

Validation of the Automation Attitude Questionnaire for airline pilots

Preven Naidoo

Leopold Vermeulen

Department of Human Resource Management
University of Pretoria
Pretoria

Preven.naidoo@hotmail.com

Abstract

This article reports on the construction of the Automation Attitude Questionnaire (AAQ), which assesses airline pilots' perceptions about operating advanced commercial aircraft. A total of 262 airline pilots from a large South African carrier participated in the validation of the instrument. A five-factor measurement model was established after applying several rounds of exploratory factor analysis. The five factors associated with pilots' perceptions of advanced flight deck automation were labelled understanding, training, trust, workload, and design. The Cronbach's alpha coefficients and the mean inter-item correlation of each factor were highly satisfactory and confirmed the homogeneity and unidimensionality of the five-factor solution for the AAQ.

Keywords: aviation, Automation Attitude Questionnaire, factor analysis, human factors, flight deck, psychometric.

1 Introduction

Advancements in computer technology, coupled with economic demands, have given rise to increased electronic flight deck designs. During the 1980s there was a rapid development of automated apparatus being incorporated into large commercial aircraft to improve efficiency, in an attempt to counter rising fuel costs (Ishibashi, Kanda and Ishida, 1999). There was thus a significant impact to both navigation and power-plant systems. This included developments in air data and inertial reference units (such as IRS), flight guidance systems, automatic control of throttle or thrust systems and strategic data management systems (FMS). Pilots also now solve onboard faults using more electronic assistance, for instance the crew-alerting systems such as EICAS (Engine Indicating and Crew Alerting System) in Boeing manufactured aircraft and ECAM (Electronic Centralised Aircraft Monitoring) in Airbus manufactured aircraft (Wiener, 1989; Sherman, 1997; Risukhin, 2011). In general terms, these examples of new technology have resulted in terminology such as 'the glass cockpit' (Wiener, 1988, p. 435), pointing to displays presented entirely by computer-generated graphics.

Over the years, research has clearly established without doubt, that the integration of advanced technology and more automation into the airline pilot's work environment (flight deck) has contributed to greater efficiency, productivity and overall safety (Wiener, 1993; Röttger, Bali and Manzey, 2009). However, the increased application of

automation in the flight deck has over the past two decades generated a strong debate *for or against* such over-reliance on technology and increase in automation (Sherman, Helmreich and Merritt, 1997; Risukhin, 2011). Subsequently, the media have clung onto this issue and produced a fair amount of disinformation amongst the general flying public on the subject. For some time now there have been concerns that as the level of automation based on technologically advanced systems increased, there was ‘a growing discomfort that the cockpit may be coming too automated’ (Wiener, 1989, p.1). Such concerns may stem from human factor issues such as poor interface design, pilot complacency and over-reliance on automation (not taking control when things go wrong), deteriorating manual handling skills and diminished situational awareness (Palmer, 1995; Billings, 1997; Parasuraman, & Riley, 1997; Wood, 2004). There is presently insufficient empirically based research to support any extreme view on the topic. Aircraft manufacturers are therefore focusing on the benefits which technology can offer, such as continually designing interactive graphic display systems so as to provide the pilot with an enhanced sense of situational awareness. Situational awareness is described as the perception of elements in the environment within time and space, the comprehension of their meaning, and the projection of their status in the near future (Endsley, 1995). However, providing pilots with *reality-on-a-plate*, can also produce unforeseen complications. For instance, recently Casner, Geven and Williams (2012) found that pilots’ situational awareness and situation control were in fact poor because pilots had little awareness of computer modes, specifically during periods of high workload. It was found that crews were easily startled when a system failure occurred or the computer did not behave as expected, resulting in fatally inappropriate control of aircraft trajectory.

2 Background

The present advanced flight deck incorporates flight data information on cathode ray tubes (CRTs) and liquid crystal displays (LCDs) – the main reason that many observers refer to these systems as ‘glass cockpits’ (Risukhin, 2011). As an example, the complete digitised flight deck system can consist of electronic attitude director indicators (EADIs); electronic horizontal situation indicators (EHSIs); data management systems (FMS) and symbol generators to drive the electronic indicators; navigation system control and display units (ND); and laser gyroscopic air data inertial reference systems. Various crew alerting systems are further incorporated on the modern flight deck to support pilots in operating aircraft more safely and precisely in today’s congested airspace. This includes technology such as Traffic Collision and Avoidance System (TCAS) and Controlled Flight into Terrain (CFIT) avoidance equipment, for example the Enhanced Ground Proximity Warning System (EGPWS), which incorporates global positioning (GPS).

Although flight deck automation has been well received by the aviation industry, a number of important ergonomic and human factors related issues have been raised (Lyall & Funk, 1998). Research suggests that the increased presence of computers, such as flight management computers (FMCs) can result in crew members spending an increasing amount of ‘heads-down’ time during critical phases of a flight, a key contributor to distractions and reduced situational awareness (Damos, John and Lyall, 2005). The general design of systems onboard the analogue (old generation) flight deck

allowed pilots to make a number of small errors, which could easily be ameliorated through manual intervention (Laudeman and Palmer, 1992; Röttger, Bali and Manzey, 2009). In contrast, human errors traditionally considered small, such as inverting a number, have resulted in catastrophe in the advanced (digital) aircraft (Edwards, 1988; Casner, Geven & Williams, 2012). Any human based mistakes committed on a modern flight deck are therefore more likely to result in a serious incident or accident due to a compounding input-output effect. For example, the use of reduced thrust take-offs has become an everyday method of reducing wear and tear on high-bypass turbofan engines. However, an error in the pilot's input of the pseudo-temperature into the flight management computer may result in disaster as the aircraft will fail to establish the necessary speed (output) during the take-off phase (for instance, when the input assumed temperature is far higher than the stipulated calculations). This highlights the fallibility of the basic computer-human dyad. Damos *et al.* (2005) refer to this erroneous input-output process as GIGO (garbage-in-garbage-out). A study conducted by the National Transportation and Safety Board (NTSB, 2009) found that human error is a significant contributor to accidents in complex transport systems. Given the potential issues generated by advanced technology, the present study was based on the premise that an understanding of airline pilots' perceptions would provide researchers with a deeper understanding of the root phenomena associated with these human factor issues.

