2007

Filling in the Gaps: An Investigation of the Knowledge Needed for Effective Human-Automation Interaction

Michael T. Curtis

Michelle Harper-Sciarini

Florian G. Jentsch

David Schuster

Ron Swanson

FILLING IN THE GAPS: AN INVESTIGATION OF THE KNOWLEDGE NEEDED FOR EFFECTIVE HUMAN-AUTOMATION INTERACTION

Michael T. Curtis
Michelle Harper-Sciarini
Florian G. Jentsch
David Schuster
Ron Swanson
University of Central Florida
Orlando, FL

Despite the decade-long focus on pilot understanding of automated systems, knowledge gaps continue to suggest that current training programs are incomplete. There are several different types of knowledge that contribute to an effective understanding of automated systems, including more complex knowledge structures. This suggests that simply training pilots how to operate automated systems will not provide optimal performance. This paper describes an empirical investigation of the importance and measurement of different types of knowledge important for effective automation use. Twenty trained commercial pilots completed several knowledge assessments of automation and airmanship and responded to situational vignettes that assessed their proficiency in automation use. Results indicated that, in addition to paper-and-pencil measures of basic airmanship, measures of structural knowledge of automation (in the form of concept maps) accounted for large and significant amounts of variance in the pilots' proficiency of automation use.

Introduction

Gaps in pilots' knowledge of automated aircraft have been cited as contributing to past accidents and incidents in advanced aircrafts (Beringer & Harris, 1997; Billings, 1997; Bowers, Jentsch, & Salas 1995; Feary, Alkin, Palmer, Sherry, McCrobie, & Polson, 1997; Sarter & Wood, 1992; Sarter & Woods, 1994). Also, researchers have found that pilots have misconceptions about the flight deck automation they interact with (Sarter & Woods, 1992; Sarter & Woods, 1994). These findings may be attributed to procedurally-focused pilot training that does emphasize the "how" of automation use, as opposed to a more conceptually driven training that facilitates a deeper understanding of the automation; an understanding that allows pilots to make quick and precise decisions.

Automation is now commonplace in complex systems. In many applications, automated technology is introduced for the purpose of reducing operator workload and increasing efficiency. With the addition of automation, actively controlling the system has been changed to passively monitoring the system. This in and of itself brings about many challenges. For example, pilots often experience problems with tracking the status and behavior of the automation in glass cockpits (Sarter et al., 2003). Consequentially, pilots lose track of the current state of the aircraft and/or cannot predict the future states or the behaviors of the automation (e.g., loss of mode awareness) (Mumaw, Sarter, & Wickens, 2001).

One factor that has consistently been identified as contributing to loss of mode awareness is the existence of gaps and misconceptions in pilots' understandings of flight deck automation (Mumaw et al, 2001; Sarter & Woods, 1997, 2000; Sarter et al, 2003, Wickens, 2007). As mentioned previously, these gaps and misconceptions may in part be attributed to training programs that focus too much on *how* to operate the system as opposed to facilitating a deeper understanding of the "*why*" of automation; specifically, an understanding of the structural relationships among automation components that should promote mode awareness.

How can a deeper understanding of flight deck automation be promoted? To answer this question, we have been investigating how the operation of complex systems is influenced by an operator's cognitive structures. This paper specifically addresses how the accuracy of a pilot's knowledge structure influences his/her understanding of how to operate flight deck automation. In this study, twenty trained pilots completed a set of knowledge measures that would allow us to compare how different types of knowledge contribute to pilot performance. The results of this study will contribute to guiding the future development of both diagnostic tools and training methods that will more accurately address gaps in pilot understanding of automation.

Knowledge Structures

One way to conceptualize how knowledge is gathered and stored in memory is through the use of relational networks (Deese, 1965), also referred to as structural representations (e.g. Bower, 1972; Collins and Quillian, 1969; Johnson & O'Reilly, 1964; Shavelson, 1972). Knowledge structures are distinct from other types of knowledge (e.g., declarative or procedural knowledge) in that they not only contain information about the "what" and "how" of a particular domain, but – by showing relationships – also provide information about the inter-dependencies and conditional relations (i.e., the "why") among the parts of the domain.

