

NASA/TP-2009-215707



Evaluation of Fused Synthetic and Enhanced Vision Display Concepts for Low-Visibility Approach and Landing

*Randall E. Bailey, Lynda J. Kramer, Lawrence J. Prinzel III, and Susan J. Wilz
Langley Research Center, Hampton, Virginia*

April 2009

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, and organizing and publishing research results.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at <http://www.sti.nasa.gov>
- E-mail your question via the Internet to help@sti.nasa.gov
- Fax your question to the NASA STI Help Desk at 443-757-5803
- Phone the NASA STI Help Desk at 443-757-5802
- Write to:
NASA STI Help Desk
NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320

NASA/TP-2009-215707



Evaluation of Fused Synthetic and Enhanced Vision Display Concepts for Low-Visibility Approach and Landing

*Randall E. Bailey, Lynda J. Kramer, Lawrence J. Prinzel III, and Susan J. Wilz
Langley Research Center, Hampton, Virginia*

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

April 2009

Available from:

NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320
443-757-5802

Table of Contents

List of Figures.....	3
List of Tables.....	4
Symbols and Abbreviations.....	5
1. Introduction.....	6
2. Background.....	7
2.1 Integrated and Fused Synthetic and Enhanced Vision Systems Concepts.....	8
2.2 Previous Research.....	9
2.2.1 Pilot-Flying Role.....	11
2.2.2 Pilot-Monitoring Role.....	12
2.3 Current Research.....	13
2.4 Experiment Objectives.....	13
3. Methodology.....	15
3.1 Experiment Method.....	15
3.2 Simulator.....	15
3.3 Head-Up Display.....	16
3.4 Auxiliary Display.....	16
3.5 Head-Down Displays.....	16
3.6 Synthetic Vision Database.....	17
3.7 Out-the-Window (OTW) Scene.....	17
3.8 Enhanced Vision System.....	17
3.9 Symbology.....	17
3.9.1 HUD Symbology.....	17
3.9.2 HUD Tunnel.....	18
3.9.3 Auxiliary Display Symbology.....	19
3.10 Display Concepts.....	19
3.10.1 Head-Up Display Concepts.....	19
3.10.2 Auxiliary Display Concepts.....	21
4. Procedures.....	24
4.1 Evaluation Task.....	24
4.2 EVS Crew Procedures.....	25
4.3 Experiment Matrix.....	26
4.4 Measures.....	26
4.5 Procedure.....	27
5. Results.....	29
5.1 Path Control Performance.....	29
5.1.1. Approach Lateral Path Error.....	29
5.1.2. Approach Vertical Path Error.....	30
5.1.3 Vertical and Lateral Path Error on Final.....	30
5.1.4 FTE on Approach.....	30
5.1.5. Path Control Results Discussion.....	31
5.2 Mental Workload.....	32
5.2.1. AFFTC Workload Estimate – PF HUD.....	33
5.2.2 AFFTC Workload Estimate – PNF-AD.....	33
5.2.3. SWORD.....	33
5.2.4 Workload Discussion.....	34
5.3 Situation Awareness.....	35
5.3.1 SART – PF HUD.....	35
5.3.2. SART – PNF-AD.....	35
5.3.3. SA-SWORD.....	35
5.3.4. Situation Awareness Discussion.....	36
5.4 Pilot Display Preferences.....	37

5.5 Head-Tracking Analysis	37
5.6 Subjective Assessments of Display Clutter	40
5.6.1 HUD Concept Ratings	40
5.6.2 PNF-AD concept Ratings	43
5.7 Fusion Controls.....	43
5.8 Non-Normals	45
5.8.1 Runway Incursions	46
5.8.2 Navigation Error	50
5.8.3. Illegal Landings	51
6. Conclusions	53
7. References	54
Appendix A: Head Tracker Data Analysis Details	58
Appendix B. Final Questionnaire Subject Ratings and Comments	62

List of Figures

Figure 1. Synthetic vision and enhanced vision comparison.....	8
Figure 2. Integration Flight Deck simulation facility with HUD and AD.....	16
Figure 3. HUD symbology with advanced pathway guidance.....	18
Figure 4. HUD runway outline symbol and glideslope reference line.....	18
Figure 5. Head-Up Display (HUD) formats on approach at 700 feet AGL.....	21
Figure 6. Four Auxiliary Display (AD) formats on final.....	23
Figure 7. Sparks Visual arrival to RNO runway 16R.....	24
Figure 8. Post-run display clutter questions.....	27
Figure 9. RMS lateral path error during approach segment for each HUD concept.....	29
Figure 10. RMS vertical path error during approach segment for each HUD concept.....	30
Figure 11. Means of SART ratings for HUD symbology and raster combinations.....	36
Figure 12. Percentage of time where PNFs directed gaze.....	39
Figure 13. Mean ratings of demand on visual attention clutter metric for each HUD concept.....	41
Figure 14. Mean ratings of supply of visual attention resources clutter metric for each HUD concept.....	42
Figure 15. Fusion control settings by altitude.....	44
Figure 16. SV/EV fusion and EV-Only setting by altitude.....	45
Figure 17. Runway incursions (baggage cart and firetruck) as seen at 60 ft AFL using various displays.....	48
Figure 18. Lateral navigation offset examples.....	51

List of Tables

Table 1. EV and SV Comparison	8
Table 2. EVS Crew Procedures	25
Table 3. Flight Path Tracking Performance Summary	31
Table 4. Percentage of Time PNF Directing Gaze at Various Head Locations	38
Table 5. Percentage of Time PNF Directing Gaze, By Display Configuration	40
Table 6. Above Field Level (AFL) Altitudes for Theoretical Detection of Runway Incursions	46
Table 7. Illegal Landing Trials	52
Table A-1. ANOVAs on Percentage Dwell Time at the Auxiliary Display Concepts for each Altitude Segment.....	58
Table A-2. Student-Newman-Keuls Test for Altitude Segment: > 1000 ft AFL.....	59
Table A-3. Student-Newman-Keuls Test for Altitude Segment: 500 < AFL < 1000 ft.....	59
Table A-4. Student-Newman-Keuls Test for Altitude Segment: 400 < AFL < 500 ft.....	60
Table A-5. Student-Newman-Keuls Test for Altitude Segment: 300 < AFL < 400 ft.....	60
Table A-6. Student-Newman-Keuls Test for Altitude Segment: 200 < AFL < 300 ft.....	60
Table A-7. Student-Newman-Keuls Test for Altitude Segment: 100 < AFL < 200 ft.....	61

Symbols and Abbreviations

AD	Auxiliary Display
AFFTC	Air Force Flight Technical Center
AFL	Above Field Level
AGL	Above Ground Level
ANOVA	Analysis of Variance
CRM	Crew Resource Management
DH	Decision Height
EFVS	Enhanced Flight Vision System
EP	Evaluation Pilot
EV	Enhanced Vision
EVO	Equivalent Visual Operations
EVS	Enhanced Vision System
FAR	Federal Aviation Regulation
FLIR	Forward Looking InfraRed
FPM	Flight Path Marker
FTE	Flight Technical Error
HDD	Head-Down Display
HUD	Head-Up Display
IFD	Integration Flight Deck
IIFDT	Integrated Intelligent Flight Deck Technologies
IMC	Instrument Meteorological Conditions
LaRC	Langley Research Center
MDA	Minimum Descent Altitude
MMWR	Millimeter Wave Radar
NASA	National Aeronautics and Space Administration
ND	Navigation Display
NextGen	Next Generation Air Transportation System
nm	nautical mile
OTW	Out The Window
PF	Pilot Flying
PFD	Primary Flight Display
PNF	Pilot Not Flying
RI	Runway Incursion
RMS	Root-mean square
RNO	FAA airport identifier for Reno/Tahoe International Airport
RNP	Required Navigation Performance
SA	Situation Awareness
SART	Situation Awareness Rating Technique
SA-SWORD	Situation Awareness – Subjective Workload Dominance
SV	Synthetic Vision
SVS	Synthetic Vision System
SWORD	Subjective Workload Dominance
VFR	Visual Flight Rules
VMC	Visual Meteorological Conditions

ABSTRACT

NASA is developing revolutionary crew-vehicle interface technologies that strive to proactively overcome aircraft safety barriers that would otherwise constrain the full realization of the next generation air transportation system. A piloted simulation experiment was conducted to evaluate the complementary use of Synthetic and Enhanced Vision technologies. Specific focus was placed on new techniques for integration and/or fusion of Enhanced and Synthetic Vision and its impact within a two-crew flight deck during low-visibility approach and landing operations. Overall, the experimental data showed that significant improvements in situation awareness, without concomitant increases in workload and display clutter, could be provided by the integration and/or fusion of synthetic and enhanced vision technologies for the pilot-flying and the pilot-not-flying. Improvements in lateral path control performance were realized when the Head-Up Display concepts included a tunnel, independent of the imagery (enhanced vision or fusion of enhanced and synthetic vision) presented with it. During non-normal operations, the ability of the crew to handle substantial navigational errors and runway incursions were neither improved nor adversely impacted by the display concepts. The addition of Enhanced Vision may not, of itself, provide an improvement in runway incursion detection without being specifically tailored for this application.

1. Introduction

The United States air transportation system is undergoing a transformation to accommodate a projected 3-fold increase in air operations by 2025.¹ Technological and systemic changes are being developed to significantly increase the capacity, safety, efficiency, and security for this Next Generation Air Transportation System (NextGen). One of the key capabilities envisioned to achieve these goals is the concept of Equivalent Visual Operations (EVO), whereby Visual Flight Rules (VFR) operational tempos and also, perhaps, operating procedures (such as separation assurance) are maintained independent of the actual weather conditions. One methodology by which the EVO-goal might be attainable is to create a virtual visual flight environment for the flight crew, independent of the actual outside weather and visibility conditions, through application of Enhanced Vision (EV) and Synthetic Vision (SV) technologies.

An experiment was conducted to evaluate the complementary use of SV and EV technologies, specifically evaluating the utility, acceptability, and usability of integrated/fused enhanced and synthetic vision technologies and its effect on two-crew operations. This work begins the development of an all-weather commercial aviation operations capability, approaching that which might create an EVO capability.

2. Background

The Integrated Intelligent Flight Deck Technologies project, under the National Aeronautics and Space Administration (NASA) Aviation Safety Program, is comprised of a multi-disciplinary research effort to develop flight deck technologies that mitigate operator-, automation-, and environment-induced hazards. Towards this objective, crew/vehicle interface technologies are being developed that reduce the propensity for, and minimize the risks associated with, pilot error while proactively overcoming aircraft safety barriers that would otherwise constrain the full realization of the NextGen. Part of this research effort involves the use of SV and EV systems and other interface modalities as enabling technologies to meet these challenges.

SV is a computer-generated image of the external scene topography, generated from aircraft attitude, high-precision navigation, and data of the terrain, obstacles, cultural features, and other required flight information. SV provides significant improvements in terrain awareness and reductions in the potential for Controlled-Flight-Into-Terrain incidents/accidents compared to current cockpit technologies.²⁻⁵

EV is an electronic means to provide a display of the external scene by use of an imaging sensor, such as a Forward-Looking InfraRed (FLIR) or millimeter wave radar (MMWR). In 2004, Section §91.175 of the US Federal Aviation Regulations (FAR) was amended such that operators conducting straight-in instrument approach procedures (in other than Category II or Category III operations) may now operate below the published Decision Height (DH) or Minimum Descent Altitude (MDA) when using an approved Enhanced Flight Vision System (EFVS) shown on the pilot's Head-Up Display (HUD). (EFVS is the terminology adopted by the FAA to indicate the operational application of an Enhanced Vision System (EVS). In the context of this report, the use of the terms EFVS and EVS are considered to be equivalent. An EFVS (and EVS) pertains to the sensor, HUD, and associated hardware and software which enable an enhancement to the pilot's natural vision during an approach and landing.) This FAR change provides "operational credit" for EV equipment. No such credit currently exists for SV.

The intended use of EV and SV technology mirror each other as they both attempt to eliminate low-visibility conditions as a causal factor to civil aircraft accidents and replicate the operational benefits of clear-day flight operations, regardless of the actual outside visibility condition. The methodologies by which this capability is achieved, however, are significantly different, leading to significantly different operational considerations. While some may consider the technologies to be competing; they are, in fact, complementary⁶ as illustrated by the system comparison in Table 1.

SV, by virtue of being weather-independent and unlimited in field-of-regard, is particularly advantageous during flight phases, such as approach, which may be obscured by clouds and precipitation of which an EV sensor may not penetrate. On the other hand, EV provides a direct view of the vehicle external environment; independent of the derived aircraft navigation solution or of a database. Under conditions of smoke, haze, and night, a FLIR/EV provides orders-of-magnitude improvement over the pilot's natural vision; greatly enhancing the pilot's situation awareness and reducing the pilot's workload. The comparison of SV and EV in Figure 1 on a night visual meteorological conditions (VMC) approach into an airfield highlights the similarities and differences in these two technologies. In the SV image, color coding is available, the features of the SV image are only dependent upon the database content and the engineering provided by the SV designer. No atmospheric or weather obscuration is present. Conversely, in the EV image, the image is critically dependent upon the EV sensor and the external environment. In this case, the FLIR provides an outstanding view on a night approach, where roads and runways are

clearly demarked from vegetation because of thermal differences. Atmospheric moisture such as clouds are clearly visible; in this case, in the far horizon of the EV image. The EV image is detail-rich and “locked” to the aircraft.

2.1 Integrated and Fused Synthetic and Enhanced Vision Systems Concepts

The complementary capabilities of SV and EV have been well-recognized⁶ with the premise that⁷ “the strengths of (the) enhanced system can compensate for the deficiencies in the synthetic system and that the strengths of (the) synthetic system can compensate for the deficiencies in the enhanced vision system.” While these goals are obvious, optimal methods and capabilities are not necessarily well defined.

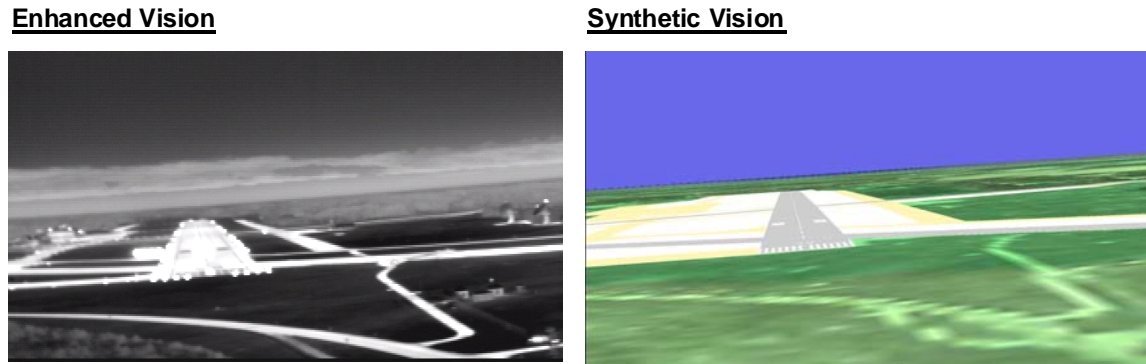


Figure 1. Enhanced Vision And Synthetic Vision Comparison.

Table 1. EV and SV Comparison

	Enhanced Vision (EV)	Synthetic Vision (SV)
Benefits For Terrain & Situation Awareness?	Significant	Significant
Affected by Outside Weather and Visibility Condition?	Yes – > Image quality dependent upon sensor characteristics and environmental conditions.	No – > Image quality independent of external conditions.
Affected By Aircraft Navigation “Solution”?	No – > System “bolted” to airframe and aligned to aircraft boresight	Yes – > Image accuracy critically depends on position and attitude information
Training Required for Image Interpretation and Understanding?	Yes – > Sensor phenomenology must be understood as image content and presentation is not predetermined	Minimal – > Image content and presentation method is predetermined by SV designer.
Real-Time Obstacle Detection?	Yes – > Real-time sensor view from aircraft; but image quality depends upon sensor and target characteristics and environmental conditions	No – > Image drawn from stored database of the terrain, obstacles, cultural features; unless augmented by real-time detection sensors.
Operational Credit?	Yes, under FAR Part 91.175.	No

2.2 Previous Research

Several studies⁸⁻¹⁰ have shown that the optimal combination of SV and EV technology provides the direct display of SV to the flight crew *without* direct display of EV, but instead, using EV “behind the glass” for navigation error detection, database integrity monitoring, and real-time obstacle/object detection. Image processing performs these functions automatically without intervention by the flight crew. This arrangement provides a highly usable display presentation (i.e., SV) that is impervious to the actual weather and visibility conditions, yet if un-charted obstacles, database errors or navigation errors are detected by the EV running in the background, the situation is annunciated, and almost “perfect” decision-making by the pilot occurs.⁸⁻¹⁰ In fact, the study conclusion from Parrish⁸ – SV “concepts should *not* be implemented without incorporating image-processing decision aides for the pilot” – launched a 5 year effort at NASA and elsewhere developing enabling technologies for database integrity monitoring and object detection (i.e., *SV Systems (SVS)*).¹¹⁻¹⁵

While degrees of success in developing these decision aids have been met, technology for “perfect” object detection and database/navigation error detection does not yet exist. Further, there may always be gaps, such as minimal radar cross-section objects or below-detection threshold errors, which may still warrant flight deck procedures and human interventions for integrity and error checks.

Research suggests that the optimal configuration of displays, controls, and formats depends upon the intended function of the system, the flight phase, and the role of the pilots^{16,17} involving a balance between:

- Understanding – how the design promotes the perception of the SV and EV information, the comprehension of its meaning, and the projection of their status into the near future.¹⁸
- Workload – how the design minimizes the physical (e.g., use of controls, visual scanning) and mental (e.g., amount of cognitive effort necessary to generate understanding) workload for effective use and understanding of the information.

These factors must not only include the “steady-state” use of EV and SV information, but the “costs” associated with transitioning to and from this visual environment and their effect during non-normal flight situations. For example, consideration must include how the design eases or impedes the transition from instrument to visual flight references and vice versa.

A key interconnection between the factors of understanding and workload is the concept of display clutter – i.e., an excessive number and/or variety of color and symbols, that obscures essential information, or presents distracting, disorganized and unnecessary information delaying visual detection.

Clutter warrants special consideration for EV/SV as the designer is intentionally increasing the volume of data presented to the flight crew. For the HUD, in particular, FAA certification policy¹⁶ formalizes this conundrum in that “clutter should be minimized” yet, “essential or critical (HUD) data must always be displayed” such as that necessary for EFVS operational credit (as per Chapter 14 of the Code of Federal Regulations, Part 91.175(l)). Thus, task-critical or essential EV/SV imagery must be displayed but the information must not excessively interfere with pilots’ ability to see the actual runway environment through the display or impede their perception, understanding, and use of the information. Declutter control in present EFVS implementations allow the pilots to selectively add or remove symbology or raster (i.e., SV or EV) imagery, but declutter controls also introduce pilot workload during task- or time-

critical situations and raise the possibility that critical information might be inappropriately removed.

A common method of combining SV/EV information is by use of an integrated single display with simultaneous presentation of unadulterated SV and EV content. Image integration in this case is defined as the use of a single display with the separate information sources being readily identifiable in the image. By visual proximity, this method minimizes the visual scan and cognitive effort in integrating the disparate information. Typically, “inset” displays of EV information have been used to take advantage of the “unlimited” field-of-regard in SV information to complement a limited field-of-regard EV sensor. Integration in this manner allows the user to retain comparison of the separate sources for improved understanding. However, careful design of the integration process must be made to prevent an unintended *loss* of understanding. Often, a control for the pilot is provided which enables modulation of the amount of EV imagery shown in the inset image compared to the SV background imagery. This functionality may be generated by many methods. One often-used method uses a simple pixel-by-pixel combination of the two information sources (a so-called “pixel-averaging” of SV and EV). As a simple illustration, the gray scale level for a inset pixel could be determined as $Inset_{ij} = aSV_{ij} + (1-a)EV_{ij}$ where the ij subscript denotes the i^{th} row and j^{th} column of the display inset image gray scale value ($Inset$), which is a function of the SV and EV image input gray scale values, and the parameter, a , is a decimal value ranging from 0 to 1, based on the pilot-control input. This technique would enable the final inset image to range between 0% SV and 100% EV to 0% SV and 100% EV). Research shows that this averaging method is problematic.¹⁰ Image degradation of the blended image was the primary factor.^{17,19} A pure pixel-averaging technique may create a situation where a poor quality (low content) sensor image obscures good synthetic data without adding any value to the image; conversely, a good quality sensor image is obscured by uncorrelated synthetic data. More sophisticated blending methods are available to minimize this problem,¹⁷ but the potential to lose information in the combination process must always be considered. Further, pilot control of the process introduces the costs of workload for control and a high potential for miss-setting of the controls.¹⁰ Simultaneously showing both images on one display also introduces clutter, degrading the effective conveyance and useage of the displayed information.

Display clutter can be mitigated by *not* displaying both EV and SV simultaneously on the same display using one of two methods:

- Spatial separation – by locating SV and EV information on different displays. This process forces the pilot to look across displays and mentally perform the information integration. This methodology has been demonstrated under NASA flight test, wherein SV information was presented on a head-down primary flight display, and EV information was presented on the HUD. The pilots transitioned to the HUD at a specific height above the landing runway. This methodology was found to be acceptable although not necessarily without improvements being desired. Sufficient information must be on the HUD to restrict the need for the Pilot Flying (PF) to go “head-down” to acquire task critical information; otherwise, HUD usage can create “a loss of situation awareness” and additional workload from visual scanning. Designs utilizing visual momentum may assist in the integration.¹⁹
- Temporal separation – by displaying *either* SV or EV information on the same display, using automatic or manual selection for which source is displayed. An automatic transition has been tested and it felt “natural” to the pilot.²⁰ Without retaining some elements of SV or EV display, however, visual momentum (i.e., the use of symbology or imagery to provide cognitive coupling which assists the user in navigating a transition from one format or context to another) and the complementary benefits of both EV and SV may be lost.

Using a single display, image fusion may also reduce clutter. Fusion, in this context, is image processing where the separate SV and EV image sources are not readily identifiable in the resultant image. The drawback to fusion is the potential loss of understanding by the user from source comparisons and contrasts. The specific fusion methodology significantly affects the utility of the resultant display.¹⁷ The use of pixel-averaging image integration of two sensors was shown to generally provide more features in the resultant image than either sensor alone but it also yielded lower overall image contrast than that of either individual input. In comparison, feature-level fusion methods – that is, image processing techniques that are specifically designed to concentrate on areas which contain higher image contrast and/or greater spatial power - generally enabled higher contrast in the resultant image while retaining the individual source information. Unfortunately these fusion methods may also pass through some undesirable noise content of the sensor inputs. In either case, registration of the two image sources is of paramount importance for fusion.

Another method of EV/SV fusion displays SV-derived information symbolically, using symbols, icons, and/or perspective shapes for conformal overlay with EV imagery. For instance, SV-derived symbology, in the form of runway outlines, with simulated EV imagery provided visual momentum yet didn't create any perceptual confusion even when flying with significant navigational (intentional) errors of 1500 ft and 6000 ft inserted into the SV-derived guidance.²² Similar results have been found elsewhere.²¹

2.2.1 Pilot-Flying Role

EV/SV integration on head-down displays permits the use of color to differentiate and declutter the respective data attributes.²³ In addition to color, head-down displays do not require conformality but their limitations must be recognized.²⁴ The head-down to head-up transition for the landing task is critical, however, and the literature shows a clear, consistent pilot preference for flight information to be displayed head-up during approach and landing operations.⁶

EV imagery must be displayed on the HUD – if it is critical or essential flight information to the piloting task – but the information must not excessively interfere with pilots' ability to see the actual environment. Excessive information might also exacerbate cognitive tunneling²⁵⁻²⁷ and clutter typically penalizes performance.²⁸ The HUD provides many other advantages due to its conformal display of symbology (i.e., scene-linked) and imagery.²⁹⁻³¹ However, its display characteristics – monochromatic, variable lighting background, and limited luminance capabilities – introduce constraints, especially on potential options for clutter control.^{24,32-33} For instance, color cannot be used to differentiate different information sources.

The presence of EV/SV imagery (e.g., terrain) for the PF has been found to *not* be a significant factor in flight path performance when compared to the overwhelming influence of flight guidance symbology^{6,34} unless, however, the guidance information is obscured or hindered by the presence of SV and/or EV imagery.²² In this latter case, performance degrades in the presence of EV/SV information. More typically, terrain imagery has been regarded as supplemental information that complements guidance information or the EV/SV imagery provides required visual references allowing certain tasks or operations, such as EFVS.³⁵

In addition to influence on flight path control, the complementary use of EV/SV information for the PF to perform navigation integrity checks and obstacle/object detection has been evaluated in simulation and, in general, performance is relatively mediocre.⁸⁻¹⁰ In an evaluation of an integrated SV/EV Head-Down Display (HDD) concepts⁸ (using an EV inset), four pilots (~10% of the non-normal trials) flew to

touchdown on the synthetic runway image, essentially ignoring a lateral navigation error in the EV (actual) inset image. Similarly, when presented with runway incursions, 3 pilots missed the incursions (~12.5% of the non-normal trials) when using this same display concept. These results would likely be influenced by the EV/SV integration and fusion methods employed but conclusive cause-and-effect has not been identified.

EV imagery has been touted as providing an element of traffic detection above and beyond unaided natural vision. While this seems intuitive, several factors must be considered to realize this capability. First, imaging sensors are not perfect. Detection and recognition is dependent upon the sensor resolution and its sensitivity to the environment and “target”, the size and contrast of the obstacle/object against the image background, and the display resolution.^{6,36-37} Second, if the PF display uses flight vector information, the flight path marker symbol and guidance symbol, if provided, may obscure an object.¹⁶ Finally, the pilot-flying’s attention is primarily focused on the landing task, and the PF may not have sufficient attentional resources for “unexpected” scenarios. Theunissen⁹ found that if the PF had to perform navigation integrity checking using EV/SV information, the situation was rated as “unacceptable” and resulted in unacceptable workload.

