Abstract

Precision navigation, display, and avionics technologies have progressed to the point that a head-up primary flight display incorporating synthetic elements such as terrain and commanded flight path is a possibility in the near future. The goal of such a display is to increase situation awareness and reduce aircrew reliance on warning systems or automation to prevent controlled flight into terrain (CFIT) mishaps. The primary flight display – and primary focus of attention – in modern fighter aircraft is a head-up display (HUD). This is becoming true also for newer transports in the U.S. Air Force inventory.

Some human factors issues associated with synthetic vision in a head-down display are different from those associated with a head-up synthetic vision display, especially when the displays are used as primary flight references. Among these issues are the use of color, ability to see through the display, symbology clutter, compatibility between head-up and head-down displays, and attentional factors. This paper reports the results of a study in which HUD-experienced pilots flew simulated complex precision approaches to landing in three visibility conditions, with and without synthetic terrain, using either pathway-in-the-sky symbology or more traditional military standard HUD symbology. Workload and situation awareness measures were collected to determine the relative workload associated with these conditions and if, as has been proposed elsewhere, flying a pathway-in-the-sky display is associated with "cognitive capture", or a decrease in situation awareness concerning things other than the pathway. It was hypothesized that including pathway and synthetic terrain in a head-up primary flight display would result in a conformal symbology set that naturally draws pilots’ attention to external events. It was also hypothesized that workload could be reduced by allowing pilots to maintain spatial orientation via preattentive processes rather than relying on instruments requiring focal vision and active interpretation.

Introduction

Controlled Flight Into Terrain (CFIT) accidents continue to be a major source of fatalities and airframe losses in both military and civil aviation, despite on-board warning systems [1,2]. Examination of evidence from the USAF Safety Center reveals that a causal factor in over half of these accidents is related to some deficit in situation awareness (SA) [1]. Current on-board warning systems (e.g., the Ground Proximity Warning System) have reduced CFIT dramatically, but are designed to prevent disaster once it becomes imminent (i.e., once the system detects that an aircraft flight path will be below some minimum safe altitude). What is needed are systems to improve pilot SA to the point where the need for warnings of imminent disaster are greatly reduced or even eliminated. Both enhanced (sensor-driven) and synthetic (database-driven) vision systems have been proposed as means to enhance pilot SA and reduce accidents caused by SA deficits. Synthetic vision systems have several features that strictly sensor-based systems do not. Among these are infinite field of regard, field of view limited only by display hardware, range up to (and even beyond) the visible horizon, and ability to portray surrounding terrain regardless of atmospheric conditions. Synthetic vision systems are only as reliable as the database, navigation, and display subsystems upon which they are built, but even with these limitations they seem a useful adjunct to traditional navigation aids and sensor systems. Research in the Air Force Research Laboratory has demonstrated substantial increases in pilot SA with the addition of synthetic terrain to a simulated HUD.
Further, this increase in SA was associated with a reduction in ground impacts during the low-level ingress scenarios used in these simulations. An example of the synthetic terrain used in this study is shown in Figure 1. The grid format shown in this figure was associated with the largest increase in SA overall (i.e., in Instrument Meteorological Conditions (IMC) both night and day) and is the format that was used in the current study. This format is also easily distinguishable from sensor imagery and therefore suitable for use in displays combining sensor and database imagery.

![Figure 1. Grid Format Synthetic Terrain.](image)

Previous research in this same laboratory [4] has shown that a pathway-in-the-sky similar to the Command Flight Path Display tested in the 1980s [5,6] allows pilots to fly complex precision approaches in IMC at night roughly twice as well compared to the current military standard head-up display (MIL-STD HUD) found in MIL-STD 1787C [7]. Pilots using the pathway were able to maintain the commanded path with roughly half the error in airspeed, lateral, and vertical deviation. Such complex approaches, with multiple curves and descent rates, may become common in the next decade as augmentation systems to the current GPS signal, both military and civilian, become operational. Later research showed that the benefit of a pathway HUD in landing these approaches did not vary with visibility in three daylight conditions, but did not directly compare the pathway to the MIL-STD HUD [8]. While clear differences have been found between pathway and MIL-STD HUD symbology with regard to performance in flying and landing complex approaches, the differential effects of these two symbology sets on situation awareness and workload have yet to be measured. Measurement of SA and workload is especially important in the current context for two reasons: 1) control/response ratio considerations, and, 2) attention management concerns.

