Concurrent Pilot Instrument Monitoring in the Automated Multi-Crew Airline Cockpit

Stephen R. Jarvis

INTRODUCTION: Pilot instrument monitoring has been described as "inadequate," "ineffective," and "insufficient" after multicrew aircraft accidents. Regulators have called for improved instrument monitoring by flight crews, but scientific knowledge in the area is scarce. Research has tended to investigate the monitoring of individual pilots when in the pilot-flying role; very little research has looked at crew monitoring, or that of the "monitoring-pilot" role despite it being half of the apparent problem.

METHODS: Eye-tracking data were collected from 17 properly constituted and current Boeing 737 crews operating in a full motion simulator. Each crew flew four realistic flight segments, with pilots swapping between the pilot-flying and pilot-monitoring roles, with and without the autopilot engaged. Analysis was performed on the 375 maneuvering-segments prior to localizer intercept.

RESULTS: Autopilot engagement led to significantly less visual dwell time on the attitude director indicator (mean 212.8–47.8 s for the flying pilot and 58.5–39.8 s for the monitoring-pilot) and an associated increase on the horizontal situation indicator (18–52.5 s and 36.4–50.5 s).

DISCUSSION: The flying-pilots' withdrawal of attention from the primary flight reference and increased attention to the primary navigational reference was paralleled rather than complemented by the monitoring-pilot, suggesting that monitoring vulnerabilities can be duplicated in the flight deck. Therefore it is possible that accident causes identified as "inadequate" or "insufficient" monitoring, are in fact a result of parallel monitoring.

KEYWORDS: visual scanning, pilot attention, instrument monitoring.

Jarvis SR. *Concurrent pilot instrument monitoring in the automated multi-crew airline cockpit*. Aerosp Med Hum Perform. 2017; 88(12):1100–1106.

For at least two decades investigators of accidents to commercial multicrew aircraft have implicated "crew monitoring" as a factor leading to the event. As far back as 1994 the NTSB reported that 31 of 37 "flight crew-involved" major accidents involved "inadequate monitoring and/or cross checking" by flight crews.[14] Bodies including regulators and advisory organizations have repeatedly highlighted the need for better pilot monitoring.[5,8,10]

Despite these efforts, descriptions of recent accidents to multicrew commercial air transport aircraft have included "*flight instruments were not monitored effectively,*"[1] "*flight crew's insufficient monitoring of airspeed indications,*"[16] "*failure of monitoring the airspeed and pitch attitude of the aircraft,*"[6] and "*a significant breakdown in their [both pilots'] monitoring responsibilities.*"[15] Such descriptions are not based on data or evidence of how the pilots were monitoring at the time, since there is no such evidence, but on the single finding that the crew failed to notice one or more pieces of critical information that, in hindsight, could have alerted them to the potential event that was about to happen.

The problem with such claims is that they implicitly assume that it is possible for crews to monitor adequately, effectively or sufficiently, but the current state of scientific knowledge is insufficient to support such assumptions. Despite claims that studies of pilot monitoring behavior are plentiful,[20] most such studies have concentrated on demonstrating the existence of a problem (i.e., what pilots fail to look at or notice) rather than investigating the nature of pilot monitoring. There are very few studies that explore how both the flying-pilot (PF) and the

From Jarvis Bagshaw Ltd, Husborne Crawley, UK.

This manuscript was received for review in March 2017. It was accepted for publication in August 2017.

Address correspondence to: Stephen R. Jarvis, Jarvis Bagshaw Ltd, Husborne Crawley, Bedfordshire, UK; steve@jarvisbagshaw.com.

DOI: https://doi.org/10.3357/AMHP.4882.2017

monitoring-pilot (PM) monitor the instruments in a multicrew automated flight deck. The development of this fundamental knowledge is a crucial missing step between awareness of the apparent problem and the guidance on how to avoid it.