Studies have shown that highly integrated knowledge structures of a domain will contribute to effective performance (Cooke, 1999; Cooke & Schvaneveldt, 1988; Klein, 1993; Orasanu, Martin, & Davidson, 2001; Stout, Salas, & Kraiger, 1997). Further, Day, Arthur and Gettman (2001) suggested that accurate knowledge structures are essential for skill-based performance. Based on these findings, we hypothesized that structural knowledge would provide unique information about a pilot's knowledge of automation and would uniquely contribute to the prediction of pilot performance; that is, we expected that the measurement of knowledge structures would give us information that would be above and beyond what other types of knowledge could contribute.

Method

Participants

Twenty active commercial pilots who were trained on the CRJ200 and Collins4200 Flight Management Systems (FMS) participated in this study. Participants had a median 2,200 hours of flight time in any aircraft (M = 2,809 hours), and between 25 and 2,500 hours of flight experience in the CRJ aircraft (median: 250 hours). Participants received a moderate amount ($25/hr) of monetary compensation for their time.

Measures

Dependent variable: Performance on a Situational Judgment Test. The dependent measure for this study was defined by the percent correct score on a paper-and-pencil situational judgment test developed at George Mason University (Smith, Prada, & Boehm Davis, 2006). The test presented situational vignettes that were based on selected flight scenarios which required proficiency in automation use and understanding. After studying the vignettes, pilots responded to 112 questions about specific courses of action in the situations described in the vignettes.

Predictor 1: Commercial Level Questionnaire. The commercial level questionnaire is a 22-item multiple-choice instrument to measure pilot knowledge of general airmanship, with items relating to aerodynamics, energy management, thrust-and-airspeed relationships, etc. Scores are based on percent correct.

Predictor 2: System Level Questionnaire. The system level questionnaire is a 19-item test of components and procedures specific to the target aircraft's automated flight deck system. Seven items were open-ended and required brief written responses from the respondents; the remaining items were in multiple-choice format. Scores are based on percent correct.

Predictor 3: Flight Management System (FMS) Level Questionnaire. The flight management system level questionnaire contains 15 open-ended response items asking about the specific function of the target aircraft's FMS system. Scores are based on percent correct.

Predictor 4 and 5: Structural Knowledge Scores. Two structural knowledge scores were derived from the scoring of two concept maps that the pilots completed as part of the study. In general, concept maps are created by making a web of connections between concepts available based on the participants perception of relationship between concepts. Scores are obtained by rating each connection made based on correctness, as compared to an expert representation or expert map. Correctness ratings for each connection could range from 0 to 4. Total rating scores were calculated by adding the ratings for all connections on each map. They represent a quality score of the overall concept map.

Predictor 4: CRJ Structural Knowledge Score. The CRJ concept map structural knowledge score was derived from a concept map that required pilots to map and describe the relationships between 15 terms relevant to CRJ the automated flight deck. Example items on this map included "AHRS," "autopilot," and "thrust levers."

Predictor 5: Training Structural Knowledge Score. The training concept map structural knowledge score was derived from scoring a concept map of 23 terms related to standard flight procedures in the CRJ aircraft. Example terms from this map included "ETA," "Crossing Restriction," and "Flight Path."

Other Materials and Apparatus

Demographic Form. The demographic form contained questions pertaining to participant's general flight experience and more specifically about their experience in the CRJ jet. In addition to the demographics form, the three declarative knowledge tests were administered via paper and pencil means.

Concept Mapping Software. The Team Performance Laboratory – Knowledge Assessment Tool Set (TPL-KATS) software designed at the University of Central Florida was used to administer the concept maps for participants. In the software, creating concept maps is accomplished by manipulating the position of target concepts ("cards") using the mouse and then connecting them in a web of connections based on perceived relationship between them, via a "draw arrow" function that requires use of the mouse and text entry via the keyboard (see Figure 1).

