Synthetic Vision Enhanced Surface Operations with Head-Worn Display for Commercial Aircraft

NASA Langley Research Center

Jarvis (Trey) J Arthur III, Lawrence J Prinzel III, Kevin J Shelton, Lynda J Kramer,

Steven P Williams, Randall E Bailey and Robert M Norman

NASA Langley Research Center

Hampton, VA

Abstract

Experiments and flight tests have shown that airport surface operations can be enhanced by using synthetic vision and associated technologies, employed on a Head-Up Display (HUD) and head-down display electronic moving maps (EMM). Although HUD applications have shown the greatest potential operational improvements, the research noted that two major limitations during ground operations were its monochrome form and limited, fixed field-of-regard. A potential solution to these limitations may be the application of advanced Head Worn Displays (HWDs) particularly during low-visibility operations wherein surface movement is substantially limited because of the impaired vision of pilots and air traffic controllers. The paper describes the results of ground simulation experiments conducted at the NASA Langley Research Center. The results of the experiments showed that the fully integrated HWD concept provided significantly improved path performance compared to using paper charts alone. When comparing the HWD and HUD concepts, there were no statistically-significant differences in path performance or subjective ratings of situation awareness and workload. Implications and directions for future research are described.

INTRODUCTION

Background

The Integrated Intelligent Flight Deck (IIFD) project, under NASA's Aviation Safety Program, comprises a multi-disciplinary research effort to develop flight deck technologies that mitigate operator-, automation-, and environment-induced hazards. Towards this objective, the IIFD project 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 (NextGen) (Joint Planning and Development Office [JPDO], 2004). Part of this research effort involves the use of synthetic and enhanced vision systems and advanced display media as enabling crew-vehicle interface technologies to meet these safety challenges.

While NextGen concepts envision the capability to handle up to a 3-fold increase in air traffic, the National Transportation Safety Board (NTSB) continues to have runway incursion prevention on its top six most wanted list for aviation safety (NTSB, 2006). The increase in air traffic could potentially result in a concomitant increase in runway incursions requiring a critical need to develop both air- and ground-based solutions. In the 4-year period between 2001 and 2004, 1,395 runway incursion events were reported to the FAA which is a rate of almost 1 runway incursion event per day (FAA, 2004). Also during this time, over 60% of the FAA towered airports reported at least one runway incursion event. These statistics and events are cause for alarm. The worst aviation accident, in terms of fatalities, occurred in 1977 when two fully loaded 747 airplanes collided on a runway at Tenerife airport. Moreover, each year there are reports of close “near-miss” runway incursions that happen with sufficient regularity at the world's busiest airports to pose perhaps the most significant hazard confronting aviation today.

One such airport plagued with runway incursions is Chicago O'Hare International Airport (FAA identifier: ORD). Chicago O'Hare is a complex airfield and represents one of the world's busiest and most challenging airports for surface operations. Current runway incursion safety mitigations employ a “layered” approach using technology, training, and awareness. The ORD airport authority has identified “hot spots” which are areas where incursions are likely to occur. In these areas, special ground traffic and aircraft handling are designed so that nominal operations minimize incursion potential. These hot spots are published and disseminated to aircraft and ground crew operators to heighten vigilance when operating in and near these areas. ORD operates an Airport Movement Area Safety System (AMASS) which provides warnings of incursions to controllers. Even with these protocols and technology implementations, there have been several close calls (FAA, 2004).

Emerging NextGen concepts may further increase the task-load on the flight deck, precipitating the need for display technologies tailored to support these new operational requirements. For instance, automated surface managements system are being developed that utilize dynamic algorithms to calculate the most efficient movement of all surface traffic to increase the efficiency with which airport surfaces are utilized. If these systems are to be implemented, pilots will be required to comply with 4-D taxi clearances, in which a pilot is required to be at a specific location at a specific time. Furthermore, pilots will be expected to “maintain separation,” even during these “super density” operations, from other aircraft regardless of visibility conditions, just as they do today during visual flight operations. These emerging “equivalent visual operations” concepts will require substantially more critical information to help pilots navigate around the airport without “natural visual” references to ensure safe separation from other aircraft. These operations and their information requirements likely may not be supportable by today's flight deck displays, lending to the need for research to investigate other alternatives that may better be tailored to what is envisioned for the air transportation system of the future.

Past Research

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. These surface operations include taxiing and maneuvering aircraft and vehicles to/from the active runways and gates.

The Taxiway-Navigation And Situation Awareness (T-NASA) concept (Figure 1) was developed to improve the efficiency and safety of airport surface operations in Category IIIB weather (no decision height, <1200 ft {366 m} Runway Visual Range (RVR)) (Foyle et al., 1996). T-NASA uses a suite of cockpit displays - a HUD and an Electronic Moving Map (EMM) concept, implemented on a Navigation Display (ND) or Electronic Flight Bag (EFB). The concepts have been shown to provide the following benefits, in various degrees of measure and success:

- Eliminated hold location errors and failure to hold errors
- Allowed increased taxi speeds
- Eliminated taxi navigation errors in low-visibility and night conditions
- Enabled better awareness of airport traffic
- Improved pilot-Air Traffic Control (ATC) communication of clearance
- Improved Captain-First Officer intra-cockpit pilot communication

Insert Figure 1 here

Under the T-NASA concept, the EMM includes a labeled airport layout, ownship position, positions of other traffic, graphical route guidance, text clearance window, and ground speed and heading indicators. The EMM depicts the cleared taxi route graphically, via the magenta path, and textually, via the text box on the bottom of the map. Hold short instructions are portrayed with a yellow hold bar, and the portion of the route beyond the hold changes from magenta to yellow. Airport traffic is depicted in real-time, and pilots can choose to view aircraft icons with or without data tags. All information is dynamic and updated in real-time.

The EMM is designed with the primary purpose of aiding navigation and situation awareness; it is not designed to support the control of the aircraft. As such, the map purposely lacks specific detail regarding the aircraft's position relative to the centerline, location of wheels, speed or braking parameters, or an accurate depiction of aircraft size and wingspan.

In contrast, the HUD uses “scene-linked” symbology for conformal display against the out-the-window environment (when visible) which theoretically leads to efficient cognitive processing of both the symbology and the environment, and mitigates problems of attentional tunneling and symbology fixation. The taxi symbology contains taxiway centerline markers and taxiway edge cones. Virtual signage aids in augmenting cleared-path awareness. Taxiway information provides enhanced situation awareness for taxi navigation.

Simulation data were analyzed to pinpoint the mechanisms by which T-NASA technology components could mitigate classes of surface operations navigation errors (pilot deviations) (Hooey & Foyle, 2006). A taxonomy of 3 error classes was used. The simulation data replicated current-day operations and also included trials with T-NASA technologies including data-link, EMM, and HUD.

The error decomposition showed that pilots committed navigation errors on 17% of current-day operations trials (in low-visibility and night), distributed roughly equally across the 3 error classes. When using T-NASA technologies, the error data showed a unique set of contributing factors and mitigating solutions:

- Planning errors were mitigated by technologies that provided an unambiguous record of the clearance (data-link and the EMM, which possessed a text-based clearance).
- Decision errors were mitigated by technologies that provided both local and global awareness including information about the distance to and direction of the next turn, current ownship location, and a graphical depiction of the route (as provided by the EMM and HUD together).
- Execution errors were best mitigated by the HUD, which disambiguated the environment and depicted the cleared taxi route.

Further enhancements to the T-NASA concepts have evolved based on follow-on research and testing. In particular, tactical turn guidance, in the form of so-called "breadcrumbs" or other manifestations, have shown to significantly aid in tactical surface operations guidance, particularly for aircraft which require over-steer to remain on the taxiway centerline (Figure 2, (Jones & Rankin, 2002)). Without non-conformal guidance information, the conformal information such as the centerline and edge markings would not be drawn on the limited HUD field-of-view or the information that was provided would be difficult to interpret.

Insert Figure 2 here

The key to preventing runway incursions is to ensure that pilots know (Young & Jones, 1998):

1. Where they are located
2. Where other traffic is located
3. Where they are cleared to go on the airport surface

The T-NASA concepts and its instantiations contribute significantly toward these elements. However, not knowing the above or deviating from clearances, the flight crew and ATC should be alerted to the situation. NASA's Runway Incursion Prevention System (RIPS) program developed methodologies for flight deck alerting, targeted toward the prediction of runway incursion to provide immediate alerting for the principal participants in the operation (i.e., the flight crews). The T-NASA concepts provide guidance and situation awareness information to mitigate many factors contributing toward runway incursions, but a final protective “wrapper” was felt to be warranted nonetheless (Jones & Prinzel, 2006). Since the objective of the present research was to focus evaluations on proactive surface operation situation awareness, none of the concepts tested included flight deck alerting (i.e., RIPS concepts).