Control/response (C/R) ratio refers to the ratio of control movement needed to achieve a given system or display response [9]. A low C/R ratio implies a great deal of display movement with little control movement (i.e., high gain) while a high C/R ratios implies the opposite (i.e., low gain). The C/R ratio for a pathway display is essentially determined by the width of the path displayed. Current GPS approaches call for tolerances as tight as ±0.15 nautical miles or roughly 1,800 feet. However, path dimensions in the literature on this topic vary widely. Theunissen has tested tunnels of roughly 75, 150, and 300 feet in width[10]; Snow and Reising used a path width of 400 feet[3]. Williams [11] used a path width of 600 feet, as did the current study, roughly simulating an approach with a Required Navigation Performance (RNP) of 0.05 nautical miles. Several studies have shown that path and tunnel displays result in less flight technical error relative to traditional flight displays, but the question becomes, “At what cost?” Is this increased precision associated with higher workload? How tight a tightrope should pilots have to walk? Putting a path or tunnel display in the cockpit is likely to be counterproductive if it means that the path is so busy maintaining the path that no time or attention can be spared for anything else. It is critical that flight technical error, situation awareness, and workload be measured in conjunction.

With regard to attention management, there is reason to be concerned about cognitive capture or attentional tunneling: the possibility that a synthetic vision display, especially one including a pathway or tunnel, will be so compelling and contain such a large proportion of the pilot’s information requirements that awareness of other displays and events will deteriorate in comparison. Some authors report no difference in situation awareness on-path versus off-path [3], while others report a decrease in awareness of events and information not contained in the synthetic vision display [11,12]. To date, most of these studies have been done with head-down displays combining sensor and database imagery.
displays and the differences found (or not found) seem to be highly task-dependent (but see Fadden et al. [16]). One purpose of the current study was to see whether using a head-up display for synthetic vision, and presumably focusing pilots' attention head-up rather than head-down would alleviate such concerns.

Method

Participants

Thirteen pilots volunteered to participate in the study. All were Air Force pilots with HUD experience. Pilot experience ranged from 1700 hours to 15000 hours with an average of 4819 hours. All pilots were male. Pilots ranged in age from 30 to 53 with an average of 39.

Experimental Design

The study used a 3 x 2 x 2 within-subjects design. However, there were only ten experimental conditions: no trials were run in the VMC Day condition with synthetic terrain.

Independent Variables. The independent variables manipulated in the study were, 1) visibility condition (Visual Meteorological Conditions (VMC) Day, VMC Night, IMC Day), 2) primary flight display (MIL-STD HUD vs. Pathway), and, 3) synthetic terrain (Grid vs. None). The IMC Day condition consisted of ¼-mile visibility with a 100-foot ceiling (equivalent to ILS CAT II minimums). Figures 2 through 4 show the MIL-STD HUD in VMC Day, the Pathway in IMC Day, and the MIL-STD HUD in VMC Night with synthetic terrain, respectively. These figures show each symbology set from the same vantage point (short final) and aircraft state.

Dependent Variables. The dependent variables included flight technical error (lateral, vertical, and airspeed deviation) and situation awareness and workload measures. The two situation awareness measures used were the Situation Awareness Global Assessment Technique (SAGAT) and the Situation Awareness adaptation of the Subjective Workload Dominance technique (SA-SWORD). The former is an objective measure based on responses to task-relevant questions [13] while the latter is a subjective paired-comparison technique [14]. The SAGAT questions asked in the study are listed in Table 1.

Figure 2. MIL-STD HUD in VMC Day.

Figure 3. Pathway in IMC Day.

Figure 4. MIL-STD HUD in VMC Night with Synthetic Terrain.
Table 1. SAGAT Questions Asked.