2.2.2 Pilot-Monitoring Role

While the pilot-flying is primarily focused on flying the aircraft, the pilot-not-flying (PNF) or pilot-monitoring in a two-crew operation *does* have the primary responsibility for monitoring and verifying flight path performance, cross-checking guidance and raw data indicators, and identifying the visual runway environment and clearing the landing area.³⁷ Unfortunately, like the PF data, the use of EV/SV information by the PNF to perform the navigation integrity, obstacle detection, and runway incursion detection results in relatively mediocre performance. The results, again, appear to be influenced by the EV/SV integration and display methods employed, but during low-visibility auto-landings¹⁰ where most included an “anomaly” consisting of SV database misalignment, runway incursions, and uncharted obstacle, pilot-monitoring evaluations showed that:

- Using an EV-only HUD, 57% of the obstacles were undetected and 50% of the runs with the database misalignment went undetected. 38% of the runway incursions went undetected. (Despite this relatively poor performance, the pilots subjectively preferred this head-up implementation overall, mirroring PF display preferences.)
- A Primary Flight Display (PFD) concept, showing SV with a pilot-controllable EV inset image, exhibited the worst performance for anomaly detection. The PFD inset concept “missed” 100% of the uncharted obstacles - all leading to hazardous situations - and a 33% rate of missed detections for the database misalignments. Detection of runway incursions showed a missed detection rate of 33%. The poor performance and low-acceptance of the head-down EVS inset concept was attributed to the clutter of the image, the small EV image size, and confusion between the SV/EV images.
- With EV-only shown on a HDD, all uncharted obstacles were successfully detected (but 20% of runs continued into hazardous conditions because the evaluation pilots (EPs) didn’t recognize the severity of the threat) yet 33% of the runs missed the database misalignment. A missed detection rate of 22% for runway incursions was found. This concept was well liked for its large image size and minimal display clutter, but disliked because of the workload in transitioning from head-down to head-up flight.

A preference for control of the EV sensor image in an inset display has been shown elsewhere⁹ but pilot workload was shown to increase because of the additional pilot control for display management.

2.3 Current Research

Some trends for optimal presentation of SV/EV information in the pilot-flying or pilot-monitoring roles may be inferred from past research cited above. However, to conclusively infer effectiveness upon crew resource management (CRM) or optimal display configurations might be misleading or inappropriate.

In particular, the effects of EV/SV usage on crew resource management and crew coordination, in a two-place cockpit, has not been studied extensively. Crew coordination in an EFVS approach operation was the primary focus of one study,³⁸ but the results of this work may lack generalizability because of issues uniquely associated with their use of millimeter wave radar as an EV sensor, such as the development of a perspective image from radar data and the lack of vertical (height) EV information. Anecdotal evidence from flight demonstrations suggest that “a centrally-located display of EVS imagery” may facilitate CRM and training.³⁹ The most extensive evaluation of PNF display requirements for SV/EV operations showed some promising findings, but the test was run with the full expectancy for failures/non-normals. Also, the test was not a “crew” experiment.³¹

2.4 Experiment Objectives

The experiment reported herein was conducted to evaluate the complementary use of SV and EV technologies. The overall objective of this experiment was to test the utility, acceptability, and usability of integrated/fused EV and SV technology concepts in a two-crew commercial or business aircraft cockpit.

The present study focused on the complementary use of SV and EV during EFVS operations as defined under 14 CFR §91.175 since this operational credit provides a logical path for evolutionary expansion toward EVO. Under these rules, EV is presented head-up to the pilot-flying (Captain) whose primary responsibilities include flight path control, recognition and identification of the EV “visual references” under §91.175(l), and recognition and identification of the “visual references” under §91.175(l)(4), without reliance on the enhanced flight vision system. In this operation, the PF remains in control of the aircraft through the approach and landing. This procedure precludes a transfer of control at a time when such a transfer could interfere with safe landing of the aircraft. (This operation contrasts with a “monitored approach” where the Captain takes on the role of pilot-monitoring and the First Officer takes on the role as pilot flying. Prior to or at minimums, the Captain (left seat pilot) either commands a go-around or takes control and lands.)

Five specific experimental objectives, presented in the following, were addressed. By conducting a two-crew experiment, the interaction of these factors on crew-coordination and decision-making were also examined. The experiment was developed based on the past research cited above and as required by the assumed EFVS operation:

1. Investigate the effect of SV information on EFVS operations. Unlike these previous works, this experiment examined two aspects of adding SV to EV. First, a Required Navigation Performance (RNP)-type approach was flown, using a curving, descending arrival into a terrain-challenged airport. Since weather and other atmospheric conditions might negate a useful EV image during

this flight phase, SV may provide potential benefit as has been shown in other works previously; however, in this case, the operation transitions to a EVFS operation, which requires that SV not degrade the EV imagery. As such, EV imagery was shown on the HUD for the PF. SV imagery was, therefore, not shown near decision height and for some period of time before that (as discussed in Section 3) to ensure that the required EV imagery was unaffected. This transition from SV to EV imagery was automated. Integration of SV with EV using inset “windows” was not used, although it has been found to be viable,^{5,7} because performance suffers from clutter and deficiencies related to manual control of the inset imagery.^{8,10} The automation methodology was designed to minimize display clutter and reduce PF workload and the potential for human-error compared to a manually controlled operation.

2. Investigate the influence of Pathway Guidance information on EFVS operations. Unless the guidance information is cluttered by SV and/or EV imagery,⁵ superior performance is expected from the addition of pathway guidance. However, unlike these previous works, this experiment transitioned to a EVFS operation from a curving, descending arrival into a terrain-challenged airport. As such, EV imagery was shown on the HUD for the PF and pathway information was decluttered near decision height and for some period of time before that (as discussed in Section 3).
3. Investigate the influence of scene-linked SV information. Since HUD clutter is a significant concern, particularly for an EFVS operation, minimal additional symbology or imagery was added to the HUD, especially near decision height and landing. This work investigated the presence or absence of retaining some elements of SV for an EV display and its influence on visual momentum.
4. Investigate the influence of adding symbology to a PNF display of EV sensor data during an EFVS operation. While the PF is primarily focused on flying the aircraft, the PNF has the primary responsibility for monitoring and verifying flight path performance, cross-checking guidance and raw data indicators, and identifying visual runway environment and clearing the landing area (FAA, 2003). Many of the design principles noted above carry-over, with some significant differences. In particular, a single pilot study, with the pilot serving as a PNF, showed significantly better object detection and runway incursion detection performance using an EV display that did not include symbology.¹⁰ This work did not, however, investigate the presence and absence of symbology. Further, this work is being run the expectancy on the part of the flight crews for failures and non-normals. Also, the test is a true “crewed” experiment.
5. Investigate the influence of adding SV information to a PNF display of EV sensor data during and EFVS operation. In a single pilot study, with the pilot serving as a PNF, relatively poor detection of database alignment errors or runway incursions occurred with EV-only HDD or an integrated EV/SV display using a pilot-controllable EV inset.¹⁰ This work contrasts this study, investigating a feature-level fusion of SV/EV which allows the PNF to tailor the presentation. Further, this work is being run the expectancy on the part of the flight crews for failures and non-normals. Also, the test is a true “crewed” experiment.

3. Methodology

3.1 Experiment Method

A fixed-based simulation experiment was conducted to evaluate the effect of adding synthetic vision information and advanced pathway guidance to an enhanced vision HUD for the PF during low-visibility approach and landing operations. In addition, the experiment evaluated the effect of adding synthetic vision information and symbology to the PNF's display of the EV sensor data. A two-crew experiment was conducted to assess the interaction of these display concepts on crew coordination during low-visibility approach and landing operations.

Twenty-four pilots, representing seven airlines and a major cargo carrier, participated in the experiment. An attempt was made to pair the pilots based on which airline or cargo carrier they were currently flying. This procedure minimized any standard operating procedure differences which might affect crew coordination during the experiment. All but one crew were successfully paired by current affiliation.

All participants had previous experience flying with HUDs. The subjects had an average of 1787 hours of HUD flying experience and an average of 13.8 years and 16.2 years of commercial and military flying experience, respectively. EV experience was not required although some pilots were familiar with imaging sensor technology from prior military flight experience. None of the subjects were currently flying EV-equipped aircraft.

Because of the unique nature of the test and the dichotomous nature of the PF and PNF equipment for a single HUD-equipped aircraft, the training and testing was designed to provide maximum exposure and experience to the pilots in all operations of the equipment, including PF- and PNF-unique operations.

The crews were trained and flew together as subjects in both the PF and PNF roles. One pilot started in the left seat, as Captain, and was trained as the PF. The other pilot, seated in the right seat, was simultaneously trained as the PNF. At the completion of their training in these roles, they exchanged seats and were simultaneously trained in the opposite roles.

Data collection was broken into four equal sessions. The first pilot flew as PF in the first and third session while the second pilot flew as PNF. The second pilot flew as PF in the second and fourth session, while the first pilot flew as PNF.

3.2 Simulator

The experiment was conducted in the Integration Flight Deck (IFD) simulation facility (see Fig. 2) at NASA Langley Research Center (LaRC). The IFD emulates a Boeing B-757-200 aircraft and provides researchers with a full-mission simulator capability. The collimated out-the-window (OTW) scene is produced by an Evans and Sutherland ESIG 4530 graphics system providing approximately 200 degrees horizontal by 40 degrees vertical field-of-view at 26 pixels per degree resolution.

The participants occupied the left (as PF) and right (as PNF) seats. The left seat included an overhead HUD projection unit and the right seat included an auxiliary display (AD) under the right side window (see Fig. 2).

3.3 Head-Up Display

The HUD subtended approximately 32° horizontal by 24° vertical field of view. The HUD presentation was written strictly in a raster format from a video source (RS-343) input. The input consisted of a video mix of symbology and computer-generated scene imagery (either EV or SV as described in Section 3.10.1). The symbology included “haloing” to ensure high-contrast symbology against the scene imagery background. Brightness and contrast controls were provided to the pilot. Also, the pilot had a declutter control, implemented as a push-button on the left hand horn of the PF yoke. The button cycled through three “declutter” states: 1) No declutter (full symbology and scene imagery); 2) “Raster” declutter (full symbology, no scene imagery); and 3) “Full declutter” (no HUD display).

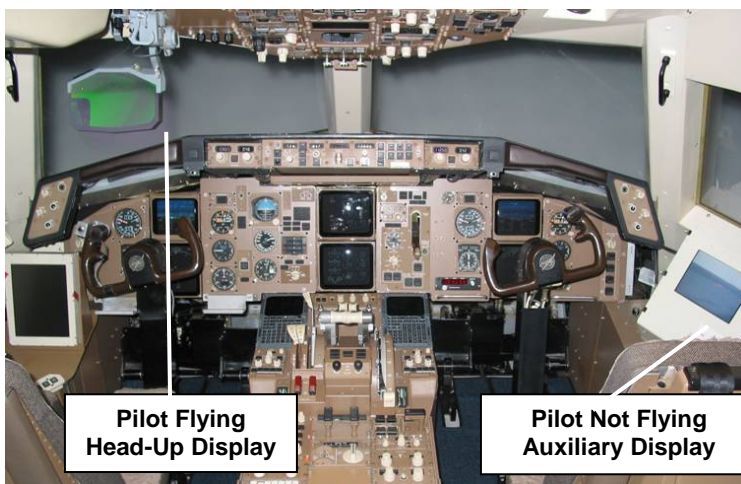


Figure 2. Integration Flight Deck Simulation Facility With HUD and AD.

3.4 Auxiliary Display

The PNF-Auxiliary Display (PNF-AD) was located outboard of the PNF location. The display was positioned as a compromise between optimal PNF viewing position, minimal display/instrument panel obscuration, and moderate installation complexity. The 8.4” diagonal display was full-color with 1024 x 768 pixel resolution. The display video source was a video mix of “haloed” symbology and computer-generated scene imagery (either EV or the output of a fused EV/SV signal as described in Section 3.10.2).

3.5 Head-Down Displays

Minimal changes were made to the PFD and Navigation Display (ND) for the experiment so they closely resembled current B-757 equipage. The PFD was only modified to include a Flight Path Marker (FPM) and guidance cue. The PFD FPM and guidance cue were driven by algorithms identical to the HUD. Standard B-757 ship’s flight director needles were disabled. No changes to the ND were made. The ND showed the Flight Management System approach but did not include Enhanced Ground Proximity Warning System “peak’s mode” nor Traffic Alert and Collision Avoidance System information.

3.6 Synthetic Vision Database

An SV database was created from a 1 arc-sec Digital Elevation Model (DEM) of a 53 by 53 nautical mile (nm) square area centered around the Reno-Tahoe International Airport (FAA airport identifier: RNO). The airport was represented by three-dimensional models of the runway, taxiways, and terminal buildings. The DEM was draped with 1 meter/pixel satellite photographic imagery within a 16 x 21 nm area centered around the airport and 4 meter/pixel imagery outside this inner region.

3.7 Out-the-Window Scene

The OTW imagery used the same source data as the SVS database but was rendered using different graphics processes and computers.

3.8 Enhanced Vision System

A physics-based FLIR simulation (using Evans & Sutherland EPX Sensors™) was created from the OTW visual database by applying materials properties to each component of the database. The characteristics of a short/mid-wave FLIR were simulated in a “white-hot” presentation. The time-of-day, time-of-year, and other diurnal properties were held constant. Atmospheric properties (cloud layer, cloud height and thickness, fog, and visibility) were varied experimentally to modulate the visibility that the evaluation pilots had in the FLIR and the OTW scene presentations (as described in Section 4.3). The EV imagery was provided in 640 horizontal by 480 vertical pixel resolution.

3.9 Symbology

3.9.1 HUD Symbology

The HUD format is shown in Figure 3. Vertical and horizontal path deviation was always shown by linear path deviation indicators (i.e., “dog-bones”) simultaneously with the standard instrument landing system (ILS) (angular) glideslope and localizer “raw data” indicators. When the dog-bone indicator was displaced one “dot” left or right of course, this corresponded to 87.5 ft vertical and 150 ft lateral linear path error. The linear dog-bone indicator was not equivalent to the angular deviation ILS indicators and the pilots were briefed and trained on the differences.

The pitch-roll guidance cue (“ball”) used modified pursuit guidance⁴⁰ along the desired path centerline, 5.5 seconds ahead of ownship. Horizontal and vertical position of the ball corresponded to the track and flight path angles to fly to the center of the desired path.

Depending upon the experimental condition, the following symbology elements were also used, as described in Section 3.10.1 A glideslope reference line was drawn (Fig. 4) using the RNO Runway 16R ILS descent angle of -3.1 degrees. Also, a runway outline symbol was drawn to conformally position the symbol based on the threshold coordinates of Runway 16R/34L based on the simulated aircraft navigation position. The symbol portrayed an 8000 ft long by 200 ft wide runway, consistent with certified Head-up Guidance Systems.

3.9.2 HUD Tunnel

As an experiment variable, advanced pathway guidance in the form of a “minimal” tunnel⁴¹ was flown (see Fig 3). The minimal tunnel concept consists of a series of “crow’s feet” which represented the truncated corners of nominally-connected rectangles spaced at 0.2 nm increments along the desired path. The tunnel portrayed a constant 300 ft wide (± 150 ft lateral) by 175 ft high (± 87.5 ft vertical) path, 1 nm ahead of ownship position, along the desired path. One dot of vertical and lateral path error (“dogbone” deviation) corresponds to the vertical and lateral extent of the tunnel, respectively.

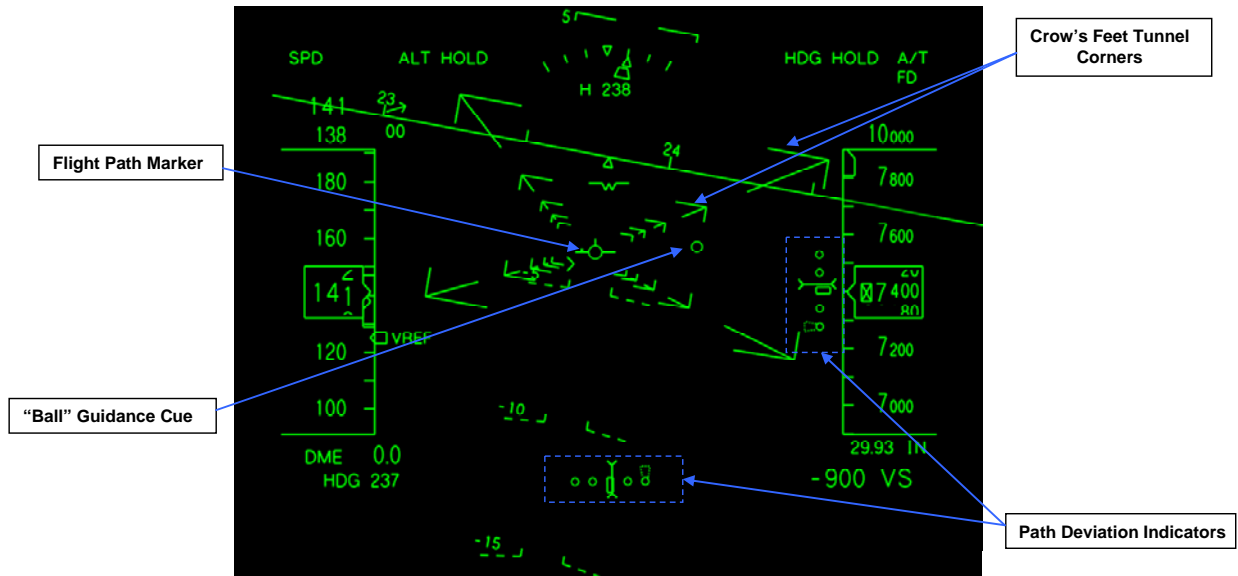


Figure 3. HUD Symbolry With Advanced Pathway Guidance.

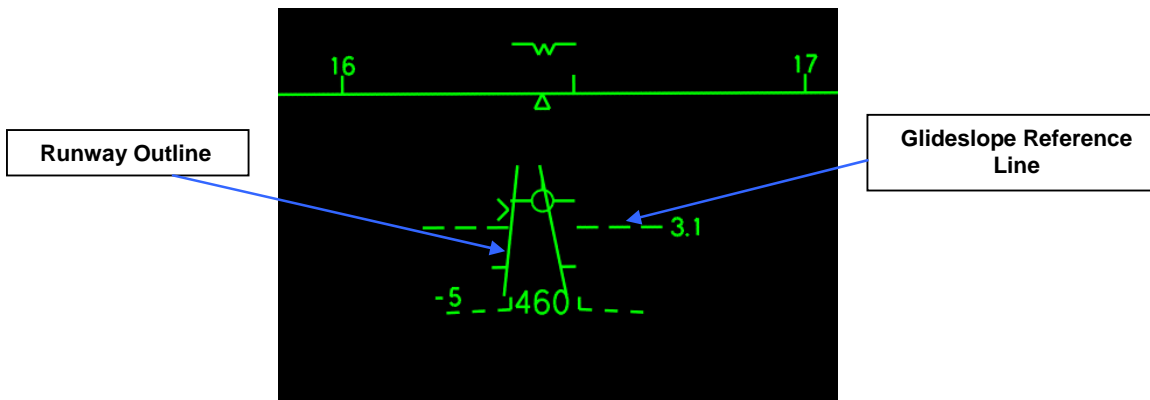


Figure 4. HUD Runway Outline Symbol And Glideslope Reference Line.

The minimal tunnel was used to minimize HUD clutter. Past studies^{2,28} have shown that sufficient path information is provided by the minimal tunnel concept – at a minimal cost of display clutter – when path deviation indicators, guidance symbology and the FPM are also provided.

3.9.3 Auxiliary Display Symbology

The auxiliary display symbology (when used) was a subset of the HUD symbology to aid the PNF in monitoring the approach without obscuring too much of the raster image. The symbology included digital readout of indicated airspeed and altitude; zero pitch attitude line (horizon line); flight path marker, pitch/roll (ball) guidance cue; path deviation indicators; ILS deviation indicators and scales; waterline; radio altitude, and event marker enunciators. (The event marker enunciators were not needed for the evaluation subjects, but were included for experimental data recording.) Alternative symbology sets to clear the center of the display were tried, but the pre-test usability results were not encouraging enough to move forward with the concepts.

3.10 Display Concepts

Four HUD concepts and four AD concepts were evaluated by the evaluation crew (PF and PNF) while flying approaches to Reno-Tahoe Airport, Runway 16R. The head-down display formats were invariant.

3.10.1 Head-Up Display Concepts

Four PF-HUD display concepts were tested, differing from each other by: 1) two types of raster background; and, 2) two types of symbology.

Two HUD raster (background) formats were flown:

- 1) EV-only (hereafter referred to as “FLIR”).
The FLIR concept represented our “baseline” HUD. In this configuration, the simulated FLIR output was exclusively displayed whether useful imagery was being provided or not.
- 2) Fused SV/EV (hereafter referred to as “Fusion”).
The Fusion raster started out as unadulterated SV imagery, transitioning through a fused SV/EV presentation beginning at 600 feet above field level (AFL), and ending with an unadulterated FLIR imagery by 500 feet AFL. Between 600 feet and 500 feet AFL, the fusion levels changed in 10% step increments from 100% SV and 0% EV ending at 0% SV and 100% EV, over a period of approximately 10 seconds.

Each raster concept showed FLIR alone below 500 ft to enable the operational credit now offered by EFVS. The 500 ft transition altitude was chosen from a usability study prior to the test and flight experience² as the altitude:

- 1) after which FLIR would be required yet with sufficient time to become accilimated to the FLIR imagery,
- 2) that provides SV imagery to assist in establishing stabilized approach parameters up to and beyond the minimum recommended Instrument Meterological Conditions (IMC) stabilized approach altitude of 1000 ft AFL; and,
- 3) at or just after the recommended minimum VMC stabilized approach altitude to allow full utilization of EV. The 100 ft transition between SV and EV imparts visual momentum between the concepts for the PF HUD.

The “fusion” concept provided the basis to evaluate the utility, acceptability, and usability of SV and EV on the HUD. This methodology was based on the literature review cited in Section 2 to maximize image legibility and limit pilot manual control requirements for an EFVS operation as the PF. In this methodology, SV and EV are shown in the flight regimes where they are most advantageous to the PF, and pilot-controllable fusion methods are not implemented as the PF is already over-burdened with controlling the HUD brightness, contrast, declutter, and FLIR control. The fusion concept and its behavior, including its transition from SV to EV, was trained to the flight crew. The HUD included either an “SVS,” “EVS,” or “Fused” annunciation in the lower left hand corner of the display for positive indication of the Fusion mode to the PF.

Two HUD symbology sets were flown:

- 1) Standard HUD symbology (hereafter referred to as “Baseline”).
The glideslope reference line was always drawn. In addition, text specifying the raster format being displayed on the HUD (“EVS”, “SVS”, or “FUSED”) was always drawn underneath the airspeed tape.
- 2) Enhanced HUD symbology (hereafter referred to as the “Tunnel” symbology set).
The baseline symbology set was enhanced with pathway guidance and a runway outline. The “Tunnel” symbology set was tailored to transition at the same altitudes as the Fusion raster. The tunnel was shown above the 500 ft above field level (AFL) transition altitude, the last tunnel segment was positioned at 500 ft AFL (thus, it was no longer visible below 500 ft), and, upon reaching 500 ft AFL, the glideslope reference line was drawn and a runway outline was projected until reaching 50 ft AFL.

The glideslope reference line is part of both the Baseline and Tunnel symbology sets as it is identified as a required symbology element under FAR §91.175 EFVS operations.

In Figure 5, all four concepts are shown - the FLIR-Baseline HUD, FLIR-Tunnel HUD, Fusion-Baseline HUD, and the Fusion-Tunnel HUD. In the FLIR-Baseline concept, a minimum of symbology is used and the FLIR does not necessarily provide terrain and runway cues, depending for instance, upon the atmospheric conditions. Conversely, the Fusion-Tunnel concept uses tunnel guidance for distinct path demarcation and SV for clear terrain and runway references, above 500 ft AFL. Below 500 ft with the Tunnel symbol set, the runway outline provides an element of SV visual momentum within the EV raster background image. The tunnel is decluttered at 500 ft AFL to minimize clutter and is replaced by the glideslope reference line to ensure approach path angle awareness. Below 500 ft, the only additional symbology in the Tunnel symbology set, over that of the Baseline set, is the runway outline symbol. Below 500 ft, the FLIR-Baseline and Fusion-Baseline configurations are identical.