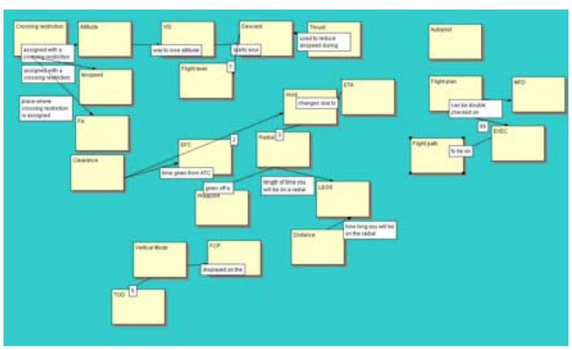

Figure 1. Sample concept map

Procedure

The study was broken into two sessions which were completed on two separate days. There were no time limits imposed on any of the experimental sessions. On day 1, participants first read and signed the informed consent form. Following this, the pilots were randomly assigned to one of two training conditions. The condition manipulation was part of a related study on effective training for automation and presented a short (15-minute) training session either with procedural or conceptual information about automation operations. The training condition manipulation is not further described here, but in all subsequent analyses, we first partialled out any training effects so that the manipulation would not influence the relationships between the remaining study variables. We also tested for possible interaction effects between condition and our target variables, but found none.

The condition manipulation was followed by the presentation of the general aviation questionnaire on a desktop computer. After completion of the questionnaire, participants were finished with the first session of the experiment and scheduled a time to participate in the second session.

The second session began with participants completing a demographics questionnaire. Upon its completion, participants were given the three paper-and-pencil tests describe above (i.e., Predictors 1 through 3) which they completed in approximately 30 to 45 minutes.

Participants were then given a short tutorial on how to create a concept map using the TPL-KATS knowledge elicitation software. Following completion of the maps, participants were debriefed and compensated for their participation.

Results

Table 1 provides information on means, standard deviations, and inter-correlations among the study variables. All analyses were performed using SPSS 12.0; the alpha level was set at .05, unless otherwise specified.

As part of data screening, a Cook's distance (Cook & Weisberg, 1982) analysis was performed which revealed an extreme outlier in the data. Further inspection of this case indicated that the pilot had the minimum number of hours in the CRJ in the entire data set and, indeed, substantially fewer hours than the next-least experienced pilot. This outlier was therefore removed from all subsequent analyses, which, thus, were based on a valid *N* of 19. Research variables were also tested for homogeneity and homoscedasticity. A log function transformation was performed on the training structural knowledge score to correct for positive skewness.

Data were analyzed using a standard multiple regression/correlation (MRC) analysis. However, because of the small *N* and the unfavorable *N*-to-k ratio that would have resulted had we entered all five predictors in addition to condition into an MRC, we first combined the three written tests (Predictors 1 through 3) into a Declarative Knowledge score, and the two concept map scores (Predictors 4 and 5) into one Structural Knowledge score.

Table 1. Study inter-correlations, means and standard deviations

Correlation	N	M	SD	1	2	3	4	5	6	7
1. Performance Measure	19	.76	.07	--	--					
2. Condition	19	.47	.51	.58**	--					
3. Commercial Level Questionnaire	19	.61	.08	.47*	.45	--				
4. System Level Questionnaire	19	.51	.10	.31	.17	.16	--			
5. FMS Level Questionnaire	19	.51	.08	.18	.35	.23	.22	--		
6. CRJ Structural Knowledge Score	19	57.16	14.19	.73**	.42	.28	.32	.37	--	
7. Training Structural Knowledge Score	19	4.44	.30	.39	-.06	.24	.10	.15	.43	--

Table 2. Results of the Multiple Regression Analysis for Performance on the Situational Judgment Test

	Unstandardized Coefficients		Standardized Coefficients		95% Confidence Interval	
Variable	*B*	*SE B*	β	*t*	*Lower*	*Upper*
Condition	0.06	0.03	0.45	2.63*	.01	.12
Declarative Knowledge	0.01	0.01	0.02	0.13	-.03	.03
Structural Knowledge	0.04	0.01	0.56	3.27*	.01	.07

Note: $N = 19$, * denotes $p < .05$

A standard multiple regression was then performed, regressing the dependent variable, (i.e., performance on the situational judgment test), on condition, declarative knowledge, and structural knowledge as the predictor variables. The model containing all three predictors accounted for a highly significant 57.4% of the variance in performance scores ($R^2 = .645$, *adj* $\underline{R}^2 = .645$), $F(3, 15) = 9.070$, $p < .001$ (see Table 2). In this model, both (training) condition ($sr^2 = .16$) and structural knowledge ($sr^2 = .25$), contributed significantly to the prediction of the performance score; the declarative knowledge score, however, did not ($sr^2 = .004$).

