GENERAL AVIATION TRAINING: INTEGRATION OF ADVANCED COCKPIT DISPLAYS

Shawn M. Doherty
Frances Atchley-Greene
Embry-Riddle Aeronautical University

Liu Chang
Gerald P. Chubb
The Ohio State University

June 2002

Technical Report
Prepared for
FAA Aero Center
Civil Aeromedical Institute
Oklahoma City, OK

FAA Contract DTFA 02-01-C-09254

Abstract

The inclusion of new global positioning systems (GPS) in general aviation aircraft has been of concern to the aviation community due to the potential for this new technology to increase workload in general aviation aircraft. A study was conducted jointly by Embry-Riddle Aeronautical University (ERAU) in Daytona Beach, Florida and the Ohio State University (OSU) in Columbus, Ohio, to investigate this issue. Pilots were evaluated on the ability to interact with two display interfaces under high workload conditions. The first interface was an advanced digital avionics system equipped with a GPS-like component in the multi-function display (MFD). The second interface was a traditional general aviation console with an onboard GPS system. Participants were also to interact with these systems with two forms of device inputs: voice and touch screen style input. These systems were expected to show differences in workload on the basis of the type of interface the pilot was confronted with and the method by which they entered data into the system.

Workload was measured through an analysis of participant performance during flight to an approach that required a change due to weather. Flight tracking performance, time to change the flight plan to the new runway, and subjective workload reports provided information about the effects of these different platforms and interfaces on pilots' workload.

Results indicated that the advanced digital avionics system did produce better flight path tracking, faster times to change runways, and a lower subjective workload compared to the more traditional general aviation console.

Table of Contents

List of Figures

List of Tables

Introduction

The implementation of new types of cockpit displays to a Global Positioning System (GPS) and their effect on pilot workload has been of primary concern to the FAA Civil Aeromedical Institute (CAMI) in Oklahoma City, Oklahoma.

Background

Workload has been a major concern since the late 1970's. In general, workload is a sense one feels of how hard they are working. Pilot workload is divided into two categories: physical workload and mental workload. Physical workload involves the amount of work that is physically affecting the pilot. It involves the heart, muscles, lungs, and other organs necessary for carrying out physical activity. Mental workload involves the amount of work the brain does. Including cognitive processes and decision making that takes place in response to task demands. Mental workload can be difficult to measure as there is much room for variability from one subject to the next. One person may indicate that a certain task requires an extremely large amount of workload, while another person may not have to make very much mental effort at all for the same task.

While automated aircraft controls have contributed to a decrease in the amount of physical workload imposed on a pilot, increasing air traffic and more stringent fuel constraints combined with stricter noise regulations have created an increased mental workload. The pilot now has to be a "flight manager" in order to fill both of the responsibilities of monitoring certain advanced displays, as well as being able to control the aircraft's flight path by the use of the yoke and rudder.

For decades, traditional cockpit displays have remained in general aviation (GA) aircraft. There is currently an effort to begin replacing these traditional displays with newer displays. One of these new displays, the highway-in-the-sky (HITS) display, provides the pilot with a perspective view of a flight path which is seen in the form of a series of boxes making up a tunnel. The research for this type of display began in the 1950's, as the result of efforts made by the Join Army-Navy Instrumentation Program.

A (HITS) primary flight display (PFD) depicts a commanded flight path through the use of connected, graphic objects representing the corridor of airspace through which the aircraft must be flown, showing the current and future position of the aircraft (Wickens, 1997). In addition, the flight path display may be more consistent with the pilot's mental model of the airspace through which the aircraft must be guided. (Fadden & Wickens, 1997). By using the Global Positioning System (GPS) and advanced computer technologies, HITS displays provide better position tracking and aircraft control when compared to more conventional types of guidance displays.

Problems with the actual implementation of the HITS display were a factor until recently because it was too expensive to be installed in an aircraft. Due to the advent of two technological breakthroughs, this display system is now a reality for most GA cockpits. The first breakthrough is an affordable Global Positioning System (GPS) receiver that provides real-time, accurate information as to the position of the aircraft. GPS is a system for navigation and position determination using line of sight signals with the technology of artificial satellites. The GPS constellation is comprised of 24 Earth-orbiting satellites, which transmit radio signals that consist of the satellite's position and the time it transmitted the signal. These signals can be received on Earth with a relatively inexpensive device. The distance between a satellite and a receiver can be computed by subtracting the time that the signal left the satellite from the time that it arrives at the receiver. If the distance to four or more satellites is measured, then a three-dimensional position on Earth can be determined (Nendick & George, 1995).

GPS is the most recent high technology navigation system to enter aviation. Civilian pilots have been using GPS in uncontrolled airspace for applications such as crop dusting, aerial photography and surveying, search and rescue, and basic point-to-point navigation for some time. On June 9, 1993, the Federal Aviation Administration (FAA) approved GPS for supplemental use in the domestic, oceanic, terminal, and non-precision approach phases of flight in controlled airspace as well. It was also approved for Primary Means IFR navigation in Australia in December 1995, with approach certification anticipated thereafter. The GPS also will be used for Category I precision approaches (Nendick & George 1995). Precision approaches are required when the weather conditions at a given airport reduce the ceiling (height of the base of a cloud layer) and the visibility (distance a pilot can see visually) to levels that are below non-precision approach criteria. It is necessary to improve GPS-derived accuracy to allow the system to be used for these types of approaches. Consequently, the FAA has not yet provided certification for use of GPS style in all areas of flight.

However, GPS has been adopted for General Aviation (GA) use before the regulatory provisions and training requirements were established. Introducing GPS into aviation in this way could mean that human factors considerations are overlooked or neglected. The positive aspect of using GPS is that users reported being more accurate with their navigation, tracking, and awareness of position relative to airspace and terrain, preventing pilots from running out of fuel through poor route following. One negative effect of using GPS is that GPS does not appear to be encouraging pilots to take more time scanning the instruments. Some pilots do not cross-check the GPS information, and fly using GPS instead of a chart. The design of most GPS units makes it difficult to easily check the entered track and distances of routes with the flight plan. This is a basic requirement for safe navigation to prevent the GPS leading the pilot to where it has been programmed but not intended to go.

The second breakthrough leading to the widespread use of HITS and GPS style systems is the fact that powerful graphic display systems are being produced inexpensively. These display systems are powerful enough to provide real-time HITS depictions in the cockpit, showing the information which is received by the GPS or other synthetic radar. In addition to HITS, the inexpensive nature of technology has led to the development of a multi-function display to be integrated with the HITS display in GA cockpits. The MFD provides the pilot with different types of information to assist in navigation tasks and contains the information already included by GPS-style systems. Not only does the MFD display terrain, traffic, and weather information, it is also used to enter and edit flight plan information.

In the future it is projected that the MFD will overtake the functionality present in current GPS units for flight plan information. The HITS display will be paired with the MFD in order to plan and execute flights. The flights could be conducted in both visual and instrument meteorological conditions. Thus, it is of some interest to compare the MFD style displays with the currently available GPS-style systems to see if there is a difference in pilot performance and workload.

Workload

What is less clear is the level of workload these new displays and technologies will have on the general aviation pilot. A number of factors, such as display types and input devices, may mediate the mental workload the pilot may experience when interacting with these new technologies.

Display Effects. Differences in presentation of information have been shown to affect pilot performance and workload. This study is interested in the comparison between an advanced avionics display, the HITS, in comparison to a more conventional avionics console. This is not a unique comparison and the virtues of the HITS displays over more conventional flight instrumentation is well known (e.g. Fadden, Ververs, & Wickens, 2001; Wickens & Long, 1995). HITS displays have been shown to be better for flight path guidance due to the nature of the integrated display.

The benefit of HITS displays appears to emerge from two primary characteristics of the HITS display: the preview provided by the tunnel symbology and the integrated information presented to the pilot in the tunnel (Doherty & Wickens, 2000). The property of preview in the HITS display is that the highway shows the pilot where they should be in position at different points in the future. This allows the pilot to anticipate and plan for course corrections prior to reaching the point in space or time where those corrections are necessary. Similarly, the HITS display integrates information about horizontal and vertical flight plan command inputs and makes them visually obvious to the pilot. This reduces the mental integration the pilot needs to perform in their head to determine their position on the flight path with traditional instruments.

However, these properties of the display indicate little about workload. It may be inferred that a reduction in workload will occur in response to HITS displays because of the benefit in performance that occurs as a function of the integrated information and preview that emerges as a natural function of the display's properties. The workload that emerges directly from interaction with the GPS systems is less clear. The MFD-style displays that are emerging frequently require a long sequence of keystrokes or button presses necessary to interface with the system to obtain direct and specific information about flight path information or change flight plans. This, then, could cause workload to be equivalent to or even potentially greater than the workload for obtaining flight path changes in traditional aircraft consoles. That issue is the purpose of this study: to address the changes in workload that occur as a function of the interaction with the new GPS systems.

Input Devices

Input from the pilot to the GPS style system may also affect the level of workload in the operator. A multiple resource model of the pilot would argue that the performance of the pilot will be limited by the number of tasks that share similar resources. These resources are limited in nature and take the form of visual, auditory, spatial, analytical, manual, or speech modalities. Tasks that share modalities increase workload and make performance worse due to "mental loading" in which multiple tasks leave each subsequent task fewer resources to accomplish the overall goal. Therefore, a pilot trying to interface with the GPS system through a touch screen interface may end up having poorer flight performance because both tasks require manual input (flight through the yoke, and GPS information through the touch screen). The more frequently the pilot has to reference the touch screen manually, the more difficulty he may have meeting optimal flight performance because the same limited resource is being spread out over more than one task.

An alternate design solution is to allow the pilots to interface with the system through voice commands. Voice commands, while not removing the maze of options characteristic of many computer-based menus, will allow the pilot to retain their hands for flying and use their voice to interface with the system. The goal of speech input is to make the pilot interface with the MFD easier, and, therefore, decrease workload.

The benefit from voice command systems is somewhat limited due to technological reasons. Unlike humans, modern computers do not have the sophistication to recognize words that are part of a series of words slurred together in a sentence. This means that speech recognition devices demand slower-than-normal conversational speech. Additionally, several other things can account for the voice recognition system not recognizing an utterance. The pilot may be under stress, either emotionally or physically, which causes change in voice pitch. There could also be noises going on around the pilot which cause the voice recognition system to be unable to recognize the command. When external noises or the computer's inability to distinguish similarly-sounding words are present the computer may confuse that one word as another of similar sound and spatial frequency components. All of these cases are things that occur under high pressure and high stress situations. These occurrences may limit the effectiveness of voice communications as a way to mediate workload effects.

Testing

Testing the pilot's workload is an important way to evaluate the overall aircraft system design. The efforts to precisely measure workload have been continuing for several decades. Since workload is a subjective concept, ways to quantify a subjective component have not always proven reliable. A subjective rating of mental workload has been the most widely used method due to its ease of use. Operators are asked to rate how hard they feel they have worked in a particular task, with the ratings collected during or after the task has been completed. This method has the advantages of being nonintrusive and easy to implement. On the other hand, its disadvantages include the lack of a theoretical framework, difficulties in comparing results between experimenters using different rating scales, and the problem of ratings yielding relative rather than absolute results. The NASA Task Load Index (TLX) is a measure of subjective workload that combines ratings from several unweighted scales into a single weighted measure (Hart and Staveland, 1988). See Appendix A for an example of the form. Despite some concerns about the NASA-TLX as a measure of workload (e.g. Bailey & Thompson, 2001) it is the most commonly used form of subjective workload assessment and will be used in the current study.

