Evaluation of a Peripherally-Located Instrument Landing Display Under Dual-Task Conditions

Zachary Spielman

Joseph Vargas

Taleri Hammack

Nathan Bulkey

Roger Lew

See next page for additional authors

Authors
Zachary Spielman, Joseph Vargas, Taleri Hammack, Nathan Bulkey, Roger Lew, and Brian P. Dyre

EVALUATION OF A PERIPHERALLY-LOCATED INSTRUMENT LANDING DISPLAY UNDER DUAL-TASK CONDITIONS

Zachary Spielman, Joseph Vargas, Taleri Hammack, Nathan Bulkley, Roger Lew & Brian P. Dyre
University of Idaho
Moscow, Idaho

Previous research found that a peripherally-located instrument landing system (ILS) embedded in a head-up display (HUD) supported equal or better control of glide-path during simulated approach and landing than the traditional centrally-located MIL-STD ILS. Here, we used a dual-task paradigm to examine whether gains in landing precision with the peripheral ILS are also accompanied by a reduction in mental workload. Participants controlled glide-path during simulated instrument landings while simultaneously performing a secondary task monitoring a head-down engine display for fault states. We varied the type of ILS (peripheral vs. MIL-STD) and assessed mental workload using the NASA-TLX and primary and secondary task performance measures: glide-path errors and engine-fault detection sensitivity, respectively. We found equivalent glide-path errors for the two displays, but the peripheral ILS produced lower subjective estimates of mental workload and significantly less dual-task decrement in engine-monitoring sensitivity, indicating that this display affords effective glide-path control with lower reduced mental demand.

Non-alphanumeric displays incorporating more naturalistic and less symbolic representations of actual flight parameters—known as *virtual displays*—can enhance aircraft pilots' spatial orientation, increase overall performance, and reduce workload (Hettinger, Brickman, Roe, Nelson, & Haas, 1996). Using simulations of approach and landing, Bulkley, Dyre, Lew, & Caufield (2009) showed that an instrument landing system (ILS) using a peripherally-located head-up display (HUD) with virtual symbology composed of moving arrows can provide superior landing precision as compared to the military standard ILS symbology (MIL-STD-1787B) presented in its normal central location within the HUD. Furthermore, Bulkley, Spielman, and Dyre (2011) found that such virtual HUD symbology can be restricted to smaller spatial extents in the far visual periphery to reduce HUD clutter without compromising landing performance. Part of the rational for development of these peripheral displays was that in addition to reducing central visual field load and clutter, the displays should also reduce the mental demands of the ILS. While such reductions in mental workload have been found for peripheral virtual displays for controlling airspeed (Cox, 2000), so far the mental demands of the peripheral ILS HUD have not been examined. Here we compare the mental workload imposed by the peripheral ILS to the traditional MIL-STD ILS using two workload assessment methods: a) subjective estimates measured with the NASA-TLX (Hart & Staveland, 1988) and b) dual-task performance, which required participants to perform a secondary visual engine monitoring task while controlling glide-path under instrument flight conditions (IFC).

Roscoe (1980) estimated that 90% of aircraft control is performed using the central visual field. To accurately read most aviation displays, pilots must directly fixate their gaze upon the display and allocate enough mental resources to interpret the alpha-numeric information presented. This workload is particularly high during approach and landing, and can be exacerbated by pilot fatigue (Hart & Hauser, 1987). A virtual display could present landing information to pilots without taxing overburdened central visual field and attentional resources by taking advantage of automatized orienting and motion coding processes that are particularly robust in peripheral vision. Cox (2000) and Schaudt, Caufield, and Dyre (2002) showed that for simulated flight, a virtual speed error indicator composed of moving fields of arrows projected to the visual periphery provided better flight path and airspeed control than a head-up display (HUD) speed indicator defined by MIL-STD-1787B, while simultaneously lowering subjective workload. These studies clearly demonstrated that the peripheral visual field can process the optical flow created by moving display elements and extract meaningful speed information with less attentional demand and central visual field resources than traditional symbolic displays.