Present Study

The present study was an extension of this previous research to evaluate if emerging synthetic vision and head-worn display technologies (Figure 3) can provide further safety and operational improvements and enable an application solution that would better enable an implementation path for synthetic and enhanced vision technologies in support of NextGen operations. This test was the second of two studies evaluating the efficacy of head-worn displays for surface operations (Arthur et al., 2007).

Insert Figure 3 here

METHODOLOGY

Simulation Facility

The Experiment was conducted in the Research Flight Deck (RFD) simulator at NASA Langley Research Center (LaRC). The RFD was a fixed-based, dual-pilot simulator consisting of a collimated 200˚ out-the-window visual. The out-the-window visual consisted of the airport, including taxiways and runways with appropriate markings, airport lighting, model aircraft representing traffic and simulated weather/lighting conditions. The visual acuity of the out-the-window visual provided a Snellen equivalent of approximately 20/80. The RFD was equipped with a 30˚ H x 24˚ V HUD on the captain's side. The HWD, worn only by the captain, was an 800H x 600V pixel, full color display with see-through capability, 60 Hz refresh and a pilot selectable brightness knob.

The subject pilots placed the HWD near the right eye so that it was visible by glancing up which maintained unimpeded stereoscopic vision for out-the-window monitoring. The resulting display was conformal to the real-world (out-the-window visual) if the pilot tilted his or her head down. This procedure was also used to minimize binocular rivalry. An optical head tracker provided the head orientation data. The RFD had 8 Size D (6.4 inch {16.3 cm} square viewable area) head-down displays typical of those found in modern "glass" cockpits: captain and first officer PFD and 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 as 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 medium- to long-haul commercial passenger aircraft, classified as an International Civil Aviation Organization (ICAO) aerodrome code D.

Evaluation Pilots

Twelve 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 15,000 flight hours with 29 years total flight time and the first officers had an average of over 9,000 flight hours with an average of 26 years total flight time. Two-thirds of the captains required corrective lenses. The subject crews were given a 45-minute briefing on the display concepts and the evaluation tasks. After the briefing, a 45-minute training session was conducted to familiarize the subjects with the RFD simulator, the HUD, the HWD device, and the piloting task. Only the captain had a head-up display; the first officer had a head-down repeater display of the captain's head-up device. A simple eye dominance test was performed after the training briefing. Of the 12 subject captains, 11 were right eye dominant. The HWD was viewed with the right eye for all subjects. The HWD is compatible with eyeglasses. Following training, 2.5 hours of data collection was conducted. The total experiment time for each subject crew was approximately 4 hours.

EXPERIMENTAL TEST

Evaluation Tasks

Crews conducted simulated taxi operations at Chicago O'Hare International Airport. The display concept and weather were experimentally varied.

A total of 25 different taxi scenarios were used in the study. The scenarios covered a range of realistic scenarios. The weather state for the out-the-window scene was varied between night-time with unlimited visibility Visual Meteorological Conditions (VMC), and daytime with 700 RVR (213 m). All tasks involved exiting the active runway and taxiing to the airport movement area. Pilots were instructed to taxi at a speed they thought appropriate for the task and to avoid other aircraft. 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.

One of the 25 scenarios was a rare-event (Foyle & Hooey, 2003) runway incursion (see below). Rare event scenarios offer the opportunity to evaluate the display concepts in off-nominal situations. For this final run, the visibility was reduced to 500 RVR (152 m) to set-up the taxiway incursion rare event scenario.

Nominal Taxi Evaluation Tasks

Before each data trial, the flight crews were briefed on their current location and expected runway turnoff. Each trial began with an initial speed of 10 or 15 knots on an active runway followed by an immediate call from the tower controller. Once clear of the runway, the first officer switched to ground frequency and called the ground controller for clearance. The ground controller provided the taxi instructions along with a data-linked message of the cleared route. If requested, the taxi instructions were repeated to the crew. All ATC calls were automated and various other ATC communications were played to simulate typical radio party-line chatter. In addition, other pre-recorded aircraft traffic taxied around the airport surface. Crews were instructed that the traffic was pre-recorded; therefore, they should give way to all traffic. Further, they were briefed that the ground controller would not provide traffic awareness cues.

Runway Incursion Rare Event Scenario

The final run of the day created a potential nose-to-nose taxiway incursion. The nose-to-nose rare event was designed to provide insight into traffic awareness between the different display concepts. A common occurrence at ORD is when the terminal area is congested, aircraft may be given a “double back” clearance to create spacing and clear other taxiways. These events provide for a situation where two aircraft become “stuck” facing each other on a taxiway without taxi clearance to maneuver. When this occurs, it requires an aircraft tug to separate the two airplanes. Such occurrences are serious because a nose-to-nose situation can significantly reduce airport efficiency to resolve the incursion (FAA Class D level incursion).

To induce this “nose-to-nose” taxiway incursion rare event, crews were asked to turn-off Runway 9R onto Taxiway Mike 7 (M7) and contact ground. Ground cleared the aircraft to follow Taxiways Mike, Delta, and hold short of Mike-6 on Delta. Taxiing parallel on the first officer’s side were two aircraft on Taxiway Bravo, holding short of Delta-4 waiting to enter Taxiway Delta. Upon arriving and holding short of Mike-6, the flight crew contacted ground and were then given a ground controller instruction for them to turn onto Taxiway Mike-6 which, unknown to the crew and ground ATC, was already occupied by a small commuter jet (Figure 4).

Insert Figure 4 here

The scenario set-up was typical of occurrence at ORD, and elsewhere, when an aircraft misses a turn and stops on the taxiway awaiting further instruction. A small commuter aircraft has turned-off Runway 9R onto Mike-6 but was told to hold short of Mike (not to continue on Mike-6). Therefore, if the commuter aircraft had complied with the instruction to turn down Mike, it would have been acceptable and provided the necessary separation. However, the aircraft accidentally crossed Mike and continued on Mike-6, then and stopped once they had realized the mistake. They were in the process of switching over to ground when the evaluation crew received the ground instruction to turn down Mike-6, creating the “nose-to-nose” situation.

The visibility was reduced to 500 RVR (152 m) for this scenario so that the traffic was difficult to see, but still detectable out-the-window. The incurring aircraft was on the left (captain's) side. The scenario also depicted two additional aircraft on the first officer's side to serve as a potential distraction to the crew’s attention. These two aircraft also lent credibility to the amended clearance to turn onto Mike-6 taxiway for clearance separation from these two B-747-400 aircraft. To further increase the workload of the first officer, a complex ground clearance was given close to the incursion point. Therefore, the prevention of the nose-to-nose situation depended almost entirely on the captain's awareness – the principal subject of the experimental display variation.

The rare event display condition was evenly distributed across flight crews between each of the 4 display concepts; therefore each of the 4 display concepts contained 3 rare event data points total. The scenario was presented on the last trial of the day but this was not communicated to the pilots, who were unaware of the number of trials to be presented.

Display Concepts

Four display concepts were tested (Figure 5):

1. A Paper Chart of the ORD airport surface representative of today's airline equipage. This concept is hereafter referred to as the "Paper,"
2. A head down only display consisting of an advanced EMM containing iconic traffic, clearance and routing information. This concept is hereafter referred to as the "Advanced EMM,"
3. A HUD concept with an advanced EMM head-down display. The scene-linked HUD symbology consisted of a 3-dimensional depiction of the cleared route by highlighting the taxiway edge lines and centerlines. In addition, non-conformal symbology was provided for precision guidance ("breadcrumb" display). The head down display was the advanced EMM described in Item 2 above. This concept is hereafter referred to as the "Advanced HUD," and
4. An advanced HWD concept coupled with a head tracker. With virtual traffic and routing information and an advanced EMM head-down display. This concept is hereafter referred to as the "Advanced HWD."

Insert Figure 5 here

Advanced Electronic Moving Map Display

With the exception of the Paper concept, all display concepts employed a head-down EMM display shown on both the captain and first officer navigation displays. The 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. Both the captain and first officer had independent range controls for the EMM, which consisted of 4 zoom levels. In addition to the perspective track-up mode, the pilot could select a north-up mode that showed the entire airport view from directly above.

Advanced Head-Up Display

The Advanced HUD display concept was based on the RIPS (Jones, 2005; Jones, Quach & Young, 2001) and T-NASA (Foyle et al, 1996; Atkins, 1999) concepts albeit without incursion alerting (Figure 6). The head-up display showed current ground speed in digital format, the current taxiway, next cleared taxiway, 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 roadway turn signs) (Hooey, Foyle, & Andre, 2001). Runway holding positions were displayed as a single solid line at the hold short locations. 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 a declutter button. The auto-throttle disconnect button was used for declutter because it was conveniently located and auto-throttles were not used in the experiment. A second press of the auto-throttle disconnect restored all of the HUD symbology. The captain also had control of the brightness level of the HUD.

