment tool capable of objectively identifying those candidates with an increased propensity for service as a future Army Aviator.

Keywords: Aviator Training Next, Virtual Reality Simulation, Assessment

1. Introduction and Background

USAACE is focused on recruiting, assessing, training, and retaining the aviators of the Army. However, current demand and lucrative job opportunities within the commercial aviation industry have made it difficult for the Army to sustain a viable population of trained pilots (Report to Congress, 2019). USAACE leadership realizes with a resulting shortage in aviators, it is imperative to improve the branch's assessment capabilities. To achieve this, USAACE leadership wants to explore the assessment capabilities of the new Aviator Training Next (ATN) technology. Utilizing ATN's technology as an assessment tool provides the Army with a means to select a higher quality candidate through assessments in an aviation-centered environment.

Aviator Training Next is an innovative approach to enhance the flying capability of current and future aviators. ATN uses commercial off-the-shelf (COTS) simulation technology, which includes an immersive virtual reality (VR) headset, cyclic, collective, and cockpit seat, to replicate a flight scenario with low to moderate fidelity. In the training environment, ATN supplements real flight hours with virtual flight hours. The primary intent is to maximize training task repetitions while minimizing the required resources commensurate with live flight training. In a controlled experiment conducted at Fort Rucker between 2019 and 2020, researchers found no significant difference in the performance evaluation scores of those student populations that trained in VR when compared to the population that trained exclusively in the live aircraft (Woody, 2019). This research validated ATN as an effective training tool. Based on these findings, USAACE prompted its research team to explore the capacity of ATN as an assessment tool.

Based on stakeholder analysis and research, increased capacity for effective decision-making during periods of high cognitive loading was identified as a key performance indicator of a high-quality aviator (O'Connor & Cohn, 2010). The current branch assessment process reduces cognitive assessments to a candidate's undergraduate academic GPA and Selection Instrument for Flight Training (SIFT) exam score. While these are effective metrics of intelligence, they are, at best, a proxy metric for determining a candidate's cognitive load capacity. To develop a more comprehensive metric for cognitive loading, this paper explored the use of ATN as an augmentative assessment to be used in conjunction with the SIFT exam score and GPA. This was accomplished by conducting stakeholder analysis, identifying key skills that suggest enhanced aptitude for aviation, and developing the framework for an aviator assessment battery. Additionally, a design of experiments was developed to prescribe how future research teams will validate the effectiveness of the assessment battery and its ability to identify candidates with a higher propensity for success in the aviation branch.

Stakeholder analysis and a review of literature highlighted the importance of cognitive ability in successful aviation officers (O'Connor & Cohn, 2010). Thus, the branching process must capture the significance of cognition. ATN offers a unique way to evaluate a candidate in an immersive aviation environment. Augmented with the SIFT exam score and other current branching assessments, ATN improves the current candidate selection process to ultimately increase the effectiveness of the branch. The following systems design process describes the process towards this conclusion.

2. Methodology

2.1 Assumptions

Currently, the aviation branch has no clear definition of what makes a good aviator. Thus, this research looked to a review of literature and stakeholder analysis to best define a good aviation officer. Based on these, this research defines a good aviator as a leader first, pilot second, with high cognitive ability to include problem solving skills, quick decision making and level headedness under stress. Next, this research assumes that ATN possesses sufficient fidelity to real flight despite its fixed-wing model software. As such, it validates ATN's assessment of high-quality aviators for the branch. Literature supports the validity of this assumption asserting low level fidelity tasks can be trained using low level fidelity systems (Stewart, 2008). Candidates will only be asked to complete the low-level tasks included in the ATN assessment scenarios. Lastly, it is assumed that minimal training of candidates on ATN will still enable valid data collection. This is necessary to assess the aviation potential of inexperienced candidates with varying learning curves in ATN. Highly experienced Instructor Pilots from the Second Aviation Detachment, including CW5 Roland, will validate the amount of training necessary for data collection on ATN.

2.2 Systems Thinking

Systems thinking is a holistic mental framework and world view that recognizes a system as an entity first, as a whole, with its fit and relationship with its environment being primary concerns. (Parnell, 2007). It captures the dynamics of systems and the systems complexities. The system environment for this project includes many operational elements within four categories: flight school, Organization and Personnel Force Development (OPFD), flight qualification, and cadet performance. These categories interact with each other to form the system that is the aviator selection process. Candidates formulate their branching profiles for review and pass the SIFT and medical exams before receiving their assignment as an aviator. Currently, ATN plays a role in flight school to allow additional flight repetitions with more simulation-based training. The ATN team is looking at using ATN in a new capacity in the earlier stages of candidate qualification and the branching profile to select higher quality aviators. In this capacity, ATN will be used like the SIFT to identify candidates who possess a greater aptitude for aviation service. One way to better understand the environment of the system is through creating a system-i-gram which can be seen in Figure 1. The information in the system-i-gram was gathered through stakeholder meetings and interviews with personnel in the Aviation branch.