2.1 Prior studies on human factor and aircraft automation issues

The term 'automation' has for some time been difficult to define, although a number of prominent scholars have agreed that the term '...generally means replacing human functioning with machine functioning,' whilst in flight deck terminology, for automation '...we generally mean that some tasks or portions of tasks performed by the human crew can be assigned, by the choice of the crew, to machinery' (Wiener, 1989, p. 121). Funk *et al.* (1999, p. 56) also indicated that 'Automation is the allocation of functions to machines that would otherwise be allocated to humans.' Evidence that the advanced flight deck and pilots' extensive use of automation creates a fertile ergonomic environment for confusion, communication-breakdown, confirmation bias and a number of flawed heuristics (as a short-cut to decision-making, a symptom of complacency), is well established (Rudisill, 1995; Mosier, Skitka, Heers & Burdick, 1998; Damos, John & Lyall, 2005; Casner, et al., 2012).

Because of the concerns of the potential negative effects of advanced automation and increasing technology on pilots' behaviour, the United Kingdom (UK) Civil Aviation Authority (CAA) conducted one of the first important studies in the field (McClumpha, James, Green and Belyavin, 1991). The study subsequently 'assess[ed] the effects of advance automation on UK pilots in order to identify possible problems' and explored the 'opinions and attitudes of UK pilots to advance flight deck automation' (McClumpha *et al.*, 1991, p. 3.2).

McClumpha and colleagues developed a questionnaire that included 78 items to explore pilots' opinions regarding aircraft automation. Ten of the items were related to a general attitude towards aircraft automation and 68 items addressed several ergonomic, human factor concerns and automation issues on advanced flight decks. This included design, reliability, flight management system input, output and feedback, skills, training, crew interaction, monitoring and procedures, workload, and overall impressions (McClumpha

et al., 1991). Additionally, the four-factor model (Table 1) developed by McClumpha *et al.* (1991), accounted for at least one third of the explained variance.

Table 1. UK pilots' perceptions of aircraft automation

Factor	Meaning
Understanding/Mastery	Comprehension, expertise, knowledge and use of the system
Workload	Workload, demand, stress and task efficiency
Design	Ergonomic efficiency, design and displays
Skills	Handling skills, crew interaction, and self-confidence

Unfortunately the authors did not report on the validity and reliability of the four factors, neither did they provide a list of all the significant items that signify the rotated factors. Nonetheless, this study provided the initial impetus and general backdrop for the present study. However, the goal of the study being reported here was to assess the situation within the South African context.

A decade later, Singh, Deaton and Parasuraman (2001) adapted a scale to assess pilot attitudes towards cockpit automation and furthermore to determine the reliability of such a measurement. Thirty items were selected from the original McClumpha *et al.* (1991) questionnaire. These items included the first 10 general attitude items and 20 items associated with human factors and various automated systems. This questionnaire was then administered to 170 pilots at Embry-Riddle Aeronautical University. One hundred and sixty three pilots with experience in advanced automation participated, of whom 111 completed the questionnaire satisfactorily. After computation, it was established that six factors accounted for 58% of the explained variance. These 6 factors were named: workload, design, skills, feedback, reliability, and self-confidence. The Cronbach co-efficient reliabilities of the six factors ranged from between .75 and .98. Although satisfactory reliability coefficients were reported, closer examination revealed the inclusion of three items that cross-loaded on more than one factor, which made the extracted components questionable. If these three items were omitted from the last two factors, both the factors of 'reliability' and 'self-confidence' would not have been included in the underlying factor solution of the questionnaire. According to Tabachnick and Fidell (2007, p. 646) the interpretation of factors defined by only one or two variables is 'not feasible'. Therefore, the underlying structure defined in the aforementioned study is unstable for practical use.

The discussion thus far clearly points out that further scientific research is needed to fully understand the human-automation interaction in today's advanced aircraft. Studies in this vein (Sherman, 1997) will always remain important because identifying the core human factor concerns and technology issues in the aviation psychology genre will enhance safety, and more importantly, save lives. The present study is an effort to aid in the valid and reliable identification and description of specific areas of pilots concerns regarding their performance in a highly advanced automated environment and their opinions about advanced flight deck systems.

3 Research design

3.1 Research Approach

In order to achieve the study objective a quantitative research approach based on the positivist paradigm was followed. A survey was conducted, using a structured questionnaire to collect the research data from a sample of airline pilots. The data was analysed in accordance with an associational design (Field, 2005). An associational design was deemed appropriate as the researchers wished to establish the correlation between items' scores on the instrument and to identify the underlying dimensions or factor structure of a hypothesised construct. The design furthermore aids in computing the internal consistency of extracted factors. The research design for the study as reported in this article was scrutinised by the Research Ethics Committee of the Faculty of Economic and Management Sciences of the University of Pretoria and was approved by this Committee.

3.2 Research Method

3.2.1 Participants

The research group represented a sample of 262 airline pilots current on advanced type aircraft. Demographic information (as independent variables) was elicited from all the participants in the first section of the questionnaire. These characteristics of the participants are summarised in Table 2.

Of this group, it was noteworthy that 245 were male pilots and 17 were female pilots. The small proportion (6.5%) of female participants was due to the fact that females have only recently started choosing professional flying as a career option. These numbers nonetheless do indeed reflect the current low proportional representation of female pilots (6.1%) engaged overall in commercial aviation in South Africa (SACAA, 2012).

Furthermore, the sample ranged from lower entry pilots (in-flight relief crew) to high-level pilots (for example, senior training captains on long-range fleets). The sample also represented diversity in terms of the type of aircraft flown, pilots' age and level of experience. Thirty six percent of the respondents had flown Boeing type aircraft and 63% had flown Airbus type aircraft. The participants' ages ranged from 25 to 65 years (a spread of 40 years, mean = 44 years, SD = 9.6). The respondents' number of years of flying experience ranged from between 4 years and 46 years (mean = 24 years, SD = 10). The mean number of flying hours of the sample was 12 231 hours (SD = 5 636). In addition, the mean digital flight hours logged by the sample was 4 691 hours (SD = 2 530). The total digital flying time logged was expected to be significantly lower than the total flying time, as the carrier only began to operate modern automated aircraft in the last ten years. Unlike in the United States where it is compulsory for airline pilots to hold some university qualification, only 25% of the respondents in this sample had any tertiary education at this level.