Discussion

The current research was intended to examine the types of knowledge that are associated with improved understanding of flight deck automation. Due to the fact that many of the pilot errors that are attributed to unexpected events in the cockpit are associated with automation, it is important to improve ways to identify what is at the root of pilot problems with automation. This also is important for the assessment of training needs and for the determination of training successes. Being able to more accurately show what pilots know and what they do not know will also help trainers to better focus their attention on the factors that might help reduce error.

By comparing pilot scores on measures of both declarative and structural knowledge, our aim was to identify how the different knowledge types contribute to the overall pilot performance (measured here as performance on a situational judgment test with vignettes describing common flight procedures). Previous findings supported the idea that deeper conceptual knowledge leads to improved performance on tasks, and that measures of structural knowledge should, therefore, be better indicators of pilot proficiency and better predictors of pilot performance. The aim of this study was to examine this phenomenon in a complex setting.

It is interesting to note that training condition and structural knowledge score both significantly predicted pilot performance, while the declarative knowledge did not. Although not described here in more detail, the training module in the conceptual training condition was intended to teach pilots the "why," instead of the "how" of autoflight operations. Thus, we expected that it would be related to higher scores on the performance measure. This finding supports the contention that improving structural knowledge can help develop pilot's performance with flight deck automation.

The most interesting finding from our research was, however, that the structural knowledge score significantly predicted performance beyond training condition, whereas declarative knowledge did not. This supports the notion that proficiency in the operation of complex systems is improved by knowing how the components of the system interact (i.e., structural knowledge), as compared to only knowing how to operate the components. As the results of this study indicated not only that the structural knowledge assessment technique accounted for significant variance beyond training, but also for a portion of performance that was not accounted for by previously used declarative knowledge assessments.

The implications of this finding suggest that there are beneficial applications to structural training implementations of which further pursuit is worth while. Although there is some question about the practicality of using and scoring conceptual knowledge assessments, such as concept mapping, as opposed to the paper and pencil tests that are available for most declarative knowledge assessments, currently available software programs such as the TPL-KATS (Harper et al., 2003) software provide a more convenient platform for assessing conceptual knowledge.

While the declarative knowledge scores did not have a significant impact on the predicted variance in this study, its value in a training setting should not be neglected. It is logical to deduce that a combination of all types of knowledge will provide a stronger knowledge base than any one type of knowledge in isolation.

Limitations

Although the study had an admittedly small *N* and used somewhat exploratory analyses, our findings provide a direction that suggests that structural knowledge assessments have a potential to capture "previously uncaptured" variance in pilot proficiency, and may even replace declarative knowledge tests. The idea that conceptual knowledge assessments provide a more complete assessment of a pilot's knowledge and subsequent performance shows promise bearing further examination.

Future Research Directions

Future research should focus on how to identify what specifically comprises these gaps in order to develop an assessment that can specifically identify conceptual knowledge deficiencies that will help trainers to better tailor training to individuals. It is also important to consider the fact that the observations of check airmen are the primary scores used to assess pilot performance. As a result of this, further examination of how conceptual knowledge assessment scores relate to these ratings will shed light on if conceptual knowledge is currently targeted in check airmen observations.

The findings from this study provide a fresh look at how structural knowledge contributes to the overall knowledge picture as applied to complex systems. This research provides a starting point for a shift in focus on the characteristics of training that might close the gaps in understanding.

References

Beringer, D.B. & Harris, H.C., Jr. (1997). Automation in General Aviation: Two studies of pilot responses to autopilot malfunctions. Washington, DC: FAA Office of Aviation Medicine, report number DOT/ FAA/AM-97/24.

Billings, C.E. (1997). *Aviation Automation: The Search for a Human-Centered Approach* Mahwah, NJ: Erlbaum, 1997.

Bower, G.H. (1972). Mental imagery and associative learning. In L. Gregg (Ed.), *Cognition in learning and memory*, 51-88.

Bowers, C., Jentsch, F., & Salas, E. (1995). Studying automation in the laLCan you? Should you? In N. McDonald, N. Johnston, & R. Fuller (Eds.), *Applications of psychology to the aviation system* (pp. 281-286). Cambridge, England: Avebury.

Collins, A., & Quillian, M. (1969). Retrieval from Semantic Memory. *Journal of Verbal Learning and Verbal Behavior* 8, pp. 240–247.

Cook, R.D. & Weisberg, S. (1982). Residuals and influence in regression. New York: Chapman & Hall.