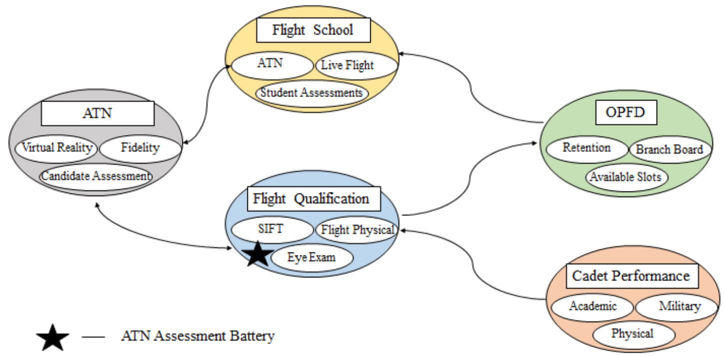

Figure 1. System-i-Gram of Aviator Training Next

2.3 Stakeholder Analysis

The first step in developing an assessment tool capable of assisting in the selection of higher quality aviators was developing a better understanding of the aviation environment and the key factors that aviators need to be successful. This was accomplished by interviewing stakeholders and researching the literature surrounding the assessment of aviators. These stakeholders included Major McFarland, Chief Warrant Officer 5 Ford, and Lieutenant Colonel Salerno. Producing a stakeholder network diagram allows the visualization of the interlinked relationships between all stakeholders in the system. The stakeholder network diagram for this system is seen in Figure 2.

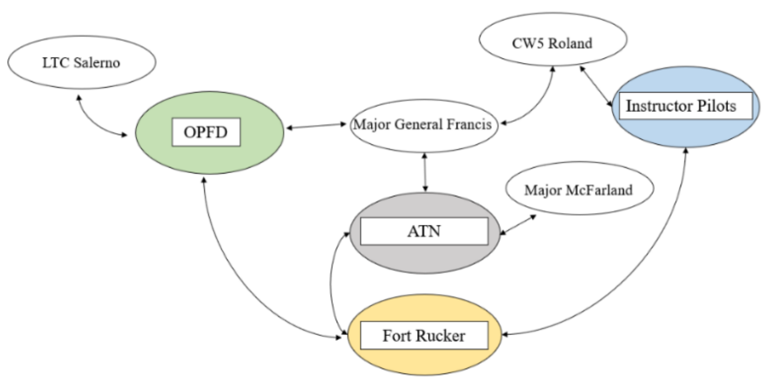

Figure 2. Stakeholder Network Diagram

The key findings from the interviews were sorted into an FCR matrix which is an integral part of systems thinking. Stakeholders indicated that the current assessments of aviators—the SIFT, Talent Assessment Battery (TAB), and branch interview—perform well in determining a candidate's aviation potential; however, the branch needs an assessment which focuses on a candidate's ability to manage and utilize their cognitive assets in complex and changing environments. Research showed that the cognitive abilities most indicative of enhanced aviator performance include motor coordination, tracking, operating with divided attention, and working memory. The literature researched included multiple meta-analyses of aviator assessments that were aimed at discovering which assessments and aviation skills were most correlated to an aviator being successful (Martinussen, M. 1996). We were able to use what we learned from these dialogues and the literature to begin the development of aviation assessments that could be conducted using the ATN flight simulator.

2.4 Assessment Generation

To generate assessments, this research team utilized a Zwicky's Morphological Box with categories: maneuver, stimulus, test type, and goal. *Maneuver* is either simple or complex and helps assess candidates' divided attention. Next, *Stimulus*, offers different signals such as radio queues or lights in the cockpit to test motor coordination, working memory, and divided attention. Candidates may also be in flying or not flying in the assessments, options for *Test Type*, which highlights the fidelity ATN simulators provide as a unique assessment tool. Finally, *Goal* addresses how each candidate will be tested in the assessments, either through flying through hoops or maintaining the nose of the helicopter on a moving target. This category highlights candidate adaptability to different motor coordination tests. The assessment generation matrix below displays final assessments chosen after feasibility screening.

