Simon Müller, Vitalij Sadovitch, Dietrich Manzey

Attitude Indicator Design in Primary Flight Display: Revisiting an Old Issue With Current Technology
Subtitle


This is an Accepted Manuscript of an article published by Taylor & Francis in The International Journal of Aerospace Psychology on 06 July 2018, available online: http://www.tandfonline.com/10.1080/24721840.2018.1486714.

Terms of Use
Copyright applies. A non-exclusive, non-transferable and limited right to use is granted. This document is intended solely for personal, non-commercial use.
Attitude Indicator Design in Primary Flight Display: Revisiting an Old Issue With Current Technology

Simon Müller, Vitalij Sadovitch, and Dietrich Manzey

Work, Engineering, & Organizational Psychology, Department of Psychology and Ergonomics, Technische Universität Berlin, Berlin, Germany

Corresponding author: Simon Müller, simon.mueller@tu-berlin.de, Department of Psychology and Ergonomics, Technische Universität Berlin, Marchstr. 12, 10587 Berlin, Germany
Abstract

**Objective:** The experiments investigated the “old issue” of the attitude indicator’s moving-horizon versus moving-aircraft format with current primary flight display technology. Of interest was whether the effects found in earlier studies, favoring the moving-aircraft format, could be replicated with most recent technology including extended horizon displays, which depict the artificial horizon extended over the whole screen with overlaying speed and altitude scales (e.g., B787).

**Background:** Although the moving-horizon format represents the standard approach in Western aviation, human factors research from the 1950s to the 1970s with round electromechanical instruments favored the moving-aircraft format with respect to better support of flight-path tracking and unusual attitude recoveries. However, recent studies using laboratory displays more similar to modern primary flight displays provided inconsistent results. This led to the assumption that the display’s design is a moderating factor of those effects.

**Method:** Thirty-two novices and 13 pilots flew several tracking and recovery tasks in a PC-based simulator equipped with moving-horizon and moving-aircraft formats in classic and extended horizon design.

**Results:** The data show that the previous effects favoring a moving-aircraft format of displaying bank information can be replicated with current primary flight display designs. However, the extended horizon design seems to reduce this effect, at least for pilots.

**Conclusion:** The results suggest reconsidering the format of the attitude indicator at least for new applications, such as control of remotely piloted aircraft.
Keywords: display-control or stimulus-response compatibility, display design principles, spatial disorientation, attitude indicator
Attitude Indicator Design in Primary Flight Display: Revisiting an Old Issue With Current Technology

Flying an aircraft in instrument meteorological conditions (IMC), for example, clouds or night skies, precludes the direct reference to the outside view, possibly contributing to an unrecognized spatial disorientation. Spatial disorientation can be defined as an “erroneous sense of one’s position and motion relative to the plane of the earth’s surface” (Gillingham & Previc, 1993, p. 77) and has been a constant contributing factor to a number of fatal aviation accidents (Comstock, Jones, & Pope, 2003; Gibb, Ercoline, & Scharff, 2011; Poisson & Miller, 2014; Roscoe, 2004). Especially untrained and beginner pilots who are not familiar with flying under IMC tend to experience difficulties maintaining proper spatial orientation, when unsuspectedly losing the natural horizon as visual reference (Roscoe, 2004). In IMC, pilots depend on the attitude indicator (AI) to assess the orientation of their aircraft. The AI is one among other instruments that offer ownship orientation information. It provides information on the aircraft’s pitch and bank angles in relation to the natural horizon and represents the central element of the primary flight display (PFD) in modern aircraft.

In aviation history, two alternative design options have primarily been used to present attitude information: the moving-horizon (MH) format and the moving-aircraft (MA) format.¹ The MH format was introduced in 1929 and has been the standard

¹ There are several other concepts proposed for displaying the flight attitude, such as frequency-separated display (Beringer, Williges, & Roscoe, 1975; Roscoe, 1968), kinalog display (Fogel, 1959), or Arc-Segmented Attitude Reference display (Self, Breun, Feldt, Perry, & Ercoline, 2002). These concepts
AI format in Western aviation ever since (Previc & Ercoline, 1999). It shows a fixed airplane symbol in the center of the display as a stable element, while the artificial horizon is moving according to the outside view (the natural horizon). That is, banking of the aircraft to the right is indicated by rotating the artificial horizon to the left and vice versa. Pitching of the aircraft is indicated by upward or downward movements of the horizon line. This attitude display format is based on the so-called principle of pictorial realism (Roscoe, 1968), because it indicates changes of bank angles by movements of an artificial horizon, as if looking through a porthole in front of the aircraft to the outside or drawing the aircraft symbol on the windscreen and viewing it against the natural horizon. It can be considered as an abstract version of a so-called contact analog display, which provides visual cues conformal to the “same laws of motion perspective as their visual-world counterparts” (Roscoe & Eisele, 1976, p. 44).

The MA format has been used for a long time in Soviet and later Russian aviation (Previc & Ercoline, 1999). It also shows an aircraft symbol in the center and an artificial horizon line. Congruent with the MH format, the pitch angle of the aircraft is indicated by an upward or downward shift of the artificial horizon line. However, contrary to the MH format, the bank angle is indicated by rotating the aircraft symbol while keeping the artificial horizon in a steady horizontal position in reference to the instrument panel. That is, a bank movement of the aircraft to the right or left is indicated by a rotation of the aircraft symbol in the AI to the same direction. Thereby, the MA format fulfills what has been referred to as the principle of moving part (Roscoe, 1968), that is, “the moving have been widely discussed, but they still lack broad adoption in civil aviation. This article is therefore limited to contrasting the two standard formats of Western and Russian aviation.
element on a display should correspond with the element that moves in the pilot’s ‘men-
tal model,’ . . . and should move in the same direction as that mental representation” (Wickens, 2003, p. 152).