Cooke, N. (1999). Knowledge elicitation. In F. T. Durso (Ed.), *Handbook of Applied Cognition*, pp. 479-509, U.K.: Wiley.

Cooke, N., & Schvaneveldt, R.: Effects of Computer Programming Experience on Network Representations of Abstract Programming Concepts. *International Journal of Man-Machine Studies*, vol. 29, no. 4, 1988, pp. 407-427.

Day, E.A., Arthur, W. & Gettman, D. (2001). Knowledge structures and the acquisition of a complex skill. *Journal of Applied Psychology,* 86(5), 1022-1033.

Deese, J. (1965). The structure of associations in language and thought. Baltimore: Johns Hopkins Press.

Donald, M. (1991). Origins of the modern mind. Cambridge, MA.: Harvard University Press.

Feary, M., Alkin, M., Palmer, E., Sherry, L., McCrobie, D., & Polson, P. (1997). Enhancing Vertical Navigation Performance in Glass-Cockpit Aircraft, *In Proceedings of the Ninth International Symposium on Aviation Psychology, Proceedings Vol. 1* (A98-28868 07-53). Columbus, OH: Ohio State University.

Harper, M.E., Jentsch F.G., Berry, D., Lau, H.C., Bowers C. & Salas, E. (2003). TPL-KATS-card sort: A tool for assessing structural knowledge. *Behavior Research Methods, Instruments, & Computers,* 35(4), 577-584.

Johnson, D.M. & O'Reilly, C.A. (1964). Concept attainment in children: Classifying and defining. *Journal of Educational Psychology,* 55(2), 71-74.

Klein, M. K. (1993). Issues in the assessment of personality disorders. *Journal of Personality Disorders* Suppl. (Spring), 18-33.

Mumaw, R., Nikolic, M., Sarter, N., & Wickens, C. (2001). A simulator study of pilots' monitoring strategies and performance on modern glass cockpit aircraft. *Human Factors and Ergonomics Society 45th Annual Meeting.* Minneapolis/St. Paul, MN: 8-12 Oct. 2001. pp. 73-77.

Mumaw, R., Sarter, N., & Wickens, C. (2001). Analysis of pilots' monitoring and performance on an automated flight deck. *Proceedings of the 11th International Symposium on Aviation Psychology,* Columbus, OH: The Ohio State University.

Orasanu J., Martin L., & Davidson J. (2001) Cognitive and contextual factors in aviation accidents. In : Klein G., Salas E., editors. Applications of naturalistic decision making. Hillsdale, NJ : Lawrence Erlbaum Associates in press.

Sarter, N., Wickens, C., Mumaw, R., Kimball, S., Marsh, R., Nikolic, M., & Xu, W. (2003). Modern flight deck automation: Pilots' mental model and monitoring patterns and performance. *To be Presented at the 12th International Symposium on Aviation Psychology.* Dayton, OH.

Sarter, N. & Woods, D. (1992). Pilot interaction with cockpit automation I: Operational experiences with the flight management system. *International Journal of Aviation Psychology.*

Sarter, N. & Woods D. (1994). Pilot interaction with cockpit automation II: An experimental study of pilots' mental model and awareness of the Flight Management System (FMS). *International Journal of Aviation Psychology.*

Sarter,N., & Woods, D. (1997). Team play with a powerful and independent agent: Operational experiences and automation surprises on the airbus A–320. *Human Factors,* 39, pp. 553–569.

Shavelson, R. (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology,* 63,3, pp. 225-34.

Smith, C., Prada, R., Boehm-Davis, D. (2006). Conceptual knowledge training. (unpublished work).

Stout, R., Salas, E., & Kraiger, K. (1997). The role of trainee knowledge structures in aviation team environments. The International Journal of Aviation Psychology, 7(3), 235-250.

Wickens, C. (2007) Cognitive factors in aviation. *Handbook of applied cognition* (Durso, F., Ed), 2nd Edition, pp. 361–389, Wiley, New York, NY.