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate your pitch.</td>
</tr>
<tr>
<td>Estimate your indicated airspeed.</td>
</tr>
<tr>
<td>Estimate your altitude AGL.</td>
</tr>
<tr>
<td>Estimate your barometric altitude.</td>
</tr>
<tr>
<td>Estimate your current bank angle.</td>
</tr>
<tr>
<td>Estimate the distance to your destination.</td>
</tr>
<tr>
<td>Estimate the bearing to your destination.</td>
</tr>
<tr>
<td>Estimate the number of your next waypoint.</td>
</tr>
<tr>
<td>Estimate your current heading.</td>
</tr>
<tr>
<td>Estimate your current descent angle.</td>
</tr>
<tr>
<td>Estimate your vertical velocity.</td>
</tr>
<tr>
<td>The terrain at 12 o'clock, midway to the horizon is: rising, descending, roughly flat, or water?</td>
</tr>
<tr>
<td>The terrain at 12 o'clock, at the horizon is: rising, descending, roughly flat, or water?</td>
</tr>
<tr>
<td>Does your CDM currently intersect sky, terrain, water, or the runway?</td>
</tr>
<tr>
<td>Estimate your drift angle.</td>
</tr>
<tr>
<td>Are you currently acceleratng, decelerating, or neither?</td>
</tr>
<tr>
<td>Estimate the bearing to the nearest terrain that is above your current altitude.</td>
</tr>
<tr>
<td>Estimate how much you are currently above or below your commanded altitude.</td>
</tr>
<tr>
<td>Estimate how far left or right you are from the centerline of your commanded path.</td>
</tr>
<tr>
<td>What will your commanded altitude be ten seconds from now?</td>
</tr>
<tr>
<td>Estimate the descent angle of your commanded flight path ten seconds from now.</td>
</tr>
<tr>
<td>Estimate your commanded heading ten seconds from now.</td>
</tr>
<tr>
<td>What direction, if any, will your commanded flight path turn in the next ten seconds (left, right, or none)?</td>
</tr>
<tr>
<td>Locate all traffic, ground and air, currently present in the environment.</td>
</tr>
<tr>
<td>Estimate distance to the nearest traffic.</td>
</tr>
<tr>
<td>Estimate the bearing to the nearest traffic.</td>
</tr>
</tbody>
</table>

The two workload measures used were NASA TLX (NASA Task Load Index) and SWORD (Subjective Workload Dominance technique). The former is a rating technique with six subscales while the latter is a paired-comparison technique. Responses to SAGAT questions and NASA TLX ratings were taken during each approach while SWORD and SA-SWORD measures were taken at the conclusion of the experiment.

**Procedure**

Participants received an introductory brief, simulator and symbology familiarization, and then flew ten practice approaches, once in each experimental condition. Data collection then proceeded with participants flying ten different approaches twice each (a total of twenty approaches during the data collection phase). Each experimental condition was replicated twice in random order. The simulation was briefly interrupted once at a random interval during each approach to ask SAGAT questions and take NASA TLX ratings. Two to four F-16s were present in the airport environment during each approach with the final (twentieth) approach differing from the rest in that one of the F-16s was stationed on the runway at the touchdown point. This last approach was always conducted in one of the two VMC conditions. Upon completion of all approaches, participants rated their workload and situation awareness using the SWORD and SA-SWORD techniques, and filled out a questionnaire to solicit subjective opinions concerning the conditions and symbology sets flown.

**Apparatus**

The cockpit used was a fixed-based simulation of a generic fighter/attack aircraft using an F-16 aeromodel. Primary control inputs consisted of throttle, force stick, touch screen, and gear handle. The out-the-window scene was viewed on three projectors surrounding the cockpit with a total lateral field of view of 110° and a vertical field of view of 30°. HUD symbology was superimposed on the center projection screen with a field of view of 30° lateral by 20° vertical. A picture of the simulator is shown in Figure 5. Head-down instruments included a moving map and a traditional primary flight display (Attitude Director.
Indicator, Horizontal Situation Indicator, Airspeed and Altitude Indicators). Participants were given an instrument approach procedure to be studied prior to each approach, which they then placed on a kneeboard for reference throughout the approach. A sample approach procedure is shown in Figure 6.