Hypotheses

The integration of advanced navigation displays with on-board flight planning displays is expected to increase the pilot's efficiency of flight operations for GA aircraft. The purpose of the present study was to determine how long it takes to make changes to the planned course, whether any errors are made doing so, and how well the airplane can be flown while updating the GPS data.

Thus it is expected that differences in flight performance will be found between the SmartDeck system and the Hawk system. Numerous studies have demonstrated that HITS displays induce better flight path tracking compared to more conventional instrumentation. The same is expected in the current study.

Of more interest is the anticipated workload difference between the SmartDeck and Hawk systems. It is expected that pilots will have lower workload when using the SmartDeck displays (HITS integrated with the MFD) than when using the conventional displays and the KLN 89B GPS unit due to the integrated information and graphical depiction of the flight plan provided by the advanced avionics displays.

It is also expected that pilots will spend less time in changing the approach runway using the SmartDeck system compared to similar changes using the KLN 89B GPS unit. This is expected since the SmartDeck system has an intuitive touch screen interface that provides a faster response than the conventional turn knob in the Hawk system. However, it is also expected that the voice input of the Verbex system will produce even faster changes in flight plans due to the reduction in workload and use of a different modality compared to the touch screen or turn knob interfaces. This effect is contingent upon the technology responding quickly enough to the voiced commands.

Method

The purpose of this study is to investigate different forms of GPS technology that are expected to affect workload for the future general aviation pilot. To investigate these aspects, a comparison of two GPS style systems was performed by Ohio State University, as well as an investigation of two different styles of command input that was conducted by Embry-Riddle Aeronautical University. Many of the technical details are similar between the two programs and differences will be elaborated where warranted.

Participants

ERAU. Fifteen participants were selected to participate. Each was FAA certified private pilots with an instrument rating. Participants were paid $30 for their participation in the study. Participants were selected through on-campus advertisements and screened for an instrument rating apriori. Participants had between 107 and 1200 hours of flight time. Thirteen of the participants were male and two were female. Pilots' average flight time was 373.8 hours.

OSU. Thirty-two experimental subjects were participants, all of whom were FAA certificated private pilots with an instrument rating. Those pilots who had prior experience with the Bendix / King 89B were paid a flat $12 per hour for their participation. Those not current or needing special training on the Bendix / King 89B were paid $20 for participation. The total payment was based upon their completion of both HITS training and the testing sessions. Participants were selected from OSU's Academic Flight Lab and Flight Training Clinic, subscribers to a local pilot bulletin board on the internet, and posters in various local flying clubs. Other subjects were obtained by word of mouth invitation of participants and aviation enthusiasts who were aware of the study and could indicate who needed to be contacted in order to arrange participation. These participants were divided into two groups: those having gone through the OSU flight program (OSU group) and those trained outside of OSU (non-OSU group). Participants had between 95 and 27,000 hours of flight time with the non-OSU pilots having an average of 3409 hours of flight time and the OSU pilots having an average of 869.5 hours of flight time. Thirty participants were male and two were female.

Apparatus

The SmartDeck portion of this experiment was conducted in the Small Aircraft Transportation System (SATS) Lab both at ERAU and OSU. The SmartDeck HITS display system platform was installed in the Airway Sciences building on the Embry-Riddle campus (see Figure 1).

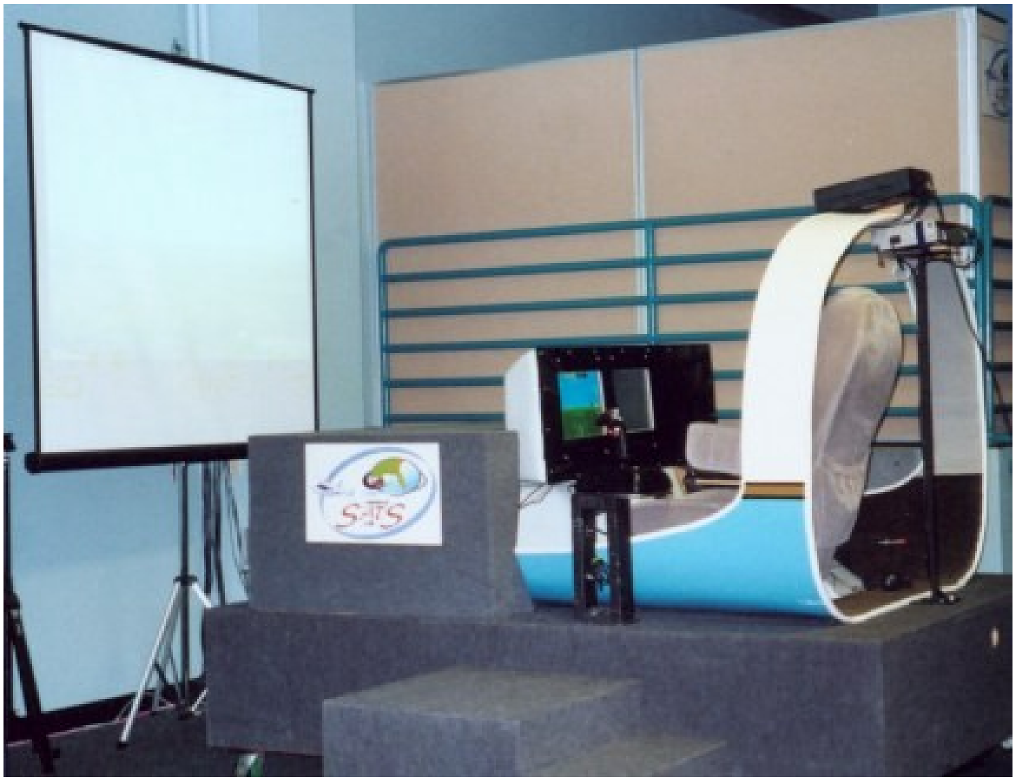

Figure 1. ERAU SmartDeck simulation platform.

The SmartDeck HITS display system simulator was installed in a renovated ATC 810 cockpit shell located in Hangar 2 at the OSU airport. (See Figure 2.) In both cases, two display screens were put on the top of the control box that had the pilot's yoke, throttle quadrant, and other simulated controls. An out-the-window scene was projected over the platform onto a screen or wall in front of the platform. To control the airplane, the subjects used a control yoke which was on the left side of the control box, the throttle which was in the middle of the control box and the rudder pedals which were located on the floor of the cockpit shell. Other controls could be ignored (flaps, landing and landing gear).

Figure 2: OSU SmartDeck simulation platform.

The Bendix/King KLN 89B GPS system is installed in OSU's Cessna 172 aircraft used for instrument flight education, and the same unit is therefore installed in the AST Hawk, since it is intended to be the full time display that is used in support of instrument flight education. The Hawk has four display screens for an out-the window scene simulation and a conventional small airplane cockpit layout, as shown in Figure 3.

Figure 3: Hawk's Cockpit Layout.

Displays

Both OSU and ERAU participants utilized the SmartDeck HITS Display System. The computer program for the new SmartDeck “Highway-in-the-Sky” (HITS) display system was designed by Goodrich Avionics. This display system is composed of two display screens which provide all the information necessary to maintain flight control, navigate, control aircraft configuration, and monitor systems’ health. The Primary Flight Display (PFD) is on the left, immediately in front of the pilot (Figure 4). The HITS format appears on the PFD. It shows a forward view of the world relative to the aircraft position, as well as aircraft configuration information and basic instrument information. The purpose of the PFD display is to provide the critical information necessary for flying and controlling the aircraft. As such, this display cannot be re-configured to portray anything but the PFD information. The PFD Page as displayed on this screen is a display only and has no pilot interaction capability.

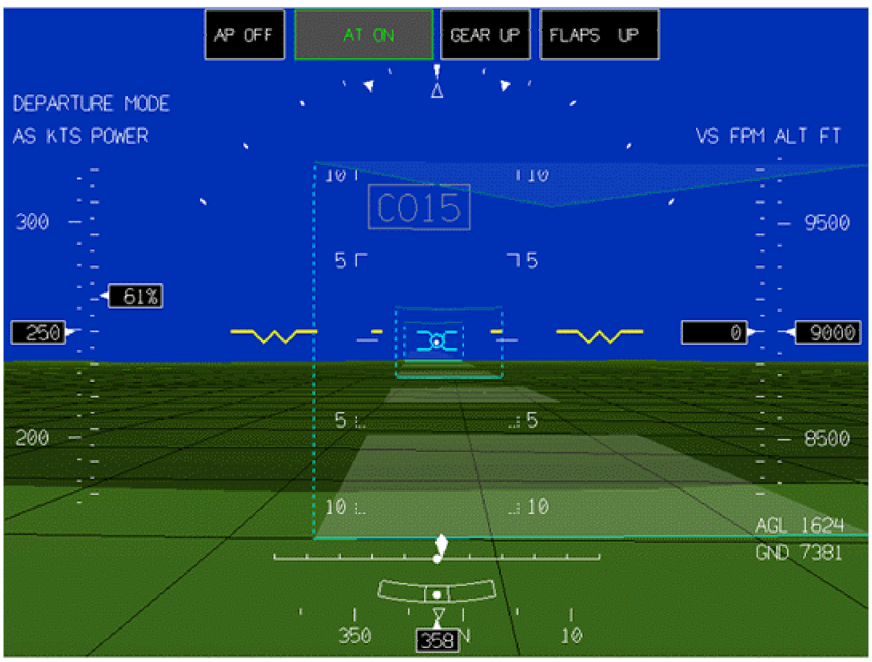

Figure 4: PFD – Primary Flight Display Format.

The second display screen in the SmartDeck system is the Multi-Function Display (MFD) shown in Figure 5. The MFD is to the right of the PFD and may be toggled between several interactive pages which provide detailed information related to navigation, systems status monitoring, and various checklists. The information displayed may be highly customized to suit pilot preference, flight mode, and specific situational needs. The MFD screen provides the primary pilot interface to the SmartDeck system using a touch sensitive panel over the display. This screen includes five top-level display pages and a series of submenus.