Table 2. Sample characteristics (N=262)

Variable	Frequency	Percentage
Gender		
Male	245	93.5
Female	17	6.5
Position		
Dedicated in-flight relief pilot	16	6.1
Co-pilot (Short Range)	60	22.9
Co-pilot (Long Range)	49	18.7
Captain (Short Range)	48	18.3
Captain (Long Range)	53	20.2
Training Captain (Short Range)	11	4.2
Training Captain (Long Range)	18	6.9
Other	5	1.9
Age		
25 – 35 years	59	22.5
36 – 45 years	88	33.6
46 – 55 years	67	25.6
56 – 65 years	48	18.3
Level of education		
High school	163	62.5
Diploma	33	12.6
Bachelor's degree	40	15.3
Post Graduate	25	9.6
Initial flying training		
Military	131	50.0
Cadet	21	8.0
Self (Part-Time)	72	27.5
Self (Full Time)	37	14.1
Total digital flying time logged		
0 to 2 000 hours	33	12.6
2 001 to 3 000 hours	53	20.2
3 001 to 4 000 hours	46	17.6
4 001 to 5 000 hours	48	18.3
5 001 to 6 000 hours	20	7.6
6 001 hours>	60	22.9
	2	0.8
Total flying time logged		
1 500 to 7 900 hours	65	24.8
7 901 to 11 200 hours	69	26.3
11 201 to 16 000 hours	56	21.4
16 001 to 27 000 hours	69	26.3

3.3 Measuring Instrument

To identify the core human factor issues related to flight deck automation and furthermore to assess airline pilots' perceptions of the phenomena involved, a measuring instrument entitled the Automation Attitude Questionnaire (AAQ) was constructed. Various research outputs in the field of flight deck automation were considered as points of departure in constructing the AAQ. Original items for the AAQ were generated by analysing seminal research undertaken by Wiener (1989), studies conducted by Funk and Lyall (2000), McClumpha, *et al.* (1991), and Sherman, *et al.* (1997). A general discussion and the findings of these background studies, which form the basis of the empirical work undertaken in the present research, were presented in section 2.1.

The original item pool of the initial AAQ included 85 items. Thirty-three of these items were reconstructed from the 78 items of the early attitude survey developed by McClumpha, *et al.* (1991). Secondly, a further 35 items were extracted from a literature survey (Wiener, 1989; Helmreich, Klinect and Wilhelm, 1999; Funk and Lyall, 2000) and adjusted to ensure clarity and relevance in the context of the airline pilots who participated in the current study. Thirdly, an additional 17 new items were generated based on consultation with subject matter experts (which included a number of experienced airline training captains, academics, and airline managers). These 17 items were included so as to provide a good coverage of the hypothesised research construct. Each of the 85 items of the initial AAQ presented one statement that covered various domains that encompass automation training, such as flying skills, workload management, ergonomics (for instance, comfort design, functional design, user-friendliness of the system), and automation performance (refer to Table 3 for examples of statements). All the statements (except for the demographic variables) were rated on a seven-point Likert-type measure to assess the perceptions of respondents at an approximate interval level. Unfavourable statements were scored on a scale ranging from strongly agree (1) to strongly disagree (7). The favourable statements were reverse coded to produce a measure where high scores indicated positive perceptions and low scores resulted in a more negative perception of flight deck automation.

In its final form, the preliminary AAQ consisted of three sections. Section 1 related to the participants' demographic information. Section 2 consisted of the 85 items related to pilots' perceptions, opinions and behaviour regarding automation systems found onboard the flight deck. This second section attempted to determine the core human factors, ergonomic issues and operator concerns related to flight deck automation. An additional part, Section 3, was added to gain qualitative input from respondents. Participants were given the opportunity to comment, either positively or negatively on operating highly advanced automated aircraft in general. Due to space constraints in this paper, a full description of the instrument can be obtained directly from the researchers via email.

3.4 Research Procedure

A list of all the airline pilots employed at a large South African carrier was obtained from the organisation's human resources department. Permission was granted from the executive and chief pilot of the particular company to distribute the questionnaires to the entire pilot population in their employment. A total of 800 questionnaires were distributed on an individual basis via a physical box-drop method.

In order to maximise the response rate, a covering letter with the endorsement from management was attached to each questionnaire. The covering letter also stated the purpose of the research and further stressed voluntary participation and anonymity. Eliminating the need to provide a name on the questionnaire ensured participant anonymity. The completed questionnaires were collected manually from a dedicated collection box. A total of 262 (33%) questionnaires which yielded viable data was received. According to Tabachnick and Fidell (2007), this number of responses is adequate for an exploratory factor analysis.

3.5 Statistical Analysis

Exploratory factor analysis (EFA) was used to explore the internal structure and validity of the AAQ. EFA was carried out by means of principal axis factoring and rotated using the promax procedure with Kaiser's normalisation to obtain an obliquely generated factor solution for the AAQ. Based on the theoretical entities of the aforementioned discussion, it was deemed feasible that an oblique extraction methodology be followed because it was believed that the calculated factors were systematically correlated. Thus the eigenvectors (factors) were rotated in an attempt to achieve a simple structure (Gorsuch, 1983). To assess compliance with the distribution requirements, Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy were used. In order to determine the number of significant item factors, a triangulation method, namely Kaiser's criterion, Horn's parallel analysis and Catell's scree-plot were used (Tabachnick & Fidell, 2007). According to Hayton, Allen and Scarpetto (2004), a parallel analysis provides the most accurate estimate of the number of true factors in a complex dataset.

