Evaluation of a Non-Traditional Aircraft Attitude Indicator

Robert J. Poisson III, Michael E. Miller, Michael W. Haas
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433

Henry P. Williams
Naval AeroMedical Research Unit – Dayton
Wright-Patterson AFB, OH 45433

Abstract
Aviation mishaps involving spatial disorientation (SD) have cost the U.S. Air Force over $2B in the past two decades. A non-traditional attitude display, the attitude stabilization display (ASD) has been proposed which may alleviate concerns with the current attitude indicator (AI) and mitigate the risks of spatial disorientation. Participants used both the proposed and current designs to recover from unusual attitudes in a desktop flight simulation. Participants completed recovery tasks approximately 2 seconds faster with the AI, on average. There was a significant difference indicating that participants also found it easier to learn how to use the AI. There was a significant effect of flight experience on recovery time difference, with more experienced pilots performing better with the AI and less experienced pilots performing better with the ASD. Since the majority of participants already had experience with the AI, these results were expected. Survey responses revealed that certain ASD design choices could be beneficial in the cockpit. Since this study did not measure the full intent of the ASD, which is to aid the pilot during SD inception and avoid SD altogether, further investigation of the ASD is warranted.

Keywords
Spatial Disorientation, Attitude Indicator, Human Factors, Flight Simulator

Disclaimer
The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of the Department of the Air Force, Department of the Navy, Department of Defense, nor the U.S. Government.

1. Introduction
Since the advent of air travel, aircraft pilots have experienced Spatial Disorientation (SD), in which the pilot’s perception of aircraft position, motion, or attitude does not correspond to reality [1]. When suffering from SD, pilots naturally tend to make aircraft inputs and controls that may create safe flight in their perceived orientation, but result in unsafe flight in reality. These inputs often cause the aircraft to enter unusual attitudes (UAAs) which may include unperceived inversions, steep climbs, and sharp dives. These unusual attitudes brought on by SD thus immensely increase the risk of a mishap. Across the U.S. Air Force, SD mishaps are both prevalent and costly. In fact, 72 spatial disorientation (SD) Class A mishaps have occurred in the Air Force since fiscal year 1993 which resulted in the loss of 101 lives and 65 aircraft for a total cost of $2.32 billion [2].

Pilots often use displays and instruments in the cockpit to determine their orientation when a view of the outside world is degraded by weather, darkness, or a perceived visual illusion. Particularly when suffering from SD, pilots are instructed to focus only on their instruments to discern their aircraft’s attitude. The first instrument to combat SD was an attitude indicator (AI) known as the Sperry Horizon, originally developed in 1928 by Elmer Sperry Jr. of the Sperry Corporation [3]. Since that time, despite some known human factors and training issues, this attitude instrument and display has become standard in most instrumented aircraft cockpits [4] and is generally replicated in electronic form within even the most modern aircraft cockpits. However, this instrument may not effectively combat spatial disorientation since mishaps involving spatial disorientation continue to occur across all forms of flight [2].

Through this research, we will seek to understand the performance of a proposed attitude display. This proposed system is termed an Attitude Stabilization Display (ASD). The ASD differs in three significant fashions from the Sperry-style AI. First, it draws the pilot’s attention by way of an auditory alert when it determines that the aircraft is entering an unexpected attitude, indicating the potential onset of SD. Second, the display employs a potentially more intuitive graphical interface (explained later) to aid pilots in determining their attitude. Finally, the ASD provides a
specific, recommended course of action to guide the pilot towards returning the aircraft to the expected attitude once it has detected the presence of the unexpected attitude. While each of these differences are intended as improvements to the Sperry-style AI, this research will focus on only the second intended improvement, comparing the graphical depiction in the ASD to the traditional AI.

2. Literature Review

2.1 Spatial Disorientation
SD is typically categorized based on the pilot’s response. Specifically, a pilot can recognize, not recognize, or become incapacitated by SD. Type I, or unrecognized, SD occurs when pilots do not realize that they are suffering from SD and fly the aircraft in an unintended attitude. Typically, Type I SD results in either a controlled flight into terrain (CFIT) or a transition to Type II SD. Recognized, or Type II, SD comes into existence when pilots recognize that they are spatially disoriented. At this stage, Type II SD typically results in a recovery and regaining of spatial orientation or a transition to Type III SD. If pilots are unable to handle the realization that they are suffering from SD and are thus unable to match their perception of motion, position, and attitude to reality (i.e. recover), this is classified as Type III SD, or incapacitating, SD [5].