More recently, Bulkley et al. (2009) found that a similar peripherally-located virtual display could be adapted to replace the flight command bars of a traditional ILS display, which are typically presented in the center of the HUD (as defined by MIL-STD 1787B). The ILS command bars instruct pilots to the direction the aircraft needs to be moved to attain proper lateral and vertical alignment with the optimal approach path during landing. The display contains a horizontally-oriented bar that moves up and down to indicate vertical angular errors from the

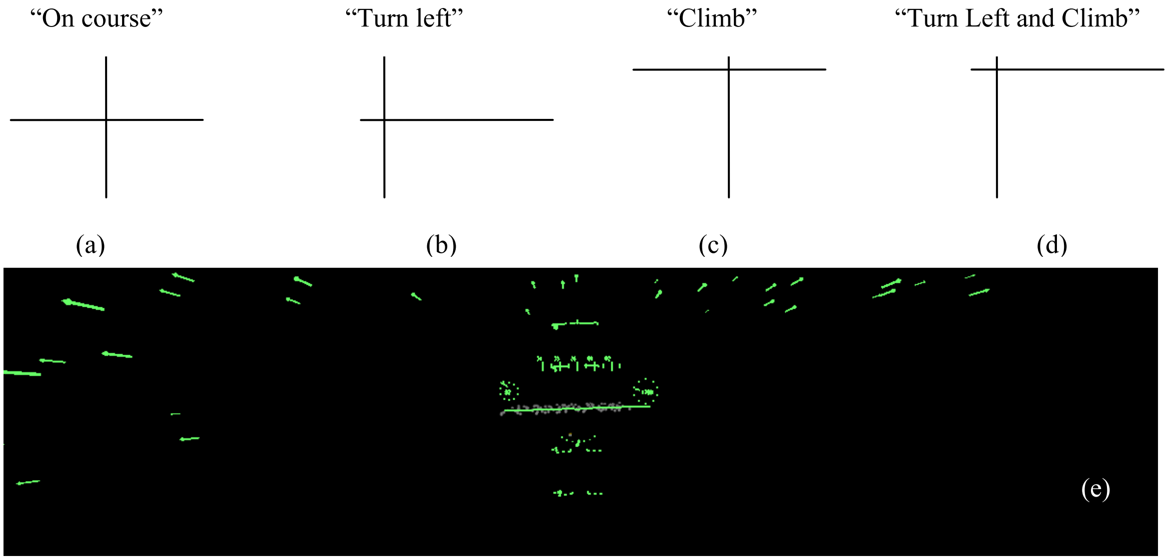

Figure 1. Panels (a)-(d) show the format of MIL-STD 1387B HUD ILS command bars. The four panels show display configurations representing: (a) zero lateral and vertical error, on course, no control correction needed, (b) rightward lateral error, the aircraft needs to turn left to reestablish the optimal approach path, (c) downward vertical error, the aircraft needs to climb to reestablish the optimal approach path, and (d) both downward and rightward errors, the aircraft needs to turn left and climb to reestablish the optimal approach path. Panel (e) shows the peripheral optical flow ILS configuration that would be equivalent to the MIL-STD ILS shown in panel (d), instructing the pilot to “Turn-left and Climb.”

ideal glide slope, and a vertically-oriented bar that moves left-to-right to indicate lateral angular errors from the ideal approach path (see Figure 1, panels a-d). When the bars form a symmetrical cross “+” the aircraft is optimally aligned with the ideal approach path. Movements of the bars from the “+” configuration indicate that control inputs are needed to regain the optimal path, and serve as command displays for how the pilot must maneuver the aircraft. In essence, pilots “chase the bars.” Bulkey et al. (2009) adapted the peripherally-located virtual HUD format developed for representing airspeed by Cox (2000; see also Schaudt et al., 2002) to an ILS display by simply replacing the centrally-located flight command bars with fields of moving arrows located in the visual periphery (See Figure 1, panel e). In essence, pilots “chase the arrows.” The magnitude of lateral and vertical errors is redundantly coded by the speed of the moving arrows and their size, which both increase as errors increase up to a maximum limit. Importantly, size changes occur in discrete steps, creating transient events that can alert a pilot to increasing errors *pre-attentively* (Egeth & Yantis, 1997), even when the displays are presented in the visual periphery. Also, to help reduce display clutter, the arrows disappear completely when errors fall below a minimum threshold.