Results

**Flight Technical Error (FTE)**

Flight Technical Error data collected were the airspeed, lateral, and vertical deviations from commanded values. These data were then broken into two groups based on occurrence of a secondary task: non-distracter and distracter. Root Mean Square Error (RMSE) was calculated for statistical analysis. Also calculated from the raw FTE data were the percentages of time spent outside the commanded corridor vertically and laterally.

All of the non-distracter FTE means were significantly lower (α = .05) for the pathway condition. Figure 7 shows the means for lateral deviation RMSE and is representative of the other FTE measures (error bars in all figures represent 90% confidence intervals).

Percent time spent off path was also significantly lower in the Pathway condition. For the main effect of the visibility variable, only mean differences in RMSE for airspeed (VMC Day vs. IMC Day) and percentage of time off-path vertically (VMC Day vs. VMC Night) were significantly lower (in the VMC Day and VMC Night conditions, respectively).

Figure 6. Sample Instrument Approach Procedure Used In The Study.

Figure 7. Effects Of Primary Flight Display, Visibility, And Synthetic Terrain On Lateral RMSE.
For the distracter data, all of the FTE means were also significantly lower in the Pathway condition (see Figure 8 for a representative graph).

![Figure 8. Effects Of Primary Flight Display, Visibility, And Synthetic Terrain On Lateral RMSE During A Secondary Task.](image)

For the visibility variable, no main effects were significant. However, for the PFD * Visibility interaction, a significant difference was found for RMSE vertical deviation. This difference occurred for both VMC Night and IMC Day vs. VMC Day. In both cases, the RMSE vertical deviation mean was lower for VMC Day in the Mil-Std. condition and higher for VMC Day in the Pathway condition.

![Image](image)

The performance of a secondary task resulted in significant differences in the means of absolute lateral deviation, airspeed RMSE, and all measures of percent time off path. Further, the interaction of PFD * Secondary Task also had a significant effect on lateral RMSE and all percent off path measures. Effects on total percent time off path are shown in Figure 9.

Synthetic terrain had no significant effect on any of the flight performance dependent variables.

A secondary analysis was conducted on FTE by adding Secondary Task to the model. Not surprisingly, results of this model reveal main effects for PFD and Secondary Task. Absolute lateral deviation, RMS airspeed deviation, and percent time off path (total, lateral, and vertical) were all significant for Secondary Task. Interestingly, the PFD * Secondary Task interaction was significant. Upon further examination, the means for absolute and RMS lateral deviation and percent time off path (total, lateral, and vertical) were significantly lower for pathway (see Figure 10 for a representative graph).

![Image](image)

Figure 10. Effects Of Primary Flight Display And Secondary Task On Lateral RMSE.

**Workload**

A repeated measures multivariate analysis of variance (MANOVA) was conducted for SWORD and NASA TLX Weighted Sum workload levels. Differences were significant for both measures for both primary flight display and visibility, with the
lower means (less workload) occurring in pathway conditions. For visibility, significantly lower SWORD means occurred under VMC Day conditions vs. VMC Night and IMC Day. The significantly lower TLX means only occurred under VMC Day when contrasted with IMC Day.

Synthetic terrain had a significant effect on the SWORD variable only, with workload rated lower when synthetic terrain was present. There was also a significant effect on SWORD ratings of the interaction between primary flight display and synthetic terrain, with synthetic terrain rated as reducing workload only in the MIL-STD HUD condition (see Figure 11). Corresponding NASA TLX data are depicted in Figure 12 (use same legend).

Situation Awareness

Included in the MANOVA above was the SA-SWORD variable. For situation awareness, the only main effects significant were primary flight display and visibility. The SA-SWORD scores had higher (more SA) means in pathway and VMC Day conditions. A significant interaction was found between primary flight display and synthetic terrain, similar to that found for SWORD data. Effects of independent variables on SA-SWORD are shown in Figure 13.