The general question of which AI format would be better suited to display the air-
craft’s attitude on head-down instruments in terms of intuitive understanding and com-
patibility has been addressed by many studies since 1945. Most of the early studies (1950s–1970s) comparing the effectiveness of the different AI formats, have investigated the performance of flight novices (i.e., nonpilots without prior knowledge of flying an aircraft) in recovery tasks. The recovery task simulates a flight situation where a pilot is surprised by a possibly dangerous change of aircraft attitude, for example, an unusually high bank angle. To recover to a horizontal attitude, it is necessary to establish quickly a proper spatial orientation and to initiate a rapid compensatory bank movement (e.g., Roscoe & Williges, 1975). What this research usually found was a clear advantage of the MA over the MH format. Especially, flight novices committed significantly fewer reversal errors—that is, initial movement away from the nearest horizon—when flying with the MA compared to the MH format (cf. reviews by Johnson & Roscoe, 1972; Previc & Ercoline, 1999). In contrast, results of studies with experienced pilots were less con-
sistent. For example, Browne (1954), Gardner, Lacey, and Seeger (1954), and Hasbrook and Rasmussen (1973) did not find significant differences between MH and MA format for pilots. Whereas in the study of Beringer, Williges, and Roscoe (1975) pilots per-
fomed better with the MH format, in the study of Dunlap and Associates (as cited in Previc & Ercoline, 1999) the pilots performed better with the MA format. However, the studies suggested at least that changing from the familiar MH format to the MA format would not lead to significant performance decrements. Altogether, these results have
been taken as evidence that the MA format of the AI is more intuitive to understand than the MH format, which directly contrasts to the current standard in most aircraft today (Previc & Ercoline, 1999).

Several theoretical explanations have been raised to explain the putative superiority of the MA format for maintaining spatial orientation. The two most common ones attribute it to effects of display-control compatibility and figure–ground relation.

Regarding display-control compatibility, two related aspects can be distinguished. The first one involves what has been referred to as response–effect compatibility (Janczyk, Pfister, Crognale, & Kunde, 2012). It concerns the compatibility of the relationship between the direction of the movement at the controls and the anticipated effect in terms of a change indicated in the display. The relationship is compatible when the movement direction within the display directly corresponds to the movement direction of the control input device that causes the movement. This sort of compatibility is fulfilled with the MA but violated with the MH format. With the MA format, a leftward (or counterclockwise) control input causes a corresponding counterclockwise rotation of the moving element in the display. With the MH format, this relationship is reversed. Based on the ideomotor theory (Greenwald, 1970), it can be expected that pilots’ control movements can be selected faster and more reliably with the MA than the MH display. Direct support for this assumption has been provided by a study by Janczyk, Yamaguchi, Proctor, and Pfister (2015). In this study, novices were required to bank their simulated aircraft from a horizontal starting position either to the left or to the right with both AI formats. Responses were quicker and more correct when conducting the task with the MA display. Further, more indirect support can be derived from the positive results of studies with the frequency-separated displays, which have been proposed to
provide display-control-compatible initial indications of flight attitude changes for the conventional MH format (Beringer et al., 1975; Roscoe & Williges, 1975).

The second aspect concerns to what extent the stimulus–response mapping—that is, the relationship between observed changes in the display and required responses at the controls—fits to the mental representation of the task. In case of recovery tasks, the goal of the control movement is to compensate a given attitude deflection indicated by a change in the display. That is, the expected relationship between the direction of the change (deflection) seen in the display and the needed direction of movements at the controls required to compensate for it, is reversed. This, again, is fulfilled by the MA format where movements in the display to the left or right require compensatory movements at the controls to the right or left, but violated with the MH display where movements in the display to the left or right have to be compensated by movements in the same direction.

A second effect that might contribute to the superiority of the MA format is proposed to be a figure–ground reversal issue (Johnson & Roscoe, 1972). Typically, an object will be perceived as a figure, when it is moving in front of a stable background or ground. Yet, if most of the visual field is moving uniformly, it also can be perceived as a stationary background, while the observer is moving (Fitts & Jones, 1947). The latter is exactly what happens when pilots look out of the cockpit windscreen when flying a turn. In this case, they see the natural horizon moving but immediately interpret it as a movement of their aircraft. However, this is different when considering the movement of the artificial horizon line in an MH formatted AI. Although, representing what would be seen if one looks outside through a small porthole, the horizon line does not fulfill the typical characteristics of a (back)ground. First, it is not presented far behind the aircraft.
symbol. Second, it represents a comparatively small moving element included in a larger (and stable) instrument panel. This can easily lead to a figure–ground reversal, where the horizon line is perceived as the figure and the instrument panel as the ground. Johnson and Roscoe (1972) suggested that flight novices, but also experienced pilots, could easily misinterpret the horizon as the moving part that is being manipulated by their control input, which then leads to exact reversed control responses compared to what is required.

