12-2015

Review of Aviation Research: A Content Analysis of Articles Published in the Collegiate Aviation Review, 2007–2012

Safak Aktemur
Florida Institute of Technology

Review of Aviation Research:

A Content Analysis of Articles Published

in the *Collegiate Aviation Review*, 2007–2012

by

Safak Aktemur

Bachelor of Arts
Aviation Management
Anadolu University
2011

A thesis submitted to the College of Aeronautics at
Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Airport Development and Management

Melbourne, FL
December, 2015

We the undersigned committee
hereby approve the attached thesis

A Content Analysis of Articles Published
in the *Collegiate Aviation Review*, 2007–2012

by

Safak Aktemur

Michael A. Gallo, Ph.D.
Associate Professor
College of Aeronautics
Committee Chair

William Rankin, Ph.D.
Associate Professor
College of Aeronautics

M. Mine Subasi, Ph.D.
Assistant Professor
Mathematical Sciences

Stephen Rice, Ph.D.
Associate Professor
Graduate Program Chair
College of Aeronautics

ABSTRACT

TITLE: A Content Analysis of Articles Published in the *Collegiate Aviation Review*, 2007–2012

AUTHOR: Safak Aktemur

MAJOR ADVISOR: Michael A. Gallo, Ph.D.

The purpose of the study was to conduct a content analysis of the methodological quality of articles published in the *Collegiate Aviation Review* (*CAR*)— a refereed journal of the University Aviation Alliance (UAA)—to determine if actual practices of aviation researchers were consistent with commonly recommended research methods and procedures. The accessible population consisted of the 76 articles published in *CAR* between 2007 and 2012 (Volumes 25–30). The sample consisted of $N = 69$ articles and excluded literature reviews, meta-analyses, studies that described the development or validation of an instrument, philosophical inquiries, position papers, or historical studies. The analysis focused on the fundamentals of research principles, measurement, and data analysis procedures including the extent to which authors gave attention to describing: purpose statements and research questions; sampling issues such the target and accessible populations, sampling strategy, sample representativeness, and sample size determination; instrumentation validity and reliability; research methodology and design; threats to internal validity; data analysis procedures; conclusions and recommendations; and limitations and delimitations. Using a

coding form with a set of predetermined categories that corresponded to these methodological issues, two coders coded the articles independently and interrater reliability was established using percent agreement.

Major findings included the following: (a) the majority of articles contained a purpose statement, but half did not include corresponding research questions; (b) about half of the articles did not contain information about the population, the most frequently used sampling strategy was convenience, more than half of the articles did not describe the sample, and nearly 90% of the articles did not address sample representativeness; (c) two-thirds of the articles did not give attention to instrumentation validity and reliability; (d) survey was the most common research methodology; (e) nearly 90% of the articles did not discuss at least one internal validity threat; (f) the most commonly used statistical procedures were descriptive; (g) only 13% of the articles gave attention to population generalizability; (h) 90% of the articles expressed conclusions by restating the study's findings; and (i) two-thirds of the articles did not specify any study limitations or delimitations. The findings indicate that the methodological quality of articles published in *CAR* for the targeted 6-year period should be of concern to the aviation research community, particularly to the editors, authors, and readers of *CAR*. The lack of thoroughness with respect to methodological quality affects both generalizability and replication studies. The reader is cautioned not to overgeneralize the findings because they apply only to the targeted articles of *CAR* published in 2007–2012.

Table of Contents

List of Tables

Chapter 3

Chapter 4

Acknowledgement

I am grateful to my thesis advisor, Dr. Michael Gallo, whose expertise and continuous support made it possible for me to complete my thesis. Dr. Gallo gave me a lot of his time and his feedback proved to be very beneficial. I also thank Dr. William Rankin and Dr. Munevver Subasi for accepting to be on my committee. I really appreciate my family and friends for being there for me even during the most difficult time of writing this thesis.

Dedication

This thesis is dedicated to the memory of my late father, Muhsin Aktemur, whom I still miss everyday. Without his inspiration, none of this would happen. I also want to dedicate this thesis to my mother Herdem Aktemur and my sister Dilek Aktemur, who gave me all the love and encouragement for my success.

Chapter 1

Introduction

Background and Purpose

Background. Content analyses of journal articles have been conducted over the past 30 years within the educational research domain to determine if actual research practice is consistent with recommended methods and procedures. Shaver and Norton (1980a) examined articles from the *American Educational Research Journal* (*AERJ*) to determine the extent to which researchers were using random samples from defined populations, and to determine if researchers were limiting their conclusions according to the sampling techniques employed. Shaver and Norton (1980b) conducted a similar content analysis of two social studies journals, *Theory and Research in Social Education* (*TRSE*) and *Social Education* (*SE*) with similar findings. Eight years later, Wallen and Fraenkel (1988a, 1988b) conducted a follow-up study of social studies education research. They reported that their findings were essentially the same as those reported by Shaver and Norton (1980a, 1980b).

Horton et al. (1993) examined the methodological quality of research articles published in the *Journal of Research in Science Teaching* (*JRST*) from 1985 through 1989. This analysis identified the: study type by methodology such as experimental, correlational, and causal comparative; the extent to which authors specified the purpose of their study; sampling and group membership issues;

qualifications of conclusions, and threats to internal validity. Horton et al. concluded, "The results of this study indicate that the methodological quality of published science education research should remain a concern for both practitioners and readers" (p. 857).

Complementing Horton et al. (1993), Hsu (2005) conducted a content analysis of research studies published in the *American Educational Research Journal* (*AERJ*), *Journal of Experimental Education* (*JEE*) and *Journal of Educational Research* (*JER*) from 1971 to 1998. Hsu focused on the frequency of subject matters, research methods/designs, data analysis procedures, and corresponding trends among these three journals. Hsu selected up to 24 articles per year per journal for each of the targeted years, and identified 18 subject matter categories, 30 research methods/designs categories, and 34 data analysis procedures categories. To assure that the same criteria were used in classifying the articles, a manual describing the characteristics of each category was created and a group of graduate students were recruited and trained to carry out the coding procedure.

Hsu (2005) reported that four of the five most frequently investigated subjects matters were identical for the three journals: psychology in education, teaching/instruction, teachers, and measurement. With respect to the most frequently used methods/designs, Hsu reported that four of the frequently used methods that were identical for all three journals were comparative approach, descriptive research, quasiexperimental, and survey. Hsu also reported that five of

the six frequently used statistical procedures were identical among the three journals: descriptive statistics, ANOVA, bivariate correlation, *t* test, and regression.

Content analyses also have been the focus of dissertation research studies in aviation. For example, complementing the qualitative approaches she applied to examine peoples' perspectives, opinions, beliefs, and attitudes, Garner (1992) used content analysis as part of her dissertation to examine safety themes presented in six newspaper articles about the Delta 1141 crash at Dallas-Fort Worth Airport in 1988. Garner found "the stories were a mix of voyeurism, affirmation and reassurance, focusing on death and destruction while affirming the safety of flying and the eventual solution to the mystery of the crash" (p. 1). Hankins (2007) examined the characteristics of aviation baccalaureate programs reported in various collegiate publications. The purpose of his analysis was to study the quality indicators, or characteristics, of nonengineering aviation baccalaureate programs. These characteristics included curriculum, students, faculty, program activities, equipment, facilities, leadership, resources, reputation, and value. Hankins applied these indicators to the catalogues and online marketing materials for aviation programs at 72 U.S. colleges and universities. Hankins found "…there is wide variation among schools in both the degree, as well as the way, in which quality characteristics are displayed to prospective students and stakeholders" (p. 115). For example, some schools chose to present information about their programs online using sophisticated imbedded video files and other multimedia presentations. Other

schools, though, relied primarily on written brochures or catalogs complemented by simple and sparsely populated web pages. Hankins also reported that although every school had positive scores for curriculum, which was considered the most important quality characteristic, more than half of the schools failed to exhibit a score greater than 50% for leadership, which was considered the fourth most significant quality characteristic.

More relevant to the current study, Bliss (2012) conducted a content analysis of all 189 articles published from 1983 to 2010 in the *Collegiate Aviation Review* (*CAR*), which is a refereed journal published by the University Aviation Alliance (UAA). The purpose of his analysis was to provide demographic information relative to the content and contributions of the articles published. This included the total number of articles published in each volume, the aggregate number of pages of each article, the number of authors, authors' institutional or organizational affiliation, article subject classification, and geographical locations partitioned by regions with respect to which region submitted the most or fewest articles. Bliss reported that universities and colleges accounted for 97% of *CAR* articles, with very few nonacademic organizations having published in *CAR* during the targeted period. Bliss also indicated most articles concentrated on collegiate flight training, collegiate aviation degree programs, collegiate aviation students, airport systems including air traffic control, and the commercial airline industry. Absent from his analysis, though, was information about the methodologies

described in the articles, or the extent to which sound research practices were incorporated within the studies.