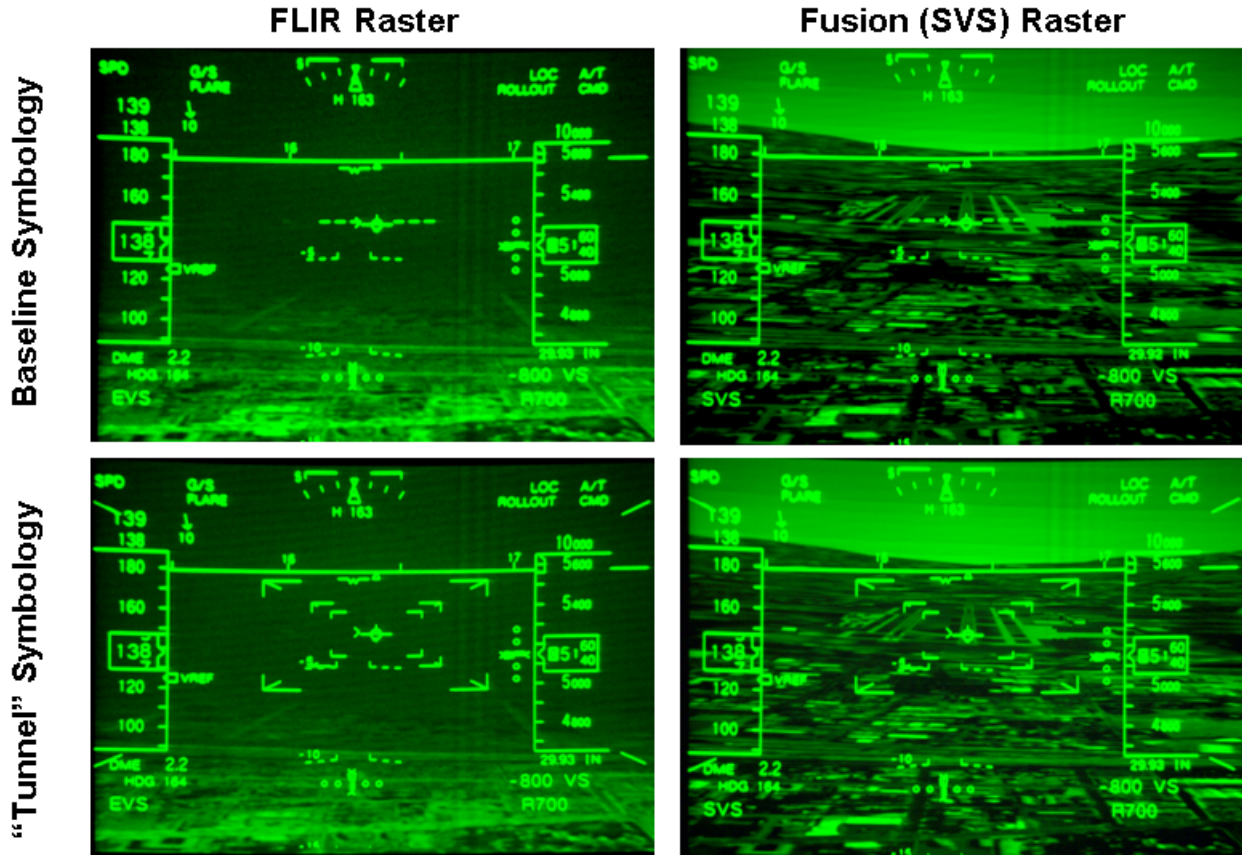


Figure 5. HUD Formats On Approach At 700 feet AFL

3.10.2 Auxiliary Display Concepts

Four PNF-AD display concepts were tested, differing from each other by: 1) two types of raster background presented; and, 2) two types of symbology presented.

Two AD raster (background) formats were flown:

1. EV only (hereafter referred to as “FLIR”);

The FLIR concept represented our “baseline” AD. In this configuration, the simulated FLIR output was exclusively displayed whether useful imagery was being provided or not.

2. Fused SV/EV (hereafter referred to as “Fused”).

In this configuration, the AD raster imagery was pilot-controllable and could be tuned at any time to one of 10 states: EV-only, SV-only, or 8 fusion combinations of EV and SVS, using an Equinox EP-3000™ fusion board. The fusion board employs a feature-level extraction algorithm with two pilot control inputs. The first control biased the feature level fusion through 8 weighting values weighting EV or SV. (A value of 1 biased the extraction to 11% FLIR and 89% SV whereas a value of 8 weighted the extraction to 89% FLIR and 11% SV). The second control modulated the false-color coding of the fusion image through 1 of 8 values. A setting of 1 did not apply any color-coding (the display was a monochromatic fused image). A setting of 8 applied maximal green shading to the features which were assessed by the fusion algorithm to be

“common” features between the two input videos and which had spatial frequency content above a threshold value.

As detailed later, the fusion controls were recorded and the evaluation pilots were trained on the use of the fusion, but not given any direct guidance as to how to use it. In this way, the experiment was designed to identify through pilot usage data, flight data, and pilot comments, a first-order cut at optimal fusion displays and controls for the PNF.

Two AD symbology sets were flown:

1) No Symbology

The absence of symbology has previously been shown to be very advantageous for clear raster imagery viewing. The absence of symbology would also likely be less costly to equip.

2) Symbology.

The symbology was derived during usability testing prior to the experiment based on a “decluttered” version of the PF HUD symbology (see Section 3.9).

In Figure 6, the four PNF-AD concepts are shown: FLIR-No Symbology AD (upper left), FLIR-Symbology AD (lower left), Fused-No Symbology AD (upper right), and Fused-Symbology AD (lower right). (The terminology “Fused” was used when the pilot controlled the blending of SV and EV imagery, such as the case for the PNF-AD. Whereas, the term “Fusion” was used when the blending was automatically controlled, such as the case for the PF-HUD.)

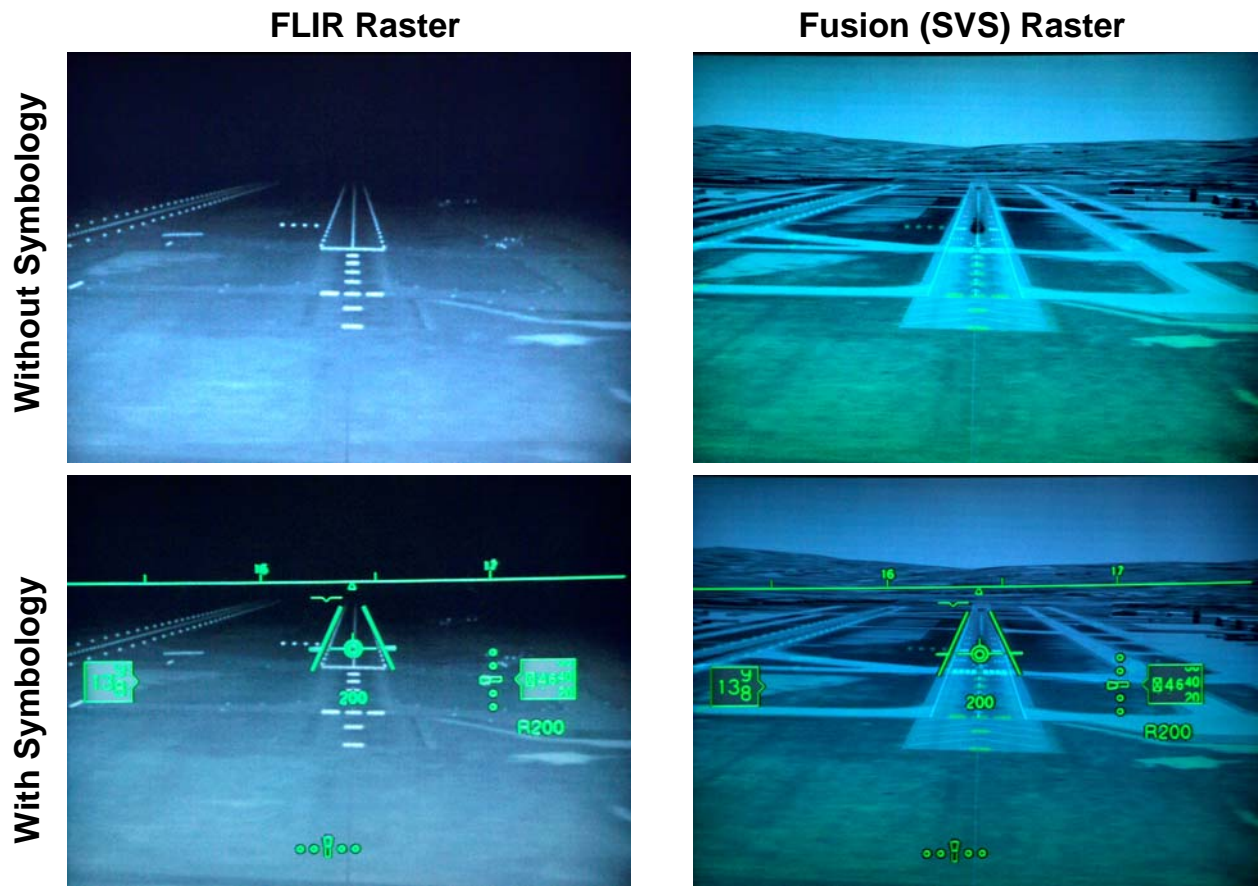


Figure 6. Four Auxiliary Display (AD) Formats at 200 ft AFL

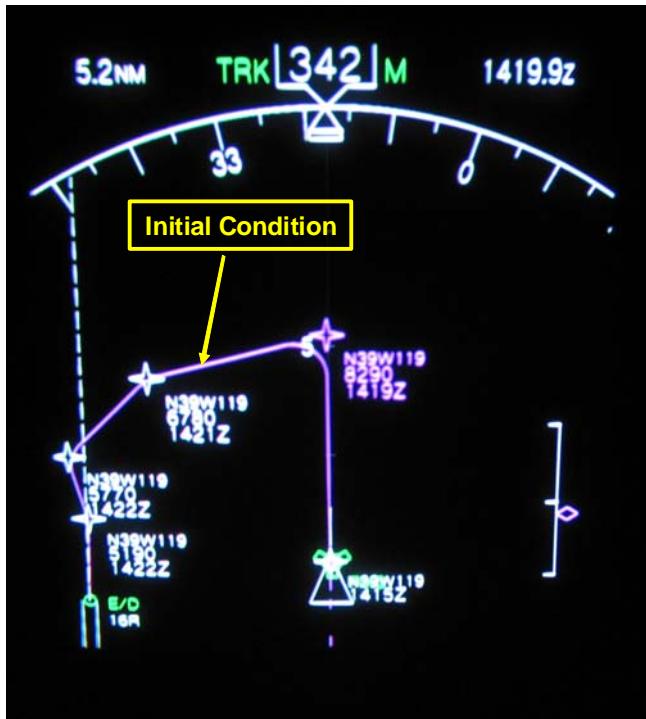
4. Procedures

4.1 Evaluation Task

The evaluation task was selected to approximate what may be typical of the emerging NextGen concept called an “equivalent visual operation.” The task was based on a published visual arrival – reflecting an efficient and preferred routing for ATC and noise-abatement – which currently requires VMC for the pilot to see-and-avoid terrain, traffic, and obstacles while navigating with respect to ground references. The approach path mimicked a RNP-type arrival, using a curved, descending path. The evaluation task tests the ability of SV and EV technology to support this type of operation by providing “equivalent visual” information into the cockpit. Further, these technologies may offer the potential for operational efficiency and minimums reduction above and beyond what can be provided by RNP.

The PF hand-flew the base and final leg portions of the Sparks Visual Arrival to RNO Runway 16R (see Fig. 7) with autothrottles engaged at an approach speed of 138 knots. The aircraft was configured for landing prior to each run (landing gear down and flaps 30 degrees). The path converged into the ILS for Runway 16R. The aircraft was established on a stabilized straight-in approach by 1000 ft AFL. The PNF monitored the approach from the right-hand side of the flight deck using standard instruments and the AD. Pilot participants were instructed that the run would end at main gear touchdown but that they should perform a go-around if either felt the landing was not safe.

Navigation Display Format



Approach Plate

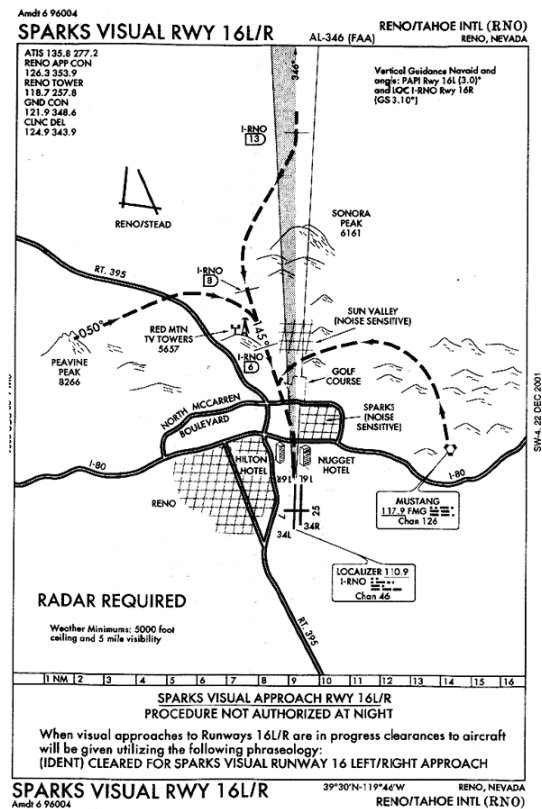


Figure 7. Sparks Visual Arrival To RNO Runway 16R.

4.2 EFVS Crew Procedures

EFVS crew procedures, adapted from those currently used in business aircraft EFVS operations, were established. Instructions in the use of the procedures (see Section 4.5) were given to each crew. An overview of these procedures is given in Table 2, including automatic call-outs. The altitude call-outs were set up assuming a 200 ft DH for the published, non-EFVS approach.

At 500 ft AFL, the “EVS Normal, System Normal” call by the PF corresponds to the point where the PF would nominally check that the FLIR was set-up properly and functioning properly on the approach. In our test, this call-out cued the PF to ensure that the HUD declutter, brightness and contrast were properly set (i.e., good raster image brightness and contrast). No EV controls, such as FLIR sensitivity or black/white hot options, were available to the flight crew in the experiment. The EPs were instructed that the FLIR was always set to its optimal setting to provide the best image possible for the flight crew under the prevailing weather, environment, and atmospheric conditions.

By the published minimums of 200 ft DH, the crew procedures dictate that the PF must have the required EFVS references or the required landing visual references (using natural vision) to continue the descent. The landing references were those published in FAR §91.175. For this test, the approach light system for RNO 16R provided the prominent EFVS references. If these EFVS references were visible, the PF was instructed to call “EVS Lights”.

If the PF saw the lights or markings of the threshold (the predominant landing visual reference for RNO 16R), the PF called “Landing.” The “landing” call was required no later than 100 ft AFL.

The PNF provided monitoring, including back-up on all decision heights, and was instructed to call “go-around” if “EVS Lights” was not called at or before 200 ft DH or if “Landing” was not called by 100 ft DH. The PNF was allowed to assist the PF in picking up the required visual cues (normal or EFVS). Transfer of control between the Captain and First-Officer was not permitted.

Table 2. EFVS Crew Procedures

Altitude-Based Events	AFL / Baro-Altitudes (ft)	Automatic Callouts	PF Tasks / Callouts	PNF Tasks / Callouts
500 feet AFL	500 / 4912	“500”	Response: “ <i>Systems Normal, EVS Normal</i> ”	Call “500 feet”
100 feet Above Minimums	300 / 4712	“ <i>Approaching Minimums</i> ”	Response: “ <i>Check</i> ”	Call “100 feet Above”
Published Minimums (200 ft AFL)	200 / 4612	“ <i>Minimums</i> ”	<u>With EVS Visual Cues,</u> Call “ <i>EVS Lights</i> ”	<u>When Visual Cues Appear,</u> Call “ <i>Lights</i> ” or “ <i>Field in Sight</i> ”
			<u>Without EVS Visual Cues,</u> Call “ <i>Going Around</i> ”	<u>Without PF Call of ‘EVS Lights’,</u> Call “ <i>Go Around</i> ”
EFVS Decision Altitude (100 ft AFL)	100 / 4512		<u>When Actual Visual Cues,</u> Call “ <i>Landing</i> ”	<u>When Visual Cues Appear,</u> Call “ <i>Lights</i> ” or “ <i>Field in Sight</i> ”
			<u>Without Actual Visual Cues,</u> Call “ <i>Going Around</i> ”	<u>Without PF Call of ‘Landing’,</u> Call “ <i>Go Around</i> ”

The crew procedures were new to all of the flight crews. Some procedures were counter to, others consistent with, their current airline Standard Operating Procedures (SOPs). In either case, the crew procedures were trained and “enforced” for the test. During the post-test debrief, questions and issues of how these procedures may or may not work within their airline operation and SOPs were discussed. These results are captured in Appendix B. Flight crews from the same airline were paired to the greatest extent possible to minimize SOP differences and to mitigate potential interference in crew interaction.

4.3 Experiment Matrix

Nominally, 40 experimental runs were completed by each evaluation crew with each pilot flying 20 approaches, as PF, evaluating the HUD concepts and with each pilot monitoring 20 approaches, as PNF, while evaluating the AD concepts.

The wind and weather varied on each run. The nominal visibility in the FLIR and OTW varied from 1 mile down to ½ mile. The required EFVS visual references became visible on the HUD between 450 ft and 250 ft AFL. Four runs per flight crew were specifically designed so the EFVS references were visible but the required runway (i.e., using normal un-aided vision) references were not. These four runs, if properly flown using the EVS crew procedures, should conclude by a go-around initiated no lower than 100 ft AFL.

The PF was instructed to fly each approach as precisely as possible, albeit as if there were passengers aboard, using the display information available, as the effect of the display information on the PF’s ability to fly the approaches would be quantitatively and qualitatively evaluated. In addition, the PF was instructed to land as close as possible to the centerline of the runway.

A significant component of the test, in addition to the nominal runs, was accomplished by measuring the ability of the flight crew to react and properly handle non-normal events. Four non-normal runs were flown by each crew. The non-normals were runway incursion (RI) scenarios and database integrity monitoring scenarios. The number and ordering of RI and database integrity scenarios were designed to avoid expectancy on the part of the flight crew.⁴²

The RI scenarios simulated an incursion with either a non-transponding baggage cart or fire truck. Both the baggage cart and fire truck were stationary. Details of the runway incursion set-up are contained in Section 5.8.

The database integrity monitoring scenarios purposefully introduced a lateral navigation solution error (of either 50 or 75 feet) with respect to the real runway. This error resulted in the synthetic vision terrain, pathway and guidance cue being misaligned from the FLIR and ILS data (which were defined in the flight crew briefing as being correct).

4.4 Measures

During each run, path error, pilot control inputs, PNF head-position and PNF-AD control inputs were recorded for analysis. In addition, an event marker was used by the EPs or the experimenter to log points of interest.

After each run, pilots completed a run questionnaire consisting of the Air Force Flight Technical Center (AFFTC) Revised Workload Estimation Scale,⁴³ Situation Awareness Rating Technique (SART),⁴⁴ and four Likert-type (7-point) questions specific to different constructs of display clutter (see Fig. 8).

After data collection was completed, pilots were administered two separate Situation Awareness – Subjective Workload Dominance (SA-SWORD)⁴⁴ and Subjective Workload Dominance (SWORD)⁴⁵ tests: one for HUD concept comparisons (FLIR-Baseline, FLIR-Tunnel, Fusion-Baseline, Fusion-Tunnel) and another for AD concept comparisons (FLIR-No Symbology, FLIR-Symbology, Fused-No Symbology, Fused-Symbology). The pilots also participated in a semi-structured interview to elicit comments on the HUD/AD concepts, HUD SVS-to-EVS transition strategy for the fusion concept, AD fusion strategy, and EVS crew procedures. These data are presented in Appendix B.

For the post-run questions, separate analyses of variance (ANOVA) analyses were conducted for the HUD concepts and the AD concepts. For the HUD concepts, there were two main factors, each with two levels (as shown in Figure 5.): 1) Raster (FLIR, Fusion) and 2) Symbology (Baseline, Tunnel). For the AD concepts, there were two main factors, each with two levels (as shown in Figure 6.): 1) Raster (FLIR, Fused) and 2) Symbology (On, Off). When the 2nd order interaction was significant, a simple main effects analysis was conducted using $\alpha=0.05$. For the post-test paired comparisons, simple ANOVA and Student-Newman-Keuls (SNK) post-hoc multiple comparison tests with alpha (α) set at 0.05 were performed.

<i>Display Clutter Ratings</i>	Low.....High						
	1	2	3	4	5	6	7
Rating #1 <u>Demand on Visual Attention</u> How much visual search time and cognitive effort was required to scan and locate task-critical display information in the display?							
Rating #2 <u>Supply of Visual Attention Resources</u> How much spare visual attention and mental ability was available to accomplish secondary task(s)?							
Rating #3 <u>Understanding</u> What was your ability to quickly and accurately understand task-critical display information?							
Rating #4 As I Performed the evaluation task, the level of display clutter on the HUD/Auxiliary Display was _____.							

Figure 8. Post-Run Display Clutter Questions.

4.5 Procedure

The subjects were given a 1-hour briefing to explain the SV/EV concepts on the HUD and AD, EVS crew procedures, and the expected evaluation tasks. After the briefing, a 2-hour training session in the IFD was conducted to familiarize the subjects with the aircraft handling qualities, display symbologies, EVS crew procedures, and controls. The ‘rare-event’ scenarios were not discussed, although the pilot’s

responsibility for maintaining safe operations at all times was stressed. Data collection lasted approximately 4.5 hours followed by a 30-minute semi structured interview. 10 minute breaks were taken approximately every hour during data collection. The pilots were also given a take-home final questionnaire. The entire session including lunch and breaks lasted approximately 9 hours.

5. Results

5.1 Path Control Performance

Root-mean-square (RMS) calculations of lateral and vertical path error were used as the measures for flight path control performance. For the HUD concepts, there were two main factors, each with two levels: 1) Raster (FLIR, Fusion) and 2) Symbology (Baseline, Tunnel). Separate ANOVA analyses were performed on RMS path error (lateral and vertical) for two segments of the run: approach and final. The approach segment began at the task starting point and ended at 500 feet AFL. The final approach segment was between 500 feet and 100 feet AFL. When the 2nd order interaction was significant, a simple main effects analysis was conducted using $\alpha=0.05$. When appropriate, SNK post-hoc tests with α set at 0.05 were performed. Two runs were excluded from these analyses due to simulation problems (e.g., lost path on navigation display during those runs). (The non-normal runs with a lateral navigation error - 50 feet or 75 feet - were not included in the final segment analyses.)

5.1.1. Approach Lateral Path Error

An ANOVA revealed that HUD symbology ($F(1,23)=12.60$, $p=0.002$) was highly significant for RMS lateral path error during the approach segment of the flight. Pilots had less lateral path error when using a tunnel (mean=28 ft) compared to the baseline (mean=41 ft) symbology configuration. The other main factor, raster type (FLIR or Fusion), and the second order interaction between symbology and raster were not significant ($p>0.05$) for this measure. In Figure 9, post-hoc tests (SNK using $\alpha=0.05$) showed two unique subsets for the approach lateral path error with the 4 HUD combinations: 1) FLIR-Baseline (mean=42 feet) and Fusion-Baseline (mean=40 feet) and 2) FLIR-Tunnel (mean=30 feet) and Fusion-Tunnel (mean=27 feet).

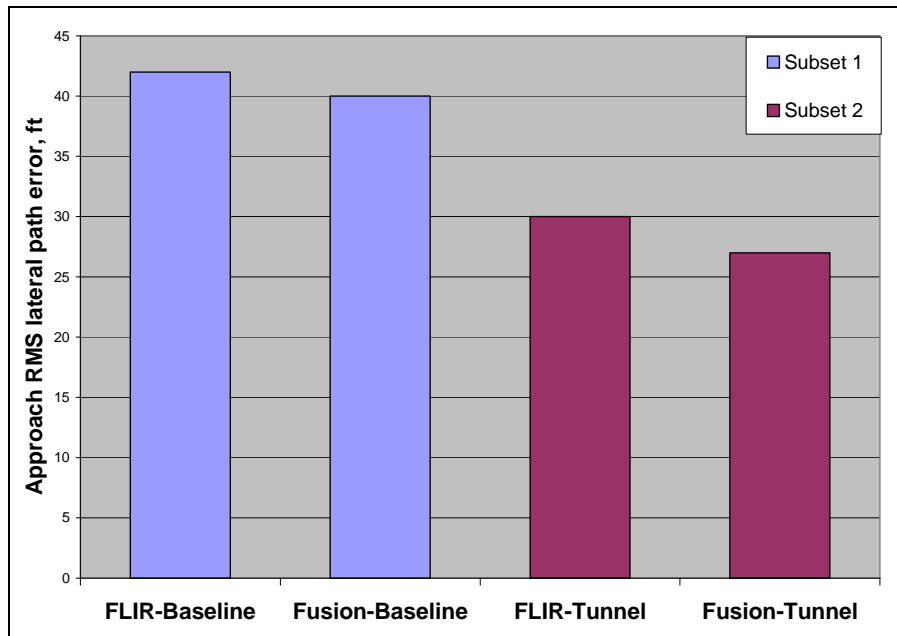


Figure 9. RMS Lateral Path Error During Approach Segment For Each HUD Concept.

5.1.2. Approach Vertical Path Error

An ANOVA revealed that both main factors, HUD symbology ($F(1,23)=5.67$, $p=0.026$) and HUD raster ($F(1,23)=4.70$, $p=0.041$) were significant for RMS vertical path error during the approach segment of the flight. Pilots had less vertical path error when using a tunnel (mean=9 ft) compared to without one (mean= 11 ft) and they had less vertical path error when using the Fusion imagery (mean=9 ft) compared to the FLIR imagery (mean=11 ft). The second order interaction between symbology and raster was not significant ($p>0.05$) for this measure. Post-hoc tests (see Fig. 10) revealed that the FLIR-Baseline (mean=12 feet) concept had significantly higher RMS vertical path error than the other three concepts: Fusion-Baseline=10 feet; FLIR-Tunnel=9 feet; & Fusion-Tunnel=9 feet.

