2007

Multivariate analysis of operational errors in air traffic control

Damir Ceric
San Jose State University

MULTIVARIATE ANALYSIS OF OPERATIONAL ERRORS IN AIR TRAFFIC CONTROL

A Thesis

Presented to

The Faculty of the Department of Psychology

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Damir Ceric

August 2007

APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

Dr. Kevin Jordan, Professor, Psychology

Dr. Kevin M. Corker, Professor, Industrial and Systems Engineering

Samm Starr, SJSUHAIL

APPROVED FOR THE UNIVERSITY

ABSTRACT

MULTIVARIATE ANALYSIS OF OPERATIONAL ERRORS IN AIR TRAFFIC CONTROL

by Damir Ceric

Improving safety by reducing the frequency of operational errors in an air traffic control environment has long been the goal of the FAA. Complexity embedded in air traffic control operations, configuration of airspace, and human information processing capacity limitations allow for multiple factors to influence incidents, creating several categories of operational error. This study examines several predictors of operational errors, and is tailored towards recognizing error patterns. A single operational error may have multiple categories. Identifying unique contributions of any of them is essential, because it leads to a better human error prediction model. Results in this study revealed the influence of a number of contextual factors in prediction of error type frequency. Furthermore, the severity conformance metric (an aircraft proximity index) predicted the frequency of error categories to some extent. The number of aircraft in a sector did not contribute significantly towards the prediction of error categories.

ACKNOWLEDGEMENTS

I want to mention and thank several people who made my thesis writing experience meaningful.

Dr. Jordan introduced me into the world of perceptual psychology research by offering me a research assistant position at NASA Ames research center during my graduate years. I became even more interested in human perception after attending his brilliantly taught graduate course at San Jose State University. He is my thesis committee chairperson and I am sincerely grateful for his expertise in the field, his patience, and support with my constantly changing thesis topics.

Dr. Corker, who has been my mentor, the principal investigator at SJSUHAIL lab, allowed me to work with unique data sets from aviation and air traffic control. I am also grateful for being able to collaborate with him on various projects, one of which became my thesis topic. I would have been lost without his guidance in this, for me a rather exotic field of study. I am flattered to have a scientist of that caliber as my thesis committee member. I am sincerely thankful for his patience and encouragement throughout this process.

Samm Starr, our lab manager at SJSUHAIL lab, helped me with my thesis writing in ways too numerous to list. She organized lab teamwork perfectly throughout various projects, the largest of which became the dataset used in my thesis. She incorporated our duties properly into the thesis requirements and flexibly arranged schedules that were balanced with school and work. She helped with the editing, formatting, and construction of my thesis. I sincerely thank her for everything.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

Introduction

Human Error

Due to their complex nature, the causes of human error have never been precisely defined. We are more likely to talk about different perspectives on human error, where each one differs in terms of the nature and cause of the problem. Although it is beneficial to have different approaches to a problem, most human error models are of a theoretical or academic nature without a significant benefit to applied practitioners (Wiegmann & Shappell, 2001). The authors mentioned above describe five different perspectives on human error pertaining to the aviation environment: cognitive, ergonomics and systems design, aeromedical, psychosocial, and organizational. This approach, however, could be applied to any human-machine interaction environment.

The cognitive approach is one of the most commonly used models for human error. Those models are usually considered in clearly distinct stages (attention, recognition, decision, and action) and are based on the information processing theories of cognitive psychology. Although popular, cognitive models are not problem-free as they do not address all the relevant issues. For example, they do not consider the influences at an organizational level or any task-related factors such as equipment design.

Contrary to the cognitive approach, in the ergonomics and systems design approach the human is never the only responsible cause of error. They are seen as yet another component of an inseparable system including software, hardware, and various environmental conditions. Within the aviation community, the focus is usually on human-hardware interaction and improvements in cockpit and air traffic management displays and procedures. The disadvantage of this approach lies in the potential overemphasis on the human-machine interaction. There is a tendency to see all problems as caused by a flawed design, thus neglecting cognitive, social, and organizational aspects.

The aero medical perspective approach emphasizes underlying physical conditions that may lead to human error. Due to the unique environment of an aircraft crew operation, there is a chance that hypoxia, dehydration, fatigue, and disorientation may cause more frequent human errors. The physiological aspect is usually underrepresented in crew training and one of the challenges is to determine the tolerance threshold of each person, beyond which the crewmember is likely to commit an error. The psychosocial perspective is seen in terms of the interaction between all aspects of air transport. Pilots, flight attendants, and air traffic control personnel, all contribute to the quality of those interactions, and most importantly human error is viewed as a function of the breakdown in team dynamics and interpersonal communications. The major drawback of psychosocial theories of human error is the lack of empirical evidence.

The organizational perspective was only recently introduced in the aviation environment because of an increasing complexity of air accidents. This perspective is based on the decisions of managers and supervisors as to the causes of the incidents. The major advantage of this approach is broadening the area of accident investigation as well as the use of industrial organizational psychology research. However, the causal distance of organizational factors from the context in which an error occurred makes it almost impossible to address any aspect of the organization as responsible for the error.

Furthermore, in case of a remedial intervention, changing procedures at the organization level does not guarantee success.

Human error in aviation. Even with great improvements in air travel safety during the last five decades, there is an increase in accidents and runway incursions (Shappell & Wiegmann, 2001). Although we may never be able to answer exactly why airplane accidents occur, there is a trend in the nature of aircraft accidents that is worth mentioning. Decades ago, failures were primarily mechanical and physical in nature, while today the most responsibility can be attributed to human error. Nowadays, people flying the aircraft are more dangerous than the planes they fly (Shappell & Wiegmann, 2000). Even though there is a 70%-80% contribution of human error in accidents, not all the errors are related to a single cause, making the whole investigation process very complicated (Shappell & Wiegmann, 2001). The multitude of factors that may or may not contribute to committing an error is too large for a simplistic interpretation of incidents. Therefore, having a valid and reliable human error classification system would be beneficial for helping understand the specific factors involved in each individual incident. That is exactly what is missing today; a unified, data driven framework that could be used for the analysis of human errors. Because, there is no final consensus on a unified definition of a human error, for practical reasons, classification systems have been developed.

