Stereoscopic Depth Cues for Enhancing Pilot Interpretation of the Artificial Horizon

Arrundell, D.A.; Landman, H.M.; Stroosma, O.; van Paassen, M.M.; Groen, Eric L.; Mulder, Max

Publication date
2023

Document Version
Final published version

Published in
22nd International Symposium on Aviation Psychology

5-31-2023

Stereoscopic Depth Cues for Enhancing Pilot Interpretation of the Artificial Horizon

Dale-Allen Arrundell

Annemarie Landman

Olaf Stroosma

M. M. van Paassen

Eric Groen

See next page for additional authors

Stereoscopic Depth Cues for Enhancing Pilot Interpretation of the Artificial Horizon

Dale-Allen Arrundell[1], Annemarie Landman[1,2], Olaf Stroosma[1], René van Paassen[1], Eric Groen[2], Max Mulder[1]
[1] Delft University Of Technology, Delft, The Netherlands, [2] Netherlands Organisation for Applied Scientific Research, Soesterberg, The Netherlands

Background. Previous studies and accident analyses have shown that pilots can make roll reversal errors when responding to bank angles shown by the artificial horizon in the Primary Flight Display (PFD). In the current study, we tested whether adding stereoscopic depth cues to the artificial horizon may lead to better bank angle representation due to an improved figure-ground separation between the symbols. **Method.** Stereoscopic depth cues were created by using a half-silvered mirror multi-layer PFD, which presented the horizon symbol on a lower layer and the aircraft symbol on a higher layer. A group of 23 non-pilots and 18 general aviation pilots were shown left or right bank angles on this multi-layer PFD as well as on a normal single-layer PFD, with the task to roll the wings level using a joystick. **Results.** In the pilot group, a similar amount of roll-reversal errors was made with both displays (median = 3.3%) with no significant difference, $p = 0.635$. In the non-pilot group, fewer roll-reversal errors were observed with the multi-layer display, but this difference did not reach significance either (median = 3.3% vs. 5.0%, $p = 0.182$). In both pilots and non-pilots, the reaction time was longer in the multi-layer display, which reached significance in the non-pilots ($p = 0.016$) but not in pilots ($p = 0.215$). Participants noticed that the depth was only visible during the start of the session. **Conclusions.** The results suggest that using stereoscopic depth cues are not a viable manner to enhance the figure-ground relation in the artificial horizon.

Spatial disorientation is still one of the major causal factors in cases of loss of control in flight (LOC-I). It was determined to have contributed to 17% of LOC-I accidents in transport and commuter aircraft between 1981–2016, with no signs of a decreasing trend (Newman & Rupert, 2020). The Primary Flight Display (PFD) is the main instrument with which pilots can prevent or counteract spatial disorientation. Several studies have indicated that the bank indication of the artificial horizon is suboptimal, as it can lead to misinterpretations of the bank angle direction as well as incorrect roll inputs known as roll-reversal errors (RREs). In simulator studies, airline pilots were shown to make RREs in 6.9-8.7% of the cases when being shown a PFD with an unforeseen bank angle and attempting to correct to wings-level flight (Müller, Sadovitch, & Manzey, 2018; Van den Hoed et al., 2022). When spatially disorienting roll cues preceded PFD presentation, this percentage increased to

20% of total cases and 39% of the first encounter (Van den Hoed et al., 2022). Examples of accident cases that have been associated with spatial disorientation-induced RREs include Kenya Airways Flight KQA507 (Cameroon Civil Aviation Authority, 2010), Flash airlines flight 602 (Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile, 2009), and Crossair flight 498 (Aircraft Accident Investigation Bureau, 2002).

These misinterpretations of the artificial horizon are thought to be caused by sub-optimal display design. According to the so-called "figure-ground" principle, we normally perceive objects moving in the foreground against the fixed horizon in the background. In standard artificial horizon displays, the horizon symbol is the part that moves, while the aircraft symbol is fixed, which hampers quick interpretation (Grether, 1947; Johnson & Roscoe, 1972; Roscoe, Corl, & Jensen, 1981). The format of current displays does allow for integration with head-up displays. Hence, it is relevant to investigate display adaptations that may enhance the PFD in its current format.

