Development of an FAA-EUROCONTROL Technique for the Analysis of Human Error in ATM

Dr. Anne Isaac
Human Factors and Manpower Unit
EUROCONTROL
anne.isaac@eurocontrol.int

Dr. Julia Pounds
Federal Aviation Administration
USA
julia.pounds@faa.gov

Abstract

Human error has been identified as a dominant risk factor in safety-oriented industries such as air traffic control (ATC). As the capacity and complexity of airspace continues to increase, and as ATC develops more advanced interfaces and computerised support technology, the importance of identifying the human factors leading to human error will increase, and the ability of traditional design practices alone to effectively mitigate human error will be strained. Therefore, appropriate methods for development of error tolerant systems are needed.

This paper reports on the project to harmonise two methods for investigating the human factors behind human errors in air traffic safety systems. The Human Factors Analysis and Classification System (HFACS) is a human factors taxonomy originally developed for the US Navy for investigation of military aviation accidents and is currently being used by the US Federal Aviation Administration (FAA) to investigate civil aviation accidents. The Human Error in ATM Technique (HERA) is a method of human error identification developed by EUROCONTROL for the retrospective analysis of airspace incidents and for prospective diagnosis during ATM system development.

Background

Human Error in Air Traffic Management

Human errors in ATM/ATC have been defined by Isaac and Ruitenberg (1999, pg. 11) as "intended actions which are not correctly executed." Further, Hollnagel, Cacciabue, and Hoc (1995) pointed out that the term *human error* can denote a cause as well as an action. Thus to comprehensively examine human error in air traffic control, one should consider the possibility of cognitive failure which may result in an incorrectly executed action. Past research has demonstrated that breakdown in cognitive processing such as attention and communication have contributed to reported operational errors (OEs) in US airspace. An OE is defined as a violation of the applicable separation minima between two or more aircraft, or between an aircraft and terrain or obstacles, or when an aircraft lands or departs on a runway closed to aircraft operations after receiving air traffic authorization (FAA Order 720.56, 2001). Early analyses by Kinney, Spahn, and Amato (1977) found that 95% of separation violations in en route centers that were classified as operational errors involved errors in attention, judgement, or communications. These same error types have repeatedly been found in other studies of air traffic control operational errors (e.g., Redding, 1992; Rodgers & Nye, 1993; Stager & Hameluck, 1990).

Analysis of Human Error in ATM

The FAA has several model-based programs of work relating to identifying and reducing human error in aviation. One of these is the work currently underway at the Civil Aerospace Medical Institute (CAMI) to adapt a previously developed method, HFACS, to the ATC environment for research on human factors related to OEs. EUROCONTROL has also recognised the need for a model-based approach to understanding human error and is pursuing similar work in the HERA project. (EATMP, 1999b). Although there are parallels between the FAA and EUROCONTROL objectives with respect to human error, there are differences in the ways the issue of human error is being addressed. For example, both the FAA and EUROCONTROL have focused on human error, their cognitive process, and other operational factors. However, the two techniques vary in distinctive ways. HERA places the air traffic incident in its ATM context by identifying the ATC behavior, the equipment used, and the ATC function being performed. HFACS examines instances of human error as part of a complex productive system which includes management and organizational vulnerabilities.

The Trans-Atlantic Partnership

Theoretical Backgrounds

Both techniques incorporated common elements from human error models and general psychological theories. Comparison of the two techniques suggests that the HERA technique was formulated after identifying useful frameworks relevant in ATM/ATC. Further, HERA development included models of ATC performance and of current and future ATC task and behavioral requirements. In contrast, development of HFACS was based on a set of models drawn from psychology, aviation, and accident and industrial human error identification. Although later adaptations have expanded the HFACS model to other domains (e.g., aviation maintenance), the original model was developed for the investigation of naval aviation accidents and incidents. Thus, the taxonomy was not originally developed to represent ATC concepts specifically, although the general concepts captured in the HFACS tiers, categories, and subcategories seemed to be generally applicable to ATC. Both HFACS and HERA include some of the same theoretical concepts, but at different levels of granularity. For example, the specificity of HERA's identification of psychological error mechanisms is not captured explicitly in the HFACS technique. The HFACS analyst is however required to perform content analyses and make inferences about cognitive processes.