Often, the vestibular system of the inner ear is to blame for SD episodes. The semicircular canals and otolith organs, which make up the vestibular system, are sensitive only to acceleration, not to sustained movement. Therefore, after sustaining a constant turn for approximately 10-15 seconds, an aircraft pilots’ vestibular organs begin to relay sensory signals which are consistent with straight and level flight, while the aircraft is continuing to turn [6]. Several other imperfections in the vestibular system can cause issues in flight. The utricle (one of the two otolith organs), for example, cannot distinguish between a tilting of the head and a linear acceleration. Therefore, under sustained forward acceleration, the utricle will provide the same signals to the brain that it would if the head was tilted backward under no acceleration [6]. Thus, the pilot may mistakenly perceive forward acceleration of his aircraft as an upward pitch (i.e. a backward tilt of the body/head/aircraft) and mistakenly pitch the aircraft down while in straight and level flight.

2.2 Auditory Alarms
The auditory alert employed by the ASD was developed in recognition of the fact that the operational concept behind the Sperry-style AI is flawed. Specifically, its weakness is that it requires the pilot to periodically focus visual attention on the instrument to determine if their perception of attitude is correct. However, focusing on this instrument is a non-intuitive action for pilots because even if type I SD has set in, they have no reason to believe that their perceived attitude is false. Thus, there is no reason to ensure that it is correct. Therefore the traditional AI violates Norman’s design principle of feedback since pilots have to actively seek feedback from the control movements that they input instead of feedback being provided to them in a way that is cognitively simple to perceive [7].

By permitting the pilot to communicate expected flight parameters to the system, the ASD automatically monitors the attitude of the aircraft and provides auditory alerts to the pilot whenever the attitude of the aircraft is outside the pilot’s expected flight parameters. These auditory alerts plausibly allow pilots to spend less time visually scanning their instruments and more time with their eyes outside of the cockpit, ensuring that their airspace is clear of hazards. Thus, the non-intuitive check of the traditional AI to ensure that a pilot is not suffering from SD is alleviated. This change may improve the pilot’s ability to become aware of spatial disorientation (i.e. transition quickly from type I to type II SD, or skip type I SD entirely) before it becomes a significant issue.

Aside from the common experiential knowledge that auditory alarms tend to capture our attention, there is some scholarly work on the subject. A primary advantage of auditory alarms over visual ones is that when we focus our visual attention, we typically see one specific item very clearly while our visual perception of non-attended items suffers. The auditory sense is quite different in that it is not as easily focused. As a result, we tend to hear certain auditory alerts even when we are not attending to them [8]. Therefore, human factors guidance often recommends that “if there is an alarm signal that must be sensed…it should be given an auditory form (although redundancy in the visual or tactile channel may be worthwhile)” [9].

Unfortunately, this capture of attention can be undesirable. Alarms, which are intended to immediately induce focus from a pilot, may inadvertently disrupt their cognitive processing, distract them, and steal their attention from a potentially more important stimulus [8, 10]. This issue occurs in the cockpit as a result of the ever-increasing number of ad-hoc auditory alarms and signals being implemented [11]. It is therefore possible that while the ASD’s auditory alarm may effectively capture the attention of pilots suffering from SD, it may also contribute to their confusion during times where many different auditory alarms may be sounding.
2.3 Command Displays
In addition to the auditory alert, the ASD employs a visual command to the pilot (e.g. “pull up”), which informs him or her of the correct action to initiate return to straight and level flight. There have been a number of robust research efforts which compared status displays, which simply provide an alert that something has gone wrong, and command displays, which additionally provide information about actions that must be taken. The underlying theory is that decision making is split into three basic steps, “(1) acquiring and perceiving information or cues relevant for the decision, (2) generating and selecting hypotheses of situation assessments about what the cues mean,…[and] (3) planning and selecting choices to take” [9].