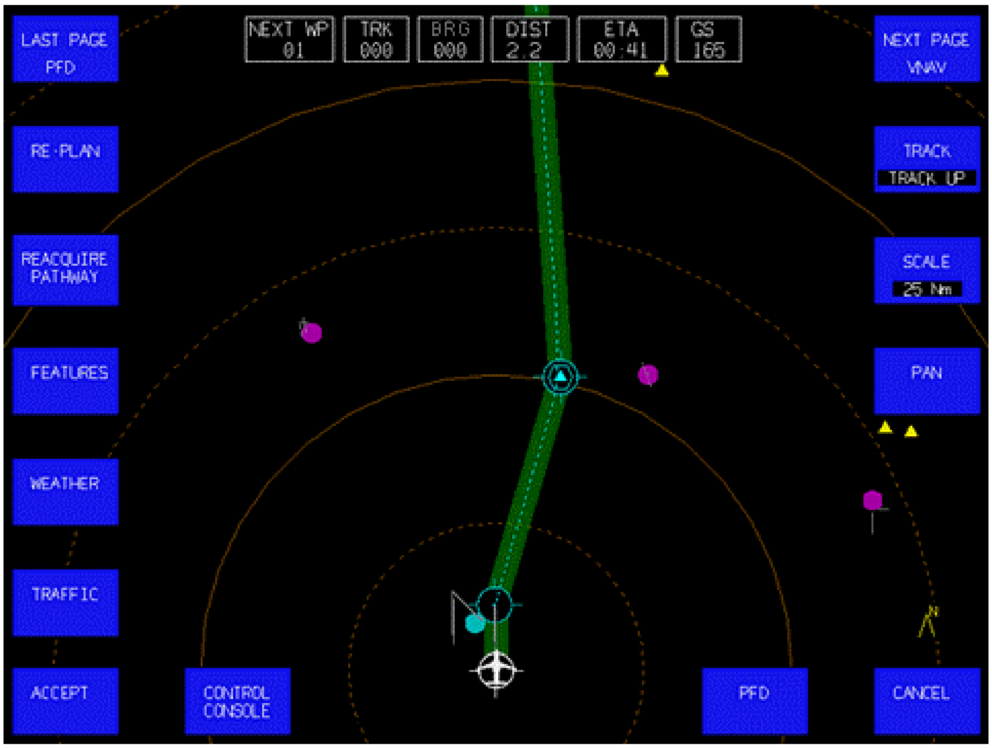

Figure 5: MFD – Multi-Function Display Format.

The top-level pages are: a horizontal navigation page (HNAV), a vertical navigation page (VNAV), a systems page for three aircraft subsystems (engine, fuel, and electrical), a checklist page, and a redundant PFD page. The default page at startup is the HNAV page. The other four pages are accessed in round robin fashion using the MFD. The MFD is also used to display ATC messages and system warning messages. These are delivered both as audible messages and as text in message windows on the MFD. The HNAV page provides a bird's-eye view of the flight path over the ground (see Figure 6). Looking at this page is like looking at the airplane and its flight path superimposed on a map. The planned route of flight and the airplane's position in relation to this route are shown. As the pilot flies on the pathway using the PFD, they will see his or her airplane move forward along the planned route on the HNAV display. Weather and contour map information can be overlaid. Although not implemented for this study, traffic information can also be added.

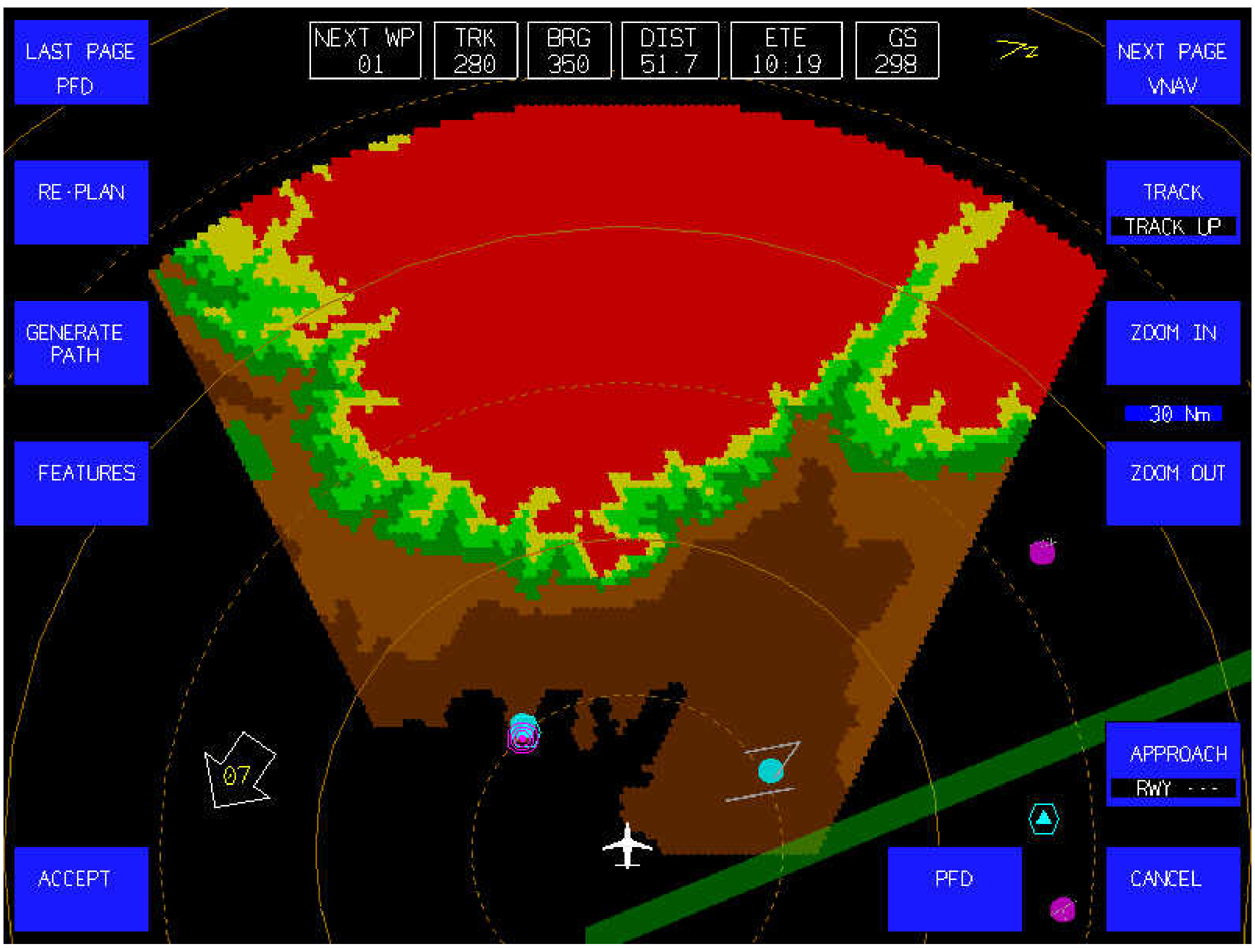

Figure 6: HNAV with Contour Map Overlay.

The VNAV page shows a profile view of the flight and lets the pilot see his or her airplane's altitude in relation to terrain elevation and planned flight path. As the pilot flies on the pathway using the PFD, he or she will see his or her airplane climb, level off at specific altitudes, and descend on this page.

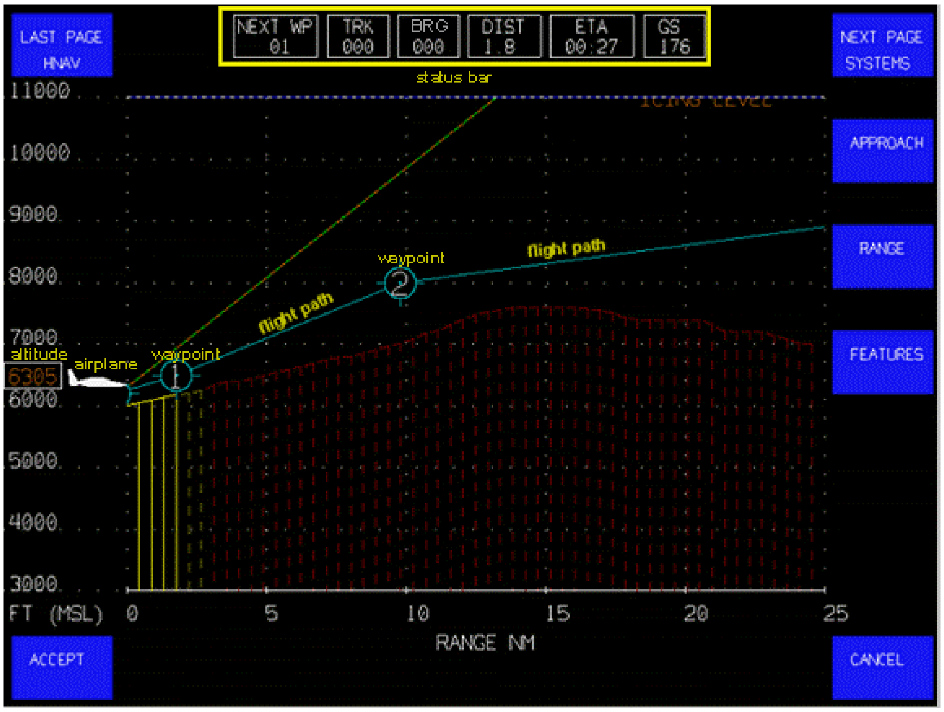

Figure 7: VNAV – Vertical Navigation Display Page.

When the pilot gets close to the destination, he or she needs to set up an approach for landing at the airport. The pilot can select a runway by starting from the VNAV page. First, the pilot will select a runway on the Approach menu and choose Accept. Next, the pilot must return to the HNAV Page and select the PFD menu. This is a critical step for pilots in this experiment and requires numerous button presses to reach this point within the MFD.

When the pilot chooses Approach Mode, the appearance of the flight course will change to show an approach path designated by a dashed line to the selected runway. The final approach portion of the path is shown by an elongated arrow head shaded on one side. A Maltese cross is shown at the final approach fix (FAF). This is the position at which the pilot should begin the final descent for landing. When the pilot reached the final approach fix, the trial ended for this study.

The pathway on the PFD changes from blue to green to designate the approach course. By steering down the pathway, just as the pilot needs to do during cruise flight, he or she will be able to fly a precise approach to the touchdown zone of the runway. This terminal phase of flight was not examined in this study.

The Systems page gives pilots information on the status of equipment in the airplane, allowing the pilot to access information about the status of the airplane's engine, fuel system, and electrical system. For example, this page shows a variety of gauges that can help the pilot determine if the engine is working properly.

The redundant PFD page lets the pilot view the Primary Flight Display on this screen. Its purpose is to serve as a backup in case the PFD monitor (to the left of the MFD) fails during a flight. Then the MFD can be used instead to display the PFD information. That was not necessary here. The Checklist page gives the pilot access to a variety of checklists to follow for performing specific airplane procedures. It was not necessary to use this page in h present study.

The Bendix / King KLN 89B GPS unit in the AST Hawk simulation platform provides a graphical display which will help the pilot navigate more easily and more accurately than standard traditional instrumentation (Figure 8). It has trip planning features, can do air data calculations, and includes other graphical information. In addition, the KLN 89B is FAA certified for En route, Terminal, and Non-precision Approach Instrument Flight Rule (IFR) operations. Notice, however, that while this unit is a step in technology between the more traditional navigation tools and the pictorial representation found in the SmartDeck systems, it can show much of the same basic information needed for navigation.