In the current study, stereoscopic depth cues are investigated as a means to improve the figure-ground relation in the artificial horizon, in order to prevent RREs. These cues are presented using a multi-layer display (MLD), presenting the aircraft symbol and horizon symbol on two different layers. As this produces a different visual image in each eye, the cues are stereoscopic. MLDs have previously been used to separate categories of information to improve users' information uptake and to prevent clutter (Dünser, Billinghurst, & Mancero, 2008; Hayes, Moore, & Wong, 2006). They also have been applied to make information more salient to improve search performance (Wong, Joyekurun, Nees, Amaldi, & Villanueva, 2005). However, to our knowledge, MLDs have not yet been used to improve figure-ground separation in cockpit displays.

Method

Design

The experimental tasks were first performed by a group of non-pilots to obtain more information on optimal depth between two layers of the MLD. These non-pilots performed the tasks with a single layer display (SLD) PFD (baseline condition) and with a MLD with either low or high depth between the two layers (two groups, randomly assigned). The effect of MLD and that of MLD depth were tested using a mixed-model design.

Using these outcomes, one layer of depth was chosen to further test with a group of private pilots. This group performed the same tasks in the baseline condition and with the MLD with the depth that had the most effect in the non-pilots. The effect of the MLD was tested using a within-subject comparison.

Participants

Of the non-pilot participants (n = 23, 18/5 male/female, mean age = 24.8, *SD* = 6.9), 5 had never seen nor used an artificial horizon, 4 knew what it is but had never used it, 10 had some experience with flight simulators at different fidelity levels, and 8 had experience with glider flying. None had experience controlling powered aircraft of any category.

The participating pilots (n = 18) held a private pilot license (PPL, all male, mean age = 43.5, *SD* = 16.1). Flight experience was on average 324 flight hours, *SD* = 245 hours, and 9.8 years of being active as a private pilot, *SD* = 7.4.

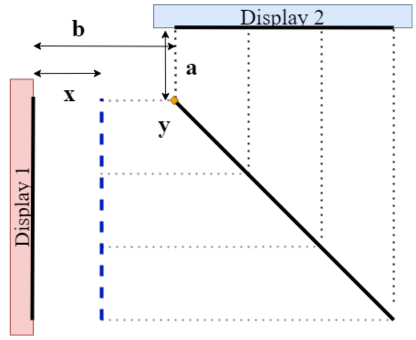

Figure 1. Left: the experimental setup with the display and joystick. Right: The configuration of the MLD with two displays and mirror (the black diagonal line).

All participants were right-handed and had normal or corrected-to-normal eyesight. This experiment was approved by the ethical committee of the Delft University of Technology, and all participants provided informed consent.

Apparatus

A desktop setup was used with no outside visuals (see Figure 1, left). A half-silvered mirror-type display was used for the MLD. A schematic of the inner workings of this display can be seen in Figure 1 (right). For each display layer, a standard LCD-type display was used with a resolution of 1152 x 864. These displays were placed perpendicular to each other, and a half-silvered mirror was placed at a 45°between the displays. This creates a virtual display (dashed blue line in Figure 2) at some distance x in front of display 1. Distance x could be set by moving display 2 to increase distance a. For the low depth condition, X was 1.6 cm, and for the high depth condition 2.1 cm. For the single-layer display (SDL; baseline) condition, display 2 was turned off and all information was displayed on display 1. A standard Logitech Extreme 3D pro joystick was used as input device. The maximum angular deflection on the roll axis of this joystick was measured to be 20°.

A Boeing 747-based PFD was used (see Figure 2). In the MLD, the aircraft symbol, sky pointer and bank angle scale are presented on the upper layer, as well as the speed and altitude tapes. The horizon and pitch ladder were presented on the lower layer. A simplified aerodynamic model was used. The model had a fixed speed of 120 knots, altitude always indicated 10,000 ft., and the attitude was controllable in the pitch and roll axis with sensitivity and dampening resembling that of a small single-engine piston aircraft.

