

Evaluation of Head-Worn Display Concepts for Commercial Aircraft Taxi Operations

Randall E. Bailey,* J.J. (Trey) Arthur III, Lance Prinzel III, and Lynda J. Kramer
NASA Langley Research Center, Hampton, VA

ABSTRACT

Previous research has demonstrated that a Head-Up Display (HUD) can be used to enable more capacity and safer aircraft surface operations. This previous research also noted that the HUD exhibited two major limitations which hindered the full potential of the display concept: 1) the monochrome HUD format; and, 2) a limited, fixed field of regard. Full-color Head Worn Displays (HWDs) with very small sizes and weights are emerging to the extent that this technology may be practical for commercial and business aircraft operations. By coupling the HWD with a head tracker, full-color, out-the-window display concepts with an unlimited field-of-regard may be realized to improve efficiency and safety in surface operations. A ground simulation experiment was conducted at NASA Langley to evaluate the efficacy of head-worn display applications which may directly address the limitations of the HUD while retaining all of its advantages in surface operations. The simulation experiment used airline crews to evaluate various displays (HUD, HWD) and display concepts in an operationally realistic environment by using a Chicago, O'Hare airport database. The results pertaining to the implications of HWDs for commercial business and transport aviation applications are presented herein. Overall HWD system latency was measured and found to be acceptable, but not necessarily optimal. A few occurrences of simulator sickness were noted while wearing the HWD, but overall there appears to be commercial pilot acceptability and usability to the concept. Many issues were identified which need to be addressed in future research including continued reduction in user encumbrance due to the HWD, and improvement in image alignment, accuracy, and boresighting.

Keywords: Head-Worn Display, Head-Up Display, Surface Operations, Latency, Simulator Sickness

1. INTRODUCTION

The Integrated Intelligent Flight Deck Technologies (IIFDT) project, under NASA's Aviation Safety Program (AvSAFE), comprises a multi-disciplinary research effort to develop flight deck technologies that mitigate operator-, automation-, and environment-induced hazards. Towards this objective, IIFDT is developing crew/vehicle interface technologies that reduce the propensity for pilot error, minimize the risks associated with pilot error, and proactively overcome aircraft safety barriers that would otherwise constrain the full realization of the next generation air transportation system (NGATS). Part of this research effort involves the use of synthetic and enhanced vision systems and other interface modalities as enabling technologies to meet these safety challenges.

To enable the NGATS, technologies and operational concepts must be developed and implemented to provide a safe, efficient and reliable air transportation system that removes many of the constraints in our current system, supporting a wider range of operations, and thus, delivering an overall system capacity up to three times that of current operating levels.¹

Many operational concepts are emerging to improve airport terminal area operations for NGATS; one of the key elements is to enable precision navigation, including 4-dimensional trajectory (space plus time) and Visual Flight Rules-like capabilities and safety in all weather conditions. A so-called Equivalent Visual Operations (EVO) capability is hypothesized as one potential operating concept whereby the safety and operational tempos of current-day Visual Flight Rules (VFR) might be obtained and, perhaps, VFR operational procedures might be employed independent of the actual weather conditions. The NASA

* Randall.E.Bailey@nasa.gov, NASA LaRC, 24 W Taylor Street, Hampton, VA 23681, phone 757 864 8682

Crew-Vehicle Interface element of the IIFD project is focusing on flight-deck-centric, flight safety issues and technologies necessary to eliminate barriers to these NGATS safety and capacity goals.

2. BACKGROUND

Previous research has shown that, while the capability may be available to take-off and land aircraft in near zero visibility and zero ceiling weather, the operational tempo and safety within the airport terminal area is significantly degraded due to limitations in surface operations – that is, taxiing and maneuvering aircraft and vehicles to/from the active runways and gates.

Over the last 5 years, NASA and its industry partners have developed and deployed Synthetic Vision (SV) technologies for commercial and business aircraft which have been shown to provide significant improvements in terrain awareness and reductions in the potential for Controlled-Flight-Into-Terrain incidents/accidents compared to current generation cockpit technologies.^{2,3} Synthetic vision is a computer-generated image of the external scene topography that is generated from aircraft attitude, high-precision navigation, and data of the terrain, obstacles, cultural features, and other required flight information.

In comparison, Enhanced Vision (EV) (or Enhanced Flight Vision System) is an electronic means to provide a display of the external scene by use of an imaging sensor, such as a Forward-Looking InfraRed or millimeter wave radar. Both SV and EV are “vision-based” technologies intended to create, supplement, or enhance the natural vision of the pilot.

The optimal fusion of Synthetic and Enhanced Vision Systems (S/EVS) technology is emerging as a cornerstone to the development of advanced flight deck information systems which can provide the flight crew with significantly improved spatial awareness, increased awareness of outside terrain and obstacle features, enhanced manual flight performance, and reduced pilot workload. S/EVS technology can potentially provide unlimited field-of-regard awareness for terrain, obstacles, traffic, and airspace constraints and establish one cornerstone to an EVO capability. A key component to these revolutionary benefits may be advanced display media which can create spatially-integrated, large field-of-view, “unlimited” field-of-regard display of this information for the pilot (Figure 1).

Helmet-mounted display (HMDs) or, more appropriately for the commercial market, head-worn displays (HWDs) are one display media which may meet these information display requirements. HMDs and HWDs are not new technology, particularly for military operations, but component miniaturization and maturation are progressing to the point where HWDs can be considered in commercial and business aircraft operations – i.e., the costs are reaching affordable levels and their use should be as unobtrusive to the pilot as wearing sunglasses.

NASA recently conducted a preliminary evaluation of a HWD as a viable technology to support an EVO concept.⁴ The HWD provided a full color, 40° field of view, fully-overlap binocular display and was coupled with a head-tracker. Approach and landing operations, commensurate with commercial airline operations and procedures were evaluated. Technical difficulties with the HWD negatively influenced the results, but nonetheless, these data showed the promise of the technology, the pitfalls for commercial applications, and the potential that new, emerging technologies may provide. Subsequent to this work, a trade study was conducted by NASA to assess the emergence of small, very lightweight HWDs and to begin the build-up of system level requirements to meet commercial and business aircraft applications. This work and others are being used by NASA to evaluate the potential of HWDs for emerging NGATS operating concepts, such as EVO.

The previous work in leveraging HWD technology as an enabler of EVO focused on approach and landing operations. While this work is relevant and more studies will continue, the worst civilian aircraft accident in history occurred during the collision of two B-747 aircraft at Tenerife during runway/surface operations in fog. Further, the National Transportation Safety Board (NTSB) has included runway incursion on its “most wanted list” of transportation safety improvements since the list began in 1990.⁵



Figure 1: Notional EVO Flight Deck

Previous research^{6,7} to develop technologies to improve aircraft surface operations efficiency and safety led to the development of the Taxiway-Navigation and Situation Awareness (T-NASA) system and the Runway Incursion Prevention System (RIPS). The T-NASA system showed the importance of taxiway awareness, route awareness, and taxi guidance on Head-Up and Head-Down Displays to enable taxi times, approaching those achievable in VFR operations, and to prevent pilot error in surface operations, including missed taxi clearances, missed understanding of taxi instructions, making wrong turns, and becoming lost on the airport surface. In particular, the T-NASA work demonstrated that a Head-Up Display (HUD) enables more efficient and safer surface operations. The Head-Up presentation of information was critical to enable the pilots to remain head-out – improving the cross-check of information and the outside world, and ensuring that the aircraft remained on the taxiways.