The internal consistency of the AAQ was assessed by calculating the Cronbach Alpha coefficient for each factor. Item-reliability indices of the individual items were calculated to establish whether the items contributed to the underlying construct of the factors (Gregory, 2004), and the average correlations between the items of each factor were calculated to examine the homogeneity and unidimensionality of the retained factors (Cortina, 1993; Clark & Watson, 1995). Frequencies and distributive statistics were used to describe the characteristics of the sample and to analyse the distribution (mean, standard deviations, skewness and kurtosis) of the responses.

4 Results

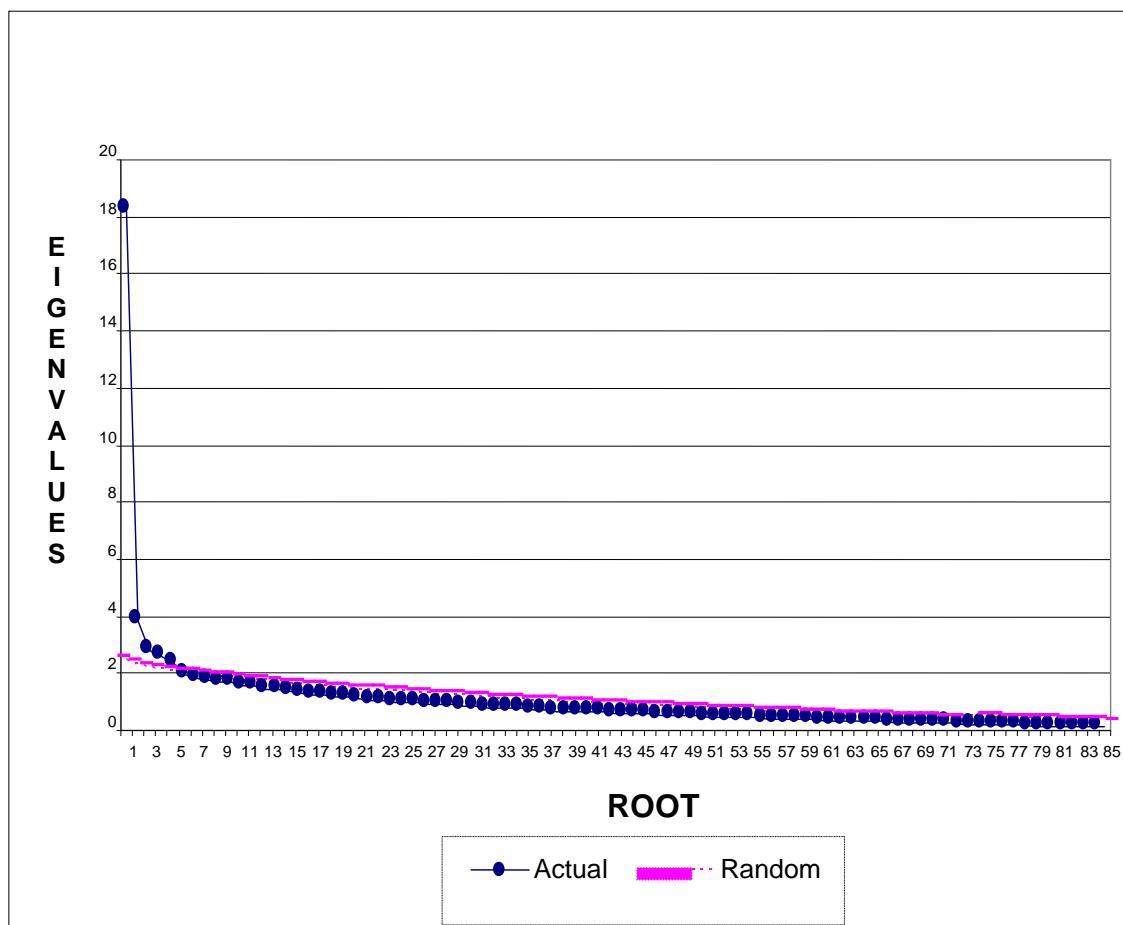
4.1 Exploratory Factor Analysis

The Exploratory Factor Analysis (EFA) was carried out by means of principal axis factoring, and was rotated using the promax procedure ($\kappa = 4$) with Kaiser's normalisation to an oblique solution (for reasons discussed earlier). This allowed the researchers to seek the lowest number of factors that account for the common variance in the set of 85 variables.

In the first round of EFA, the 85 items of the AAQ were inter-correlated and rotated to form a simple structure by means of the promax rotation. Owing to the size (85 X 85), the inter-correlation matrix is not reported in this paper. Based on Kaiser's (1961) criterion (eigenvalues larger than unity) 25 factors were postulated. The 25 factors explained 68% of the variance in the factor space of data. The factor analysis yielded more factors in the real test space than was expected. It was surmised that this is probably due to the presence of differentially skew items (Schepers, 2004). However, the results of Horn's parallel analysis and the scree-plot (Figure 1) confirmed that there were five significant constructs in the dataset. Parallel analysis indicated a break in the scree-plot between roots six and five. However, the curve of the eigenvalues of the random data set (the broken line) intersected the curve of the eigenvalues for the real data (the solid line) at root six. To avoid under-factoring, it was decided to include all the items of the six factors in the second round of EFA.

The items included in the six factors were first scrutinised; and the items that had factor loadings lower than .35 were omitted. A total of 33 items were retained and were subjected to the second round of EFA with promax rotation. The Kaiser-Meyer-Olkin (KMO) test for measuring sampling adequacy and Bartlett's test of sphericity displayed satisfactory results. Both diagnostic tests confirmed that the data was suitable for factor analysis. The calculated KMO value of .902 was greater than .7. Bartlett's test of sphericity [$\chi^2 (528) = 3470.758, p < .01$] confirmed that the properties of the correlation matrix of the item scores were suitable for factor analysis.