Conceptual Coverage

Both HFACS and HERA view the individual operator as an element in a larger safety system. Conceptually, both techniques analyse the error event by considering the relationships between elements in the system. Both techniques also examine individual errors and the situational and organizational factors surrounding the event. The strength of the HFACS technique is that it forces the analyst to capture the conceptual depth and breadth of the system view by moving from the individual act to the preconditions, supervision, and organizational influences. HERA's strength is that the technique provides a fine-grained analysis of the individual's cognitive processes to identify those that lead to the error event. Thus the conceptual similarities and differences between the techniques are not so much related to which concepts are included, but rather the differences between where the primary analytic effort is invested.

Analytic Methods

The two techniques differ in several ways in the conduct of analysis. Because of these differences the two techniques result in somewhat different types of output data. However, both techniques rely on summary data from other investigators and in both cases the analysts using these techniques are urged to resist making assumptions beyond the data given.

Reliability of the Methodologies

Any useful error framework should be broadly applicable. That is, different users analyzing the same error event using the method should identify similar factors. Both HERA and HFACS were developed against the criteria of Cohen's Kappa, an index of agreement between multiple coders which is corrected for chance. Both methods had successful validation trials. A large validity and reliability study was conducted to test the HERA technique for consistency across users and across reports originating from different nations (EATMP, 2000). Several reliability studies have been conducted to test the HFACS model. Initial studies were conducted using lists of causal factors from US Navy, Marine Corp, and Air Force aviation mishap reports (Shappell & Wiegmann, 1997).

Thus the techniques have some similarity and several differences. More detailed comparisons between the two techniques are presented in the Appendix.

The Harmonisation Process

The harmonisation process to create a technique using the strengths of both techniques was undertaken in three separate but associated phases. Phase 1 analyses compared techniq;ues and developed materials for Phase 2 analyses. In Phase 2 operational personnel provide their opinions about the relative utility of concepts from each technique. Phase 3 resulted in the harmonized technique.

Phase 1 – Comparing the Two Techniques

Method

Air traffic control subject matter experts (SMEs) with experience in operational incident investigations and familiar with both techniques were recruited to participate in the analysis: two representatives from EUROCONTROL and two from the US. The incident cases (10 European and 10 US) were selected to represent different types of possible scenarios, e.g., from terminal, en route, etc. Not surprisingly the formats were different, but allowed the analyses to be undertaken.

The recording form for HERA had 7 classification categories (i.e., Task, Equipment, and Information, Error Type, Cognitive Domain, Internal Error Mode, Psychological Error Mechanism, and Performance Shaping Factors). The data recording template for HFACS had 3 classification categories (i.e., Tier, Category, Sub-category).

The differences between HERA and HFACS can be seen in the different levels and concepts used by the two techniques. Together, the HERA and HFACS techniques represented 454 concepts (terms) -- 414 from HERA (91%) and 40 (9%) from HFACS. These were distributed as follows. The HFACS Categories (C) and Subcategories (S) within each Tier (T) are not listed here.

HERA Conceptual Groups
28 - Tasks
85 - Equipment and Information items
27 – External Error Modes/Violations
4 - Cognitive Domain levels (CDs) and Error
Mechanisms encompassing 67 concepts :
24 - Perception and Vigilance
17 - Memory
15 - Planning, & Decision Making
11 - Response Execution
207 – Performance Shaping Factors (PSFs)
HFACS Tiers
13 - Unsafe Acts
9 - Preconditions
5 - Unsafe Supervision
13 - Organizational Influences

All participants first worked independently to analyse each of the 20 incident cases prior to convening at the FAA Civil Aerospace Medical Institute in Oklahoma City for a three-day joint meeting to compare results of the individual analyses. At the meeting 5 US and 5 European of the original 20 incident cases were re-analysed using the SMEs' individual analyses as the starting point. Both HERA and HFACS data were discussed for each incident and disagreements were resolved.

Results

Output from the analysis of each incident was a list of items (identified error events), and the associated human factors terms resulting from the SMEs' analyses. An illustration of the output from the analysis of one error item is shown in Example 1. Terms 1-9 are the concepts resulting from the SMEs' HERA analysis; terms 10-13 are the concepts output from their HFACS analysis.