Previc and Ercole (1999) added a neuropsychological explanation for the figure–ground reversal effect based on the assumption that objects in close proximity to the pilot, like cockpit instruments, are perceived and processed differently than information that is farther away, for example, the natural horizon. Specifically, they assumed four major brain systems are involved when interacting with the external three-dimensional world (Previc, 1998). The first and closest system is called the *peripersonal* system. It is involved in manipulating and interacting with objects near our bodies. Accordingly, movements of objects in this space are usually perceived as what they are (movements of objects) but not consequences of a self-motion. The other extreme is the *ambient extrapersonal* system. It processes information in far distances of the field of view and is mainly involved in monitoring, controlling, and stabilizing one’s position in reference to Earth. Perceived large-scale movements in this domain are usually interpreted as consequences of self-motion. Based on this framework, the main problem of the MH format and the basis for the figure–ground reversal is that movements of the natural horizon, which usually are perceived as large-scale changes in the far domain and processed by the ambient extrapersonal system, are visualized by a small instrument (i.e., the AI) positioned in the peripersonal space. Consequently, it can be expected that the
artificial horizon in the MH format intuitively is perceived as the controllable element, instead of a consequence of self-motion as implicitly assumed by the principle of pictorial realism (Previc & Ercoline, 1999).

However, the results of more recent studies are less consistent and could not always replicate the advantages of the MA compared to the MH format (Gross & Manzey, 2014; Yamaguchi & Proctor, 2006, 2010). This might be caused by two factors. Firstly, some of the latter studies have used only continuous tracking tasks instead of discrete recovery tasks (Yamaguchi & Proctor, 2010). Tracking tasks simulate an attitude-holding task with atmospheric disturbances; that is, participants need to compensate for disturbances in pitch and bank to maintain or regain a stable horizontal flight over a distinct period of time (e.g., Cohen, Otakeno, Previc, & Ercoline, 2001; Yamaguchi & Proctor, 2010). In contrast to recovery tasks, which request quick discrete movements in response to a sudden change in the AI, tracking movements represent continuous movements controlled by continuous visual feedback. This feedback might make it easier to adapt to the different formatted AIs without any visible performance differences. Second, most of the early evidence revealing advantages of the MA stemmed from studies using small round electromechanical instruments, as they were common in cockpits at that time. In the more recent studies, usually considerably larger, computer-generated, rather abstract laboratory AI displays expanding over the whole screen were deployed, which often did not correspond to any real cockpit display (Gross & Manzey, 2014; Yamaguchi & Proctor, 2006, 2010).

This suggests that general aspects of AI format like display size or presentation on monitors might also make a difference with respect to the MA versus MH issues.
However, just a bigger display does not seem to better support the pilot’s spatial orientation when using an MH-formatted AI (Ding & Proctor, 2017; Previc & Ercoline, 1999). More likely, the specific display design itself might be the key factor here. Especially, most recent AI designs, as used in the B787, have the potential to be superior to small electromechanical indicators or the current AIs integrated into the standard glass-cockpit PFDs (e.g., A320). These new designs, which we refer to as extended horizon designs, comprise an artificial horizon that is extended over the whole screen behind speed and altitude scales. This might better support an interpretation of the artificial horizon as (back)ground, and, thus, reduce the figure–ground reversal issue. This is also suggested by some early approaches of AI enhancements that extended the artificial horizon even beyond the actual display or instrument and successfully improved the spatial awareness of pilots (e.g., Liggett, Reising, & Hartsock, 2009; Malcolm, 1983).

Based on these considerations, this research addresses to what extent the issue of MA versus MH format persists with the typical head-down PFDs usually found in current commercial aircraft (A320) or more recent versions (B787), how the formats affect the performance in different flight tasks (flight-path tracking, attitude recovery), and what difference expertise (novices vs. pilots) makes. Two experimental studies are reported. The first one included flight novices, whereas the second one included experienced pilots. For novices, it was hypothesized that the earlier results of a better flight performance with the MA format compared to the MH format can be replicated with the classic PFD design, at least for the recovery task. However, we also assumed that the putative superiority of the MA format would be reduced with the extended horizon design. For pilots, predictions were more difficult. The fact that the pilots have extended train-
ing and experience with their MH format must be considered and might affect performance in favor of the MH display. Yet, it has been shown that a significant portion of MH trained pilots (about 33%) still have a mental model of attitude changes that conforms to the MA format (Kovalenko, 1991). This would suggest that even the effects for MH trained pilots might be similar to what is expected for novices, primarily for tasks that are not frequently trained on the job and thus require some spontaneous and intuitive behavior (e.g., recoveries from unusual attitudes). In the following, first the general method of both experiments is described. Subsequently, for each experiment, the specific methods and results are presented and discussed. The article closes with a summarizing discussion of both studies in context and some conclusions for applications and further research.

**General Method**

**Apparatus**

The experiments were conducted in a PC-based flight simulator. It consisted of a cockpit panel mock-up (Cessna 172 Skyhawk SP G1000) with an integrated screen displaying a PFD and an outside-view projection on the wall approximately 1.2 m in front of the mock-up cockpit. The PFD design corresponded in almost all aspects to the PFD currently used in the A320. Some adaptations were made with respect to the implementation of the two AI formats (MA and MH) and both design types; that is, classic and extended horizon (see Figure 1). All PFDs were 12.6 cm high and 19.9 cm wide. The AI of the classic PFDs was 8.3 cm high and 7.1 cm wide. The participants were placed in usual seating distance to the PFD (approximately 60 cm). The input device consisted of a commercially available Logitech Extreme 3D Pro joystick. The simulation was a reduced lin-
ear flight model with two degrees of freedom, one each in pitch and bank. The input deflec-
tions of the joystick were linearly transferred into pitch and bank rates. There was no
need for thrust control. The outside view was generated using the X-Plane 10 flight sim-
ulation.