Procedure and tasks

After a briefing and an intake questionnaire, participants were first showed the SLD and MLD version of the PFD. Non-pilots were explained the symbols and how the display is used to control an airplane. Participants then familiarized themselves with the flight dynamics for 5 minutes by flying several turns and level changes. The familiarization was followed by 10 practice runs of the experimental task with the SLD and 10 with the MLD. Each run started with a black screen displayed for 5 seconds. Then, the display showed a PFD

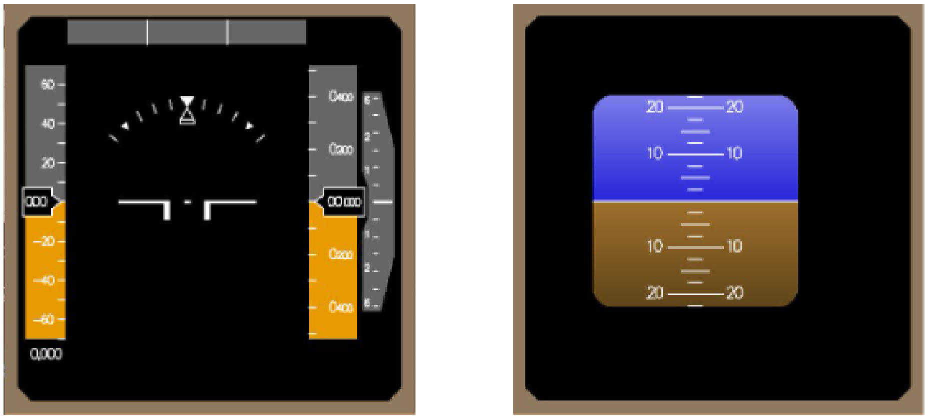

Figure 2. Left: The image presented on the upper layer of the MLD. Right: the image presented on the lower layer of the MLD.

indicating a bank angle at 30°to the left or to the right. When the display appeared, the participant was to respond immediately by rolling the wings level using the joystick. Ten seconds after the appearance of the PFD, the run ended. This was immediately followed by the black screen at the start of the next run. The participant's hand was always placed on the joystick. Following practice, this task was performed in two blocks of 30 runs with the SLD and with the MLD, with a short break in between. The order of conditions was counterbalanced between participants.

Dependent measures

Data on the roll angle and roll control input were logged and analyzed with MATLAB. From this, the following dependent measures were obtained:

- Error rate. A roll reversal error is recorded when the first roll control input was towards the same side of the displayed bank angle. The threshold of input detection was set at 1.5°stick deflection. The error rate is the percentage of runs with an error over the total number of runs.
- Reaction time. This is the time between PFD presentation and the start of the first roll control input.

Lower error rates and faster reaction times were expected with the MLD compared to the baseline condition, as the MLD was expected to facilitate quicker and more accurate recognition of the figure-ground relationship.

Statistics

In both groups, the effect of MLD was tested using a paired-samples *t*-test between MLD and baseline. If not normally distributed, the Wilcoxon signed rank test was used.

In non-pilots, the effect of the MLD depth between layers was additionally tested using a mixed-model ANOVA with display type (baseline, MLD) as the within-subject factor, and MLD depth (low depth group, high depth group) as the between-subject factor.

Results

Data collection

A number of runs (1% of total) were excluded from analysis due to incorrect detection of an input immediately following PFD presentation due to the stick position not being centered within the limits of 1.5°stick deflection.

Non-pilots

The Wilcoxon signed rank test showed that there was no significant difference in error rates between the baseline (median = 5% errors) and MLD condition (median = 3.3%, $Z(22)$ = -1.33, $p = 0.182$). The mixed-model ANOVA showed no significant difference in improvement between the two MLD depths used ($F(1{,}21) = 1.059$, $p = 0.315$). Nevertheless, as the low depth showed the largest improvement, we decided to use this depth for the pilot group. The paired-samples t-test indicated a significant difference in reaction time between the baseline condition (mean = 591 ms) and the MLD condition (mean = 605 ms, $t(22)$ = -2.608, $p = 0.016$), a difference which was opposite to the expected direction.

Pilots

A Wilcoxon Signed Rank test showed that there was no significant difference in error rates between the baseline (median = 3.3% errors) and MLD condition (median = 3.3%, $Z(18)$ = -0.475, $p = 0.635$). A Wilcoxon Signed Rank test t-test indicated no significant difference in reaction time between the baseline condition (median = 625 ms) and the MLD condition (median = 644 ms, $Z(18)$ = -1.24, $p = 0.215$).

Discussion

The results did not indicate a significant improvement in performance when using the MLD compared with the SLD. In contrast, reaction times in non-pilots were longer when using the MLD than the SLD, suggesting that it was more difficult to read the bank angle quickly with the MLD. Participants reported that they indeed perceived depth in the MLD, although this depth perception was mostly present at the start of the experiment or any time they moved their head. The requirement of head motions indicates that the MLD was unsuccessful in presenting stereoscopic cues, as no head motions should be required for such cues. This also makes it impractical for use in the cockpit.