However, this research also noted that two of the major HUD limitations during ground operations were their monochrome form and limited, fixed field-of-regard. A monochromatic display precludes the use of color for information decluttering and information cuing so the display area of a HUD must be carefully designed to provide the pilot with enough information without saturating it with clutter. Further, HUD imagery, while allowing head-out operations, is restricted to its forward, fixed field-of-view and the use of conformal imagery is, consequently, limited.

These HUD limitations, coupled with the importance of keeping the aircrew “head-out”, naturally points to an ideal application for HWDs.⁸ In fact, Rediess⁹ proposed such a concept and developed an architecture under a Small Business Innovative Research (SBIR) contract. Similar concepts have been proposed for Air Traffic Controllers.¹⁰ The work by Rediess, however, deviated from the present work and did not address some very key elements:

- The encumbrance of the HWD to the commercial and business aircrew is paramount to success. Without unobtrusive technology, the application is likely unacceptable for commercial aviation use.
- Technology in Synthetic and Enhanced Vision has advanced to the point where these systems are not only realizable, but extremely capable and beneficial.
- Latency is critical.¹¹ Any potential that latency may induce visual illusion or spatial disorientation that will, again, doom the application.

Significant work by Rediess involved datalink technology for communications between ground control and the aircrew. While datalink technology is clearly the path to the future, the importance of aural communications and its criticality in today’s operations cannot be dismissed. This communication modality should be utilized even in the future since the visual modality can be overloaded and voice communications may provide an avenue for error detection and enhanced situation awareness. NASA is

investigating methodologies and technologies which capitalize on data-link communications, yet ensure that modality changes retain the good features of previous auditory channels (e.g., retaining “party-line” situation awareness), maximize the benefit of new technologies (e.g., improved non-native English language communications, automatic route entry into a Flight Management System) and create new and improved functionalities (e.g., a “Culture-Neutral/Language-Neutral” flight deck).¹² NASA has published a Technology Invention disclosure¹³ for a crew-vehicle interface system concept that couples a HWD and speech interface system for greatly improved surface operations, particularly in single-crew operations.

3. EXPERIMENTAL EVALUATION

Two ground-based simulation experiments were conducted to evaluate if a head-worn display could provide benefits in performance and safety over a HUD during aircraft surface operations at Chicago, O’Hare International Airport (FAA identifier: ORD). (More extensive details for these studies are contained in a companion paper, presented in the paper by Arthur, et al within the Enhanced and Synthetic Vision Conference.¹⁴ This paper primarily details the first experiment – “Experiment One” – with respect to the HWD design and evaluation issues from this study.) The previous research conducted by NASA and others during the RIPS and T-NASA system developments were mirrored accordingly. Some repetition of the display concepts and scenarios evaluated whether the experimental device or protocol biased the results and comparisons. The HWD was hypothesized as being able to overcome two major HUD limitations shown in previous research, while retaining the HUD’s respective advantages: 1) the monochrome form; and, 2) a limited, fixed field of regard.

Both studies were conducted in the Research Flight Deck (RFD) simulator at the NASA Langley Research Center (LaRC). The RFD is a dual-pilot simulator consisting of a collimated 200-degree out-the-window visual. The RFD is equipped with a 30° H x 24° V HUD on the captain’s side. The HWD was an 800H x 600V pixel, full color display with see-through capability, 60 Hz refresh and a pilot-selectable brightness knob. The HWD field of view was approximately 23° horizontal by 16.5° vertical. The subject pilot (Captain) placed the HWD over the right eye so that it was visible by glancing up and therefore, the pilot maintained unimpeded stereoscopic vision for out-the-window monitoring. An optical head tracker provided the head orientation data. The RFD has 8 Size D (6.4 inch square viewable area) head-down displays: Captain and First-Officer primary flight displays (PFD) and navigation displays (ND), two engine displays on the center aisle and two outboard auxiliary displays. For both experiments, the first officer’s outboard auxiliary display was used a repeater display of the captain’s head-up display. The pilot controls were a tiller, throttles, rudder pedals (nose wheel steering) and differential toe brakes. The simulated aircraft for both experiments was a Boeing 757.

The results for the first of two experiments are synopsized in this paper with respect to the HWD design and evaluation issues. The experiments evaluated four different display concepts (Fig. 2):

- 1) Baseline: A head-down electronic moving map (EMM) without routing or traffic information. There was no head-up display with this concept.
- 2) Intermediate HWD: A HWD concept with a head tracker that displayed a virtual airport environment but no traffic, routing or clearance information. This concept included the baseline EMM head-down display,
- 3) Advanced HUD: A HUD concept with an advanced EMM head-down display. The HUD symbology consisted of a 3-dimensional depiction of the cleared route by highlighting the taxiway edge lines and centerlines. The advanced EMM contained iconic traffic, clearance and routing information; and,
- 4) Advanced HWD: An advanced HWD concept with virtual traffic and routing information and an advanced EMM head-down display.

All display concepts had a head-down EMM display. The baseline EMM consisted of a perspective, track-up view of the airport showing an ownship symbol, ground speed, heading, surface movement areas,

centerlines, airport surface labels, and current range selection. The advanced EMM contained the same information as the baseline EMM with the addition of a route display, clearance information, distance to the next taxiway, and traffic icons. Both the captain and first officer had an EMM display with independent range controls (4 zoom levels). The EMM was shown on the ND.

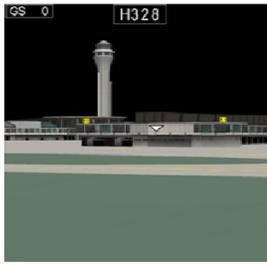
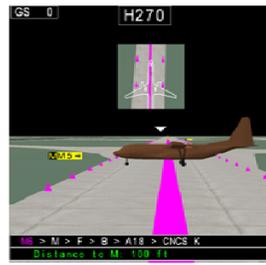
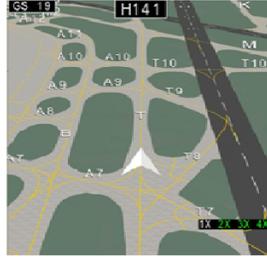
Baseline	Intermediate HWD	Advanced HUD	Advanced HWD
 <p data-bbox="315 726 451 747">No head up display</p>	 <p data-bbox="613 726 750 768">Baseline HWD (no traffic or route)</p>	 <p data-bbox="932 726 977 747">HUD</p>	 <p data-bbox="1192 726 1308 747">Advanced HWD</p>
 <p data-bbox="315 1050 451 1092">Baseline EMM (no traffic or route)</p>	 <p data-bbox="613 1050 750 1092">Baseline EMM (no traffic or route)</p>	 <p data-bbox="886 1050 1003 1071">Advanced EMM</p>	 <p data-bbox="1192 1050 1308 1071">Advanced EMM</p>

Figure 2. Display Concepts: Baseline, Intermediate HWD, Advanced HUD and Advanced HWD.