Pilots' visual attention is driven by a mixture of bottom-up and top-down processes,[24] meaning that it is either drawn to elements (attracted by them) or driven to them (i.e., directed attention). Likelihood of attention allocation to a specific visual area has been modeled using the SEEV model.[23–25] The model's name is derived from the four factors assumed to drive visual attention: salience (S) of events in terms of capturing attention (for example a flashing light is more salient than a steady light), the effort (E) required to switch attention to the new area, the expectancy (E) that the location will contain information, and the value (V) of the information to be obtained there.[23] Whereas eye movement is said to be driven by salience, expectancy and value, it is inhibited by the effort required to redirect attention (e.g., rotate the head and eyes to fixate on the new area).[23,24] One equation for how these four factors model the probability of attention to an area has been proposed as $P(A) = sS–efE + ev * EV$.[24]

When manually flying, pilots have been consistently found to exhibit a visual scan that centers on the primary attitude information and samples the main flight instruments.[7,17,19] Usually called the "T-scan,"[11,21] a major characteristic is the very high total dwell time on the primary attitude instrument compared to all other areas of interest.[7,19] This occurs due to the nature of the manual flying task because the attitude display (usually in the form of an Attitude and Direction Indicator, ADI) is the most data-rich and highest bandwidth instrument, displaying both pitch and bank information[22] and provides the most important information about vehicle status.[11]

Observations of scan behavior when an autopilot is engaged show different characteristics to manual flying,[4,19] most notably a large reduction in time spent on the attitude indicator when an autopilot is engaged. For pilots flying a multiphase sector, ADI dwell time was only about 20% with the autopilot engaged compared to over 50% without.[7]

Such findings run contrary to current regulatory guidance. The FAA states that a pilot-flying should always be engaged in flying the aircraft even when the aircraft is under autopilot control[9] and UK CAA guidance advises pilots to "monitor the flight instruments just as you would when you are manually flying the aircraft."[5] In light of the scientific knowledge, such advice may be unrealistic.

Reductions in attitude scanning are often accompanied by increased attention to navigational displays.[7,19] For example, 330 pilots flying realistic arrivals exhibited 18.2% of dwell time on the navigational display (ND) and only 10% on the attitude direction indicator.[2] The reason appears to be related to task prioritization. Pilots often use aviate-navigate-communicate-systems (ANCS) to prioritize flight deck tasks including monitoring;[18] the "aviate" priority being aligned with pilot attention to the attitude indicator or primary flight display, and the "navigate" priority being aligned with attention to the navigational displays.[12,23] Hence it appears that when the autopilot is engaged, some of the visual resource previously used for the aviate task moves to the navigate task.

In multicrew aircraft, further questions arise in terms of how the aviate and navigate tasks should be shared between the crew, and specifically whether the PM should prioritize differently to the PF. It has been claimed that current practice is for PF and PM roles to similarly prioritize tasks.[12] In line with this, FAA guidance puts responsibility for "the current and projected flight path and energy of the aircraft at all times" onto both pilot roles, designating responsibility for its management to the PF while making the PM responsible for its monitoring.[9] However the same FAA document hints at different prioritizations between pilots:

"The PF is always engaged in flying the aircraft… and avoids tasks or activities that distract from that engagement. If the PF needs to engage in activities that would distract from aircraft control, then PF should transfer aircraft control to the other pilot, and then assume the PM role."[9]

This suggests that the aviate task is expected to be a higher priority for the PF than the PM. Notably however there is no explicit guidance on the prioritization of the navigational task (either in relation to other tasks or between pilots). If the PF must prioritize more attention than the PM to the aviate task then there is a question of whether the PM should compensate by prioritizing their monitoring to navigation and systems elements. This has been suggested for the modeling of NextGen operations.[12]

There is very little scientific knowledge relating to how pilots monitor instruments or prioritize when in the PM role. Visual scans of PMs observed during go-around maneuvers[3] were broadly spread with relatively low dwell time on the attitude (horizon). If similar to PF autopilot-engaged scans, then the result could be low levels of joint crew attention to the aviate task.

The lack of scientific knowledge around monitoring practices in multicrew automated aircraft, particularly for the PM role, has not kept up with the growing international focus on crew monitoring. This research aimed to begin the process of understanding multicrew monitoring practices, particularly relating to overall task prioritization of the different roles under different automation levels.

METHODS

Subjects

There were 34 Boeing 737 pilots from two large UK airlines; 17 captains and 17 first officers. Pilots were paired into properly constituted crews, so that each crew contained a captain and a first officer. Data from two captains were rejected due to eye-tracking problems, so demographic data are for the remaining 32 pilots group. There were 3 women and 29 men. Age brackets were 20–29 yr ($N = 4$), 30–39 yr ($N = 12$); 40–49 yr ($N = 7$) and 50–59 yr ($N = 9$). Mean airline flying experience was 14.97 yr (SD 7.84), 8493.75 flying hours (SD 4239.97). Mean type experience (Boeing 737) was 5718.75 (SD 2660.76). All pilots were in current practice; mean time between their previous flight and experiment participation was 4.25 d (SD = 6.04). Participation was voluntary; pilots were fully informed of the purpose of the study, anonymity of data collection, and pilots were made aware of their right to withdraw at any time. Pilots were informed that they would be told the aim of the study after the trials, and this was done. No company personnel other than the crew were present throughout the process. The companies involved had no sight of any individual data. The study protocol was approved in advance by the ethics review board of Jarvis Bagshaw Ltd, consisting of one medical practitioner, three qualified human factors personnel and two lay-persons. Subjects provided informed consent before participating.