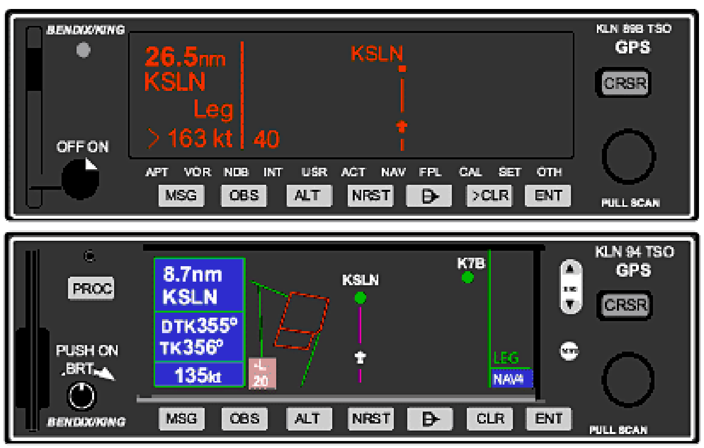

Figure 8: KLN-89B GPS unit.

Scenario

For the experimental task, participants were to plan, enter, and execute an instrument approach to a prescribed airport using the MFD and/or Bendix/King 89B GPS. During the flight, participants were given tasks that required them to interact with the MFD to gather information about weather, terrain, and traffic in the area. Shortly before beginning the initial approach, participants received weather and ATC messages requiring that they use a different runway from the one planned. This required pilots to change the flight plan so that the new approach could be executed. See Appendix B for examples of the tasks required of pilots during the scenario and Appendix C for air traffic control (ATC) messages given to pilots during the flight.

The experimental scenario flight (Scenario 1) originated over the Aurora airport, designated by the code (01V). The destination airport was Centennial, designated by the code (APA). The filed flight plan departed from Aurora Airpark (O1V), flew direct to Falcon VOR (FQF)—waypoint1, then along a Victor Airway (V95) to the HOHUM intersection. On the way to HOHUM, the pilot would turn to the initial approach fix (IAP) NERXY at the waypoint—TURN1—waypoint2. During the leg from FQF to TURN1, the subject was asked to do two tasks in order to prevent them from being bored. The GPS approach for runway 35 was pre-programmed into each of the GPS units (SmartDeck and KLN 89B). After NERXY—waypoint 3, the fourth waypoint is HOHUM, and then the fifth waypoint CASSE, the Final Approach Fix (FAF) for GPS runway 35. Finally, the pilot will land on runway 35 at APA if they follow the flight plan.

However, in order to create a situation where the runway needed to be changed from this plan, a weather report was issued that dropped the ceiling below the FAA specified minimums for runway 35 but above the minimums specified for the GPS approach to runway 28. The alert pilot should have recognized this fact and requested the GPS approach to runway 28 when the ATC specialist asked for "intentions" after announcing the change in weather that implied the runway 35 minimums were no longer met.

All of the non-OSU subjects shot the approach to runway 35R (something commercial pilots would not be allowed to do, but private pilots can do legally). This precluded or eliminated any change in runways, when changing runways had been the intent of the study (as implied by the instructions subjects were given – See Appendix D). In order to force pilots to change the runway, a second scenario (Scenario 2) was created by OSU and given to an additional 16 participants (the OSU pilot group). The second scenario was exactly the same as the first scenario, but the participants were explicitly instructed to select an approach, instead of the pilot having to make a decision about where to go once the minimums had dropped.

In either case, after the subject got the clearance for the GPS runway 28, the subject should start the procedure for changing the runway (in either the SmartDeck or the KLN 89B), and follow the new generated pathway leading toward NIDLY, the FAF for runway 28.

Right after the subject finished the procedure of changing the runway, the experiment was frozen for as long as 5 to 6 minutes which allowed the subject to remember workload information presumably without substantial memory decay.

During this scenario pilots also had a suggested altitude profile to follow which was provided to subjects through the VNAV page and also provided in a flight chart.

- Depart Aurora Airpark at 11,000ft
- Maintain 11,000ft Direct to Falcon VOR (FQF)
- At FALCON descend to 10,000ft while on V95 to HOHUM
- When turning South off V95 to NERXY descend to 9,000ft
- After the clearance to runway28 is issued descend to 8,000ft to NIDLY

Design

Three independent variables were examined between the two campuses in a mixed-model design. The first independent measure was a within-subject variable that included the type of display shown to participants at OSU, at two levels: The SmartDeck simulation platform and the Bendix / King 89B GPS unit in the AST Hawk Flight Training Device. The presentation of this variable was counterbalanced across participants. A second, between-subject, quasi-variable was the type of pilot participating, also with two levels: pilots from Ohio State University, and non-Ohio State University pilots. The third independent variable in this study was conducted at Embry-Riddle. This between-subject variable also contained two levels. Participants were exposed to either touch screen input devices or voice command input devices. This led to the creation of a 2x2 mixed model created by OSU and a 2x1 mixed model design, as demonstrated in Figure 9. Of these six different conditions, the two cells shown in bold were to be pooled in all analyses using those conditions in order to increase the power to find workload effects.

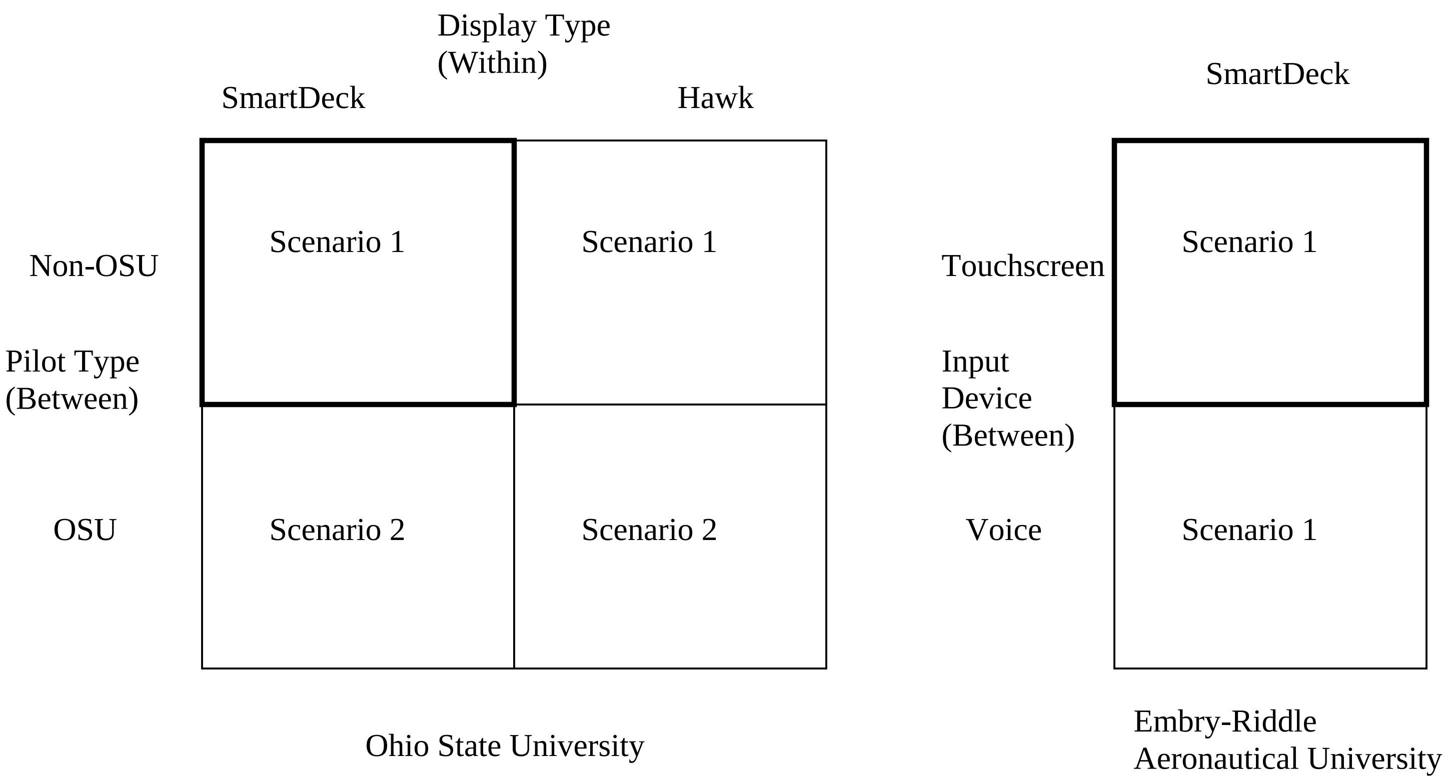

Figure 9. Experimental design of the study.

A number of dependent measures were automatically collected by the SmartDeck and Hawk simulation platforms. Most of this data included navigation errors relative to the flight path and was collected to assess that the pilot was focusing on the flight task as well as interacting with the GPS systems. Data was collected every 2 seconds on the SmartDeck platform and every second on the Hawk platform. Dependent measures included altitude, heading, throttle, roll and pitch changes. Workload measures were also collected after the scenario was completed.

Procedure

The procedure for both ERAU and OSU was very similar. In each, the experimenter read the subject instructions (see Appendix D) to the participant. The instructions briefly outlined the experimental requirements. Subjects were then asked to complete a consent form (Appendix E). Following confirmation of a subject's intention to participate, each subject was given their copy of the "consent for participation" form at OSU while participants from ERAU were offered a copy of the consent form.

The subjects also were asked to provide their personal flight information. The demographic information from subjects can be found in Appendix F.

Participants at OSU were asked to participate in both of two experimental sessions (one session for each of the SmartDeck and Hawk platforms) and the order was counterbalanced across participants. Participants at ERAU were asked to remain for one session.

Pilots were then requested to follow a computer-based training on the procedures used in the SmartDeck avionics. After the training, pilots were permitted to ask whatever questions they wished as well as practicing on a practice scenario. Once participants felt comfortable with the equipment, the real experimental scenario was conducted. For OSU participants, the second session was implemented by asking the subject to fly the same route using the HAWK with Bendix/King 89 GPS simulation. If the subject did not have any former experience with Bendix/King 89 GPS system, there was an extra session for training the subject to get familiar with it before the real experiment was run.

After each of the experimental sessions, the subject was asked to complete the NASA-TLX workload survey. Finally, the subjects were thanked for their participation.

Results

Analyses were performed on three sets of data: 1) performance data (altitude and heading deviations), 2) time data (duration of implementing the runway change procedures), and 3) workload data (NASA-TLX workload survey ratings). The current results focus on the workload and performance data in conjunction by comparing the independent measures of simulation type (Hawk or SmartDeck) and pilot group. Performance data were only analyzed for Leg 2 and Leg 3 of the scenario, as Leg 1 contained only familiarization data, and the scenario was terminated after Leg 3 of the flight path in the scenario as this was when the runway change occurred. Altitude and heading information is used to provide an indirect index of pilot performance under workload conditions as well as an indication of attention to the primary task of flight during the scenario. The time measure reflects the duration of the tasks associated with changing the runway from 35R to 28. The workload measure is probe of subjective workload based on performance in the two display forms.

Due to technical difficulties, data from the voice condition was not obtained and cannot be reported. Additional details regarding this limitation will be covered in the discussion section.

Table 1 shows the basic descriptive statistics for the six conditions outlined in Figure 9. In order to assess whether the common condition between the two institutions could be pooled, a t-test of the four dependent measures was performed. Figure 13 shows the results of this analysis. It is clear that in all four cases, there is a wide difference between the OSU pilots and the ERAU pilots. Since these two groups are significantly different from each other, the data was not pooled for analysis. Since the data for the voice condition is also not available, the right hand cells of the design model (Figure 9) cannot be assessed and compared.