To understand the relative contribution of each method, the terms generated in Phase 1 from the analysis of all 10 cases were compiled into one list. Many of the terms had been selected in more than one case analysis. The resulting list contained 1818 data points representing the terms used: 1156 (63.6%) from the HERA analyses and 662 (36.4%) from the HFACS analyses. Because the conceptual contribution of each method to the consensus analysis was the interest here, duplicate items were removed to eliminate double counting. This resulted in a list of 126 unique concepts: 98 (77.8%) from HERA and 28 (22.2%) from HFACS.

Thus, although the percentage of HERA concepts relative to the HFACS concepts from these analyses differed from the initial availability of 91% and 9% respectively, these results were not conclusive and only revealed that both techniques contained useful elements upon which the harmonized technique could be built.

Phase 2 – Analysing the Two Techniques

The purpose of the second phase was to use the output from Phase 1 to identify the most useful concepts from each technique for operational investigations, the depth of detail most useful for retrospective analysis of incidents, and the advantages and disadvantages of each technique as a tool for use by operational personnel. To accomplish this a panel of experts with experience in both operational investigations and associated mitigation strategies was convened.

Method

The meeting to conduct the Phase 2 analyses was held at the Institute of Air Navigation Services (IANS) in Luxembourg. Five SMEs were selected to participate in this expert forum based on their operational expertise, knowledge and experience about operational needs relative to mitigation of operational incidents. They had no prior knowledge of either technique. Two SMEs from the first meeting also attended to clarify any questions about the data.

Data from the 10 cases analysed in Phase 1 were presented as follows. An example from the materials is shown in the boxed section of Example 1. For reference, the example also lists the technique and its concept for each Term. (The participants did not have this information, however.) First, an *Incident Situation* statement summarized the overall event. In the example, the Situation was that the *Arrival a/c was descended to an altitude that put it in conflict with an overflight a/c*. The critical points identified and analyzed in Phase 1 were listed as *Items*. In the example, the first critical point Item was that the *Controller missed an incorrect altitude readback.* A total of 40 Items were presented to the SMEs for analysis. The number of Items within Incident Situations ranged from 2 to 7 (mean = 4, mode = 5). The *Terms* output from the Phase 1 analysis were listed under each Item. In the example, there were 13 Terms.

The number of Terms to be ranked within Items over all Situations ranged from 2 to 26 (mean = 9.1, mode = 13). Overall, there was a total of 363 Terms -- 228 Terms from HERA (62.8%) and 135 (37.2%) Terms from HFACS appeared across Incidents for ranking. On the right, the example shows the results from the later Phase 2 analysis for each Term.

Example 1. Results from HERA and HFACS Analyses.

Incident: 11, Situation: Arrival a/c was descended to an altitude that put it in conflict with an overflight a/c.	Technique		Phase 2 Analysis	
Item 1: Controller missed an incorrect altitude readback.	HERA	HFACS	Mean Rank	Mean Score
1. R/T Communications -- read-back	Task		5	.05
2. Descent	Keyword		10.2	.11
3. Clearance	Keyword		8.2	.09
4. Altitude	Keyword		9.2	.10
5. Incorrect information received/recorded	External Error Mode		8.8	.10
6. Perception and Vigilance	Cognitive Domain		4.6	.05
7. Hearback/No Detection-auditory	Internal Error Mode		5.8	.06
8. Expectation bias	Psychological Error Mechanism		6.6	.07
9. Pilot breach of R/T Standards	Performance Shaping Factor		10	.11
10. Skill-based error		T1, C1, S2	6	.07
11. Attention error		T1, C1, S2, Failure	2.2	.02
12. Error		T1, C1	6.4	.07
13. Unsafe act		T1	8	.09

Note. HFACS Levels: T = Tier, C = Category, S = Subcategory. *N* = 5.

The members of the expert forum worked individually and ranked the Terms according to how important each would be (relative to the other Terms in the set) in understanding the Incident Situation using the following method: 1 = Most Important to N = Least Important. Because the number of Items under each Incident Situation was not held constant, N, the upper limit on the range of scale values, was dependant upon the number of other Terms in its list. Each technique was not equally represented in each list and 11 Items did not have any HERA Terms listed for ranking. These Items had 2-3 Terms listed and examination revealed that they identified primarily supervisory and organizational vulnerabilities.