Apparatus

All trials were conducted in a certified full-motion Boeing 737-400 simulator in current airline use for crew training and checking. SMI 30-Hz eye-tracking glasses were worn by all pilots. These were calibrated at the start of the session then checked and recalibrated (if required) between trials. Headsets were not worn; a speaker system was used to play Air Traffic Control audio files to the crew during each trial.

Procedure

For this study, crews flew four trials in the simulator. Each simulator session included eight scenarios in total with four being relevant to this research and four being for a different set of analysis (however these were also nonemergency, normal scenarios). The experimental session lasted between 2 and 3 h, and each scenario was about 12 min in length.

Before each trial, pilots heard the Air Traffic Information Service broadcast (ATIS), and were then allowed time to complete their own predescent briefing with the simulator in flight-freeze at the scenario start point, prior to being unfrozen when crews were ready.

All trials were radar-vectored arrivals starting 2 mi (3.2 km) inbound to either Bovingdon or Ockham VOR radio beacons at FL100 to an ILS approach and landing onto runway 09L or 27L at London Heathrow. However in this particular study only the arrival was used in the analysis; the data from the approach and landing were not included. The scenario (adjusted for each trial) was created and tested after consultation with each company's flight data departments and pilots, in order to realistically emulate a normal, busy segment of a line flight.

The automation configuration was stipulated as either autopilot-engaged or autopilot disengaged (manual flying). Flight directors remained switched on in both cases. Each pilot took the role of PF and PM once for each automation configuration, making four experimental conditions (levels of the independent variable): PF, Autopilot-Engaged, PF, Autopilot-Disengaged, PM, Autopilot-Engaged, and PM, Autopilot-Disengaged. The sequence in which crews were allocated these four conditions was counter-balanced (different for each crew).

Realistic Air Traffic Control (ATC) audio tracks provided radar vectors and clearances throughout. Other traffic was included in order to create a realistic environment in which crews had to determine which ATC calls were relevant, as they must do normally in a line flight. Each trial used a different audio track. The tracks were exactly equivalent in terms of timing, amount of maneuvering, number of ATC commands, weather and quantity of other traffic, but the arrival routings were different for each. Tracks were tested multiple times by two type-rated pilots and adjusted in an attempt to make all of them equivalent in terms of pilot workload. Additionally, the track sequence was counterbalanced. Hence each crew was presented with a different order of tracks matched to different automation levels and roles from other crews. This mitigated for any possible remaining workload differences between scenarios, familiarization effects and practice effects.

The cloudbase was always broken 300 ft (91 m) and overcast 200 ft (61 m) with a visibility of 5 km, meaning that crews were in instrument meteorological conditions until late in the approach.

Data Treatment

Eye movement data were collected and coded onto an instrument panel template using SMI BeGaze software (New Biotechnology Ltd., Jerusalem, Israel). Before coding the eye tracking data, two experimenters rechecked the calibration and offset the original calibration if required for greater accuracy.

All coding and analyses were conducted on a 375-s segment starting from 2 min inbound to the beacon for all tracks. The end of the segment (375 s) occurred approximately 1 min prior to intercepting the localizer. Total dwell time on the EADI (Electronic Attitude Direction Indicator; primary attitude information) and EHSI (Electronic Horizontal Situation Indicator; primary navigational information) were compared between the PF and PM roles for both automation levels, for each of the 375-s segments. "Total dwell time" is the total amount of time (in seconds) that a pilot's eyes spend looking at a particular instrument throughout the 375 s of the scenario, and is calculated by adding together the lengths of all the fixations that the pilot made on that instrument during that period. Additionally, the measures of PF and PM dwell time were combined to form a new measure termed "crew dwell time" in order to compare the total crews' dwell time between the two automation levels.

IBM SPSS v24 software was used for all statistical analysis.