Table 1. Assessments

	Maneuver	Stimulus	Test Type	Goal
Complex Coordination	Complex	Out of Cockpit	Flying	Moving Target
Discrimination Reaction Time	Complex	In Cockpit	Flying	Hoops
Rotary Maneuver	Complex	Mixed	Flying	Hoops

For the Complex Coordination assessment, a candidate must keep the aircraft pointed at a lead helicopter throughout flight (North & Griffin, 1977). The lead helicopter will move more erratically as the test progresses, challenging the candidate's ability to effectively keep the aircraft on target through lever and medal manipulation. Thus, this assessment tests motor coordination. Distance veered off target and time spent on target throughout the assessment serve as performance measures for this assessment.

In the Discrimination Reaction Time assessment, a candidate will fly through a simple pattern of hoops while receiving light stimuli in the cockpit to adjust switches and buttons on the controls (North & Griffin, 1977). The stimulus prompts will progressively become more intricate, requiring the candidate to produce correct responses in a specific order quickly while still maneuvering through hoops. This assessment tests a candidate's ability to effectively manage cognitive load and work with divided attention. Response time to stimuli serves as the performance measure for this assessment.

In the final assessment, Rotary Maneuver, a candidate will again maneuver the aircraft through hoops at different altitudes in the air. First, the candidate will receive no stimuli. However, the candidate will fly the course again, this time while receiving the stimuli from the Discrimination Reaction Time assessment as well as listening and responding to radio prompts. This assessment tests a candidate's ability to maneuver their aircraft while responding to additional stimuli. The performance measure of this assessment is the difference in time between the first and second course completion times.

3. Design of Experiment

After generating assessments, this experiment will be integrated and evaluated in the West Point environment. Although ideally this test would be available to every potential candidate for aviation, constraints such as the number of ATN devices available, time available to train and test candidates, and the style of tests programmed in the ATN devices limit the scope of this experiment. As such, this experiment to assess future aviators must be carefully designed within these limits.

To assess the validity of this experiment, this research must compare the assessed potential for flight school success with actual flight school performance. To do so, researchers will first determine candidates for the experiment. For proper experimentation, subjects must be randomized and organized to limit the impact of outside variables on performance. Since stakeholders are interested in people committed to branching Aviation, tests will begin for Second-Class cadets who have enrolled in the SIFT. Although this limits the initial pool of subjects, the overall weighted sum metric for assessing candidates includes SIFT data. Furthermore, stakeholder analysis identified an interest in Aviation as an important factor of a good aviator. Therefore, this research will only select individuals who have shown a vested interested in the branch through participation in the SIFT.

Candidates will be randomly blocked into three homogeneous groups based on the number of training hours they will complete on the simulator: no hours, 0-1 hours, and more than 1 hour of training time. This will control for variation caused by differences in training time. Based on initial analysis of the device and constraints on device availability, most cadets will receive limited training hours. After familiarization with the device, each candidate will conduct Complex Coordination, Discrimination Reaction Time, and Rotary Maneuver assessments. Assessments will be given in the same sequence to candidates to limit additional variability in results.

This model builds a composite score for each candidate based on their performance in the assessments as well as data from other assessments the Aviation branch identifies as important in determining future aviators such as the GPA, TAB, SIFT, and branch interview scores per candidate. The resulting model scores will rank candidates' aptitude for the branch based on this research's definition of a good aviation officer.

First, each performance measure from the different tests will be normalized to be scored against the others. Since many performance measures have different units or scales, as seen in Table 1, each performance measure must be normalized to make direct comparisons between them. To normalize, each individual score per candidate is evaluated as its percentage of the ideal value, either the maximum or minimum. Taking "Distance from Target" as an example, a better aviator is someone who stays on target. Therefore, the minimum distance from target is preferred. Table 2 identifies the ideal values for each performance measure further below.

Table 2. Description of Performance Measures

Assessment	**Performance Measure**	**Units**	**Ideal Value**
Complex Coordination	Distance from Target	Meters	Minimum
	Time on Target	Seconds	Maximum
Discrimination Reaction Time	Response to Stimulus Accuracy	Percentage	Maximum
Rotary Maneuver	Difference in Time between Flights	Seconds	Minimum
Other Assessments	GPA	Constructed	Maximum
	TAB Score	Constructed	Maximum
	SIFT Score	Constructed	Maximum
	Branch Interview Score	Constructed	Maximum

Since preliminary candidates for this experiment will have no flight school experience, this model lacks a response variable. Therefore, this model uses the rank order centroid (ROC) method to determine weights for each performance measure. Based on stakeholder analysis and review of literature regarding cognitive load, this model ranks the performance measures as follows: response to stimulus accuracy, difference in time between flights, distance from target, time on target, GPA, SIFT score, TAB score, and branch interview score. The following formula represents the ROC method:

$$w_j = \left(\frac{1}{M}\right)\sum_{n=j}^{M}\frac{1}{n} \tag{1}$$

where M is the number of performance measures and w_j is the weight for the j^{th} performance measure as in Table 3.