Purpose. Although content analyses have been conducted in a wide range of areas and applications including aviation, none have examined aviation research articles published in refereed journals from a methodological quality perspective. The purpose of the current study was to conduct a content analysis of articles published in *CAR* from 2007 to 2012 to determine if actual practices of aviation researchers were consistent with commonly recommended research methods and procedures. This assessment focused on the fundamentals of research principles, measurement, and data analysis procedures. As reflected in the coding form (see Appendix A), this included: purpose statements and research questions; sampling issues, including specifying the target and accessible populations, sampling strategy, sample representativeness, and sample size determination; instrumentation issues, including attention given to validity and reliability; research methodology and design; attention to internal validity; data analysis procedures; conclusions and recommendations; and limitations and delimitations.

The reason *CAR* was selected was because Johnson, Gibson, Hamilton, and Hanna (2006) identified *CAR* as one of the three most important peer-reviewed journals in aviation education (the other two were *Journal of Aviation/Aerospace Education & Research* and the *Journal of Air Transportation*). The reason for targeting the period 2007–2012 was because (a) *CAR* began publishing articles

biannually (spring and fall) in 2007, and (b) the number of articles available for review ($N = 76$) during the 6-year period seemed sufficient to identify patterns relative to methodological quality.

Definition of Terms

Key terms and phrases used in the current study were operationally defined as follows:

1. *Data analysis procedures* were defined as both quantitative and qualitative approaches used for analyzing numerical or contextual data. Quantitative procedures included descriptive statistics such as mean, median, standard deviation, range, frequencies, and percentages, and inferential statistics such as *t* test, ANOVA, covariance, correlation, regression, chi-square, confidence intervals, and effect size.
2. *Measurement issues* referred to the description of a data collection instrument, including the attention given to instrumentation validity and reliability.
3. *Research principles* referred to all the components commonly associated with methodology and design issues, and included both quantitative and qualitative approaches. These included, but were not limited to: purpose statement, research questions, and hypotheses (if applicable); sampling issues and sample size; type of research methodology/design (e.g., experimental, ex post facto, correlational, survey, historical, content analysis, case study, narrative, grounded theory, phenomenological, and ethnographic); threats to

internal/external validity and their qualitative counterparts of credibility, transferability, dependability, and confirmability; conclusions and recommendations; and limitations and delimitations.

Research Questions

The primary research questions that guided this study were as follows:

1. To what extent do aviation researchers state the purpose of their study and corresponding research questions?
2. To what extent do aviation researchers define their target and accessible populations?
3. What sampling strategies do aviation researchers commonly use?
4. To what extent do aviation researchers describe their sample, including representativeness, sample size, and assignment?
5. To what extent do aviation researchers address instrumentation issues such as validity and reliability?
6. What research methodologies/designs do aviation researchers commonly use?
7. To what extent do aviation researchers give attention to threats to internal validity?
8. What data analysis methods do aviation researchers commonly use?
9. In what way do aviation researchers report their conclusions?
10. To what extent do aviation researchers state their recommendations, limitations, and delimitations?

Study Design

As noted in the purpose statement, the research methodology of the current study was content analysis. This research methodology was appropriate because "content analysis … is a research method applied to written or visual materials for the purpose of identifying specified characteristics of the material" (Ary, Jacobs, & Sorenson, 2010, p. 457). It enables researchers to study human behavior indirectly through an analysis of their communications (Fraenkel, Wallen, & Hyun, 2011). For the current study, I described the prevailing research practices presented in articles published in *CAR* to determine the type of information authors are or are not including in their presentations. I systematically analyzed the methodological quality of these articles to determine if actual research practice is consistent with commonly accepted standards of research as reported in various educational research methods textbooks such as Ary et al. (2010), Gall, Gall, and Borg (2007), and Fraenkel, Wallen, and Hyun (2011).

Significance

The rationale for the current study was grounded in the commonly held belief that a periodic review of common research practices in a scholarly discipline can help guide and facilitate improving such practices. As a result, the current study's primary focus was to inform the aviation research community about the state of the art in articles published in the targeted volumes of *CAR*. Because no similar content analyses have been conducted with respect to aviation research

published in refereed journals, the current study was the first step to determine if the actual practices of aviation researchers are consistent with commonly recommended research methods and procedures. If other content analyses of refereed aviation journals are conducted periodically, then the current study will contribute to the aviation research community for the development of scholarly publications in the area. The researchers planning to publish their studies and the editors of refereed aviation journals can benefit from the findings of the study. In light of the findings of the current study, the *CAR* editors will have the opportunity to criticize their standards toward the quality of articles published in the targeted 6-year period. Also, researchers in the aviation field who endeavor to publish their research will have the opportunity to design their studies relative to the findings of the current study.

Study Limitations and Delimitations

Limitations. A limitation in a research study refers to circumstances, events, or other study conditions that are outside the control of the researcher, but can have an impact on the results and generalizability of a study. In the current study, there were several limitations as follows:

1. *Journal editor influence*. Although the current study determined if actual practices of aviation researchers are consistent with commonly recommended research methods and procedures, it is possible that certain methods and procedures were not reported because of editorial influence. In other words, I

will have no control over what details of a study were removed or restricted by the journal editor, or what details an editor did not require.

2. *Paucity of prior studies.* Because there are not any similar content analyses of *CAR* articles, I was not able to compare the findings of the current study to prior studies. This also limits the generalizability of the study's findings.

Delimitations. A delimitation in a research study refers to circumstances, events, or other study conditions that the researcher imposes on a study to make the study feasible to conduct, but further limits the generalizability of the study's findings. In the current study, there were several delimitations as follows:

1. *Targeted journal*. The current study focused on *CAR* articles and therefore the results are restricted to *CAR*. Thus, a similar content analysis conducted using different aviation research journals might not get similar results.
2. *Targeted years.* The current study targeted *CAR* articles from 2007–2012. Therefore, a similar content analysis of *CAR* articles conducted during a different time period might not get similar results.
3. *Coding form.* The current study analyzed *CAR* articles using a predeveloped coding form (Appendix A). Therefore, a similar content analysis that uses a different coding form might not get similar results.
4. *Personal interpretations*. The findings of the current study are limited to my interpretations. Therefore, the same results might not be found if different researchers coded the articles.

Chapter 2

Review of Related Literature

Introduction

This chapter is organized and presented in two main sections. The first section contains a review of the literature of past content analyses that have been conducted in various contexts, including education and aviation. The last section contains a discussion of the related literature's implications to the current study.

Review of Past Research Studies

As noted in Chapter 1, content analysis as a research methodology is used to analyze various types of media, including written, visual, and audio recordings. The purpose for conducting a content analysis is to identify specific characteristics of the material being analyzed, compare communications, and to determine if there are any trends in the communication content. Weber (1990) noted content analysis provides a systematic approach for making valid inferences from text (p. 9).

Although content analysis as a research methodology has been applied across many disciplines, including sociology, political science, psychology, business, and education, its use within the aviation research community has been limited mostly to aviation safety. For example: (a) Jones and Endsley (1996) analyzed 143 aviation incidents and compared the levels at which the flight crew and pilots made errors; (b) Wiegmann et al. (2005) analyzed over 14,000 general aviation accident records from 1990–2000 to identify aircrew errors; and (c) Garner

(1992) conducted a content analysis as part of her dissertation research that examined the safety themes presented in various newspaper articles surrounding the Delta 1141 crash at Dallas-Fort Worth Airport on August 31, 1988. Independent of aviation safety, Hankins (2007) conducted a content analysis to examine the quality indicators, or characteristics, of non-engineering aviation baccalaureate programs reported in college catalogs and related collegiate publications including online marketing materials. These aviation-related content analyses, however, are neither appropriate nor relevant to the current study.

After an exhaustive literature review, I found three articles that were not only appropriate and relevant, but they also helped guide the current study. The first two articles were from the education field and involved content analyses of articles published in refereed educational journals. The third article was a content analysis of articles published in *CAR*, but focused on article demographics and not methodological quality. A discussion of each article follows.

Horton et al. (1993). Horton et al. conducted a content analysis of articles published in the *Journal of Research in Science Teaching* (*JRST*) from 1985–1989. Their primary objective was to determine if actual practice, as manifested by the published articles, was consistent with what is regarded as commonly accepted research methods and procedures. Horton et al. patterned their study after previous content analyses by Shaver and Norton (1980a, 1980b) who examined the methodological quality of articles published in the *American Educational Research*

Journal (*AERJ*), *Theory and Research in Social Education* (*TRSE*), and *Social Education* (*SE*). Horton et al. also reported that Wallen and Fraenkel (1988a, 1988b) followed up Shaver and Norton's studies, but no such content analysis had been conducted with respect to articles published in *JRST*.

Horton et al. (1993, p. 858) posited five research questions that focused on the extent to which science education researchers (a) selected their samples randomly from defined and/or described accessible population, (b) defined their target populations and describe their samples, (c) used replication as a research strategy, (d) restricted their conclusions based on the limitations of their sampling techniques or with respect to possible differences between their accessible and target populations, and (e) provided alternative explanations for positive findings relative to threats to internal validity.