The intermediate display concept consisted of the baseline EMM head-down display and the HWD with head tracker. The HWD was a virtual airport view from the pilot's eye perspective. The virtual airport consisted of the ORD airport, buildings, surface movement areas and centerlines. Taxi signage was displayed in the HWD. This signage was modeled to appear to be actual airport surface signage; however, the HWD signage was placed on the side of an upcoming turn and did not necessarily correlate with the actual sign placement out the window. The HWD displayed the ground speed, heading and an aircraft-heading pointer. The aircraft-heading pointer was used to aid the pilot in determining the aircraft heading during head movement.

The Advanced HUD display concept was based on T-NASA and RIPS concepts. The HUD format is depicted in Figure 3. The display showed current ground speed in digital format, the current taxiway (shown as M6 in Fig. 3), next cleared taxiway (shown as M in upper right of Fig. 3), centerline markers and virtual cones on the taxiway edge. Additional cues were given for turns. These cues consisted of turn flags and virtual turn signs (similar to road way turn signs). Hold short lines were displayed with a single drawn line at the hold short location. Also, a virtual stop sign was placed in the middle of the hold short line. A non-conformal taxi director display provided an intuitive display of the relationship between the taxiway centerline and the aircraft's landing gear. The captain could remove all the symbols from the HUD display by pressing the auto throttle disconnect button. A second press of the auto throttle disconnect would restore all of the HUD symbology. The captain also had control of the brightness level of the HUD.

The advanced HWD concept contained all of the information in the intermediate concept with the addition of traffic and routing information. The advanced HWD employed a 3-dimensional generic aircraft model to depict traffic, the cleared route was shown as a magenta overlay on the taxiway centerline, text was

written for the cleared route and for the distance to the next taxiway, and virtual taxiway edge cones depicted the cleared route (Fig. 4). Like the HUD, virtual turn signs were used as an additional turn cue and hold short lines were denoted by virtual stop signs. Similar to the HUD, a non-conformal insert display showed the airplane outline, landing gear, and the airport as if viewed from directly above (Fig. 4). HWD declutter controls were provided by re-using the auto-throttle disconnect button



Figure 3. HUD Depiction showing centerline and edgeline cones of the cleared route. Bottom box shows the taxi director non-conformal display.

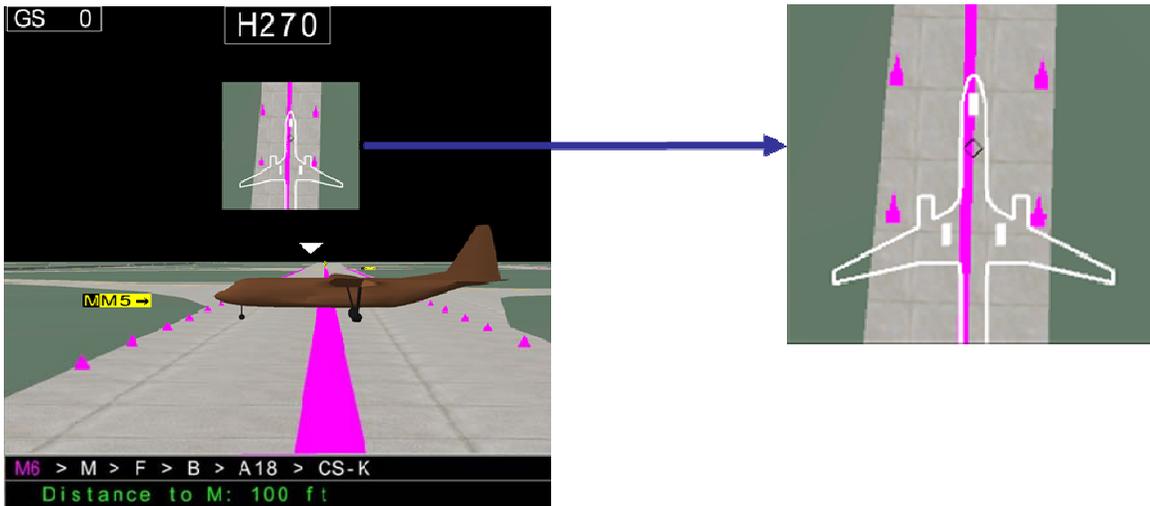


Figure 4. The non-conformal insert display concept including main gear, nose gear and aircraft outline.

Evaluations Pilots

Commercial flight crews (a captain and first officer) participated in the experiment. Each flight crew flew for the same company to ensure crew coordination and cohesion with regard to surface operation procedures. The captains had an average of over 16,000 flight hours with 22.3 years of commercial flying and the first officers had an average of over 11,000 flight hours with 13.5 years of commercial flying. Forty-four percent of subject captains required corrective lenses. The subjects were given a 45-minute briefing to explain the display concepts and the evaluation tasks. After the briefing, a 1 hour training session was conducted to familiarize the subjects with the RFD simulator, HUD, the HWD device, and the piloting task. Only the captain had a HUD; the first officer had a head-down repeater display of the captain's head-up device. An eye dominance test was performed after the training briefing. All subject

captains were right eye dominant. The HWD was viewed with the right eye. Following training, 5 hours of data collection was conducted. The total experiment time for a subject crew was approximately 8 hours.

Evaluation Task

Pilots conducted taxi operations at Chicago O'Hare International Airport. The display concept and weather were varied. A total of 27 different taxi scenarios were used in the study. All taxiing tasks were exiting the runway and taxi to the airport movement area. The weather state for the out-the-window scene was varied between night time with unlimited visibility (visual meteorological condition; VMC), and 700 foot runway visibility range (RVR). The subject crews were briefed to follow their company guidelines as far as taxi speeds and procedures. Further, crews were instructed that the safety of the aircraft should never be compromised.

Before each data trial, the pilots were briefed on their current location and expected turnoff. Each trial began with an initial speed of 10 knots followed by an immediate call from the local controller. Once cleared of the runway, the first officer switched to ground frequency and called the ground controller for clearance. At this point, the ground controller provided the ground clearance along with a data linked message of the route. During the data run, ATC calls to other aircraft were played to simulate typical radio party line chatter. In addition, other aircraft traffic were taxiing around the airport surface. Crews were instructed that the traffic was pre-recorded, thus they should give way to all traffic. Further, they were briefed that the ground controller would provide no cues for traffic awareness.

4. COMMERCIAL ACCEPTANCE OF HWD

The results of this experiment are interpreted with respect to the acceptance of a HWD for commercial and business aircraft operations, focusing on taxi operations. "Acceptance" depends upon numerous factors, in this case, the focus will be on the following characteristics which are believed to be key ingredients: a) Situation awareness and workload associated with the HWD; b) HWD latency; c) Simulator sickness; and, d) Usability and preference.

4.1 Situation Awareness and Workload

The flight crew gave ratings using the NASA Task Load Index (NASA-TLX) scale¹⁵ of their mental workload after each run. An analysis of variance test was performed on these data and no significant differences were found ($p < 0.05$). The pilot workload using the baseline EMM, intermediate HWD, Advanced HUD, and Advanced HWD were indistinguishable. The workload data suggests that the HWD displays did not increase, nor reduce, mental workload demands compared to the T-NASA concepts.