Insert Figure 6 here

Advanced Head-Worn Display

The Advanced HWD concept provided a conformal (head-tracked) virtual airport view from the pilot's eye perspective (Figure 7). The virtual airport view consisted of the ORD airport, buildings, surface movement areas and centerlines. (The information elements and formatting for this Synthetic Vision scene were not established experimentally, but were chosen as an initial basis for concept evaluation. Follow-on studies will be used to establish these information requirements.) 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, for better visibility, and did not necessarily correlate with the actual out-the-window sign placement. 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 displayed for the cleared route and for the distance to the next taxiway, and taxiway edge cones depicted the edge lines of the cleared route. Like the HUD, virtual turn signs were used as an additional turn cue and hold short cues were denoted by virtual stop signs. Similar to the HUD, a non-conformal insert depicted a plan-view of the cleared route, together with the airplane outline and location of the nose and main gear. The pilot could remove this non-conformal display by pressing the auto-throttle disconnect button. A second press of the auto-throttle disconnect button removed all symbology in the HWD. A third press restored all HWD symbology . Also, the captain could control the brightness of the display via a rotary knob, located on the center pedestal. The Advanced HWD format represents the most complex, but also the most preferred configuration based on a usability study conducted prior to the simulation experiment (Arthur, Prinzel, Williams, & Kramer, 2006).

Insert Figure 7 here

The HWD concept was implemented using commercial off-the-shelf equipment and was not necessarily representative of the envisioned operational concept. For instance, the HWD concept required alignment (known as boresighting) before the start of the data trials. A fielded concept, however, should not require a manual alignment procedure and should instead be aligned through an automated process. An alignment grid was displayed on the out-the-window visuals and the HWD. The captain boresighted the HWD by aligning the grids through head movement. Once the grids were aligned, the captain verbally called “alignment” and the boresight parameters were saved. The conformality of the virtual airport view was then dependent upon the accuracy of the boresighting. The conformal virtual airport view was shown whenever the captain slightly tilted his/her head down. When properly aligned, the virtual taxiways overlayed the actual taxiways.

RESULTS

Quantitative Results

Taxi Performance for All Data Runs

A Multivariate Analysis of Variance (MANOVA) on Root Mean Square (RMS) path error, taxi speed, and time-to-taxi yielded significant effects for display condition, $F(9, 363)=4.18$, $p<0.001$; visibility condition $F(3, 149)=5.71$, $p=0.001$; and their interaction $F(9,363)=3.60$, $p<0.001$. Follow-up univariate Analysis of Variance (ANOVA) indicated that time-to-taxi was not significantly ($p>0.05$) affected by display condition but taxi speed and RMS path error were. Post-hoc Student-Newman-Keuls (SNK) tests revealed 2 unique subsets of the display conditions for taxi speed and for RMS path error (Figure 8).

Insert Figure 8 here

Pilots taxied significantly slower and had more path error when using Paper charts than with the Advanced EMM, Advanced HUD, and Advanced HWD. On average, pilots had 5.4 ft less path error and taxied 0.4 knots quicker during Day 700 RVR conditions than in Night VMC.

Navigational Errors

Navigational errors, when they occurred, were divided into 2 categories: major and minor (McCann et al, 1998). A major navigation error is defined as a loss of navigational awareness, which resulted in a wrong turn or a failure to turn. A minor navigation error is defined as a failure to remain on route but it was immediately noticed and corrected by the crew. A navigation error, which involved an incursion with other aircraft, was accounted for in a different measure and not captured as a navigational error. A total of 14 navigational errors were made, where 7 were classified as major errors and 7 were classified as minor. Most of the errors (8 of 14 total errors) occurred with the Baseline Paper Chart concept (Figure 9).

Insert Figure 9 here

An ANOVA was performed on the number of navigational errors committed by the crew with the display concept (Baseline Paper, Advanced EMM, Advanced HUD, Advanced HWD), navigational error category (major, minor), and visibility (Night VMC, Day 700 RVR) as the main factors. Display concept, $F(3,33)=2.93$, $p=0.048$; navigational error category, $F(1,11)=9.43$, $p=0.011$; and the interaction between the display concept and navigational error category, $F(3,113)=4.12$, $p=0.008$) were significant. The remaining main factors and interactions were not significant ($p>0.05$) for the number of navigational errors committed. Post-hoc tests (SNK, using $\alpha=0.05$) showed 2 unique subsets for the display concept: 1) Advanced EMM, Advanced HUD and Advanced HWD (fewest navigational errors) and 2) Paper. The crews had fewer minor navigational errors than major errors with the concepts tested, most notably within the Baseline Paper concept.

Taxi Incursions

A taxiway incursion event was defined as a collision with another aircraft or making a turn in front of another aircraft and creating a close call. A total of 2 incursion events occurred, one with the Advanced EMM concept and one with the Advanced HUD concept. An ANOVA was performed on the number of taxi incursions committed by the crew with the display concept (Baseline Paper, Advanced EMM, Advanced HUD, Advanced HWD), and visibility (Night VMC, Day 700 RVR) as the main factors. There were no significant differences ($p>0.05$) among the main factors or their interactions for this measure.

Required Navigation Performance

To quantify path performance from this experiment, the principles of Required Navigation Performance (RNP) for surface operations were employed (Cassell, Smith, & Hicok, 1999). For this experiment, the visibility conditions were such that RNP requirements for an ICAO code D aircraft stipulates path deviation within ±7.2 ft (±2.2 m) of the route centerline 95% of the time (Figure 10). RNP was assumed in this analysis to be solely a function of path error; that is, path performance was quantified as Flight Technical Error (FTE) and RNP was computed assuming no errors contributions due to path definition or positioning error.

None of the display concepts met RNP requirements (i.e., none were within ±7.2 ft {±2.2 m} of route centerline 95% of the time).

Analysis showed that time-of-day was not significant, $F(1,10)=7.563$, $p=0.094$. However, an ANOVA revealed that display type was significant, $F(3,10)=3.719$, $p=0.05$. Post-hoc tests on display type show two overlapping subsets: 1) Advanced EMM (mean=12.5 ft {3.8 m}), Advanced HUD (mean=13.5 ft {4.1 m}) and Advanced HWD (mean=13 ft {4 m}) and 2) Advanced HWD (mean=13 ft {4 m}), Advanced HUD (mean=13.5 ft {4.1 m}), Paper (mean=15.75 ft {4.8 m}). Paper had significantly worse lateral RNP during surface operations than the Advanced EMM; however there was no significant differences between Paper and the Advanced HUD and Advanced HWD.

Insert Figure 10 here

Rare Event

For the Paper Chart display concept, which did not have path or traffic information, all 3 crews got into a nose-to-nose condition. For the display concepts that had iconic traffic display (Advanced EMM, Advanced HUD and Advanced HWD), all crews were able to avoid the nose-to-nose situation.

Qualitative Results

Post-Run Questionnaires

Several questionnaires were given at the end of each data run. At the end of the day, paired-comparison questionnaires were given to both the captain and first officer.

Situation Awareness Rating Technique

An ANOVA was performed for the dependent variable, situation awareness (SA), derived from flight crew ratings on the SART where SA = Understanding - (Demand - Supply). Analysis found a significant effect for display condition, $F(3,15) = 3.77$, $p < 0.05$. A SNK test revealed two unique subsets: (1) Advanced HWD (135.25), Advanced HUD (142.16), or Advanced EMM (142.38) (no significant differences between) - Highest SA and (2) Baseline Paper condition (82.0) - Lowest SA.

Taxi Situation Awareness Questions

Flight crews were administered a Likert post-run experimental questionnaire (1 to 5 scale; 1 = ``not at all''; 5 = ``very much'') after each run which asked the pilots to rate the display conditions contribution to: a) taxi efficiency, b) overall navigation awareness, c) route awareness of local controller clearance, d) route awareness of ground controller clearance, e) surface traffic awareness, f) direction awareness, and g) taxi safety (McCann et al., 1998). Advanced concepts (path, clearance and traffic displayed) were rated significantly higher than the Baseline Paper display concept for all questions. An ANOVA revealed significant effects for all dependent variables, $p < 0.05$. Post-hoc SNK tests were performed on these dependent variables resulting in two unique subsets: (1) no significant differences between Advanced EMM, Advanced HUD, and Advanced HWD, and (2) Baseline Paper condition. Only flight crew ratings of display contribution to taxi efficiency were found not to be significant.