EDUCATIONAL MINIMUMS AND STANDARDIZED INTELLIGENCE TESTING IN PILOT SCREENING AND SELECTION

Diane L. Damos
Damos Aviation Services, Inc.
Gurnee, IL

Pilot shortages have resulted in changes in pilot screening and selection at air carriers. One of the widespread changes in the screening process is the use of low or no minimum educational requirement for pilot applicants. An informal survey of major and national air carriers revealed that few carriers require a college degree and many have no minimum educational requirement. The elimination or reduction of educational standards may result in changes in aptitudes in the applicant population. Two air carriers that did not require a college degree administered standardized intelligence tests to their pilot applicants. These data were compared to intelligence tests scores obtained from a sample of pilot candidates in the United States Air Force (USAF). The data from the air carriers showed a lower average full-scale IQ, a larger standard deviation, and a lower range of scores than the USAF data. Potential causes for these differences and their implications for training are discussed.

Background

For at least the last 15 years, pilot shortages have been the focus of attention both by the government and the popular press (Fiorion, 2007; Hopkins, 2001; *Pilots and Aviation Maintenance Technicians for the Twenty-First Century*, 1993). The regional and national air carriers may have felt these shortages more acutely than the major air carriers for several reasons. One reason is that the regional and national air carriers have expanded their fleet while many of the major air carriers have decreased the size of their fleets. A second reason is that regional and national carriers have lower pay scales and benefits than the major carriers. They are, therefore, not as attractive to highly experienced pilots, such as ex-military pilots, as the major carriers (Lehman, 2003). Finally, the number of employers competing for the pool of qualified pilots has been increasing.

These factors, combined with recent case law, have resulted in changes in pilot screening and selection at air carriers. These changes have occurred gradually, attracting little notice. Their implications have not been completely understood. This paper focuses on two interrelated issues—the minimum educational requirements for pilot applicants and lack of use of standardized intelligence tests in pilot selection.

Minimum Educational Requirements

Federal Aviation Regulations (FAR) Part 61.123 and Part 61.153 have no educational requirements for either a commercial pilot certificate or an airline transport pilot certificate. To satisfy both of these FARs, the applicant must only “be able to speak, read, write, and understand the English language.” An air carrier may, however, set a minimum educational standard for its applicants.

Air carriers publish their minimum requirements for pilots on their websites and on the websites of companies that provide employment information to prospective applicants. The minimum requirements for pilot applicants published on one company offering employment information, fltops.com, was examined on February 14, 2007. Only information from carriers operating under Part 121 of the FARs and that were actively hiring pilots on that date were considered. If no information was posted about educational requirements, the website of the carrier was searched.

Seven major air carriers were actively hiring. Two carriers required a bachelor’s degree and two others preferred candidates with bachelor’s degrees. Eighteen national airlines with jet fleets were hiring. Of these, none required a bachelor’s degree. One carrier preferred a bachelor’s degree. Two carriers required an associate’s degree or the equivalent and one of these preferred a bachelor’s degree. Seven carriers listed no educational requirements and the remaining eight required a high school education.

Two reasons for the low minimum educational requirements may be given. First, a company may set high minimum requirements to reduce the number of applications that must be reviewed. During a pilot shortage, high minimum requirements may be counterproductive. Second, setting no or low minimum requirements may provide some protection in case of legal proceedings involving adverse impact against the carrier. Although this reason is speculative, few processes have been validated for setting minimum requirements and approved by a federal court (Buster, Roth, & Bobko, 2005), leaving an air carrier using higher minimum requirements vulnerable to legal actions.

Standardized Intelligence Tests

The effects of a reduction or elimination of educational requirements can be counteracted by the use of standardized intelligence tests as part of the selection process. Two clients of Damos Aviation Services, Inc. (DAS) agreed to administer standardized intelligence tests to their pilot applicants. One of the air carriers had never used any form of standardized testing in its selection process. The second had used some standardized testing previously but had eliminated the tests for economic reasons. Both of the carriers were regional airlines that historically had high selection ratios (a large proportion of the qualified applicants receive offers of employment). Both carriers required a high school degree for pilot applicants.

Data from one air carrier were collected in the summer and fall of 2005. Data from the other were collected in the spring and summer of 2003. Both air carriers administered the tests as part of the selection process, i.e. all of the test takers had passed the screening portion of the hiring process. The first airline administered the Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 2001). Scores on the Raven's were transformed into full-scale IQ scores using a conversion table provided in Raven et al. (2001). The second airline administered the Multidimensional Attributes Battery (MAB) (Jackson, 2003), which produces a full-scale IQ. The full scale IQ scores from both airlines (N= 82) were used to create the smoothed frequency distribution seen in Figure 1.