Reason (1990) proposed a four-layer model that describes human failure where the layers influence each other. The first layer is called active failure and is directly responsible for the accident, while the other three layers are termed latent failures. Unsafe acts, the actions that directly lead to an accident, are the point where the investigators find the majority of "causes." An example would be improper scanning of an instrument. The next layer is called preconditions for unsafe acts (including psychosocial and physiological breakdowns such as communication problems or fatigue). Crew resource management (CRM) is the term commonly used to refer to those issues. For example, if the crew member is fatigued, bad decisions are made and an accident may occur. Further, the third layer, termed unsafe supervision, happens in complex, multi-factor situations. For example, two inexperienced pilots who do not know each other are paired and sent through bad weather. Therefore, there is a lack of supervision that may lead to the system entering preconditions for unsafe acts and eventually to an incident or accident. The fourth and last layer is at the organizational level. For example, the airline budget cut translates into more modest training and flight time, eventually leading to an accident.

While this model encompasses many aspects of human error, it is mainly a theoretical framework that lacks details of a real-world application. Our goal is to know the "holes" in each layer and identify them during an accident investigation or detect and prevent them from happening in the first place.

Human error in the air traffic control environment. The current study focuses specifically on human error and concentrates on errors committed by air traffic controllers. Human error in the air traffic control environment shares many of the characteristics of human errors in general aviation. Although the collisions of two or more aircraft are rare (Rodgers, 1998), reducing the number of human errors that occur in air traffic control is a meaningful endeavor.

Air Traffic Control Environment

Considering the technical nature of this study, it is essential to present a brief summary of the operating procedures used in the air traffic control system.

Airspace sectors. The US airspace consists of 21 zones and there is one Air Route Traffic Control Center (ARTCC) for each zone. For increased safety and efficiency, the 21 zones are divided into smaller sectors (Davis, Danaher, & Fischl, 1963).

Air traffic control centers. Tower control is in charge of airport runways, taxiways, and the immediate surroundings of the airport. Landings, takeoffs, and handoffs to the other sectors are handled by this type of air traffic control. Consequently, the size of the airport and the frequency of the traffic will dictate the complexity of the tower operations.

TRACON. Terminal Radar Approach Control (TRACON) is a type of control service that exists only around major airports. It serves as an intermediate step in charge of maintaining climbing and descending traffic between the airport and the en-route services. One of the primary functions of TRACON is to slow down the aircraft when approaching the major airport.

ARTCC. Air Route Traffic Control Centre (ARTCC) is an en-route service providing pilots with control services on their way between the airports. Usually, this air traffic control occurs at high-level altitudes in airspace divided into sectors. Every sector is under the responsibility of its own air traffic controller team. The two-member team includes one radio controller or R-side (executive controller) and one data controller or D-side (planner controller). The only exception to this setup occurs during the high volume traffic situations when the third controller (tracker) may be added.

Similarities and differences between TRACON and ARTCC. In order to understand this two-part study, it is essential to outline the specific nature of both types of air traffic centers. The more dynamic centers are TRACON centers where constant altitude change, more variety in aircraft types, and greater reliance on visual flying is encountered. The main difference between ARTCC and TRACON is the attitude towards dealing with emergencies. In ARTCC, where the traffic flow is more predictable, there is sufficient time to develop long-range strategies for dealing with the problems that may arise. However, In TRACON emergencies must be handled in a more immediate way or tactically, as they arise suddenly in this less predictable environment (Wickens, 1997).

Separation minimums for TRACON and ARTCC. Another important set of differentiating characteristics are mandatory safety separation minima. ARTCC separation restricts the aircraft's proximity to five miles laterally and 2000 feet vertically. In TRACON, however, the minimum standards are three miles and 1000 feet (Wickens, 1997).

The air traffic controller's responsibilities. According to the Air Traffic Controller Manual of controller duties, 7110.65L (DOT/FAA/OED, 1996), the first priority is to separate aircraft and issue safety alerts if necessary. That is accomplished by using good judgment based on the situation at hand. The controller should utilize all available information in guiding the aircraft until it is handed off to the adjacent sector. Although there are many similarities among controllers (depending on the kind of the air traffic controller involved), the perceptual and cognitive demands may vary, which in turn may have an impact on errors. For instance, ARTCC controllers rely on a radar display and paper strips, while tower controllers can use visual cues to determine the position of an aircraft.

Operational Errors

Operational error can be described as any activity or state of one or more aircraft in the airspace that eventually leads to the breach of the specified minimum separation distance. Additionally, human operators such us pilots, air traffic controllers, and other personnel involved in air operations may contribute to occurrences of operational error. The Federal Aviation Administration (FAA) and The US Department of Transportation (DOT) are the two executive organizations responsible for the defining of procedures and regulations of the operations. Air Traffic Controller Manual 7110.65M (DOT/FAA/OED, 1996) strictly classifies operational errors into critical ones affecting separation minima and less critical ones dealing with the deviations (Rodgers & Nye, 1993).