Here, we are interested in determining whether the peripheral ILS does indeed reduce attentional demand and lower mental workload as compared to the MIL-STD ILS. We report results from one experiment that measured mental demands of the two displays, both objectively, by assessing dual-task performance, as well as subjectively, by administering the NASA-TLX workload questionnaire (Hart & Staveland, 1988). Our primary task required non-pilot participants to fly a simplified simulation of approach and landing at night with poor visibility and significant wind disturbances under instrument flight rules (IFR) using one of the two ILS displays, which we varied within-subjects. We measured glide-path performance—lateral and vertical deviations from glide-path—on the primary task alone and when it was combined with a secondary engine-monitoring task. Performance on the engine monitoring task was assessed as the sensitivity (A’) of detection of double engine-faults. Primary and secondary task performance and the unweighted NASA-TLX score or RawTLX (Hart, 2006) served as our measures of mental workload.

Method

Participants

The experiment tested 16 volunteer participants from the undergraduate population of the University of Idaho who received course extra credit as compensation. Participants reported normal or corrected-to-normal acuity and no previous aviation piloting experience. Bulkley et al. (2009) showed that non-pilots using simplified flight controls fly similar flight paths to those flown by pilots for simulated visual approaches under “blackhole” conditions (Gibb, Schvaneveldt, & Gray, 2008), which suggests that pilots and non-pilots are able to perceive their dynamic spatial orientation relative to the runway in a similar manner. Hence, we believe that our sample of non-pilots validly represents the same visual processes that pilots use to control their aircraft during landing.

Apparatus and Stimuli

Visual displays simulating flight were created by a set of five personal computers running ViEWER v2.23 (Dyre, Grimes, & Lew, 2009). One computer served as the simulation host, which coordinated the activity of four graphics channels over a local area network. Three of the graphics channels rendered the main forward view of the environment to three display projectors, which front-projected images with a spatial resolution of 1024 x 768 pixels (H x V) at a refresh rate of 60 Hz onto three large screens arranged as three sides of an octagon with the design viewpoint at the center of the partial-octagon, 1.8 m from the center of each screen. Together, the three main screens subtended 135 x 33.75 degrees of visual angle (H x V).

The fourth graphics channel rendered a head-down display of two pairs of engine status gauges on a 0.17m diagonal LCD screen. The screen was located approximately 1m directly in front of the participant and below the center of three main display screens, and projected a 4 x 3 (H x V) degree image centered at a declination angle of 20 degrees. This position ensured the display was not readable when the eyes fixated the HUD depicted in the forward view, though it remained visible in the periphery. Conversely, when the head-down engine display was fixated, the forward view displayed on the three main screens was visible only in the periphery. The head-down display, shown in Figure 2, consisted of temperature and pressure gauges for two engines. The red status indicators moved throughout the trial according to a unique pseudo-random disturbance determined for each trial. The tic marks labeled “high” and “low” indicated the thresholds at which that engine would be considered to have a critical fault. The combination of movements across the four indicators was programmed to produce ten single faults—only one engine showed a critical fault (or two faults)--and ten double faults—both engines showed at least one critical fault simultaneously. We instructed participants to respond whenever they detected a double fault by pulling the trigger button on the joystick.