Figure 1. All primary flight display (PFD) configurations used in the experiments. The
left side shows the moving-horizon (MH) format and the right side the moving-aircraft
(MA) format, whereas the upper PFDs are in classic horizon design and the bottom
PFDs are in extended horizon design. All PFDs show a bank angle of 45° to the right and
pitch up of 10°. Note. AI = attitude indicator.
Tasks

**Tracking.** The participants had to maintain a stable horizontal flight with pitch and bank angle of 0°, thereby compensating for preprogrammed disturbances by proper corrections on x- and y-axes of the joystick. The disturbances were simulated by two separate disturbance functions for each axis, based on the sum of five sine functions with input frequencies of 0.1705 Hz, 0.2885 Hz, 0.4918 Hz, 0.8333 Hz, and 1.4286 Hz vertically, as well as 0.1304 Hz, 0.2222 Hz, 0.3750 Hz, 0.6383 Hz, and 1.1111 Hz horizontally (cf. Fracker & Wickens, 1989). The amplitude in pitch direction was reduced to a third of the amplitude of the bank function.

**Recovery.** The participants had to perform unusual-attitude recoveries to maintain a horizontal flight attitude. The unusual-attitude stimuli included a sudden discrete skip of the AI in one frame to the left or right, indicating the change of bank angle of the aircraft by 45°, 90°, and 135°. The pitch angle initially stayed at 0°, but could be altered by the participants during recovery. Participants were instructed to recover to a stable horizontal attitude as quickly as possible.

Design

The experiment included three factors. The first factor comprised the two PFD design approaches (classic vs. extended horizon). The second factor represented the two formats of the attitude reference, MA and MH (see Figure 1). The third factor, only used for the investigation of recovery-task performance, was the bank angle of unusual attitudes (45°, 90°, and 135°).

Dependent Measures

Performance in the tracking task was assessed by means of deflections in relation to 0° for both axes, bank and pitch, separately recorded with a frequency of 60 Hz.
Based on these data, the root mean square error (RMSE) was calculated across trials to assess the tracking error for bank and pitch movements. Note that the tracking error for pitch movement was just calculated as a control variable. No effects were expected for this measure because the AI formats did not differ in depicting pitch movements. Performance in the recovery task was assessed by two measures. The first one included the percentage of reversal errors, counted whenever the initial joystick input to an unusual attitude change was initiated to the wrong direction; that is, an initial input that amplifies instead of compensates for a given attitude change. The second measure included the response time needed to respond to a given attitude change. It was defined as the time between the occurrence of the unusual attitude stimulus and the first input detected at the joystick. Only correct trials without reversal errors were considered for this measure. In addition, for both tasks, the subjectively perceived workload was assessed by means of the unweighted mean score of the NASA–TLX (Hart, 2006; Hart & Staveland, 1988).

Data Analysis

Prior to data analyses, outlier corrections were made. For analyses of tracking performance, participants’ data were excluded if their bank RMSE exceeded 3 standard deviation (SDs) from the mean of the respective condition. Regarding the recovery task, only successfully completed recovery trials were considered in the analysis. A recovery was defined as successful if the bank and pitch angles of the aircraft were restabilized within 10 s and remained stable for at least 2 s within a range of ±2°. Furthermore, trials for which response time was shorter than 100 ms were excluded from both measures of the recovery task. Data for participants who could not successfully finish more than 25% of the recovery trials in one of the conditions were entirely removed from the analysis.
Analyses of variance (ANOVAs) with repeated measures were used to analyze the dependent measures for tracking and recovery tasks. Percentage data of reversal errors were arcsine transformed to achieve better distribution characteristics (Sokal & Rohlf, 1981). We report the back-converted descriptive statistics for reversal errors in percent to facilitate interpretation. An alpha level of 5% was defined for considering effects as significant. In case of violations of the sphericity assumption (Mauchly, 1940), degrees of freedom of the $F$ test were corrected according to the Huynh–Feldt procedure (Huynh & Feldt, 1976).

**Experiment 1**

**Method**

**Participants.** A total of 36 participants (16 female, 20 male) took part in the study. None of them had any prior knowledge of flying a real aircraft whatsoever. Eighteen participants had some limited experiences based on casually flying in flight simulators of different fidelity (including PC-based games). They were randomly assigned to two groups constrained by an equal distribution of gender. The first group performed all tasks with the classic PFD design as used, for example, in the A320. The mean age of the participants of this group was 27.0 years ($SD = 4.3$). The second group performed the tasks with the extended horizon PFD. Their mean age was 25.7 years ($SD = 3.8$). For participation they received a compensation of 10€ or course credits.

**Design.** The experiment included a 2 (horizon design) × 2 (AI format) mixed-factor design. The first factor representing the classic versus extended horizon design was defined as a between-groups factor. The second factor representing the two AI for-
mats was defined as a within-subject factor. For investigating recovery task performance, a third factor was added, defined as within-subject factor representing the different bank angles of unusual attitudes (45°, 90°, and 135°) used in the recovery task.

**Procedure.** Prior to the data collection, every participant read a brief standardized introduction including information on the test procedure and the tasks. This was followed by a 4-min accommodation phase to familiarize the participants with the simulation, aircraft controls, and flight displays. During this phase, only the outside view was displayed as reference to control the aircraft, and participants were requested to make several flight maneuvers, including different turns and level flights.