Post-Test Paired Comparisons

A MANOVA statistical procedure was performed on four paired comparison scales administered to the captain and first officer of each flight crew. The paired comparison scales asked the pilot to evaluate each of the four display concept in comparison to one another on four constructs: Situation Awareness (SA-Subjective Workload Dominance (SWORD)), Mental Workload (SWORD), Taxi Efficiency, and Surface Operations and Taxi Safety. The analyses were conducted separately for captain and first officer responses. Significant results reported are at the $p < 0.01$ significance level.

Situation Awareness

For the Captain ratings, there was a significant main effect found for situation awareness, $F(3,30) = 17.37$. A post-hoc test revealed three overlapping subsets: (a) Paper was rated significantly lower for situation awareness (SA-SWORD) than the other three display concepts; (b) the Advanced EMM was rated significantly lower than the Advanced HWD but not significantly different from the Advanced HUD; and (c) no significant differences between the Advanced HUD and Advanced HWD concepts.

For the First Officer ratings, there was a significant main effect for situation awareness, $F(3,30) = 17.9$. A post-hoc test revealed that Paper was rated significantly lower for situation awareness (SA-SWORD) than the other three display concepts. No other effects were found to be significant.

Mental Workload

For the Captain ratings, there was a significant main effect found for mental workload, SWORD, $F(3,30) = 366.69$. A post-hoc test revealed that Paper was rated significantly higher for mental workload than the other three display concepts. There were no differences between the other 3 display conditions for mental workload, as assessed by SWORD.

For the First Officer ratings, there was a significant main effect found for mental workload, $F(3,30) = 91.33$. A post-hoc test revealed that Paper was rated significantly higher for mental workload than the other three display concepts and there were no differences between these concepts.

Taxi Efficiency

For the Captain ratings, there was a significant main effect found for taxi efficiency, $F(3,30) = 25.76$. A post-hoc test revealed three overlapping subsets: (a) Paper was rated significantly lower for taxi efficiency than the other three display concepts; (b) the Advanced EMM was rated significantly lower for taxi efficiency than both the Advanced HUD and Advanced HWD; and (c) no significant differences between the Advanced HUD and Advanced HWD concepts.

For the First Officer ratings, there was a significant main effect found for taxi efficiency, $F(3,30) = 32.96$. A post-hoc test revealed that Paper was rated significantly lower for taxi efficiency than the other three display concepts. No other effects were found to be significant.

Surface Operations and Taxi Safety

For the Captain ratings, there was a significant main effect found for taxi safety, $F(3,30) = 4.9$. However, subsequent post-hoc pair-wise comparison (Bonferroni) failed to find any mean difference significant at the $\alpha = 0.05$ level.

For the First Officer ratings, there was a significant main effect found for taxi safety, $F(3,30) = 14.74$. A post-hoc test revealed that Paper was rated significantly lower for taxi safety than the other three display concepts. There were no significant differences for taxi safety between the other three display concepts.

HWD Usability

To get a general appreciation of the HWD usability for surface operations, a relatively simple, but broad-based usability tool was used (Brooke, 1996). 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 pilot rated statements of the HWD usability were:

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

The 10 usability questions are scored (equal weighting) to give an overall rating between 0 and 100 (Figure 11). The positive statements (the odd statements, Figure 12) were weighted by a factor of 2 while the negative statements (the even statements, Figure 13) were reversed scored and weighted by a factor of two. Figures 12 and 13 are boxplots (Tukey, 1977) that present the smallest and largest values, the lower and upper quartiles and the median for each usability statement. For the experiment, the average score for the HWD concept was a 75 for 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 scores can be determined from responses to the individual questions.

Insert Figure 11 here

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 certain pilots. 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.)

Insert Figure 12 here

Insert Figure 13 here

Another source of disagreement and negative ratings was in response to whether the pilots would "use the system frequently." Four 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.

DISCUSSION

Taxi Performance Results

The results with the Advanced HUD in this experiment are similar to the results from previous surface operations research conducted by NASA Ames and NASA Langley (e.g., McCann et al., 1998). This lends credence to our experimental methods.

The taxi performance data showed no significant differences between the Advanced EMM, Advanced HUD, and Advanced HWD display concepts for the dependent variables measured, but pilots taxied at significantly faster speeds and more accurately with these displays than when taxiing with just paper charts alone. On average, pilots were able to complete the taxi route 15% faster with the advanced concepts compared to Paper. Previous T-NASA research reported taxi speed increases in the range of 16% to 26% (Hooey et al., 2001). Additionally, the crews made significantly more navigation errors with the paper charts than with any of the other three advanced display concepts. As one would expect intuitively, the data support that advanced display concepts provided information (e.g., cleared route, ownship position, taxi guidance cues) and enabled faster, safer, and more efficient taxi than paper.

No quantitative taxi performance differences were found in this study differentiating head-up versus head-down display concepts (when displaying essentially the same or similar information). The taxi performance data showed no significant differences between the advanced concepts for route accuracy or taxi speeds. This result is somewhat counter-intuitive since the EMM is designed to promote strategic taxi information needs, whereas the head-up concepts (HUD and HWD) are designed to promote tactical taxi information needs. These results suggests that the flight crew could effectively use the EMM concepts for tactical taxi decision-making despite their head-down location.

Comparing the Advanced HWD and Advanced HUD concepts, there were no statistically significant taxi performance differences. This result was expected since the information content of the two concepts was essentially identical for taxi performance.

Required Navigation Performance

As NextGen concepts emerge, the importance of Required Total Performance (Communication, Navigation, and Surveillance) will certainly permeate through all future operational scenarios. The data from this experiment were cast into RNP to assess how this philosophy may apply in surface operations. The data shows that none of the display concepts were within surface RNP requirements as proposed by Cassell et al. (1999). This result raises several issues.

The taxi routes used in this experiment were very challenging by design especially when considering the given visibility conditions. Nonetheless, future operations require VFR-like safety and operational tempos in all weather conditions, especially in surface operations. The display concepts did not include an explicit display of Surface Operations RNP / Actual Navigation Performance (ANP) information, like that required for flight operations. These concepts should be evaluated in future work to assess if they are required or practical.

Even with concepts that explicitly showed taxi path, the crews were instructed to get to the gate as quickly as practical. In addition to RNP display concepts for surface operations, the

FTE component to RNP requirements for surface operations need to be validated and carefully considered with taxi efficiency (speed) and safety.

Subjective Ratings

The results of the paired comparisons showed that the addition of a head-up or head-worn display subjectively increased taxi efficiency compared to just having an advanced EMM alone. These results agree with past research conducted at NASA Ames Research Center demonstrating that the combination of head-up display and head-down display taxi concepts provides superior taxi performance. This subjective preference did not, however, manifest itself in quantitative performance. However, when the subject captains were asked to rate their overall impressions of situation awareness, the SA-SWORD results did reveal that both the Advanced HUD and Advanced HWD provided significantly higher SA than the Advanced EMM concept.

The mental workload results as measured by the SWORD post-run questionnaire, revealed that the advanced display concepts were rated as having significantly lower mental workload than the Paper concept. There were no significant differences found between the three advanced display concepts for the SWORD construct.

No significant differences were found across the any of the subjective constructs between Advanced HUD and Advanced HWD, suggesting that these two display concepts provide for equal amounts of perceived situation awareness, mental workload, taxi efficiency, and taxi safety. This result – and the quantitative performance shown above – suggest that the unlimited FOV provided by the HWD and its color characteristics did not provide any advantages over the HUD in this test. Further analysis of this result will be conducted in association with the HWD design, investigating the influence of HWD latency, display FOV, monocular/biocular display design and HWD format/information.

Runway Incursions

The nose-to-nose rare event was designed to highlight traffic awareness by the crew. All crews with display concepts that had traffic information (Advanced EMM, Advanced HUD, and Advanced HWD) were able to avoid the nose-to-nose situation. Crews had information available within the cockpit that contradicted the controller's clearance. Crews contacted the ground controller to resolve the discrepancy to avoid costly and potentially hazardous mistakes. Further, crews commented that the information presented on the HUD provided no additional benefit in detecting this rare event in contrast to the information available on the HWD (e.g., traffic). In other words, the Advanced HWD presented another source for displaying traffic information that supplemented information being presented on the Advanced EMM head-down display.