Using a rating team that consisted of a faculty member and five graduate students, Horton et al. (1993) examined 130 articles published in Volumes 22, 24, and 26 of *JRST*. Following Krippendorf's (1980) recommendation, Horton et al. used a purposive sampling strategy to insure balanced coverage of the time period of interest. This strategy also was appropriate for descriptive studies such as a content analysis. Because their focus was on methodological quality, Horton et al. did not include "literature reviews, meta-analyses, instrument development or validation studies, philosophical inquiries, position papers, or historical studies" (p. 859). Horton et al. also used Campbell and Stanley's (1966) designations of

preexperiments, true experiments, quasiexperiments, correlational, survey, causal-comparative, and ethnographic to classify the articles they analyzed. To facilitate their analysis, Horton et al. developed a category evaluation sheet that was based on the coding form used by Shaver and Norton (1980a) and Wallen and Fraenkel (1988a). The categories developed for the coding sheet were based on commonly accepted standards of research as reported in educational research methods textbooks. The rating team rated several articles from earlier volumes of *JRST* for both training purposes and instrument refinement. Pairs of raters then used the final instrument, and Spearman correlation coefficients ranging from .83 to .91 were sustained among the different pairs of raters. Horton et al. also used Scott's pi as a measure of the reproducibility of raters' coding. Of the 130 articles analyzed, Horton et al. (1993) found the following:

- All of the authors provided a purpose/problem statement for their study, with 90% of the studies containing either (a) an explicit argument of the worth of their study (72%) or (b) an implied worth (18%). The remaining studies either were a test of theory (5%), direct or systematic replication of previous research (3%), or an extension of the findings of previous work (5%).
- With respect to sampling issues: 5% defined the target population, 12% described an accessible population, 62% used convenience sampling,

57% reported some demographic information about their sample, 24% used random assignment, and 12% randomized treatments.

- With respect to internal validity, although 90% of the articles reported or discussed at least one threat, only 5% discussed the threats satisfactorily and 19% discussed it marginally.
- With respect to external validity, the most common problem was overgeneralization: only 35% restricted conclusions relative to the sampling procedure but 48% of the studies overgeneralized their findings.

Horton et al. (1993) recommended encouraging researchers to use replications as a solution to limited generalizability. They also recommended a follow up study to evaluate the quality of the studies reported after their study. Based on their findings, Horton et al. concluded, "The results of this study indicate that the methodological quality of published science education research should remain a concern for both practitioners and readers" (p. 857).

Hsu (2005). Hsu conducted a content analysis of research articles published in the *American Educational Research Journal* (*AERJ*), *Journal of Experimental Education* (*JEE*), and *Journal of Educational Research* (*JER*) from 1971 to 1998. Hsu sought to determine (a) the subject matters educational researchers frequently investigated; (b) the research methods/designs educational researchers frequently used; (c) the data analysis procedures educational researchers frequently employed; and (d) if there were any trends with respect to the these subject matters, methods,

and analyses. Her purpose in targeting these three journals was because they "solicited manuscripts of original scientific research on practical educational problems " (p. 112), and she was interested in showing "the similarities and differences of methods/analyses used in articles published over the years" (p. 112).

Hsu's (2005) data set came from a database she developed as part of a cross-cultural study of educational and psychological research methods. She constructed this database by first identifying 10 prestigious journals from the U.S., China, and Taiwan. She then selected up to 24 articles per year per journal for each of the targeted years (1971–1998). She also used a stratified random sampling strategy if a particular year had more than 24 articles per journal. These articles were then reviewed and classified with respect to the three main areas of subject, methodology/design, and analysis. Ultimately, 18 subject matter categories, 30 research methods/designs categories, and 34 data analysis procedures categories were identified. In total, Hsu analyzed 2,226 articles: $N = 713$ from *AERJ*, $N = 638$ from *JEE*, and $N = 875$ from *JER*. Hsu classified each article into only one primary subject matter category. For those articles in which more than one research method applied, Hsu used what she felt was the most important of the three methods. For data analysis, all procedures in the articles were recorded only once even if they were used more than once. To assure that the same criteria were used in classifying the articles, Hsu created a manual that described the characteristics of each

category and a group of graduate students were recruited and trained to carry out the coding procedure.

Hsu (2005) defined "most frequently investigated" if at least 5% of the articles were classified as belonging to a particular category, and "least frequently investigated" if less than 1% of the articles were classified as belonging to a particular category. With respect to subject matters, Hsu reported that the most frequently investigated subjects were psychology in education, teaching/instruction, teachers, and measurement, and the least frequently investigated subjects were agencies and institutions related to education, counseling/medical services, occupational education, and policy-making areas. With respect to research methods/designs, Hsu reported that the most frequently used methodologies were comparative approach, descriptive research, quasiexperimental design, and survey. With respect to statistical procedures, Hsu reported that the five of the most frequently data analysis strategies were descriptive statistics, ANOVA, bivariate correlation, *t* test, and regression. Hsu also noted there was a shift from quantitative experimental methods to qualitative methods during past 2 decades.

Hsu (2005) reported that the results of her study should encourage students to understand and interpret concepts related to those methods/designs that are identified as the most frequently used. She also recommended that these methods/designs should be considered as basic cores of knowledge for graduate students as researchers besides methods/designs related to their interests and

specialties. Hsu recommended research educators should strengthen the instruction of qualitative-related methods. She also highly recommended that a study should be conducted to investigate the reason why qualitative research methods were least used by agencies and institutions related to education, counseling/medical services, occupational education, and policy-making areas during the targeted years. Hsu encouraged the editors of the targeted journals to review their respective journal editorial policies and scopes of publications relative to her findings.

Included with her findings, Hsu (2005) acknowledged the affect of editorial members' role and the journals' editorial policy in determining the types of articles that were published during the targeted years of her study. Hsu reported that the relationship between the trends and the editorial policies were speculative and readers should use the data to make their own judgments about the feasibility of the interpretations. Hsu also warned the reader not to overgeneralize her findings because the results apply only to the three journals she targeted and most likely would be different if other journals were examined.

Bliss (2012). As noted earlier, Bliss conducted a content analysis of all 189 articles published in *CAR* from 1983 to 2010. Unlike the current study, which will focus on methodological quality, Bliss's content analysis focused on article demographics, which included the total number of articles published in each volume, the aggregate number of pages of each article, the number of authors, authors' institutional/organizational affiliation, article subject classification, and

geographical locations partitioned by regions with respect to which region submitted the most or fewest articles. Bliss' primary purpose was to (a) review the individual contributions to *CAR* and (b) to reflect on what the future might hold for *CAR* (pp. 2–3). His rationale for conducting his study was because he believed the findings would help "UAA and its membership (be) academically positioned to meet the ever-changing aviation/aerospace industry and address the continuing challenges of the 21st century" (p. 3). The salient findings of Bliss' study included the following (pp. 11–12):

- *CAR* contributors included authors from 54 different institutions or organizations and one non-affiliation. The top five institutions were Southern Illinois, Middle Tennessee State, Purdue, Auburn, and Embry-Riddle.
- The top 10 contributors were all educational institutions and accounted for 72% of the total *CAR* articles. Universities and colleges accounted for 97% of *CAR* articles, with very few non-academic organizations having published in *CAR* during the targeted period.
- The top five institutions with respect to the aggregate number of journal articles were Southern Illinois University (79 articles), Middle Tennessee State University (31), Purdue University (24), Auburn University (22), and Embry-Riddle Aeronautical University (18).

- The top five institutions with respect to the aggregate number of pages by a lead author were Southern Illinois University (557 pages), Embry-Riddle Aeronautical University (194), Auburn University (175), Middle Tennessee State University (158), and Purdue University (134).
- Although topics varied widely, they were mostly concentrated in descriptions/issues with respect to collegiate flight training (37 articles), descriptions of collegiate aviation degree programs (20), demographics of collegiate aviation students (20), airport systems including air traffic control (15), and the commercial airline industry (14). The topics that received the least attention were U.S. government and military (5 articles) and international aviation issues (3).
- Among the nine regions into which he partitioned the country (Alaska, Northwest Mountain, Western Pacific, Great Lakes, Central Southwest, New England, Eastern and Southern), Bliss (2012) reported that of the 343 *CAR* articles he analyzed during the targeted years, 159 (46%) came for the Great Lakes region and 90 (26%) came from the Southern region.
- During 2001–2010, 60% of all *CAR* articles were published, with 39% of them (73 articles) being published between 2006–2010. Bliss (2012) surmised that this increase in the number of published articles was most likely due to the editors' decision to publish *CAR* biannually in the spring and fall beginning in 2007.

Bliss (2012) reported that he believes that the large contribution of articles to *CAR* from the Great Lakes and Southern regions will continue because they have well-established and possess large numbers of degree programs, aviation students, and faculty members. He also surmised that contributions to *CAR* would continue to come from academia because faculty are encouraged and rewarded to engage in scholarly activity. As for the small influence international aviation has in *CAR*, Bliss believes that this is not a concern because the percentage of *CAR* articles that focused on international aviation was commensurate with international memberships to UAA.