RESULTS

Before separating the data into the four experimental conditions (PF/PM, autopilot engaged/disengaged), a one-way repeated-measures ANOVA was run to compare pilot dwell times (in seconds) between the six major displays (altimeter, speed tape, EADI, EHSI, heading indicator and vertical speed indicator) over the total 1500 s (four 375-s segments) of flying done by each pilot. One subject's data were removed due an extreme outlier on speed tape dwell time. Five marginal outliers were retained, after testing with them removed showed the same conclusion as with them included. Assumptions of normality were upheld. Epsilon (ϵ) was 0.41 as calculated according to Greenhouse and Geisser,[13] and this correction was used. The dwell time was statistically significantly different between the six AOIs; $F(2.032, 60.966) = 207.956$, $P < 0.0005$, partial $\eta 2 = 0.874$. Pairwise comparisons showed that all combinations of AOIs were significantly different to each other ($<$ 0.0005) with the exception of the altimeter and speed tape ($P =$ 0.645).

All subsequent analysis was performed on the AOI dwell time across the four levels of the independent variable. **Fig. 1** shows the extent to which the EADI was used by the PF when manually handling; mean 212.8 s (SD 58.05 s) as opposed to when the autopilot was engaged, mean 47.8 s (SD 24.6 s). The EADI accounted for over half the mean dwell time in the manual handling condition. All 32 pilots had a higher mean dwell time on the EADI than on the EHSI when flying manually.

Two one-way repeated measures ANOVAs were run (EADI and EHSI) to compare dwell-time across the four conditions (four levels of the independent variable). Four nonextreme outliers were found for EADI dwell time and three for EHSI dwell time. All were kept after testing found that their removal did not change the main result of either ANOVA. All dwell times were normally distributed as assessed by Shapiro-Wilk's test (> 0.05), except the EHSI dwell time for the PF manual-flying condition (0.015). For additional confidence that the ANOVA results were not adversely impacted, two nonparametric Freidman's tests (with post hoc pairwise comparisons) were run in parallel to the ANOVAs and matched the equivalent output of both, confirming that the outliers and the single violation of normality did not adversely affect the ANOVA results.

ϵ was 0.51 for the EADI and 0.82 for the EHSI, as calculated according to Greenhouse and Geisser,[13] and used to correct the one-way repeated measures ANOVAs. **Table I** shows the descriptive statistics for the EADI ANOVA and the EHSI ANOVA.

EADI mean dwell time was statistically significantly different across the four conditions; $F(1.503, 46.582) = 252.604$, $P <$ 0.0005, partial $\eta 2 = 0.891$. With the autopilot disengaged, the PF EADI mean dwell time significantly increased by165 s (95% CI, 139.70 to 190.35, $P < 0.0005$) and the PM EADI mean dwell time significantly increased by 18.67 s (95% CI, 7.27 to 30.07), $P < 0.0005$, compared to when the autopilot was engaged. With the autopilot disengaged, PF EADI mean dwell time was 154.36 s higher than PM EADI mean dwell time, which was statistically significant; 95% CI, 125.62 to 183.01, $P < 0.0005$. With the autopilot engaged, there was no significant difference between the PF and PM mean EADI dwell time ($P = 0.21$). No significant difference was found in mean EADI dwell time between the PF with the autopilot engaged and the PM when the autopilot was disengaged ($P = 0.15$).

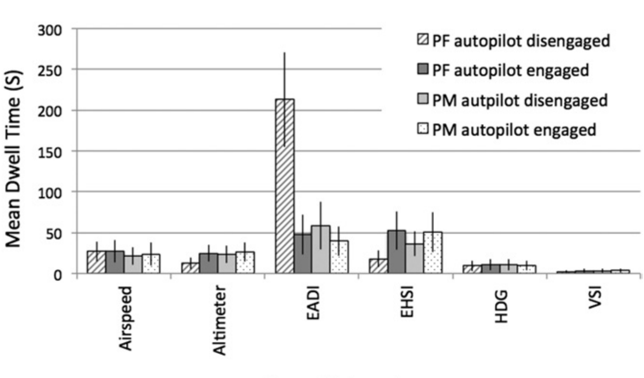

Fig. 1. Mean total dwell time (in seconds) for eight main instrument AOIs in the four conditions, for the 375-s segment. Error bars represent 1 SD.