Head Tracking

Two issues that influenced this work were the head-tracker size and its alignment/accuracy. For this experiment, the HWD was installed on a helmet to provide a stable mounting location for the head-tracker. This configuration resulted in significant pilot encumbrance and head-borne weight. Also, the HWD was aligned with the scene by displaying a grid pattern in the HWD and the same pattern in the out-the-window visuals. For actual operations, the alignment process must be quick, reliable and with a pre-determined degree of integrity and assurance. Further, the HWD image stability and alignment must be maintained during operation. (With a HUD, this boresighting procedure is done once and “hard-mounted” into the aircraft.) Current research efforts are exploring the use of optical head tracking techniques that would minimize or eliminate these HWD “costs.” Otherwise, any dollar savings, derived by weight reductions for HWD-equipage would be out-weighed by the cost in developing a robust procedure for HWD alignment, image correlation, and pilot “encumbrance.”

CONCLUSIONS

At the start of a data collection day (without experiencing any of the concepts), most crews commented that the head down Advanced EMM was the only display needed. Subject pilots commented that the Advanced EMM concept was a “quantum leap” for surface awareness compared to the “baseline” (i.e., paper charts). However, it became clear to them over the course of the experiment that even greater enhancements to situation awareness were provided by the head-up displays (HUD and HWD). This observation was clearly reflected in their ratings.

The results suggest that the Advanced HUD and Advanced HWD are comparable to each other with regard to mental workload, taxi efficiency, taxi performance, and perceived taxi safety. There were a few limitations of the implementation of the HWD concept that may have reduced its full potential to demonstrate marked differences between the capabilities of the HUD and HWD concepts.

Although no statistically significant quantitative efficiency or safety advantages of the advanced HWD were shown over the HUD, there are other considerations that argue for a HWD solution. On the ground, one of the main tasks of the crew is to survey all around the aircraft to avoid collisions with other airplanes or objects on the airport surface. The typical viewing area of a HUD is 30˚ H by 24˚ V, which is sufficient for flight but not necessarily for surface operations. This limitation was especially evident in the present experiment when the flight crew attempted turns but the path was only displayed as virtual turn flags in the HUD due to required over-steering. This attribute will be evaluated in future work. Further, the HWD provides potential weight savings that would have significant cost advantages to operators.

FUTURE RESEARCH ISSUES

The experiments revealed numerous future directions to better optimize and develop these concepts. One future direction involves the integration of enhanced vision sensor technology for surface operations. The capability for “equivalent visual operations” in-flight and on the surface is a goal of NextGen to obtain the safety and operational tempos of Visual Flight, independent of the actual weather conditions. HWD concepts will be evaluated for their efficacy to support “Equivalent Visual Operations” and evaluated in scenarios representative of emerging NextGen concepts such as 4D surface operations. Applications will test the necessity for unlimited field-of-regard and color capability while trading-off HWD design issues, such as display format, display FOV, and monocular/biocular optics. RNP requirements for surface operations will be validated and tested against speed and safety of 4D surface operations.

Further, for these experiments, the routing and clearance information was relayed to the aircraft displays via a simulated controller data-link. Currently, the IIFD/Crew-Vehicle Interface team is conducting research employing voice recognition technology to quickly and accurately enter routing information during read-back. The potential also exists for conducting analysis of the speech and airport information for route awareness and route / track crew-error analysis.

A significant body of research has shown that runway incursions can be mitigated or even prevented via flight deck alerting. For this experiment, however, the crew's situation awareness in the absence of alerting was of most interest. Alerting, in conjunction with these displays, would clearly add significantly to enhancing further the safety of surface operations. Future research will evaluate the additive effects of including such alerting algorithms, derived from the NASA RIPS research, to determine whether further safety enhancements to airport surface operations are possible.

References

Arthur III, J.J., Prinzel III, L.J., Shelton, K.J., Bailey, R.E., Kramer, L.J., Williams, S.P. & Norman, R.M. (2007). Design and testing of an unlimited field-of-regard synthetic vision head-worn display for commercial aircraft surface operations. Enhanced and Synthetic Vision 2007. eds. Jacques G. Verly and Jeff J. Guell. April. 6559.

Arthur III, J.J., Prinzel III, L.J., Williams, S.P., & Kramer, L.J. (2006). Synthetic vision enhanced surface operations and flight procedures rehearsal tool. Enhanced and Synthetic Vision Proceedings of SPIE. eds Jacques G. Verly and Jeff J. Guell. 6226. page 62260I.

Atkins, M. L. (1999). Head-Up Display Symbology For Surface Operations: Comparisons Among Scene-Linked Symbology Sets for Optimum Turn Navigation. 10th International Symposium on Aviation Psychology. Ohio State University. Columbus, OH.

Brooke, J. (1996). SUS: A Quick and Dirty Usability Scale. Usability Evaluation in Industry. Eds P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland. London, UK. Taylor and Francis. p. 189-194.

Cassell, R., Smith, A. & Hicok, D. (1999). Development of Airport Surface Required Navigation Performance (RNP). NASA/CR-1999-209109. NASA Langley Research Center. June.

FAA (2004). Runway Incursion Trends and Initiatives at Towered Airports in the United States, FY 2001 through FY 2004. FAA Runway Safety Report. US Department of Transportation. Washington DC, USA. August.

Foyle, D.C., Andre, A.D., McCann, R.S., Wenzel, E., Begault, D. & Battiste, V. (1996). 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. Vol 105. p. 1411-1418.

Foyle, D.C & Hooey, B.L. (2003). Improving Evaluation and System Design Through The Use of Off-Nominal Testing: A Methodology for Scenario Development. 12th International Symposium on Aviation Psychology. Dayton, OH. Pages 397-402.

Hart, S. G. & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. Human mental workload. Amsterdam, The Netherlands. Ed. P. A. Hancock and N. Meshkati. Elsevier. p. 139-183.

Hennessy, R.T., Sharkey, T.J., Matsumoto, J.A. & Vorrhees, J.W. (1992). Simulator Induced Alteration of Head Movements (SIAHM). AIAA/AHS Flight Simulation Technologies Conference. Hilton Head Island, SC. August. p. 29-36. AIAA-1992-4134.

Hooey, B.L. & Foyle, D.C. (2006). Pilot Navigation Errors on the Airport Surface: Identifying Contributing Factors and Mitigating Solutions. The International Journal of Aviation Psychology. Lawrence Erlbaum Associates, Inc. Vol 16(1). pages 51-76.

Hooey, B.L., Foyle, D.C. & Andre, A.D. (2001). The design of aircraft cockpit displays for low-visibility taxi operations. Ed. A.G. Gale. Vision in Vehicles IX. Holland. Elsevier Science Publishers.

Joint Planning and Development Office (2004). Next-Generation Air Transportation System Integrated Plan (Version 1). U.S. Department of Transportation. December.

Jones, D. (2005). Runway Incursion Prevention System Testing at the Wallops Flight Facility. Enhanced and Synthetic Vision Proceedings of SPIE. Ed. Jacques G. Verly. Vol 5802. p. 47-58. Bellingham, WA.

Jones, D.R. & Prinzel, L.J. (2006). Runway Incursion Prevention for General Aviation Operations. 25th Digital Avionics Systems Conference. October.

Jones, D.R., Quach, C.C. & Young, S.D. (2001). Runway Incursion Prevention System - Demonstration and Testing at the Dallas/Fort Worth International Airport. 20th Digital Avionics Systems Conference. October.

Jones, D.R. & Rankin, J.M. (2002). System for Preventing Runway Incursions. Journal of Air Traffic Control. Vol 44(3). July – September.

Kennedy, R.S., Lane, N.E., Berbaum, K.S. & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. International Journal of Aviation Psychology. p. 203-220.

McCann, R.S., Hooey, B.L., Parke, B., Foyle, D.C., Andre, A.D. & Kanki, B. (1998). An Evaluation of the Taxiway Navigation and Situation Awareness (T-NASA) system in high-fidelity simulation. SAE Transactions: Journal of Aerospace. Vol 107. p. 1612-1625.

National Transportation Safety Board (2006). Most Wanted Transportation Safety Improvements. List Brochure. November. http://www.ntsb.gov.

Nelson, W.T., Bolia, R.S., Roe, M.M. & Morley, R.M. (2000). Assessing Simulator Sickness In A See-Through {HMD}: Effects Of Time Delay, Time On Task, And Task Complexity. IMAGE 2000 Conference. Scottsdale, AR. July.

Stanney, K., Kennedy, R.S. & Drexler, J.M. (1997). Cybersickness is not simulator sickness. Human Factors Society and Ergonomics Society 41st Annual Meeting. p. 1138-1142.

Tukey, J.W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Williams, J.L., Hooey, B.L. & Foyle, D.C. (2006). 4-D Taxi Clearances: Pilots' Usage of Time- and Speed-Based Formats. Proceedings of the AIAA Modeling and Simulation Technologies Conference. AIAA-2006-6611.