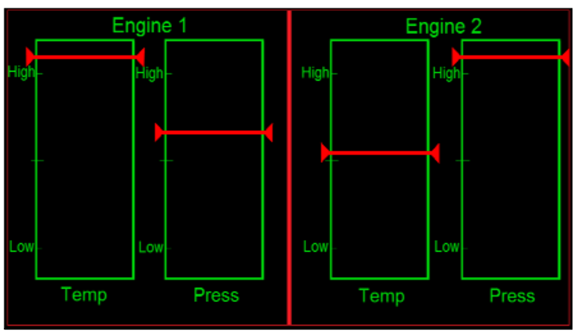

Figure 2. Engine status indicators displayed on the head-down monitor used for the secondary engine fault detection task. These indicators are consistent with a “double-fault” requiring

The simulated environment displayed on the three main screens consisted of a large island extending 288,000 m in width (x) and 72,000 m in depth (z). A 30m wide runway started at the center of this island [x, y, z] = [0, 0, 0] and extended 950 m in depth [0, 0, -950]. Green, white, and red lights were spaced every 50m along the sides of the runway in a standard runway lighting configuration. The runway was texture-mapped with a pavement texture and the surrounding ground surface was texture-mapped with a high spatial frequency seamless texture resembling dirt and grass. At the far end of the runway the ground began to slant upward at 5 degrees for 18,000m before leveling off at a maximum elevation of 1600 m. The virtual environment simulated ambient lighting equivalent to a foggy, overcast, moonless night such that the unlit ground became barely visible at an altitude of 160 m and the runway marker lights became visible at a distance of approximately 2000m..

The simulated view from the vehicle was open, no windscreen frame or other artifacts were used to create a simulated cockpit. The vehicle started 9889m short of the runway at an altitude of 720m, aligned laterally with the center of the runway and moving forward on a level trajectory at 45 ms^{-1}. The target glide-slope angle of 7.3 degrees was defined via two invisible waypoints. The first was located at the vehicle starting point [0, 720, 9889], and the second was centered on the runway, 2m from its near end [0, 0, -2].

Simulated lateral and vertical wind disturbances were produced by translating the vehicle along the lateral (x) and vertical (y) axes independently. Each disturbance was defined by a sum-of-sines with five prime frequencies. Amplitudes were chosen such that the maximum acceleration did not exceed 1 G (9.8 ms^{-2}). The disturbance frequencies for the x- and y-axes were 0.055, 0.085, 0.145, 0.185, 0.215, and 0.035, 0.065, 0.115, 0.155, 0.205 Hz, respectively; amplitudes were 10.47, 2.34, 1.37, 1.08, 0.93 and 25.85, 7.49, 2.39, 1.32, 0.75 m, respectively.

Since participants were non-pilots, simplified flight controls were used to control the lateral and vertical positions of the simulated aircraft. Left-right movements of a CF-F-16 joystick with the right hand controlled lateral velocity with a first-order transfer function with gain = 25 ms^{-1} at maximum stick deflection and exponential lag constant = 1.0 s^{-1}. For simulation of banking while moving laterally, left-right stick movements also produced a zero-order roll with a gain = 15° at maximum deflection and exponential lag constant = 0.5 s^{-1}. Movements of a CF Pro Throttle configured as a first-order controller with gain = 25 ms^{-1} and exponential lag constant = 1.0 s^{-1} with the left hand controlled vertical velocity. Forward movement of the throttle caused the vehicle to go up (simulating more thrust); backward movement caused the vehicle to go down (less thrust).

Similar to Bulkley et al (2009), monochromatic green HUDs were superimposed over the terrain and environmental objects on the display and included the following indicators from MIL-STD-1787B: aircraft pitch reference symbol, climb/dive marker, climb/dive ladder, airspeed indicator, target airspeed indicator, altitude indicator, heading indicator, and bank indicator. Two ILS formats were implemented. One ILS format was the MIL-STD 1787B represented as vertical and horizontle bars refferred to as course deviation and vertical deviation indicator bars respectively (hereafter referred to as *MIL-STD*). The other ILS was implemented within the HUD as a peripherally-located virtual display of fields of moving arrows (hereafter referred to as *Peripheral OF*). Fields of moving arrows randomly arranged within a volume of space appeared in the upper, lower, left or right peripheral areas of the HUD to provide control commands to overcome lateral and vertical deviations from the optimal flight path (see Figure 1). The visual fields subtended by the arrows were equivalent to the far periphery condition of Bulkley et al (2011), since they found no decrement in landing precision when the field of arrows was constrained to the far periphery. Similar to the MIL-STD ILS, the Peripheral OF ILS used a command format that informed pilots which direction to move to correct their course—in effect participants needed to “chase the arrows” to maneuver toward the optimal glide path. The arrows coded lateral and vertical deviation error magnitudes redundantly using both size and speed of movement. Zero course error resulted in zero speed and size--nothing was displayed. Small course errors resulted in small arrows moving slowly. As course errors increased, the speed of the arrows increased proportionately, and the size of the arrows increased in step-wise increments to produce sudden size-changes that naturally captured attention.