The remaining four cells in the model can be analyzed and the altitude and heading analyses will be separately reported.

Display Type	Pilot Group	N		Mean Alt. Leg 2	Std. Deviation Alt. Leg 2		Mean Hdg. Leg 2	Std. Deviation Hdg. Leg 2
SmartDeck	Non-OSU	16		10151.84	45.88		203.61	1.30
	OSU	16		10157.10	73.89		204.56	.72
	ERAU	15		10305.84	341.40		197.79	13.61
Hawk	Non-OSU	16		10872.80	402.62		206.62	5.48
	OSU	16		11018.05	571.93		211.08	2.68
				Mean Alt. Leg 3	Std. Deviation Alt. Leg 3		Mean Hdg. Leg 3	Std. Deviation Hdg. Leg 3
SmartDeck	Non-OSU	16		9238.77	84.57		182.98	1.71
	OSU	16		9288.70	73.89		183.50	.62
	ERAU	15		9395.30	589.66		189.66	14.00
Hawk	Non-OSU	16		10000.05	586.29		184.69	7.41
	OSU	16		10297.98	797.84		173.83	20.96

Table 1. Descriptive statistics for the conditions reported in the study.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Altitude-SmartDeck-Leg2	Equal variances assumed	8.885	.006	-1.789	29	.084	-153.9937	86.07362	-330.034	22.04665
	Equal variances not assumed			-1.732	14.474	.104	-153.9937	88.89316	-344.066	36.07883
Altitude-SmartDeck-Leg3	Equal variances assumed	6.962	.013	-1.052	29	.302	-156.5317	148.85839	-460.981	147.9179
	Equal variances not assumed			-1.018	14.540	.325	-156.5317	153.70972	-485.061	171.9977
Heading-SmartDeck-Leg 2	Equal variances assumed	3.994	.055	1.704	29	.099	5.8176	3.41488	-1.16664	12.80178
	Equal variances not assumed			1.649	14.239	.121	5.8176	3.52880	-1.73904	13.37417
Heading-SmartDeck-Leg 3	Equal variances assumed	7.807	.009	-1.895	29	.068	-6.6781	3.52424	-13.88595	.52981
	Equal variances not assumed			-1.834	14.393	.087	-6.6781	3.64041	-14.46601	1.10986

Table 2. T-test results for assessing pooling conditions between OSU and ERAU.

Leg 2 Altitude

Figure 10 graphically depicts the performance differences from the altitude measure in Leg 2. The optimal target for this leg of the scenario was 10,000 ft. The overall analysis of variance of the altitude for Leg 2 (Table 3) indicates a significant main effect for the within-subjects factor of display type ([F (1,30)= 85.47, p < 0.001]). This suggests a difference in altitude maintenance between the SmartDeck platform and the Hawk platform. There is no significant interaction effect ([F (1, 30)=.67, p = 0.420) between the pilots source (OSU versus Non-OSU) and performance. The main effect for the between-subjects factor (Table 4) was also not significant ([F (1, 30)=.70, p = 0.410]).

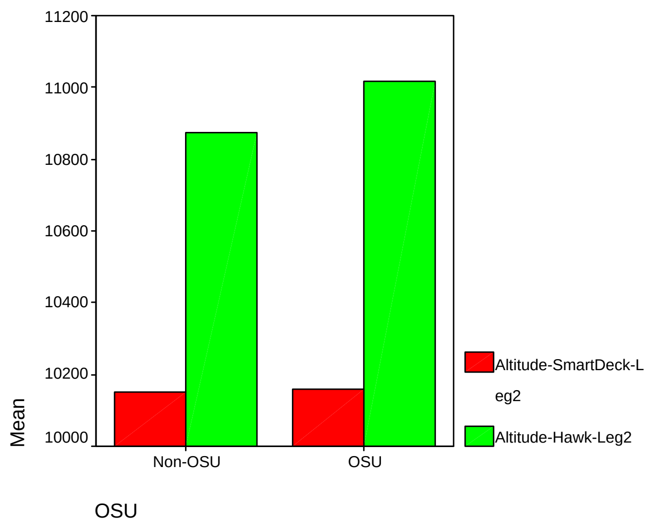

Figure 10: Altitude data for Leg 2.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
DISPLAY	Sphericity Assumed	10009674.9	1	10009674.86	85.466	.000	.740	85.466	1.000
	Greenhouse-Geisser	10009674.9	1.000	10009674.86	85.466	.000	.740	85.466	1.000
	Huynh-Feldt	10009674.9	1.000	10009674.86	85.466	.000	.740	85.466	1.000
	Lower-bound	10009674.9	1.000	10009674.86	85.466	.000	.740	85.466	1.000
DISPLAY * OSU	Sphericity Assumed	78386.807	1	78386.807	.669	.420	.022	.669	.124
	Greenhouse-Geisser	78386.807	1.000	78386.807	.669	.420	.022	.669	.124
	Huynh-Feldt	78386.807	1.000	78386.807	.669	.420	.022	.669	.124
	Lower-bound	78386.807	1.000	78386.807	.669	.420	.022	.669	.124
Error(DISPLAY)	Sphericity Assumed	3513576.567	30	117119.219					
	Greenhouse-Geisser	3513576.567	30.000	117119.219					
	Huynh-Feldt	3513576.567	30.000	117119.219					
	Lower-bound	3513576.567	30.000	117119.219					

a. Computed using alpha = .05

Table 3: Altitude Deviations in Leg 2—Main Within-Subjects and Interaction Effects.

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	3561644218	1	3561644218	54946.455	.000	.999	54946.455	1.000
OSU	45301.542	1	45301.542	.699	.410	.023	.699	.128
Error	1944608.177	30	64820.273					

a. Computed using alpha = .05

Table 4: Altitude for Leg 2—Main Between-Subjects Effect.

Leg 2 Heading

Figure 11 graphically depicts the performance differences from the heading measure in Leg 2. The optimal target for this leg of the scenario was 197 degrees. The overall analysis of variance of the heading for Leg 2 (Table 5) indicates a significant interaction between display type and pilot group ([F(1, 30)=4.99, p=.033]). The significant main effects of display type ([F(1,30)=36.84, p=.000]) and group type (Table 6) ([F(1,30)=11.89, p=.002]) can be seen within the interaction. The interaction suggests that the SmartDeck facilitated performance more for the OSU pilots more than for the non-OSU pilots.

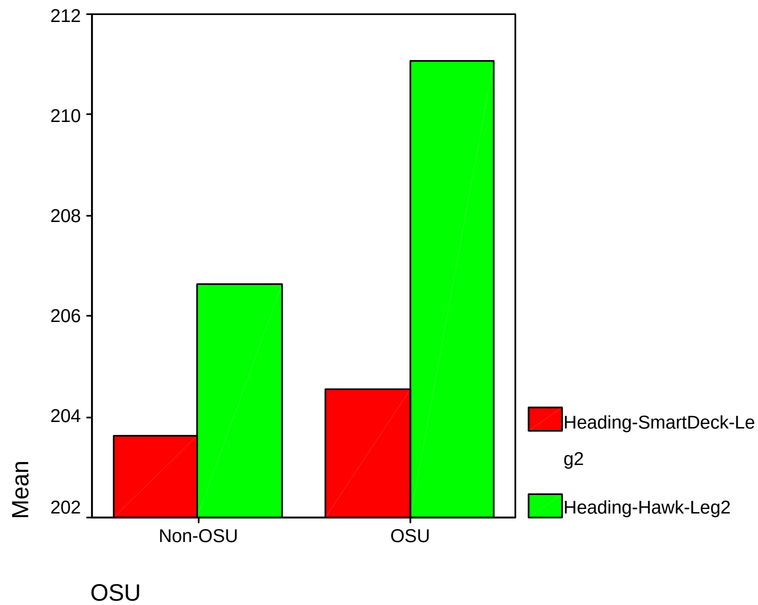

Figure 11: Heading data for Leg 2.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
DISPLAY	Sphericity Assumed	363.604	1	363.604	36.839	.000	.551	36.839	1.000
	Greenhouse-Geisser	363.604	1.000	363.604	36.839	.000	.551	36.839	1.000
	Huynh-Feldt	363.604	1.000	363.604	36.839	.000	.551	36.839	1.000
	Lower-bound	363.604	1.000	363.604	36.839	.000	.551	36.839	1.000
DISPLAY * OSU	Sphericity Assumed	49.226	1	49.226	4.987	.033	.143	4.987	.580
	Greenhouse-Geisser	49.226	1.000	49.226	4.987	.033	.143	4.987	.580
	Huynh-Feldt	49.226	1.000	49.226	4.987	.033	.143	4.987	.580
	Lower-bound	49.226	1.000	49.226	4.987	.033	.143	4.987	.580
Error(DISPLAY)	Sphericity Assumed	296.099	30	9.870					
	Greenhouse-Geisser	296.099	30.000	9.870					
	Huynh-Feldt	296.099	30.000	9.870					
	Lower-bound	296.099	30.000	9.870					

a. Computed using alpha = .05

Table 5: Heading in Leg 2—Main Within-Subjects and Interaction Effects.

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	1364098.414	1	1364098.414	278119.6	.000	1.000	278119.604	1.000
OSU	58.317	1	58.317	11.890	.002	.284	11.890	.916
Error	147.142	30	4.905					

a. Computed using alpha = .05

Table 6: Heading in Leg 2—Main Between-Subjects Effect.

Leg 3 Altitude

Figure 12 graphically depicts the performance differences from the altitude measure in Leg 3. The optimal target for this leg of the scenario was 9,000 ft. The overall analysis of variance of the altitude for Leg 3 (Table 7) indicates a significant main effect for the within-subjects factor of display type ([F (1,30)= 49.48, p<0.001]). This suggests a difference in altitude maintenance between the SmartDeck platform and the Hawk platform for altitude in Leg 3. There is no significant interaction effect ([F (1, 30)=.97, p = 0.332]) between the pilots source (OSU versus Non-OSU) and display. The main effect for the between-subjects factor (Table 8) was also not significant ([F (1, 30)=1.99, p = 0.168]).