Young, S.D. and Jones, D.R. (1998). Flight Testing of an Airport Surface Guidance, Navigation, and Control System. Institute of Navigation Conference. January.

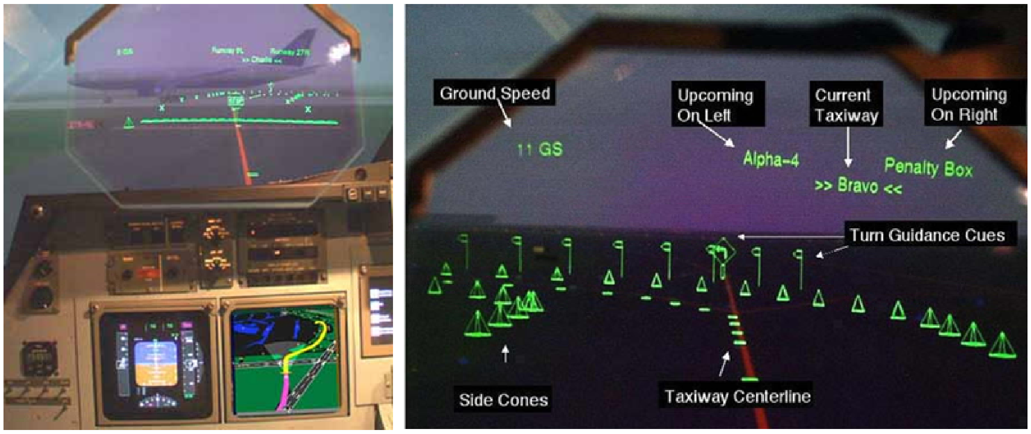

Figure 1. Illustration of the T-NASA concept.

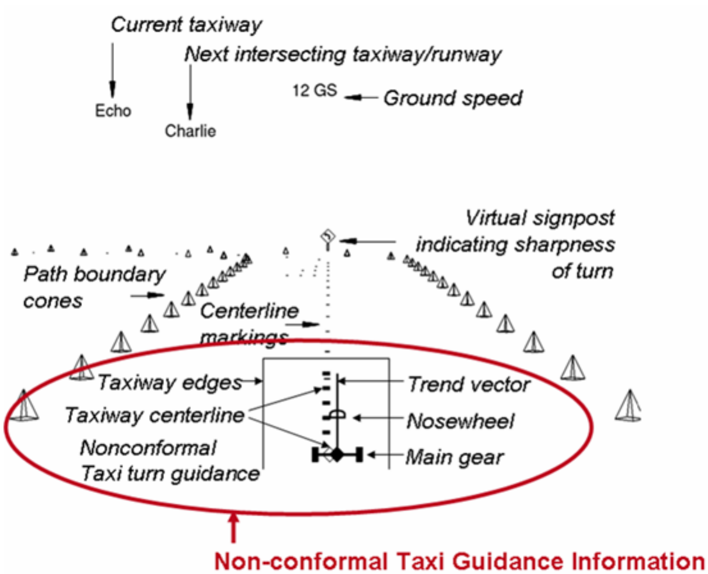

Figure 2. RIPS HUD concept.

Figure 3. The HWD concept and envisioned use in future advanced cockpits.

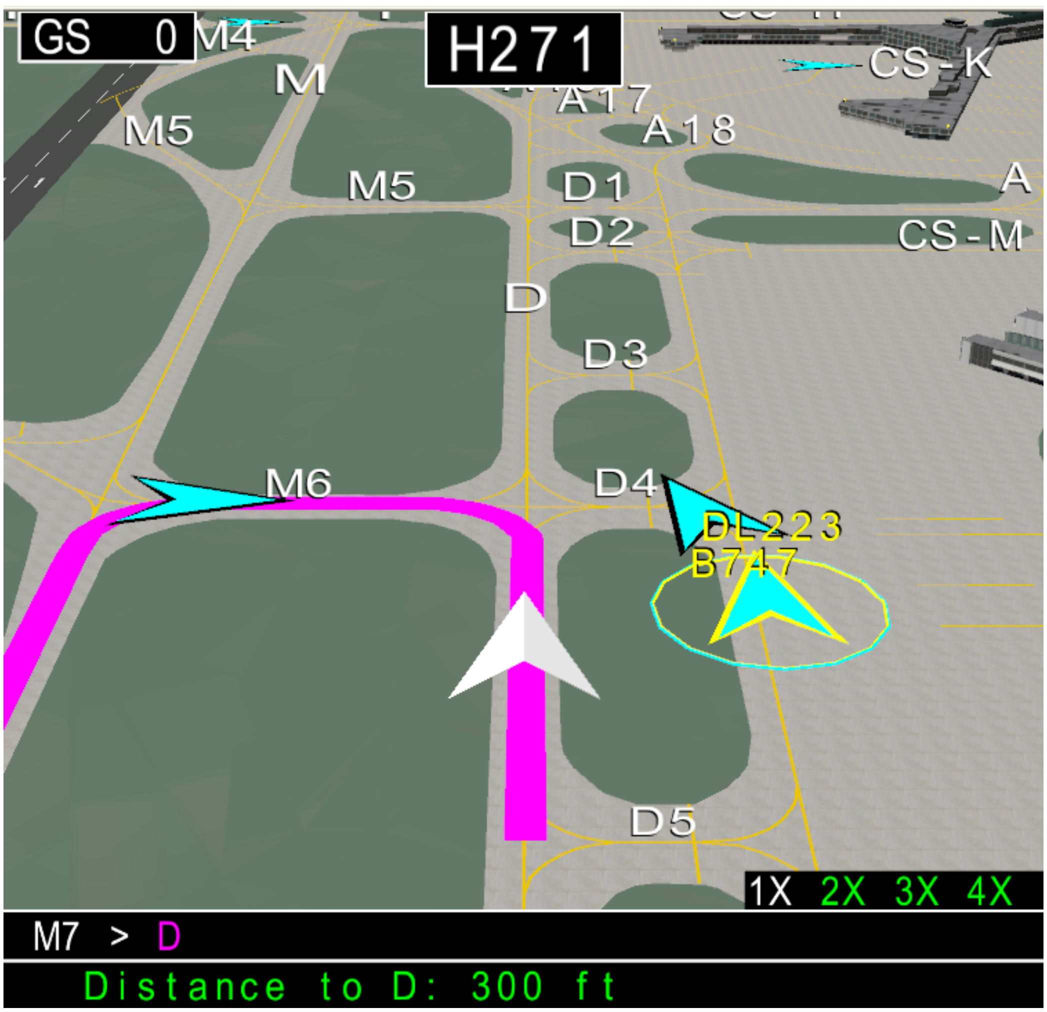

Figure 4. The nose-to-nose scenario on the EMM.

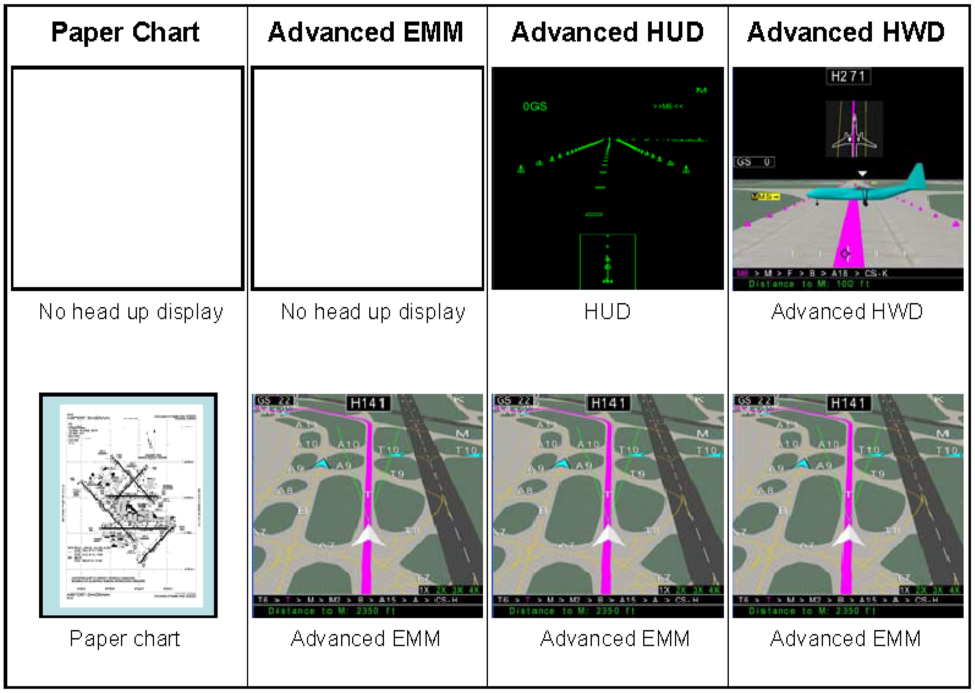

Figure 5. The display concepts evaluated in the experiment.