Figure 13. Effects Of Primary Flight Display, Visibility, And Synthetic Terrain On SA-SWORD Ratings.

The only significant effect on SAGAT scores was the effect of visibility on overall SAGAT score (i.e., total number of correct responses). A test of within-subject contrasts showed that this effect was due to a difference between VMC Day and the other two visibility conditions. SAGAT scores throughout the study were low (pilots typically answered between a third and half of the questions correctly) and it is difficult to know whether the lack of significant findings for SAGAT scores was due to low statistical power (typical observed power was 0.2), a floor effect, insensitivity of the measure, true invariance of SA as independent variables were manipulated, or some combination of these. Even when groups of related questions were analyzed together (e.g., terrain questions), results were still not significant, although trends were in expected directions. For example, of responses received to SAGAT terrain questions across all pilots, 57 were correct when synthetic terrain was present, versus 40 correct when it was not. Similarly, pilots answered 31 of these questions correctly in the VMC Day condition, but only 17 correct responses were received in the IMC Day condition without synthetic terrain.

Two other results of note were not analyzed statistically: incidence of CFIT, and response to
runway incursion. Of the 260 approaches flown during the data collection phase of the study (i.e., excluding practice), seven resulted in controlled flight into terrain. All of these occurred in the IMC Day condition without synthetic terrain. Six of the seven occurred when the pilot was using the MIL-STD HUD. The seventh occurred after the pilot had lost the pathway.

Software/simulation problems invalidated two of the runway incursion approaches (wingtip and tail lights on the simulated F-16 failed to operate during VMC Night). Of the remaining eleven runway incursion approaches, a correct response to the runway incursion (initiation of a missed approach) was observed on eight of these. In the three approaches in which a missed approach was not observed, pilots executed a normal landing on or very near the simulated F-16 at the touchdown point. Post-experiment questioning revealed that these pilots were totally unaware of the simulated runway incursion. All of the runway incursion approaches in which a missed approach was not initiated occurred in the VMC Night condition, using the MIL-STD HUD, without synthetic terrain.

Discussion

With respect to flight technical error, the results of this study replicate those of previous comparisons between traditional flight directors and pathway displays. Pilots are much better at maintaining the commanded flight path, including airspeed, when using a pathway-in-the-sky display. Indeed, pilots spent roughly four times as much time outside the commanded corridor when flying the MIL-STD HUD as they did when flying the Pathway. The absolute magnitude of these effects is very likely driven by what was, in retrospect, an extraordinarily challenging task: time spent learning the symbology sets, simulator flight control characteristics, and studying approach plates was much less than one would expect of an operational environment. Further, the paths flown represented something of a “worst case” in terms of corridor dimensions and were designed to test the limits of pilot ability in a precision navigation environment unhindered by the interception and tracking of radio navigation aids. While not a focus of the study, the results support the current practice of allowing pilots to fly a stabilized approach in IMC (one in which requirements for control inputs and configuration changes are minimized). Several pilots commented on the difficulty imposed by the variety and number of changes in descent angle in the approaches flown.

Despite the demanding task, pilots were able to successfully fly the approaches when using the pathway, even in solid IMC and even in the presence of distracter tasks. The workload and SA data provide insight into why this was the case. Better flight performance in the pathway condition was achieved — not at the expense of increased workload — but because use of the pathway reduced workload. Indeed, pilots commented that increased situation awareness regarding the upcoming path (and associated control inputs) allowed them to better manage secondary tasks. This is evident in the effects on flight performance of the interaction between primary flight display and secondary task: performance with the MIL-STD HUD worsened more with addition of a secondary task than did performance with the pathway. The NASA-TLX data show that workload decreased roughly 20% with use of the pathway, a decrease that may be practically significant. Experience with this metric indicates that a “redline”, a point at which performance begins to deteriorate significantly, may be around 50 (e.g., [15]). In the current study, pilots rated their workload near this value when using the pathway, but well beyond it when using the MIL-STD HUD.