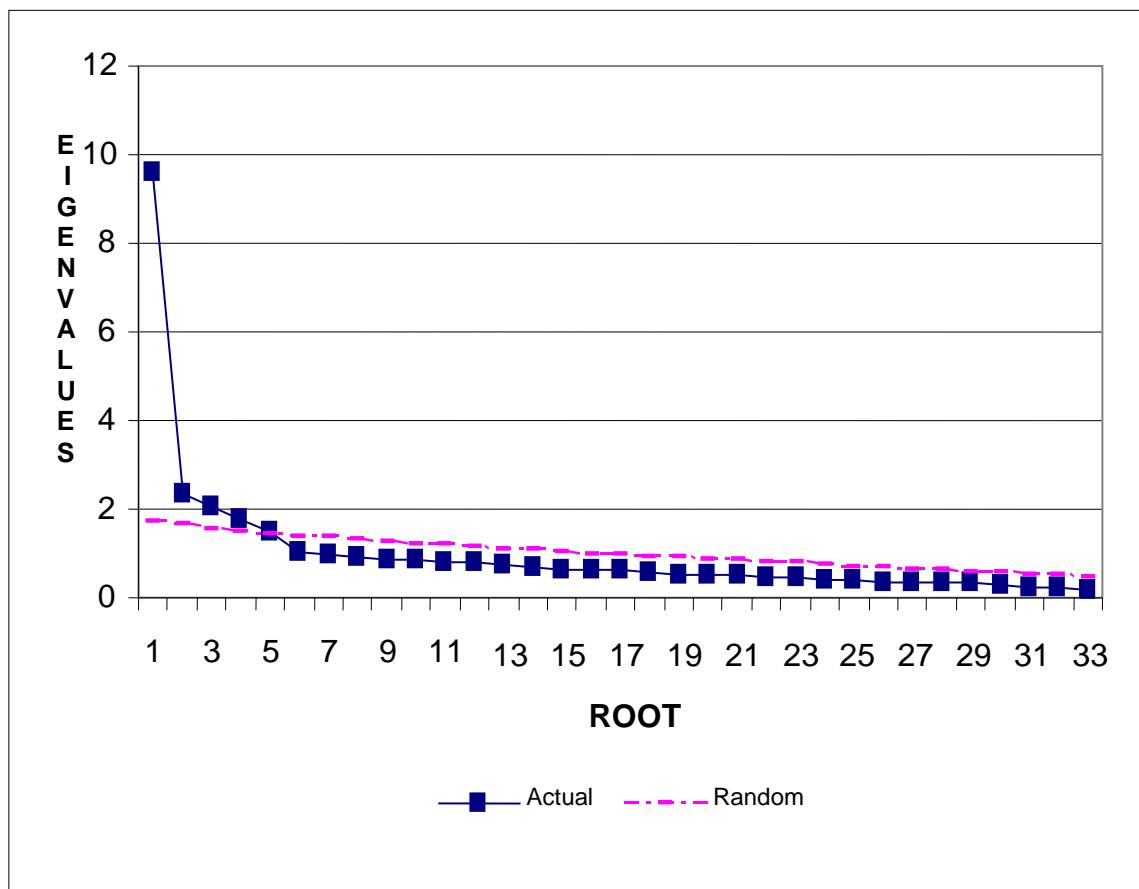
Figure 1. Scree plot of actual and random data of 85 factors



Six factors with eigenvalues greater than one were extracted in the second round of EFA. The six factors explained 55% of the total variance in the data. However, based on an inspection of the results of the parallel analysis presented in Figure 2, a five factor solution clearly seemed more appropriate. Only one noteworthy item with a loading of .369 was associated with Factor Six. According to Tabachnick and Fidell (2007, p. 646), the interpretation of factors defined by only one or two variables is 'risky', under even the most exploratory of factor analyses. Consequently, Factor Six was disregarded. This resulted in a 32-item pool that measures (reliably and validly) five factors related to perceptions of flight deck automation. Of the 85 items included in the preliminary AAQ, 13 of the 33 original items and 14 of the 35 adjusted items from the McClumpha *et al.* (1991) survey, and five of the 17 new items were retained. The five factors that

associated with the core issues or demands of operating an automated flight deck or 'glass cockpit' were labeled Understanding, Training, Trust, Workload and Design.

Figure 2. Scree plot of actual and random data of 33 factors



Understanding consisted of 8 items and included issues such as how a pilot interprets and understands the capabilities, limitations, modes, and operating principles and functioning of the automated flight deck system. This factor included pilot's competence to interpret the flight mode annunciator (FMA) and manage automation 'surprises' (Parasuraman & Riley, 1997).

Training, the second dimension, was made up of seven items that referred to the training and learning required to get a pilot to an adequate standard or to the level needed to operate the automation system. The elements of this factor refer to quality time spent in classroom training, on simulator training, recurrent training, route training, line training and transition training on advanced aircraft.

The third factor was labelled *Trust* and included six items that dealt with the level of belief and assurance a pilot has in the performance of automated devices. It measured pilots' identification with the automation system; feelings of increased exposure to risk and stress due to automation, feelings that the aircraft is ahead of him or her; and being detached from the human-machine loop. A specific item of this factor also refers to the impedance in crew co-ordination due to system trust issues.

The fourth factor looked at perceptions of *Workload* and included six items. The primary issues covered in this factor were increases in workload during critical phases of flight. Elements of the workload factor consisted of the amount of time spent instructing the automation computer via the flight management system (heads-down time) and thereafter having it accomplish a specific task correctly. Other elements also included the procedures required for safely operating the aircraft and the ability to maintain adequate situational awareness.

The fifth factor consisted of five items related to the *Design* characteristics and reliability of automation systems. This included the ergonomic features (control systems, flight deck layout, colour coding, position of controls and displays) and general design of the flight deck. Elements of the display design included the adequate presentation of accessible, useful, understandable and diagnostic visual and sound information, as well as the ease in utilising the information.

The factor loadings and corrected item-total correlation of the items in each of the five factors of the AAQ are summarised in Table 3. The corrected item-total correlation of each item in the five factors was satisfactory (DeVellis, 2003; Field, 2005). In retaining items within a specific behavioural scale, the following argument was taken into consideration:

DeVellis (2003) views an item with an item-total correlation of more than .20 as generally acceptable to be included. Field (2005) however, posits that if an item-total correlation is less than .3, that a particular item should not be included as a variable in a scale. The values of the corrected item-total correlation in the five factors were all above .3. In addition, the percentage variance, sums of squared loadings, squared multiple correlations and factor correlations are reported in Table 4.