Experimental Design and Procedure

A 2 x 2 x 3 within subject factorial design tested the effects of ILS format (MIL-STD, Peripheral OF), block (1 or 2), and task (Single-Task Engine Monitoring, Single-Task Landing, Dual-Task) on glide-path errors (vertical and lateral), engine-monitoring performance, and subjective workload. Testing occurred over two 90-minute sessions, with the ILS format blocked across session. We tested half of the participants first with the MIL-STD format then the Peripheral OF; the other half of participants received the reverse order. Each testing session started with verbal instructions on the physics of flight, relation of control movements to aircraft movements, a review of all display elements, and a description of the engine monitoring task. Following this training, half the participants were first tested on the engine monitoring single task, then the landing single task, while the other half received the reverse order. We tested the engine monitoring single-task across three trials, and the first trial was considered training and not included in the data analysis. To become acquainted with the flight controls for the landing task, participants performed a day-time landing with no ILS using visual flight rules (VFR). Following this, the experimenter demonstrated an IFR landing using the ILS assigned to the participant for that session to familiarize the participant with the ILS format and night-time conditions. We then tested the landing single-task across six trials, and the first four trials were considered practice and not included in the data analysis. At the end of each session, all participants performed three trials of the dual-task condition, of which the first trial was considered practice and not included in the data analysis. Each trial lasted approximately 3 minutes and 40 seconds and participants were able to complete all 14 trials, plus instructions and debriefing within two sessions scheduled on successive days. At no time were the participants instructed to give priority to either of the tasks. After each block of trials for a task, participants completed the NASA TLX.

Results

Glide-path (Primary) Task Performance

Altitude errors were defined as the difference between actual altitude and the target altitude defined by the linear glide slope at a particular point along the approach. Lateral errors were defined as the difference in position of the simulated aircraft relative to the runway centerline. We computed the root-mean-squared (RMS) of these errors for each trial to represent overall error glide-path errors.

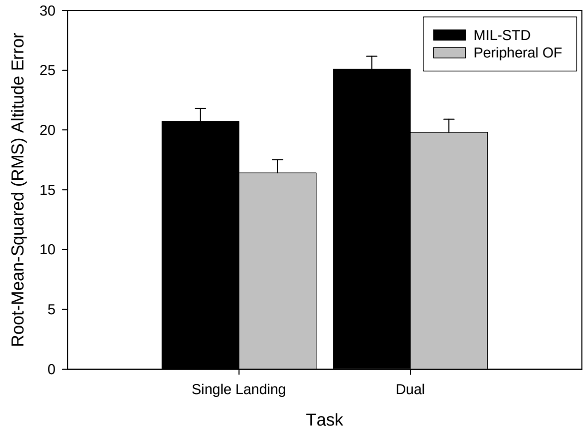

Figure 3. Mean RMS calculated for all participant vertical deviation error. Error bars report average standard error across participants for each category. Error bars represent within-subjects standard errors of the mean calculated from the MSE term of the task x display interaction.