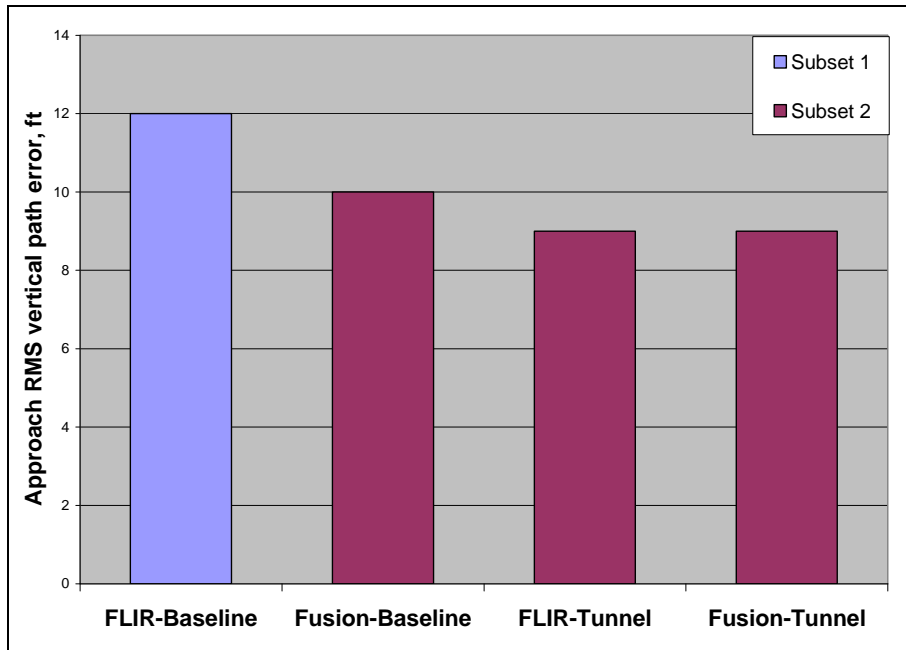


Figure 10. RMS Vertical Path Error During Approach Segment For Each HUD Concept.

5.1.3 Vertical and Lateral Path Error on Final

As expected, there were no significant differences ($p>0.05$) for HUD concept for the RMS lateral path error (mean=7 feet) or RMS vertical path error (mean=7 feet) during the final segment of the flight. Once on final, the only difference between the configurations was the presence or absence of the runway outline. This symbology element should be (and was found to be) inconsequential to flight performance.

5.1.4 Flight Technical Error on Approach

While the RMS measures for flight path control show statistically-significant differences on approach, the operational significance of these differences is questionable. To tease out the possible operational differences, the lateral and vertical tracking error data on approach was analyzed in terms of Flight Technical Error (FTE) as if it were a part of Required Navigation Performance (RNP).

Lateral and vertical path error data were evaluated using a histogram analysis for the nominal Sparks 16R approach runs for the four HUD concepts (FLIR-Tunnel, FLIR-Baseline, Fused-Tunnel, and Fused-Baseline). The horizontal and vertical path steering error components of the RNP calculation⁴⁶ include both FTE and display error. For this analysis, it was assumed that display error was negligible, so FTE was the only component of path steering error. It was also assumed that the other two components (path definition error and position estimation error) of the horizontal RNP calculation would be equivalent across the display concepts evaluated. Similarly, it was also assumed that the other three components (altimetry system error, vertical path definition error, and horizontal coupling error) of the vertical RNP performance calculation would be equivalent across the display concepts evaluated.

A summary of the flight path tracking performance is shown in Table 3. Using the FTE analysis, all HUD concepts yielded a horizontal FTE navigational accuracy of 0.02 nmi at least 95% of the time. No statistically significant differences were found between display concepts. These results are consistent with, and show even better performance than other NASA studies^{41,47} that showed HUD guidance concepts, using flight path-centered symbology, can enable manual RNP operations with lateral FTE of 0.05 nmi.

For vertical FTE, it was assumed that the pilot was flying a specified vertical profile so that the required vertical navigation performance accuracy was 300 feet.⁴⁶ With these assumptions, all HUD concepts yielded a vertical FTE navigational accuracy of 300 feet at least 99.7% of the time. In fact, the Fused - Tunnel, Fused-Baseline, and FLIR-Tunnel concepts yielded a vertical FTE navigational accuracy of 80 feet at least 99.7% of the time. The FLIR-Baseline HUD concept yielded a vertical FTE navigational accuracy of 120 feet at least 99.7% of the time.

Table 3. Flight Path Tracking Performance Summary for Approach

Raster	Symbology	Lateral rms mean	Vertical rms mean	Horizontal FTE	Vertical FTE
	Significance?	2 Unique Subsets	2 Unique Subsets	None	2 Unique Subsets
FLIR (No SV)	Baseline (Guidance Only)	42 ft*	12 ft*	0.02 nm	120 ft*
Fusion (SV)	Baseline (Guidance Only)	40 ft*	10 ft [†]	0.02 nm	80 ft [†]
FLIR (No SV)	Tunnel (Guid. + Tunnel)	30 ft [†]	9 ft [†]	0.02 nm	80 ft [†]
Fusion (SV)	Tunnel (Guid. + Tunnel)	27 ft [†]	9 ft [†]	0.02 nm	80 ft [†]

* Member of Subset 1 † Member of Subset 2

5.1.5. Path Control Results Discussion

As summarized in Table 3, the path control results show that the tunnel concepts have lower RMS lateral path error than the baseline, non-tunnel HUD concepts on the approach segment. The primary difference between the configurations was presence or absence of the tunnel and the turn anticipation cues that it provides. However, while statistically significant, the operational significance of the differences was not

found. This conclusion is drawn based on the absence of significant differences using an analysis of the FTE - a component of Required Navigation Performance. In this analysis, the data did not reveal any statistically-significant differences. All HUD concepts could be effectively used to 0.02 nm FTE levels.

Similarly, the approach vertical path error also showed statistical significance, but the very small differences imply little operational significance. In this case, statistically-significant vertical FTE differences were found, where the baseline (FLIR-Baseline) configuration showed the worst performance. However, the FTE difference (see Table 3) was only 40 ft (80 ft vs. 120 ft), implying little operational significance.

The pilot performance results are consistent with past research^{41,47} that showed HUD guidance concepts, using flight path-centered symbology, can enable manual RNP operations with lateral FTE of 0.05 nmi. Minimal performance differences were expected since each display concept utilized the same pursuit guidance control laws and symbology (i.e., the flight path marker, integrated cue guidance symbol, and path deviation indicators). However, it might be conjectured that the pilots with a Tunnel and/or SV information in the raster background were better able to attend to the dual task of vertical and lateral path tracking. For the baseline-FLIR configurations, pilots may have been concentrating on lateral tracking at the expense of vertical path performance.

Subjectively, the EPs felt that the tunnel provided good turn anticipation cues and the SVS background, when present, also improved flight path control performance because the database imagery in the background provided stronger roll reference visual cues. These features also limited a tendency to overcontrol in roll in the fixed-base B-757 simulator, particularly when compared to flying the baseline symbology set (i.e., compensatory guidance symbol only). These differences are subtly apparent in the data. In all cases tested, the pilots were able to meet operational performance requirements, such as FTE for RNP, flying manually.

5.2 Mental Workload

Mental workload was assessed after each experimental run, using the AFFTC workload estimate tool, and post-test, using SWORD.

The AFFTC workload estimate is a self-reported survey tool with significant operational crew acceptance. The scale assesses the immediate condition, using 7 point descriptors:

1. Nothing To Do; No System Demands
2. Light Activity; Minimum Demands
3. Moderate Activity; Easily Managed; Considerable Spare Time
4. Busy; Challenging But Manageable; Adequate Time Available
5. Very Busy; Demanding To Manage; Barely Enough Time
6. Extremely Busy; Very Difficult; Non-Essential Tasks Postponed
7. Overloaded; System Unmanageable; Important Tasks Undone.

The SWORD technique uses a paired-comparison of the experimental conditions to quantify their relative pilot workload. The SWORD is administered at the conclusion of the test; thus, reflecting the entire duration and experiences of the experiment and the experimental conditions.

5.2.1. AFFTC Workload Estimate – PF HUD

In the post-run data, the main factors of HUD raster ($F(1,23)=5.46$, $p=0.028$) and symbology ($F(1,23)=17.05$, $p<0.001$) were significant for workload. The second-order raster by symbology interaction was not statistically significant ($p>0.05$). Post-hoc tests (SNK using $\alpha=0.05$) showed two unique subsets:

Subset 1) Fusion-Tunnel (mean=3.1); and,

Subset 2) FLIR-Tunnel (mean=3.4), Fusion-Baseline (mean=3.5), and FLIR-Baseline (mean=3.6).

Pilots rated the Fusion-Tunnel HUD concept as having significantly less workload than the other 3 HUD concepts tested. While statistically-significant, the mean values show little difference since the mean values don't cross a rating boundary. On the AFFTC Workload Scale, a value of "3" indicates "Moderate Activity – Easily Managed; Considerable Spare Time" and a value of "4" indicates "Busy – Challenging but Manageable; Adequate Time Available". On average, all HUD configurations were rated as requiring "Moderate Activity – Easily Managed; Considerable Spare Time" but the second subset tended more toward being a "busy" activity.

5.2.2 AFFTC Workload Estimate – PNF-AD

For the PNF-AD concepts, there were no significant ($p>0.05$) differences between raster type (FLIR, Fused), symbology (Off, On), or their interaction for post-run workload. A mean pilot rating of 2.6 was given for the AD concepts by the pilots. On the AFFTC Workload Scale, a value of "2" indicates "Light Activity; Minimum Demands" and a value of "3" indicates "Moderate Activity – Easily Managed; Considerable Spare Time."

This result indicates that the presence or absence of symbology and the presence or absence of fusion controls for the PNF does not have a measurable effect on pilot monitoring workload.

5.2.3. SWORD

Pilots were administered the pair-comparison SWORD test that enabled ratings of mental workload across the four display concepts (raster * symbology) for both the PF and PNF displays. The definition of mental workload was given as "the amount of cognitive resources available to perform a task and the difficulty of that task."

The post-test SWORD data indicate that there were no significant ($p>0.05$) differences among the PF HUD concepts for the SWORD ratings of mental workload.

The post-test SWORD data show that AD concept was highly significant ($F(3, 69)=15.02$, $p<0.001$). Post-hoc tests (SNK using $\alpha=0.05$) showed three unique subsets for the mental workload ratings with the 4 PNF-AD concepts: Subset 1) FLIR-Symbology and Fused-Symbology (lowest workload); Subset 2) Fused-No Symbology; and, Subset 3) FLIR-No Symbology (highest workload).

5.2.4 Workload Discussion

The workload data do not show substantial differences associated with PF-HUD concepts. Pilots ranked the Fusion-Tunnel HUD concept as having significantly less workload than the other 3 HUD concepts tested, post-run. However, the post-test SWORD data indicate that there were no significant ($p>0.05$) differences among the HUD concepts for the SWORD ratings of mental workload. Operationally, the workload data suggest that the Fusion-Tunnel concept reduced PF workload to the extent that the average workload is “easily managed”, whereas the other concepts elicited workload ratings tending toward “challenging but manageable” workload levels. Pilot commentary suggested that the workload when flying the tunnel symbology concepts was easier (less scanning between the HUD and ND, easier to anticipate the turns), but the differences were not of a magnitude to warrant a workload penalty in the opinion of the pilots. Since the definitions of workload and the types of tests differ between the post-run (AFFTC Workload estimate) and the post-test (SWORD), it is difficult to draw a conclusion that the PF workload was affected by the HUD configuration differences. This result suggests that the addition of SV (i.e., Fused vs. FLIR condition) did not increase PF workload on the approach nor did the addition of the tunnel increase PF workload. If anything, the post-run results suggest that the addition of both SV and the tunnel symbology *reduces* PF workload.

Similarly, the post-run AFFTC workload ratings for the PNF-AD concepts showed no statistically significant differences, but post-test, pilots ranked the two AD concepts with symbology as requiring significantly less mental workload in their SWORD ratings. (There were no appreciable differences between the FLIR-Symbology and Fused-Symbology AD concepts for mental workload ratings.)

Pilot commentary typically noted the advantage of symbology in reducing the need to visually scan and cognitive task of integrating the different display information. The PNF-AD symbology set contained two key components of information:

- Flight path marker and guidance cue superimposed on the outside world (either SV or EV);
The FPM and guidance cue directly indicates to the PNF that the guidance is leading the PF to the correct location (guidance cue superimposed on a terrain background) and PF is following the guidance (FPM is following the guidance cue). Without this symbolic and terrain information, this information must be inferred from the PNF’s scan of the head-down instruments and the out-the-window view.
- Path error in the form of “dogbones;”
On the approach, no path error indication was provided on the head-down displays (neither PFD nor ND) although the presence of a lateral path error could be inferred from the ND path and ownship symbol. On final, path error was shown by the ILS raw data on the PFD. These data are not as precise nor as easily interpreted as the dogbones display.

The PNF workload data also suggest that the physical workload induced by adding fusion controls was either minimal or it was offset by the reduction in mental/cognitive workload provided by adding EV and SV information.

The location of the PNF-AD was often noted as being “non-optimal.” It was too far away from the forward field of view / instrument panel and comments indicated that it added workload to the PNF because it required head and neck movements to see the display. The EPs would have preferred that the

PNF-AD was within a comfortable visual field-of-regard of the primary instrument displays and the forward view outside the aircraft .

5.3 Situation Awareness

Situation awareness was assessed after each experimental run using the post-run SART, and after each test, using the post-test SA-SWORD measures. Situation awareness was assessed in both PF and PNF roles. SART is a multi-dimensional rating technique using the constructs of: 1) demand on attentional resources; 2) supply of attentional resources; and, 3) understanding. From these components, the SART rating is “understanding” reduced by the difference of “demand” minus “supply” (i.e., $SART = \{(\text{understanding}) - (\text{demand} - \text{supply})\}$).

Similar to the SWORD described above, the SA-SWORD is a paired comparison technique that provides relative situation awareness ratings across the four display concepts for both the PF and PNF displays. For these comparisons, situation awareness (SA) was defined as “the pilot’s awareness and understanding of all factors that will contribute to the safe flying of their aircraft under normal and non-normal conditions.” The SA-SWORD measure differs from the post-run SART measure construct as it is a pair-comparison test, it was administered post-test, and the underlying definition and construct for the SA ratings are different.

5.3.1 SART – PF HUD

An ANOVA revealed that both HUD raster type ($F(1,23)=5.41$, $p=0.029$) and symbology type ($F(1,23)=18.85$, $p<0.001$) and their interaction ($F(1,320)=5.31$, $p=0.022$) were significant for PF-HUD SART ratings. Pilots rated their SA significantly higher when the HUD symbology included pathway/tunnel guidance and when the HUD used the Fusion imagery. A simple main effects analysis revealed that the effects of symbology type on SA ratings was stronger within the Fusion raster type than within the FLIR raster type (Fig. 11).

5.3.2. SART – PNF-AD

An ANOVA on SART ratings across the PNF display conditions found no significant results for the main effects (raster, symbology) or the interaction ($p>0.05$).

5.3.3. SA-SWORD

An ANOVA revealed that the PF HUD concepts were highly significant ($F(3, 69)=43.61$, $p<0.001$) for the SA-SWORD ratings. Post-hoc tests (SNK using $\alpha=0.05$) showed three unique subsets for situation awareness ratings with the 4 HUD concepts: 1) Fusion-Tunnel (highest SA); 2) Fusion-Baseline and FLIR-Tunnel; and 3) FLIR-Baseline (lowest SA).

The PNF-AD concepts were also highly significant ($F(3, 69)=37.78$, $p<0.001$) for the SA-SWORD ratings. Post-hoc tests (SNK using $\alpha=0.05$) showed three overlapping subsets for situation awareness ratings with the 4 AD concepts: 1) Fused-Symbology; 2) FLIR-Symbology and Fused-No Symbology; and, 3) Fused-No Symbology and FLIR-No Symbology. Pilots ranked the Fused-Symbology AD concept as having significantly higher SA than the other 3 AD concepts tested.

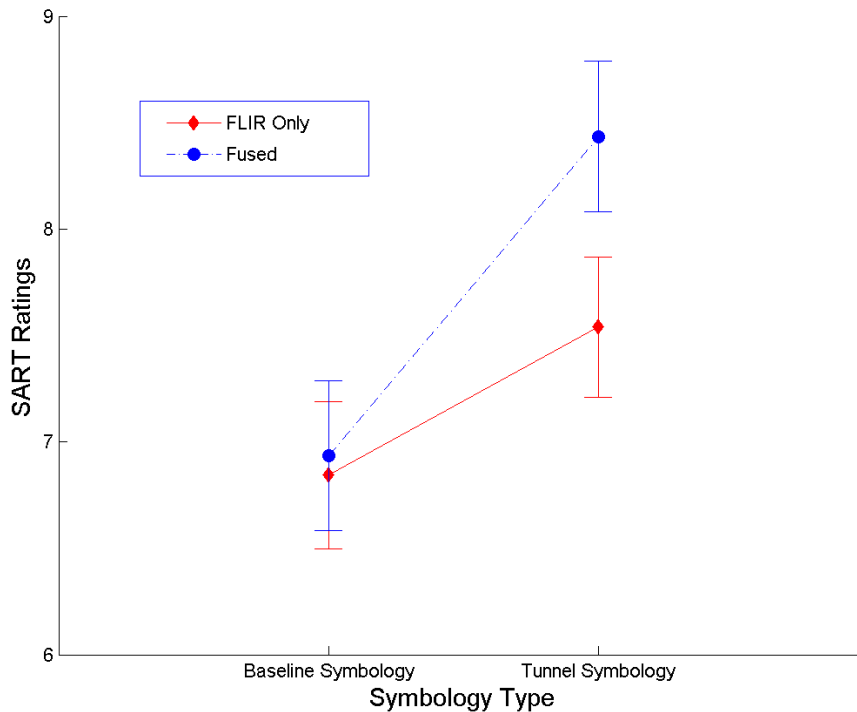


Figure 11. Means of SART Ratings With 95% Confidence Intervals for HUD Configurations

5.3.4. Situation Awareness Discussion

Pilot commentary consistently noted that the presence of the tunnel gave the PF a better understanding and appreciation of the curving, descending visual arrival path. This commentary was reflected in the SART and SA-SWORD data. Without the tunnel guidance, pilots commented that they had to use the head-down ND more frequently for path (turn) guidance. (Head-tracking data was unfortunately not recorded for the PF.)

Also, the SV component in the Fusion HUD concept provided significant terrain information unavailable in any other cockpit displays thereby providing an intuitive, conformal display of terrain and pathway information. These SA features emerged in the SA data, both post-test and post-run. Superior SA was rated for the Fusion-Tunnel HUD concept whereas the FLIR-Baseline was rated inferior for SA. Interestingly, by the SA measures, the Fusion-Baseline and FLIR-Tunnel concepts were not significantly different. This result would imply that the pilots felt the SV contribution to SA was essentially equivalent to the tunnel contribution to SA.

The benefits of Fused imagery and symbology on the PNF-AD only emerged within the post-test data. The post-test and post-run SA differences might be attributed to the fact that SA-SWORD asks for a *general* appraisal whereas the SART asks for ratings from what was experienced for *that* pilot on *that* run. SA can be high - it was high in all conditions, including the baseline, as they were all highly skilled pilots - and the task really wasn't extremely demanding of the PNF. But, when asked to compare the PNF display concepts post-test to each other, SA differences emerged.

Similar to the workload data, the pilots commented that they felt SA was impacted by several issues. SA was significantly improved with Fused imagery on the PNF-AD by providing a way to better monitor the EV and navigation system performance and improve their understanding of their flight path with respect to terrain. Symbology on the PNF-AD provided two key SA benefits. First, the FPM and guidance cue (with FLIR and/or SVS imagery) provided visual evidence that the PF was flying to the proper point on the ground (i.e., flying to the intended runway, touchdown point) and the raw data displays on the PNF-AD symbology was a direct indication of path error for the PNF (i.e., the “dog-bones” were not shown on the PNF PFD.) Without the dogbones, the PNF had to use the ND to monitor approach-tracking performance. These PNF-AD attributes may not have been critical to the experiment on each run (i.e., minimal impact on post-run SART) but they can contribute significantly to SA - in general - for a PNF in this type of operation.

5.4 Pilot Display Preferences

Separate post-test paired comparisons for pilot display preferences were made on the HUD and AD concepts after data collection was completed.

The preferred HUD concept rankings showed highly significant differences ($F(3, 69)=73.17, p<0.001$). Post-hoc tests (SNK using $\alpha=0.05$) showed three unique subsets for the pilot-preferred display with the 4 HUD concepts: 1) Fusion-Tunnel; 2) Fusion-Baseline and FLIR-Tunnel; and 3) FLIR-Baseline. Pilots ranked the Fusion-Tunnel HUD concept as being preferred significantly more than the other 3 HUD concepts tested. Interestingly, the absence of statistically-significant differences between the Fusion-Baseline and FLIR-Tunnel HUD concepts indicates that the presence of SV on the HUD is equivalent in preference to the addition of tunnel/pathway information. Pilots prefer both to be added, but they didn't prefer one more than the other.

The preferred PNF AD concept rankings also showed highly significant differences ($F(3, 69)=23.74, p<0.001$). Post-hoc tests (SNK using $\alpha=0.05$) showed three overlapping subsets for the pilot-preferred display with the 4 AD concepts: 1) Fused-Symbology; 2) FLIR-Symbology and Fused-No Symbology; and 3) Fused-No Symbology and FLIR-No Symbology. Pilots ranked the Fused-Symbology AD concept as being preferred significantly more than the other 3 AD concepts tested. In this case, the overlapping subsets in statistically-significant differences mildly suggests that symbology on PNF-AD was more preferred over Fusion, if a choice must be made and only one “feature” could be provided. Post-test commentary, while strongly favoring the addition of symbology to the PNF-AD, also indicated that the PNFs wanted a symbology declutter control so they could selectively turn on or off the symbology. (This feature was not available to them.)

5.5 Head-Tracking Analysis

A head-tracker was used to quantify what display the PNF was using and when. An eye tracker would have been preferable for this task, but its installation was impractical. The head-tracker was unable to differentiate whether the PNF was looking OTW or at the instrument panel HDDs because very small head-movements, if any, were used by the PNFs to look between the HDDs and OTW. Fortunately, the head-tracker could very effectively distinguish when the PNF was looking at the AD or at the OTW/HDDs. The head-tracker statistical data analysis is detailed in Appendix A.

In Table 4, the percentage of time that the PNF looked at either the AD, the OTW/HDD (also referred to as the “Forward” location), or elsewhere is shown as a function of the aircraft altitude (in feet), Above Field Elevation (AFL). Simple ANOVA calculations were conducted on the percentage of time spent looking at the different head locations (Auxiliary Display, Forward, and Other) for seven phases of flight (or altitude segments). For each altitude segment, the factor - head location - was highly significant ($p < 0.001$) within each altitude bins and post-hoc tests (SNK, $\alpha = 0.05$) showed 3 unique subsets for percentage of dwell across the head locations (Forward, AD, or Other).

The mean percentages, also shown in Figure 12, indicates that the forward location (i.e., HDD and OTW) always had the greatest percentage for dwell time, followed by the Auxiliary Display and then “Other” (i.e., neither looking Forward nor at the Auxiliary Display). The data also shows that the PNF was using the AD only about 14% of the time (collapsed across PNF-AD concept) above 1000 ft but increased their usage (up to 30%) as the aircraft descended until it reached 400 ft AFL. Below 400 ft, AD usage quickly reduces to about 10% of the time below 300 ft AFL. The PNF was mostly looking forward (HDD/OTW) below 200 ft (~90%) and pilot commentary suggested that the PNFs were looking OTW for the runway visual references.

Table 4. Percentage of Time PNF Directing Gaze at Various Head Locations

Head Location	Alt > 1000	500 < AFL < 1000	400 < AFL < 500	300 < AFL < 400	200 < AFL < 300	100 < AFL < 200	AFL < 100
Forward (HDD/OTW)	80.0%	64.7%	59.4%	72.1%	84.0%	88.7%	85.8%
Auxiliary Display	14.1%	26.4%	29.9%	20.2%	11.2%	8.0%	9.5%
Other than AD or Fwd	5.9%	8.9%	10.7%	7.7%	4.8%	3.3%	4.7%

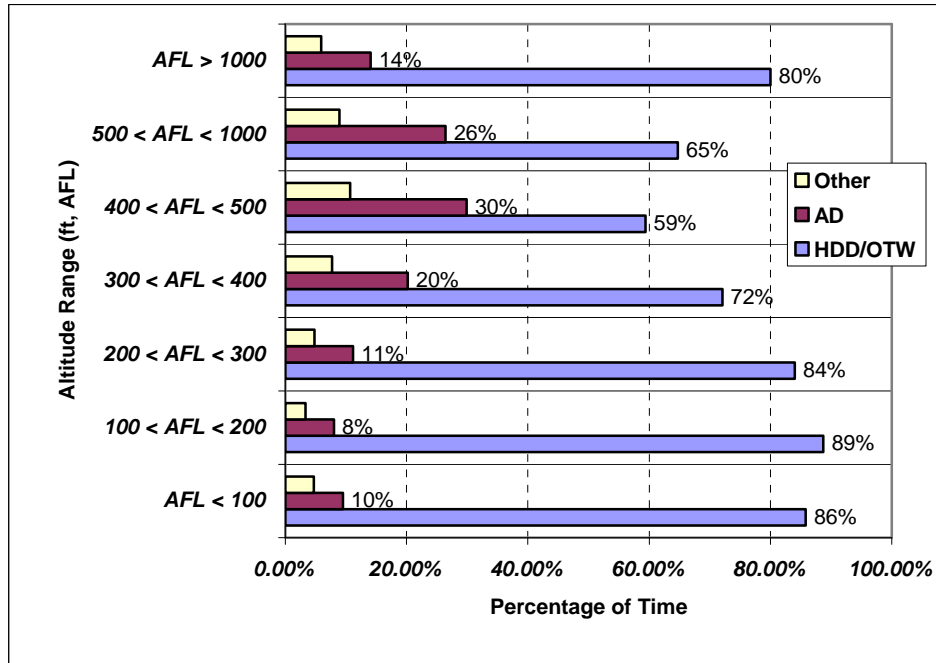


Figure 12. Percentage Of Time Where PNF Directed Gaze.