It has been hypothesized that command displays, such as the one found in the ASD significantly reduce or eliminate the time and cognitive effort needed to perform steps two and three [12]. The claim is that in high stress situations, such as an in-flight emergency, pilots experience a high temporal and cognitive demand. Therefore, the automation of this process can aid the pilot in returning his or her aircraft to the desired orientation. Recognition-primed decision making could be happening when recovering from UAs if the pilot in question has been in similar situations [13]. As discussed above, the cognitive effort required to access long-term memory and compare the current situation to past experiences can take some time to perform. Command displays attempt to bypass that time by providing pilots with a decision instead of waiting for them to make their own.

These hypotheses were empirically tested using pilot response to simulated in-flight icing of an aircraft. With the participation of 27 commercial pilots from the University of Illinois, pilot response time and accuracy to the first indication of icing when using either a status or command display was measured [12]. Additionally, the accuracy of information provided was manipulated to determine any effects of pilot trust or distrust in automation. A lack of reliability of automation can result in the user distrusting the automation. On the other hand, a very high reliability may cause the user to become complacent and not check the work of the automation. As a final caveat, humans are so unpredictable that they may display some form of mistrust, in which their trust level of the automation is not related to reliability at all [9].

The University of Illinois study found that pilots using the command display trended towards better performance in terms of response accuracy, though there was no significant effect of display type on response time. However, the most interesting results were the interactions between display type and information. Inaccurate information was linked to a much larger performance decrement in command displays than it was in status displays [12]. The experimenters appear to have validated their hypothesis that command aids help to eliminate decision-making steps for the pilot. The larger performance decrement seems to indicate that pilots are more likely to blindly follow the instructions of the command display while cognitively analyzing the status display before acting. Ostensibly, this blind following saves vital seconds in response time. Clearly, though, if the wrong instructions are presented to a blindly obedient pilot, the results may be catastrophic.

Importantly, in the realm of manned aviation, it has been shown that the use of aural commands may have the capability to dramatically aid the pilot in recovering from unusual attitudes. In a 2008 experiment, 12 U. S. Air Force fighter pilots were presented with unusual attitudes in an F-16 flight simulator. Experimental conditions varied the presence of certain attitude display aids, with the control condition utilizing only a standard heads-up display (HUD) and other experimental conditions using a command visual icon, the icon and an auditory command, or the icon and a tactile command. When pilots were given the auditory command aid, they were approximately 15% faster in leveling their wings under a moderate inversion (approximately 120° of roll, and varying pitch angles), and approximately 20% faster when under a severe inversion (approaching 180° of roll, and varying pitch angles). Additionally, pilots input one quarter the number of incorrect control movements when using the auditory aid than when using the HUD only. Subjectively, 80% of the pilots who indicated preferring one aid over another selected the auditory commands as their most preferred aid [14].

2.4 Attitude Indicator Graphical Layouts
The ASD’s graphical interface is also a point of interest. First, it employs a moving-aircraft symbol, stationary horizon (also known as outside-in) construct instead of the moving-horizon, stationary aircraft (also known as inside-out) construct of the Sperry-style AI. The selection between these two structures has been hotly debated since before the Sperry Horizon was patented. In support of the moving-aircraft displays, the principle of the moving part is often cited. According to this principle, the best displays employ movement in a manner that accurately and intuitively represents that movement in reality. When the principle of the moving part is applied to the aviation
domain, “one might say that when a pilot moves a control, he knows he is controlling his aircraft, not the outside world relative to his aircraft, and therefore he expects his aircraft symbols to move” [4]. Thus, this principle theoretically favors a moving-aircraft AI.

Interestingly, the argument is not entirely theoretical. There is a substantial body of research which indicates that inexperienced pilots learn to use the moving-aircraft display more quickly and that experienced pilots quickly achieve higher levels of performance when transitioning to the moving-aircraft display. In fact, in 1960, Donald Bauerschmidt and Stanley Roscoe simulated an air-to-air attack task and compared pilot performance on the two display types. Average steering errors calculated at the end of the task with the moving-aircraft display were approximately one fifth the size of those calculated with the moving-horizon display. Additionally, the pilots made approximately 18 times the number of control reversals when using the moving-horizon display as they did when using the moving-aircraft display. Perhaps the most intriguing discussion point is that all of these results were found despite the fact that all participants’ flight experience had included the traditional moving-horizon display [15].