The required distance between layers in the MLD to obtain a stereoscopic effect would cause the pitch and roll indications to become inaccurate, as the position of the symbols relative to each other would then shift greatly depending on head position. Several participants mentioned that they thought the aircraft model in the MLD condition was slower to react than that in the SLD condition, which was not the case. We do not know what may have caused this perception. The results of the experiment lead us to conclude that stereoscopic depth cues achieved thought MLDs are not suitable for enhancing the figure-ground representation in the attitude indicator.

Subsequent research into optimizing the PFD for attitude representation could focus instead on monoscopic cueing. There are several types of monoscopic cues that could be implemented without making serious changes to the PFD designs currently in use in commercial aviation. Examples of this are the use of a horizontal color gradient to simulate "aerial perspective" (Gibson, 1950), linear perspective lines, ground texture, extending the horizon behind the speed and altitude tapes, or adding a dark line under the aircraft symbol to simulate shadow. With each of these additions, it is important to ensure that the salience of the horizon, aircraft, and pitch and roll ladders remains intact. Additions to the sky and ground should not contrast too much with the colors of these surfaces, and thickness of added lines should be minimal. Empirical studies are needed to evaluate the effectiveness of these design changes, and possibly to fine-tune the optimal use of added symbols.

References

Aircraft Accident Investigation Bureau. (2002). *Final Report of the Aircraft Accident Investigation Bureau on the Accident to the Saab 340B Aircraft, Registration HB-AKK of Crossair Flight CRX 498 on 10 January 2000 Near Nassenwil/ZH.* https://www.sust.admin.ch/inhalte/AV-berichte/1781_e.pdf.

Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile. (2009). *Final report on the accident on 1st June 2009 to the Airbus A330-203, registered F-GZCP, operated by Air France, Flight AF 447 Rio de Janeiro–Paris.*

Cameroon Civil Aviation Authority. (2010). *Technical Investigation into the Accident of the B737-800 Registration 5y-Kya Operated by Kenya Airways that Occurred on the 5th of May 2007 in Douala.* https://reports.aviation-safety.net/2007/ 20070505-0_B738_5Y-KYA.pdf.

Dünser, A., Billinghurst, M., & Mancero, G. (2008). Evaluating visual search performance with a multi layer display. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat* (pp. 307–310).

Gibson, J. J. (1950). *The Perception of the Visual World.* Boston (MA): Houghton Mifflin.

Grether, W. F. (1947). Discussion of pictorial versus symbolic aircraft instrument displays. *USAAF AMC, Engineering Division, Aero Med Lab., TS EAA.*

Hayes, J., Moore, A., & Wong, B. L. (2006). Information layering to de-clutter displays for emergency ambulance dispatch. In *Proc. 13th European conference on Cognitive ergonomics: trust and control in complex socio-technical systems* (pp. 10–16).

Johnson, S. L., & Roscoe, S. N. (1972). What Moves, the Airplane or the World? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *14*(2), 107-129.

Müller, S., Sadovitch, V., & Manzey, D. (2018). Attitude indicator design in primary flight display: Revisiting an old issue with current technology. *The International Journal of Aerospace Psychology*, *28*(1-2), 46–61.

Newman, R. L., & Rupert, A. H. (2020). The magnitude of the spatial disorientation problem in transport airplanes. *Aerosp. medicine and hum. perf.*, *91*(2), 65–70.

Roscoe, S. N., Corl, L., & Jensen, R. S. (1981). Flight Display Dynamics Revisited. *Human Factors*, *23*(3), 341–353.

Van den Hoed, A., Landman, A., Van Baelen, D., Stroosma, O., van Paassen, M. M., Groen, E. L., & Mulder, M. (2022). Leans illusion in hexapod simulator facilitates erroneous responses to artificial horizon in airline pilots. *Human Factors*, *64*(6), 962–972.

Wong, B. W., Joyekurun, R., Nees, A., Amaldi, P., & Villanueva, R. (2005). Information layering, depth and transparency effects on multi-layered displays for command and control. In *Proc. Human Factors and Ergonomics Society* (Vol. 49, pp. 352–356).