In Table 5, the percentage of time the PNF spent looking at the AD is shown to be dependent upon the AD display concept. Simple ANOVAs were conducted on the percentage of time spent looking at the different AD concepts (FLIR, with and without symbology and Fused, with and without symbology). Every altitude segment, except the AFL < 100 feet segment, showed significant differences for the AD concept.

The data shows that on approach, above 500 ft AFL, pilots focused on display concepts with symbology more than those with terrain information (SV or EV). (Given the simulated weather conditions, the FLIR did not provide any terrain information above 1000 ft AFL.) For instance, the FLIR-Symbology configuration induced more gaze, as estimated from the PNF head tracker data, on the part of the PNFs than the Fused-No Symbology configuration (terrain information but no symbology). This result is likely due to the PNF-AD symbology which included the path error – this information was not readily derived from the other PNF displays.

Between 500 and 200 ft AFL, the usage of all the display concepts increased. The presence of symbology primarily and SV secondarily induced attention to the AD. While statistical significance is shown in Table 5 below 200 ft, the differences in display concept are operationally insignificant. (10% percentage represents approximately 1 second of time looking at the PNF between 100 and 200 ft AFL.)

Table 5. Percentage of Time PNF Directing Gaze, By Display Configuration

AD Concept	Altitude Range (ft, AFL)						AFL < 100
	AFL > 1000	500 < AFL < 1000	400 < AFL < 500	300 < AFL < 400	200 < AFL < 300	100 < AFL < 200	
Significance?	4 Unique Subsets	4 Unique Subsets	2 Unique Subsets	2 Unique Subsets	2 Overlapping Subsets	2 Overlapping Subsets	None
FLIR - No Symbology	5%	18%	24%*	15%*	9%*	6%*	9%
Fused - No Symbology	13%	22%	24%*	15%*	10%*	7%*†	10%
FLIR-Symbology	16%	30%	36%†	25%†	12%*†	10%†	11%
Fused - Symbology	23%	35%	36%†	27%†	14%†	9%†	9%

* Member of Subset 1 † Member of Subset 2

Based on the quantitative gaze data in Table 5 and from the pilot debriefs, below about 300 ft AFL, all the PNFs were predominately “eyes-out” looking for visual landing references and preparing for landing. On occasion, a PNF would state that they tried to take a glance at the PNF-AD. They noted that if the PNF-AD were installed on the forward instrument panel, in a convenient location, they might have had a tendency to use it more. But likely, not significantly much more.

5.6 Subjective Assessments of Display Clutter

After an experimental run, each pilot gave ratings for the 4 Likert-type questions on display clutter (Figure 8) for the display (HUD or AD) concept they had just flown.

The first three clutter questions (see Figure 8) mirror the SART rating paradigm and are not a validated methodology. The questions were an attempt to quantify clutter effects as it affects a pilot’s SA. In essence, this creates a “Clutter-SA” rating – that is, it attempts to quantify the visual demand, supply, and understanding by the pilots during the experimental trials. The last question was the pilot’s subjective assessment of display clutter. This data serves as baseline for the “Clutter-SA” data. (These data are being used primarily in another related study to assist in the development of validated quantitative and qualitative display clutter metrics.)

5.6.1 HUD Concept Ratings

For the HUD concepts, the data were as follows:

- **PF-HUD Clutter (Rating #1 Demand on Visual Attention)**

This rating corresponds to the pilot’s subjective assessment of their visual effort to scan and locate task-critical display information. Symbology type ($F(1,23)=18.87, p<0.001$) was significant for this rating with the Tunnel concept resulting in a lower rating than the Baseline symbology. The raster type main effect and the raster-by-symbology interaction were not significant ($p>0.5$) for this measure. Post-hoc tests (SNK using $\alpha=0.05$) showed three unique

subsets (Fig. 13) for demand on visual attention ratings with the 4 HUD combinations: 1) Fusion-Tunnel (mean=3.3); 2) FLIR-Tunnel (mean=3.5); and, 3) Fusion-Baseline(mean=3.8); and FLIR-Baseline (mean=3.9).

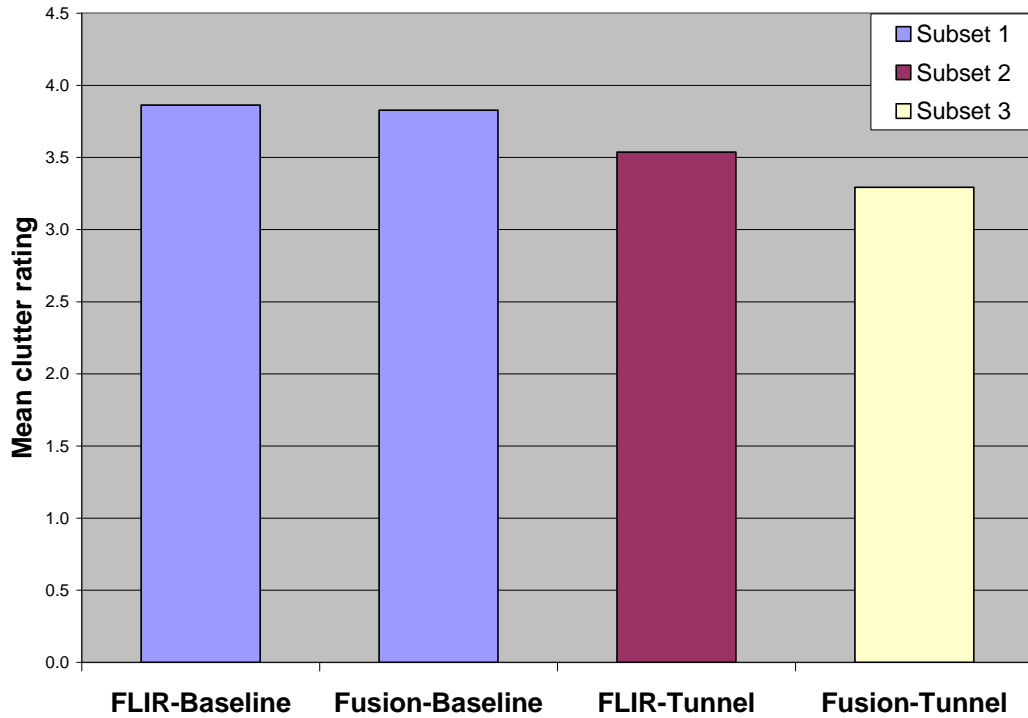


Figure 13. Mean Ratings Of Demand On Visual Attention Clutter Metric For Each HUD Concept.

- **PF-HUD Clutter (Rating #2 Supply of Visual Attention Resources)**

This rating corresponds to the pilot’s subjective assessment of their supply of visual attention available to attend to secondary tasks. The data shows that raster type ($F(1,23)=4.90, p=0.037$) and symbology type ($F(1,23)=12.26, p=0.002$) were statistically significant for the HUD. Pilots rated that they had a highest supply of visual attention resources when flying with the Fusion raster and the Tunnel symbology HUD concept compared to the other three concepts. Post-hoc tests (SNK using $\alpha=0.05$) showed three overlapping subsets (Fig. 14) for supply of visual attention resources ratings with the 4 HUD combinations: 1) Fusion-Tunnel (mean=5.1); 2) FLIR-Tunnel (mean=4.8) and Fusion-Baseline (mean=4.6); and, 3) Fusion-Baseline (mean=4.6) and FLIR-Baseline (mean=4.5).

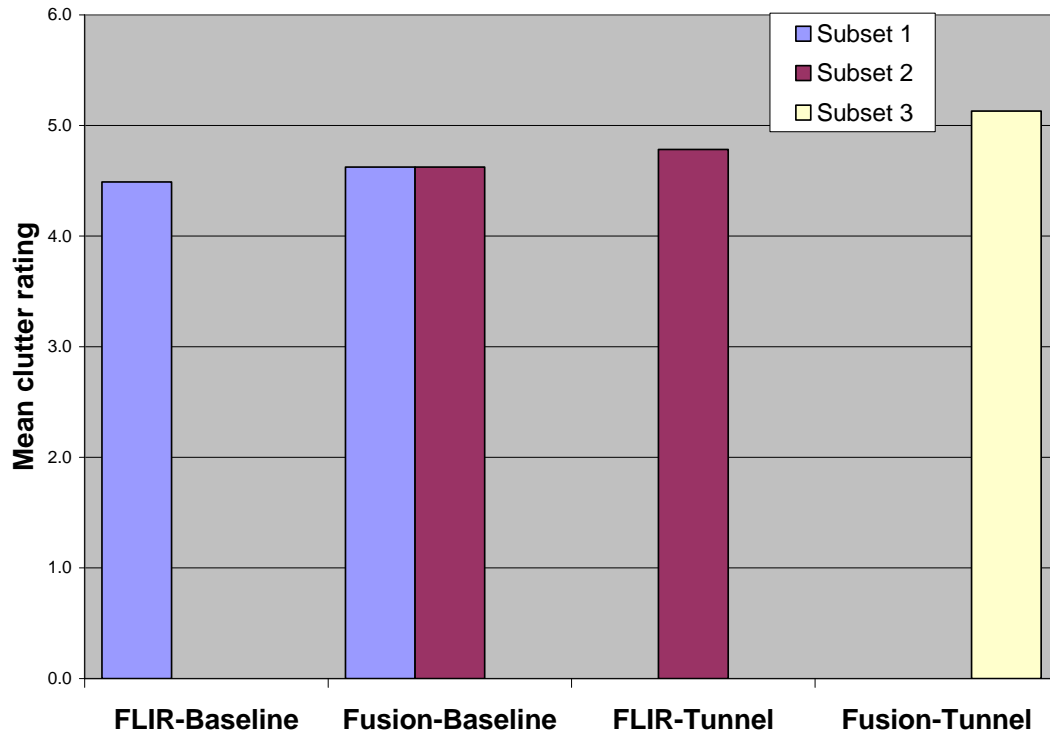


Figure 14. Mean Ratings Of Supply Of Visual Attention Resources Clutter Metric For Each HUD Concept.

- **PF-HUD Clutter (*Rating #3 Understanding*)**

This rating corresponds to the pilot’s subjective assessment of their understanding of task-critical display information. Symbology type and raster type main effects and the raster by symbology interaction were not significant ($p>0.05$) for this measure. The average rating for the HUD concepts was 6.1 which correspond to a moderately high understanding of the task-critical information on the HUD.

- **PF-HUD Overall Clutter (*Rating #4*)**

This rating corresponds to the pilot’s subjective assessment of the degree of display clutter. The rating data showed no statistically significant differences for raster type, symbology type, or their interaction. The average rating for the HUD concepts was 3.3 which correspond to a moderate amount of display clutter.

Intuitively, one might infer that adding more information to a flight critical Primary Flight Display, like adding a pathway guidance or tunnel to the HUD, would negatively increase display clutter. On the contrary, these ratings suggest that the addition of the Tunnel to the HUD did not impact clutter. No difference in the overall clutter rating or in the pilot’s understanding of task critical information was shown. Further, the “Clutter-SA” construct ratings showed statistically-significant *improvements* in “Visual SA” by the addition of the tunnel. The ratings showed statistically-significant reductions in the demand on visual attention resources and increases in the supply of visual attention resources to attend to

secondary tasks with the addition of the tunnel. Fusion (i.e., the presence of SV information) also increased the pilot's supply of visual attention resources so that he/she could perform secondary tasks.

5.6.2 PNF-AD concept Ratings

For the PNF-AD concepts, the data were as follows:

- **PNF-AD Clutter (*Rating #1 Demand on Visual Attention*)**

There were no significant differences ($p>0.05$) for raster type, symbology, or their interaction on the pilot's subjective assessment of the demand on visual attention. The average rating for the AD concepts was 2.5 which correspond to a moderately low demand of visual attention to scan and locate task-critical display information on the PNF-AD.

- **PNF-AD Clutter (*Rating #2 Supply of Visual Attention Resources*)**

There were no significant ($p>0.05$) differences for raster type, symbology, or their interaction in the pilot's subjective assessment of their supply of visual attention resources. The average rating for the AD concepts was 5.7 which corresponded to a moderately high supply of visual attention resources to effectively attend to other secondary visual tasks.

- **PNF-AD Clutter (*Rating #3 Understanding*)**

There were no significant ($p>0.05$) differences for raster type, symbology, or their interaction in the pilot's subjective assessment of their understanding of task-critical PNF-AD information. The average rating for the AD concepts was 6.4 which corresponded to a very high understanding of task-critical display information.

- **PNF-AD Overall Clutter (*Rating #4*)**

Symbology type ($F(1,23)=6.30$, $p=0.020$) was statistically significant for the pilot's subjective assessment of the degree of PNF-AD clutter. The "No-symbology" PNF-AD concepts (mean=1.8) were rated as having less overall clutter than the PNF-AD concepts with symbology (mean=2.2). Both of these mean ratings for the PNF-AD concepts correspond to a moderately low level of display clutter; thus, the operational significance of these differences is minimal.

The pilots noted that the symbology on the PNF-AD was beneficial to SA, but contributed to clutter (based on their "overall clutter" ratings). The three subjective elements of the "Clutter-SA" metric, however, did not show statistically significant differences. The absence or presence of PNF-AD symbology did not affect the post-run PNF-AD ratings for visual understanding, demand or supply.

The pilots, in their post-test debriefing, clearly advocated that they wanted symbology *and* a completely clear FLIR or Fusion raster on the AD to promote better readability and understanding of the imagery. Their proposed solution was to include symbology on the PNF-AD and also, include a PNF-AD "declutter" button, analogous to the PF-HUD, so the symbology could be toggled on and off as needed.

5.7 Fusion Controls

By providing a plethora of controls to the PNF for the Fused AD concept, the experiment data provided first-order determination if: a) a "fusion" concept was viable in the commercial cockpit; b) allowing the

PNF to control their presentation was viable or desirable; and, c) SVS and EVS was necessary for the PNF. No “guidance” was provided to the pilots for optimal usage. Instead, the system operation was explained and they were given the training runs to “play” with the controls. After training, the pilots were allowed to use whatever control settings that they felt necessary and appropriate. The EPs quickly learned how the fusion worked and what the most effective means to employ the controls were.

In Figure 15, the percentage of time that a fusion control setting was used by the PNFs during the experiment is plotted by altitude range. The altitude ranges correspond to the approach segment (“APP”: run start to 500 ft AFL), and 100 ft increments from 500 ft AFL to touchdown (or go-around). A fusion setting of 1 corresponds to SVS-only on the PNF-AD, a fusion setting of 10 corresponds to EVS-only, and a setting of 2 through 9 corresponds to the feature-level extraction algorithm bias as described in Section 3.10.2.

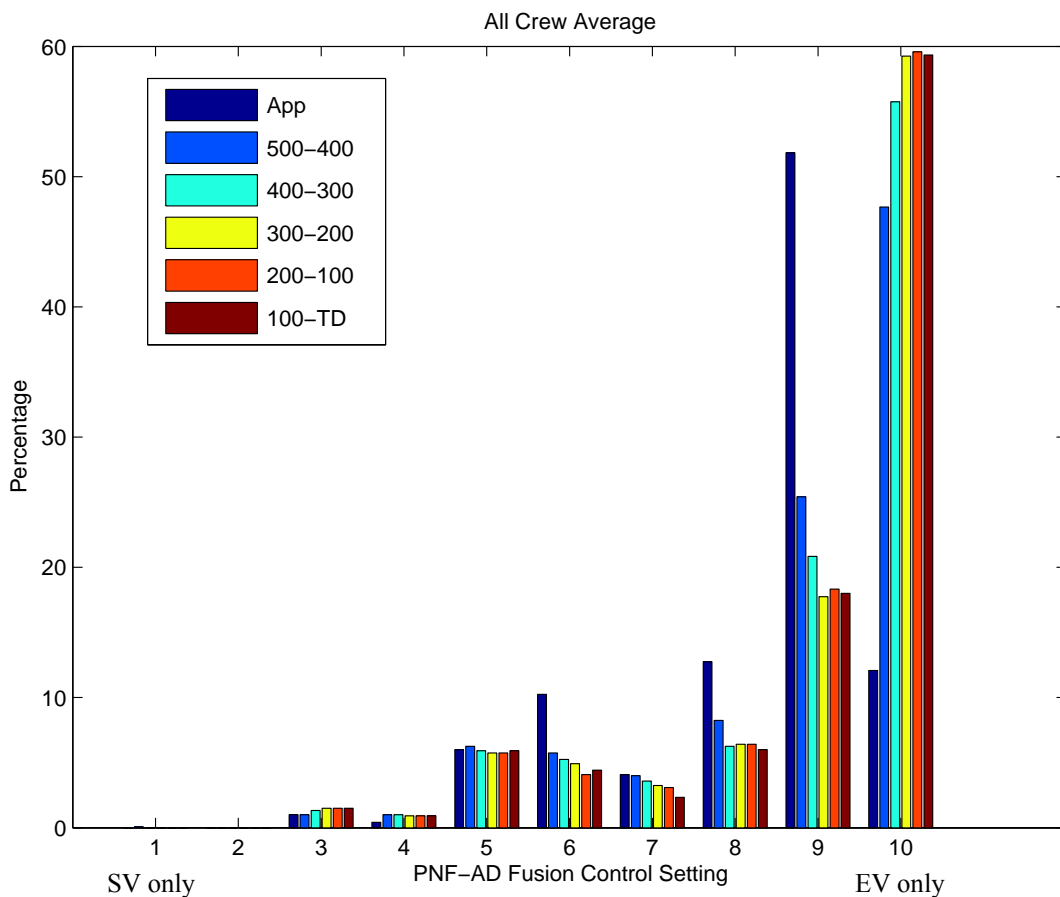


Figure 15. Fusion Control Settings By Altitude.

The data indicate a consistent trend. At altitude (on approach down to approximately 500 ft AFL), the PNFs tended to use the feature-level fusion of EV and SV. The most prevalent settings were heavily weighted toward EV but still contain SV content (i.e., settings of 8 and 9 in Figure 15). On the approach segment, the EVS did not have any information content because of simulated clouds on the approach. With feature-level extraction, the fusion image shows the SV database image without significant

alteration or contrast reduction. The PNFs often used an intensity (false color-coding) value of 8 so when color appeared on the PNF-AD, this cued the pilot that the EV was starting to show useful information. The color signalled that they could effectively begin using an “EVS-Only” setting. The data indicate that fusion control was used - albeit not to its full-range - and the PNF gathered significant information that assisted in their monitoring function. The EPs quickly learned how the fusion worked and what the most effective means to employ the controls were.

In Figure 16, these data are collapsed to highlight the percentage of time that the PNF used any fusion settings (i.e., “SVS-Only” or feature-level fusion values, “SV-Fusion,” fusion control settings 1 through 9) or “EV-Only” (i.e., fusion control setting 10) when the PNF-Fused AD concept was flown. The data shows that, on the approach, fusion or SV information was displayed more than 85% of the time. Below 500 ft, in the final approach segment to landing, EV-only was used 60% of the time and Fusion (i.e., SV) reduced to 40%. The 60-40 distribution in EV-only and “SV-Fusion” settings, respectively, suggests that the PNF used both information sources cooperatively and effectively. The common PNF strategy was to cycle between EV-only and the highest level of SV/EV fusion. They could do this quickly because these were adjacent switch positions.

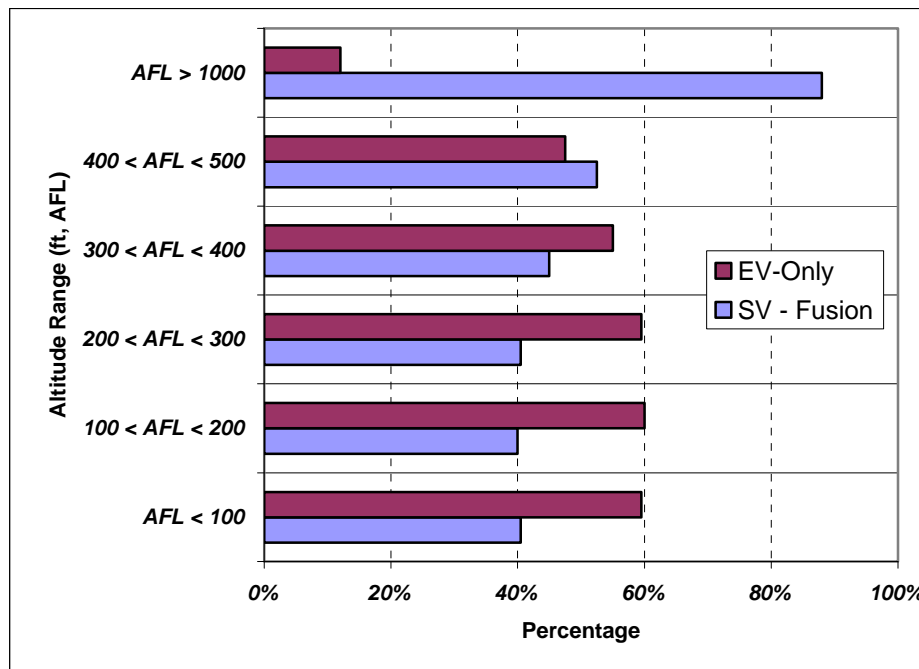


Figure 16. SV/EV Fusion And EV-Only Setting By Altitude AFL.

5.8 Non-Normals

Non-normals were injected into the test unbeknownst to the evaluation subjects. The non-normals were two runway incursions and four lateral offsets for each flight crew.

5.8.1 Runway Incursions

The runway incursions were created by a baggage cart and a fire truck. Both vehicles were positioned in the same location, approximately 850 ft from the RNO Runway 16R landing threshold and just slightly offset from the centerline. They were both positioned perpendicular to the runway (i.e., they were facing toward the runway edgelines.) The weather on the runway incursions was held constant at 2400 ft RVR (OTW) with the lowest cloud layer at 500 ft AFL. The FLIR visibility was very good in this condition – approximately 4 times the OTW RVR.

The baggage cart runway incursion was always performed before the fire truck incursion. The baggage cart was much more difficult to see due to its small size. This ordering tested for “just noticeable differences” for runway incursion detection.

The view of the baggage cart and firetruck incursions using several of the displays and display concepts available to the PFs and PNFs are shown in Figure 17. Each of the pictures was taken at 60 ft AFL, approximately on glideslope and localizer to landing. A yellow oval is added to Fig. 17(a) and Fig. 17(d) to highlight the location of the incurring vehicles. The incurring vehicles are also contained in the other images in the approximately same location within the other display concepts shown in this figure.

While innumerable issues factor into the data analysis for runway incursion, four of the key attributes that must be considered in analyzing these data are: display resolution, target size, display color (presence and absence), and the presence or absence of symbology. These factors will be discussed in relation to the experimental data.

To put the RI results into perspective of display resolution, the EVS and OTW (simulated) visual scene resolutions were used to compute the theoretical altitude (AFL) that the RIs might be reasonably expected to be observed by the crew. While target detection is critically dependent upon visibility, lighting, target contrast, color, gray scales, etc., resolution only is used in this example. It is assumed that 10 pixels (scan lines) are required for a human observer to recognize a target/object in this example.²⁹ The baggage cart consisted of a tug and a cart. The tug was approximately 7.5 ft tall and 10 ft long, tied to a cart 6.5 ft tall and 10 ft long. The Fire Truck was 31.6 ft long and 13 ft tall. An operating rotating beacon was depicted atop the Fire Truck.

In Table 6, the height of the aircraft (AFL) when 10 lines (pixels) draw the incurring vehicles is shown. The analysis assumes a 3 degree glideslope, while the limiting resolution of the EV shown on the HUD and PNF-AD was that of the simulated FLIR (640x480 resolution). The HUD and PNF-AD FLIR resolution was 20 pixels per degree. The OTW resolution provides 26 pixels per degree. The vertical and horizontal resolutions (i.e., pixels per degree) were identical.

The theoretical detection ranges show that the Fire Truck should be detectable at almost twice the distance as the Baggage Cart, particularly its vertical extent (i.e., height). The OTW provided better detection capability because of its higher resolution. Finally, none of the vehicles were theoretically “detectable” in this analysis above 200 ft AFL using the FLIR (HUD or PNF-AD). It should be noted, however, that the Fire Truck was detectable above 200 ft AFL on the EV, *if* the observer was cued to its existence and studied the display.

Table 6. Computed AFL Altitudes for Theoretical Detection of Runway Incursion

Object	Dimension / Scan Direction	AFL Altitude for EVS “Detection”	AFL Altitude for OTW “Detection”
Baggage Cart	Height	42 ft	58 ft
	Width	120 ft	156 ft
Fire Truck	Height	78 ft	101 ft
	Width	190 ft	246 ft

The experimental results showed that, for the 12 flight crews, only one crew member (PNF) saw the baggage cart (they saw it in the OTW) and initiated a go-around. The other 11 crews had a runway incursion with the baggage cart. From the analysis of Table 6, the baggage cart should be “detectable” between 50 and 150 ft AFL but this leaves only 5 to 15 seconds before landing – not very much time to spot a small object, parked slightly short of the intended touchdown zone.

Eleven crews saw the Fire Truck OTW (7 by the PNF, 3 by the PF, and 1 simultaneously by the PF and PNF) and the one remaining crew saw it on the PNF-AD. Upon seeing the incursions, all crews initiated a go-around (all lower than 50 feet AFL).