This accommodation phase was followed by two experimental blocks corresponding to the two AI format conditions. Each block started with a familiarization phase of the respective AI format. This phase first included flying with both outside view as well as PFD. Yet, after a couple of minutes, the outside view was removed and only the PFD remained as a reference to control the attitude of the simulated aircraft. The outside-view projection was only enabled at the beginning of each AI training, not during the experiment. To ensure that all participants gained a similar knowledge about the aircraft’s reaction to the control inputs, all participants needed to complete several defined flight tasks and a free flight phase. Each familiarization phase lasted 4 min. Besides the PFD, there were no other displays active during the following experimental tasks. First, the participants performed the tracking task for 2 min. Subsequently, they had to provide the NASA–TLX ratings for this task. Then 24 trials (3 bank angles × 2 directions × 4 replications) of the recovery task were performed with a random time interval of 5 to 20 s between two successive trials. After 12 trials, a short break was taken where the partici-
pants provided initial NASA–TLX ratings for this task. After the second set of 12 recovery trials, another sampling of NASA–TLX ratings followed. Performing the 24 recovery trials with a given AI format lasted about 10 min. The order of experimental blocks, corresponding to the two AI format conditions, was counterbalanced across participants. Overall, each experimental session lasted about 1.5 hr.

Results

Tracking task. No participant was considered an outlier in the tracking task. Thus, all 36 participants were included in the following analysis. The participants of both groups were significantly better in maintaining a stable bank attitude with the MA format \((M = 4.63^\circ, SE = 0.37^\circ)\) than the MH format \((M = 5.53^\circ, SE = 0.65^\circ)\), in terms of the RMSE of bank angle, \(F(1, 34) = 6.67, p = .014, \eta_p^2 = .16\). However, the ANOVA revealed neither a significant main effect of the horizon design, \(F(1, 34) = 2.23, p = .144, \eta_p^2 = .06\), nor a significant interaction effect of Horizon Design × AI Format, \(F(1, 34) = 2.21, p = .146, \eta_p^2 = .06\). As expected, no significant effects emerged, when considering the RMSE of pitch, all \(F < 2.5, p > .12, \eta_p^2 \leq .07\).

The participants rated their perceived workload in the NASA–TLX significantly lower in condition MA \((M = 37.6, SE = 2.8)\) compared to MH \((M = 42.3, SE = 3.2)\), \(F(1, 34) = 4.99, p = .032, \eta_p^2 = .13\). In addition, they rated their workload somewhat lower when flying with the extended horizon \((M = 35.0, SE = 4.0)\) than when flying with the classic PFD design \((M = 45.0, SE = 4.0)\). However, this latter effect just failed to reach statistical significance, \(F(1, 34) = 3.11, p = .087, \eta_p^2 = .08\). No significant interaction effect for Horizon Design × AI Format was found, \(F(1, 34) = 0.11, p = .744\),
\[ \eta_{p}^2 < .01. \] Inspection of the different NASA–TLX subscales revealed that the mental demand and effort contributed subscales most to these results.

**Recovery task.** Data of 3 participants were excluded from analysis in the recovery task. All three outliers were found in the group flying with the classic MH display. Therefore, the sample size was reduced to 33 participants. For these participants, 3.7% of all individual trials were discarded due to unsuccessfully finished recoveries or response time constraints.

**Reversal error.** The mean percentage of reversal errors committed by the participants of both groups for both AI formats and the three bank angles are shown in Figure 2A. The 2 (horizon design) \( \times \) 2 (AI format) \( \times \) 3 (bank angle) mixed-factor ANOVA revealed that the participants of both groups were significantly better able to avoid this sort of error with the MA format \( (M = 4.3\%, \ SE = 0.9\%) \) than with the MH format \( (M = 13.1\%, \ SE = 2.6\%) \), \( F(1, 31) = 11.90, p = .002, \eta_{p}^2 = .28 \). In addition, the main effect of bank angle became significant, \( F(1.92, 59.66) = 5.60, p = .006, \eta_{p}^2 = .15 \). As becomes evident from Figure 2A, the participants made fewer reversal errors the smaller the bank angle of the stimulus was. Although it seems that this effect was stronger for the MA than the MH format, the AI Format \( \times \) Bank Angle interaction just failed to become significant, \( F(2, 62) = 2.78, p = .070, \eta_{p}^2 = .08 \). No other effect became significant, all \( F < 1.6, p > .22, \eta_{p}^2 \leq .05 \).

**Response time.** The mean time needed to respond to a given attitude change did not differ significantly over all conditions, which can be seen in Figure 2B. Neither any main effect nor any interaction effect became significant, all \( F < 1.9, p > .16, \eta_{p}^2 \leq .06 \).
Figure 2. Novices’ means of (A) reversal error and (B) response time over both horizon design groups for both attitude indicator (AI) format conditions, moving horizon (MH) and moving aircraft (MA), and for each bank angle condition, 45°, 90°, and 135°. Error bars represent standard errors.

**Workload.** Generally, the group performing the recovery tasks with the extended horizon design rated their perceived workload on the NASA–TLX lower \((M = 23.9, SE = 2.7)\) than the classic horizon group \((M = 34.1, SE = 3.0)\), \(F(1, 31) = 6.42, p = .017, \eta_p^2 = .17\). In addition, the mean subjective workload was also lower with the MA format \((M = 26.4, SE = 1.7)\) compared to the MH format \((M = 31.6, SE = 2.7)\), \(F(1, 31) = 7.09, p = .012, \eta_p^2 = .19\). However, the Horizon Design × AI Format interaction effect did not become significant, \(F(1, 31) = 0.12, p = .736, \eta_p^2 < .01\). Considering the subscales of the NASA–TLX, the horizon design effect was primarily observable in mental demand, temporal demand, effort, and frustration, whereas the format effect was observable in physical demand, temporal demand, and performance.
Discussion

This research provides evidence that the superiority of the MA versus MH format of AIs persists also with the AIs integrated in the typical PFDs of glass cockpits in modern commercial aircraft. This holds true for both flight-path tracking as well as quick recoveries from unusual flight attitudes. Contrary to our expectations, essentially the same pattern of effects was found for the classic PFDs and the new generation of PFDs with extended horizon designs.