Table 3. ROC Weights for Performance Measures

Performance Measure	Rank	Weight
Response to Stimulus Accuracy	1	0.34
Difference in Time between Flights	2	0.21
Distance from Target	3	0.15
Time on Target	4	0.11
GPA	5	0.08
SIFT Score	6	0.05
TAB Score	7	0.03
Branch Interview Score	8	0.02

Once the performance measures are normalized, they are aggregated into a metric to determine who is a good future aviator. The total score of the i^{th} candidate, S_i, is calculated using a weighted sum:

$$S_i = \sum_{j=1}^{n} s_{ij} w_j \tag{2}$$

where s_{ij} is the score of the i^{th} candidate on the j^{th} performance measure, and w_j is the importance of the j^{th} performance measure. With this information, this research can compare candidates' likelihood to be good aviators.

ISBN: 97819384962-0-2

4. Preliminary Results and Future Research

A proof-of-concept trial including current USMA cadets selected for Aviation, the ATN system, and Instructor Pilots (IPs) from the 2nd Flight Detachment at USMA validated the design of experiment research. The testing population of approximately twenty cadets received forty minutes of flight instruction from the IPs followed by a twenty-minute assessment phase. This assessment required cadets to fly through a series of hoops while answering questions to induce heightened cognitive load in flight. After, IPs subjectively scored cadets based on their overall flight performance while managing additional load and stressors throughout flight. Through this proof of concept, this research validates that ATN can be utilized as an assessment tool for West Point cadets. Moving forward, to evaluate candidates as described in the design of experiment above, ATN technology must be adjusted, to include new testing scenarios and additional audiovisual cues, to support the three proposed assessment batteries. With these adjustments, future research teams can specifically assess the desired traits of high-quality aviators and provide the Aviation branch additional insight into potential future aviation officers.

5. Conclusion

In recent years, Army Aviation has faced numerous challenges, including declining retention amongst aviation officers; in response to this, we have been tasked with designing a measure of assessment to answer, "what makes a good future aviator?" Drawing from key findings in stakeholder analysis, we determined that ATN has the potential to augment the current accessions process by measuring a candidate's cognitive abilities. From this and further research into cognitive load, we determined that working memory, complex coordination, and operating with divided attention are three measures we would like to leverage in our assessment. Thus, three assessments were designed with these three measures in mind. Utilizing statistical analysis, we developed a weighted sum equation that gives a composite score for each candidate based on their ATN performance and other assessment scores, such as GPA, TAB, SIFT, and branch interview. From this, we offer a first step for Army Aviation to develop the highest quality officers to lead the branch to better face the challenges of tomorrow.

6. References

Martinussen, M. (1996). Psychological Measures As Predictors of Pilot Performance: A Meta-Analysis. *The International Journal of Aviation Psychology*, *6*(1), 1–20. https://doi.org/10.1207/s15327108ijap0601_1

Myers, M. (2019, August 15). The pilot shortage: The Army's struggle to fix its aviation problems. Retrieved November 16, 2020, from https://www.armytimes.com/news/your-army/2019/08/07/the-pilot-shortage-the-armys-struggle-to-fix-its-aviation-problems/

O'Connor, P. E., & Cohn, J. V. (2010). *Human performance enhancement in high-risk environments: Insights, developments, and future directions from military research*. Santa Barbara, CA: Praeger Security International/ABC-CLIO

Parnell, G. (2007). Decision Making in Systems Engineering and Management.

Randel, B. (2020, June 24). The Army Needs a Better Solution for its Pilot Shortage. Retrieved November 18, 2020, from https://mwi.usma.edu/army-needs-better-solution-pilot-shortage/

Report to Congressional Armed Services Committees on Initiatives For Mitigating Military Pilot Shortfalls (Rep.). (2019). Office of Publication and Security Review.

North, R., & Griffin, G. (1977). *Aviator Selection 1919-1977* (Rep.). Pensacola, Florida: Naval Medical Research and Developmental Command.

Stewart. J. (2008, April). *Fidelity Requirements for Army Aviation Training Devices: Issues and Answers.* US Army Research Institute.

Woody, C., MAJ, Goldie, C., LTC, O'Brien, K., Aura, C., & Delgado-Howard, C. (2020). *Aviator Training Next Technical Report #1: Performance and Demographic Analysis.* U.S. Army Aeromedical Research Laboratory & Operations Research Center.