Legal procedure for handling operational errors. FAA Air Traffic Quality Assurance Order 7210.56c (FAA, 2002) is what determines and describes the procedure of incident reporting (Bailey, 2005). When an operational error occurs, the air traffic center conducts a thorough investigation immediately after the incident. Normally, a supervisor detects the occurrence of an error. The first phase is a completion of a preliminary report document FAA 7210 (FAA, 2000). Next, the incident is presented to safety staff and air traffic managers. The air traffic controller under whose supervision the incident happened is temporarily removed from the position, the recorded tapes are analyzed and all relevant information surrounding the circumstances of the event are gathered by air traffic managers (Rodgers, 1998). In some centers equipped with SATORI (Systematic Air Traffic Operations Research Initiative) system, the whole situation is re-enacted in order to search for the main cause of the error. That method, however, is used only after an operational error occurred (Pounds & Ferrante, 2003). Nevertheless, SATORI provides an accurate description of factors that influenced the incident. At the end, the final operational error or deviation report (containing all information obtained during the investigation) is filed by ATD (Air Traffic Division) manager. Refresher training courses, recommendations for future work, possible penalties, corrective actions, and even decertification finalize the process. For an example of a final operational error report see Appendix A.

Operational errors classification. Due to the complexity of circumstances involved in an incident, the need for a useful framework defining specific categories of operational reports was recognized by FAA evaluators and investigations staff members. However, the most crucial feature of an operational error seems to be "severity." Severity of an operational error is defined as the physical proximity of two or more aircraft, the sector characteristics, weather, and altitude. (Pounds & Ferrante, 2003). The most current severity index (SI) is based on post hoc computations of the data from an incident in an operational error report. There is a point scale ranging from 0 to 100. Several factors such as vertical and horizontal separation, level of control, and flight paths influence the point value assigned to each incident. In addition, there are three basic levels of severity, high, moderate and low (Bailey, 2005). A statistical analysis performed by Bailey (2005) targeted at the assessment of SI as a useful severity classification tool for operational errors. The results showed that, overall, SI served as a reasonable metric for categorizing Air Traffic Control operational errors.

The present study will use the classification frameworks of operational errors developed by San Jose State University Human Automation and Integration Lab (HAIL) in 2005. The classification was based on the analysis of official FAA operational error reports. Corker and Garcia-Chico collected and examined a set of 539 operational error reports that occurred in the first six months of the year 2004 (Corker & Garcia-Chico, 2007). There are 24 different categories of operational error types (Appendix B). The authors developed the framework for both, operational error types and contributory factors. The basis for the evaluation and classification of errors was provided by the objective statements in reports without the interference of the reviewers' interpretations of the facts (Garcia-Chico, 2006). The focus of this study is the number of error categories in the operational error reports. Each operational report with more error categories contained within each incident indicates more complexity, more factors to deal with, and deserve more attention. There are three reasons for having less categories of operational error in a report. The first is entirely of practical nature. Any intervention necessary to separate aircraft will require relatively safe and straightforward solution if less categories are involved. The second would be important in terms of developing a causal model as to why the incident occurred. And finally the fewer variables included in an analysis, the more reliable the results.

Concurrent factors classification. A wide variety of factors closely or remotely related to an incident are collectively called concurrent factors. They may be present prior to, during, and after an operational error occurred. In addition to the classification of operational errors, factors concurrent to operational errors classified into 32 categories (Appendix C) were used in this study (Corker & Garcia-Chico, 2007). The focus of this study primarily lies in the number of contextual factors in each operational error report. As was the case with each of the operational errors, the presence of multiple factors in a report indicated more complexity, more severe operational errors report, and a greater need for immediate intervention.

Predictor variables. In this study, three independent variables are used: contributory factors, severity conformance metric, and number of aircraft. Independent variables will be referred to as "predictors" in the study.

Criterion variables. The study uses one dependent variable, the sum of different categories of operational error in each report. The dependent variable will be referred to as a "criterion" in the study.

Hypothesis 1

There will be statistically significant overall relationship between the three predictors: (1)contributing factors, (2)number of aircraft, and (3)the severity conformance metric and the criterion (the number of categories of operational error).

Rationale for hypothesis 1. The first hypothesis serves as a preliminary indication of the level of proposed predictors' contribution in explaining the variance in criterion. Essentially, the significant support for the first hypothesis allows further investigation into temporal order of predictors.

Hypothesis 2

After statistically controlling for the number of concurrent factors, the severity conformance metric will contribute significantly to the prediction of the number of the categories in each operational error report.

Rationale for hypothesis 2. When analyzing operational errors in general, particular attention should be devoted to the causal chain of events. Even a seemingly benign set of errors if occurring in a particular order, may result in events that cannot be undone (Ferrante & Pounds, 2003). Therefore, entering predictors in a pre-determined logical order allows for a more precise investigation of temporal precedence of predictors. Controlling for the first predictor in some way standardizes the reports, bringing them to the same level and enables measuring a unique contribution of the following variable, severity conformance metric.

Hypothesis 3

After statistically controlling for the number of concurrent factors and the severity conformance metric, the number of aircraft (under the supervision of an air traffic controller) will help predict the number of the categories in each of the operational error reports.

Rationale for hypothesis 3. A surprising, counterintuitive finding (Rodgers & Nye, 1993) revealed that the number of aircraft and traffic complexity were not significant predictors of operational error severity. However, the severity was defined in terms of horizontal and vertical separation in their study. This study takes a rather different approach. Testing the number of categories of operational error in each report has different results in terms of the importance of the number of aircraft that may be expected.

Method

Participants

The operational error reports for the study were collected and recorded in paper and electronic format by the FAA. Participants in the study were the air traffic controllers from various ARTCC and TRACON centers in the United States. However, for the purpose of this study operational error reports symbolize participants. Due to the sensitive nature of the data, demographic or any other information about the air traffic controllers is not identified in any way. This study was conducted at San Jose State University Human Automation Integration Laboratory (SJSUHAIL). In 2006, after receiving the data from the FAA, the lab analyzed and classified air traffic controllers' operational error reports that occurred during the first six months of 2005 over the US air space. The study was an extension of the previous work that utilized the FAA data from the same period of year 2004. The goal was to further analyze FAA-supplied reports and search for patterns of frequency of occurrence, as well as to identify the causal relationships among the predictors of operational errors.