Let us first consider the results for flight-path tracking and recoveries with the classic PFD design. When required to maintain a horizontal flight attitude towards external disturbances (flight-path tracking), flight novices were better able to correct continuously for bank-angle deflections with the MA than the MH format. Likewise, the participants reported less workload involved in flight-path tracking with the MA compared to the MH format. As expected, no such differences were found for corrections of pitch deflection, which were depicted in the same way with both formats. Thus, the results suggest that not only the effectiveness but also the efficiency of flight-path tracking benefits from depicting bank deflections in terms of a moving airplane compared to a moving artificial horizon line. These findings support the results of Cohen et al. (2001), but are in contrast with recent results reported from a study by Yamaguchi and Proctor (2010), who did not find such an effect with comparable groups of flight novices (i.e., undergraduate students). The reasons for this partial inconsistency are difficult to assess. One possible reason might be related to the sort of external disturbance functions used to produce random deflections of the indicated bank and pitch angle from a horizontal flight. Perhaps only relatively large deflections as used in this study are sufficient to produce the performance difference between the two AI formats. Unfortunately, no
detailed descriptions of the disturbance functions as used by Cohen et al. (2001) and Yamaguchi and Proctor (2010) are available. Thus, no decisive conclusion can be reached in this respect, and certainly, other factors, such as different displays, other control input devices, or different instructions might have contributed to finding different effects.

More clearly and in line with the majority of earlier research are the findings with respect to recovery task performance. First, as expected, the rate of reversal errors increased with increasing bank-angle changes for both display formats. However, more importantly and largely independent of the degree of bank-angle changes, the participants committed a higher rate of reversal errors with the MH compared to the MA format. The fact that the response times until the initiation of recovery movements did not differ significantly between conditions eliminates the possibility that the differences in reversal error were only caused by a sort of speed–accuracy trade-off. Rather, it directly supports the hypothesis of a real difference between both AI formats in terms of better support for quick and correct recovery performance by the MA format. This is further mirrored by the NASA–TLX data, which show that the participants perceived lower workload when flying with the MA than the MH format. This pattern of results directly confirms the results of early studies with recovery tasks (Browne, 1954; Gardner et al., 1954). It indicates that the classic design of PFDs in glass cockpits does not appear to change much of pilots’ performance and compatibility issues compared to electromechanic AI instruments used in earlier studies, primarily of the 1950s to 1970s. Obviously, the principle of the moving part still represents the dominant compatibility principle, guiding intuitive and quick responses in current glass cockpits.

Contrary to expectations, the extended horizon design did not change this effect much, either. Actually, the same pattern of findings was observed for both tracking and
recovery when using the extended horizon design. Originally, we had expected that the extended horizon design would ease the appropriate perception of figure and ground relation and, thus, facilitate the correct interpretation of the MH format. Consequently, differences in performances between the MH and MA format were expected to decrease. However, neither the results of the tracking task, nor the results of the recovery task provide evidence for this assumption based on the performance measures. A general benefit of the extended horizon design is only reflected in the assessments of subjective workload after performance of the recovery task. This effect emerged independently of the AI format, though.

Why the expectations concerning the extended horizon format were not supported is not clear at this moment. A possible explanation is that in an extended horizon design the artificial horizon is still displayed in the peripersonal system. Thus, according to the neuropsychological theory of Previc and Ercoline (1999), even with a better figure–ground representation, the artificial horizon of the MH format will not be interpreted as a stable reference system. Additionally, the familiarization phase might not have been sufficient to provide novices with a proper and good understanding of the basic logic of the MH format and its relationship to the natural horizon. Lacking this understanding, any design features making the figure–ground relationship more intuitive might not have been effective for them.

This leads to a general limitation of the first experiment, namely the use of flight novices. It might be questioned to what extent the results favoring the MA over the MH format might be generalized to pilots. Experienced Western pilots trained with the MH display have knowledge that novices do not have and, thus, can be expected to have a better understanding of the AI format reference with respect to the basic principle of
pictorial realism. Earlier findings indeed suggest that the superiority of MA might not emerge with such pilots (Browne, 1954; Gardner et al., 1954; Hasbrook & Rasmussen, 1973). This is expected especially for tasks they perform in daily flying (e.g., tracking). However, results also suggest that pilots would not have much difficulty switching from the MH to the unfamiliar but putatively more intuitive MA format (Previc & Ercoline, 1999). Thus, a second experiment was conducted including pilots as participants.

**Experiment 2**

**Method**

**Participants.** Thirteen certified pilots (2 female) participated in the study. The pilots’ age ranged from 23 to 33 years ($M = 27.4, SD = 2.8$). All pilots had experience with flying according to instrument flight rules (IFR), ranging from 50 to 1,400 flight hours with a mean of 379.8 hr ($SD = 404.1$). Three of them were helicopter pilots. All 13 pilots obtained their IFR training and experience with MH formatted displays or instruments. They volunteered their time to participate in the study.

**Design.** The same factors were used as in the first experiment. However, instead of one between-subject factor both factors were defined as within-subjects factors.

**Procedure.** The experimental procedure corresponded in most aspects to the first experiment. However, the standardized introduction was shortened, and the phases to adapt to the simulator and to familiarize with the different display configurations were condensed to one session displaying both outside view and PFD at the same time. However, during data collection the outside view was removed as in the first experiment.
Each participant performed both tasks with all four possible display conditions. However, the factor horizon design was always sequenced en bloc to resemble a procedure similar to the first study. The resulting possible combinations were counterbalanced over all participants. Between switching the horizon design, the participants had a break of about 15 min.