After completing the ranking task, the participants were divided into teams, each having both US and European experts. They were given general instruction about conducting HERA and HFACS analyses. Two of the experts from the Phase 1 meeting monitored the groups to answer any technical questions about the methods. Each team analysed two incident cases (one European, one US) using each method. The order of cases and method used were counterbalanced. This activity was designed to give the participants hands-on experience with both techniques before they were asked for feedback about overall strengths and weaknesses of the techniques.

Results

Utility of Terms

The rank of each Term was converted to a score that both represented the number of options competing for ranking with it under that Item and which could also be compared across Items.[1] Scores range from 0 to 1, with lower scores indicating a higher ranking adjusted for number of possible Terms competing for that ranking.

Ranking data from all experts resulted in scores for 1818 Terms--1156 Terms from HERA and 662 from HFACS. Of these, some Terms received no rankings by the experts and were assigned a ranking of 0. Table 1 shows the overall scores for each technique. In general, the expert forum rated the HERA items as more important to understanding the incident.

Table 1. Analysis of Ratings for Each Technique Over All Items

	N Scores	Mean Score
HERA	1156	.08
HFACS	662	.16

[1] We thank Dr. David J. Weiss from California State University—Los Angeles for this technique.

Relative Utility of Techniques

The mean scores for each HERA concept is shown in Table 2. They have been ordered from lowest (Cognitive Domain), to highest (Internal Error Mode, e.g., no detection-auditory). The expert forum appeared to prefer those HERA Terms that were descriptive of information about the general stage of information processing associated with each critical point (e.g., Perception and Vigilance) and the associated cognitive mechanisms (e.g., visual search failure), but ranked HERA Terms describing how the error was manifested internally (e.g., no detection--visual) as being less useful.

Table 2 . Mean Scores for HERA Categories

Category	N	Mean
1. Cognitive Domain	95	0.04
2. Psychological Error Mechanism	110	0.06
3. External Error/Violation Type	130	0.08
4. Task	151	0.08
5. PSFs	300	0.09
6. Information and Equipment	265	0.10
7. Internal Error Mode	105	0.54

The same method was used to compare Terms from the HFACS technique at the Tier and Category levels. Table 3 shows the mean scores associated with HFACS Terms at these levels. Note that not every HFACS tier and category were represented in the data. Some were eliminated during Phase 1.

The mean scores for HFACS Terms suggest that the participants in the expert forum preferred concepts describing the individual (e.g., Unsafe Act). However, the rankings suggest that the HFACS categories ranked as less useful include those addressing preconditions and supervision.

To compare both techniques, the Terms from Table 2 and Table 3 were ordered from lowest to highest. An equivalency between HERA Categories and HFACS Tiers/Categories was presumed. Based on these scores it appeared that most of the HERA Terms were rated as more useful than the HFACS Terms.

Table 3. Mean Scores for HFACS Terms by Tier and Category		
	N	Mean
1. Unsafe Acts	120	.11
1.a Errors	335	.07
1.b Violations	20	.21
Tier Emphasis	475	.09
2. Preconditions for Unsafe Acts	5	.43
2.a Substandard Conditions of Operators	10	.28
2.b Substandard Practices of Operators	--	--
Tier Emphasis	15	.33
3. Unsafe Supervision	26	.42
3.a Inadequate Supervision	5	.53
3.b Planned Inappropriate Operations	31	.42
3.c Failed to Correct Problem	5	.08
3.d Supervisory Violations	--	--
Tier Emphasis	67	.40
4. Organizational Influences	35	.31
4.a Resource Management	40	.26
4.b Organizational Climate	10	.35
4.c Operational Process	20	.36
Tier Emphasis	105	.30

Overall Strengths and Weaknesses of the Techniques

After both teams completed all analyses everyone was asked for their oral and written feedback about how useful each of the techniques would be, their strengths/weaknesses, and usability. The questions asked and a summary of the most important results in rank order appear in Tables 5, 6, and 7.