Summary and Study Implications

Although content analyses have been conducted in a wide range of areas and applications including aviation, none have examined aviation research articles published in refereed journals from a methodological perspective. The current study addressed this gap in the literature by focusing on the methodological quality of published aviation research articles in *CAR*. Similar to Horton et al. (1993) and Hsu (2005), I examined a purposive sample of *CAR* articles published between 2007 and 2012. My reason for beginning with 2007 was because this was when *CAR* began publishing on a biannual basis, spring and fall (Bliss, 2012). Because I concentrated on methodological quality, I followed Horton et al.'s protocol and excluded articles that consisted of "literature reviews, meta-analyses, instrument development or validation studies, philosophical inquiries, position papers, or

historical studies" (p. 859). Although I did not include these types of articles as part of data analysis, I still reported the number of such articles from a demographic perspective to complement Bliss' (2012) findings.

The implications of the findings of the current study should be informative to the aviation research community, including the editors and reviewers of *CAR*, the authors who have published or seek to publish in *CAR*, and the readers of *CAR*. For example, the findings of the current study could serve as an impetus for *CAR* editors to review their performance, reviewers, editing, and content policies. The current study's findings also could give educators in aviation colleges and universities the opportunity to review the strengths and weaknesses of their research classes, and alert potential *CAR* contributors to what constitutes sound research practices. It also could enlighten the readers of *CAR* about the methodological quality of the studies being reported.

Chapter 3

Methodology

Population and Sample

Population. The target population was all articles published in *CAR* between the time of its first publication, 1991, and 2014. The accessible population was all articles published in *CAR* between 2007 and 2012 (Volumes 25–30). The reason for starting with 2007 was because the editors of *CAR* began publishing *CAR* biannually (spring and fall) beginning with the Spring 2007 volume (Bliss, 2012).

Sample. The sampling strategy was purposive. I reviewed all the articles published in *CAR* during the targeted time period and selected only articles that described either quantitative or qualitative studies. Because the focus of the current study was to assess the methodological quality of research studies and the coding form was developed purposely for quantitative or qualitative studies, I did not include literature reviews, meta-analyses, studies that described the development or validation of an instrument, philosophical inquiries, position papers, or historical studies. The final sample size was 69, but the aggregate number of articles in the accessible population was 76.

Instrumentation

The primary data collection instrument was a category-based coding form, which my advisor developed with a group of graduate students as part of a research

practicum. This coding form consisted of a set of predetermined categories that corresponded to the 10 research questions cited in Chapter 1. The categories were based on commonly accepted standards of research as reported in various educational research methods textbooks such as Ary et al. (2010), Gall et al. (2007), and Fraenkel et al. (2011). The initial coding form was augmented inductively during the coding of *CAR* articles during the practicum. The earlier studies of Shaver and Norton (1980a, 1980b), Wallen and Fraenkel (1988a, 1988b), and Horton et al. (1993) provided guidance in the initial development of the coding form. By aligning the current coding form to the coding forms used in these previous studies provided a certain level of content validity.

The coding form had 10 main categories covering the author's purpose and/or justification for the study, research questions and/or hypotheses, sampling issues, instrumentation issues, type of design, discussion related to internal and external validity, data analysis approaches and corresponding statistical measures reported, recommendations, limitations, and delimitations. A copy of the coding form is provided in Appendix A. Although the categories of the current form were predetermined, I supported an emergent design framework to allow new categories to be added to the form based on the content of the articles I reviewed.

Procedures

Research methodology. The research methodology for the current study was content analysis, which is a " a research method applied to written of visual

materials for the purpose of identifying specified characteristics of the material" (Ary et al., 2010, p. 457). Because I reviewed articles published in several volumes of a refereed journal to determine the methodological quality of the articles, content analysis was the most appropriate research methodology.

Human subject research. Following university protocol, I submitted an application to Florida Institute of Technology's Institutional Review Board (IRB). This study satisfied the exempt criteria for research involving human subjects because (a) it involved the collection and study of existing documents publicly available to members of UAA, (b) did not present a risk to human subjects, (c) did not use any special populations, and (d) was conducted in an established or commonly accepted educational setting involving normal educational practices. A copy of the approved IRB application is provided in Appendix B.

Study implementations. According to Ary et al. (2010, p. 457) a content analysis is implemented using the following steps: (a) specify the phenomenon to be investigated, (b) select the media from which the observations are to be made, (c) formulate exhaustive and mutually exclusive coding categories, (d) decide on the sampling plan to be used, (e) train the coders, and (f) analyze the data. Steps (a) through (d) were described in previous sections, step (e) is discussed in the following paragraph, and step (f) will be presented as a separate section at the end of this chapter.

To address step (e)—train the coders—I read and coded five articles from the 2006 journal of *CAR*. My advisor and a group of graduate students in a research practicum previously coded these articles. I then met with my advisor and reviewed my coding with those of his and his former students. This process continued until there was a greater than 90% correlation between the two sets of coding forms. This then ended the training session for me.

Beginning with 2007 articles, my advisor and I coded each article independently and then compared our coding results to assess our level of agreement. As part of this process, we calculated a percent agreement between our respective ratings by summing the number of cases that were coded the same way and dividing by the total number of cases. The overall quotient, which served as measure of intercoder reliability, was greater than 90%. As noted by Cohen (1960), the problem with a percent agreement approach is it does take into consideration that raters are expected to agree with each other a certain percentage of the time simply based on chance. To address this problem, Cohen recommended calculating reliability by using Cohen's Kappa. One of the assumptions of this reliability index, though, is the raters must operate independently of each other. Although my advisor and I coded the articles independently, we occasionally worked together to come to a consensus about what a particular rating should be given when there were inconsistencies in our coding. As a result, Cohen's Kappa was not calculated and percent agreement was the only approach used for intercoder reliability.

Threats to internal validity. The general concept of internal validity is the extent to which changes in a dependent variable (DV) are related directly to an independent variable (IV). Campbell and Stanley (1966) initially presented a set of eight threats to internal validity, which if not controlled, could provide reasonable alternative explanations for a study's outcome other than the targeted IVs. These threats are history, maturation, testing, instrumentation, statistical regression, selection bias, experimental mortality (attrition), and diffusion. Ary et al. (2010) subsequently added four additional threats: selection-maturation interaction, experimenter effect, subject effects, and location. Because the concept of internal validity is more closely associated to empirical studies (both intervention and observational) and not descriptive studies, many of these threats were not applicable to the current study. Nevertheless, following is a brief discussion of these threats that describes whether or not they were applicable to the current study. For those threats that were applicable, I included a description of how they might have impacted the current study and how I controlled for or mitigated these threats.

***History*.** A history threat refers to whether an event independent of the treatment occurs between measurement periods (i.e., prior to treatment and after treatment), which could then provide an alternative explanation to the results of the study. This threat was not applicable to the current study because it did not involve the administration of a treatment.

Maturation. Maturation refers to biological or psychological changes within participants that may occur over time. As a result, it is possible for participants to perform differently on the dependent variable because they are older, wiser, more fatigued, or less motivated (Ary et al., 2010). This threat was not applicable to the current study because it did not involve human participants.

Testing. A testing threat is possible when participants of a study are administered a preassessment prior to treatment and a post-assessment after treatment. The concern is that it is possible for participants' post-assessment scores to be influenced by their exposure to a preassessment instead of the treatment. This threat was not applicable to the current study because it did not involve human participants and there were no pre- and post-assessments.

Instrumentation. An instrumentation threat refers to changes in a data collection instrument that are made during the course of a study, changes in the way assessments are scored, or using different scorers/coders. This threat was possible to the current study if changes were made to the coding form or if different coders were used. For example, if I had made changes to the coding form based on the emergence of new categories, it would be possible that the coding associated with articles based on an earlier form might no longer be consistent. Additionally, if another person were to code some of the articles, it would be possible that the coding could be inconsistent. To mitigate this threat, (a) I re-examined previously coded articles against any subsequent categories that might be added to the coding

form, and (b) my advisor and I were the only ones who coded the articles. We also ensured that we had a percent agreement of at least .90.

***Statistical regression*.** This threat refers to the possibility a participant is selected on the basis of extreme scores. This threat was not applicable to the current study because it did not involve human participants or selection based on scores on an assessment.

***Selection bias*.** This threat refers to the possibility that participants might have personological traits or characteristics related to the study's variables and it is these traits/characteristics that account for changes in the dependent variable. This threat was not applicable to the current study because it did not involve human participants.

***Mortality*.** The mortality threat refers to a differential loss, or attrition, of participants. For example, if all poor performing students drop out of a study, the results would reflect a different sample and the results would not truly reflect a treatment. This threat was not applicable to the current study because it did not involve human participants.

***Diffusion*.** This threat refers to the concern that members of a treatment group might share information about their "treatment" with members of the control group. This threat was not applicable to the current study because it neither involved human participants nor involved placing participants into treatment and control groups.

Selection-maturation interaction. This threat refers to the combined influence of selecting participants who have specific characteristics and might mature faster than another group over the course of a study. This threat was not applicable to the current study because it did not involve human participants.

Experimenter effect. This threat refers to the concern for how a researcher's personality, enthusiasm, or personological characteristics such as age and gender might unintentionally affect or influence the performance of participants. This threat was not applicable to the current study because it did not involve human participants.