The flight crew gave ratings using the Situation Awareness Rating Technique (SART)¹⁶ after each run. For SART, situation awareness is defined as Understanding plus the difference between attentional supply minus the attentional demand). Analysis found a significant effect for display condition ($F(3, 15) = 4.16, p < 0.05$). A Student Newman-Keuls test revealed two unique subsets: (1) Highest SA for the Advanced HWD (133.04) and Advanced HUD (130.23); and, (2) Lowest SA for the Baseline (104.07) and Intermediate HWD (112.30). The SA provided by the Advanced HWD and Advanced HUD were not significantly different from each other nor were the Baseline and Intermediate HWD significantly different from each other. The results suggest that the cleared route information was likely the biggest factor in SA for the flight crews in this task. (The Baseline and Intermediate HWD did not include this information whereas the Advanced HUD and HWD did.) No significant differences were found between Advanced HUD and Advanced HWD suggesting that these two displays provide for equal amounts of perceived situation awareness.

Flight crews were administered a Likert post-run experimental questionnaire (1 to 5 scale; 1 = "not at all"; 5 = "very much") to rate the display conditions contribution to taxi efficiency, overall navigation awareness, route awareness of local controller clearance, route awareness of ground controller clearance, surface traffic awareness, directional awareness, and taxi safety. An ANOVA revealed significant effects for all dependent variables, $p < 0.05$. Post-hoc Student Newman-Keuls tests were performed on these

dependent variables resulting in two unique subsets: (1) no significant differences between Advanced HUD and Advanced HWD, and (2) no significant differences between Intermediate HWD and Baseline. Means for each dependent variable are presented in Table 1.

Table 1. Means for Post-Run Questionnaire By Display Condition

	Taxi Efficiency	Navigation Awareness	Route Awareness (local)	Route Awareness (Ground)	Traffic Awareness	Direction Awareness	Taxi Safety
Baseline	3.56	3.62	3.62	3.47	2.62	3.17	3.49
Intermediate HWD	3.63	3.67	3.72	3.59	2.81	3.26	3.58
Advanced HWD	4.49	4.46	4.35	4.42	3.9	3.76	4.42
Advanced HUD	4.56	4.48	4.37	4.43	4.04	3.82	4.44

The grouping in these data between the Advanced HUD and Advanced HWD suggest that the information shown on these devices, coupled with their head-up characteristic, significantly influence SA and subjective awareness measures, the key information being the cleared taxi route. A HWD concept without cleared route information – even though it is full field-of-regard – does not provide improvement over a fixed field-of-view HUD with cleared route information. When cleared route information was provided on the HWD, it was rated as being no better or no worse than the HUD in these measures. (While not shown here, the Advanced HWD *did* show significant improvements in surface traffic awareness, as measured under “rare event” conditions, over the Advanced HUD.¹⁴)

The Advanced HWD might also provide economic benefits over the HUD such as potential weight savings that would have significant cost advantages to commercial airlines.

4.2 Latency

System time delays or latencies inherent to the head-worn (and helmet-mounted) displays have been shown to critically influence performance, utility, usability, and acceptability. Consistent definitions and measurement techniques as well as preliminary latency (time delay) requirements have been proposed for Head-Worn Display S/EVS applications.^{11,17} Based upon the most stringent requirements for HMD applications of S/EVS (i.e., demanding tasks using a high resolution, large field-of-view head-worn display), the required system latency might be as low as 20 msec.

An in-situ latency measurement technique – the so-called “windshield washer” test¹¹ – was used to measure the HWD latency. This test used the conformal runway scene (i.e., space-stabilized, boresighted symbology) and a grid pattern of 5 degree lines projected on the out-the-window simulator visual. The grid pattern is used for simulator visual scene alignment and calibration. In the presence of HMD latency, the boresight symbol cannot remain perfectly space-stabilized.

The test required that the user smoothly oscillate his head in azimuth at a rate which caused the runway centerline symbol to touch the outer, target box. The aircraft was headed directly down the runway and positioned on the runway centerline. The head-movement rate data divided by the size of the target box defines the equivalent time delay (at one frequency). (Note that the latency in the head movement data is immaterial to this computation; only the average rate is needed for the equivalent delay calculation.)

The head-tracker allows “prediction” or lead compensation to be applied to the tracker data, attempting to minimize the system latency. The prediction method is not described by the head-tracker manufacturer. The latency data was taken at three prediction values: a) none (0 msec); b) 25 msec prediction; and, c) 50 msec prediction. The experiment was run using the 50 msec prediction value. The selection of a prediction

value is a compromise between system latency (i.e., time delay) and noise or jitter in the HWD-displayed scene. At the 50 msec prediction value, the latency was minimized while the jitter and noise of the head-tracker output was judged during the “pre-test” trials to be “reasonable.”

In Figure 5, the measured latency values are shown. Using the “windshield wiper” test, the measured latency values are unequivocally “end-to-end.” It represents the total lag between the pilot’s head movement until the conformal imagery is finally drawn. This latency includes, not just the head-tracker, but also the communications delay between the head-tracker and the SV image generator (RS-232) and the SV image generator computational time, to name just two elements. The test was conducted three times to measure repeatability and it also used a 5 degree target spacing and 2.5 degree spacing.

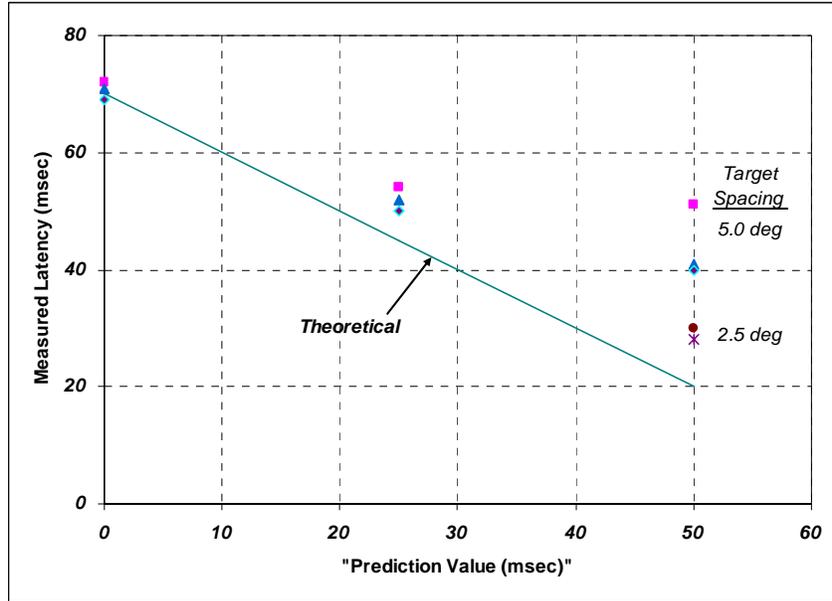


Figure 5: Measured End-to-End Latency

Without “prediction” from the head-tracker, the end-to-end latency is approximately 70 msec. Less than 2 msec variation was measured between the 3 samples. With 25 msec prediction, the measured latency was approximately 52 msec with very little variation in measurements. With 50 msec prediction, the measured latency varied considerably depending upon the target spacing used. The measured latency was 45 msec for the 5 degree spacing and only 30 msec for the 2.5 degree spacing. This plot also includes the theoretical latency from the 70 msec nominal value if the prediction method was “perfect.” The data do not match these “perfect” predictions.