The debate is not one-sided, however, and there are many advocates of the moving-horizon AI. Nearly all of them discredit the results of any experiment performed on the ground because the utility of the moving-horizon display, they claim, is only achieved in actual flight [16]. The validity of ground-based results can certainly be called into question when researching a realm where airborne accelerations and the vestibular cues that they provide will no doubt influence the pilots’ perception of their orientation. In support of this, many cite Col. James Doolittle, who influenced the design of the Sperry Horizon. Doolittle claimed that the pilot and the aircraft function as one, and the pilot’s main frame of reference is indeed the aircraft [16]. This concept can be corroborated nearly verbatim in modern literature [6]. With this in mind, Doolittle claimed that the real aircraft never moves with respect to the pilot, so it makes no sense that the display’s aircraft symbol should move with respect to the pilot, and thus he requested that the Sperry Horizon employ a moving-horizon, stationary-aircraft construct [16].

Putting the debate to the test, the Federal Aviation Administration (FAA) took to the sky with each of 32 FAA-certified male pilots, a Beech T-34 military trainer, and a safety pilot. They performed an airborne experiment which compared the two types of displays in their ability to aid the pilot in recovering from UAs. In terms of bank angle recovery, pitch angle recovery, and number of control reversals, there were no overall trends that indicated either the moving-aircraft or moving-horizon indicator was superior. In general, it appeared that low-experience pilots tended towards better performance with the moving-aircraft AI while high-experience pilots tended towards better performance with the moving-horizon display. This effect held true when measuring the number of control reversals, with both groups performing at about the same level when using the moving-aircraft AI [16].

2.5 Literature Summary

Based on this literature review, the ASD may offer a significant benefit in terms of enabling recovery from and/or preventing SD. First, it appears that command displays, such as those utilized in the ASD, in comparison to status displays utilized in traditional AIs, may decrease the time necessary for pilots to recover from UAs. Additionally, it is widely held that auditory alerts such as those employed by the ASD, are more effective at capturing attention than are visual signifiers such as those passively displayed by a traditional AI. Finally, the ASD’s moving-aircraft display, when compared to a traditional moving-horizon AI, has the potential to be effective in decreasing the time needed to respond to UAs and in decreasing the number of control reversals during recovery from them. Thus, the ASD merits further investigation and analysis.

While the combination of the ASD’s attributes is interesting, the current research was focused to understand the effect of the graphical depiction of aircraft attitude in the ASD as compared to the traditional AI. This limitation was due to unforeseen issues with auditory command lagging and the unavailability of a moving-based simulator. However, it is likely that the other attributes of the ASD, either singly or in combination will have benefit beyond those investigated within the current experiment. Additionally, this study is intended to contribute to the general body of knowledge of AIs. Thus, through data analysis and a survey process, this study will unfold the utility of certain differences between the AI and the ASD. In so doing, the goal is to determine why various aspects of the ASD may or may not be beneficial to pilots.
3. Method

3.1 Overview
A desktop computer based, non-moving flight simulator was used to compare the graphical depiction of the ASD to the traditional AI. Participants had a variety of flight experience levels and were asked to recover from already in-progress unusual attitudes in the flight simulation. Metrics included the number of control errors, the time to complete recovery, and the root mean square error from perfect recovery.

3.2 Participants
Participants were 28 male Wright-Patterson Air Force Base personnel, ranging in age from 21 to 65 years with a mean of 30. Previous flight experience ranged from 0 to 5000 flight hours with a mean of 600 hours, and 0 to 2000 unmanned flight hours (including flight simulator, remotely piloted aircraft, and radio control aircraft) with a mean of 263 hours. Of the 28 participants, 6 were instrument-rated pilots, 8 had experienced SD in flight, and 9 had undergone SD training. For the purpose of data analysis, participants were categorized based on their flight experience levels. There were 5 “experienced” pilots who had over 1000 hours of flight time, 9 “unmanned only” pilots who had no manned flight time, and 5 “total novice” participants who had no manned or unmanned flight experience. These were three separate binary categories with all 28 participants being categorized three times as either a member or non-member of each category.