ESHI mean dwell time was statistically significantly different across the four conditions; $F(2.455, 76.107) = 45.419$, $P <$ 0.0005, partial $\eta 2 = 0.594$. With the autopilot disengaged, the PF EHSI mean dwell time significantly decreased by 34.42 s (95% CI, 24.69 to 44.14), $P = 0.001$, and the PM EHSI mean dwell time significantly decreased by 14.14 s (95% CI, 23.64 to 46.44), $P = 0.001$, compared to with the autopilot engaged. There was no significant difference in EHSI mean dwell time between PF and PM with the autopilot engaged ($P = 1.0$). However, without the autopilot engaged PF EHSI mean dwell time was 18.32 s lower than mean PM EHSI mean dwell time, which was statistically significant; 95% CI, 12.50 to 24.15, $P < 0.0005$. A significant mean difference of 16.97 s was found between the PF EHSI mean dwell time with and without the autopilot engaged (CI, 7.04 to 25.16), $P < 0.0005$.

Paired samples t-tests were used to compare the EADI and EHSI dwell time under each of the four conditions, using a Bonferroni adjusted confidence interval of 0.0125 (0.05 / 4).

For the PF manual flying condition, due to one violation of normality (Shapiro Wilks = 0.015) a logarithmic transformation (Log10) was applied to the EHSI and EADI data. Both transformed variables met assumptions of normality (Shapiro Wilks = 0.1 and 0.65). The transformed EADI data contained no outliers, but because the EHSI data contained two (one of which was extreme) the t-test was rerun with both these outliers removed (Shapiro Wilks = 0.86 and 0.57). In both cases the t-tests were found to be highly significant (without outliers $t(31) = 25.613$, $P < 0.0001$; with outliers $t(29)=24.45$, $P <$ 0.0001).

In the PF autopilot-engaged condition there were two outlying data points, neither was extreme. Assumptions of normality were upheld (Shapiro Wilks = 0.213 and 0.227). The paired samples t-test was not significant ($t(31)=-0.89$, $P = 0.380$). A parallel t-test with the outliers removed yielded the same nonsignificant conclusion, so the result from the initial t-test (including outliers) was used.

For the PM manual flying condition assumptions of normality were upheld (Shapiro Wilks = 0.333 and 0.369) and there were no outliers. The paired samples t-test was significant ($t(31)=3.673$, $P = 0.001$).

In the PM autopilot-engaged condition there was one extreme outlier, removal of which created two marginal outliers. With all outliers removed and assumptions of normality upheld (Shapiro Wilks = 0.945 and 0.195) the subsequent t-test was highly significant ($t(28)=-3.010$, $P = 0.005$).

Table I. Descriptive Statistics for EADI and EHSI Dwell Time for Repeated Measures ANOVAs ($N = 32$ in All Cases).

AREA OF INTEREST (AOI)	EXPERIMENTAL CONDITION	MEAN DWELL TIME (s)	SD (s)
EADI	PF Autopilot Engaged	47.81	24.59
EADI	PF Autopilot Disengaged	212.84	58.06
EADI	PM Autopilot Engaged	39.81	18.04
EADI	PM Autopilot Disengaged	58.48	28.91
EHSI	PF Autopilot Engaged	52.45	23.57
EHSI	PF Autopilot Disengaged	18.03	10.04
EHSI	PM Autopilot Engaged	50.49	24.26
EHSI	PM Autopilot Disengaged	36.347	15.21

The PF and PM dwell time were combined for each segment to create a new measure termed "crew dwell time." Paired samples t-tests were run for both the EADI and the EHSI, to compare overall crew dwell time between the autopilot engaged and autopilot disengaged conditions for each. Because there were two such tests, a Bonferonni adjustment was made, giving a confidence interval of 97.5%.

For the EADI, assumptions of normality were upheld (Shapiro Wilks = 0.524 and 0.896) and there were no outliers. The paired samples t-test showed that crew dwell time on the EADI was higher with the autopilot disengaged (271.32 ± 71.31 s) than with it engaged (87.63 ± 37.94 s), a statistically significant increase of 183.69 s (95% CI, 161.53 to 205.86 s, $t(31) = 19.524$, $P < 0.0005$).

For the EHSI, assumptions of normality were upheld (Shapiro Wilks = 0.09 and 0.152). There was one outlier that was retained, after running the t-test without it obtained the same level of significance (also with normality assumptions upheld). The paired samples t-test showed that crew dwell time on the EHSI was lower with the autopilot disengaged (54.37 ± 22.97 s) than with it engaged (102.94 ± 42.81 s), a statistically significant decrease of 48.53 s (95% CI, 35.69 to 61.43 s, $t(31) = 8.889$, $P < 0.0005$) (**Fig. 2**).