Table 5: "What are the good/well-liked aspects of these approaches?"	
HERA	N
A very comprehensive and detailed approach	10
Questions and flow-charts are good	4
Provides specificity	4
Leads the analyst through the process	4
Considers all errors in an event equally	3
Leads you back if you go wrong	2
Does not blame	1
HFACS	
The process is simple to understand and quick to use	9
Less time needed for analysis	2
It describes items well	2
There is a distinction between error and violation	2
Adverse supervision is considered a variable	2
Easier to train someone in this method	1
Includes causal factors	1

TABLE 6: "What are the poor / disliked aspects of these approaches?"	
HERA	N
Too much paper to go through	3
The Internal Error Modes, Psychological Error Mechanisms and Performance Shaping Factors are quite complex without training	3
Adverse supervision should not be a Contextual Condition	2
Too much human factors jargon	2
Too subjective	2
The causal categories are difficult to establish	1
The pro-forma should be redesigned	1
Overlooks non-compliance from the controller	1
HFACS	
Oversimplification which could lead to wrong conclusions	8
Misunderstanding the tiers/categories/sub-categories	4
Limited nature of error classification	3
References to the pilot environment	3
Academic wording not suitable	2
Definitions are not clear and specific enough	2
Too easy to be subjective	2
No cross checking in the technique	2
Technique seems incomplete	2

Table 7: "What would you like to see included (✓) or excluded (x) in future technique development?"	✓	x
HERA		
Recording Form	5	0
Task lists	6	0
Information and Equipment lists	4	2
Cognitive Domain flow charts	7	1
External Error Mode/Violation tables	7	1
Internal Error Mode flow charts	7	1
Psychological Error Mechanism flow charts	7	1
PSFs tables	7	1
HFACS		
Unsafe Acts categories	6	0
Error categories	6	2
Violation categories	5	3
Preconditions for Unsafe Acts	3	4
Unsafe Supervision categories	5	3
Organizational Influence categories	6	2

Phase 3--JANUS: A harmonized technique

The goal of the third phase was to agree on an acceptable technique which included the best aspects from both approaches. In order to fulfil this goal, the research leaders from the FAA and EUROCONTROL met for a four-day meeting in Brussels to discuss the findings from the previous two phases of work. During this meeting the harmonised technique emerged.

It was clear from this work that each of the two technqiues was developed for quite different specific objectives in two different environments. Their commonality, however, is that they both draw from some of the same foundational literature of cognition and human error, albeit to different degrees and to different ends. Also, both attempt to improve the way human error is identified and analysed in the aviation environment.

There are several points which should be mentioned about the objectives of these two methods. Firstly, HFACS was designed for the military flying environment and the HERA technique was specifically designed in the ATC environment. Secondly, and perhaps more importantly, is the fact that HFACS was designed to investigate the human error embedded in aspects of an incident/accident broadly defined, whereas the HERA technique concentrates most specifically on the human error causal factors in the incident. Although there are typically aspects of other human factors issues within the HERA technique itself, i.e., in the PSFs, HFACS specifically tries to capture those categories, i.e., Unsafe Supervision by Planned Inappropriate Operations. Another factor which obviously influences the usability and acceptability of the two techniques is the precise nature of the HERA technique which was designed to find the specific underlying cognitive failure within the human--the controller in this case. The categorical HFACS technique, on the other hand, seeks to establish the chain of events to link the system vulnerabilities which result in failed human performance.

The different initial objectives and development of the two techniques has led quite naturally to different methodologies. Neither is better or worse; they are simply different. Although both techniques seek to address similar human factors issues, the method for identifying the issues and the grain of analysis are different between the two techniques. The goal of this work was to find the common threads.

It is clear from the work reported that the harmonized technique would benefit from incorporating the HERA technique's detailed, comprehensive, complex and more specific methodology at the individual level. This should lend increased precision to incident investigation. The users reported their appreciation that its logical and structured approach reduces subjectivity. However the analysts also report that its complexity, and often specialised use of language would make use of the technique more difficult without special training for the users of this technique.

Similarly, the harmonized technique would benefit from incorporating the system-wide approach from HFACS. Users reported that HFACS is a simple, easy to comprehend technique, which lacks the cognitive specificity of HERA but is comprehensive and defines contextual factors at the supervisory and organizational level. Contextual factors are often found more distal from the final incident or accident but are often no less contributory. Its broader categorical approach to analysis allows quicker analyses and possibly less training to use effectively.