3.3 Apparatus
Flight simulation took place on a Hewlett-Packard Z820 workstation running X-plane 10 Professional on a 30” Samsung Syncmaster 30ST monitor. Manipulation of the simulated aircraft was accomplished with a Saitek X-52 joystick and throttle combination. The software simulated an F-22 Raptor flying at 450 knots at 20,000 feet above ground level.

The ASD is shown in the left side of Figure 1, depicting a descending left turn. To interpret the display, participants were instructed to concern themselves only with pitch and bank. To determine their pitch, participants used the vertical scale in the center of the display. The green upside-down “V” symbol represented the nose of their aircraft and the thick white bar represented the horizon. Thus, when the “V” was above the white bar, they were pitched up, and if the “V” was below the white bar, they were pitched down. To determine their bank, participants used the rounded scale occupying the uppermost portion of the display. The white aircraft symbol represented their aircraft. When this symbol was at the top of the rounded scale, the aircraft was straight and level. As participants banked left or right, the symbol would slide along the scale to the left or right, respectively.

![Figure 1: ASD (left) and AI (right) Depicting a Descending Left Turn](image)

The AI is shown in the right side of Figure 1, depicting the same descending left turn. To interpret the display, participants were instructed to concern themselves only with pitch and bank. To determine their pitch, participants used the vertical scale in the center of the display. The black upside-down “V” symbol represented the nose of their aircraft and the thin white bar which separates the blue and brown areas represented the horizon. The blue area represented the sky and the brown area represented the ground. Thus, when the “V” was above the white bar into the blue, the simulated aircraft was pitched up. When the “V” was below the white bar into the brown, the aircraft was pitched down. To determine their bank, participants either referenced the horizon bar to ensure that it was completely horizontal, or used the rounded scale occupying the uppermost portion of the
display. The white arrow on this scale always points directly towards the sky. When this arrow was at the top of the rounded scale, the aircraft was straight and level. As participants banked left, the symbol would slide along the scale to the right, as shown in the right side of Figure 1. As they banked right, the symbol would slide to the left.

3.4 Procedure
After giving informed consent and completing a demographic survey, participants read an instruction document which explained the tasks they were to perform. The two displays (AI and ASD) were explained in detail to the participant and any necessary clarifications were made. Participants were given instruction on how to best interpret the two displays, but not on specific recovery techniques. The participant was then free to fly the simulator with the first display (one of two within-subjects conditions) for up to ten minutes. This free-fly session was used to familiarize the participants with the controls, the display, and the behavior of the simulated aircraft. Next, two practice trials were performed to familiarize participants with the task. In each trial, participants were placed in an already-in-flight situation. In each situation, the simulated aircraft was started in one of six unusual pitch/bank attitudes. These attitudes included three levels of bank (moderate bank of 45°, moderate inversion of 120°, and severe inversion of 165°) and two levels of pitch (moderate pitch 30°, and severe pitch of 60°).

Participants had no visual reference except for the display being used, which occupied a 3” by 3” square on the otherwise black screen. This was intended to simulate a pilot who was experiencing SD and, in accordance with his training, was focusing solely on his instruments to regain his perception of orientation. Participants began each trial looking at an entirely black screen, with their hand neutral on the joystick. On the experimenter’s command “ready, go!”, the simulation was unpaued, the first display being used appeared on the screen, and the participant began the task of returning the aircraft to straight and level flight (+/- 5° of pitch, +/- 10° of bank). Once the simulated aircraft stayed within these parameters for at least 2 seconds, the trial was terminated. After two practice trials and six experimental trials, the second display was explained in detail and the entire process was repeated.