For comparison purposes, data from King and Flynn (1995) were used to create the second distribution shown in Figure 1. King and Flynns administered the MAB to 208 pilot candidates before they entered undergraduate pilot training (UPT). The participants were either graduates of the United States Air Force Academy, an Officer Training School, or the Reserve Officer Training Corps. Both the graduates from the Academy and the Officer Training Corp had completed a college education. The participants from the Reserve Officer Training Corps were either seniors in college or entering their senior year. Each frequency cell in the air carrier distribution was multiplied by 2.54 to compensate for the smaller number of cases in the sample.

Although the shapes of both distributions are similar and appear to be basically normal, the distributions have different characteristics. King and Flynn report a mean of 119.7 for the full-scale IQ with a standard deviation (SD) of 6.7 and a range of 96-141. In contrast, the mean for the air carrier full-scale IQ scores is 101.0 with a standard deviation of 10.4 and a range of 72-127.

The differences between these two distributions warrant comment. Probably the most striking difference between the two distributions is the difference in the means. In terms of the USAF distribution, the air carrier mean is 2.8 SDs lower. Indeed, with a mean of 101.1 and a range of 72 to 127, the air carrier distribution resembles that of an unselected U.S. population.

More problematic, perhaps, are the extremes of the distributions. Only 3 (3.6%) air carrier applicants scored at or above the mean of the USAF distribution. Twelve air carrier applicants (14.6 %) scored at or below a full-scale IQ of 90, whereas only 2 USAF pilot candidates (less than 1%) scored this low. Similarly, the top 5% (8 applicants) of the air carrier distribution had IQs that ranged from 114 to 127, whereas the top 5% for the USAF distribution ranged from 120-141.

Discussion

The low minimum educational requirements are particularly interesting given the results of Spurlock v. United Airlines ("Spurlock v. United Airlines," CA 10 1972). In this case, United Airlines had established a bachelor's degree as a minimum educational qualification for airline pilot trainees. Spurlock sued based on adverse impact, but the 10th Circuit Court upheld the need for a college degree.

Arguably, educational requirements could be lowered if standardized intelligence testing were used to identify those individuals less likely to pass the ground school portion of air carrier training in the allotted training time. To compare the distribution of air carrier pilot applicants to those of USAF pilot candidates, two regional carriers administered standardized intelligence tests to the applicants who had passed the screening portion of the hiring process.

In examining the two distributions shown in Figure 1, three facts should be noted. First, the air carrier distribution is based on a relatively small sample size (N=82). Thus, the properties of the distribution may change somewhat with more cases. Second, the data were obtained from relatively small regional carriers and, therefore, may not be representative of the major U.S. carriers. Third, the data were obtained from job applicants. The distribution may not be representative of those individuals who received and accepted a job offer. Nevertheless, the USAF distribution was taken from candidates before they entered undergraduate pilot training. Arguably, the lower-scoring individuals would fail or leave UPT, resulting in a larger difference between the two distributions.

The two distributions have very similar shapes but differ in their means, SDs, and ranges. These differences raise some interesting questions. The first question pertains to the large difference in the means. Several factors may contribute to this difference. The first is the difference in the minimum educational requirements. By requiring only a high school degree, the air carriers may be facilitating a shift in their applicant pool toward that of an unscreened population. The second is the omission of standardized tests in the selection process. Again, this omission may facilitate a shift in the pilot applicant pool toward that of an unscreened population. Additionally, the omission may be attributed partially to a misunderstanding of litigation involving adverse impart and partially to the cost associated with test administration. The third factor is the presence of companies and websites that help candidates pass the airline selection process. These companies train applicants on interviewing techniques and on the specific interview questions. Both companies and websites may provide information on the specific selection instruments being used by a carrier and may give examples of questions and correct answers. This information may increase the likelihood that candidates who previously would have been rejected are now successfully completing the hiring process. This information may also contribute to decreasing the predictive validity of the selection system.

The second question pertains to the implications of these data for training. Training developers need to take the characteristics of the learner into account when developing training. Very few studies have been published describing the aptitudes of air carrier pilots. The extent to which training developers use the existing data is unknown.

In summary, the intelligence scores from two small regional carriers were compared to the of USAF pilot candidates. The distributions show marked differences. The extent to which these differences are taken into account in the development of training material is currently unknown.