Materials

The study has used methodical descriptions of operational error incidents that occurred over US airspace between January and June 2005.

Procedure

Complete reports containing operational error reports were reviewed and all potentially relevant indicators of controller's behavior before, during and after the occurrence of the incident were documented. The reports were then divided into

TRACON and ARTCC divisions of centers, forming separate datasets. It is crucial to emphasize that TRACON and ARTCC are two completely independent sources of information. The study therefore is a two-part analysis. In order to provide support for each of the hypotheses, two distinct types of analyses were conducted. For the first hypothesis, standard multiple regression analysis was implemented to gain a general understanding of the contributions of the predictors of the variance in the criterion. In order to test the temporal order of influence of the predictors stated in the remaining two hypotheses, a hierarchical multiple regression analysis was performed.

There are two independent sets of results, TRACON and ARTCC datasets. Both use standard and hierarchical multiple regression analyses. Study 1 describes TRACON and Study 2 describes ARTCC results respectively.

Data Organization

Based on the FAA definitions of the categories of the operational errors, researchers analyzed reports and different concurrent factors and classes of operational errors were identified. It is important to mention that a single operational error report could contain one or more categories of operational error or concurrent factors. Regardless of the nature of a particular operational error or concurrent factor, a higher number would represent a higher occurrence of complexity in airspace and is therefore examined in more detail. Furthermore, the operational error reports included additional information. Time on duty when an incident occurred, training information, and location are some of the examples.

Results

This study represents two independent analyses of two very different kinds of airspace, TRACON and ARTCC. The number of observations that entered analysis for the TRACON was 254 (N=254) and for the ARTCC the number was somewhat larger (N=337).

Analysis 1 TRACON

To gain a general sense of the relationships among predictors (as well as the relationships between the predictors and the criterion) Table 1 is presented here.

Table 1

Means, Standard Deviations, and Correlations among Variables for TRACON

Variable Description	*M*	*SD*	*1*	*2*	*3*	*4*
1. Categories of Operational Error	1.70	1.26		.48**	-.33**	.02
2. Number of Concurrent Factors	1.48	1.58			-.23**	.06
3. Severity Conformance Metric	78.19	22.45				-.07
4. Number of Aircraft	6.00	2.15				

Note. Listwise exclusion (N=254). $^*p<.05$. $^{**}p<.01$.

The overall mean for the criterion was relatively low (M=1.70, SD=1.26). The mean for the concurrent factors was similar with somewhat more variability (M=1.48, SD=1.58). The number of aircraft in the sector (M=6.00, SD=2.15) and severity conformance metric (M=78.19, SD=22.45) means were higher.

It should be noted that the criterion (categories of operational error) and concurrent factors were expected to have relatively low means, as well as low minimum and maximum values simply due to the fact that the number of categories of operational error was limited and there are only so many concurrent factors that a controller can encounter in any particular incident. The number of aircraft present when an operational error occurs is somewhat limited as well. The severity conformance metric, however, ranges from zero to well over one hundred allowing for more variance and a higher mean value.

The pearson correlation coefficients are presented in Table 1. The criterion is positively correlated with the concurrent factors (r=.48, p<.001) which is consistent with the study's expectations. However, the number of categories of operational error criterion is negatively correlated with the severity conformance metric (r=-.33, p<001) indicating that the lower the severity conformance metric index, the higher the number of categories of operational error in each report there are. Considering the nature of this variable, with lower score representing higher proximity of the aircraft, the correlation is consistent with the rationale for the study's hypothesis.

Standard multiple regression correlation analysis. To identify the significant predictors of the criterion, a standard multiple regression was conducted as the first analysis. In order to test our first hypothesis all three predictors were entered into the analysis at the same time and the results indicate that they contributed to 28% of the variance in the criterion, R=.53, R^2=.28, F(3, 250)=32.37, p<.001. Next, to assess unique contribution of the predictors, the betas are examined. Table 2 shows that only two predictors are significant. The number of concurrent factors (β=.43, p<.001) is positive, indicating that the higher number of concurrent factors leads to the higher number of categories of operational error. Severity conformance metric shows negative significance with the criterion (β=-.23, p<.001) meaning that the lower value of the metric leads to the higher number of the categories of operational error. The number of aircraft predictor is non-significant (β=-.02, p>.68) showing that the number of aircraft was not a good predictor of the number or categories of operational error in each report.

Table 2

Standard Multiple Regression for TRACON

Variable Description	B	SE B	β
Number of Concurrent Factors	.34	.04	.43**
Severity Conformance Metric	-.01	.00	-.23**
Number of Aircraft	-.01	.03	-.02

Note. *p<.05. **p<.01.

Hierarchical multiple regression correlation analysis. The second hypothesis is that severity conformance metric has a significant predictive value after the number of concurrent factors variable is controlled for. The third hypothesis proposes that after statistically controlling for the number of concurrent factors and the severity conformance metric, the number of aircraft adds to the prediction of the criterion. The hierarchical multiple regression (Table 3) was conducted to test the second and the third hypotheses.

Table 3

Hierarchical Multiple Regression for TRACON

Variable Description	B	SE B	β
Step 1			
Number of Concurrent Factors	.38	.04	.48**
Step 2			
Conformance Severity Metric	-.01	.00	-.23**
Step 3			
Number of Aircraft	-.01	.03	-.02

Note. $*p<.05$. $**p<.01$.

In step one of the regression analysis, concurrent factors accounted for 23% of the variance in criterion, R=.48, R^2=.23, R^2_{adj}=.23, $F(1, 252)$=74.98, $p<.001$. Coefficient beta was significant (β=.48, $p<.001$) continuing to show that the larger number of concurrent factors per report leads to the larger number of the categories of operational error.