3.5 Experimental Design
Demographics collected via survey were age, gender, manned and unmanned flight hours logged, previous instrument ratings, previous experience of SD in flight, and previous SD training. The only independent variables which were manipulated were the display being used, and the order of displays used within-subjects. The order of displays used was alternated between participants to counterbalance any practice effect which might increase performance on the second display used. However, it is acknowledged that the trained pilots had significant experience and training using the traditional AI, training beyond that received by any participant using the ASD. The counterbalancing of order was also performed to minimize any perception that the ASD was “new and/or improved” while the AI was “old technology”.

The order of situations was counterbalanced to minimize any learning effect from one display to the next. With the first display, participants went through situations 1-6 in numerical order. With the second display, situations 2 and 3 were swapped with situations 5 and 6. This difference in order was applied to reduce the likelihood that participants would predict the next situation based on the experience they had with the first display. This particular arrangement was chosen because it did not alter the order of severity of banks/pitches and thus allowed analysis to be performed regarding each display’s performance in varying unusual attitude severities.

Dependent variables included time to complete recovery, RMS error of recovery, and initial control error count. After the end of an experimental session, the data was analyzed in Microsoft Excel to calculate these variables. The program used in Excel allowed the experimenter to see the elapsed time of the simulation at a precision of one tenth of a second, the simulated aircraft’s pitch and bank at a precision of one hundredth of a degree, as well as the participant’s joystick inputs at a precision of one hundredth of a percent. To determine time to recovery, the experimenter found the first instance that the aircraft’s pitch and bank was within the acceptable tolerances for straight and level flight (+/- 5° of pitch, +/- 10° of bank) for at least two seconds. The elapsed time between the beginning of a trial and the beginning of this two second portion was recorded as the recovery time. To tally initial control errors, the experimenter determined whether the participant made the initial joystick input in the correct direction of bank (i.e. a left input for right banks and vice versa). An initial control error was recorded when the participant made a control input of 10% or more in the incorrect direction for any length of time or a control input of 5% or more in the incorrect direction for at least two tenths of a second. These criteria were adopted to prevent misclassification of any unintended stick movements.
Poisson, Miller, Haas, & Williams

A post-experimental survey was used to elicit participant’s subjective thoughts towards each display. These included preference of one display over the other, the perceived best and worst aspects of each display, the ease of learning how to use each display (with a 5-point rating scale), the strategies used, whether either display was misinterpreted during the trials, and any recommended improvements. The goal of these survey items was to provide some background for effects seen in the experimental data. For instance, if participants communicated that one display was more preferred and easier to use, there would be an expectation that performance in the experiment would be superior with this display. Also, if one display performed poorly, survey responses indicating frequent misinterpretations and complicated strategy for using this display may explain the poor performance.

3.6 Data Cleaning
It was noted that in the unusual attitudes that included an upwards pitch, there was a large difference in recovery technique between experienced and novice participants. Novice pilots tended to push the stick forward to bring the nose of the aircraft to the horizon. Experienced pilots avoided this technique as it would incur negative Gz forces, which would cause the pilot to rise out of the seat and press against the seatbelts, potentially reducing the pilots’ ability to control the aircraft. Because of this varying technique, recovery times and RMS values differed greatly for reasons that had nothing to do with the effect of the display. Thus, all upward pitched situations were excluded from data analysis. Additionally, one participant lost control of the aircraft during several recoveries and was excluded as an outlier.

4. Results

4.1 Preference
In terms of preference of one display over the other, approximately 26% (7) of the participants preferred the ASD over the traditional AI. Although this is a relatively low proportion, it is an impressive finding since nearly all of the participants had some level of prior experience with the AI, none of them had experience with the ASD, and participants were instructed to select the display they would fly a real aircraft with if they had the option. In fact the most experienced participant to prefer the ASD had 850 actual flight hours with the AI. However, as previous flight experience increased, preference for ASD generally decreased, as expected. In fact, 81.8% of those who had manned or unmanned flight experience preferred the AI, as shown in Table 1. Thus, the odds of preferring the AI were 6.75 times higher for those with some experience than for those with none. However this effect only neared statistical significance ($\chi^2(1, 27) = 3.710, p=0.0541$).