Baggage Cart

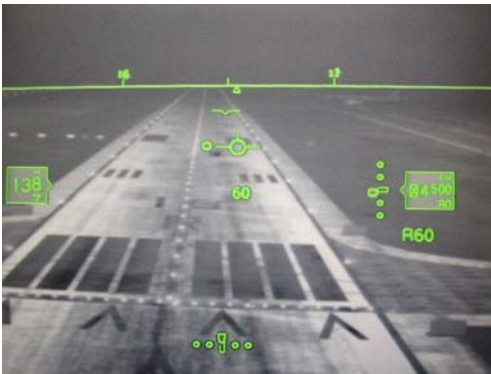
a) Out-The-Window



b) Head-Up Display



c) Auxiliary Display



Fire Truck

d) Out-The-Window



e) Head-Up Display



f) Auxiliary Display

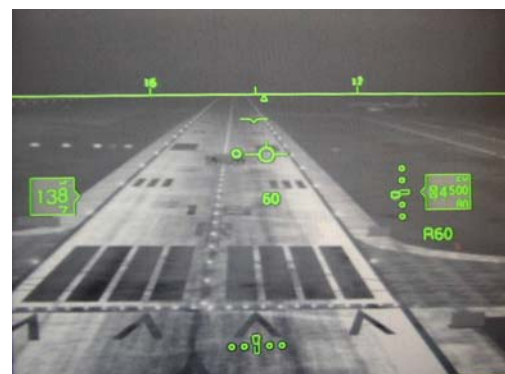


Figure 17. Runway Incursions (Baggage Cart And Firetruck) As Seen At 60 Ft AFL Using Various Displays.

These data highlight several important factors in RI detection by flight crews:

- The vehicles were not moving so motion effects which might aid target detection were absent.

- The incurring vehicle scenarios were intentionally designed as low luminance and color contrast targets. Acuity metrics, such as those touted in Table 6, will provide optimistic detection performance since the metrics typically assume high contrast targets. Whatever color contrast was available between the incurring vehicles and the runway background was only available to the pilots when using the OTW information for detection (see Figure 17a and 17d).
- The incurring vehicles were visible in the PNF-AD and HUD, yet the experimental data suggests that EVS imagery on the HUD and PNF-AD were not useful for RI detection.
- In the HUD, the incurring vehicles were largely occluded by symbology on the HUD (FPM and guidance cue) – see Figure 17b and 17e as examples. Further, the experimental HUD, like all HUDs currently manufactured, have limited contrast ranges. The image contrast is also a function of pilot control. Additionally, the PF is consumed with the task of flying the aircraft and identifying the proper visual references for landing. Therefore, successful runway incursion detection at this late stage on the approach by the PF should not be considered as a likely outcome. The data confirmed this hypothesis in that the PF detected only 4 out of the 24 runway incursions.
- In the PNF-AD, the vehicles were much more apparent than in the HUD, but the data does not indicate significantly greater detection success. The PNF-AD display is more attuned to the RI detection task because of the lack of symbology (in some cases) and its improved gray scale contrast performance (see Figure 17c and Figure 17f as examples of PNF-AD with Symbology). But, as the theoretical data shows, the vehicle size and EVS resolution make detection on the PNF-AD moderately difficult above 200 ft AFL. Below 200 ft AFL, the vehicles were much more obvious in the image, but the PNF noted that they were head-out the vast majority of the time. The head tracking measurements shown above quantify this pilot comment. The PNF was head-out the vast majority of the time, ranging from 86% to 100% of the total time below 200 ft. AFL. Based on these data and the pilot comments, the use of the PNF-AD for incursion detection was not probable. The presence of symbology on the PNF-AD could also obscure the vehicles, particularly if the PNF only used cursory looks at the PNF-AD.
- The use of PNF Fusion controls when available (i.e., toggling between SV and EV imagery) did not statistically help or hinder detection.
- Unlike the PF, the PNF is tasked with monitoring the approach and the landing area. 9 out of 24 incursions were detected by the PNF. But, starting around 500 ft AFL and below, the PNF is going “eyes-out” so runway incursion detection using a head-down display may not be practical (only one PNF saw the incursion on the AD). The off-boresight design of the tested PNF-AD was noted by the pilots as hampering their use of the display on final. Current flight crews are not familiar with using head-down displays on short final to check for incursions. This was not part of the pre-experiment flight crew instructional briefing.
- The results suggest that EV on the HUD and PNF-AD were not useful for RI detection. In the HUD, the incurring vehicles were largely occluded by symbology on the HUD (flight path marker and guidance cue) and the small size and relatively low resolution of HUD made vehicle detection extremely difficult for the PF. In the PNF-AD, the targets were not highly visible in the flight regime where the PNFs were engaged with the display.

The display concepts tested in this experiment – typical of current and future PF-HUD and PNF-AD displays – showed that requirements for display and sensor technology for runway incursion detection should be developed. The requirements for this function should span the breadth of the problem, including human perception, sensor design and detection theory, crew procedures, and crew interface issues. Current flight crews are not necessarily familiar with using head-down displays on short final to check for incursions. Flight crews should be trained on effective use of EV-capable HUDs and more importantly, on the PNF’s use of EV-capable head-down displays to check for incursions, including the use of declutter controls, complementary SV information, and their implications on CRM. The displays are not necessarily optimized for this role. For instance, the capability to provide EV image “zoom” - to increase the EV resolution for object detection before the final phase of the approach – should be considered as it could increase the probability for object detection and recognition before the PNF goes “eyes-out.” The capability to declutter symbology from the displays should be provided and used for a runway “check” before landing.

5.8.2 Navigation Error

The navigation errors were either a 50 foot or 75 foot lateral offset (see Figure 18a for PF-HUD with Tunnel Symbology and in Figure 18b for PNF-AD with FLIR-Symbology examples). The offsets could be detected by either the PF or the PNF. The errors were noticeable from one of several principal ways (depending upon the display configuration) using numerous display indications:

- By a disagreement between the lateral path error and the localizer deviation symbology (HUD and PNF-AD concepts with symbology).
- By a non-zero localizer deviation on the PFD when the PF is flying on the final approach path centerline.
- By differences between the SV and the EV registration using the PNF-AD Fusion controls.
- By differences between the runway outline and the EV imagery of the runway (HUD and PNF-AD concepts with symbology).
- By differences in the pitch/roll guidance symbol and the EV imagery (PF-HUD and PNF-AD concepts with symbology)

a) 50 ft Lateral offset as seen on PF-HUD



b) 75 ft Lateral offset as seen on PNF-AD

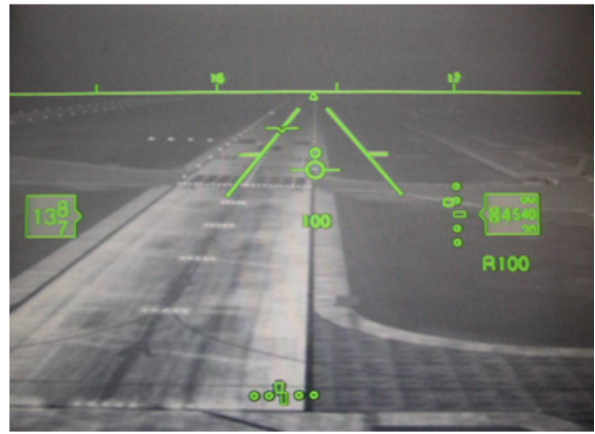


Figure 18. Lateral Navigation Offset Examples.

The majority of flight crews verbally noted the presence of the 50 foot offset (15/24) and 75 foot offset (19/24) during the approach. None of the pilots executed a go-around with this anomaly. Each performed a lateral correction and landed near the runway centerline.

Video analysis showed that navigation errors were predominately noted by the PF (~85%) when they noticed that the pitch/roll guidance symbol was leading them to the left or right of the runway. The “real” runway was detectable in the EV image. One person (flying as the PNF) noted the non-zero localizer deviation on the PFD presentation while tracking the path centerline.

The flight crews were not instructed on the course of action to take when confronted with a navigation error, and the pilots had relatively little training and experience with the system. Despite this, the study showed that lateral navigation errors were verbally acknowledged a significant percentage of time and, even when unrecognized (i.e., not explicitly verbalized), all flight crews landed safely and accurately on the runway. These results suggest that dissociations between raw data, sensor, and/or database presentation should be easily recognized and managed by experienced pilots. Pilot training to recognize these discrepancies could further improve operations in the event of this anomaly.

5.8.3. *Illegal Landings*

Although not technically a “non-normal,” each flight crew was confronted with four trials where weather conditions obscured the required visual cues to complete the landing as defined by FAR §91.175. Of the 48 “illegal landing” rare event trials, during only six of these trials did pilots continue and land the aircraft. (See Table 7.)

Table 7. Illegal Landing Trials

Crew	Observation	HUD	AD
3	Go-around considered	FLIR No Tunnel	Fused No Symbology
4	Threshold lights called at 80 ft AFL	FLIR No Tunnel	FLIR/ Symbology
4	Threshold lights called at 60 ft AFL	FLIR No Tunnel	Fused Symbology
7	Threshold lights called at 100 ft AFL	FLIR Tunnel	Fused Symbology
8	Threshold lights called at 120 ft AFL	FLIR No Tunnel	Fused No Symbology
11	Threshold lights called at 90 ft AFL	FLIR No Tunnel	FLIR No Symbology

The results demonstrated a small but finite potential (12.5%) for flight crews to continue approaches to a landing during visibility conditions that instead require a go-around under the new §91.175 operational rules for EFVS. On all of the illegal landing scenarios, the pilot flying had excellent visibility of the runway using the FLIR on the HUD, but didn't see the runway visual landing references until passing through 100 ft AFL. (The current §91.175 EFVS rule requires visual acquisition of the runway references without use of the EFVS.)

No statistically significant effects of symbology type on the HUD concept was observed (nor was statistical significance expected due to the low statistical power of the rare event) although 5 of the 6 events occurred when the runway outline was not present.

The operational procedures necessary to follow the §91.175 regulation (i.e., for seeing the runway with “un-aided” vision) were generally found to be awkward for the PF, requiring the PF to declutter the HUD or to look-around the HUD combiner. The radio altitude shown on the HUD could be used for judging height above touchdown (HAT).

The PNF was typically “eyes-out” and not closely monitoring the altitude at 100 ft AFL. An aural call-out of 100 ft may have aided adherence to the DA. The aural call-outs were set to Cat. I decision heights.

The few occurrences of “below minimums” landings suggest that the current regulations can be operationally viable. However, an aural call-out at 100 ft AFL is recommended. Nonetheless, there still exists an awkwardness in the transition from EV/HUD-to-visual runway references. The PFs typically commented that the EFVS provided suitable visual references to complete the flare and landing. Future research should examine the effects of eliminating the natural vision/visual segment requirement of FAR §91.175 in other than Category II or III operations while conducting low visibility approach and landing with an EFVS.

6. Conclusions

An experiment was conducted to evaluate the complementary use of SVS and EVS technologies, specifically focusing on new techniques for integration and/or fusion of synthetic and enhanced vision technologies and crew coordination while operating under the newly adopted FAA rules which provide operating credit for EVS.

These data show that significant improvements in SA can be provided by the integration and/or fusion of synthetic and enhanced vision technologies for the PF and PNF. Workload for the PF and PNF was not substantially different when flying the tested concepts. Increasing the “informational complexity” of the HUD by adding SVS and tunnel data, and increasing the number of controls and symbology on a PNF-AD did not affect PF or PNF workload.

In contrast, SA for the PF and PNF was *improved* by the addition of tunnel and SVS on the HUD and by adding fusion control and symbology on the PNF-AD. For the PF-HUD, the overall clutter was not rated as being any worse by the addition of SV and tunnel information. Clutter-SA ratings – a new construct introduced in the test - showed statistically-significant improvements by the addition of the tunnel and SV.

Analysis of the PF-HUD concept data showed that adding tunnel/pathway information reduces the RMS lateral path error on the approach compared to the baseline, no-tunnel HUD concepts. The tunnel provided turn anticipation cues for the turning, descending arrival flown in this test. However, while statistically significant, the differences were not judged to be operationally significant.

The ability of the flight crew to handle a substantial navigational errors was not impacted by the display concepts. They were verbally acknowledged a significant percentage of time and for all these runs, the pilots landed safely and accurately on the runway despite the large lateral navigation error. These results lend convincing evidence to the assertion that the potentially compelling display of pathway information (in the form of a runway outline in this test) does not adversely capture attention nor induce pilots to follow erroneous display information in the presence of real-world visual information or other display cross-checks.

The ability of the flight crew to handle a runway incursion was neither impacted nor significantly aided by the display concepts tested. Although the increase in near-domain symbology information (runway outline) did not degrade pilot response to the Fire Truck runway incursion event, there was also not an observed enhancement in incursion detection as hypothesized for the FLIR. The display concepts and scenarios tested in this experiment – typical of current and future PF HUD and PNF-AD displays - did not show adequate incursion detection functionality. All but one of the runway incursion scenarios were detected without the use of the cockpit displays. Sensor and display design must be tailored to this function and corresponding crew procedures and interfaces developed to support RI detection.

Symbology on the PNF-AD was strongly preferred and rated highly, but the presence of symbology degraded the readability of the raster, particularly of the runway and touchdown point. The PNFs strongly suggested that a declutter capability on the PNF-AD should be included.

7. References

- [1] Joint Planning and Development Office. *Next-Generation Air Transportation System Integrated Plan, Version 1*. Dec 2004.
- [2] Kramer, L.J., Arthur, J.J., Bailey, R.E., and Prinzel, L.J. III, *Flight Testing An Integrated Synthetic Vision System*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2005, Editor: Jacques G. Verly, Volume 5802, Paper 1, April 2005.
- [3] Arthur, J.J., Prinzel, L.J., Kramer, L.J., Bailey, R.E., and Parrish, R.V. *CFIT Prevention Using Synthetic Vision*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2003, Editor: Jacques G. Verly, Volume 5018, Paper 16, April 2003.
- [4] Schiefele, J. et al. *Human Factors Flight Trial Analysis for 3D SVS: Part II*, In Proceedings of SPIE, Enhanced and Synthetic Vision 2005, Editor: Jacques G. Verly, Volume 5802, Paper 18, April 2005.
- [5] Schnell, T., Theunissen, E., and Rademaker, R. *Human Factors Test & Evaluation of an Integrated Synthetic Vision and Sensor-Based Flight Display System for Commercial and Military Applications*. Paper presented at the NATO Research and Technology Organization, Human Factors and Medicine panel workshop entitled "Toward Recommended Methods for Testing and Evaluation of EV and E/SV-Based Visionic Devices", Williamsburg, VA, USA, 26-27 April 2005.
- [6] Arthur, J.J., Kramer, L.J., and Bailey, R.E., and Prinzel, L.J. III. *Flight Test Comparison Between Enhanced Vision (FLIR) and Synthetic Vision Systems*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2005, Editor: Jacques G. Verly, Volume 5802, Paper 3, April 2005.
- [7] Craig, G., Jennings, S., Link, K., and Kruk, R.: *Flight Test of a Helmet-Mounted, Enhanced and Synthetic Vision System for Rotorcraft Operations*. Paper presented at the RTO SET Workshop on "Enhanced and Synthetic Vision Systems" held in Ottawa, CA, Sept. 2002, and published in RTO-MP-107.
- [8] Parrish, R.V., Busquets, A.M., Williams, S.P., & Nold, D.E.: *Evaluation of Alternate Concepts for Synthetic Vision Flight Displays with Weather-Penetrating Sensor Image Inserts During Simulated Landing Approaches*. NASA TP-2003-212643, October 2003.
- [9] Theunissen, E., et al: *Integration of Imaging Sensor Data into a Synthetic Vision Display*. The 23rd Digital Avionics Systems Conference, 2004. DASC 04. Volume 1, 24-28 Oct. 2004. pp. 4.A.1 - 41-10.
- [10] McKay, D., Guirguis, M., Zhang, R. and Newman, R.L.: *Evaluation of EVS for Approach Hazard Detection*. Paper presented at the RTO SET Workshop on "Enhanced and Synthetic Vision System", held in Ottawa, Canada, September 2002, published as NATO RTO-MP-107.
- [11] Harrah, S.D., et al. *The NASA Approach To Realize A Sensor Enhanced-Synthetic Vision System (SE-SVS)*. Paper presented at 21st IEEE / AIAA Digital Avionics Conference, October 2002.
- [12] Uijt De Haag, M., et al. *Flight Test Results of a Synthetic Vision Elevation Database Integrity Monitor*. Paper presented at SPIE's 15th Annual International Symposium on Aerospace / Defense Sensing, Simulation and Controls, AeroSense, Orlando, FL, April 16-20, 2001
- [13] Jones, D. R., Quach, C. C. and Young, S. D.: *Runway Incursion Prevention System – Demonstration and Testing at the Dallas/Fort Worth International Airport*. Proceedings of the AIAA/IEEE 20th Digital Avionics Systems Conference, 2001.

- [14] Young, S.D.: *On the Development of In-flight Autonomous Integrity Monitoring of Stored Geospatial Data Using Forward-looking Remote Sensing Technology*. PhD Dissertation, Ohio University, March 2005.
- [15] Ganoe, G., and Young, S.D.: *Utilization of GPS Surface Reflected Signals To Provide Aircraft Altitude Verification For SVS*. SPIE. In Proceedings of SPIE, Enhanced and Synthetic Vision 2005 Editor: Jacques G. Verly, Volume 5802-12.
- [16] Federal Aviation Administration: *HUD Certification Criteria - Working Paper*. PS-ANM100-2001-00085 (Draft), Dated January 13, 2005.
- [17] Bender, E.J., Reese, C.E., and van der Wal, G.S.: *Comparison of Additive Image Fusion vs. Feature-Level Image Fusion Techniques for Enhanced Night Driving*. Low-Light-Level and Real-Time Imaging Systems, Components, and Applications, C. Bruce Johnson, Divyendu Sinha, Phillip A. Laplante, Editors, Proceedings of SPIE, Vol. 4796 (2003), pp. 140-151.
- [18] Endsley, M. R.: *Situation Awareness in Aviation Systems*. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), *Human factors in Aviation Systems* (pp. 257-276). Hillsdale, N.J.: Lawrence Erlbaum. 1999.
- [19] Simard, P., Link, N.K., and Kruk, R.V.: *Feature Detection Performance with Fused Synthetic and Sensor Images*. Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, pp. 1108-1112, 1999.
- [20] Andre, A.D., Wickens, C.D., and Moorman, L.: *Display Formatting Techniques for Improving Situation Awareness in the Aircraft Cockpit*. International Journal of Aviation Psychology, Vol. 1, No. 3, 1991, pp. 205-218.
- [21] Greig, I., Shaw, G., and Lumsden, B.: *Augmented Vision Systems for the All Weather Autonomous Landing Guidance of Fixed Wing Aircraft: An Investigation of the Key Parameters by Piloted Flight Simulation*. From Proceedings of the 10th CEAS Conference on Free Flight: 1997. pp, 32-1 to 32-15
- [22] Mann, T.L.: *Autonomous Landing Guidance Concept - The Effects of Video and Symbology Dynamics on Pilot Performance*. Paper presented at the 6th SAE Aerospace Behavioural Engineering Technology Conference, Paper 872390, 1988, pp. 131-138.
- [23] Schnell, T.; Ellis, K.; and Etherington, T.: *Flight Simulator Evaluation Of An Integrated Synthetic And Enhanced Vision System For Terrain Avoidance*. The 24th Digital Avionics Systems Conference, DASC 2005. Volume 1, 30 Oct.-3 Nov. 2005, pp: 4.E.4-1 - 4.E.4-18.
- [24] Bailey, R.E., Parrish, R.V., Kramer, L.J., Harrah, S.D., and Arthur, J.J. (Trey) III: *Technical Challenges In the Development of a NASA Synthetic Vision System Concept*. Paper presented at the RTO SET Workshop on “Enhanced and Synthetic Vision System”, held in Ottawa, Canada, September 2002, published as NATO RTO-MP-107.
- [25] Crawford, J. & Neal, A.: *A Review Of The Perceptual And Cognitive Issues Associated With The Use Of Head-Up Displays In Commercial Aviation*. The International Journal Of Aviation Psychology, 16(1), 2006. pp. 1–19.
- [26] Prinzel III, L. J. & Risser, M.: *Head-Up Displays And Attention Capture*. NASA TM-2004-213000. Washington, DC: NASA, 2004.
- [27] Wickens, C. D., Ververs, P. M., & Fadden, S.: *Head-Up Displays*. In D. Harris (Ed.), “Human Factors For Civil Flight Deck Design.” Burlington, VT: Ashgate Publishing, Ltd. 2004, pp. 103-140.

- [28] Yeh, M., Merlo J. L., Wickens, C. D. & Brandenburg, D. L.: *Head Up Versus Head Down: The Costs Of Imprecision, Unreliability, And Visual Clutter On Cue Effectiveness For Display Signaling*. Human Factors, 45(3), 2003, pp. 390-407.
- [29] Snow, M.P., & French, G.A.: *Human Factors In Head-Up Synthetic Vision Displays*. In proceedings of the 2001 Aerospace Congress (ACE2001CD-01-2652). Warrendale, PA: Society of Automotive Engineers (2001)
- [30] Ververs, P. M., & Wickens, C. D.: *Head-Up Displays: Effects Of Clutter, Display Intensity, And Display Location On Pilot Performance*. International Journal of Aviation Psychology, 8(4), 377-403, 1998.
- [31] Snow, M.P., Reising, J.M., & Liggett, K.K.: *Flying Complex Approaches Using A Head-Up Display: Effects Of Visibility And Display Type*. Air Force Research Laboratory, Human Effectiveness Directorate Report No. AFRL-HE-WP-TR-2000-0011. 2000.
- [32] Ercoline, W.: *The Good, The Bad, and The Ugly in Head-Up Displays*. IEEE Engineering in Medicine and Biology, 19, 2000. pp. 66-70.
- [33] Alexander, A. L., Wickens, C. D., & Hardy, T. J.: *Synthetic vision systems: The effects of guidance symbology, display size, and field-of-view*. Human Factors, 47, 2005. pp. 693-707.
- [34] Kramer, L.J., Prinzel, L.J., Arthur, J.J., & Bailey, R.E. *Pathway Design Effects On Synthetic Vision Head-Up Displays*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2003, Editor: Jacques G. Verly,, Volume 5424-08.
- [35] Larimer, J., Pavel, M., Ahumada, A. and Sweet, B.: *Engineering a Visual System for Seeing Through Fog*. AeroTech '92, SAE Paper No. 921130, pp. 428-432. 1992.
- [36] Boff, K.R. and Lincoln, J.E., (Editors). *Engineering Data Compendium: Human Perception and Performance*. H.G. Anderson Medical Research Laboratory, Ohio.
- [37] Federal Aviation Administration Advisory Circular: *Standard Operating Procedures for Flight Deck Crewmembers*, AC 120-71A, Dated February 27, 2003.
- [38] Lorenz, B. and Korn, B.: *Crew Coordination Issues Of EVS-Approaches*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2004, Editor: Jacques G. Verly, Vol. 5424, pp. 89-97.
- [39] Tuttle, T.T., Imrich, T., Etherington, T.J., and Theunissen, E.: *Terminal Area Operations with Enhanced and Synthetic Vision: Experience in the Boeing Technology Demonstrator*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2003, Editor: Jacques G. Verly, Vol. 5081, pp. 136-145.
- [40] Merrick, V.K. and Jeske, J.A., *Flightpath Synthesis and HUD Scaling for V/STOL Terminal Area Operations*. NASA Technical Memorandum 110348, April 1995.
- [41] Kramer, L.J., Prinzel III, L.J., Arthur III, J.J., & Bailey, R.E. (2005). *Advanced Pathway Guidance Evaluations On A Synthetic Vision Head-Up Display*. NASA Langley Research Center: NASA Technical Paper TP-2005-213782.
- [42] Foyle, D.C and Hooley, B.L., *Improving Evaluation and System Design Through The Use of Off-Nominal Testing: A Methodology for Scenario Development*. Proceedings of the 12th International Symposium on Aviation Psychology, pp. 397-402. Dayton, OH.
- [43] Ames, Lawrence L. & George, Edward J.: *Revision and Verification of a Seven-Point Workload Estimation Scale*. Air Force Flight Technical Center: AFFTC-TIM-93-01, 1993.

- [44] American Institute of Aeronautics and Astronautics: *Guide to Human Performance Measurement*. Washington D.C.: American Institute of Aeronautics and Astronautics. 1993.
- [45] Vidulich, M.A. & Hughes, E.R.: *Testing A Subjective Metric Of Situation Awareness*. Proceedings of the Human Factors & Ergonomics Society, 35, 1307-1311, 1991.
- [46] RTCA: *Minimum Aviation System Performance Standards: Required Navigation Performance for Area Navigation*. RTCA DO-236A. Washington, D.C.: RTCA, Incorporated, 2000.
- [47] Kramer, L.J., Prinzel, L.J., Bailey, R.E., & Arthur, J.J.: *Synthetic Vision Enhances Situation Awareness And RNP Capabilities For Terrain-Challenged Approaches*. Proceedings of the American Institute of Aeronautics and Astronautics Third Aviation Technology, Integration, and Operations Technical Forum, AIAA 2003-6814, 1-11, 2003.

Appendix A: Head Tracker Data Analysis Details

A head-tracker was used in an attempt to quantify what display the pilot not flying (PNF) was using and when. Because an eye-tracker was not used, the head-tracking data are only an estimate for this. For this reason, the head-tracker data could only be reliably used for significant spatial location differences; that is, whether the pilot was looking forward (i.e., looking out the window or the head-down displays), looking at the AD (outboard of the PNF), or neither.

In this appendix, the head-tracker statistical data analysis is detailed.

In Table A-1, the simple ANOVAs of the percentage of time that the PNF's head was directed at either the AD, out the window/head-down display (OTW/HDD), or elsewhere is shown. The data are shown as a function of the aircraft altitude (i.e., Above Field Elevation (AFL), in feet). (The mean percentages are shown in Table 4 in the text.)