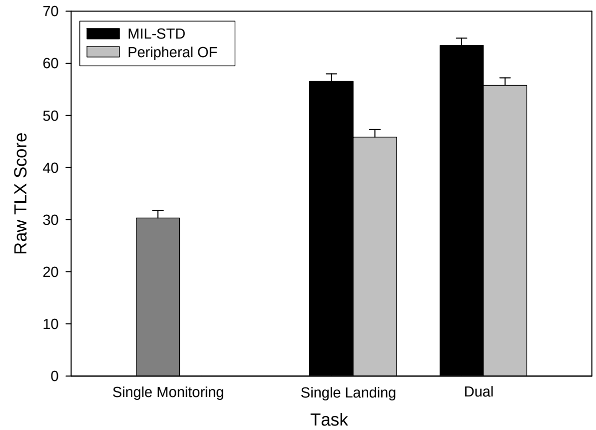

Figure 4. Raw TLX subjective measure workload. Error bars report average standard error across participants for each category. Error bars represent within-subjects standard errors of the mean calculated from the MSE term of the task x display interaction.

RMS lateral and altitude errors were each analyzed using a 2 x 2 x 2 x 2 factorial mixed analysis of variance (ANOVA) with ILS configuration (Peripheral OF, MIL-STD), task (single, dual), and block (1-2) as within-subjects factors and order of ILS configuration as a between-subjects factor. Though the pattern of RMS Altitude error in Figure 3 suggests main effects for ILS and Task, these trends were non-significant, $F(1, 14) = 3.733$, $p = .074$, $MSE = 208.14$, and $F(1, 14) = 3.613$, $p = .078$, $MSE = 50.77$, respectively. No reliable effects were found for lateral error.

Engine Monitoring (Secondary) Task Performance

We computed sensitivity in dual engine fault detection as A' and analyzed these values using 3 x 2 x 2 mixed ANOVA with Task-Display Condition (Single Task Engine Monitoring, Dual-Task Peripheral OF, Dual-Task MIL-STD) and block as the within-subjects factors and order as a between-subjects factor. We found a significant main effect of task-display condition, $F\ (2, 28) = 24.29$ $p = 0.001$, $MSE = .002$. According to this effect, detection performance was best in the Single-Task condition, followed by the Dual-Task Peripheral OF condition, with worst performance in the Dual-Task MIL-STD condition (see Table 1).

Subjective Workload

Data for two participants was lost due to technical glitches. We computed Raw TLX scores for the resulting 14 participants by taking the average score across the six subscales and analyzed these scores using a 2 x 2 x 2 x 2 factorial mixed analysis of variance (ANOVA) with ILS configuration (Peripheral OF, MIL-STD), task (single, dual), and block (1-2) as within-subjects factors and order of ILS configuration as a between-subjects factor. We found significant main effects of Display and Task condition, $F\ (1, 12) = 6.13$ $p = 0.033$, $MSE = 164.35$ and $F\ (1, 12) = 11.88$ $p = 0.006$, $MSE = 71.31$ respectively. The pattern of means, listed in Table 1 and shown in Figure 4, demonstrate that the dual-task condition was significantly more demanding than the single task conditions and that higher demand occurred with the MIL-STD versus the Peripheral OF. We also found a significant ILS-Order interaction, $F\ (1, 12) = 16.73$ $p = 0.002$, $MSE = 164.35$, which showed overall lower workload ratings for testing on Day 2 as compared to Day 1.

Discussion

Our peripheral optical flow ILS clearly out-performed the traditional MIL-STD ILS format in providing precise glide-path control with lower cognitive demand. The peripheral ILS provided at least equivalent performance to the MIL-STD ILS while significantly improving performance on a secondary visual monitoring task, indicating that participants experienced less resource demand while landing with the peripheral ILS. Significantly lower subjective assessments of mental workload for the peripheral ILS confirmed that participants were even aware they experienced lower workload with the peripheral ILS.

Table 1

Glide-path Root-mean-squared (RMS) Errors, Detection Sensitivity (A'), and RAW-TLX Scores by Task/Display Condition

Task/Display	RMS Lateral Error (m)	RMS Altitude Error (m)	Detection A'	RAW-TLX Composite
Single-Task Engine Monitoring	-	-	.964 (.009)	35.01
Single-Task Landing				
MIL-STD ILS	9.60 (0.52)	20.42 (1.09)	-	56.53 (1.44)
Periperhal OF ILS	9.22 (0.52)	16.18 (1.09)	-	45.83(1.44)
Dual –Task Landing				
MIL-STD ILS	9.54 (0.52)	23.5 (1.09)	.887 (.009)	63.40 (1.44)
Periperhal OF ILS	9.03 (0.52)	17.88 (1.09)	.920 (.009)	55.76(1.44)

Note. Numbers in parentheses represent within-subject standard errors (SEs) of the mean. For RMSE these SEs are based on the Display x Task x Subject mean square error (MSE). For detection A' these SEs are based on the Task for the comparison of display configurations.