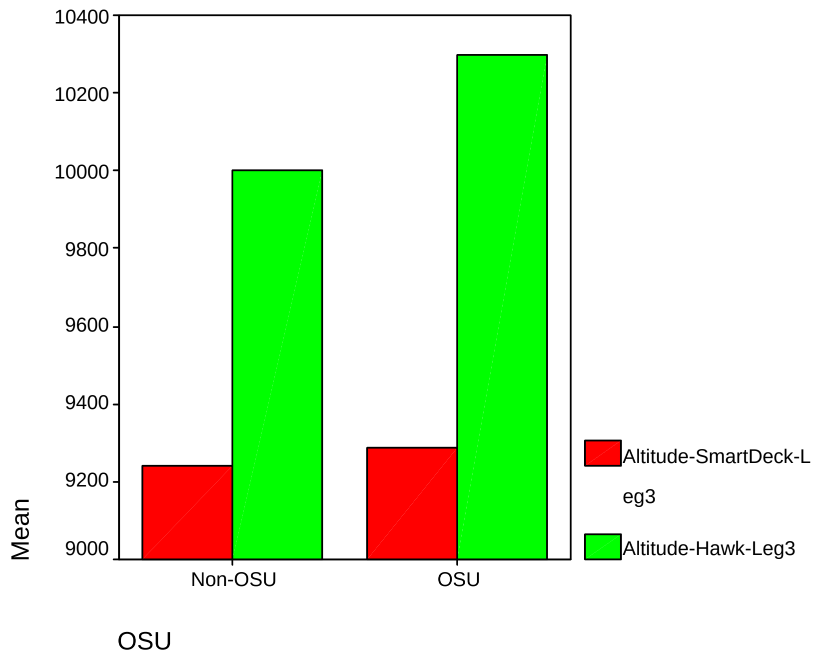

Figure 12: Altitude data for Leg 3.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
DISPLAY	Sphericity Assumed	12539691.3	1	12539691.30	49.483	.000	.623	49.483	1.000
	Greenhouse-Geisser	12539691.3	1.000	12539691.30	49.483	.000	.623	49.483	1.000
	Huynh-Feldt	12539691.3	1.000	12539691.30	49.483	.000	.623	49.483	1.000
	Lower-bound	12539691.3	1.000	12539691.30	49.483	.000	.623	49.483	1.000
DISPLAY * OSU	Sphericity Assumed	246014.584	1	246014.584	.971	.332	.031	.971	.159
	Greenhouse-Geisser	246014.584	1.000	246014.584	.971	.332	.031	.971	.159
	Huynh-Feldt	246014.584	1.000	246014.584	.971	.332	.031	.971	.159
	Lower-bound	246014.584	1.000	246014.584	.971	.332	.031	.971	.159
Error(DISPLAY)	Sphericity Assumed	7602477.954	30	253415.932					
	Greenhouse-Geisser	7602477.954	30.000	253415.932					
	Huynh-Feldt	7602477.954	30.000	253415.932					
	Lower-bound	7602477.954	30.000	253415.932					

a. Computed using alpha = .05

Table 7: Altitude in Leg 3—Main Within-Subjects and Interaction Effects.

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	3014838701	1	3014838701	24809.947	.000	.999	24809.947	1.000
OSU	242021.614	1	242021.614	1.992	.168	.062	1.992	.277
Error	3645520.086	30	121517.336					

a. Computed using alpha = .05

Table 8: Altitude in Leg 3—Main Between-Subjects Effect.

Leg 3 Heading

Figure 13 graphically depicts the performance differences from the heading measure in Leg 3 (Table 9). The optimal target for this leg of the scenario was 181 degrees. There is marginally significant interaction effect ([F (1, 30)=4.12, p = 0.051]) between the pilots source (OSU versus Non-OSU) and display in which the performance by the OSU pilots was considerably lower in the Hawk compared to performance in the SmartDeck. The main effect of the within-subjects factors was non-significant ([F(1,30)=2.01, p=.166]). The between-subjects effect (Table 10) was also non-significant (F(1,30)=3.48, p=.072]).

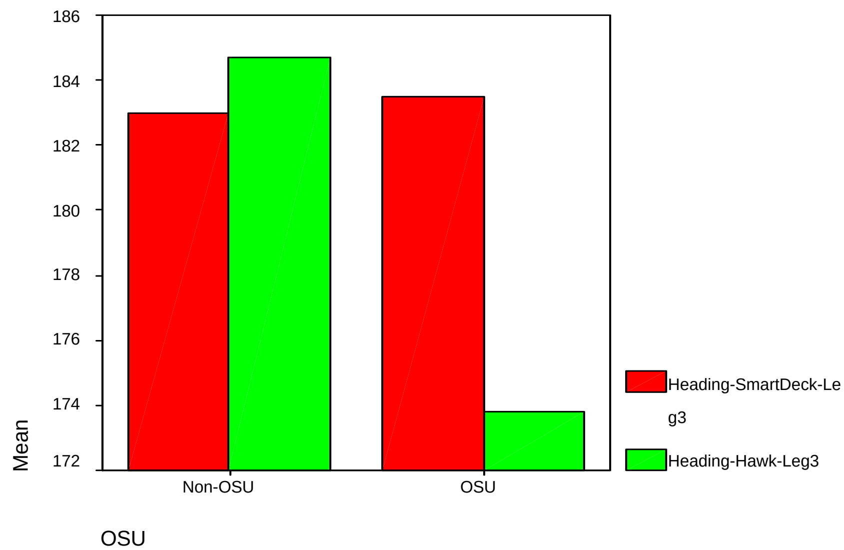

Figure 13: Heading data for Leg 3.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
DISPLAY	Sphericity Assumed	253.190	1	253.190	2.013	.166	.063	2.013	.279
	Greenhouse-Geisser	253.190	1.000	253.190	2.013	.166	.063	2.013	.279
	Huynh-Feldt	253.190	1.000	253.190	2.013	.166	.063	2.013	.279
	Lower-bound	253.190	1.000	253.190	2.013	.166	.063	2.013	.279
DISPLAY * OSU	Sphericity Assumed	517.897	1	517.897	4.118	.051	.121	4.118	.502
	Greenhouse-Geisser	517.897	1.000	517.897	4.118	.051	.121	4.118	.502
	Huynh-Feldt	517.897	1.000	517.897	4.118	.051	.121	4.118	.502
	Lower-bound	517.897	1.000	517.897	4.118	.051	.121	4.118	.502
Error(DISPLAY)	Sphericity Assumed	3773.283	30	125.776					
	Greenhouse-Geisser	3773.283	30.000	125.776					
	Huynh-Feldt	3773.283	30.000	125.776					
	Lower-bound	3773.283	30.000	125.776					

a. Computed using alpha = .05

Table 9: Heading in Leg 3—Main Within-Subjects and Interaction Effects.

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	1051230.661	1	1051230.661	17081.622	.000	.998	17081.622	1.000
OSU	213.874	1	213.874	3.475	.072	.104	3.475	.438
Error	1846.249	30	61.542					

a. Computed using alpha = .05

Table 10: Heading in Leg 3—Main Between-Subjects Effect.

Time

Task duration in the SmartDeck system was calculated from when the subject pressed the button—“next page” after the runway clearance had been issued to the subject’s pressing the “Accept” button on the MFD (after first choosing the “Approach Mode”). For the KLN 89B GPS unit, task duration was calculated from the subject going to the page “Flight Plan” (after the runway clearance had been issued) to the subject coming back to the “NAV” page. None of the ERAU participants or the Non-OSU participants that were presented with the first scenario (that did not explicitly ask for a runway change) made the correct change to the new runway. Occasionally, pilots indicated they intended to change the runway (by saying “switch to runway 28” after being asked to “say intentions”) but never completed the runway change.

After the scenario was altered (and applied to the second group—OSU pilots), there were 10 subjects out of 16 who completed the procedure of switching the runway with SmartDeck. And 8 out of 16 subjects changed the runway successfully with the Hawk. The time that they spent on changing the runway was recorded. The issue of unequal sample size has to be considered. Since there is no function to deal with the unequal sample size analysis of variance in SPSS package, the method of analyzing unequal sample size introduced by Keppel (1991) was used.

The result of the analysis of variance showed that there is a significant effect ($F = 16.50$, $p < 0.05$) on the within-subjects factor. The average of the time that the subjects spent on the procedure on SmartDeck (mean=43.31 seconds, SD=23.26) was less than that on the Hawk (mean=87.875 seconds, SD=71.12). From this result, we can conclude that the SmartDeck method and procedures did significantly reduce the time needed to change runways.

Workload

Table 11 provides the descriptive statistics for the workload dependent variable as measured by the NASA-TLX rating scale (Figure 14). An ANOVA (Table 12) performed on this within-subject factor of display type indicated a main effect ($F[1, 14]=23.63$, $p<.000$]). This suggests that OSU pilots subjectively felt that they had less workload using the SmartDeck system compared to the Hawk system while changing the runway.

Descriptive Statistics

	Mean	Std. Deviation	N
SMARTDEC	53.6467	15.52191	15
HAWK	74.2333	8.73725	15

Table 11: Descriptive Statistics for Workload.

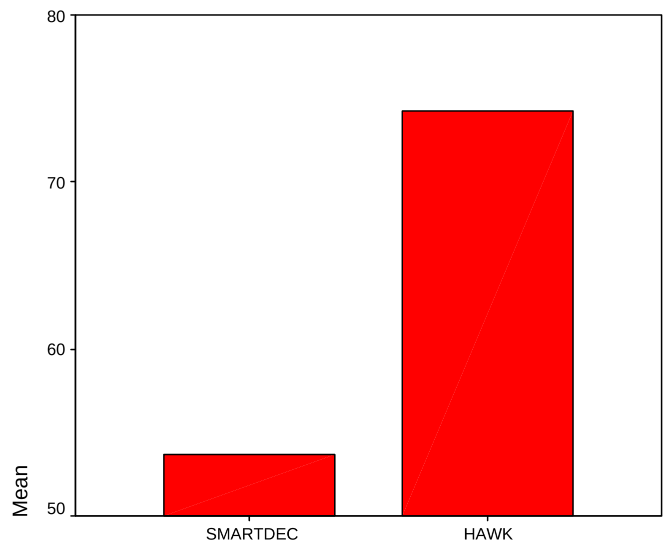

Figure 14: Mean Workload Ratings.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
WORKLOAD	Sphericity Assumed	3178.581	1	3178.581	23.625	.000	.628	23.625	.995
	Greenhouse-Geisser	3178.581	1.000	3178.581	23.625	.000	.628	23.625	.995
	Huynh-Feldt	3178.581	1.000	3178.581	23.625	.000	.628	23.625	.995
	Lower-bound	3178.581	1.000	3178.581	23.625	.000	.628	23.625	.995
Error(WORKLOAD)	Sphericity Assumed	1883.639	14	134.546					
	Greenhouse-Geisser	1883.639	14.000	134.546					
	Huynh-Feldt	1883.639	14.000	134.546					
	Lower-bound	1883.639	14.000	134.546					

a. Computed using alpha = .05

Table 12: Workload—Main Within-Subjects Effect.

Discussion

Workload studies are notoriously fickle in their ability to demonstrate effective findings. To avoid this instability, this experiment was designed to increase the confidence of finding workload effects by combining findings through multiple efforts. This study began with the best of intentions but fell short in results. This study was plagued by issues that limited the ability to combine findings along the way. For example, the original intent of this study was to coordinate common simulation platforms included at both ERAU and OSU in order to address workload issues between the SmartDeck platform (available to both campuses) and two similarly configured general aviation aircraft (Hawk at OSU and Cessna 172 at ERAU). The lack of access to the general aviation aircraft at Embry-Riddle prevented this pooling of efforts. Differences in performance in subject populations also prevented pooling of data as well. Other obstacles for the project included technological issues, communication problems, and coordination of common protocols. Despite these issues, certain claims may be made based on the data.