In step two, the severity conformance metric was entered in the regression analysis. As predicted by the second research hypothesis, the variable added a significant amount of additional variance (5%) in the criterion, ΔR^2=.05, $F(1, 251)$=17.39, $p<.001$. However, the β value, although significant, is negative (β=.-23, $p<.001$). That finding is consistent with the nature of the variable. The lower the value of the severity conformance metric is, the higher the proximity of the aircraft conflicts.

In the third and final step of the regression analysis, after controlling statistically for concurrent factors and the severity conformance metric variables, the number of aircraft predictor was entered. Contrary to our hypothesis, adding another variable did not significantly contribute to an explanation of the variance in the criterion, ΔR^2=.00, F(1,250)=.17, p>.05. The β value was not significant (β=.-02, p>.68). This finding is rather surprising and further research and analysis is necessary to show if the number of aircraft in each incident in TRACON is indeed not significant. It seems almost counterintuitive that the increased capacity of the sector is not relevant, because the first reaction would be to believe that at some point a controller overwhelmed with a large number of aircraft would have to make an error simply because of their human cognitive limitations. However, the data in the study show the opposite.

Analysis 2 ARTCC

Descriptive statistics and a correlation table help in obtaining an overview of the relationships among the predictors and the criterion (Table 4). The overall mean for the criterion was (M=2.59, SD=1.11). The means for the concurrent factors are similar with a little more variability (M=2.34, SD=1.62). The number of aircraft in the sector (M=9.55, SD=3.84) and the severity conformance metric (M=74.87, SD=17.71) were somewhat higher.

Table 4

Means, Standard Deviations, and Correlations among Variables for ARTCC

Variable Description	*M*	*SD*	*1*	*2*	*3*	*4*
1. Categories of Operational Error	2.59	1.11		.19**	-.03	.00
2. Number of Concurrent Factors	2.34	1.62			-.05	.15
3. Severity Conformance Metric	74.89	17.71				-.09
4. Number of Aircraft	9.55	3.84				

Note. Listwise exclusion (N=337). *p<.05. **p<.01.

The pearson correlation coefficients are presented in Table 4. Criterion is mildly positively correlated with concurrent factors (r=.19, p<.001) which is consistent with the study's expectations. The number of categories of operational error criterion is not significantly correlated with the severity conformance metric (r=-.03, p>.05). The number of aircraft and the criterion are not correlated (r=-.00, p>.05).

Standard multiple regression correlation analysis. To provide support for the hypotheses and to determine the extent to which the independent variables serve as good predictors of the number of categories of operational error, a standard multiple regression analysis was conducted as the preliminary investigation. To test the first hypothesis, all three predictors were entered into the analysis at the same time and the results indicate that these predictors contribute to only 4% of the variance in the criterion, R=.20, R^2=.04, F(3, 333)=4.46, p<.01. In order to assess the unique contribution of the predictors, the β values are examined. The Table 5 shows that only one predictor was significant.

Table 5

Standard Multiple Regression for ARTCC

Variable Description	B	SE B	β
Number of Concurrent Factors	.14	.04	.20**
Conformance Severity Metric	-.00	.00	-.03
Number of Aircraft	-.01	.02	-.04

Note. *p<.05. **p<.01.

The number of concurrent factors (β=.20, p<.001) is positive, indicating that the high number of concurrent factors leads to the high number of categories of operational error. Severity conformance metric predictor (β=-.03, p>-.48) and number of aircraft (β=-.4, p>.52) are non-significant.

Hierarchical multiple regression correlation analysis. The second hypothesis was that the severity conformance metric has a significant predictive value after the number of concurrent factors variable is controlled for. The third hypothesis was that after statistically controlling for the number of concurrent factors and the severity conformance metric, the number of aircraft adds to the prediction of the criterion. The hierarchical multiple regression was conducted to test both hypotheses. As shown in Table 6, the three predictors were separately entered into the analysis.

Table 6

Hierarchical Multiple Regression for ARTCC

Variable Description	B	SE B	β
Step 1			
Number of Contributing Factors	.13	.04	.19**
Step 2			
Conformance Severity Metric	-.00	.00	-.02
Step 3			
Number of Aircraft	-.01	.02	-.04

Note. $^*p<.05$. $^{**}p<.01$.

In step one of the regression analysis concurrent factors accounted for 4% of the variance in criterion, $R=.19$, $R^2=.04$, $R^2_{adj}=.03$, $F(1, 335)=12.83$, $p<.001$. Coefficient beta was significant ($\beta=.19$, $p<.001$). This finding indicates that reports with a large number of concurrent factors lead to the reports with a large number of categories of operational error.

In step two, severity conformance metric was entered in the regression analysis. Contrary to the second research hypothesis, the variable did not add significant amount of additional variance in the criterion, $\Delta R^2=.00$, $F(1, 334)=.18$, $p>.05$. The beta coefficient was not significant ($\beta=-.02$, $p>.67$).

In the third and final step of the regression analysis, after controlling statistically for concurrent factors and severity conformance metric variables, number of aircraft predictor was entered. This step did not confirm the third hypothesis, as adding another variable did not significantly contribute to an explanation of the variance in the criterion, ΔR^2=.00, F(1,333)=.42, p>.05. Beta coefficient was not significant (β=-.04, p>.52). This finding is rather surprising and further research and analysis is necessary to show that number of aircraft in each incident in ARTCC is not a significant predictor of the number of categories of operational error.

Discussion

The overarching goal of the study was to continue the exploration of operational errors and circumstances they occur in. More precisely, the study searches for possible predictors that may influence the frequency of occurrence of operational error categories per incident and possibly their temporal order. Additionally, this study will contribute to the exploration of multiple regression correlation statistical method as a possible tool when analyzing operational errors. That method is used to develop relationships between variables collected by the FAA and will help in accident investigation. The usefulness of the novel statistical approach has three aspects. The first is that the multiple regression has not been used before in operational error analysis. The second is that it enables developing of a causal model for operational errors. The third aspect, depending on the results of the analysis, is focusing the FAA's attention to only significant predictor variables of operational errors.