Subject effects. This threat refers to the possible changes in participants' attitudes relative to the Hawthorne effect and John Henry effect. The Hawthorne effect is where participants in a treatment group might perform well because of the attention or recognition they are receiving and not because of any treatment. The John Henry effect, also known as compensatory rivalry, is where participants in a control group intentionally perform poorly. This threat was not applicable to the current study because it does not involve human participants.

Location. This threat refers to changes in the setting at which a study is implemented or an assessment is administered. This threat was not applicable to the current study because I did not anticipate any changes in the location where my advisor and I coded the articles.

Data Analysis

Data analysis involved descriptive statistics only. I reported frequencies and percentages in the targeted categories. An example of how I presented my findings is given in Table 3.1, which shows how the table was structured for the presence of research questions. In addition and when appropriate, I also reported any comments that accompanied the coding of a category.

Table 3.1
Sample Table for Research Questions Category

Categories	***N***	**%**
RQ(s) clearly specified		
RQ(s) specified as objectives		
RQ(s) implied in purpose statement		
None given		

Chapter 4

Results

This chapter is organized into two sections. The first section provides a summary of the number of articles published each spring and fall for the targeted years 2007–2012. The second section provides a summary of the results of the content analysis relative to each research question.

Overall Summary

As reported in Table 4.1, 76 articles were published during the targeted 6-year period: 40 were published during the spring terms and 36 were published during the fall terms. Of the 76 articles, 69 were reviewed in this content analysis: 35 were from the spring terms and 34 were from the fall terms. The years with the highest and lowest frequency of published articles were 2008 and 2012, respectively. For the most part, the number of articles published each year declined

Table 4.1
Summary of Number of Articles by Year

	Spring		Fall		Overall	
Year	***N***	**R**	***N***	**R**	***N***	**R**
2007	8	6	6	5	14	11
2008	10	8	9	8	19	16
2009	6	6	7	7	13	13
2010	8	8	6	6	14	14
2011	4	4	6	6	10	10
2012	4	3	2	2	6	5
Total	**40**	**35**	**36**	**34**	**76**	**69**

Note. *N* = Total number of articles. R = Total number of articles reviewed.

after 2008. Seven articles were not reviewed because they were either literature reviews, meta-analyses, studies that described the development or validation of an instrument, philosophical inquiries, position papers, or historical studies.

Results Relative to Research Questions

Research question 1. The first research question examined the extent to which aviation researchers stated the purpose of their study and corresponding research questions. As reported in Table 4.2, overall the majority of articles (84%) clearly contained a purpose statement. With respect to research questions, though, only half (51%) of the article also had clearly specified research questions, which means that half of the 69 articles reviewed did not include research questions.

Table 4.2
Summary of Problem Statement Information

	Year						
Category	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**Overall**
Purpose Statement							
Clearly stated	81%	88%	92%	71%	90%	80%	84%
Not clearly stated	18%	12%	8%	29%	10%	20%	16%
Research Questions							
Clearly specified	36%	63%	46%	43%	70%	40%	51%
None given	64%	37%	54%	57%	30%	60%	49%
Hypotheses							
Research	9%	0%	0%	0%	0%	0%	2%
Statistical	9%	19%	15%	14%	10%	0%	13%
None given	55%	56%	31%	50%	10%	20%	41%
Not applicable	18%	25%	54%	36%	80%	80%	44%
Other Information							
Test of theory	0%	0%	8%	0%	0%	0%	2%
Replication	0%	19%	8%	0%	20%	0%	9%
	$n = 11$	$n = 16$	$n = 13$	$n = 14$	$n = 10$	$n = 5$	$n = 69$

Note. This table corresponds to Research Question 1.

The years with the lowest and the highest percentages were 2007 and 2011, respectively. With respect to research hypotheses, of the studies in which hypotheses were applicable (56%), 2% were written as research hypotheses, 13% were written as statistical hypotheses, and 41% did not have any at all. Of the 69 articles reviewed, 6 (9%) were replication studies and 1 was a test of theory. Replication studies were conducted in 2008, 2009, and 2011, and a test of theory occurred in 2009.

Research questions 2 and 3. Table 4.3 contains a summary of findings relative to Research Questions 2 and 3. The second research question examined the extent to which aviation researchers defined their target and accessible populations. As reported in Table 4.3, overall 45% of the articles contained information about the target population, 54% had information about the accessible population, and 33% did not have any information about the population. The overall pattern across the 6-year period indicates only about 50% of the studies reported the accessible population. The year with the smallest frequency of information about the target population was 2010 (7%), but the articles published in 2010 also had the highest frequency of information about the accessible population (71%).

The third research question examined what sampling strategies aviation researchers commonly used. As reported in Table 4.3, the most popular sampling strategies were convenience (43%) and volunteer (39%). Only 7% of the studies

Table 4.3
Summary of Sampling Issues Part 1: Population and Sampling Strategy

	Year						
Category	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**Overall**
Population							
Target	18%	25%	54%	7%	60%	40%	45%
Accessible	55%	50%	46%	71%	50%	40%	54%
None given	45%	44%	31%	21%	20%	40%	33%
Sampling Strategy							
Random selection	0%	6%	23%	0%	10%	0%	7%
Convenience	73%	50%	38%	50%	10%	20%	43%
Volunteer	73%	31%	38%	50%	10%	20%	39%
Purposive	9%	44%	15%	24%	10%	20%	20%
Census	0%	6%	0%	0%	0%	20%	3%
Can't tell / None given	0%	6%	0%	7%	10%	0%	4%
	$n = 11$	$n = 16$	$n = 13$	$n = 14$	$n = 10$	$n = 5$	$n = 69$

Note. This table corresponds to Research Questions 2 and 3.

included random selection, which infers that the vast majority of the studies published in the targeted 6-year period consisted of biased samples. The reader will note, though, that the percentage of studies that involved convenience or volunteer samples decreased considerably after 2010.

Research question 4. The fourth research question examined the extent to which aviation researchers described their samples, including how representative the sample was to the parent population, sample size, and sample assignment. As reported in Table 4.4, overall 42% of the studies contained an adequate description of the sample, 36% of the studies contained a marginal description of the sample, and 22% of the studies did not contain any sample description. The years with the highest frequency of adequate descriptions were 2010 (65%) and 2011 (60%). The

Table 4.4

Summary of Sampling Issues Part 2: Sample Description, Representativeness, Size, and Assignment

	Year						
Category	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**Overall**
Sampling Description							
Adequate	9%	31%	46%	65%	60%	40%	42%
Marginal	73%	44%	31%	14%	30%	20%	36%
None given	18%	25%	23%	21%	10%	40%	22%
Sample Representativeness							
Author's claim	18%	25%	0%	14%	10%	20%	14%
None given	82%	75%	100%	86%	90%	80%	86%
Sample Size							
Attention given to minimum sample size[a]	9%	6%	0%	7%	0%	20%	6%
Sample size specified	100%	94%	85%	93%	100%	100%	94%
Sample Assignment							
Random	0%	6%	8%	7%	0%	0%	4%
Nonrandom	9%	6%	8%	0%	0%	0%	4%
Treatment assigned	0%	0%	0%	14%	0%	0%	3%
None given	18%	25%	8%	7%	10%	20%	15%
N/A	73%	63%	77%	72%	90%	80%	74%
	n = 11	n = 16	n = 13	n = 14	n = 10	n = 5	n = 69

Note. This table corresponds to Research Question 4.
[a]Attention to minimum sample size includes power analysis for quantitative studies, information about margin of error for surveys, and following appropriate protocols for qualitative studies.

year with the lowest frequency of adequate sample description was 2007, though the same year had the highest marginal sample description.

With respect to sample representativeness, overall 86% of the studies did not provide any information about the extent to which the sample was representative of the parent population. In 2009, none of the studies mentioned sample representativeness. The highest frequency of author's claim was in 2008

with 25%. With respect to sample size, overall 94% percent of the studies contained information about sample size, however, only 6% gave attention to power analysis or margin of error for determining the appropriate minimum sample size. Although studies conducted in 2012 had the highest frequency regarding sample size specified (100%) and attention to minimum sample size (20%), studies conducted in 2009 had the lowest frequency regarding sample size specified (85%) and attention given to minimum sample size (0%).

With respect to sample assignment, overall 74% of the studies were not applicable relative to sample assignment because the methodology used (e.g., correlation or survey) did not require a group membership variable. Of the remaining 26% of the overall studies, 4% used random assignment, 4% used nonrandom assignment, 3% assigned treatments, and 15% did not specify what type of assignment was used. Of the targeted 6-year period, 2011 had the highest frequency (90%) in which sample assignment was not relevant, which suggests the studies mostly were descriptive. Also in 2011, of the remaining 10% of the studies, none gave attention to sample assignment.