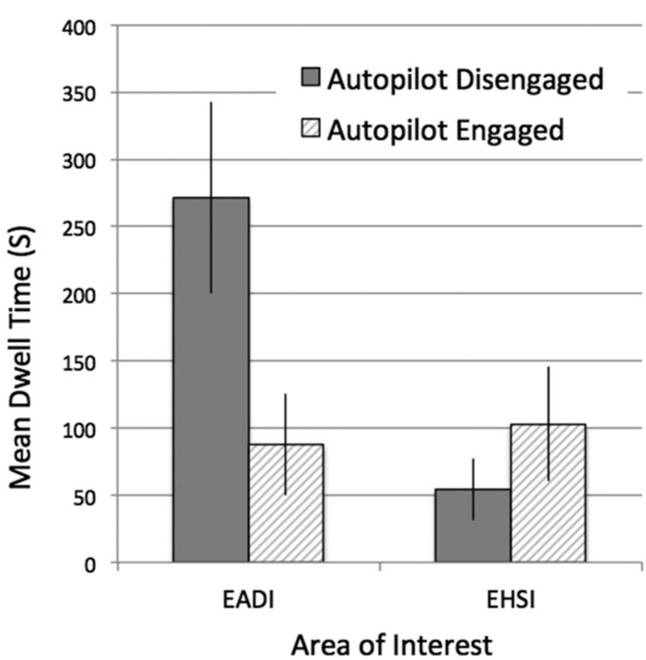

Fig. 2. Mean crew dwell time (in seconds) on the EADI and EHSI under the two automation conditions (autopilot disengaged and autopilot engaged). Error bars show 1 SD.

In summary, the results showed that the EADI and EHSI were the most used AOIs overall. When flying manually, PFs monitored the EADI significantly more, and the EHSI significantly less, than for the three nonhandling conditions. In common with PFs, when the autopilot was disengaged PMs spent significantly more time looking at the EADI and significantly less time looking at the EHSI. With the autopilot engaged, PMs spent significantly more time on the EHSI than the EADI, but for PFs there was no evidence of a difference between the two areas of interest. Crews (combined PF and PM dwell time) looked at the EADI nearly three times more, and the EHSI only about half as much, when the autopilot is disengaged, compared to when it is engaged.

DISCUSSION

In line with previous research,[7,17,21] the attitude indicator (EADI) was the dominant area of interest in the visual scan of pilots when flying manually, accounting for over half of all PF dwell time, significantly more than in the other conditions. This is because the ADI provides essential information for manual instrument flying.[11,22]

A sharp reduction in attention to the attitude indicator was found when the autopilot was engaged, supporting previous findings from commercial jet transport aircraft.[7,19] According to the SEEV model, salience, expectancy, and perceived value are key drivers in attention allocation, while effort is the inhibitor.[23,25] Since the salience of the EADI and the effort required to view it remain the same regardless of the autopilot status, the results suggest that the perceived value and/or expectancy must be considerably less in the nonhandling tasks. Unlike when engaged in the skill of manual flying, looking at the EADI when the autopilot is engaged will rarely reveal an unexpected or perceived-valuable event requiring activity, and repeated exposure will therefore lead to low levels of expectancy and perceived value. Such unconscious (skill-based) "de-valuing" of the attitude information over time inevitably leads to the observed result. This probably cannot be changed by advising pilots to "*monitor the flight instruments just as you would when you are manually flying the aircraft*,"[5] since that would require continual and unsustainable conscious effort to overcome the continually reinforced low levels of expectancy and perceived value. Hence current advice from both FAA and CAA is unrealistic, despite being well intentioned.

Interestingly, as well as the PF, the PM was found to monitor the EADI less when the autopilot was engaged. It is possible that PMs' expectancy and perceived value of the EADI is higher when a PF is flying manually than when the autopilot is engaged (i.e., PM glances at the EADI will more often result in a valuable event when the autopilot is disengaged, than when it is engaged).