The latency results do, however, match expectations because of the nature of the test. Using the “windshield wiper” technique, the frequency of the head movement (i.e., the input) to measure latency depends on the target spacing and the inherent system latency. For instance:

- For the no-prediction test, relatively slow head-movements were needed to achieve a 5 degree target spread; thus, it was easy to get repeatable latency data. Also, without prediction, while there is a lag in the display due to the inherent latency, there isn’t amplitude modification between the head-movement and the resultant conformal image. A 5 degree head-movement will result in a 5 degree image displacement.
- When prediction is used, the predictor modifies the system phase response (i.e., it provides compensation for the lag) but it also introduces amplitude modification. While the exact prediction method for the head-tracker in this test isn’t known, the frequency response of a “typical” prediction method¹⁸ is shown in Figure 6 to illustrate the effect. The predictor attempts

to keep the phase response close to 0 degrees phase across a large frequency range; thus, minimizing the lag (delay) between the head-movement and resultant image response. However, the amplitude response of the predictor is no longer 0 dB (amplitude ratio of 1) so a 5 degree head movement will result in a image response of greater than 5 degrees for frequencies where the amplitude response is greater than 0 dB.

- For a typical predictor, the higher input frequency yields a greater resultant amplification. Since the “windshield wiper” test relies on the amplitude response of the image, a non-unity frequency response from a predictor will distort the delay (phase) measurement.
- This amplitude distortion is particularly evident in the “50 msec prediction” data in Figure 5. For the 2.5 degree target, lower system latency was measured, but for the 5.0 degree target, more latency was measured. This difference is due to the greater amplitude magnification with the higher input frequency needed for the 5.0 degree target. It is also due to the predictor having less phase compensation at the higher frequencies.
- The higher propensity for noise and jitter in the image with prediction is also due to this amplitude distortion in the predictor.

A single latency value cannot be associated with the system used in this experiment due to the amplitude and phase characteristics with a predictor in the loop, but the data shows that the tested latency was between 30 and 50 msec.

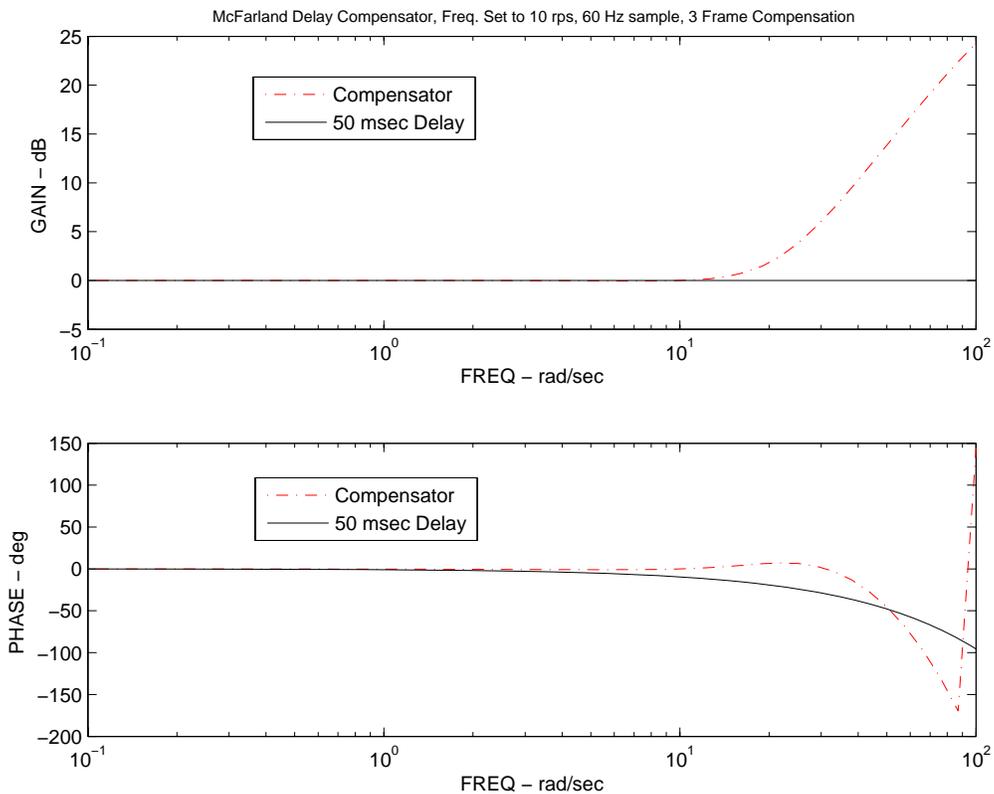


Figure 6: Typical Delay Compensator Frequency Response

4.3 Head Tracking

A comparative evaluation of the pilot's head movement with and without an HWD was not performed in the experiment. The experiment was designed to be a "pure" comparison of head-down, head-up, and head-worn displays so no head-tracking apparatus was worn by the pilot during the evaluation of head-up and head-down display concepts. Non-contact head-tracking methods were neither economically nor practically viable for implementation in this experiment. Subsequent experiments will address whether a HWD affects a pilot's head-movement compared to present-day operations.

The Captain's head movement when wearing the HWD was analyzed. Collapsing the data across all subjects and all maneuvers, the percent of time the Captain was looking in azimuth and elevation is shown in Figures 7 and 8, respectively. Both figures show a high concentration of time within ± 45 degrees of the aircraft centerline, particularly for elevation. 98% of the time, the pilot's head was position within ± 25 degrees of elevation. 97% of the time, the pilot's head was position within ± 50 degrees of azimuth. The azimuth data shows a slightly longer "tail" reflecting the need for the crew to clear a target and maneuver at angles exceeding 45 degrees off-boresight, but the data is clearly different than that shown for a typical fighter aircraft head-worn display application. As could be expected, the need for off-boresight capability is for commercial applications.

The rates of head movement when wearing the HWD were analyzed. Collapsing the data across all subjects and all maneuvers, the percent of time the Captain used head movement rates in azimuth and elevation is shown in Figures 9 and 10, respectively. Both data show a high concentration of time within ± 45 degrees/sec, particularly for elevation. 97% of the time, the rate of head movement was within ± 30 degrees per second in elevation. 97% of the time, the rate of head movement was within ± 60 degrees per second in azimuth. Note that the percentage of time when rates were greater than 100 deg/sec or less than -100 deg/sec are plotted in Figure 9 and 10, at the ± 100 deg/sec point (for readability). The azimuth rate data show that the pilots typically used relatively slow head-movements to clear for traffic in maneuvers but about 3% of the time, the pilots exceeded 100 deg/sec head-movement. Maximum head rates in azimuth of 200 to 300 deg/sec were found. These data indicate much slower head movement for commercial rather than the military applications, but a small, yet significant percentage of time still requires very fast head-rates and tracking capability. This same characteristic was not found in the elevation data.