Research question 5. The fifth research question examined the extent to which aviation researchers discussed instrumentation issues, including describing the data collection instrument and giving attention to validity and reliability. As reported in Table 4.5, overall 68% of the studies contained a description of the data collection instrument. The overall pattern regarding instrumentation description

Table 4.5
Summary of Instrumentation Issues

Category	Year						Overall
	2007	2008	2009	2010	2011	2012	
Description Given							
Yes	91%	62%	62%	50%	100%	40%	68%
No	9%	38%	38%	50%	0%	60%	32%
Attention to Validity							
Yes	18%	19%	31%	36%	50%	20%	29%
No	82%	81%	69%	64%	50%	80%	71%
Attention to Reliability							
External	0%	0%	0%	0%	0%	0%	0%
Internal	18%	19%	15%	24%	30%	0%	17%
Interrater / Interscorer	0%	13%	0%	7%	0%	20%	6%
None given	82%	50%	77%	64%	70%	60%	67%
N/A (archived data)	0%	19%	8%	14%	0%	20%	10%
	n = 11	n = 16	n = 13	n = 14	n = 10	n = 5	n = 69

Note. This table corresponds to Research Question 5.

across the targeted 6-year period, however, had a decreasing trend except for 2011 when 100% of the studies contained a description of the instrument.

With respect to instrumentation validity, overall 71% of the research studies did not indicate whether any attention was given to validity such as face, content, criterion related, or construct. In addition, between 2007 and 2011 the percentage of studies in which attention was given to instrumentation validity increased from 18% to 50%, respectively, but then decreased to 20% in 2012. As for reliability, overall two thirds of the studies (67%) did not contain any information about instrumentation reliability. Of the remaining one third, 17% focused on internal reliability (split half and Cronbach's alpha), 6% were interrater/interscorer, and

10% of the studies involved archived data and therefore used no formal data collection instrument.

Research question 6. The sixth research question examined what research methodologies/designs aviation researchers commonly used. As reported in Table 4.6, among the 63 quantitative studies, survey research (43%) was the most frequently used methodology, and preexperimental research design (2%) was the least frequently used with 2008 being the only year in which this latter design was used. Even though survey overall was the most commonly used methodology, none of the studies published in 2012 used survey research. Instead, more than half of the studies (60%) in 2012 used content analysis, and the remaining 40% were split evenly between quasiexperimental and correlational. As also reported in Table 4.6, 12 (17%) of the 69 articles reviewed were qualitative studies. The most frequently used qualitative methodology was content analysis (50%) followed by case study (25%), narrative (17%), and phenomenology (8%).

Research question 7. The seventh research question examined the extent to which aviation researchers gave attention to threats to internal validity. As noted in Table 4.7, internal validity threats were examined separately for quantitative and qualitative designs. The reader should note that the focus of this research question was not related to specific threats to internal validity, but only whether or not authors discussed threats to internal validity. With respect to quantitative designs, overall 86% of the studies did not contain any discussion about threats to internal

Table 4.6

Summary of Methodology / Design Types

	Year						
Category	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**Overall**
Quantitative							
Preexperimental	0%	8%	0%	0%	0%	0%	2%
True experimental	8%	8%	8%	8%	0%	0%	6%
Quasiexperimental	8%	17%	15%	17%	11%	20%	14%
Ex post facto	8%	8%	0%	8%	22%	0%	8%
Correlational	8%	8%	23%	33%	0%	20%	16%
Survey	58%	42%	46%	33%	67%	0%	43%
Content analysis	8%	8%	8%	0%	0%	60%	10%
Total	**12**	**12**	**13**	**12**	**9**	**5**	**63**
Qualitative							
Case Study	50%	29%	0%	0%	0%	0%	25%
Narrative	0%	14%	0%	0%	0%	100%	17%
Grounded theory	0%	0%	0%	0%	0%	0%	0%
Phenomenology	0%	14%	0%	0%	0%	0%	8%
Ethnographic	0%	0%	0%	0%	0%	0%	0%
Content analysis	50%	43%	0%	100%	100%	0%	50%
Other	0%	0%	0%	0%	0%	0%	0%
Total	**2**	**7**	**0**	**1**	**1**	**1**	**12**
Unknown / unclear	0	1	0	1	0	0	2
	$n = 11$	$n = 16$	$n = 13$	$n = 14$	$n = 10$	$n = 5$	$n = 69$

Note. This table corresponds to Research Question 6. Some of the studies had multiple designs.

validity. The only years in which at least one threat was discussed were 2008, 2009, and 2010. With respect to qualitative studies, the concept of internal validity was addressed via credibility (Lincoln & Guba, 1985). In this context, the lack of attention to internal validity threats observed in the quantitative studies continued with the qualitative studies. The concept of credibility was addressed in only two

Table 4.7

Summary of Threats to Internal Validity

	Year						
Category	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**Overall**
Quantitative Designs							
At least 1 threat discussed	0%	13%	8%	7%	0%	0%	6%
No threats discussed	100%	69%	92%	79%	90%	100%	86%
Qualitative Designs							
Credibility discussed	0%	13%	0%	7%	0%	0%	4%
Credibility not discussed	0%	6%	0%	7%	10%	0%	4%
	n = 11	*n* = 16	*n* = 13	*n* = 14	*n* = 10	*n* = 5	*n* = 69

Note. This table corresponds to Research Question 7. Specific threats to internal validity were not examined.

years, 2008 (13%) and 2013 (7%). Thus, 2008 had the highest frequency in which attention was given to internal validity for both quantitative and qualitative designs.

Research question 8. The eighth research question examined what data analysis methods aviation researchers commonly used. As reported in Table 4.8, overall 88% of the studies used descriptive statistics, which included measures of central tendency, measures of variance, frequencies, and percentages. The most commonly used inferential statistical procedures overall were *t* test for independent samples (19%), correlation (16%), chi-square (14%), and oneway ANOVA (12%), and the least commonly used were matched pairs ANOVA (1%) and regression (1%). In 2012, in addition to descriptive statistics (80%), only two statistical procedures were used: *t* test for independent samples (20%) and chi-square (20%).

With respect to corresponding statistical measures such as confidence intervals, effect size, and power, 80% of 69 studies reviewed did not report any

Table 4.8

Summary of Data Analysis Procedures and Corresponding Statistical Measures

	Year						
Category	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**Overall**
Statistical Procedures							
Descriptive	100%	75%	92%	86%	90%	80%	88%
t test Ind. samples	0%	13%	31%	0%	60%	20%	19%
t test matched pairs	9%	0%	15%	14%	0%	0%	7%
ANOVA-Oneway	18%	0%	23%	14%	10%	0%	12%
ANOVA-Factorial	18%	19%	0%	0%	0%	0%	7%
ANOVA-Matched pairs	9%	0%	0%	0%	0%	0%	1%
Covariance	0%	0%	0%	0%	20%	0%	3%
Correlation	18%	13%	8%	43%	0%	0%	16%
Regression	0%	0%	0%	7%	0%	0%	1%
Chi-square	9%	31%	31%	14%	10%	20%	14%
Unclear/Unknown	0%	13%	8%	7%	10%	0%	7%
Statistical Measures Reported							
Confidence intervals	0%	0%	0%	7%	10%	0%	3%
Effect size	0%	0%	0%	7%	0%	0%	1%
Correlation coefficient	18%	13%	8%	43%	0%	0%	16%
Regression coefficient	0%	0%	0%	14%	0%	0%	3%
Power	0%	0%	0%	0%	0%	20%	1%
Other[a]	0%	0%	0%	14%	20%	0%	6%
None given	91%	88%	92%	50%	80%	80%	80%
	$n = 11$	$n = 16$	$n = 13$	$n = 14$	$n = 10$	$n = 5$	$n = 69$

Note. This table corresponds to Research Question 8. Involves quantitative studies only. Some of the studies used multiple data analysis methods.

[a]Other = Wilks' lamda or standard error.

statistical measures. Among the various statistical measures listed in Table 4.8, the most commonly reported measure was the correlation coefficient (16%) and the least commonly reported measures were effect size (1%) and power (1%). Of the targeted 6-year period, the 14 articles published in 2010 had the highest frequency in which statistical measures were reported.

Research questions 9 and 10. Table 4.9 contains a summary of the study's findings relative to Research Questions 9 and 10. The ninth research question examined the ways aviation researchers addressed generalizability and reported their conclusions. As reported in Table 4.9, 100% of the studies addressed generalizability including population generalizability (13%) ecological generalizability (87%). In 2012, though, none of the studies discussed population generalizability. With respect to conclusions, the studies were reviewed from five perspectives. The first three perspectives were relative to Rath (1973): conclusions as truth, which indicate overgeneralization; conclusions as trivia, which indicate under generalization; and conclusions as findings, which make no conclusions but simply summarize the results. The remaining two perspectives were conclusions as plausible explanations and conclusions as implications. As reported in Table 4.9, 100% of the studies reported conclusions. However, 90% of the studies reported conclusions as findings, 14% reported conclusions as plausible explanations, and 13% reported conclusions as implications.