The first hypothesis stated that there would be statistically significant overall relationship between the three predictors (concurrent factors, number of aircraft, and the severity conformance metric) and the criterion, number of categories of operational error. The first analysis, standard multiple regression, was conducted to test the hypothesis. The results confirmed the prediction almost entirely. For TRACON analysis, two predictors, concurrent factors and conformance severity metric were statistically significant explaining a great portion of variance in the dependent variable, while the number of aircraft was not. In ARTCC analysis, however, all predictors account for a small percentage of variance. Only one predictor, the number of concurrent factors per report, was statistically significant. Therefore, the difference is two-fold: the amount of variance explained by all three predictors and the lack of statistical significance of severity conformance metric variable in ARTCC part of the study. The difference in variable contribution between ARTCC and TRACON is an important finding. Usually, the FAA collects a multitude of variables considering the obtained data as predictive of operational errors. If, for example, there is only 4% of variance accounted for, as in the part of this study, FAA may consider shifting the focus from all variables to the most significant ones only.

Possible explanation of the discrepancy between TRACON and ARTCC results may lie in the different minimum separation criteria between the two types of airspace control and that is exactly what severity conformance metric variable is based on. It is plausible that ARTCC having larger separation minimum than TRACON simply does not account for enough variance because of the nature of some of the categories of operational error that by definition have greater likelihood of happening in TRACON only. There seems to be a difference in airspace operations between ARTCC and TRACON because variables are differentially predictive for the two types of airspace. The TRACON requires three miles and 1000 ft minimum separation, while ARTCC is much larger in volume of airspace, and allows five miles laterally and 2000 ft vertically. Further, the TRACON operations require constant changes and adjustments and are in general more dynamic. Finally, TRACON more often deals with altitude-related categories of operational error and it is possible that there are more of those categories present in the TRACON dataset. For example, a category of the operational error named "altitude inadequate" is more likely to occur frequently in TRACON than in ARTCC airspace. That does not mean severity conformance metric should be completely abandoned as an indication of the number of categories of operational error in ARTCC. The difference in sample size (337 for ARTCC compared to 254 for TRACON), however, could not have been something that led to the discrepancy in the results. Although having smaller number of observations, TRACON accounted for more variance than ARTCC. Practical application of this finding is the reduction in man hours needed for data collection. Currently, all data is collected and all variables treated as having the same relevance. Identifying the most important predictors could shorten the time needed for data-collection, thus saving resources as well. This new "Data Economy" approach would mean considering and collecting the data for relevant variables only. The following two hypotheses tested for possible temporal order of the predictors. Hierarchical multiple regression analysis was used to control the order of entry of variables specified in the hypotheses.

The second hypothesis stated that after statistically controlling for the number of concurrent factors, severity conformance metric will contribute significantly to the prediction of the number of the categories in operational errors. The TRACON analysis confirmed the hypotheses and after controlling for the number of concurrent factors, adding severity conformance metric explained small but statistically significant amount of additional variance. However, in ARTCC analysis, the results were negative. These findings are similar to the first part of the study and the explanation is quite similar.

Sample size could not possibly be the issue, because there are more reports for ARTCC collected.

The third hypothesis was that after statistically controlling for the number of concurrent factors and severity conformance metric, adding the number of aircraft in the sector (under the supervision of an air traffic controller) would help predict the number of the categories in operational errors. The data (TRACON and ARTCC) do not show support for this hypothesis. These findings are consistent with previous research that number of aircraft is not a significant predictor of the severity of operational error (Rodgers & Nye, 1993). It is especially surprising in this study, because other two variables have been statistically controlled for prior to adding number of aircraft variable. The sample size could not be one of the factors in this case, especially not in ARTCC that shows larger sample size and where we see a greater number of aircraft in each incident occurrence than in TRACON (see tables 1 and 4). Another alternative explanation for the lack of number of aircraft variable significance could be that only three predictor variables were used and there may be additional factors that were not taken into account that mediate relationships between predictors and criterion.

Finally, some suggestions for future research based on the results of this study deserve to be mentioned. Most importantly, a novel approach of using multiple correlation regression analysis used in this study has allowed for a more diverse approach in operational error research. Future work could incorporate more predictors in search of a causal pattern of operational error precursors. Essentially, this is the beginning of a modeling process for relating error classification to error complexity. In order to substantiate this approach, more relevant data that would contribute to more significance of the results would need to be collected and analyzed. One way to materialize that could be the multivariate approach used on the archival data spanning over several years.

References

Bailey, L., Schroeder, D. & Pounds, J. (2005). *The Air Traffic Control Operation Errors Severity Index: An Initial Evaluation* (DOT/FAA/AM-05/5) Washington, DC: Office of Aerospace Medicine.

Corker, K. & García -Chico, J. L. (2006). Human Factors Characterization of Operational Errors in ATC. The Case Study of TCAS Alerts. In F. Reuzeau, K. Corker, & G. Boy (Eds.) *Proceedings of International Conference on Human-Computer Interaction in Aeronautics* (pp. 152-155). Seattle, WA.

Corker, K., & Garcia-Chico, J. L. (2007). An Analysis of Operational Errors and the Interaction with TCAS RAs. *Paper submitted to the 7th USA/Europe Seminar on Air Traffic Management Research and Development. Spain: Barcelona.*

Davis, C. G., Danaher, D. Q., & Fischl, M. A. (1963). *The influence of selected sector characteristics upon ARTCC controller activities* (FAA/BRD-301). Arlington, VA: Federal Aviation Administration.

Federal Aviation Administration. (1996). *Crewmember Qualifications* (14CFR/121 Subpart O). Washington, DC: Department of Transportation.