Table A-1. ANOVAs on Percentage Dwell Time at the AD Concepts for each Altitude Segment

Altitude Segment		Sum of Squares	df	Mean Square	F	Significance
AFL > 1000	Between Groups	59848.188	3	19949.396	101.631	0.000
	Within Groups	264799.155	1349	196.293		
	Total	324647.344	1352			
500 < AFL < 1000	Between Groups	63680.988	3	21226.996	29.231	0.000
	Within Groups	979616.097	1349	726.179		
	Total	1043297.085	1352			
400 < AFL < 500	Between Groups	44004.709	3	14668.236	14.692	0.000
	Within Groups	1346777.572	1349	998.353		
	Total	1390782.281	1352			
300 < AFL < 400	Between Groups	39969.357	3	13323.119	18.876	0.000
	Within Groups	952177.825	1349	705.840		
	Total	992147.183	1352			
200 < AFL < 300	Between Groups	6085.649	3	2028.550	4.808	0.002
	Within Groups	569158.532	1349	421.911		
	Total	575244.181	1352			
100 < AFL < 200	Between Groups	2895.730	3	965.243	2.914	0.033
	Within Groups	443897.097	1340	331.266		
	Total	446792.827	1343			
AFL < 100	Between Groups	813.379	3	271.126	.664	.574
	Within Groups	520396.461	1274	408.474		
	Total	521209.840	1277			

Post-hoc multiple comparison tests (SNK, $\alpha=0.05$) revealed 4 subsets, shown in Table A-2, when ownship was above 1000 ft AFL, based on percentage of dwell time for each AD concept. The four unique subsets for percentage of dwell time indicate that more dwell time was exhibited on formats having symbology, and those formats having synthetic and enhanced vision information (“fused”). The most interesting comparison is that the pilots focused on display concepts with symbology more so than those with terrain information (SV or EV). (Given the simulated weather conditions, the FLIR did not provide any terrain information above 1000 ft AFL.) For instance, the FLIR-Tunnel configuration induced more gaze on the part of the PNFs than the Fused-No Symbology configuration (terrain information but no symbology).

Table A-2. Student-Newman-Keuls Test for Altitude Segment: > 1000 ft AFL

Aux. Display	N	Subset for $\alpha=0.05$			
		1	2	3	4
FLIR – No Symbology	339	4.757			
Fused – No Symbology	354		12.962		
FLIR - Symbology	309			15.556	
Fused - Symbology	351				23.213

Post-hoc tests (SNK, $\alpha=0.05$) revealed 4 subsets, shown in Table A-3, when ownship was above 500 ft but less than 1000 ft AFL. Based on percentage of dwell time for each AD concept in this altitude band, the same four unique subsets - and order - were found as previous. The percentage of dwell times indicate again, that more dwell time was exhibited on formats having symbology, and those formats having synthetic and enhanced vision information (“fused”). The pilots focused on display concepts with symbology more so than those with terrain information (SV or EV).

Table A-3. Student-Newman-Keuls Test for Altitude Segment: 500 < AFL < 1000 ft

Aux. Display	N	Subset for $\alpha=0.05$			
		1	2	3	4
FLIR - No Symbology	339	17.796			
Fused - No Symbology	354		22.340		
FLIR - Symbology	309			30.324	
Fused - Symbology	351				35.330

As shown in Table A-4, post-hoc tests (SNK, $\alpha=0.05$) revealed just two unique subsets for percentage of dwell time in the altitude band of 400 ft AFL to 500 ft AFL. More dwell time was exhibited on formats having symbology than those without.

Table A-4. Student-Newman-Keuls Test for Altitude Segment: 400 < AFL < 500 ft

Aux. Display	N	Subset for $\alpha=0.05$	
		1	2
Fused - No Symbology	354	24.292	
FLIR - No Symbology	339	24.330	
FLIR - Symbology	309		35.519
Fused - Symbology	351		35.891

Post-hoc tests (SNK, $\alpha=0.05$) revealed just two unique subsets for percentage of dwell time in the altitude band of 300 ft AFL to 400 ft AFL, in Table A-5. Identically to the previous altitude band, more dwell time was exhibited on formats having symbology than those without.

Table A-5. Student-Newman-Keuls Test for Altitude Segment: 300 < AFL < 400 ft

Aux. Display	N	Subset for $\alpha=0.05$	
		1	2
FLIR - No Symbology	339	14.934	
Fused - No Symbology	354	15.003	
FLIR - Symbology	309		24.564
Fused - Symbology	351		26.764

For the altitude band from 200 to 300 ft AFL, post-hoc tests, shown in Table A-6, (SNK, $\alpha=0.05$) revealed two overlapping subsets for percentage of dwell time. The fused/symbology AD format had statistically significant (but perhaps not operationally significant) more dwell time than the 2 AD concepts without symbology. Again, more dwell time tended toward formats having symbology than those without.

Table A-6. Student-Newman-Keuls Test for Altitude Segment: 200 < AFL < 300 ft

Aux. Display	N	Subset for $\alpha=0.05$	
		1	2
FLIR - No Symbology	339	8.466	
Fused - No Symbology	354	10.358	
FLIR - Symbology	309	11.531	11.531
Fused - Symbology	351		14.252

For the altitude band from 100 to 200 ft AFL, post-hoc tests (SNK, $\alpha=0.05$), shown in Table A-7, revealed two overlapping subsets for percentage of dwell time. The FLIR/No symbology AD format had statistically significant (but perhaps not operationally significant) less dwell time than the 2 AD concepts with symbology.

Table A-7. Student-Newman-Keuls Test for Altitude Segment: 100 < AFL < 200 ft

Aux. Display	N	Subset for $\alpha=0.05$	
		1	2
FLIR –No Symbology	336	5.938	
Fused - No Symbology	351	7.367	7.367
Fused – Symbology	351		9.246
FLIR – Symbology	306		9.599

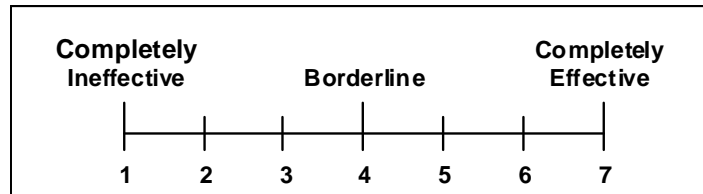
Below 100 ft AFL, post-hoc tests (SNK, $\alpha=0.05$) showed no statistically-significant differences between display configurations.

Appendix B. Final Questionnaire Subject Ratings and Comments

In the following, the responses and subsequent statistical analyses to the post-test questionnaires administered to the flight crews are presented. The questions are written in italics.

Display Concept Questions:

1. *On a scale of 1 to 7, please rate the effectiveness of each concept in terms of conducting an approach in IMC as the Pilot Flying (from the left seat with the HUD).*

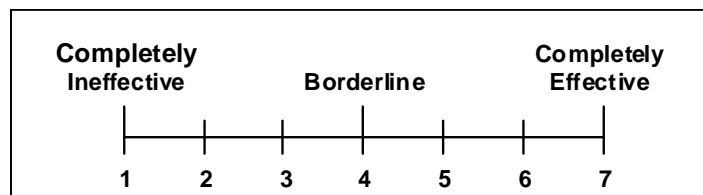


An ANOVA revealed that the HUD display concepts were significantly different ($F(3, 63) = 15.96$, $p < 0.0001$) for their effectiveness in conducting an approach in IMC for the PF. Post-hoc tests (SNK, $\alpha = 0.05$) revealed 3 unique subsets. Pilots were able to effectively conduct an approach in IMC with all of the HUD concepts but they rated the Fused HUD-Tunnel concept as being the most effective of the 4 tested.

Display Concept	Mean Subject Ratings		
	Subset 1	Subset 2	Subset 3
HUD EVS without tunnel	5.0		
HUD EVS with tunnel		5.7	
HUD Fused EVS&SVS without tunnel		5.7	
HUD Fused EVS&SVS with tunnel			6.6

Seven pilots provided comments. With SVS, the horizon is inertially stabilized which is a major improvement in terms of conducting an approach in IMC. The HUD fused image (SVS/EVS) provides the best situation awareness and the tunnel provided the most precise path control during the approach. However once stabilized on approach, the tunnel seems to add clutter and detract from EVS and threshold light acquisition capability. Four out of the 7 pilots commented that fused HUD with tunnel provided the greatest SA. One pilot commented that he found the tunnel only useful for turn anticipation cues.

2. *On a scale of 1 to 7, please rate the effectiveness of each concept in terms of conducting an approach in IMC as the Pilot Not Flying (from the right seat with the Auxiliary Display).*

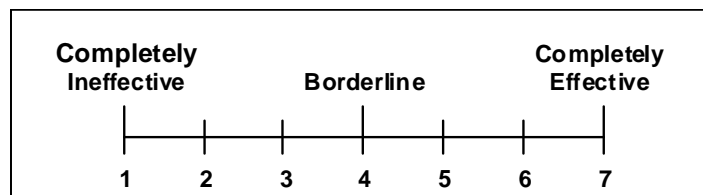


An ANOVA revealed that the AD display concepts were significantly different ($F(3, 60) = , p < 0.0001$) in their effectiveness when conducting an approach in IMC as the PNF. Post-hoc tests (SNK, $\alpha = 0.05$) revealed 3 unique subsets. The PNFs were able to effectively conduct an approach in IMC with all the AD concepts but they rated the Fused AD – With Symbology concept as being the most effective of the 4 tested.

Display Concept	Mean Subject Ratings		
	Subset 1	Subset 2	Subset 3
AD with EVS only	4.5		
AD with EVS and symbology		5.0	
AD with Fused EVS/SVS only		5.3	
AD with Fused EVS/SVS and symbology			6.0

Eight pilots added comments. Would like to see wind vector and bank angle indicator included in AD symbology and symbology should be the same as that found on HUD for crew resource management (CRM) correlation. Symbology helped with situation awareness and CRM. PNF’s highest priority is to back up the PF by cross-referencing position, confirming EVS acquisition, and confirming threshold acquisition; all while communicating in a concise standardized format that builds situation awareness. SVS on the AD is nice for overall SA but not essential on the AD. Once established on glideslope, EVS and symbology are the most effective elements on the AD. Fused image with symbology enhanced CRM and SA for the PNF. One pilot commented that the symbology blocked the image he was most interested in seeing.

3. *On a scale of 1 to 7, please rate the overall effectiveness of each concept for flight path awareness and performance as you performed the evaluations as the Pilot Flying (from the left seat with the HUD).*

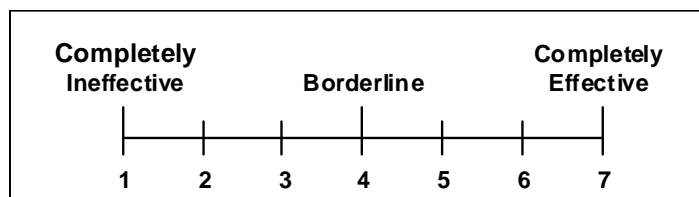


An ANOVA revealed that the HUD display concepts were significantly different ($F(3,63) = 19.50, p < 0.0001$) in their effectiveness for the PF’s flight path awareness and performance. Post-hoc tests (SNK, $\alpha = 0.05$) revealed 3 unique subsets. All display concepts were effective for flight path awareness and performance but they rated the Fused HUD-Tunnel concept as being the most effective of the 4 tested.

Four pilots provided comments. One pilot felt there was considerable excess information on all HUD concepts and that a minimum symbology set of AD on the HUD would be a vast improvement. The tunnel can become “cluttersome” at times (e.g., once wings-level on final), but it definitely assists in anticipating what is approaching in the flight path course (e.g., turns). Flight director and flight path marker are the best symbolic tools; while, the SVS/EVS gives a perception of reduced workload and easier correction and detection of deviations.

Display Concept	Mean Ratings		
	Subset 1	Subset 2	Subset 3
HUD EVS without tunnel	4.7		
HUD EVS with tunnel		5.8	
HUD Fused EVS&SVS without tunnel		5.5	
HUD Fused EVS&SVS with tunnel			6.5

4. On a scale of 1 to 7, please rate the overall effectiveness of each concept for terrain and obstacle awareness as you performed the evaluations as the Pilot Flying (from the left seat with the HUD).



An ANOVA revealed that the HUD display concepts were significantly different ($F(3,63) = 24.06$, $p < 0.0001$) in their effectiveness for the PF's terrain and obstacle awareness. Post-hoc tests (SNK, $\alpha = 0.05$) revealed 2 unique subsets. Terrain and obstacle awareness was very effectively provided by the Fused HUD concepts (mean=6.2) but the FLIR-only HUD concepts (mean=4.0) were rated on average as being only borderline effective.

Display Concept	Mean Subject Ratings	
	Subset 1	Subset 2
HUD EVS without tunnel	3.8	
HUD EVS with tunnel	4.1	
HUD Fused EVS&SVS without tunnel		6.0
HUD Fused EVS&SVS with tunnel		6.4

Six pilots commented. Inertial stabilized horizon and pitch scale improve pilot output in high gain tasks. EVS allowed last minute awareness of terrain and obstacles on very short final approach but by this time the pilot's attention was focused on approach light and runway environment acquisition. Tunnel did not affect terrain awareness and SVS is most effective for terrain and obstacle awareness. With or without tunnel, on short final to landing, the pilot is so focused on landing that it is difficult to see any obstacles with EVS and SVS.

5. What features did you like the best on the HUD display concept? Which features did you like the least on the HUD display concept?

Best features:

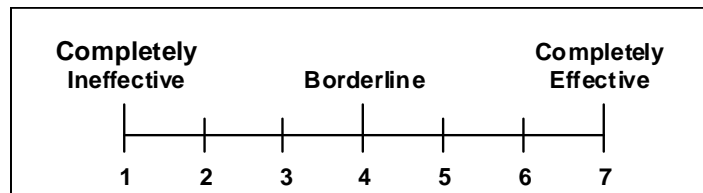
- Declutter option.
- Tunnel (10 pilots).

- Fused EVS/SVS (3 pilots) makes it easier to fly safer approaches
- Tunnel with flight path marker and guidance cue (2 pilots).
- Flight data, synthetic terrain, and tunnel.
- Fusion with no tunnel.
- Flight director and flight path marker on the HUD; phasing out SVS at 500 feet; ability to see runway and landing environment in EVS; SVS/EVS workload reduction turning to final; HUD runway outline and flight director good cross check for positional error.
- SVS
- Flight path marker with pitch reference.

Worst features:

- Inability to adjust HUD symbology brightness independently (3 pilots).
- Didn't like having to declutter HUD or look around HUD at minimums to determine if you should continue approach or go-around as this is a critical time in the approach. (3 pilots)
- Need provision for barometric decision altitude to be set and indicated at DA +100 feet.
- Too much symbology on the HUD, needs to be simplified (2 pilots).
- Remove tunnel once localizer and glideslope are captured or at wings level and < 1000 ft Above Ground Level (AGL) (2 pilots).
- HUD symbology is too bright – obstructs threshold acquisition (3 pilots).
- Have transition from SVS to EVS later in approach (~350 ft AGL)
- Large size of display – information was too spread out for precision instrument task.

6. *On a scale of 1 to 7, please rate the overall effectiveness of each concept for flight path awareness and performance as you performed the evaluations as the Pilot Not Flying (from the right seat with the Auxiliary Display).*

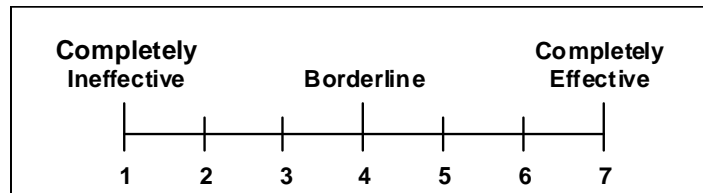


An ANOVA revealed that the AD display concepts were significantly different ($F(3, 63) = 24.3$, $p < 0.0001$) for the PNF's flight path awareness and performance. Post-hoc tests (SNK, $\alpha = 0.05$) revealed 3 unique subsets. The AD concept with Fused EVS/SVS and symbology was rated as being the most effective of the 4 tested. Similarly, the AD with EVS and symbology and the AD concept with Fused EVS/SVS (but without symbology) were somewhat effective for the monitoring pilot's flight path awareness. An AD concept with only EVS was rated as being borderline to marginally ineffective for the PNF's flight path awareness and performance.

Display Concept	Mean Subject Ratings		
	Subset 1	Subset 2	Subset 3
AD with EVS only	3.4		
AD with EVS and symbology		4.9	
AD with Fused EVS/SVS only		4.7	
AD with Fused EVS/SVS and symbology			6.0

Three pilots commented. Fused is fine but once wings-level on final less than 500 feet AGL, EVS and symbology is most effective. PNF mission is to back-up PF performance with AD symbology and to clear runway of obstacles.

7. *On a scale of 1 to 7, please rate the overall effectiveness of each concept for terrain and obstacle awareness as you performed the evaluations as the Pilot Not Flying (from the right seat with the Auxiliary Display).*



An ANOVA revealed that the AD display concepts were significantly different ($F(3,63) = 24.06$, $p < 0.0001$) for their effectiveness for the non-flying pilot’s terrain and obstacle awareness. Post-hoc tests (SNK, $\alpha = 0.05$) revealed 2 unique subsets. A non-flying pilot’s terrain and obstacle awareness was very effective (mean=5.9) when using the Fused AD concepts but were only borderline effective with the FLIR-only AD concepts (mean=3.9).

Display Concept	Mean Subject Ratings	
	Subset 1	Subset 2
AD with EVS only	3.6	
AD with EVS and symbology	4.1	
AD with Fused EVS/SVS only		5.6
AD with Fused EVS/SVS and symbology		6.2

Three pilots commented. Having a reliable 3D view, with or without symbology, is hard to beat. Fused EVS and SVS for terrain and obstacle detection. Once wings level on glide path, just EVS and symbology.

8. *What features did you like the best on the PNF Auxiliary Display concept? Which features did you like the least on the PNF Auxiliary Display concept?*

Best features:

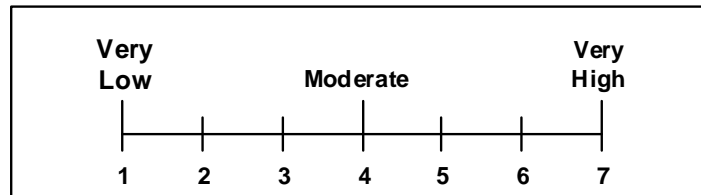
- Uncluttered presentation of symbology (4 pilots). Helpful in effective monitoring of the pilot flying.
- Symbology, fused (2 pilots).
- SVS display for situation awareness.
- Used as VMC cross-check of instruments
- Fused SVS and EVS (4 pilots)
- Fused above 500 ft AGL, EVS only on landing for situation awareness and safety,
- Knowledge of what pilot is seeing.
- Uncluttered display of EVS and symbology in acquiring EVS lights and threshold lights and instant flightpath awareness on final approach (2 pilots).

- Pilot’s ability at will to select fused or EVS only.
- Color-coding (“greening”) of imagery when EVS and SVS agree.

Worst features:

- Location – should be in pilot’s normal scan as moving head back and forth from side to front could cause spatial disorientation (9 pilots).
- Add wind vector and bank angle to symbology.
- No symbology condition (3 pilots) as had to scan flight instruments for cross-check.
- Inability to declutter symbology.

9. For each display concept, please indicate the workload as you performed the evaluations as the Pilot Flying (from the left seat with the HUD).

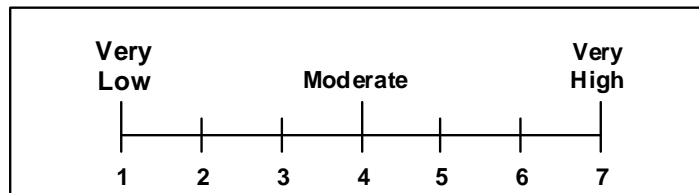


An ANOVA revealed that the HUD display concepts were significantly different ($F(3,63)=10.19$, $p<0.0001$) as they affect PF workload. Post-hoc tests (SNK, $\alpha=0.05$) revealed 2 unique subsets. Pilots workload was moderately low (mean=3.6) when using the tunnel HUD concepts but was between moderate and moderately high with the non-tunnel HUD concepts (mean=4.6).

Display Concept	Mean Subject Ratings	
	Subset 1	Subset 2
HUD EVS without tunnel	4.8	
HUD EVS with tunnel		3.7
HUD Fused EVS&SVS without tunnel	4.3	
HUD Fused EVS&SVS with tunnel		3.5

Six pilots commented. Stabilized horizon from SVS helps keep workload from increasing during high-gain tasks. Tunnel added workload due to its high precision but definitely raised SA.

10. For each display concept, please indicate the workload as you performed the evaluations as the Pilot Not Flying (from the right seat with the Auxiliary Display).



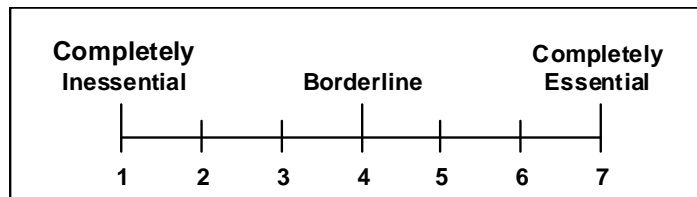
An ANOVA revealed that the AD display concepts were significantly different ($F(3,63)=5.91$, $p=0.001$) for PNF workload. Post-hoc tests (SNK, $\alpha=0.05$) revealed 2 overlapping subsets. The AD FLIR with no

symbology had the highest workload of the 4 concepts tested but it was still rated moderately low by the non-flying pilots.

Display Concept	Mean Subject Ratings	
	Subset 1	Subset 2
AD with EVS only	3.6	
AD with EVS and symbology		2.8
AD with Fused EVS/SVS only	3.2	3.2
AD with Fused EVS/SVS and symbology		2.7

Five pilots commented. Workload was lower with EVS only as there wasn't much to look at until short final. Symbology was difficult to look through on final. Workload was increased with FLIR only as pilot had to continuously check when FLIR could be seen. Increased workload due to choosing alpha and beta values but that choice is a good thing. Decluttered HUD symbology on AD was very effective for helping the pilot flying get the EVS lights call by 200 feet. SVS reduced discrimination and added clutter to the AD. AD presentation provided comfort of visual approach.

11. Please rate the essentialness of having a the tunnel concept depicted on the HUD in the presence of raster information (either SVS or EVS)



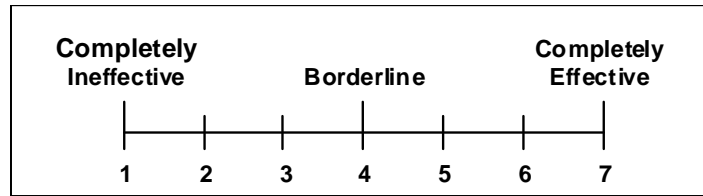
The mean “essentialness” rating: was 5.0 with a standard deviation of 1.6. On average, pilots felt that it was essential to include a tunnel on a HUD in the presence of EV and/or SV imagery.

Please provide any comments with regard to how the tunnel concept was displayed. Is this tunnel concept a useful one? If so, how? If not, please provide a reason and any suggested improvements.

Eighteen pilots commented. Twelve of the pilots felt that the tunnel, especially using it with the guidance cue and flight path marker, was a very useful concept. With the tunnel, pilots were able to anticipate upcoming course changes and understand the future track of the aircraft and it increased their ability to scan other instruments. Two of the pilots felt that the tunnel was useful but not an essential element of the HUD symbology set. Pilot suggestions included the ability to turn the tunnel off; to increase the tunnel spacing to 0.2 or 0.3 nmi from the current 0.1 nmi separation; and making the tunnel symbology dimmer than the other symbology. Two of the pilots felt that if turn anticipation cues could be provide by other HUD symbology than the banking tunnel elements that the tunnel would not be as necessary. 3 pilots felt that at times the tunnel added too much clutter to the HUD.

Auxiliary Display Symbology Questions

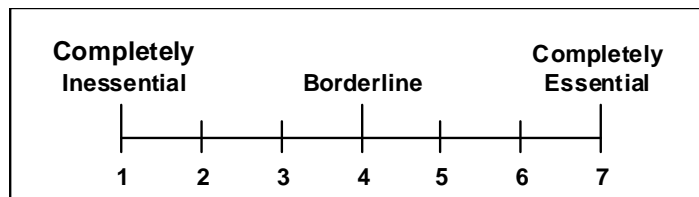
12. Please rate the effectiveness of the auxiliary display symbology for:



- a) *Awareness of future flight path:* Mean Rating of 4.23, Standard Deviation of 1.9.
- b) *Awareness of flight path errors:* Mean Rating of 4.95, Standard Deviation of 1.7.
- c) *Recognition of obstacles:* Mean Rating of 4.77, Standard Deviation of 1.9.

On average, pilots rated the symbology on the AD as being somewhat effective for awareness of future flight path and effective for awareness of flight path errors and recognition of obstacles.

13. *Please rate the essentialness of having symbology depicted on the auxiliary display in the presence of raster information*



The mean rating for the “essentialness” of AD symbology was 5.45 with a standard deviation 1.7. On average, pilots rated having symbology on the AD with imagery as being essential to highly essential.

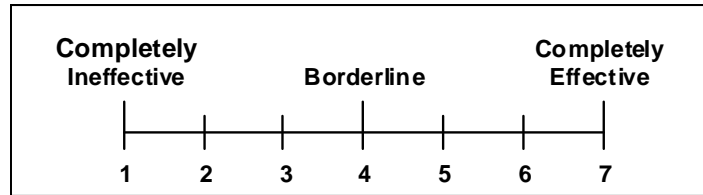
Please provide any comments with regard to how the symbology was displayed. Is this symbology set a useful one? If so, how? If not, please provide a reason and any suggested improvements.

18 pilots commented. In general, pilots felt that it was essential to have symbology on the AD and that a pilot-selectable declutter capability should be available. Pilot commentary indicated that the minimum symbology set was absolutely necessary in understanding any perception problems the PF might be encountering and that the PNF’s job is to monitor the approach both from an aircrew performance standpoint and a systems performance standpoint. In order to have effective crew resource management and data correlation, pilots felt that the AD symbology should be a subset of the symbology used on the PF HUD. Pilots commented that the current AD symbology set was uncluttered, provided just the right amount of information, and was easy to reference positional error while optimizing the capability of the EVS. Pilots felt that the AD should be mounted more in the forward view. Only one pilot commented that there was no usefulness in having symbology on the AD.