The last research question examined the extent to which aviation researchers stated their recommendations, limitations, and delimitations. As reported in Table 4.9, with respect to recommendations, the general trend over the 6-year period indicates that authors are providing recommendations with respect to future research and to practice. However, overall in 64% of the studies the authors did not provide any information about limitations and delimitations. Of the studies

Table 4.9

Summary of Generalizability Issues, Conclusions, Recommendations, and Limitations

	Year						
Category	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**Overall**
Generalizability							
Population	18%	13%	8%	14%	20%	0%	13%
Ecological	82%	88%	92%	86%	80%	100%	87%
None given	0%	0%	0%	0%	0%	0%	0%
Conclusions Reported							
As truth	27%	6%	23%	7%	10%	20%	14%
As trivia	0%	0%	0%	0%	0%	0%	0%
As findings	82%	94%	85%	93%	100%	80%	90%
As plausible explanation	9%	13%	15%	21%	0%	20%	13%
As implications	55%	0%	8%	0%	20%	0%	13%
None given	0%	0%	0%	0%	0%	0%	0%
Recommendations							
W/respect to future research	55%	69%	46%	64%	70%	80%	62%
W/respect to practice	73%	56%	69%	43%	60%	60%	59%
None given	0%	0%	8%	7%	0%	0%	3%
Limitations / Delimitations							
Adequate	18%	13%	8%	7%	20%	20%	13%
Marginal	46%	25%	31%	14%	10%	0%	23%
None given	36%	62%	61%	79%	70%	80%	64%
	n = 11	n = 16	n = 13	n = 14	n = 10	n = 5	n = 69

Note. This table corresponds to Research Questions 9 and 10.

in which the authors did address limitations and delimitations, 13% addressed them adequately whereas 23% discussed them marginally.

Chapter 5

Discussion and Recommendations

This chapter is organized in four sections. The first section contains a discussion of the results presented in Chapter 4. The second section contains a set of recommendations for future research. Based on the study's findings, the third section contains recommendations for practice relative to methodological quality. The last section contains an overall summary of the study and its findings.

Discussion

As reported in Table 4.1, there was a declining trend in the number of articles published in *CAR* for the 5-year period 2008 to 2012. One plausible explanation for this could be the lack of article submissions for publication considerations. A second plausible explanation could be the articles being submitted lacked the methodological rigor and therefore were rejected. A third plausible explanation could be due to the lack of articles that were commensurate with the journal's publishing agenda. A fourth plausible explanation could be a function of *CAR*'s financial resources or support personnel. Although not shown in Table 4.1, this trend appeared to bottom out because the number of articles published in 2013 and 2014 was 20 and 13, respectively.

As noted in Table 4.2, more than 80% of the articles provided a clear purpose statement, which described the objective or primary focus of the study. However, approximately 50% of the articles did not include corresponding research

questions (RQs). Although the vast majority of the articles included a purpose statement, this lack of RQs was surprising because every research study begins with a research question, which ultimately is answered by conducting the study. Research questions also drive research methodology, which then leads to appropriate data analysis methods. For example, the RQ "What is the relationship between *x* and *y*?" infers an associational type research methodology such as a correlational study, which in turn leads to corresponding data analysis procedures such as a *t* test, ANOVA, multiple regression, or logistic regression. Thus, it is difficult to determine if appropriate methodology and corresponding data analysis procedures are being used if RQs are not specified. The lack of RQs also makes it difficult to conduct replication studies, which are needed to extend and confirm the results of a particular study.

Relative to Research Question 2 (see Table 4.3), about one third of the articles did not provide any information about the corresponding population. This omission also was surprising from two perspectives. The first is with respect to generalizability. Without knowing the parent population from which a sample was selected, the results of a study cannot be generalized beyond the sample. This leads to what Rath (1973) refers to as "conclusions as trivia," which implies undergeneralization. The second perspective is with respect to replication studies. It is difficult to replicate a study without knowing the study's population.

Relative to Research Question 3 (see Table 4.3), the vast majority of the studies employed a convenience or volunteer sampling strategy. The reader should note that although "volunteer" is not a specific sampling strategy, it reflects a sample that consisted of participants who volunteered to complete a questionnaire. Given the frequency of using convenience and volunteer samples, many of the research studies published in *CAR* during the targeted 6-year period reflected biased samples. Because the samples were not representative of their respective parent populations, this limits generalizability of a study's results, which in turn increases the need for replication studies. However, as noted above, this is not possible if a study's RQ and corresponding population are not specified.

The concept of sampling was extended in Research Question 4. As reported in Table 4.4, less than 50% of the articles published in *CAR* during the targeted 6-year period had adequate descriptions of their samples, and 86% of the articles provided no information about how representative the sample was to the parent population. This latter finding is not surprising given the dearth of studies that described the parent population. Sample description and representativeness notwithstanding, more than 90% of the articles reviewed specified sample size. However, only 6% of the articles provided any information about what attention was given to determining the minimum sample size required, including reporting power analysis for quantitative studies are needed to find a desired effect is extremely important because researchers "can determine the probability that the

results will be significant ... before investing time and effort in the actual research" (Gravetter & Wallnau, 2013, p. 265). As Cohen (1992) observed, "There is no controversy among methodologists about the importance of power analysis..." (p. 155). Similarly, reporting a margin of error for surveys is equally important because it provides a measure of how precise the results will be in the population (Patten, 1998). Lastly, although three-fourths of the articles did not address sample assignment, this was expected because most studies were descriptive and therefore did involve a group membership variable to warrant sample assignment.

Relative to Research Question 5 (see Table 4.5), nearly 70% of the articles contained an adequate description of the data collection instrument. However, 71% and 67% of the articles did not provide any information about what attention was given to instrumentation validity and reliability, respectively. These latter two findings were of concern because a study's conclusions and recommendations are a function of the study's results, which come from the data that were collected. Thus, if the data collection instrument is flawed, then it follows that the corresponding data, results, and conclusions and recommendations also would be flawed.

Relative to Research Question 6 (see Table 4.6), the majority of articles reviewed were quantitative (63 of 69), and of the most frequently employed research methodology was survey, which represented nearly half of the quantitative studies (43%). Although the results of survey research studies can be informative, the majority of surveys reviewed were cross-sectional, which means that the results

were only reflective of that specific point in time when the data were collected. Moreover, survey research studies cannot be used to determine the effectiveness of a treatment or to understand the relationship among variables. This will require intervention or associational methodologies such as experimental, ex post facto, or correlational research.

Of the qualitative studies, reviewed, the most frequently used methodology was content analysis (50%) followed by case study (25%). Although I did not apply Creswell's (2013) criteria for determining a "good" content analysis or narrative study (p. 259 and p. 265, respectively), it was difficult to determine the level of rigor associated with these studies. Overall, most (but not all) of the authors of these studies neither applied standards of validation and evaluation (Creswell, 2013) nor provided any information about what "biases, values, and experiences" they brought to the study. This latter issue is what Creswell refers to as reflexivity (p. 216) and is used to assess how a qualitative researcher's past experiences shaped their interpretation of what they were studying.

Relative to Research Question 7 (see Table 4.7), 90% of the articles did not address threats to internal validity. This finding is alarming because threats to internal validity provide alternative explanations for the results of a study. If these threats are not controlled this could lead to many plausible reasons for the outcome of a study other than the targeted variables. Therefore, if attention is not given to internal validity, then the results are problematic because the researcher will not

know what actually led to the results. Although the current study did not identify the specific threats that were relevant to a particular study, the fact that the authors of nearly of all of the articles reviewed failed to discuss at least one threat to internal validity or provide alternative explanations to their study's findings implies that the results reported might have been spurious.

Relative to Research Question 8 (see Table 4.8), the most frequently reported statistical procedures used were descriptive, including measures of central tendency, variability, and position. This finding is consistent with the most frequently used methodology, namely, survey. Based on the results reported in Table 4.8, very few articles published in *CAR* employed inferential statistical procedures such as a *t* test, ANOVA, correlation, and regression. It is uncertain why this is the case. One plausible explanation might be that contributors to *CAR* are not interested in pursuing more elaborate studies that require these types of statistical procedures. It also is conceivable that such types of studies are not consistent with the *CAR* editors' publishing agenda. Consistent with this finding are the statistical measures reported. As noted in Table 4.8, the authors of 80% of the articles reviewed did not provide any statistical measures such as confidence intervals, effect size, correlation or regression coefficients, and power. Based on these and relative to the targeted 6-year period, it appears that *CAR* is not an appropriate journal for authors of more sophisticated quantitative research studies.

Relative to Research Questions 9 and 10 (Table 4.9), it appears that *CAR* authors are not providing sufficient information to assess the extent to which their findings may be generalized. For example, 87% of the articles did not have any information about population generalizability and the authors of 64% of the articles did not include any information about the limitations and/or delimitations of their studies. The absence of this information makes it difficult to assess the degree to which the findings of a study may be generalized to the accessible and target populations, and it makes it difficult for replication studies to be conducted. Another critical finding was that the vast majority of the articles (90%) reported their conclusions as findings. As Rath (1973) noted, conclusions are not findings; they are interpretations of what the findings mean within a study and within the context of the corresponding literature review. This approach to reporting conclusions was mitigated somewhat because the majority of the articles also contained recommendations to future research as well as to practice.

Recommendations for Future Research

Based on the experience of conducting the current study, the following list of recommendations for future research are offered:

1. The current study's focus on instrumentation validity and reliability was restricted to determining if authors discussed or did not discuss what attention they gave to validity and reliability. Therefore, a recommendation for future research is to determine specifically what attention is given to instrumentation

validity (e.g., face, content, criterion, construct) and the specific type of reliability (e.g., test-retest, equivalent forms, split-half, KR-20/21, Cronbach's alpha).