Input Devices

Differences in workload due to the input interface into the simulation platform were not able to be determined due to limitations in technology. The voice recognition system for the SmartDeck simulation platform, Verbex, never became operational. Despite the fact that the Verbex system is only a few years old, changes in technology in that period of time have made this voice recognition system obsolete. One of the predominant limitations on the voice recognition hardware was the fact that its interface required a DOS operating system. The Verbex hardware required the system to be trained in DOS on a computer separate from the SmartDeck system, then ported to the SmartDeck platform. Researchers at ERAU tried the Verbex system on eight different computers (e.g. 286 and 386 architectures as well as more contemporary models) in an attempt to make the system operational. During this search, calls were also made to Goodrich, who had worked with the system to integrate it with their displays, but this assistance was limited as the personnel had not worked with this system themselves in over a year and even then had only limited success. A computer was finally discovered that could interface with the Voxware system but quickly led to problems in training the voice recognition hardware. In order to make the system operational, separate voice profiles needed to be created for both male and female pilots. In order to train the system, the voice recognizer had to be initialized and retrained. The best efforts of ERAU and Goodrich were not able to initialize the system, nor train the system. The original voice recognizer, that had been previously loaded into the software but was inadequately trained, was lost during some of the effort to make the system operational. During this process, contacts were attempted with the manufacturer of the Verbex system, Voxware, on a number of occasions, with no results. In addition, the manuals for the system were cryptic and insufficient to provide guidance on how to initialize or train the hardware. The Verbex system does appear to have potential as a voice-recognition system that could interface with the SmartDeck system, but it would require a newer model in order to be effective at this point in time.

Display Type

It was expected that participants from both OSU and ERAU could be combined to increase detection of performance differences between display types on performance and workload measures. This was not able to be accomplished. The t-tests comparing the non-OSU pilots and the ERAU pilots indicated a wide difference in performance, negating the ability to pool the data. This distinction is most likely due to the large difference in flight time between the two groups. The non-OSU group had an average of over 3,000 hours of flight time while the same group at ERAU had only an average of around 850 hours of flight time. This led to a much higher variability in performance from the Embry-Riddle participants when compared to the OSU participants.

In the leg 2 data, differences were seen from the prescribed altitude of 10,000 feet and heading of 197 degrees. This is to be expected, given that it takes some time to descend and turn from the prior waypoint. The deviations from the prescribed altitudes and headings were in the expected direction. There were differences seen between the different platform types in both the altitude and heading data. This suggested that the display format found in the SmartDeck system provided better support to the pilots in maintenance of their altitude during the flight plan, despite two extra tasks being given to them during this leg of the flight plan. The interaction found in the heading data (that was absent in the altitude data) is most likely due to the cost incurred from the inexperience of the pilots in the OSU condition when they encountered the Hawk. That is, while the Hawk does not contain the pictorial representation of the HITS display that the SmartDeck system has for heading information, the greater experience of the pilots in the non-OSU group mediated the display cost from the Hawk compared to the OSU group, even though both groups have had exposure to the Hawk platform. Notice, also, that performance in both groups was near the expected heading with the SmartDeck displays.

In the leg 3 data, a similar pattern was seen in the altitude performance. A main effect for display type was found yet again for the third leg of the scenario. This is especially important as the number of tasks during this portion increased, as well as events leading to the runway change were required. A large difference was again found between the SmartDeck and Hawk for both between-subjects groups, indicating a benefit from the displays on the SmartDeck platform, especially given that both groups had some exposure to the Hawk platform already. The interaction found in the heading data for leg 3 of the scenario is more interesting. However, this interaction can be easily explained by pointing that the non-OSU group did not change the runway as expected, therefore was able to concentrate more on flying. These findings lend some support to the fact that additional tasks during leg 3 (especially the messages and requirement to change the runway) did influence maintenance of heading and altitude in flight. Why the OSU participants were so far under (rather than over) the correct angle in the Hawk condition (compared to the other conditions) is not clear.

Thus, overall, it was clear that the HITS display in the SmartDeck simulator provided assistance in flight path tracking compared to the Hawk simulation. This is not surprising since HITS is a natural way of presenting horizontal and vertical flight path information according to the pilots' visualizations of their position in 3D space. Thus, it reduced the need for pilot's mental integration of several representations and reduced workload. The better performance also can be attributed to the provision of preview. The preview indicates where the pilot should be, providing directly perceivable information about the future required status of the aircraft. In this experiment, HITS provides the preview of both the boxes and the "highway". When the airplane is going to be turning, the pilots could see the turn trend in the pathway. In this experiment, the vertical course deviation on SmartDeck was much smaller than that on Hawk. Without preview, errors in course deviations will increase.

Workload

The purpose of this study was to investigate workload differences. Differences in workload were difficult to determine in this study but the data that is present is clear. ERAU did not have the complete NASA-TLX form with which to score data and this error was not realized until long after participants had finished the experiment. That is, the page of relative weighting scales was missing (see Appendix A). Therefore, workload measures from this group could not contribute to the overall findings. However, data from the OSU participants did indicate, both in the timing data and the subjective workload measures, that the SmartDeck platform was better for performance over the Hawk. Although it is necessary to consider that there may also be differences in the difficulty level of flying each of these two very different simulations, it can be concluded that the procedures for switching runways provided by SmartDeck are easier and quicker than that provided by the KLN-89B GPS unit as indicated by the timing data and subjective workload measures. It should also be noted that two more OSU pilots succeeded in changing the runways with SmartDeck than succeeded using the KLN-89B, even though the pilots were used to using the Hawk system.

Additionally, most subjects liked using the touch screen on the MFD because it has a bigger screen and the description on the buttons made a lot of sense to them compared to the instrumentation in the Hawk. For example, the "WEATHER" button means that when they press this button, they could get information about weather. With the Hawk, the screen is quite small, and the pilots have to use a bigger outer knob and smaller inner knob to interact with the GPS unit, which are really confusing to some pilots. These pilots always went to the wrong pages because they turned the wrong knob. There are also some buttons under the screen. Because the instructions on the screen are not very clear, some subjects made a lot of mistakes by trying to use the "Direct" button.

Although it took less time to complete the procedure of switching the runway with SmartDeck, some subjects still complained that there were too many steps in order to change the runway, especially if they had to go to different pages to choose the runway. The fact is that most of them could not remember what to do after they chose the runway on the VNAV page and got back to the HNAV page. At least one subject suggested that the procedure should be automated, which means that after choosing the runway and hitting the ACCEPT, the system should be automatically changed to another runway by showing the new pathway. As for the Hawk, the procedure is more complicated because the pilots had to go to other pages too, by using the knobs.

This data provides at least some small support to the fact that the next-generation of displays may be robust under higher workload conditions.

Further directions

It is clear from this study that a common frame of reference is not always easy in coordination. One problem that this study contains was the lack of communications between contributors to this study. ERAU was not aware that OSU had decided to alter the protocol and perform a task in which they directly indicated to participants to change a runway. If both agencies had done this, additional data may have supported the workload findings more strongly. Despite these communications issues and other technological problems described elsewhere, the findings are still relatively clear that the SmartDeck system requires less time to make corrections under high workload conditions and ratings of subjective workload by pilots is also reduced with the SmartDeck platform.

Training, also, seemed to play a role in confusion for pilots at ERAU. After completion of the computer-based training, many pilots were still unsure of what to do with the system despite having ample time to explore the system. The link between the training and actual implementation in the simulator was tenuous. Additionally, many pilots wished to divert to another runway in the scenario, rather than recognize that another option in the form of a GPS-based descent was available to them even though they had the approach plate to refer to during the experiment. This could be a lack of knowledge or a prescribed procedure for pilots to follow. The explicit request performed by OSU appears to address this issue by increasing the number of pilots that utilized the GPS system to alter the new runway in both platforms, yet in both cases only around half of the participants still managed to do so. Thus given this scenario the issue may be less of workload and more an issue of knowledge of the system. A better crafted scenario would help to address this issue or perhaps just using participants with a higher flight time, despite the instrument rating.

References

Bailey, L. L., Thompson, R. C. (2001). The TLX: One or more constructs. *Proceedings of the Eleventh International Symposium on Aviation Psychology*, Columbus, Ohio.

Doherty, S. M., & Wickens, C. D. (2001). Effects of preview, prediction, frame of reference, and display gain in tunnel-in-the-sky displays. *Proceedings of the Eleventh International Symposium on Aviation Psychology*, Columbus, Ohio.

Fadden, S., & Wickens, C. D. (1997). Improving Traffic Awareness with a head-up flight path highway display. *Proceedings of the Ninth International Symposium on Aviation Psychology (329-333).*

Fadden, S., Ververs, P., & Wickens, C. D. (2001). Pathway HUDs: Are they viable? *Human Factors, 43*, 173-193.

Hart, S. G., & Staveland, L. E. (1988). Development of the NASA task load index (TLX): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload, 139-183.* Howell, D.C., (1997). Statistical methods for psychology. Belmont CA. Wadsworth Publishing Co.

Keppel, G, (1991). *Design and Analysis: a Researcher's Handbook*. Prentice-Hall, Inc.

Nendick, M., & George. R. (1995). GPS: Developing a human factors training course for pilots. In B.J. Hayward and A.R. Lowe, *Applied Aviation Psychology.*

Wickens, C. D. (1997). Attentional issues in head-up displays. In *Engineering Psychology and Cognitive Ergonomics: integration of Theory and Application* (pp. 3-22). London: Avebury Technical.

Wickens, C. D., & Long, J. (1995). Object- vs. space-based models of visual attention: Implications for the design of head-up displays. *Journal of Experimental Psychology: Applied, 1*, 110-135.

Appendix A
NASA TLX Workload Index: Survey Form

Subject ID: ____________
Name: ________________
Date: _________________

Rating Sheet

MENTAL DEMAND

Low High

PHYSICAL DEMAND

Low High

TEMPORAL DEMAND

Low High

PERFORMANCE

Good Poor

EFFORT

Low High

FRUSTRATION

Low High

You are presented with a series of pairs of rating scale titles and asked to choose which of the items was more important to your experience of workload in the task that you performed.
<u>Select the Scale Title that represents the more important contributor to workload for the specific task you performed in this experiment.</u>

Mental Demand
Physical Demand

Performance
Effort

Frustration
Mental Demand

Effort
Physical Demand

Temporal Demand
Performance

Mental Demand
Effort

Temporal Demand
Effort

Physical Demand
Frustration

Mental Demand
Temporal Demand

Physical Demand
Performance

Temporal Demand
Frustration

Performance
Frustration

Temporal Demand
Physical Demand

Performance
Mental Demand

Effort
Frustration

Rating Scale Definitions

Title	Endpoints	**Descriptions**
MENTAL DEMAND	*LOW/HIGH*	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	*LOW/HIGH*	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	*LOW/HIGH*	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	*GOOD/POOR*	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
EFFORT	*LOW/HIGH*	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION LEVEL	*LOW/HIGH*	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Appendix B
Scenario Tasks

Experimental Scenario 1

Trigger Location	Event	Trigger	Criteria
10 NM from Second Waypoint	Ask participants for elevation of KAFF	Experimenter, Manual	If they give the correct elevation prior to the next event
6 NM from Second Waypoint	Ask participants to change the range on VNAV page.	Experimenter, Manual	If they changed the range of VNAV page prior to the next event.
4 NM from Second Waypoint	ATC Messages	Press ATC Button5: "Convective Weather Alert" 3: "Traffic Alert for KAFF" 4: "Traffic alert for C015"	None
9 NM from the Third Waypoint	ATC Messages	Press ATC Button6: "Runway change"	Participant indicates intention to change to runway 28
Participant indicates intention	ATC Message	Press ATC Button1: "Cleared for the approach"	Participant correctly completes selection of new runway

Experimental Scenario 2

Trigger Location	Event	Trigger	Criteria
10 NM from Second Waypoint	Ask participants for elevation of KAFF	Experimenter, Manual	If they give the correct elevation prior to the next event
6 NM from Second Waypoint	Ask participants to change the range on VNAV page.	Experimenter, Manual	If they change the range on VNAV page prior to the next event.
4 NM from Second Waypoint	ATC Messages	Press ATC Button5: “Convective Weather Alert” 3: “Traffic Alert for KAFF” 4: “Traffic alert for C015”	None
9 NM from the Third Waypoint	ATC Messages	Press ATC Button6: “Runway change” Press ATC Button 2: “Select an Approach” Press ATC Button1: “Cleared for the approach”	Participant correctly completes selection of new runway

Appendix C
ATC Messages

This is the sequence of weather and air traffic control (ATC) messages leading to the necessity of the pilot to change runways at the final destination.