Along with the significantly more attitude monitoring when flying manually, flying-pilots (PF) showed a significant decrease in monitoring on the primary navigational display (EHSI) compared to when the autopilot was engaged, supporting previous research.[7,19] Yet, given the counterbalancing and the high level of similarity of the scenarios flown under different conditions, the factors in the SEEV model (expectancy, value, and salience) relating to the EHSI would have remained equivalent regardless of autopilot engagement. Hence the additional dwell time on the EHSI was not caused by an increase in navigational task need, and this suggests that the PF would have preferred to spend more time on the navigational task when manually flying than they were able to. The results suggest that when flying manually, pilots are unable to devote the ideal amount of dwell time to both the aviate and navigate task, and the navigational task gets deprioritized. This is particularly important given the PM result under the same condition and suggests that both pilots prioritize similarly, in line with previous observations.[12] Hence, rather than compensating for PFs' reprioritization of attention between the two automation levels, PM attention appears to parallel it.

Nonhandling roles (both PM roles and the PF autopilot-engaged role) were not found to be different from each other. The PF with the autopilot engaged showed no significant difference in dwell time in either of the PM roles, on either the EADI or EHSI, and total dwell time means were similar. This suggests that the sort of scan practiced by PFs with the autopilot engaged is more closely aligned to PM scanning than to the PF handling scan. The implication of this finding is that autopilot engagement results in both the PF and PM using parallel monitoring strategies, with the result that the aviate task may not be prioritized by either pilot, contrary to the FAA[9] and the CAA[5] advice. This was confirmed by assessing the combined crew dwell time under the two different automation conditions: with the autopilot engaged the crew (both pilots together) monitored the EADI only about a third as much as with the autopilot engaged, and the EHSI nearly twice as much.

These findings may be characteristic of the particular task and segment (high navigational task load) but it is also possible that similar phenomenon could occur with any task or situation that results in expectancy and/or value on instruments or displays other than those used for the aviate task. A potential risk would manifest itself in situations where control is not maintained by the autopilot or the autopilot disconnects without alerting the crew. The likelihood of such situations being noticed due their subsequent effect on the immediate flight path will be reduced where parallel attention allocation by both crewmembers is occurring that prioritizes task concerns other than the aviate task.

Overall, this research indicates a potential risk that the critical aviate task is relatively lightly monitored by both pilots at the same time when another task (such as navigation) demands attention, as suggested by the comparison of "crew dwell time" between the EADI and EHSI. A coincidental and unintentional change in flight path or critical parameter may go unnoticed, not because both pilots' monitoring was insufficient or inadequate, but because their attention coincided on a different task area due to the expectancy and value of that area's visual elements being higher than those of the aviate task. Essentially, with the autopilot engaged, the aviate task is vulnerable to lower relative levels of monitoring by both pilots (PF and PM), whereas the opposite is true with the autopilot disengaged (manual flying) where the navigation task is potentially deprioritized by both pilots. Potential solutions lie in researching and implementing methods by which pilots can complement each other's monitoring role.

Based on this research, an important consideration related to current industry thinking is whether accidents previously explained with terms such as inadequate pilot monitoring may have been caused by parallel pilot monitoring, where both pilots were actively monitoring, but in the same areas.

ACKNOWLEDGMENTS

This research was commissioned, funded and facilitated by British Airways PLC, Harmondsworth, London, and Thomson Airways Ltd., Luton, Beds., UK.

Author and affiliation: Steve Jarvis, M.Sc., Ph.D., Jarvis Bagshaw, Ltd., Husborne Crawley, Bedfordshire, UK.

REFERENCES

1. AAIB. Report on the accident to AS332 L2 Super Puma helicopter, G-WNSB on approach to Sumburgh Airport on 23 August 2013. Department for Transport. Aldershot (UK): Air Accidents Investigation Branch; 2016.
2. Anders G. Pilot's attention allocation during approach and landing eye-and head-tracking research in an A330 full flight simulator. Berlin (Germany): Institute for Aeronautics and Astronautics Technische Universität Berlin; 2001.
3. BEA. Study on airplane state awareness during go-around. Le Bourget, Paris (France): Bureau d'Enquetes et d'Analyses pour la securite de l'aviation civile; 2013.
4. Brown DL, Vitense HS, Wetzel PA, Anderson GM. Instrument scan strategies of F-117A pilots. Aviat Space Environ Med. 2002; 73:1007–1013.
5. Civil Aviation Authority. Monitoring Matters CAA. Guidance on the development of pilot monitoring skills. CAA Paper 2013/02. London (UK): UK Civil Aviation Authority; 2013.
6. Dutch Safety Board. Crash during approach, Boeing 737–800, near Amsterdam Schiphol Airport, 25 February 2009. The Hague: Dutch Safety Board; 2010.
7. Edwards RE, Tolin P, Jonsen GL. Pilot visual behavior as a function of navigation and flight control modes in the Boeing 757/76. Proceedings of the Human Factors Society 26th Annual Meeting, Seattle, WA. Santa Monica (CA): HFES; 1982; 26:441.
8. FAA. Advisory Circular C No 120-51E; Crew Resource Management Training. Washington (DC): U.S Department of Transportation, Federal Aviation Administration; 2004.
9. FAA. SAFO (Safety Alert for Operators) 15011. Washington (DC): U.S. Department of Transportation, Federal Aviation Administration, Flight Standards Service; 2015.
10. Flight Safety Foundation. A practical guide for improving flight path monitoring. Final Report of the Active Pilot Monitoring Working Group. Alexandria (VA): Flight Safety Foundation; 2014.