While the results with respect to primary flight display are not unexpected, the results with respect to synthetic terrain do contain surprises. Synthetic terrain appeared to have no effect on night performance and selected only the most sensitive measures of workload and SA (SWORD and SA-SWORD), and these only in the MIL-STD HUD condition. As described previously, analyses of SAGAT results were problematic, but these data did trend toward an increase in terrain SA with the inclusion of synthetic terrain. While caution should be exercised in basing any conclusions on such a small number of events, the fact that no CFITs occurred when synthetic terrain was present is promising.

The subjective questionnaire asked pilots to rate the usefulness of the MIL-STD HUD, the pathway, and synthetic terrain on a scale from -3 to

11.C.5-8
+3, with -3 being “Extremely Low” and +3 being “Extremely High”. The average ratings were -1.2, 2.4, and 1.2, respectively. A majority of pilots commented that the synthetic terrain would be most useful in IMC or at night. Five of the thirteen commented that they would want control of the brightness or contrast for the synthetic terrain (to include decluttering it entirely), especially on short final. Several of the pilots would have decluttered the follow-me aircraft if they had had the option, and there were several suggestions for adding reference markers (e.g., airspeed, altitude, and heading “bugs”) to the pathway display to support awareness of basic approach parameters.

While pilots’ ratings of pathway utility were quite high, two of the thirteen made comments that indicated a concern relative to non-pathway SA:

Pilot A: “Very useful in helping the pilot predict where the flight path of the aircraft would be relative to current ownship attitude and heading. However, SA on actual path segment headings or commanded altitudes was not high. Could get complacent "following the path", leading to pilots not maintaining overall orientation to the approach.”

Pilot B: “The pathway format reduced my workload greatly once I acquired a better understanding of the system. Easy to interpret the displayed information. The danger I see here is a pilot can very easily tune out the world around him while totally focusing on the pathway.”

While these comments would support a hypothesis of cognitive capture (or attentional tunneling) associated with the pathway, they are somewhat belied by the runway incursion results: none of the incorrect responses to the runway incursion occurred in the pathway condition. Rather, the results support an alternate hypothesis that – even for an unexpected event – the conformal nature of pathway HUD symbology (especially in combination with a synthetic runway outline) and its head-up location facilitate SA, at least relative to events in the far domain overlaid by the symbology [16, 17].

Conclusions

As applied research comes closer to actual application, care must be taken in overgeneralizing results. Among things to bear in mind in interpreting the results of this study and related studies that have been published in recent years is that pilots were flying a part-task simulation. Pilot performance, workload, and SA are greatly influenced by several factors common in real aviation environments that are not common in part-task simulations. Important among these are communications with air traffic management, the presence of other crewmembers, and physical stresses (e.g., G, cold/heat, turbulence).

However, with respect to flight technical error, the results of this study replicate those of previous comparisons between traditional flight directors and pathway displays both in simulation and in flight [10, 16]. In comparison to standard 2D flight directors, pilots are much better at maintaining a commanded path when using a pathway display and it seems likely that this effect is magnified with increasing path complexity. When commanded corridor dimensions (and associated C/R ratios) are held constant, workload is reduced by use of a pathway display and situation awareness is increased. The increase in awareness of future path-related events, especially those associated with control inputs, reduces pilot workload and allows better management of secondary tasks.

The current study certainly does not rule out the phenomenon of cognitive capture or attentional tunneling associated with pathway displays. However, it does support a hypothesis that any such detriment can be alleviated via the placement of conformal symbology in a head-up location. Extrapolating, concerns about decreased traffic awareness associated with pathway use should not be addressed by using symbology other than a pathway, but by including conformal HUD overlays highlighting traffic. In pursuing this strategy, research is needed concerning trade-offs between conformal overlays and clutter. Given all the potentially useful database-based information that could be displayed to a pilot (e.g., terrain, traffic, flight path, atmospheric phenomena, airspace boundaries), there is the potential to render a HUD informationally opaque and not useful for its intended purpose.

References

Symposium on Aviation Psychology. Columbus, OH: Ohio State University.