<table>
<thead>
<tr>
<th></th>
<th>AI Preferred</th>
<th>ASD Preferred</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Novice</td>
<td>18</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>Novice</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>7</td>
<td>27</td>
</tr>
</tbody>
</table>

4.2 Ease of Learning
Participants used a 5-point rating scale to indicate how easy it was to quickly become confident using the ASD, and the majority of participants (10) chose “easy”. The majority of participants (17) chose “very easy” in response to the same question with the AI. When coded numerically with -2 representing “very easy” and +2 representing “very hard”, the ASD mean response was -0.48 and the AI mean response was -1.41. The average difference in responses was therefore 0.93 lower (i.e. easier) for the AI. A Wilcoxon signed rank test revealed that the AI (Mdn = -2) was rated as easier to learn than the ASD (Mdn = -1), $z = 2.19, p = 0.0285, r = -0.287$. This rating difference ranged from 4 higher (i.e. harder) with the ASD to 3 higher with the AI. All of these results were expected since the majority of participants had already learned how to use the AI but were unfamiliar with the ASD. With these facts in mind, it is noteworthy that more than one third of the participants found the ASD at least as easy to learn as the AI, and 20% found it easier to learn than the AI.

4.3 Time to Recovery
A repeated measures analysis of variance (ANOVA) revealed that the starting orientation had a significant main effect on the time to recovery, $F(2, 25) = 15.09, p < 0.0001$. This was an expected main effect since the more drastic unusual attitudes started participants far from straight and level flight and required longer duration control inputs to recover than the less extreme starting attitudes. Display type also had a significant main effect on time to recovery,
F(1,25) = 15.03, p = 0.0007. Participants averaged 7.89 seconds to recover the aircraft using the ASD and 5.97 seconds using the AI, as shown in Figure 2. The average time difference was therefore 1.92 seconds faster with the AI. As expected, there was a significant interaction between the presence of flight experience and display type, F(1, 25) = 10.41, p = 0.0035. Because of their previous use of the AI, participants with flight experience completed the recovery task an average of 2.78 seconds faster with the AI than with the ASD, while participants with no flight experience were an average of 0.22 seconds faster with the AI that with the ASD.

Figure 2: Bar Graph of Flight Experience and Display Type Interaction Effect on Time to Recover

4.4 Root Mean Square Error
The root mean square (RMS) error from 0° of pitch and bank was calculated for each recovery, with lower values indicating a more accurate (i.e. less deviation from perfect) recovery. An ANOVA showed that the starting orientation of the simulated aircraft had a significant main effect on RMS error, F(2, 25) = 126.84, p < 0.0001. This finding was expected since the RMS values were dependent on deviations from 0° of pitch and bank. Thus, the more severe unusual attitudes necessitated higher RMS values regardless of recovery time or accuracy. Flight experience levels had no statistically significant effect on RMS error. Participants averaged 72.27 degree*seconds with the ASD and 69.20 degree*seconds with the AI. The average RMS difference was therefore 3.07 lower with the AI. However, there was no statistically significant effect of display type on RMS error.

4.5 Initial Control Error Count
It was postulated that one of the two display symbologies might allow pilots to more accurately determine their orientation at a glance, and therefore produce fewer errors in the initial recovery process. Thus, the number of initial bank errors was calculated for each participant. In other words, if the participant should have banked left to achieve the quickest recovery but banked right instead, this was coded as an initial control error. An ANOVA reported that there was a significant three-way interaction effect between display type, order of displays used, and starting orientation, F(2, 50) = 4.49, p = 0.0161. This interaction can be seen in the large differences between the left and right side of Figure 3. A potential explanation for this effect lies in the counterbalancing scheme used for the starting orientations. One of the two orders had participants engaged in a severe inversion before the less severe orientations, while the other increased in severity with each trial. Since the majority of participants had been previously exposed to the AI, any learning effect present would have been more drastic when the ASD was being used. It is likely that those who both used the AI first and had the building severity situations had the maximum amount of time to learn how to best interpret the ASD and complete the tasks. Those who either used the ASD first or had the initially severe situations had less time to learn before being tested by the severe inversion and thus committed more errors. Interestingly, none of the two-way interactions involving these variables were statistically significant. Participants averaged 0.26 errors per trial with the ASD and 0.48 errors per trial with the AI. The average error count difference was therefore 0.22 more errors per trial with the AI. However, display type did not have a significant effect on the number of errors made.
5. Discussion and Conclusions