Federal Aviation Administration. (2000). *Air Traffic Control* (Order 721011.65L). Washington, DC: Author.

Federal Aviation Administration. (2002). *Air Traffic Quality Assurance* (FAA Order 7210.56c). Washington, DC: Author.

Ferrante, A. & Pounds, J. (2003). *FAA Strategies for Reducing Operational Error Causal Factors* (DOT/FAA/AM-03/19). Washington, DC: Office of Aerospace Medicine.

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

Rodgers, M. D. (1998). The Relationship of Sector Characteristics to Operational Errors. (DOT/FAA/AM-98/14). Washington, DC: Federal Aviation Administration.

Rodgers, M. D., & Nye, L. G., (1993). *Factors associated with the severity of operation errors at air route traffic control centers* (DOT/FAA/AM-93/22). Washington, DC: Federal Aviation Administration.

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

Shappell, S.A., & Wiegmann, D.A. (2001). *A Human Error Analysis of Commercial Aviation Accidents Using the Human Factors Analysis and Classification System (HFACS).* Washington, DC: Federal Aviation Administration Office of Aviation Medicine. DOT/FAA/AM-01/3.

Shappell, S.A. & Wiegmann, D.A. (2001). Air Traffic Control (ATC)- Related Accidents and Incidents: a Human Factor Analyses. International Symposium on Aviation Psychology, Columbus, OH: The Ohio State University.

Shappell, S.A. & Wiegmann, D.A. (2001). Human Error Perspectives in Aviation. *The international journal of aviation psychology, 11*(4), 341-357.

Wickens, C. D., Mavor, A. S., & McGee, J. P. (1997). *Flight to the future*: Human factors in air traffic control. Washington, DC: National Academic Press.

Appendix A

Operational Error Report Example

N90-R-04-E-003: CPC providing OJT on combined satellite positions in the EWR area for 0+05 minutes responsible for 9 aircraft, no previous OE/D. The DEV is certified on the majority of the positions and the positions were combined to add to the complexity to determine if the DEV was ready for a certification. N90KC/WW24 was level at 3,000 feet heading southbound for a VOR alpha approach into TEB. N628CC/F2TH was east bound descending out of 4,000 for 3.000 feet overtaking the West-wind. The DEV turned N90KC to a 130 heading to intercept the final and turned N628CC south bound for base leg. The OJTI was aware of the conflict and was coaching the DEV, however when the DEV issued the last turn, the OJTI took over the frequency but was too late to maintain separation. Closest proximity 400 feet vertical, 2.0 miles lateral Moderate Severity Category "C." Required 1,000 feet or 3 miles. Staffing, 2 OS's assigned, 1 on break, 14 CPC's 3DEV's on duty, 6 CPC's on break, 1 CPC on OT.

In this report, a novice (developmental controller, DEV) fails to notice two aircraft on a collision course. The expert (certified controller, OJTI) notices them but did not have time to intervene to prevent the loss of separation.

Appendix B

Operational Error Categories List

Operational Error (OE) Category	Definition
Aircraft overlapped:	The label or aircraft symbols were overlapped on the radar screen, preventing the controller to perform adequately his/her duties and identify the conflict.
Airspace violation:	An aircraft penetrated airspace that was delegated to another position of operation or another facility without prior coordination and approval. Also, if controller issues a clearance that violates a restricted area, e.g., military, Minimum Vectoring Area (MVA).
Altitude inadequate:	New issued flight level coincides with the conflicting aircraft's flight level. The vertical paths do not cross each other, but both aircraft are flying towards the same altitude or one of them flying towards the other's altitude. Flight level change is given by the current controller.
Cleared below minimum altitude:	Aircraft cleared below minimum allowed altitude (MVA). This is always coupled with airspace violation, but not vice versa. The altitude change had to be given by the current controller.
Climb through:	The aircraft is cleared to climb to a flight level that crosses other aircraft's path. The current sector controller issues the clearance for the level change. Changing flight level to the same altitude that other aircraft has is considered "Altitude Inadequate".
Control coordination:	Failure to correctly coordinate (incomplete or absence of coordination between controllers) either inter- or intra- facilities, e.g., dismissing coordination or passing incomplete information.
Datablock misentered information:	The controller enters wrong information, e.g., he or she issued a flight level or heading and he or she entered a different one.
Descend through:	The aircraft is cleared to descend to a flight level that crosses other aircraft's path. The current sector controller issues the clearance for the level change. Changing flight level to the same altitude that other aircraft has is considered "Altitude Inadequate".

Operational Error (OE) Category	Definition
Fail converging:	Failure to detect a conflict between aircraft with converging horizontal routes. Both aircraft need to have converging headings issued in other sectors or by the CPC (certified professional controller) in control but vary in advance to the conflict. That is, the controller fails to detect the convergence. It is possible to have initially both aircraft at different levels.
Fail identification altitude climbing/ descending:	Failure to identify a conflict when one of the aircraft is descending or climbing, but the clearance was issued by other sector. If it was issued by the current sector, it is classified under "climb through", "descend through" or "altitude inadequate".
Fail overtaking traffic:	Failure to separate overtaking traffic. It includes violations of standard separation when aircraft is caught behind the other aircraft with the same course or when the CPC gives a vector to locate the aircraft in a parallel path or to resume heading (not including parallel tracks issued in other sector).
Flight Paper Strip (FPS) misentered information:	The controller writes wrong information on the flight paper strip or omits subsequent revisions, e.g., he or she issued a flight level or heading and he or she entered a different one on the flight paper strip.
Hear/readback:	The pilot reads back incorrectly the instruction given by the controller, and it goes undetected by the controller. It includes no pilot readback errors too.
Instruction not intended:	Controller issues an instruction he/she stated as non intended (e.g., slips of tongue when giving a flight level, speed).
Letter-of-agreement (LOA) misapplication:	Misapplication of letter of agreement between sectors. This is a particular case of "misapplication of procedure". If there is a letter of agreement to apply and the controller did not do it.
Misapplication of procedure:	Controller applies a procedure that is not longer valid or it is not possible to apply under current circumstances.
Misread information:	Controller reads wrong information; either he or she reads a wrong field or states reading incorrect information in the report. It does not include short term failures when retrieving information.