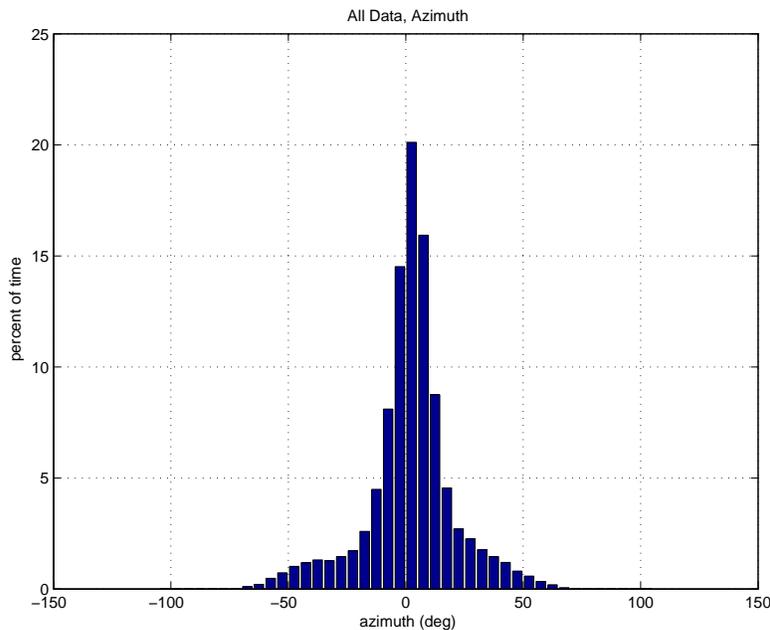


Figure 7: Percentage of Time for Head Direction in Azimuth (deg)

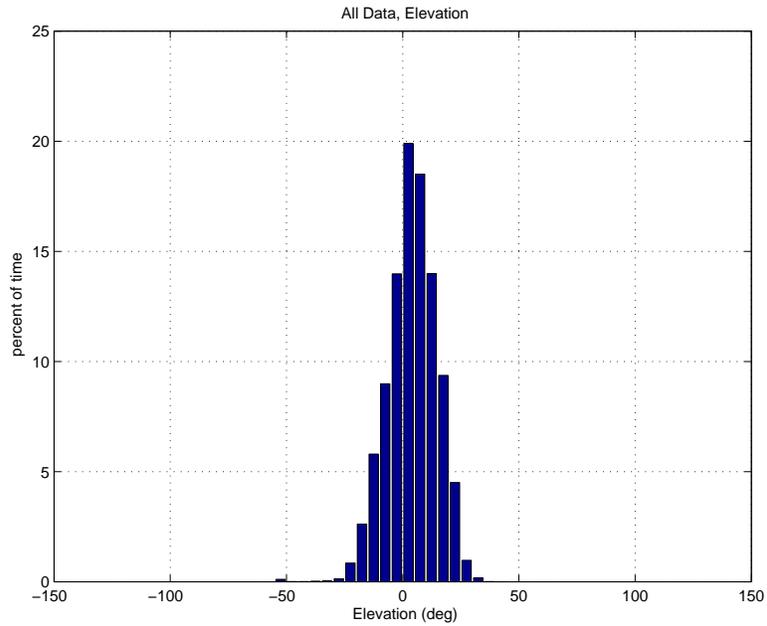


Figure 8: Percentage of Time for Head Direction in Elevation (deg)

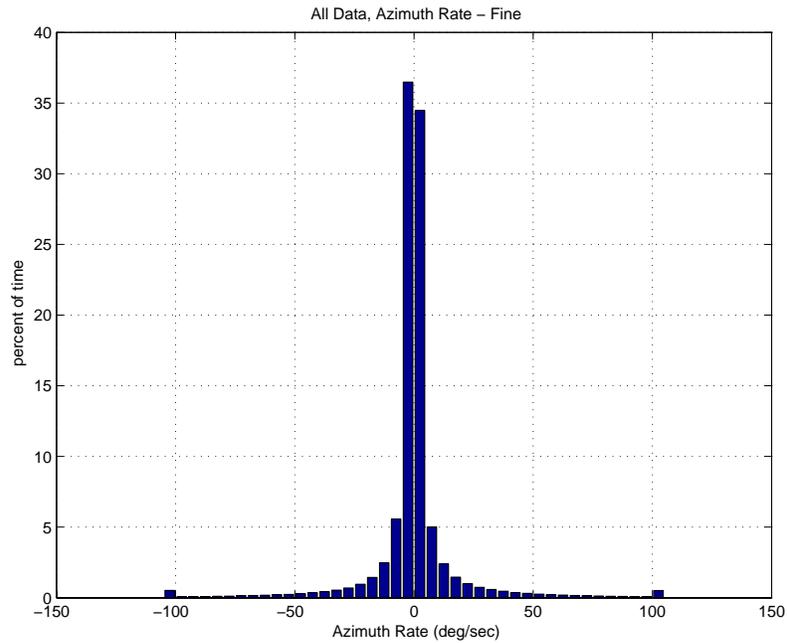


Figure 9: Percentage of Time for Head Movement – Azimuth Rates (deg/sec)

4.4 Simulator sickness

HWDs have been found to induce symptoms of motion or simulator sickness. The (potential) causes are many (e.g., see Nelson et al¹⁹). The occurrence of simulator sickness in operation or training with HWDs for commercial aviation applications would be problematic and detrimental to the commercial applications of HWDs if this were the case (e.g., see Hennessey²⁰).

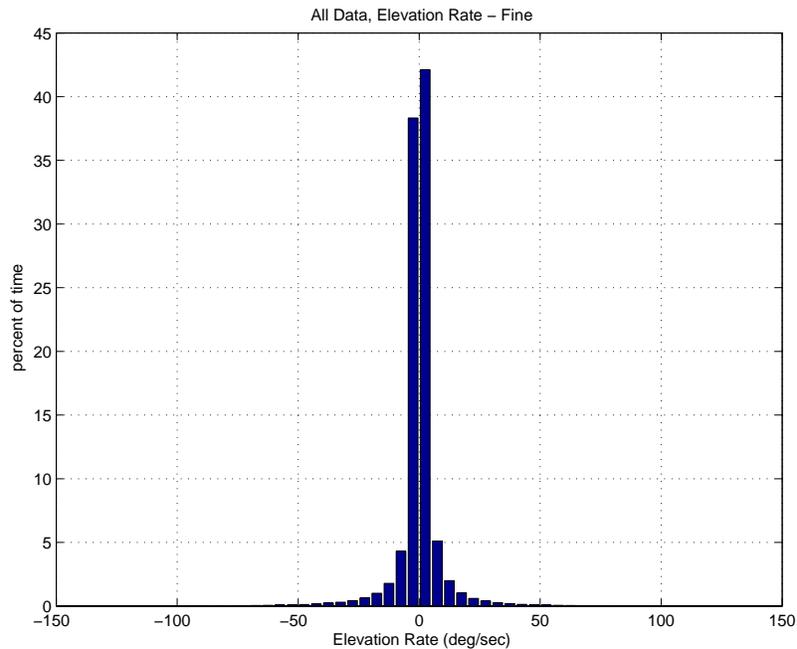


Figure 10: Percentage of Time for Head Movement – Elevation Rates (deg/sec)

During the usability study prior to this experiment, several cases of simulator sickness were experienced (but objective data of the sickness severity was not taken). For the experiment, simulator sickness questionnaires²¹ were used to track the presence and trends in simulator sickness as it might be induced by the HWD through the course of the experiment. The simulator sickness questionnaire considers 16 symptoms, belonging to the three factors of oculomotor disturbance, disorientation, and nausea. The computation of a simulator sickness questionnaire (SSQ) “score” was obtained by weighting the three factors and the 16 symptoms according to Kennedy to yield a total score.