These results have important implications for the design of HUDs. Clearly, the peripheral visual field is a potentially important visual resource that is underused with current HUDs and may be a particularly valuable resource for processing flight parameters related to spatial orientation, such as ILS control of landing approaches. Peripheral virtual displays of ILS commands appear to provide our non-pilot participants with information in a more natural, pre-attentive manner that lessens resource demand, central visual field load, and display clutter, while still affording performance that is equal to or better than traditional ILS displays. Further research is needed to determine whether these benefits generalize to actual pilots.

References

Ashford, R. (1998). *Approach and Landing Fatal Accidents 1980–1996*, Safety Analysis Report No. SA-98-002, Civil Aviation Authority, London.

Boeing Commercial Airplane Group. (2002). Statistical summary of commercial jet airplane accidents, worldwide operations, 1959-2002. Retrieved December 8,2002, from http://www.boeing.com/news/techissues/pdf/statsum.pdf

Bulkley, N., Dyre, B. P., Lew, R., & Caufield, K. (2009). A Peripherally-located Virtual Instrument Landing display affords more precise control of approach path during simulated landings than traditional instrument landing displays. *Proceedings of the 53rd Annual Meeting of the Human Factors and Ergonomics Society,* 31-35. San Antonio, TX: HFES

Bulkley, N. K., Spielman, Z., & Dyre, B. P. (2011) Peripherally-located Virtual Instrument Landing Displays. In *Proceedings of the 16th International Symposium on Aviation Psychology (ISAP)*, Dayton.

Cox, E. M. (2000). The effect of a virtual indicator on speed control. (Unpublished Master's Thesis, University of Idaho).

Dyre, B. P., Grimes, J. G., & Lew, R. (2009). ViEWER 2.23: A Virtual Environment Workbench for Education and Research. *http://www.webpages.uidaho.edu/~bdyre/viewer.htm.*

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269-297.

Gibb, R., Schvaneveldt, R., & Gray, R. (2008). Visual misperception in aviation: glide path performance in a black hole environment. *Human Factors, 50*(4), 699-711.

Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 904-908). Santa Monica, CA: Human Factors & Ergonomics Society

Hart, S. G., & Hauser, J. R. (1987). Inflight application of three pilot workload measurement techniques. *Aviation, Space, & Environmental Medicine, 58*(5),402-410.

Hart, S., G. & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North Holland.

Hettinger, L. J., Brickman, B. J., Roe, M. M., Nelson, W. T., & Haas, M. W. (1996). Effects of virtually-augmented fighter cockpit displays on pilot performance, workload, and situation awareness. *Proceedings of the Human Factors Society 40th Meeting*, 30-33. Philadelphia, PA: HFES.

Kaber, D. B., Alexander, A. L., Stelzer, E. M., Kim, S., Kaufmann, K., & Hsiang, S. (2008). Defining and psychological modeling of perceived clutter in advanced cockpit displays. *Aviation Space and Environmental Medicine, 79,* 1007-1018.

Khatwa R., Collins R., & Helmreich, R.L. (1999). Data acquisition and analysis working group final report. *Flight Safety Digest,*1-77.

Roscoe, S. (1980). *Aviation Psychology.* Ames, IA: Iowa St. University Press.

Schaudt, W.A., Caufield, K.J., and Dyre, B.P. (2002). Effects of a virtual speed indicator on guidance accuracy and eye movement control during simulated flight. *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society,* 318-322. Baltimore, MD: HFES

Wickens, C. D. (1991). Processing resources and attention. In D. Damos (Ed.), *Multiple task performance.* London: Taylor & Francis.