Table 3. Factor loadings and corrected item-total correlation

Factor and Relevant Items	Factor Loading	Corrected Item-total Correlation
<u>Factor 1</u>		
Q38. I'm often confused about why the aircraft's automatics respond in the way it does.	0.831	0.724
Q36. I am often surprised by the aircraft's response to my FMS inputs.	0.816	0.722
Q40. I often tend to question the output from the automation system.	0.624	0.638
Q41. I find myself trying to guess what this aircraft is going to do next.	0.610	0.610
Q23. In the event of a partial system failure, it is never obvious which part of the automatic system failed.	0.567	0.475
Q37. I feel that the amount of feedback I get from the automatics is excessive.	0.557	0.557
Q42. The feedback I get in response to my inputs is usually too slow.	0.546	0.512
Q39. Even after receiving adequate feedback from the system, I still won't correct my fault.	0.433	0.519
<u>Factor 2</u>		
Q56. I think that there should be more simulator training for the conversion onto this aircraft.	0.831	0.698
Q55. The computer based-training was insufficient for me to fully understand this aircraft.	0.694	0.556
Q57. I feel that a lot more hours can be devoted to route training on this aircraft.	0.641	0.577
Q54. I think that there should have been a lot more classroom training for the conversion onto this aircraft.	0.631	0.644
Q58. There is insufficient recurrent training on this aircraft.	0.589	0.544
Q59. The training I received was inappropriate to line operations.	0.444	0.498
Q60. My transition onto this aircraft was extremely difficult.	0.367	0.443
<u>Factor 3</u>		
Q78. I feel detached from the aircraft.	0.813	0.678
Q79. I feel exposed to risk by the automation.	0.745	0.683

Q77. The aircraft is always ahead of me.	0.671	0.642
Q80. Whenever I fly this aircraft, I feel a lot more stress than when I flew traditional aircraft.	0.605	0.594
Q75. The automation system greatly decreases my confidence as a pilot.	0.509	0.569
Q64. Automation impedes crew co-ordination.	0.495	0.651

Factor 4

Q73. The automation actually increases workload during critical phases of flight.	0.797	0.641
Q72. In the event of a flight plan change, the ‘heads-down’ time required is much more than in traditional flight decks.	0.733	0.591
Q69. I've noticed that there is much more ‘heads-down’ time in this cockpit.	0.575	0.465
Q71. It is very difficult for the crew to maintain a good look-out when flying this aircraft.	0.567	0.583
Q74. In general the overall workload on this flight deck has increased.	0.524	0.528
Q70. The procedures used to operate this aircraft don't suit it at all.	0.367	0.494

Factor 5

Q16. I find that the aircraft automatics are extremely unreliable.	0.646	0.540
Q13. The displays in my aircraft make very poor use of colour.	0.590	0.438
Q17. The level of reliability and redundancy of the automatics is insufficient to conduct extended range operations.	0.522	0.445
Q14. I'm extremely unhappy with the set-up of the displays in my aircraft.	0.500	0.424
Q21. If the automatics fail, most of the time I don't try to restore the system.	0.421	0.404

Table 4. Percentage variance, sums of squared loadings, squared multiple correlations and factor correlations

Factors	1	2	3	4	5
Eigenvalues	9.577	2.315	2.049	1.780	1.473
Percentage Variance	29.022	7.014	6.210	5.393	4.463
Sums of Squared Loadings (SSL)	7.071	5.453	6.797	5.450	3.872
Squared Multiple Correlation (SMC)	0.991	0.974	0.977	0.980	0.930

Factor inter-correlation matrix

Factors	1	2	3	4	5
1. Understanding	-	0.509	0.603	0.515	0.488
2. Training	0.509	-	0.501	0.419	0.250
3. Trust	0.603	0.501	-	0.569	0.412
4. Workload	0.515	0.419	0.569	-	0.351
5. Design	0.488	0.250	0.412	0.351	-

4.2 Factor Reliability

The reliability of the factors of the Automation Attitude Questionnaire (AAQ) was determined by making use of Cronbach's coefficient alpha (Field, 2005). The average mean correlations between the items of each factor were also calculated to examine the internal homogeneity and unidimensionality of the five factors (Cortina, 1993; Clark & Watson, 1995). The means, standard deviation, skewness and kurtosis, average mean correlations and Cronbach's Alphas for the five factors are provided in Table 5. According to Table 5, the Cronbach alpha coefficients for the five factors of the AAQ were satisfactory. Compared to the guideline for alpha $\geq .70$, recommended by Nunnally and Bernstein (1994), the alpha coefficient for the five factors yielded acceptable values ($F_1 = .844$; $F_2 = .817$; $F_3 = .845$; $F_4 = .786$; $F_5 = .700$). Furthermore, none of the items, if deleted, increases the internal consistency of a factor. All the mean inter-item correlations of the five factors were within the range of .15 to .50 as suggested by Clark and Watson (1995). The high mean inter-item correlations of .45 to .63 is probably the result of the specificity of the target constructs. According to Clark and Watson (1995),

a much higher average inter-item correlation can be expected when one is measuring a narrow construct. The scores on the five factors of the AAQ appear to satisfy the requirements of homogeneity and unidimensionality and can be considered to be representative of the specific factor that they are assessing.

Table 5. Descriptive statistics and reliability results

Descriptive Statistics	Under-standing	Training	Trust	Workload	Design
Scale Mean Value	46.267	38.821	37.798	32.034	31.053
SD	7.079	7.401	4.606	6.224	3.992
Skewness	-1.265	-0.675	-1.584	-0.728	-2.085
Sk error	0.150	0.150	0.150	0.150	0.150
Kurtosis	2.124	0.011	3.323	0.302	7.276
Ku error	0.300	0.300	0.300	0.300	0.300
r(Mean)	0.59	0.56	0.63	0.55	0.45
Alpha	0.84	0.82	0.85	0.79	0.70

5 Discussion

The implementation of advanced aircraft flight deck automation has for some time been the root of many human factor debates within the aviation fraternity. The introduction of highly computerised technology has presented airline organisations with interesting human resource and ergonomic challenges. These challenges need to be met so as to maintain an efficient and optimal operational front. The perceptions of airline pilots with regard to flight deck automation issues have not previously been researched in South Africa and one of the challenges facing airlines is to determine what impact these perceptions have on successfully training and converting competent pilots to new advanced aircraft from the traditional (analogue) *older* generation aircraft (Barnett, 2005; Laudeman and Palmer, 1992).