11. Gainer CA, Obermayer RW. Pilot eye fixations while flying selected maneuvers using two instrument panels. Hum Factors. 1964; 6(10): 485–501.
12. Gore BF, Hooey BL, Mahlstedt E, Foyle DC. Evaluating NextGen closely spaced parallel operations concepts with validated human performance models: scenario development and results. Document NASA/TM-2013-216503. Washington (DC): National Transportation Safety Board; 2013.
13. Greenhouse SW, Geisser S. On the methods in the analysis of profile data. Psychometrika. 1959; 24(2):95–112.
14. NTSB. a review of flightcrew-involved major accidents of U.S. carriers, 1978 through 1990. Document NTSB/SS-94/01. Washington (DC): National Transportation Safety Board; 1994.
15. NTSB. Loss of control on approach Colgan Air, Inc. operating as Continental Connection flight 3407 Bombardier DHC-8-400, N200WQ Clarence Center, New York, February 12, 2009. Document NTSB/AAR-10/01 PB2010-910401. Washington (DC): National Transportation Safety Board; 2010.
16. NTSB. Descent below visual glidepath and impact with seawall Asiana Airlines Flight 214 Boeing 777-200ER, HL7742 San Francisco, California, July 6, 2013. Document NTSB/AAR-14/01 PB2014-105984. Washington (DC): National Transportation Safety Board; 2014.
17. Rinoie K, Sunada Y. Efficient eye-scanning for reducing pilot workload–single pilot IFR and VFR flight tests. 23rd Congress of International Council of the Aeronautical Sciences, 8–13 September 2002, Toronto, Canada; 2002.
18. Schutte PC, Trujillo AC. Flight crew task management in non-normal situations. Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting, September 2–6, 1996, Philadelphia, PA. Santa Monica (CA): HFES; 1996.
19. Spady AA. Airline pilot scan patterns during simulated ILS approaches. NASA Technical Paper 1250. NASA Langley Research Center (VA): NASA; 1978.
20. Sumwalt RL. Statement on Asiana Airlines flight 214. In: NTSB, Report No NTSB/AAR-14/01 PB2014-105984. Washington (DC): National Transportation Safety Board; 2014.
21. Verma S, Corker K, Jadhav A. Modelling visual behavior of pilots in the human performance model – MIDAS. Proceedings of the Sixth International Conference on Cognitive Modelling, July 29 – August 1, 2004, Pittsburgh, PA. Mahwah (NJ): Lawrence Erlbaum; 2004:396–399.
22. Wickens CD. Scanning strategy. In: Hancock PA, editor. Human performance and ergonomics: perceptual and cognitive principles. Cambridge (MA): Academic Press; 1999:209–211.
23. Wickens CD, Goh J, Helleberg J, Horrey WJ, Talleur DA. Attentional models of multitask pilot performance using advanced display technology. Hum Factors. 2003; 45(3):360–380.
24. Wickens CD, McCarley J, Alexander A, Thomas L, Ambinder M, Zheng S. Attention-Situation Awareness (A-SA) Model of pilot error. Technical Report AHFD-04-15/NASA-04-5. Champaign (IL): University of Illinois, Aviation Human Factors Division; 2005.
25. Wickens CD, McCarley JS, Alexander AL, Thomas LC, Ambinder M, Zheng S. Attention-situation awareness (A-SA) model of pilot error. In: Foyle D, Hooey BL, editors. Human performance modelling in aviation. Boca Raton (FL): CRC Press; 2007:213–237.