References

Buster, M., Roth, P., & Bobko, P. (2005). A process for content validation of education and experienced-based minimum qualifications: An approach resulting in federal court approval. *Personnel Psychology*, *58*(3), 771-799.

Fiorion, F. (2007, February 5, 2007). Who's flying the airplane? *Aviation Week & Space Technology, 166,* 52.

Hopkins, G. E. (2001, February). A short history of pilot shortages. *Air Line Pilot,* 18-54.

Jackson, D. N. (2003). *Multi-dimensional Aptitude Battery--II Manual*.

Lehman, C. (2003). Another year of challenge. *Cat Sourcebook 2003-2004,* 11-18.

Pilots and Aviation Maintenance Technicians for the Twenty-First Century. (1993).): Department of Transportation, Federal Aviation Administration.

Raven, J., Raven, J. C., & Court, J. H. (2001). *Advanced Progressive Matrices*. Oxford, England: Oxford Psychologists Press.

Spurlock v. United Airlines, 475 F.2d 216 CA 10 1972).

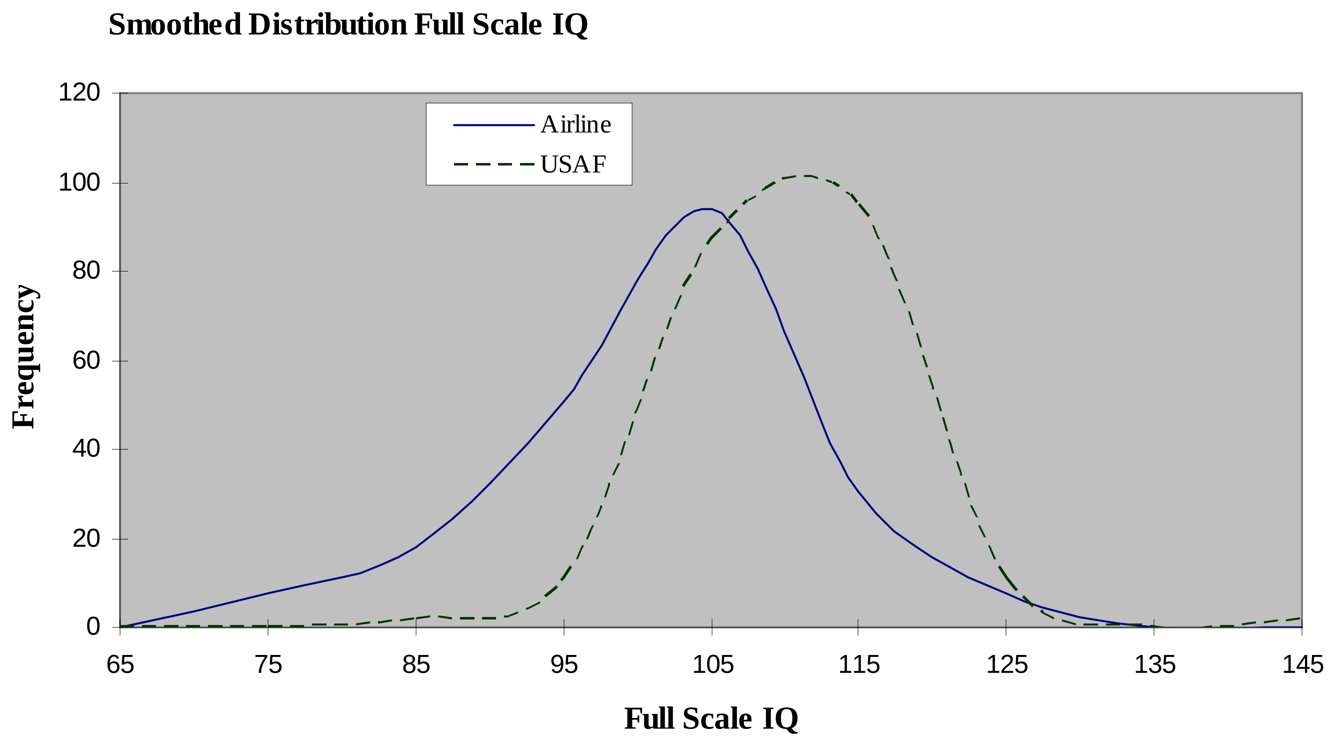

Figure 1. Distribution of Full-Scale IQ scores for pilot applicants at two airlines and the United States Air Force.