Runway Incursion Questions

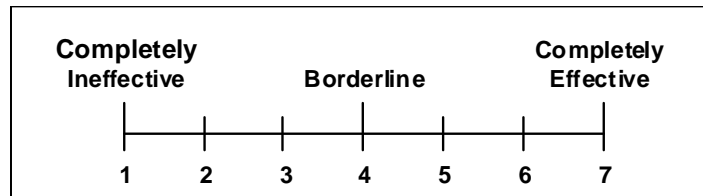
The FAA (2006) defines runway incursion as, "any occurrence in the airport runway environment involving an aircraft, vehicle, person, or object on the ground that creates a collision hazard or results in a loss of required separation with an aircraft taking off, intending to take off, landing, or intending to land."

14. On a scale of 1 to 7, please rate the effectiveness of the display concept for situation awareness of aircraft/obstacles on the runway as you performed the evaluations as the Pilot Flying (from the left seat with the HUD).



Pilot comments indicated that in general they did not see either obstacle on the HUD in the EVS imagery. They commented that since they were required to use their natural vision to see the runway environment that they were not using the EVS imagery to see the fire truck or baggage care late in the approach. It was suggested that part of the Standard Operating Procedures (SOPs) should be for the PNF to use the AD to check for obstacles on the runway after the PF has been assured of the EVS lights and raw visual of the runway environment. In essence, this would become a major function of PNF because the PF is in a difficult visual transition from short sight HUD scan to far sight runway-touchdown scan. Pilots also commented that the HUD symbology (flight path marker and guidance cue) might have been occluding the obstacles but that the tunnel did not as it was decluttered from the symbology by 500 feet AGL.

15. On a scale of 1 to 7, please rate the potential effectiveness of the HUD display concept for prevention of runway incursions in real-world operations.

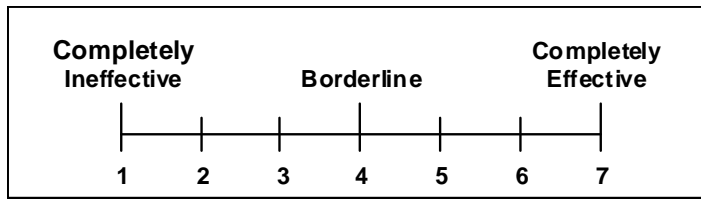


Pilots commented that maybe the HUD symbology should be dimmer and that the symbology might require more of their attention so that they have less attention to detect obstacles. Other pilots also commented that seeing obstacles is probably going to be a function of the PNF. Another pilot commented that with training and proper fusion algorithms for the EVS and SVS that these elements could be effective tools in preventing mishaps caused by runway incursions.

16. While making the approaches as the Pilot Flying, please comment on the effect of symbology and/or raster on the Head-Up Display during a runway incursion?

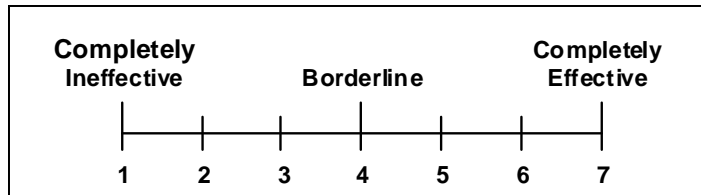
Pilots suggested that having the symbology dimmer and the EVS returns brighter would help detect a runway incursion. They also commented the flight path marker and guidance cue tend to obscure the runway and that HUD tends to reduce their focal scan especially under 100 ft AGL. They commented that concentrating on the symbology and raster image might distract them from detecting runway incursions.

17. On a scale of 1 to 7, please rate the effectiveness of the display concept for situation awareness of aircraft/obstacles on the runway as you performed the evaluations as the Pilot Not Flying (from the right seat with the Auxiliary Display).



Pilots commented that late in the approach the PNF is concentrating out the front of the plane and looking at the AD is not “natural”. In general, the pilots did not look at the AD below 500 feet on the approach. Two different training/SOP techniques were offered by the pilots: 1) have the PNF assure “EVS lights” and "landing" calls by PF and then have the PNF scan the FLIR down the runway and require a "clear deck" call if RVR<4000; or, 2) require PNF to be head-up and forward after "EVS lights" is announced by PF.

18. *On a scale of 1 to 7, please rate the potential effectiveness of the PNF Auxiliary Display concept for prevention of runway incursions in real-world operations.*



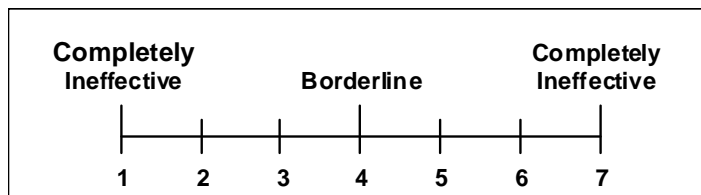
Pilots commented that with training and development of proper procedures for the PNF that the AD could be a great advancement for runway incursion protection. Pilots also suggested that the AD be mounted in a forward location and that there be a way to adjust the symbology brightness and to declutter the symbology when needed.

19. *While making the approaches as the Pilot Not Flying, please comment on the effect of symbology and/or raster on the Auxiliary Display during a runway incursion?*

Pilots commented that the AD needs to be mounted in a forward location and that the PNF needs to be able to declutter the symbology when necessary. Pilots commented that the presence of symbology could prevent detecting an incurring vehicle but was necessary to monitor the PF’s flying of the approach. One pilot thought this was the greatest benefit in the whole experiment as the PNF can use the tool to scan the touchdown zone on short final.

Head-Up Display Transition Questions

20. *Please rate the effectiveness of the HUD SVS/EVS transition that lasted 6 seconds and was finished at 500 feet. Please comment on this transition strategy and point.*



The mean “effectiveness” was 5.38 with a standard deviation of 1.4. On average, pilots rated the SVS/EVS transition that lasted 6 seconds and was finishes at 500 feet to be effective to highly effective.

Ten pilots commented. The pilot opinions varied with regard to the SVS/EVS transition altitude: 1) Three of the pilots thought the transition altitude should be higher than 500 ft. They suggested that 1000 ft AGL, the typical stabilized approach altitude, might be appropriate. Another pilot pointed out that removing the SVS image at 1000 ft might not be a good idea as the FLIR might not be painting a picture of the runway environment yet due to atmospheric conditions; 2) Two of the pilots suggested that the transition altitude should finish at a lower altitude than 500 feet (350 feet was suggested); and, 3) Three of the pilots liked the transition altitude strategy employed in the experiment. One pilot suggested removing the FLIR image between 200 ft and 100 ft AGL to force the pilot to acquire the runway environment with natural vision.

21. *If you like the gradual transition as a concept. What duration would you use (in seconds) and at what altitude (in ft AFL) would you initiate (or conclude) the transition?*

Duration (sec) _____ Transition point (feet AFL) _____ Other _____

22. *Is the height above touchdown an appropriate point for transitioning from SVS to EVS on the HUD?*

Here are the pilot suggested SVS/EVS transition duration and altitude:

- duration of 3 seconds with transition point of 1000 ft AFL
- duration of 2 seconds with transition point of 500 ft AFL
- duration of > 6 seconds with transition point of <200 ft AFL
- duration of 6 seconds with transition point of 350 ft AFL
- duration of 6 seconds with transition point of 500 ft AFL
- duration of 2 seconds with transition point of 1000 ft AFL
- duration of 6 seconds with transition point finishing before 500 ft AFL
- duration of 5 seconds with transition point of 700 ft AFL
- duration of 5 seconds with transition point starting at 500 ft and concluding at 400 ft AFL
- duration of 6 seconds with transition point of 500 ft AFL
- duration of 3 seconds with transition point of 500 ft AFL
- duration of 6 seconds with transition point of 500 ft AFL
- duration of 4 seconds with transition point of 400 or 300 ft AFL
- Fuse any time there is good correlation with transition point at 1000 feet.
- duration of 3 seconds with transition point of 500 ft AFL
- duration of 6 seconds with transition point of 500 ft AFL
- duration of 3 seconds with transition point of 500 ft AFL
- duration of 6 seconds with transition point of 500 ft AFL
- duration of 6 seconds with transition point of 500 ft AFL
- duration of 10 seconds with transition point of 500 ft AFL

The majority of pilots (11/20) wanted the duration time to be similar to that tested in the experiment (4-6 seconds); six of pilots wanted a shorter duration time (2-4 seconds) for the transition; two pilots wanted a greater duration time of greater than 6 seconds; and 1 pilot thought above 1000 ft there should be no transition as long as there was good correlation between the EVS and SVS. With regard to the transition altitude, again the majority of pilots (12/20) thought the tested transition altitude of 500 ft was

appropriate; three pilots wanted a transition altitude greater than 500 ft (values ranged from 700-1000 ft AFL; and three pilots wanted a transition altitude less than 500 ft (values ranged from <200 ft to 400 ft).

Auxiliary Display Fusion Control Questions

23. Describe your strategy for using the Alpha (spatial compatibility) and Beta (percent of EVS/SVS fusion) controls on the Auxiliary Display

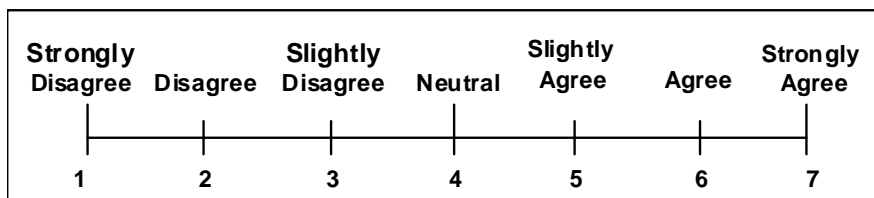
Seventeen out of 22 pilots commented. The majority of pilots (9 out of 17) wanted to have a fused image until 1000-500 ft AFL (range of values given by pilots) and then a pure FLIR image after this altitude to help obtain the approach lights and to make sure the runway was clear. 8 of the 9 pilots want max alpha above the FLIR only transition point. The other pilot (of the 9 that wanted fusion) wanted alpha fixed at 5 (mid range) always and beta to be 5 (mid range) for initial and intermediate approach segments and then on 8 (highest value) from 1000 ft AFL to 500 ft AFL after which he wanted pure FLIR from 500 ft AFL to landing. One pilot felt that FLIR under 100 ft AFL was key for runway incursion protection. 4 pilots wanted max alpha and max beta and one of the four commented that it helped with differentiating between objects and with depth perception. 3 of the pilots wanted mid-range values of alpha and beta.

24. Can you suggest any changes or improvements for controlling the EVS/SVS fusion on the Auxiliary Display?

Eleven out of 22 pilots commented. Four of them wanted the AD moved to the forward panel. In general, the pilots like the fusion control implementation. One pilot recommended a dual HUD configuration in the cockpit instead of an AD for the right seat but still liked the AD set-up and functionality as tested in this experiment.

Database Integrity Questions

25. Please indicate your level of agreement with the following statements:



a) There is enough information on the HUD to detect and correct for the gross database offset errors of 50 feet and 75 feet, respectively?

Display Concept	Mean Subject Ratings	
	50 feet	75 feet
HUD EVS without tunnel	5.0	5.6
HUD EVS with tunnel	5.0	5.6
HUD Fused EVS&SVS without tunnel	5.2	5.8
HUD Fused EVS&SVS with tunnel	5.3	5.9

AN ANOVA revealed no significant HUD concept effects ($p>0.05$) for the ability to detect and correct either a 50 ft or 75 ft database error. On average, pilots slightly agreed that they could detect and correct the 50 ft offset (mean=5.1) and agreed that they could detect and correct the 75 ft offset (mean=5.7) with the HUD concepts.

b) *There is enough information on the PNF Auxiliary Display to detect and correct for the gross database offset errors of 50 feet and 75 feet, respectively?*

AN ANOVA revealed no significant AD concept effects ($p>0.05$) for the ability to detect and correct either a 50 ft or 75 ft database error. On average, pilots slightly disagreed that they could detect and correct the 50 ft offset (mean=3.8) and were borderline in their agreement/disagreement that they could detect and correct the 75 ft offset (mean=4.1) with the AD concepts.

Display Concept	Mean Subject Ratings	
	50 feet	75 feet
AD with EVS only	3.6	3.8
AD with EVS and symbology	3.7	4.2
AD with Fused EVS/SVS only	3.8	4.1
AD with Fused EVS/SVS and symbology	4.0	4.3

26. *If there was a difference in ratings for 50 and 75 feet, please comment on the reason(s) why.*

Thirteen out of 22 pilots commented. In general, pilots stated that it was harder to detect the 50 foot offset as compared to the 75 foot offset. Reason for this difference was that the 50 foot error was “subtle” while the 75 foot error was an “attention grabber”. One pilot commented that he really didn’t see either error as his attention was transferred to the actual runway and he mentally queued out the error. Another pilot commented that as the PNF his attention was not focused on the AD but out-the-window monitoring the approach. The location of the AD was said to affect the pilot’s ability to detect and correct offset errors.

Crew Procedures

The crew procedures you used in the simulator are based on established EVS procedures used in today’s operating environment.

27. *Are the aural callouts clear and understandable as you perform the approach?*

Twenty-one out of 22 pilots commented. Twelve of the pilots thought the callouts were clear and understandable but some noted that these callouts were different than their airline’s standard operating procedures. Two pilots commented that the procedures were reversed from the challenge/response philosophy used in their current airline operations and another commented that he’d prefer Cat II/III monitored approach procedures. Others commented that the “minimums” call was counterintuitive because in today’s operations a “minimums” call requires the pilot to make a decision to land or go-around and not continuing with the runway environment in site as was the case in this test. Pilots suggested that an additional call be made to indicate EVS minimums (or 100 ft HAT) at been reached.

28. *Were the aural callouts acceptable or do have any recommendations for improvements?*

Pilots commented that the callouts were unacceptable because 1) crew callouts should not be stepping on each other and the computer callouts (e.g., have both 500 ft AGL computer and PNF callouts); 2) repetitive callouts should be removed; and, 3) calls should only be for critical actions. Pilots suggested that either a computer-callout or a PNF-callout should be added for EVS minimums of 100 ft HAT since this is where the critical decision to continue the landing or go-around is made by the PF. Pilots commented that it was confusing to have 2 decision heights when the DA was 200 ft. They suggested replacing the “minimums” call at 200 ft AGL with a “continue” or “continuing to 100 ft” PF call if lights are in sight.

Two alternatives for the crew procedures were scripted by evaluation pilots.

Script 1:

<u>PNF Challenge</u>	<u>PF Response</u>
"500 ft"	"Final Flaps 30, EVS Normal"
"Approaching Minimums"	(No Response)
"Minimums"	"Roger"
"100 ft"	"Landing (EVS lights)" or "Going Around"

Script 2:

<u>Auto Call</u>	<u>PF Call</u>	<u>PNF Call</u>
"Approaching Minimums" (@300 ft)	(No Response)	"300 ft"
"Minimums" (auto minimums call)	(No Response)	"200 ft"
"100 ft" (EVS lights)	"Lights/Land"	"Lights"
	"Missed Approach"	

29. *Are the crew procedures clear and understandable as you perform the approach?*

Thirteen out of 20 pilots found the procedures to be clear but three others thought revision was necessary. These pilots thought that the crew should use the “Challenge/Response” procedures and callouts as is done in normal instrument procedure operations. One suggested procedures script was:

<u>Auto Call</u>	<u>PNF Call</u>	<u>PF Call</u>
"Approaching Minimums" (+100 ft)	[Gives Situation Report]	(no response)
	"Approach Lights"	
	"No Approach Lights"	
	"Lights or Threshold Lights"	
	"Field in Sight"	
	"Runway in Sight"	
"Minimums"	[Gives Situation Report]	(no response)
	"No Lights – Go Around"	
	"Lights Continue"	

“100 ft”	[Gives Situation Report] “Lights” or “Land”	“100 ft” “Land” or “Go Around” based on PNF’s situation report
----------	---	---

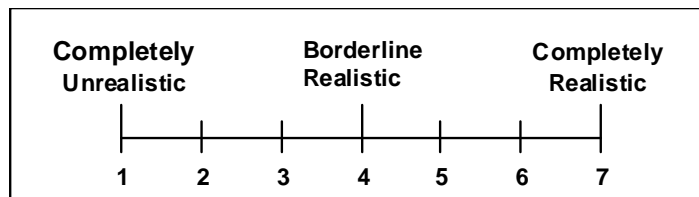
30. Were the crew procedures acceptable or do you have any recommendations for improvements?

Eight out of 20 pilots found the crew procedures to be acceptable as they were. The following recommendations to the crew procedures were offered by the pilots:

- 1) Change the “Runway in Sight” or “Landing” callouts to either one or the other since it’s a required call-out and not a pilot choice. [Current procedures allow pilot to say either callout.].
- 2) Either the PNF call real lights in sight so that the PF can switch off the HUD image to comply with the FAR requirement to actually see the lights using natural vision or there needs to be a better implementation of the control/contrast of the display so that you can see the lights with the naked eye.
- 3) If PF calls “EVS Lights” prior to decision altitude then the PNF should call “Alert” at 200 feet to remind the PF that he can legally go down to 100 feet HAT.
- 4) Use Challenge/Response procedures and callouts as is done in normal instrument procedure operations.
- 5) Amend the callouts so that they are linked to a decision altitude, or cued a response for PF.
- 6) Use a “monitored procedure” where the PNF would make the landing (take control at decision height) or call the go-around.
- 7) Need to adjust the wording of the minimums call at 200 feet and have the PNF make a call at 100 feet.
- 8) Crew should not be “stepping” on each other’s callouts and the computer callouts.
- 9) Repetitive callouts should be eliminated.
- 10) PF should say “continuing to 100 feet” upon reaching the decision altitude (if he called “EVS lights” earlier in the approach).
- 11) Have an automatic callout at 100 feet.
- 12) Have PNF call “lights” and PF call “land” or “missed approach”.

Simulator Fidelity Questions

31. On a scale of 1 to 7, please rate the fidelity of the simulation to real-world operations. Please consider whether the OTW realistically depicted actual traffic on taxiways and runways during CAT IIIa approaches.



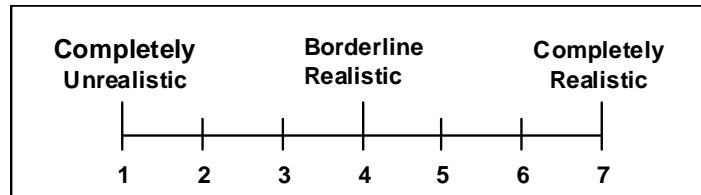
The mean simulation fidelity rating was 5.23 with a standard deviation of 1.1. On average, pilots rated the simulator fidelity to be realistic compared to real-world operations.

You gave a rating of _____. Please explain the reason for the rating. Do you think the aircraft/obstacles were unrealistically difficult to see or approximated the difficulty in real-world operations?

Pilot comments included:

- 1) Have been in sophisticated simulations before and this is close to those. The lighting from the aircraft to the ground would be a little brighter at the breakout point.
- 2) There is not enough sample population to make a realistic assessment. You would need to view the same obstacle/traffic several times in the same place with varying visibility conditions to better assess viability of EVS for acquisition response behavior.
- 3) Obstacles were reasonably accurate.
- 4) Too much computational time as symbology floats. Yoke has significant amount of free play indicating input to control time constants are too high.
- 5) Obstacles would have more contrast visually (day/and on FLIR)
- 6) Not too far off from training simulator
- 7) Weather seemed realistic though at some of the low RVRs the visibility seemed better than actual.
- 8) Very clear, almost lifelike – great fidelity.
- 9) Obstacles approximated real-world operations.
- 10) Close to real-world for low visibility operations.

32. If you have FLIR operational experience, on a scale of 1 to 7, please rate the fidelity of the FLIR simulation to actual FLIR imagery. Please consider whether the FLIR realistically depicted terrain, runways, traffic, obstacles, etc. during instrument approaches.



The mean FLIR simulation fidelity rating was 6.13 with a standard deviation of 0.86. On average, pilots rated the FLIR fidelity in the simulator to be very realistic as compared to actual FLIR imagery.

You gave a rating of _____. Please explain the reason for the rating.

Pilots provided the following comments:

- 1) Artificial – fidelity of lights drives the landing decision.
- 2) FLIR seems completely realistic
- 3) Realistic for only having hot on cold display. Having hot/cold selectable would have improved fidelity.
- 4) All but obstacles on the runway which need to have some thermal contrast.
- 5) Appeared very realistic except there was not an IR signature for the baggage cart or fire truck, which it would have.
- 6) FLIR display was realistic particularly in the final approach of flight.

General Comments

Please provide any further comments or suggestions, which may help to improve these synthetic vision/enhanced vision concepts.

The following comments were offered by the pilots:

- 1) Put PNF display in front of the PNF above the HSI.
- 2) Give the pilots the capability to select levels of declutter.
- 3) PF and PNF should be looking at same display depiction (e.g., tunnels, guidance cue, SVS and EVS pictures) so if there is a question between the two they are referencing the very same scope objects.
- 4) At around 500 feet, the auxiliary display is almost taken completely out of the picture as the PNF is searching out the front of the airplane for the actual runway and preparing for the go-around procedure if necessary.
- 5) Overall good concept and integration for PD. Ideally, PF and PNF would each have SVS/EVS/Symbology/Tunnel – with ability for PF and PNF to deselect any one feature.
- 6) Auxiliary display is only good for snapshot situation awareness for PNF. PNF primary job is to monitor instrument displays for function/malfunction and monitor approach progress (altitudes) for callouts. Going head-down to look at auxiliary display requires head movement and eyeball focal length shift.
- 7) Crew procedures and callouts have significant impact and need to be changed.
- 8) Most critical area in need of improvement is the transition from the full EVS to visual runway acquisition by 100 feet HAT.
- 9) Provide the capability to dim the intensity and brightness of the HUD symbology. Often it was so bright it would mask the EVS lights and/or the runway threshold lights.
- 10) Like the tunnel but it is not needed after glideslope/localizer capture.
- 11) Felt like could use this system in real-world operations with less than one day of training.
- 12) In terms of actual weather that the pilot flies in during transport approach operations, he estimates that would use the system 5% for weather related landings and 50% for night landings.
- 13) We need full symbology, SVS and EVS HUDs for Part 121 operations.
- 14) Effectiveness would have been much higher with a null error pitch and roll guidance (similar to flight director system on Boeing 777) rather than the fly-to command for the flight path marker.
- 15) Generally, pilots flying a path-directed approach, missed approach, arrival, or departure will presume that if they remain within required parameters of the path, that terrain and obstacles are deconflicted. Therefore, the SVS was of value only as confirmation that the directed path was accurate.
- 16) Important flight information along with terrain information on one display is a very useful tool as it greatly reduces pilot workload of viewing and processing information but overload of information such as with a tunnel display is distracting.
- 17) Having to look around or through the HUD (after declutter) at 100 feet presents problems for the PF – better to have the PNF call landing or go around.
- 18) SVS is redundant for the PNF if PF has it and could trap the crew into complacency and breaking FARs for raw visual requirements.
- 19) Symbology on AD is a must – include vertical speed information if possible.
- 20) Two HUDs are better than one.
- 21) Tunnel and guidance cue a must.
- 22) Provision for callouts using programmable radar altitude and barometric altitudes will be needed for actual operations.
- 23) No SVS for PF approaching 200 feet and no SVS or EVS for PF approaching 100 feet, PNF clears runway with EVS less than 100 feet, and PF has full EVS/SVS for flare/landing/rollout.

- 24) Even though had no previous FLIR or tunnel experience, once I grasped the concept of the tunnel I was a true convert - having SVS/FLIR and tunnel significantly reduced workload and increased scan level and situational awareness.
- 25) Minimums call at 200 feet and no call at 100 ft by PNF is not the best approach.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 01-04-2009		2. REPORT TYPE Technical Publication		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Evaluation of Fused Synthetic and Enhanced Vision Display Concepts for Low-Visibility Approach and Landing			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Bailey, Randall. E.; Kramer, L:ynda J.; Prinzel, Lawrence J., III; Wilz, Susan J.			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER 609866.02.07.07.02		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NASA Langley Research Center Hampton, VA 23681-2199			8. PERFORMING ORGANIZATION REPORT NUMBER L-19504		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001			10. SPONSOR/MONITOR'S ACRONYM(S) NASA		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) NASA/TP-2009-215707		
12. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category 03 Availability: NASA CASI (443) 757-5802					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT NASA is developing revolutionary crew-vehicle interface technologies that strive to proactively overcome aircraft safety barriers that would otherwise constrain the full realization of the next generation air transportation system. A piloted simulation experiment was conducted to evaluate the complementary use of Synthetic and Enhanced Vision technologies. Specific focus was placed on new techniques for integration and/or fusion of Enhanced and Synthetic Vision and its impact within a two-crew flight deck during low-visibility approach and landing operations. Overall, the experimental data showed that significant improvements in situation awareness, without concomitant increases in workload and display clutter, could be provided by the integration and/or fusion of synthetic and enhanced vision technologies for the pilot-flying and the pilot-not-flying. Improvements in lateral path control performance were realized when the Head-Up Display concepts included a tunnel, independent of the imagery (enhanced vision or fusion of enhanced and synthetic vision) presented with it. During non-normal operations, the ability of the crew to handle substantial navigational errors and runway incursions were neither improved nor adversely impacted by the display concepts. The addition of Enhanced Vision may not, of itself, provide an improvement in runway incursion detection without being specifically tailored for this application.					
15. SUBJECT TERMS Aviation Safety; Crew Resource Management; Enhanced Vision System; FLIR; Head-Up Display; Image Fusion; Runway Incursions; Synthetic Vision System					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			STI Help Desk (email: help@sti.nasa.gov)
U	U	U	UU	84	19b. TELEPHONE NUMBER (Include area code) (443) 757-5802