Figure 6. Example of the HUD concept used in the experiment.

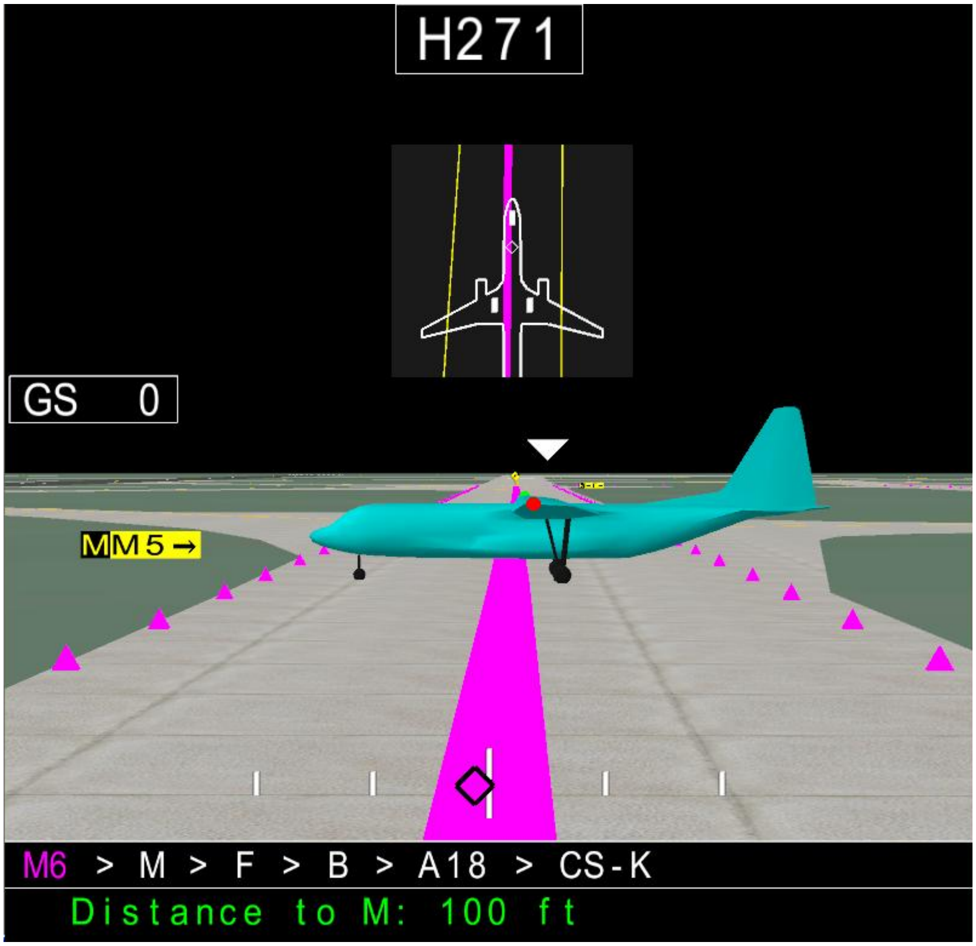

Figure 7. Displayed image for the HWD concept.

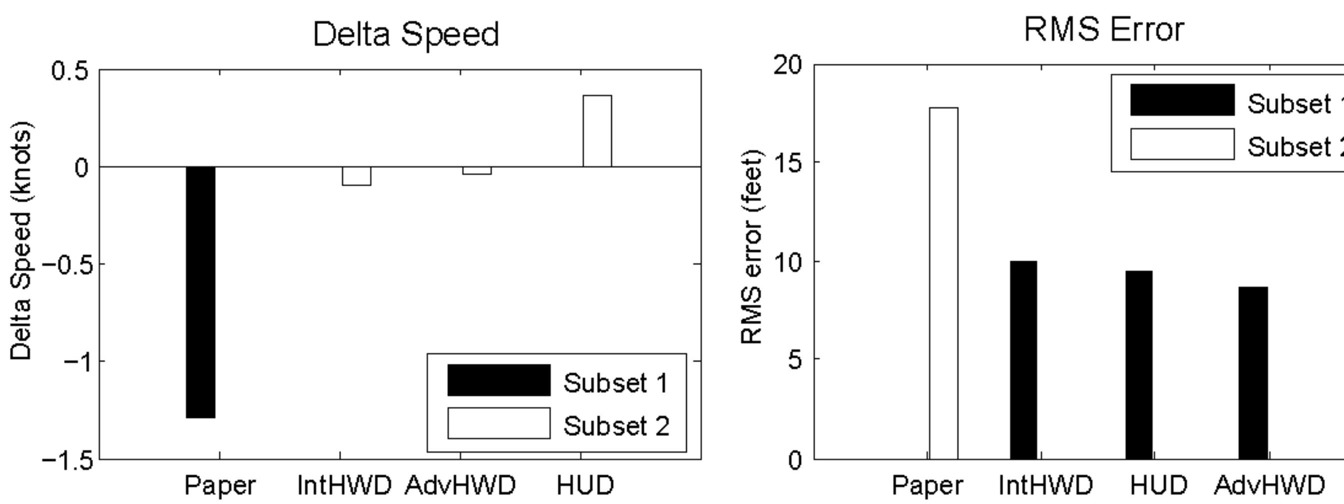

Figure 8. Speed and RMS path error per display concept.

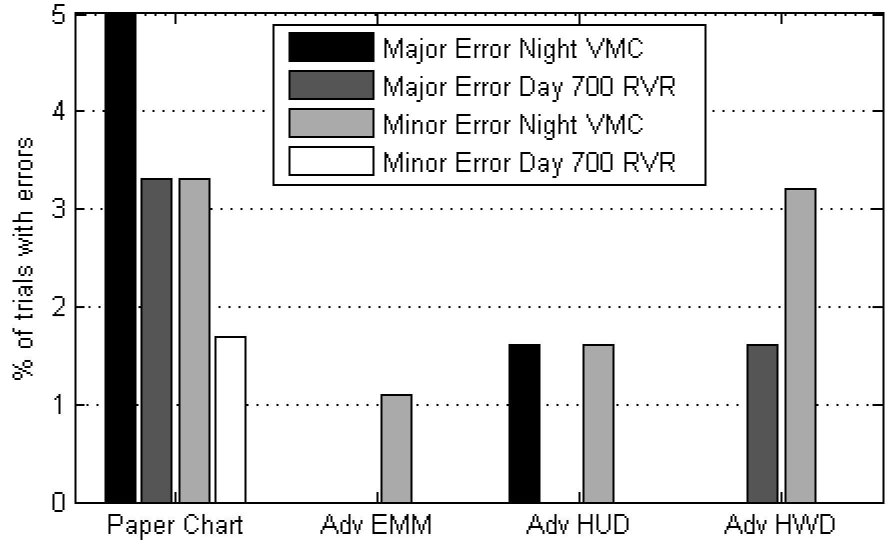

Figure 9. Navigational errors made per display concept.

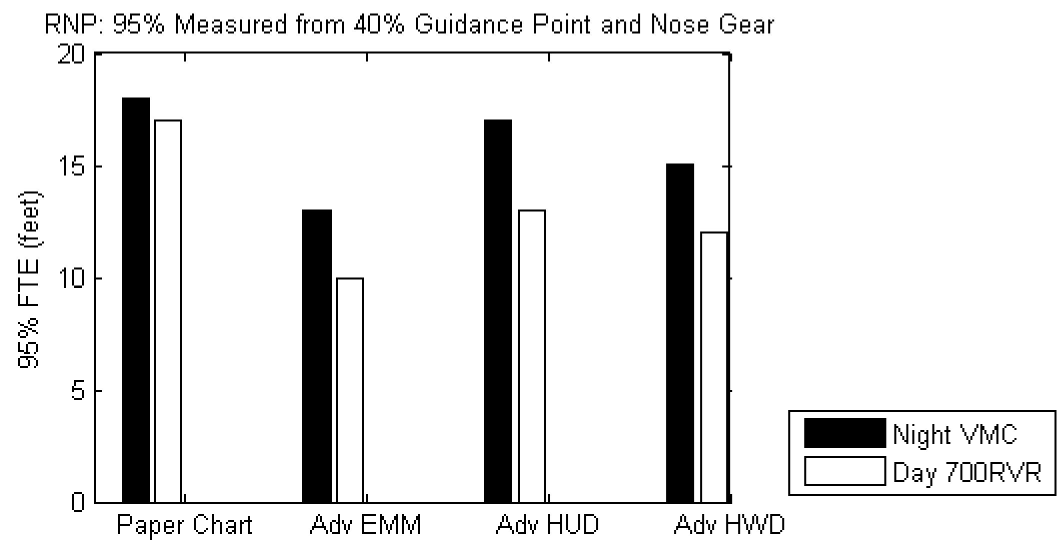

Figure 10. Surface RNP.