Conclusion

In sum, this project revealed that the two techniques, HERA and HFACS, were as complementary as they were different. Thus, the ability of the HFACS technique to capture supervisory and organizational vulnerabilities was combined with the specificity of the HERA technique to generate the harmonized technique called JANUS, named for the mythological guardian of citizens, who looked into both the past and the future, representing the philosophy of learning from past error situations in service for future aviation safety. The technique is diagnostic at the level of the individual's cognitive processes but also views the individual as part of the larger human-computer-organizational system.

JANUS is now being trialled in the US ATM system and within 7 European Countries as a tool to increase the information about causal factors related to operational errors. JANUS will be tested with operational error incidents by investigators and human factors researchers. At the completion of this testing phase the technique will be again tested and evaluated for its validity and utility as an investigatory tool.

As new systems are developed for ATM to meet future capacity demands, it is critical to have an understanding of the points at which human and system error might impact outcomes. It is likely that these tools will place increasing demands on the controller's cognitive processes to safely expedite air traffic. In addition to the known set of possible types of errors, new strategic planning tools are likely to introduce new types of errors. Once validated, the JANUS technique will provide a more sensitive means to identify and assess human and ATM system errors associated with using these tools than those techniques currently available.

References

Bird, F. (1974). *Management guide to loss control*. Atlanta, GA: Institute Press.

Edwards, E. (1988). Introductory overview. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 3-25). San Diego: Academic.

EATMP. (1999a). *Technical review of human performance models and taxonomies of human error in air traffic management*. (HUM.ET1.ST). Brussels: EUROCONTROL.

EATMP. (1999b). *Short Report on Human Performance Models and Taxonomies of Human Error in Air Traffic Management*. (HUM.ET1.ST). Brussels: EUROCONTROL.

EATMP. (2000). *Validation of the human error in ATM (HERA) classification system*. Draft report. Brussels: EUROCONTROL.

Federal Aviation Administration. (2001). *Air Traffic Quality Assurance* (FAA Order 7210.56). Washington, DC: U.S. Department of Transportation.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.

Heinrich, H., Petersen, D., & Roos, N. (1931). *Industrial accident prevention: A safety management approach* (1st ed.). New York: McGraw-Hill.

Hollnagel, E., Cacciabue, P. C., & Hoc, J. (1995). Work with technology: Some fundamental issues. In J. Hoc, P. C. Cacciabue, and E. Hollnagel (Eds.), *Expertise and technology* (pp. 1-15). Hillsdale, NJ: Lawrence Erlbaum.

Isaac, A. R., & Ruitenberg, B. (1999). *Air traffic control: Human performance factors.* Hants, England: Ashgate.

Kinney, G.C., Spahn, M.J. & Amato, R.A. (1977). *The human element in air traffic control: Observations and analyses of the performance of controllers and supervisors in providing ATC separation services* (MTR-7655). McLean, VA: MITRE Corporation: METRIEK Division.

Jensen, R. (1995). *Pilot judgment and crew resource management*. Brookfield, VT: Avebury Aviation.

Martiniuk, R. G. (1976). *Information processing in motor skills*. New York: Rinehart and Winston.

Shappell, S.A., & Wiegmann, D.A. (1997). A human error approach to accident investigation: The Taxonomy of Unsafe Operations. *International Journal of Aviation Psychology, 7*(4), 269-291.

Shappell, S.A., & Wiegmann, D. A. (2000). The Human Factors Analysis and Classification System--HFACS. (DOT/FAA/AAM-00/07). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.

Stager, P., & Hameluck, D. (1990). Ergonomics in air traffic control. *Ergonomics, 33,* 493-499.

Rasmussen, J. (1982). Human errors: A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents*, 4, 311-333.

Reason, J. (1990). *Human error*. New York: Cambridge University Press.

Redding, R. E. (1992). Analysis of operational errors and workload in air traffic control. *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 1321-1325), Vol. 2. Santa Monica, CA: Human Factors Society.