This study was designed to determine if the graphical representation provided by the AI or ASD more intuitively communicated orientation in 3-D space to the pilot. The goal was to see if a pilot who was attending to his instruments would be able to more quickly and accurately return to straight and level flight with one of these two displays. It was found that participants recovered about 2 seconds faster with the traditional AI than with the ASD, on average. Participants also rated the AI as significantly easier to learn than the ASD. However, it should be noted that nearly all of the participants came into this study already having at least some experience with the AI and thus would be expected to perform better with it. With that in mind, it is interesting to see the effect of number of flight hours on recovery times. More experienced participants tended to perform better with the AI while less experienced participants tended to perform better with the ASD.

While RMS error and initial error count comparisons yielded non-significant results, participants had lower RMS values with the AI yet higher error counts with the AI. The lower number of errors seen in the ASD may be due, in part, to its use of a visual command display. The ASD textually displayed the correct initial control action to the participants, while the AI left it up to the participants to decide on their own. While there are inherent drawbacks to the use or non-use of these commands, they may well be the reason for the fewer ASD errors. Performance aside, over 25% of participants preferred the ASD over the AI. This is an interesting fact considering the lack of experience with the ASD at the outset of the study.

Open-ended survey responses yielded results with potential for generalization to other attitude displays. The main theme behind the responses was that quick and accurate comprehension was the single most important factor to the display’s perceived effectiveness. Participants noted that the use of colors, words, and symbols can all be used in various manners to achieve this speed and accuracy. For example, the AI used blue and brown colors to distinguish current pitch direction, while the ASD used textual messages to communicate the correct control input. When deciding how to combine these possibilities effectively, it is important to remember that simplicity was mentioned many times as a key aspect to display design. While rich information can be helpful during times of low workload, designers should temper the urge to provide extra informational stimuli with the knowledge that pilots may be viewing these displays in less than ideal circumstances, such as when suffering from SD.

It should be noted that the main advantage of the ASD may not be in unusual attitude recovery. The intent behind the ASD design is to aid the pilot by alerting the pilot and drawing his or her attention to the instruments in certain SD-inducing situations. That being said, eliminating SD entirely is a difficult task which may not be possible for a single instrument. Therefore, it is important that the symbology and alerting systems be laid out in a way that allows for quick and accurate recovery from unusual attitudes. To further test the claim that the ASD may eliminate or minimize the actual occurrence of SD, future research should be performed in a high-fidelity simulator. If a moving-base simulator was used, the vestibular and visual inputs which cause or increase the likelihood of SD could be portrayed. This could allow the pilot’s actions and performance to be tracked with each display during the possible inception of SD and the researcher could see if one display caused pilots to be disoriented while the other did not.
Since fiscal year 1993, there have been 72 SD Class A mishaps in the Air Force which have claimed 101 lives and 65 aircraft for a total cost of $2.32 billion [2]. It has been hypothesized that the current technology may be one of the many factors contributing to this deadly trend. This study set out to determine whether the newly proposed ASD permits the pilot to return their aircraft to level flight more easily and efficiently than the traditional AI. Ultimately, the AI had faster recovery times and lower RMS error values. However, it is important to note that fewer initial control errors were made with the ASD. Additionally, experienced pilots in this study believed the ASD has potential in the field of SD minimization and mitigation. Several of these went far as to say that they would prefer to fly with the ASD despite their years of experience with the AI. Additionally, over half of the participants used the ASD’s visual commands in their recovery strategies and found them to be helpful. With the huge costs of SD to the Air Force, in terms of dollars, aircraft, and lives, the ASD merits further investigation as a potential path to a safer future.

Acknowledgements
The authors would like to acknowledge the assistance of Pilot Disorientation Prevention Technologies, especially Mr Jerry Marshall. The authors would like to further acknowledge the assistance of Dr. Henry Williams and Mr. Charles Powell of the Naval Aeromedical Research Unit-Dayton.

References