The display configurations were experimentally blocked. After training, the first block consisted of evaluations using “conventional” displays (head-down and head-up display concepts), followed by a second block consisting of evaluations using the HWD display concepts. This pattern was repeated twice more. Simulator sickness questionnaires were administered at the end of each data run.

For 11 out of the 16 crews, no simulator sickness symptoms were noted in any of the runs.

For the other 5 crews, marked simulator sickness symptoms were experienced by the captain (who was wearing the HWD) – relatively severe in two instances. (Interestingly, several cases of simulator sickness were also experienced by the first officer. The first officer was not using a HWD. The reasons for the first officer sickness was probably due to a combination of simulator latency, the lack of motion cues, fatigue over the course of the experiment, and the need to repeatedly go head-down to write taxi clearances and then go head-out to monitor taxi performance.)

In the most severe case, the session had to be stopped because of pilot discomfort. The contributing factor was not felt to be related to the HWD in this case. A couple of “mid-run” simulator resets were unfortunately performed and the sudden visual “rush” from these resets caused the pilot to experience severe symptoms. On Run 17 (out of a planned 27 runs), the pilot rated a total SSQ score of 29.9 with a Nausea score of 57, followed on the next run with a total SSQ score of 45 with a Nausea score of 76. Prior to these runs, no symptoms were reported. The test was subsequently stopped for this crew. More careful resets were subsequently enforced in the simulation protocol (i.e., stopping the aircraft before the reset and warning the pilots of an up-coming reset so they could close their eyes, if desired).

For the remaining 4 affected crews, the mean of the total simulator sickness scores by the Captain within the display block are shown in Figure 11. For these subjects, simulator sickness was generally “triggered” by the first HWD evaluation block. The severity of the sickness symptoms also generally increased over time, analogous to other research findings.¹⁹Error! Bookmark not defined.

For these four captains, the symptoms were primarily related to oculomotor disturbances; however, for the more severe symptoms noted by the captains of Crew 2 and Crew 4, commensurate nausea and to a lesser extent, disorientation symptoms were also noted.

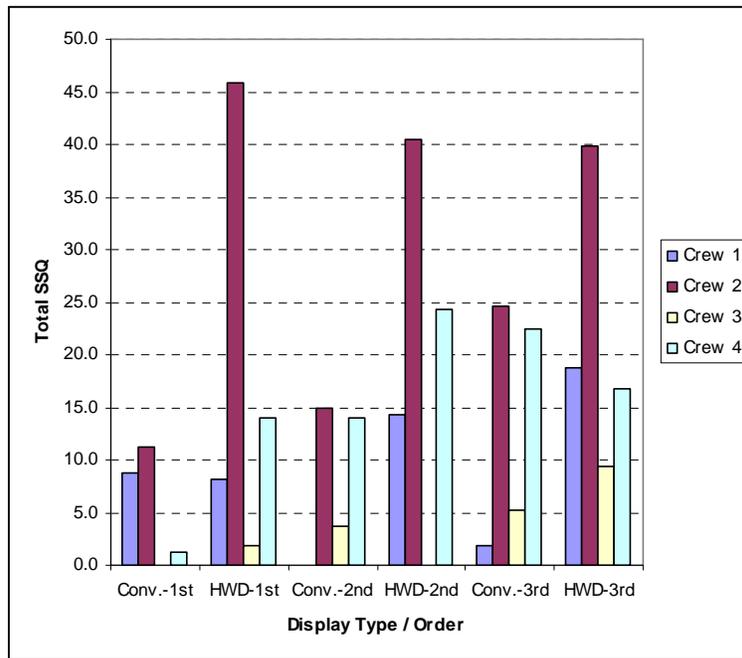


Figure 11: Mean Total Simulator Sickness Questionnaire (SSQ) Score within an Evaluation Block

The SSQ data suggests a “good news / bad news” scenario for the commercial viability of HWDs. The good news is that 75% of the evaluation pilots who used the HWDs experienced *no symptoms* of simulator sickness. The bad news is that in 12.5% of the evaluation pilots who used the HWD did experience notable and in one case, relatively severe symptoms of simulator sickness. Although there is no guidance as to what an acceptable level of the affected pilot population would be, 12.5% seems intuitively unacceptable.

The general and individual-specific differences that influence the propensity for simulator sickness when using an HWD needs to be further explored. Potential influences that the research team felt may have contributed in this experiment were the subject’s most recent flight experience, their fatigue and general health at the time of the simulation, and their propensity for developing virtual “presence” and visual-vestibular conflict. Further, the lack of motion cues in the simulator may have contributed to the simulator sickness results (e.g., several First-Officers experienced simulator sickness symptoms even though they didn’t wear the HWD). Most importantly, however, future investigations must evaluate whether there is any propensity for “simulator sickness” with these devices in operational use versus simulator evaluations.

4.5 Usability

To get a general appreciation of the HWD usability for surface operations, a relatively simple, but broad-based usability tool was used.²² After the completion of the experiment, the captains completed a questionnaire addressing a variety of technology usability issues. By using this broad-brushed scale, a large range of issues were addressed from complexity to usefulness. A comparative evaluation with the head-down or head-up displays was not conducted.

The 10 usability questions are scored (equal weighting) to give an overall rating between 0 and 100. For the experiment, the average score for the HWD concept was a 75 (n=12 pilots). The scores ranged from a maximum of 95 to a low of 52.5 with a 10 point standard deviation around the mean. The rationale for these grades can be determined from responses to the individual questions.

Overall, the system was given high marks by almost all the crews for being easy to use, not being overly complex, and being well integrated.

The poor marks were due primarily to some strong negative opinions by select users. For these questions, a “bi-polar” response was given to whether the system operation could be easily learned (7 pilots strongly agreed that it could, but 3 pilots were neutral to this question) and whether the pilots thought the system was cumbersome to use (3 strongly disagreed with this statement, but 4 were neutral to moderately agreeing to it.) Clearly, if the weight and encumbrance of the system were improved, more positive statements could be obtained to this question.

Another source of disagreement and negative ratings was in response to whether the pilots would “use the system frequently.” 4 pilots strongly agreed with this statement but 3 pilots moderately disagreed. This question should have been better posed since it could be interpreted several ways. For instance, if taken in the context of everyday operation, the pilots might have been rating how often poor weather and limited visibility necessitate the need for HWD taxi assistance. Or, they might have interpreted the question as asking whether they felt this system improved their ability to safely and efficiently conduct surface operations in general.