- At Aurora (01V): 33010KT 007 BKN 010 OVC 15/14 A2995
- This meant that at 01V- the wind was coming from 330 degrees at 10 knots. The ceiling was 700 feet and broken, with an overcast at 1000 feet. The temperature was 15 Celsius with a dew point of 14 Celsius. The altimeter setting was 29.95 inches of mercury.

- At Centennial (APA): 34012KT 006 SCT 008 BKN 010 OVC 15/14 A2994
- This meant that at APA—the wind was coming from 340 degrees at 12 knots. The scattered clouds were at 600 feet, and the ceiling was 800 feet broken, with an overcast at 1000 feet. The temperature was 15 Celsius with a dew point of 14 Celsius. The altimeter setting was 29.94 inches of mercury.

- A current Notice to Airmen (NOTAM) was issued: APA ILS35R OTS
- This meant that the Instrument Landing System for runway 35R was out-of-service.

- At 4NM from TURN1: *"Convective Sigmet 4 Central is valid for 150 nautical miles of Colorado Springs, Severe thunderstorms moving from 270 degrees at 25 knots. Tops to 6 5 0. Hail."*
- This was an ATC message warning about the weather.

- At this time the cloud ceiling was dropped to 650 feet.

- At 9NM from NERXY: *"Ceiling at Centennial has now dropped to 650 feet. Say intentions."*
- Since the ceiling is now lower than the approach minimum of runway 35R, the pilot needs to change the approach runway to Runway 28 whose approach minimum is 600 feet, which is lower than the approach minimum of runway 35—666 ft.

- The first scenario included in the design does not indicate which approach the pilot is to use because they are legally not allowed to do so. The change of approaches must be initiated by the pilot. The subject is expected to say something like: "Denver Approach, Jet 123, requests GPS runway 28"

- *"Jet 123, proceed direct, Runway 28 GPS Final Approach Fix, NIDLY."*
- This message from ATC indicates the pilot may proceed to the final approach fix.

- *"Jet 1 2 3, Cleared to land, GPS Runway 28."*
- This message if for final clearance by ATC to land.

Appendix D
Subject Instructions

SUBJECT INSTRUCTIONS

GA Training & Integration of Advanced Cockpit Displays

Welcome to the General Aviation Training and Integration of Advanced Cockpit Displays research program. Because of your instrument rating, you have been asked to participate in this study of a new avionics system that integrates Global Positioning System (GPS) functions into a more general and capable moving map display that includes weather and traffic information, as well as your planned route of flight. The purpose of the study is to determine how long it takes to make changes to the planned course, whether any errors are made doing so, and how well the airplane can be flown while updating the GPS data.

To begin this study, you will be asked to train on a computer-based training device that will instruct you on the procedures used in updating the SmartDeck avionics navigation display in the Small Aircraft Transportation System (SATS) simulation. You will be permitted to ask whatever questions you wish about the system or its operating procedures during the training. Please make certain you touch on **ALL** the sections in the training **EXCEPT** for sections entitled

“**Your Guide to AT^2**” (at the beginning)
“**Monitoring Systems**” (at the beginning)
Reviewing the Checklists” (at the beginning)

Please try your utmost to learn the material presented in the training. There is no penalty for returning to review any of the sections of training and it is encouraged that you do so if it will assist in your learning.

When you are finished with the computer-based training, please inform the experimenter.

SUBJECT INSTRUCTIONS, Part II

GA Training & Integration of Advanced Cockpit Displays

Now that you have completed the computer-based training, we will test you on your knowledge of your skills learned through the training in a series of two flight scenarios using the Small Aircraft Transport System (SATS) simulation and the HAWK simulation. You will be using the flight displays similar to what you have just seen in the training.

During these flights, we ask that you keep in mind the following items:

1. For the SATS simulation, once the scenario has begun, please IMMEDIATELY press the button on the left side of the yoke twice to turn off the “autopilot” and press the button on the right side of the yoke once to turn off the “autothrottle” functions.
2. Please maintain a constant airspeed of 150 knots throughout the duration of the flight for both scenarios.
3. From time to time, additional information or instructions for tasks may be given to you either from the experimenter or via air traffic control (ATC). If you are asked to perform a task, please accomplish the task as quickly and accurately as you can. If you find that another task is given to you before you have completed the previous task, disregard the first task and focus on the new one as that will now have higher priority. Do not return to the previous task.
4. If you need to reference the Approach Plate provided to you, please do so.
5. When you are finished with the first task, please let the experimenter know and the second scenario will begin.

Any further questions? Please ask them now, as during the scenarios you will unable to ask additional questions.

Thank you!

Appendix E

Consent Form

EMBRY-RIDDLE AERONAUTICAL UNIVERSITY
RESEARCH SUBJECT INFORMED CONSENT FORM
"General Aviation Training and Integration of Advanced Cockpit Displays"

Prospective Research Subject: Read this consent form carefully and ask as many questions as you like before you decide whether you want to participate in this research study. You are free to ask questions at any time before or after your participation in this research as well as certain times during the research. These times will be made clear to you as the research progresses.

The purpose of this research is to investigate the use of the Small Aircraft Transport System (SATS) displays for flying a next-generation general aviation aircraft. You will be asked to complete a training tutorial, and then "fly" the SATS simulator through two scenarios to test your learned skills. No deception is involved in this experiment. The entire experiment is expected to take approximately three hours to complete or less. We forsee no risks or discomforts with participation. The $30 for your participation will be paid to you upon completion of the experiment.

The results of this experiment are for research purposes, and may be presented at professional meetings or published in research literature. Your name will not be used in the reporting of results. Only group data will be used; all personal information will be kept completely confidential.

If you wish to withdraw from the experiment, you may do so at any time without penalty.

Following the experiment, I will discuss the reasons for the experiment with you. If you have any questions, please feel free to ask me or feel free to contact Dr. Shawn Doherty, Department of Human Factors and Systems, at 226-6249.

Thank you!

I, ________________________________, understand that my participation in this experiment is completely voluntary and that I may refuse to participate, or withdraw from the experiment, at any time without penalty.

____________________________________ ________________ ________________
Participant Signature Date ERAU Colleage #

____
ERAU Mailbox #

____________________________________ ________________
Experimenter Signature Date

----------------------------------Experimenter Use Only--
Licenses: _______________________________________ Hours ______ Age ____
CBT Start: ___________ CBT End: ____________ CBT break: _________

CONSENT FORM

CONSENT FOR PARTICIPATION IN SOCIAL AND BEHAVIORAL RESEARCH

I consent to participating in research entitled:
General Aviation (GA) Training & Integration of Advanced Cockpit Displays

Dr. Gerald P. Chubb and / or Dr. Richard S. Jensen
(*Principal Investigator(PI) and CoPI*)

or their authorized representative: Ms. Chang Liu

explained the purpose of the study, the procedures to be followed, and the expected duration of my participation. Possible benefits of the study have been described as have alternative procedures, if such procedures are applicable and available.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction. Furthermore, I understand that I am free to withdraw consent at any time and to discontinue participation in the study without prejudice to me.

Finally, I acknowledge that I have read and fully understand the consent form. I sign it freely and voluntarily. A copy has been given to me.

Date: ______________________________ Signed: ______________________________
(Participant)

Signed: ______________________________
(Principal Investigator or his authorized representative)

Signed: ______________________________
(Person authorized to consent for participant – if required)

Witness: ______________________________

HS-027 (Rev. 12/97) – (To be used only in connection with social and behavioral research.)

Appendix F
Demographic Data

Subject	Pilot group	Highest Rating	Fhours	age	gender
1	Non-OSU	INSTRUMENT	400	60	Male
2	Non-OSU	ATP	10000	71	Male
3	Non-OSU	INSTRUMENT	1040	68	Male
4	Non-OSU	COMMERCIAL	550	39	Male
5	Non-OSU	MULTI-ENGINE	1010	27	Male
6	Non-OSU	COMMERCIAL	1030	50	Male
7	Non-OSU	CFI	1500	36	Male
8	Non-OSU	COMMERCIAL	842	44	Male
9	Non-OSU	MEI	2550	39	Male
10	Non-OSU	MEI	2100	27	Male
11	Non-OSU	CFI	400	28	Male
12	Non-OSU	CFI	650	52	Male
13	Non-OSU	CFI	520	46	Male
14	Non-OSU	ATP	27000	74	Male
15	Non-OSU	INSTRUMENT	450	41	Male
16	Non-OSU	INSTRUMENT	4500	59	Male
17	OSU	CFII	1400	30	Male
18	OSU	INSTRUMENT	200	29	Male
19	OSU	INSTRUMENT	147	20	Male
20	OSU	CFII	1050	24	Male
21	OSU	CFI	1250	30	Male
22	OSU	INSTRUMENT	170	21	Male
23	OSU	CFI	550	23	Male
24	OSU	INSTRUMENT	170	23	Male
25	OSU	INSTRUMENT	205	23	Male
26	OSU	INSTRUMENT	175	22	Female
27	OSU	CFII	3300	53	Male
28	OSU	COMMERCIAL	1020	26	Female
29	OSU	CFII	1900	35	Male
30	OSU	COMMERCIAL	1800	25	Male
31	OSU	INSTRUMENT	95	20	Male
32	OSU	CFI	480	27	Male

Subject	Pilot group	Highest Rating	Fhours	age	gender
33	ERAU	CFII	800	24	Male
34	ERAU	COMMERCIAL	675	50	Male
35	ERAU	INSTRUMENT	130	19	Male
36	ERAU	CFI	400	21	Female
37	ERAU	MULTI-ENGINE	370	28	Male
38	ERAU	INSTRUMENT	107	19	Male
39	ERAU	CFI	300	22	Male
40	ERAU	COMMERCIAL	270	23	Female
41	ERAU	COMMERCIAL	215	20	Male
42	ERAU	COMMERCIAL	195	20	Male
43	ERAU	COMMERCIAL	185	21	Male
44	ERAU	COMMERCIAL	250	21	Male
45	ERAU	INSTRUMENT	1200	29	Male
46	ERAU	MULTI-ENGINE	260	21	Male
47	ERAU	CFII	250	19	Male