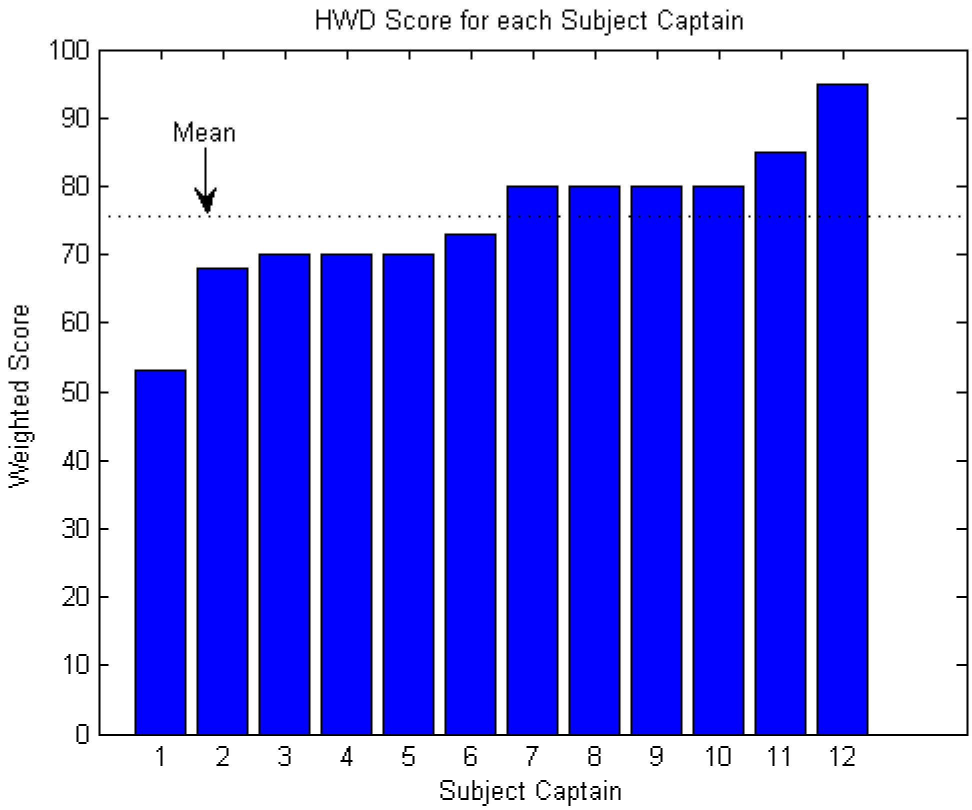

Figure 11. HWD scores for Usability Questionnaire.

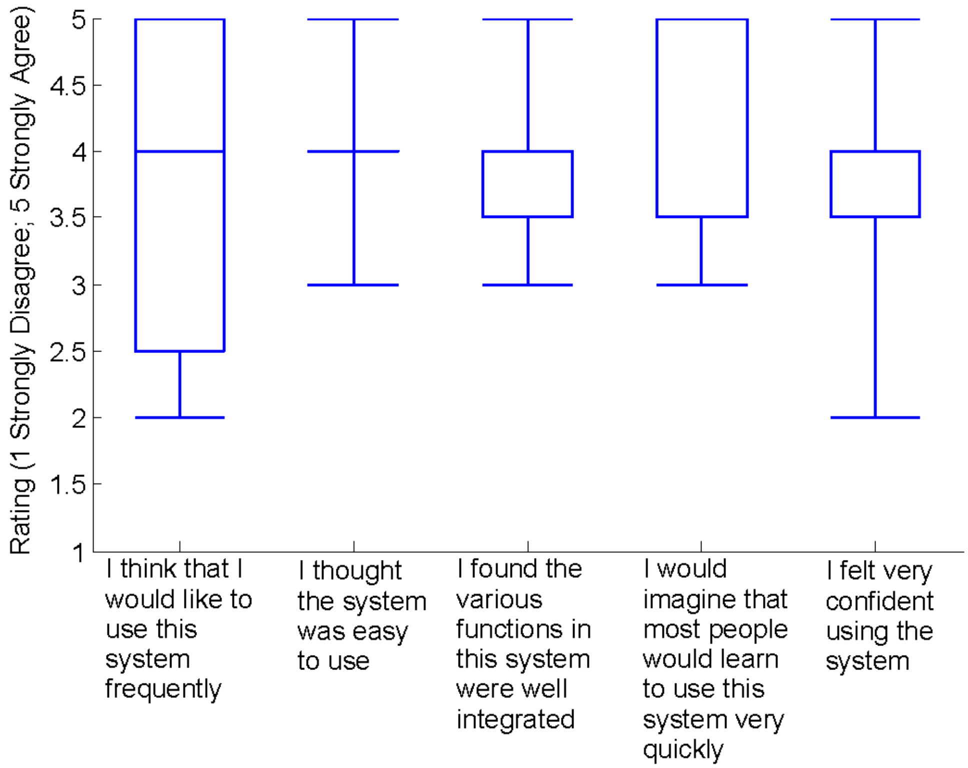

Figure 12. Subject captains' responses to positive statements.

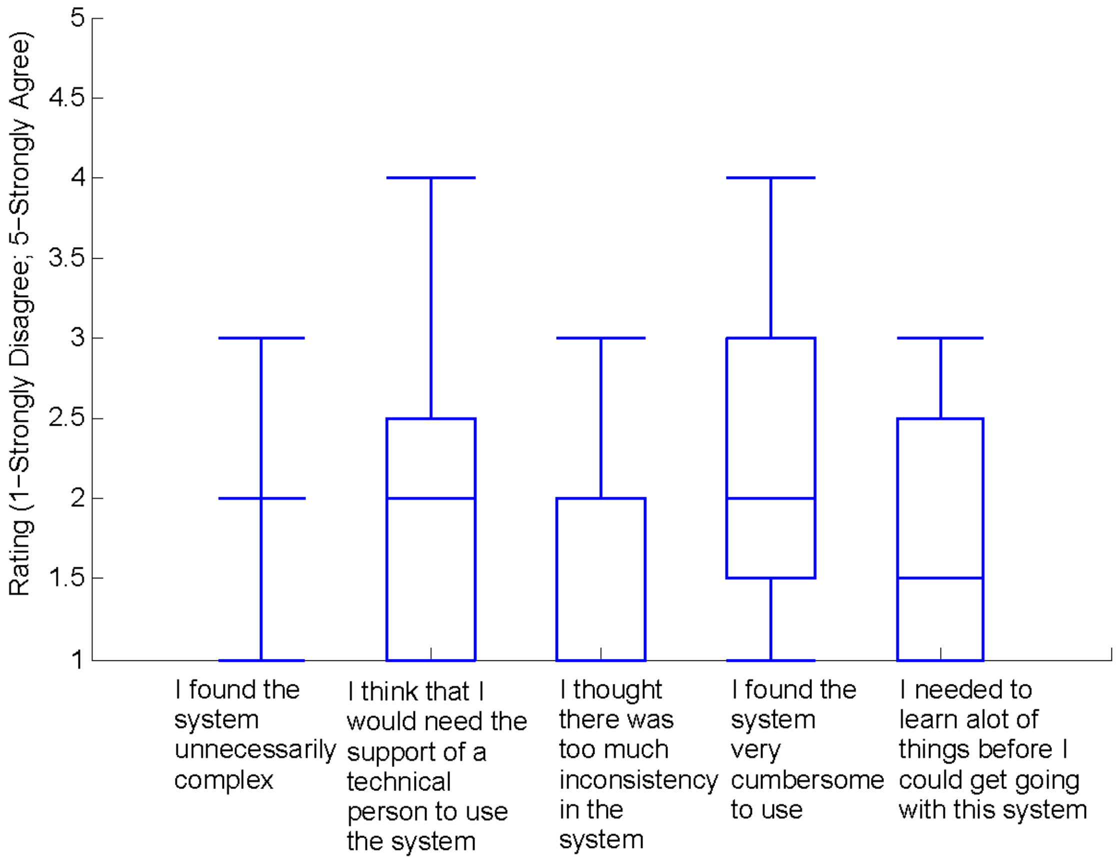

Figure 13. Subject captains' responses for negative questions.