Rodgers, M. D., & Nye, L. G. (1993). Factors associated with the severity of operational errors at air route traffic control centers. In M.D. Rodgers (Ed.), *An examination of the operational error database for air route traffic control centers* (pp. 11-25). Washington, DC: Federal Aviation Administration (Report No. DOT/FAA/AM-93/22) ADA 275986.

Wickens, C.D., (1992) *Engineering Psychology and Human Performance* (Second Edition). New York: Harper-Collins.

Wickens, C. D., & Flach, J. M. (1988). Information processing. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 111-155). San Diego: Academic.

Appendix

	HERA	HFACS
Origin	Developed for incident analysis of human errors in ATM	Developed for incident analysis of human errors in aviation accidents
Theoretical Base	Human Error Taxonomies Human Performance Models Task-based Taxonomies Error Mode Taxonomies Communication System Models Information Processing Models Symbolic Processing Cognitive Simulations Other Models and Taxonomies, e.g., SA, control system, SDT, commission errors approaches, violations Other Domain Approaches e.g., accident theory, root cause analysis, nuclear risk assessment, maritime operations, flight operations, ATM Models of Error in ATM performance	Human Error Taxonomies Human Performance Models Industrial Safety Information Processing Models Crew Resource Management
Conceptual Coverage	Ranges from the organizational level to individual internal psychological mechanisms (e.g., expectation bias).	Ranges from the organizational level to the individual's error (i.e., decision, skill, misperception).
Data for Analysis	Incident reports and narrative summaries.	Lists of causal factors from incident databases.
Analytical Process	Each human error point within the incident description is subjected to the entire HERA analysis. Error points are identified by working from the beginning of the incident report to the final event.	The Unsafe Act is identified as well as each related classifiable act in the incident description. Each is then categorized by working backwards from the Unsafe Act. Classifiable acts are identified as "holes in the cheese."
Inter-coder Reliability Values of k = .40 or less are considered "poor" agreement while values of k = .75 or greater are considered "excellent" levels of agreement (Fleiss, 1981).	At the level of Cognitive Domains, Kappa ranged from .44 to .50. With extended training, coders showed overall increased agreement (Kappa = .52), compared to .38 with only basic training. By job function the incident investigators showed highest agreement (Kappa = .61). ATM and researchers agreement was .23 and .43 respectively. Agreement between coders declined as the level of analysis becomes finer-grained, although psychological specificity increases.	Pair-wise comparisons of inter-rater agreement using Cohen's Kappa ranged from .60 in early studies to .95 later in development of the model. Using all categories, Kappa ranged from .65 to .75. Inter-rater agreement was lowest for the Supervisory and Organizational tiers.
Output Data	Each human error can be described by a cognitive domain, internal error mode, and psychological error mechanism. Each error can also be identified by the associated task, information, and a variety of situational performance shaping factors.	Each classified act can be labeled by HFACS tier, category within tier, and subcategory within category if available. Each data point has an associated description which can be subjected to content analysis.

Author Biographies

Dr. Anne Isaac gained undergraduate qualifications in England before moving to New Zealand. There she gained her post graduate qualifications in neuropsychology and taught at Otago University in human performance. Her research work in the ATC environment started in 1983 and continued when she transferred to lecture at the Massey University aviation school. Her main research interests are in mental models, the mental picture and situational awareness and how it relates to errors in ATM. She moved to Dedale in Paris in 1997 to create and produce the Team Resource Management (TRM) programme for European ATM, after which she joined EUROCONTROL in the Human Factors team where she has led the work in Error within the ATM system.

Dr Julia Pounds is an Engineering Research Psychologist with the Human Resources Research Division of the FAA's Civil Aerospace Medical Institute in Oklahoma City, Oklahoma. Before joining the FAA, she worked with the US Army Research Institute at Ft. Leavenworth, Kansas on projects to identify and enhance the problem solving skills of battlefield commanders. Since joining the FAA, her primary work has focused on the cognitive processes involved with judgment and decision making in human error, particularly those associated with expert performance of air traffic control and the mitigation of operational errors. Current projects include analysis of operational error data to identify causal factors related to human performance, development of a 360-degree human factors method of investigating system failures related to ATC, pilot, and ground personnel, and programmatic research to develop metrics of expert performance.