Operational Error (OE) Category	Definition
Other:	Other situations/errors not mentioned in the list before.
Overlooked traffic:	Controller misses or ignores the traffic situation for some reason (e.g., distraction, preoccupation, other conflict). It should be clearly stated in the operational error description that the controller overlooked the traffic.
Speed inadequate:	Speed issued is not adequate to prevent loss of separation. It might be misestimated, or coupled with other operational errors, such as issuing clearance at a wrong time.
Temporal error-issue:	Failure to issue timely the instruction, i.e., the instruction is correct, but it is not issued in time to prevent loss of separation. Operational error is only categorized as temp error-issue if this fact is written in the report.
Transpose information:	Alter the sequence of saying a clearance for two different aircraft, mixing callsigns, numbers, symbols and/or instructions. The confusion involves two aircraft that exits in the sector (or one was recently handed off), or same airline but wrong number (e.g., AA1345 instead of AA1365, which is flying as well).
Vector inadequate:	Controller is giving an incorrect or inaccurate heading (vector) to the aircraft that creates the conflict. It is sometimes coupled with "fail overtaking".
Wrong aircraft:	The aircraft callsign used is wrong. There is no confusion among aircraft within the sector, but the controller uses a callsign that does not either exist or is not in the sector.

Appendix C

Concurrent Factors Categories List

Contextual Factor	Definition
Climb slow/fast:	Aircraft did not climb as quickly or climbed faster than controller anticipated.
Combined sectors:	Controller was working with combined sectors. This does not mean merging both R-side and D-side (controllers) roles in one position.
Descend slow/fast:	Aircraft did not descend as quickly or descended faster than controller anticipated.
Distraction:	Some event attracted controller's attention and diverted him/her from the evolving conflict.
D-side absent:	D-side controller was not on duty and R-side controller assumed responsibilities of both.
D-side coordination:	D-side coordinated with other sectors a change in the aircraft profile, which conflicted with other aircraft in the vicinity. CPC (certified controller) was not able to separate, either because of lack of internal coordination or lapse on the application of further measures.
Flight plan not checked:	Controller was not aware of the flight plan and aircraft made an "unexpected" maneuver (it is unexpected because the controller is not aware, but it is compliant with the flight plan).
Inadequate relief briefing:	Inadequate position relief briefing. Either coordination or communication problem arises as the result of the position relief (incomplete or erroneous briefing). It must be explicitly stated.
Lapse in coordination:	Controller forgot completely or partially to coordinate with the adjacent sector.
Lapse in delivery clearance:	Controller forgot delivering a clearance s/he intended.

Contextual Factor	Definition
Mishearing:	Either pilot or controller misheard a voice message.
Misjudgment:	The maneuver was miscalculated or misjudged by the controller, because the complexity of the traffic or distraction. It is present when the operational error is related with timing.
Misread information:	Some misread information triggered the sequence that led to the error.
Multiple aircraft:	A conflict involving several aircraft, or the operational error was created due to a cascade effect when avoiding a conflict elsewhere.
No pilot response/deviation:	There was no pilot's response to a controller request or his/her action did not follow the corresponding clearance, if any (e.g., due to a pilot hear back error). This includes encounters where a different aircraft is responding to a controller's call when the clearance was not addressed to him/her.
Not enough information:	Not enough information to identify any factor relevant at the time of the error.
OS absence/CIC present:	OS (operating supervisor) was not in the room (e.g., on a break), he/she was engaged in other tasks at other sector, or working a radar position at the same time. The report states clearly the circumstances. It also includes the cases when there is a CIC (controller in charge).
Other complexity factors:	Traffic situations that create unusual conditions and therefore increase complexity of controller operations, which are not gathered under other category (e.g., MIT or military operations).
Other:	Other factors identified.
Pilot/Other controller requests a conflicting clearance:	Pilot requested a conflicting clearance that was critical to lose separation, and controller did not realize.

Contextual Factor	Definition
Point out aircraft:	One controller points out an aircraft to the other sector. It is consider that the traffic situation is more complex.
Recent combination/ decombination:	Sectors were recently combined/re-combined. "Recently" is considered within five minutes.
Slip in clearance:	Slip on delivering, the flight level, the callsign or the heading.
Stuck mike:	The channel or the microphone failed in the transmitter mode because other microphone blocked the frequency. There was no transmission of the message.
TCAS maneuver:	TCAS (Traffic alert and collision avoidance system), present in the operational error, might have contributed to the conflict and/or increased the complexity. Not all TCAS are considered in this category.
Temporal airspace complexity:	Operational error occurs with some unusual airspace configuration that might have impacted on the complexity. (e.g., change of runway configuration due to weather, or use of unusual airways due to high traffic volume).
Traffic complexity:	Traffic complexity rating. This is rated by controllers, but because usually the summary report does not have this information available, traffic is considered complex if the reports states the MAP (monitor alert parameter) number and the number of aircraft in the sector is higher than 2/3 of the MAP.
Training in progress:	There was training being conducted at the position where the error took place.
Transmission stepped on:	Here the transmission was partially-blocked (stepped on). One part of the message was audible, and the other part was not.
Turbulence complexity:	Turbulence was created due to another aircraft presence.
Turn wide/sharp/early/late:	Aircraft turn was wider/sharper/was initiated earlier/later than the controller anticipated.

Contextual Factor	Definition
Weather complexity:	Weather or conditions caused by weather increase the complexity of operations, (e.g., route deviation, closure of a runway, icing affecting climb).