The good scores are encouraging but the negative scores point to needed areas for improvement. Captains felt the HWD showed high potential but that refinement is clearly needed.

5. OPEN RESEARCH ISSUES

This work and others are being used by NASA to evaluate the potential of HWDs for emerging NGATS requirements. Concepts such as EVO and all-weather surface operations with the safety and capacity of visual flight operations are being evaluated. Requirements to meet these emerging capabilities must be developed to enable the appropriate technology development.

The challenges of HWD operations for surface operations are formidable. Practically speaking, the concept of zero-zero (i.e., zero visibility and zero ceiling) does not often occur. In actuality, some usable visibility is normally available, which means that the HWD must provide an augmented reality capability without significantly obscuring the outside world. NASA is looking at the potential for conformal as well as “semi-conformal” concepts which can provide conformal information without outside visibility obscuration. The need for bi-ocular, see-through and non-obscured displays must be addressed accordingly.

Work must continue toward reducing the system latency and, most critically, the encumbrance of HWDs to make them a valuable technology for commercial and business aircraft. Additionally, improved methods for boresight, alignment, and head-tracking for precise alignment and stability are needed. Finally, these results were obtained using a sophisticated ground simulation facility with no motion effects. These results and others must be validated during actual operations, focusing on daylight readability (particularly for evaluating the utility and capabilities for color), disorientation and illusion issues (i.e., “simulator sickness”), and augmented reality requirements.

6. CONCLUDING REMARKS

A head-worn display concept, coupled with a head tracker, was used to create full-color concepts which capitalize on an unlimited field-of-regard capability. A ground simulation experiment evaluated the efficacy of head-worn display applications to improve operations and safety in surface operations. The simulation experiment was conducted at the O’Hare airport using airline crews and various displays and display concepts in an operationally realistic environment. The results show that HWD concepts may

emerge as viable devices for future operations. However, further requirements definition and concept development work must occur to improve efficiency and safety. The effectiveness and utility of the display concepts, as well as the display device criticality, must be validated in the flight test.

7. REFERENCES

1. Joint Planning and Development Office. *Next-Generation Air Transportation System Integrated Plan*, Version 1. Dec. 2004.
2. 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.
3. 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.
4. Arthur, J.J., Williams, S.P., Prinzel, L.P., Kramer, L.J., and Bailey, R.E.: *Flight Simulator Evaluation of Display Media Devices for Synthetic Vision Concepts*. In Proceedings of SPIE, Helmet- and Head-Mounted Displays IX: Technologies and Applications Editors: Clarence E. Rash and Colin E. Reese, Volume 5442, Apr 2004, pp. 213-224.
5. National Transportation Safety Board, September 2006, Most Wanted Transportation Safety Improvements, www.nts.gov/recs/mostwanted/index.htm
6. Foyle, D.C., Andre, A.D., McCann, R.S., Wenzel, E., Begault, D. and Battiste, V. *Taxiway Navigation and Situation Awareness (T-NASA) System: Problem, Design Philosophy, and Description of an Integrated Display Suite for Low-Visibility Airport Surface Operations*. SAE Transactions: Journal of Aerospace, 105, 1996. pp. 1411-1418.
7. Jones, Denise R., C. C. Quach, S. D. Young, *Runway Incursion Prevention System – Demonstration and Testing at the Dallas/Fort Worth International Airport*. Proceedings of the 20th Digital Avionics Systems Conference. Oct 2006.
8. Arthur, J.J. III, Prinzel, L.J. III, Williams, S.P., and Kramer, L.J.: *Synthetic Vision Enhanced Surface Operations and Flight Procedures Rehearsal Tool*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2006, Editor: Jacques G. Verly, Volume 6226, April 2006.
9. Rediess, H.A.. *An Augmented Reality Pilot Display For Airport Operations Under Low and Zero Visibility Conditions*. AIAA Guidance, Navigation, and Control Conference, New Orleans, LA, Aug. 11-13, 1997, In “Collection of Technical Papers. Pt. 2,” Paper No. AIAA-97-3680, pp. 912-929.
10. Ruffner, J.W., Fulbrook, J.E., and Foglia, M.: *Near-to-Eye Display Concepts for Air Traffic Controllers*. In Proceedings of SPIE, Helmet- and Head-Mounted Displays IX: Technologies and Applications Editors: Clarence E. Rash and Colin E. Reese, Volume 5442, Apr 2004, pp. 120-131.
11. Bailey, R.E., Arthur, J.E., and Williams, S.P.: *Latency Requirements for Head-Worn Display S/EVS Applications* In Proceedings of SPIE, Enhanced and Synthetic Vision 2004 Editors: Jacques G. Verly, Volume 5424, Apr 2004, pp. 98-109.
12. Young, S.D. and Quon, L., *Integrated Intelligent Flight Deck: Research Proposal*. 28 April 2006.
13. NASA Technology Invention Disclosure. *Taxi Operations with Advanced Displays*. April 2006.
14. Arthur, J.J., et al: *Design and testing of an unlimited field-of-regard synthetic vision head-worn display for commercial aircraft surface operations*. In Proceedings of SPIE, Enhanced and Synthetic Vision 2007, Paper 12, April 2007.
15. Hart, S. G., and Staveland, L. E.: *Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research*. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload*. Amsterdam, The Netherlands: Elsevier. 1988. pp 139-183.
16. Taylor, R. M.: *Situational awareness rating technique (SART): The development of a tool for aircrew systems design*. *Situational Awareness in Aerospace Operations*, AGARD-CP-478, 1990, pp. 3-1 - 37.
17. Bailey, R.E., Arthur, J.J. III, Williams, S.P., and Kramer, L.J.: *Latency in Visionic Systems: Test Methods and Requirements*. Paper No. 8 presented at the RTO HFM Workshop on “Toward

Recommended Methods for Testing and Evaluation of EV and E/SV Based Visionic Devices”, held in Williamsburg, Virginia, April 2004, published as NATO RTO-MP-HFM-125.

18. McFarland, R.E.: *CGI Delay Compensation*. NASA TM-86703, 1986.
19. Nelson, W.T., Bolia, R.S., Roe, M.M., and Morley, R.M.: *Assessing Simulator Sickness In A See-Through HMD: Effects Of Time Delay, Time On Task, And Task Complexity*. Paper presented at the IMAGE 2000 Conference, Scottsdale, AR, 10-14 July 2000.
20. Hennessy, R.T., Sharkey, T.J., Matsumoto, J.A., and Vorrhees, J.W.: *Simulator Induced Alteration of Head Movements (SIAHM)*. paper presented at AIAA/AHS Flight Simulation Technologies Conference, Hilton Head Island, SC, Aug 24-26, 1992, Paper No. AIAA-1992-4134, pp. 29-36.
21. Kennedy, R.S., Lane, N.R., Berbaum, K.S., and Lilienthal, M.G.: *Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness*. *The International Journal of Aviation Psychology*, Vol. 3, No. 3, 1993, pp. 203-220.
22. J. Brooke, *SUS: A 'Quick and Dirty' Usability Scale*. *Usability Evaluation in Industry*, P. W. Jordan, et al, Editors, Taylor & Francis, Inc., London (1996), pp. 189-194.