The objective of this study was therefore, to construct a valid and reliable instrument to measure the perceptions of airline pilots towards the core automation issues linked with operating advanced automated aircraft, to assess the psychometric properties of the measure and to refine the instrument. A questionnaire named the Automation Attitude Questionnaire or AAQ was constructed to survey airline pilots' perceptions regarding automation issues linked with operating advanced automated aircraft.

From the original item pool, two applications of exploratory factor analysis yielded a five-factor solution. The magnitudes of the factor scores of the items in each of the five factors were all larger than .35 with factor scores ranging from .36 to .83. The mean inter-item correlations ranged from .45 to .63 and the alpha coefficients from .73 to .85. These results therefore provide sufficient support to claim that the AAQ is a highly

reliable and valid measure of advanced aircraft automation issues. Furthermore supporting the psychometric adequacy of the AAQ.

The five factors that associated with the core issues or demands of operating an advanced automated flight deck or *glass cockpit* were labelled as understanding, training, trust, workload and design. It is noteworthy that the 5 factors derived in the present study are closely linked to a number of human factor and ergonomic issues that were raised in the introduction of this paper (Rudisill, 1995). The difference in the specific factors extracted from the present data, however, was shown to be valid and reliable statistically, with a fair ability for practical application. So as to maintain brevity, only the main issues discovered in the present research are mentioned here and comprise the following: poor interface design; pilots' lack of understanding of the automated equipment; breakdown in attention and knowledge due to system complexity; demands in mode awareness and so-called automation surprises (Funk *et al.*, 1999; Parasuraman, Molloy and Singh, 1993); uneven distribution of workload; over-trust in the ability of the computer (autopilot) and decreased vigilance; pilot complacency and over-reliance on automation; loss of situational awareness; reduction of manual flying skills and proficiency; communication and coordination demands; and the need for new approaches to training.

The fact that the five factors identified in the present study also correspond with the ten prominent automation issues originally identified by Funk and Lyall (1998); Funk *et al.* (1999) and Parasuraman, *et al.* (1993) is encouraging. After intensive evidence based research, using various sources and criteria, Funk *et al.*, (1999, p. 120) listed the following five automation issues as the most important concerns that require solutions: 'understanding of automation may be inadequate; behaviour of automation may not be apparent; pilots may be overconfident in automation; displays (visual and aural) may be poorly designed; and training may be inadequate'.

The results from this survey also indicate a strong similarity with the factors identified by McClumpha *et al.* (1991) and Singh *et al.* (2001). Understanding, mastery, workload, skills and design are common labels that consist of similar elements. Feedback, reliability and trust also appear to share common items. Overall the results indicate that common threads permeate pilots' perceptions of automated flight decks and these appear to be consistent over time. The results of this study resonate with the capability of the AAQ to measure and assess airline pilots' perceptions of the most prominent issues and concerns in operating advance automated aircraft.

However, there are some limitations to the current research. Firstly, while the findings obtained in this study indicated that the psychometric properties of the AAQ are statistically robust, further confirmatory studies are required to support the derived factor structure. Additional research by means of confirmatory factor analysis can be of great value to refine the AAQ, if a larger sample is utilised. The response rate of 33%, which is expected in studies of this nature, may have contributed to the differentially skew distributive properties of the sample frame. Secondly, many of the participants endorsed response options at the higher end of the Likert measure used for the AAQ, possibly in an attempt to please the researchers. Consequently the scores on all the factors were non-normally distributed. Although assumptions on distribution are not in force in factor analysis (Tabachnick & Fidell, 2007), this present situation may add

complexity to *post hoc* analyses and thus call for the employment of non-parametric statistical techniques.

5.1 Practical application

The results of the statistical analyses of the responses on the AAQ suggest that the instrument constructed in this research is sufficiently reliable and valid to capture the present sample of airline pilots' perceptions of flight deck automation. Consequently aviation human factor specialists and aviation psychologists can use the tool with confidence to gather valid and reliable data about automation perceptions held by airline pilots in South Africa. Understanding key concepts and fundamental issues associated with perceptions, attitudes and behaviour that exist within the sphere of advanced flight decks, has significant benefits for the aviation industry at large. A concise understanding of this topic will benefit airlines and other organisations to design and develop specifically targeted training material and to positively influence their pilots in accepting and working effectively with automation. However, elements that influence overall perceptions of automation may depend on the type of organisation, nature of flight training, flying experience, type of aircraft, computer literacy, and operational position. Further research should endeavour to identify those variables that may have an effect on the perceptions of airline pilots.

A final practical note based on the results of this research is that because of the nature and complexity of the modern advanced automated flight deck, it is recommended that the pilots of such aircraft exercise the internal locus of control principle. In other words, it is up to each and every individual to take positive control of their learning environment, to study and understand his or her aircraft voluntarily, without being prompted to do so, and on a regular basis. Such a rich enthusiasm for knowledge is the basic building block for safety and the overall competence of the advanced aircraft pilot. By taking responsibility for their own learning, individuals subsequently mitigate latent flaws in the structured organisational training system, aircraft design and operational environment.

Acknowledgement

This article is part of a research project on flight deck automation by Dr. Preven Naidoo (The University of Pretoria) and co-ordinated by Professor Dr. Leopold P. Vermeulen (The University of Pretoria).

References

- Barnett, J. S. (2005). Training people to use automation: Strategies and methods. *Journal of Systemics, Cybernetics and Informatics*, 3(5), 73-76.
- Billings, C.E. (1997). *Aviation automation*. Mahwah, NJ: Lawrence Erlbaum.
- Casner, S. M., Geven, R. W., & Williams, K. T. (2012). The effectiveness of airline pilot training for abnormal events. *Journal of the Human Factors and Ergonomics Society*, 20(3), 22-35.

- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104.
- Damos, D.L., John, R.S., & Lyall, A.E. (2005). The effect of level of automation on time spent looking out of the cockpit. *The International Journal of Aviation Psychology*, 9(3):303-314.
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. Thousand Oaks, CA: Sage.
- Edwards, E. (1988). Introductory overview. In E.L. Wiener & D.C. Nagel (Eds.), *Human factors in aviation*, (pp. 3-25). San Diego, CA: Academic Press.
- Endsley, M. R. (1995). Towards a theory of situation awareness in dynamic systems. *Human Factors*, 37(1): 32-64.
- Field, A. (2005). *Discovering statistics using SPSS*. (2nd ed.). London: Sage.
- Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Suroteghuh, C., & Owen, G. (1999). Flight deck automation issues. *The International Journal of Aviation Psychology*, 9(2), 109-123.
- Funk, K., & Lyall, B. (1999). The evidence for flight deck automation issues. *Proceedings of the Tenth International Aviation Psychology Symposium Conference (CD-R)*. Columbus, OH: The Ohio State University.
- Funk, K., & Lyall, B. (2000). A comparative analysis of flight decks with varying levels of automation. *Final Report prepared for the FAA Chief Scientific and Technical Advisor for Human Factors*, (pp. 1-17). Washington DC: Federal Aviation Administration.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gregory, R. J. (2004). *Psychological testing: history, principles, and applications*, (4th ed.). Boston. Pearson Education Group.
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 7(2), 191-205.
- Helmreich, R. L., Klinec, J. R., & Wilhelm, J. A. (1999). Models of threat, error, and CRM in flight operations. In *Proceedings of the Tenth International Symposium on Aviation Psychology* (pp. 677-682). Columbus, OH: The Ohio State University.

- Ishibashi, A., Kanda, N., & Ishida, T. (1999). Analysis of aircraft accidents by means of variation tree. *Proceedings of the Tenth International Aviation Psychology Symposium Conference (CD-R)*. Columbus, OH: The Ohio State University.
- Laudeman, I. V., & Palmer, E. A. (1992). Measurement of automation effects on aircrew workload. In *The Third Annual ASIA Program Investigator's Meeting*. Moffett Field, CA: NASA-Ames Research Center.
- Lyall, B., & Funk, K. (1998). Flight deck automation issues. In M.W. Scerbo & M. Mouloua (Eds.), *Proceedings of the Third Conference on Automation Technology and Human Performance*, (pp. 288-292). Norfolk, VA, March 25-28, Mahwah, NJ: Lawrence Erlbaum Associates.
- McClumpha, A. J., James, M., Green, R. G., & Belyavin, A. J. (1991). Pilots' attitudes to cockpit automation. In *Proceedings of the Human Factors Society 35th Annual Meeting*, (pp. 107–111). Santa Monica, CA: Human Factors and Ergonomics Society.
- Mosier, K.L., Skitka, L.J., Heers, S., & Burdick, M. (1998). Automation bias: Decision making and performance in high tech cockpits. *The International Journal of Aviation Psychology*, 8(1):47-63.
- NTSB, National Transportation Safety Board. 2009. *Accident and Incident Report for Part 121 Operators*. Retrieved March 17, 2009, from <http://www.ntsb.gov/ntsb/AVIATION/>
- Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory*. (3rd ed.). New York: McGraw-Hill.
- Palmer, E. (1995). Oops, "It didn't arm." A case study of two automation surprises. *Proceedings of the Eighth International Symposium on Aviation Psychology*, (pp. 227-232). Columbus, OH: The Ohio State University.
- Parasuraman, R., Molloy, R., & Singh, I. (1993). Performance consequences of automation induced complacency. *The International Journal of Aviation Psychology*, 3(1), 1-23.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *The International Journal of Aviation Psychology*, 39(2):230-253.
- Risukhin, V. (2011). *Controlling pilot error. Automation*. New York: McGraw-Hill.
- Röttger, S., Bali, K., & Manzey, D. (2009). Impact of automated decision aids on performance, operator behaviour and workload in a simulated supervisory control task. *Ergonomics*, 52(5), 512–523.
- Rudisill, M. (1995). Line pilots' attitudes about and experience with flight deck automation: Results of an international survey and proposed guidelines. *Proceedings of the Eighth International Symposium on Aviation Psychology*, (pp. 288-293), Columbus, OH: The Ohio State University.

SACAA. (2012). *South African licensed pilots' information. [Statistics per license by gender, XLS]*. Licensing office of the South African Civil Aviation Authority (SACAA), July 27, 2007.

Schepers, J. M. (2004). Overcoming the effects of differential skewness of test items in scale construction. *SA Journal of Industrial Psychology*, 30(4), 27-43.

Singh, I., Deaton, J., & Parasuraman, R. (2001). Development of a scale to assess pilot attitudes towards cockpit automation. *Journal of the Indian Academy of Applied Psychology*, 27(1-2), 205-211.

Sherman, P.J. (1997). *Aircrews' evaluations of flight deck automation training and use: Measuring and ameliorating threats to safety1*. Technical Report 97-22. FAA Grant 92-G-017.

Sherman, P.J., Helmreich, R. L., & Merritt, A.C. (1997). National culture and flight deck automation: Results of a multi-nation survey. *The International Journal of Aviation Psychology*, 7(4), 311-329.

Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics*, (5th edn.). Boston, MA: Allyn & Bacon.

Wiener, E. L. (1988). Cockpit automation. In E.L. Wiener & D.C. Nagel (Eds.), *Human factors in aviation*, (pp. 433-461). San Diego, CA: Academic Press.

Wiener, E. L. (1989). *Human factors of advanced technology ("Glass Cockpit") transport aircraft*. NASA Contractor Report 177528, Moffett Field, CA, USA.

Wiener, E. L. (1993). *Crew coordination and training in the advanced cockpit. Cockpit resource management*. San Diego, CA: Academic Press.

Wood, S. (2004). *Flight crew reliance on automation*. CAA Paper 2004/10. Research Management Department, Safety Regulation Group, Civil Aviation Authority, UK, Gatwick Airport South, West Sussex.