**Results**

**Tracking task.** One participant was considered an outlier in the tracking task, thus reducing the sample size to 12 participants for this analysis. Mean bank tracking performance was somewhat better when using the MH format ($M = 2.22°$, $SE = 0.08°$) compared to the MA format ($M = 2.64°$, $SE = 0.15°$), $F(1, 11) = 20.32, p < .001, \eta_p^2 = .65$. Yet, no further effect became significant when considering the RMSE of bank, all $F < 1.4, p > .26, \eta_p^2 \leq .11$. A similar effect emerged for the pitch error, which was slightly smaller in the condition with the MH format ($M = 0.87°$, $SE = 0.02°$) compared to the MA format ($M = 0.94°$, $SE = 0.04°$), $F(1, 11) = 5.64, p < .037, \eta_p^2 = .34$. Again, no other effect was significant, as all $F$ values $< 1.0$. No significant effects were found when analyzing the NASA–TLX data, all $F < 3.3, p > .10, \eta_p^2 \leq .23$.

**Recovery task.** No participants’ data were removed from the recovery analysis due to the outlier definition, but 2.1% of all individual trials were disregarded due to unsuccessfully finished recoveries or response time constraints.

**Reversal error.** The percentages of reversal errors when recovering different bank angles with both AI format conditions and both horizon design conditions are presented in Figure 3A. As becomes evident, the pilots performing the recovery task with the classic horizon design committed considerably more reversal errors when using
their familiar MH format \((M = 4.0\%, SE = 1.7\%)\) than when using the MA format \((M = 0.8\%, SE = 0.5\%)\). However, no comparable difference emerged for the extended horizon design \((MH: M = 1.2\%, SE = 0.9\%; MA: M = 1.6\%, SE = 1.3\%)\). In the ANOVA, this was reflected in a significant Horizon Design \(\times\) AI Format interaction effect, 
\[F(1, 12) = 6.48, p = .026, \eta_{p}^2 = .35.\] Neither of the main effects became significant: AI format, \(F(1, 12) = 2.30, p = .155, \eta_{p}^2 = .16\); horizon design, \(F(1, 12) = 0.45, p = .516, \eta_{p}^2 = .04\). These findings suggest that the benefits of the MA display mainly emerged when flying with the classic PFD, compared to the condition with the extended horizon design.

In addition, the main effect of bank angle also became significant, 
\[F(1.43, 17.18) = 8.05, p = .006, \eta_{p}^2 = .40,\] reflecting that the number of reversal errors was, as expected, higher in the \(135^\circ\) condition than the other two conditions (compare Figure 4A). No interaction effect involving the bank angle became significant; all \(F\) values were \(\leq 1.0\).

**Response time.** The mean response times corresponding to the effects found for reversal errors are shown in Figure 3B and Figure 4B. None of the three main effects—that is, AI format, \(F(1, 12) = 0.31, p = 0.588, \eta_{p}^2 = .03\); horizon design, \(F(1, 12) = 0.23, p = .641, \eta_{p}^2 = .02\); and bank angle, \(F(2, 24) = 0.46, p = .639, \eta_{p}^2 = .04\)—nor the interaction effects became significant. Only the AI Format \(\times\) Bank Angle interaction at least approached the usual level of significance, 
\[F(1.57, 18.86) = 3.49, p = .061, \eta_{p}^2 = .23,\] reflecting that differences between response times for the different bank angles were somewhat larger in the MH compared to the MA condition. For all other interaction effects, the \(F\) values were \(< 1.0\).
**Figure 3.** Pilots’ means of (A) reversal error and (B) response time for both horizon design conditions, classic and extended, and both attitude indicator (AI) format conditions, moving horizon (MH) and moving aircraft (MA). Error bars represent standard errors.

**Figure 4.** Pilots’ means of (A) reversal error and (B) response time over both horizon design conditions for both attitude indicator (AI) format conditions, moving horizon (MH) and moving aircraft (MA), and for each bank angle condition, 45°, 90°, and 135°. Error bars represent standard errors.
**Workload.** No significant effects whatsoever were found for the NASA–TLX data, all $F \leq 1.4$, $p \geq .26$, $\eta^2 \leq .10$.

**Discussion**

The results of the second experiment suggest that the MA format provides performance advantages compared to the MH format even for pilots who are familiar and well trained with the MH format. However, these advantages only emerged in the recovery task whereas for flight-path tracking a reverse effect was found. In addition, the advantages seem to be largely reduced when using an extended horizon compared to the classic horizon design.

Let us again first consider the results in condition classic horizon design. In contrast to the first experiment, no benefit of the MA compared to the MH format was found for flight-path tracking. Instead, a reverse effect emerged, favoring the MH format. Most likely, it reflects the fact that our pilots all have gained extensive practice in conducting such tracking tasks with the MH format from flight training and on-the-job experience. This obviously helped to more than compensate for the disadvantages of this format in terms of compatibility. Given this, it is remarkable that the differences between both AI formats were small (RMSE of 2.22° vs. 2.64°), and hardly of any practical significance. This corresponds to other studies, which often have found MH trained pilots performing tracking tasks almost as well with the MA as with the MH format (Beringer et al., 1975; Cohen et al., 2001).

In contrast, the results of recovery task performance with the classic PFD design directly mirrored the results of novices. Even though the pilots were trained with the MH format and had a mean of about 380 flight hours of IFR experience, they committed
a higher rate of reversal errors when performing the recovery task with the MH compared to the MA format in the classic horizon condition. The rate of reversal errors with the classic MH format directly corresponds to what has been found in previous research with pilots (Previc & Ercoline, 1999). The finding of an advantage of the MA compared to the MH format was almost surprising because previous research with pilots was somewhat inconsistent in this respect, with most studies not reporting a clear advantage for either format (cf. Previc & Ercoline, 1999). A straightforward explanation for this finding is that the recovery task used in this study with sudden and distinct changes of the attitude by 45° to 135° represents a rather unusual and less practiced task, even for pilots. Hence, our finding confirms the assumption that the MA format is generally more intuitive and better able to support tasks one is not specifically trained for than the MH format. It also suggests that a transfer from MH to MA would be possible without many problems even for MH trained pilots. Overall, this again supports the assumption that the mental representation of most pilots rarely includes a world moving around their aircraft, but rather an aircraft moving in reference to a stable world, which is better represented by the MA than the MH format (Johnson & Roscoe, 1972; Kovalenko, 1991; Previc & Ercoline, 1999).

However, with the extended horizon PFD, the previously discussed effect is reduced or even eliminated in the recovery task. It is interesting that this expected reduction of differences in performance between the AI formats was only observable with pilots, but not with novices. This finding supports what we suspected when discussing the results of the first experiment. We assumed that the novices might not have acquired a proper understanding of the basic idea of the MH display and, therefore, could not benefit from an improved figure–ground representation provided by the extended horizon
design. The pilots, on the contrary, had much experience with visual as well as instrument flying and thus had a much better understanding of the relationship between the perceived “movements” of the natural horizon induced by a banking aircraft and the movements of their MH AI display. Consequently, they could benefit from the better figure–ground separation achieved by the extended horizon display than novices, despite it still being presented in the peripersonal space. Thus, extended horizon displays at least seem to be able to reduce or even eliminate the differences between the two AI formats, although they still cannot be expected to reverse the usually found performance differences between these formats. This seems true for pilots who have a proper understanding of what the basic idea of the MH format is.

**Summary and Conclusion**

The intention of the experiments presented in this article was to investigate to what extent the superiority of the MA format over the MH format in maintaining spatial orientation persists with current PFD technology. Four findings seem to be important in this respect.

First, the overall results of our experiments show that even with modern PFD technology found in today’s glass cockpits of civil aircraft, the findings of the early studies with small round electro-mechanical instruments can be replicated in most aspects. This provides strong evidence that the superiority of the MA format over the MH format with respect to supporting spatial orientation and intuitive understanding of the depicted attitude changes still persists with current PFDs. Thus, Western civil aircraft still seem to use an inferior AI format as part of the PFD.

Second, the fact that this was not only true for novices, but also experienced pilots, provides evidence that even for this latter group the principle of the moving part
represented by the MA format is the more efficient principle for supporting quick and correct responses to unexpected attitude changes than the competing principle of pictorial realism underlying the MH format.

Third, and related to the conclusion before, these findings suggest that even pilots trained and experienced to fly with an MH formatted AI would probably be able to switch to the MA format without any negative performance consequences. Thus, our findings support the conclusion of Previc and Ercoline (1999) that the aviation community might seriously reconsider an implementation of the MA concept. However, we frankly acknowledge that the chances for such change are probably close to zero, given the enormous effort in terms of investments and new certifications needed. However, for rather new applications, such as control stations for remotely piloted aerial systems, an implementation of AIs corresponding to the MA format should seriously be taken into consideration. Here it is especially advisable, because the principle of pictorial realism does not seem to be appropriate in any way for remote controllers not sitting in the aircraft they control (cf. Previc & Ercoline, 1999).

Fourth, the extended horizon PFD used in our experiments seemed to reduce or even eliminate the effect of a superior MA format at least for pilots. Thus, if the use of the MH format is inevitable, it is recommended at minimum to implement the extended horizon design, due to the better support of a proper figure–ground separation when interpreting the display.

Some limitations of this research should be considered along with these conclusions. First, our study involved only novices and pilots trained with the MH format. Given the latter, we only could investigate possible effects involved in a transfer from the MH to the MA format but not vice versa. Yet, it would be interesting to investigate
the reverse transfer as well. If our conclusions are correct, we would suggest that transferring from the MA to the MH format should be associated with much more severe performance consequences. This has been supported by observations of Kovalenko (1991) and Ponomarenko, Lapa, and Lemeshchenko (1990) but has rarely been addressed in systematic studies. The only exception we are aware of is the training study of Yamaguchi and Proctor (2010), although they did not find asymmetric transfer effects.

Second, the pilots we were able to recruit for the experiment, were relatively young. It cannot be excluded that pilots with a year-long experience would produce different results.

Third, the classic recovery task used in our studies only involved discrete bank deflections as stimulus. In aviation, however, dynamic deflections or disturbances are usually occurring. These could even increase time pressure on recoveries from unusual attitudes and exacerbate the problems of the MH format. A current study in our lab is addressing this issue.

Fourth, the studies were conducted in a simplified fixed-base flight simulator. It is not certain that the results obtained here can be generalized to real flying situations, which provide further cues for spatial awareness (e.g., vestibular feedback). Finally, with the extended design, one of the most recent design variants of civil aviation PFDs was considered in our experiments. However, other new developments are already available in some modern cockpits, which might create entirely new circumstances for the evaluation of proper AI formats, such as synthetic vision and head-up displays. With these displays, the issue of MH versus MA could be different (e.g., Beringer & Ball, 2009) and more research will be needed to see whether this issue eventually will become moot.
Thus, it seems that questions concerning proper AI design will remain an important topic of human factors research also in the future.

Acknowledgments

The authors thank all experimental participants and Stephan Pietschmann for his software support.
References