2. The current study's focus on threats to internal validity was restricted to determining if authors discussed at least one threat or none at all. A recommendation for future research is to determine what threats are applicable and the extent to which they are discussed satisfactorily.
3. The current study's focus was restricted to the 6-year period 2007–2012. Therefore, a recommendation for future research is to conduct a follow-up study to evaluate the methodological quality of *CAR* articles published after 2012.
4. The percent agreement approach, which entails adding the number of cases that were coded the same way by two raters and dividing by the total number of cases, was used in the current study for intercoder reliability. Given the limitations of this approach as noted by Cohen (1960), a recommendation for future research is to use Cohen's Kappa or Scott's pi (Krippendorff, 2004a, 2004b) as measures of intercoder reliability.

Recommendations for Practice

Based on the findings of the current study and the preceding discussion, the following list of recommendations are offered:

1. Given the absence of clearly worded research questions, a recommendation is that *CAR* authors and editors ensure that RQs are provided to help the reader understand what questions the study is trying to answer.
2. Given the lack of information about the target and accessible populations, a recommendation is that *CAR* authors and editors ensure that information about the study's population be provided to help readers assess the degree of population generalizability as well as conduct replication studies.
3. Given the prevalence of convenience sampling, a recommendation is that *CAR* authors and editors pursue probabilistic sample selection strategies to increase the likelihood that studies have a representative/unbiased sample.
4. Given the dearth of information relative to the minimum sample size needed, a recommendation is that *CAR* authors and editors ensure that attention is given to a priori power analyses for studies employing experimental and associational methodologies, and that attention is given to the margin of error for survey studies.
5. Given the lack of attention being given to instrumentation validity and reliability, a recommendation is that *CAR* authors and editors present information about the attention given to ensure that the data collection instrument produced valid and reliable data. This includes, when appropriate, attention to face and content validity, as well as reporting corresponding Cronbach alpha coefficients.

6. To increase the diversity of the types of articles published in *CAR*, it is recommended that *CAR* editors solicit studies that incorporate true experimental studies, which involve random assignment, as well as ex post facto studies—both effects type and cause type designs.
7. Given that the majority of articles did not discuss at least one threat to internal validity, a recommendation is for *CAR* authors to review their studies from an internal validity perspective and consult a research design textbook for what threats to internal validity might be relevant based on the corresponding methodology/design used. Furthermore, given the lack of detailed information associated with the concept of reflexivity relative to qualitative studies, a recommendation is that *CAR* researchers who conduct a qualitative study follow Creswell's (2013) guidance for reporting a qualitative study.
8. To increase the diversity of data analysis methods, it is recommended that *CAR* editors solicit studies that have research questions that must be answered by more sophisticated quantitative data analysis procedures such as factorial ANOVA, multiple regression, logistic regression, and mediation analyses.
9. Given that the vast majority of studies reported conclusions as findings, a recommendation is that *CAR* authors and editors focus on the distinction between findings, which refer to the results of a study, and conclusions, which are interpretations of what the findings mean within a study and within the context of the literature review.

10. Given the absence of any detailed information regarding limitations and delimitations, a recommendation is that *CAR* authors specify what events, conditions, or circumstances that were beyond their control (limitations), and what events, conditions, or circumstances they imposed to make their study feasible to implement.

Overall Summary and Concluding Remarks

The results of the current study were similar to Horton et al. (1993) relative to the methodological quality of articles published in *CAR*. Paraphrasing Horton et al., the current study's findings indicate that the methodological quality of articles published in *CAR* relative to the targeted 6-year period (2007–2012) should be of concern to the aviation research community, particularly to the authors, editors, and readers of *CAR*. It appears that *CAR* authors and editors need to have a greater appreciation for the need for replication studies within the aviation research community. By having an increased awareness of the role replication studies have in research, authors and editors presumably would be more sensitive to issues such as generalizability, defining target and accessible populations, describing samples, giving attention to instrumentation validity and reliability, and specifying limitations and delimitations. The reader should not overgeneralize the findings of the current study, however, because they apply only to the targeted articles of *CAR* for the targeted 6-year period.

References

Ary, D., Jacobs, L. C., & Sorensen, C. (2010). *Introduction to research in education* (8th ed.). Cengage: Belmont, CA.

Bliss, T. J. (2012) Twenty-eight years of the collegiate aviation review: An exploratory study of academic research in aviation education**.** *Collegiate Aviation Review, 30*(1), 1–15.

Campbell, D., & Stanley, J. (1966). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally.

Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed.). Los Angeles, CA: Sage.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, pp. 37–46.

Cohen, J. (1992), A power primer. *Psychological Bulletin,* 112(1), 155–159.

Fraenkel, J., Wallen, N., & Hyun, H. (2011). *How to design and evaluate research in education* (8th ed.). McGraw-Hill Higher Education: Boston, MA.

Gall, J. P., Gall, M. D., & Borg, W. R. (2007). *Applying educational research: How to read, do, and use research to solve problems in practice* (6th ed.). New York, NY: Allyn & Bacon.

Garner, A. C. (1992). *The disaster news story: The reader, the content and the construction of meaning.* (Doctoral dissertation). Retrieved from http://search.proquest.com.portal.lib.fit.edu/docview/304009360?accountid=27313

Gravetter, F. J., & Wallnau, L. B. (2013). *Statistics for the behavioral sciences* (9th ed.). Belmont, CA: Wadsworth, Cengage Learning.

Hankins, A. S. (2007). *Quality characteristics in aviation baccalaureate programs: A content analysis of collegiate publications*. (Doctoral dissertation). Retrieved from http://search.proquest.com/docview/304839231?accountid=27313

Horton, P. B., McConney, A. A., Woods, A. L., Barry, K., Krout II, H. L., & Doyle, B. K. (1993). A content analysis of research published in the journal of research in science teaching from 1985 through 1989. *Journal of Research in Science Teaching*, *30*, 857–869.

Hsu. T. C. (2005) Research methods and data analysis procedures used by educational researchers. *International Journal of Research & Method in Education*, *28*(2), 109–133, DOI: 10.1080/01406720500256194

Johnson, R., Gibson, B. J., Hamilton II, R. A., & Hanna, J. B. (2006). Collegiate aviation education and periodical usefulness: What aviation educators say. *Collegiate Aviation Review*, *24*(1), 82-100.

Jones D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation Space Environmental Medicine, 67*(6), 507–512.

Krippendorff, K. (1980). *Content analysis.* Beverly Hills, CA: Sage Publications.

Krippendorff, K. (2004a). *Content analysis: An introduction to its methodology*. Thousand Oaks, California: Sage.

Krippendorff, K. (2004b). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, *30*, 411–433.

Patten, M. L. (1998). *Questionnaire research: A practical guide*. Los Angels, CA: Pyrczak Publishing.

Rath, J. (1973). The emperor's clothes phenomenon in science education. *Journal of Research in Science Teaching*, *10*(3), 201–211.

Shaver, J., & Norton, R. (1980a). Randomness and replication in ten years of the American educational research journal. *Educational Researcher, 443*(3), 265–291.

Shaver, J., & Norton, R. (1980b). Population, samples, randomness, and replication in two social studies journals. *Theory and Research in Social Education*, *8*(2), 1–10.

Wallen, N., & Fraenkel, J. (1988a). An analysis of social studies research over an eight-year period. *Theory and Research in Social Education*, *16*, 1–22.

Wallen, N., & Fraenkel, J. (1988b). *Toward improving research in social studies education*. Boulder, CO: Social Science Education Consortium.

Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Newbury Park, CA: Sage Publications.

Wiegmann, D. A., Shappell, S. A., Boquet, A., Detwiler, C., Holcomb, K., & Faaborg, T. (2005). *Human error and general aviation accidents: A comprehensive, fine-grained analysis using HFACS*. Final Technical Report AHFD-05-08/FAA-05-03, May 2005. FAA.

Appendix A

Coding Form

Coding Form for Content Analysis of *Collegiate Aviation Review* (*CAR*)

Volume _____ Issue _____ Month/Year _____ Pages _____ Coder _____

Article Title: _____

Article Author: _____

Item	Yes	No	N/A	Comments
A. Author's Purpose / Justification for Study (choose one most obvious)				
Clearly stated				
Test of theory				
Replication (extension or follow-up)				
Grounded in literature				
Not clearly stated				
B. Research Question and Hypothesis (check all that apply)				
Research Question (RQ)				
RQ(s) clearly specified				
RQ(s) specified as objectives				
RQ(s) implied in purpose statement				
None given				
Hypothesis				
Research hypothesis				
Null hypothesis				
Alternate hypothesis				
Hypothesis provided as a prediction				
None given				
C. Sampling Issues				
Population				
Target population				
Accessible population				
None given				
Sampling strategy				
Random selection				
Convenience				
Volunteer				
Quota				
Purposive				
Systematic				
Snowball				
Census				
Other				
Can't tell				
Description of sample				
Adequate				
Marginal				
Sample representativeness				
Author's claim				
Based on random selection strategy				
None given				

Sample size				
Power analysis performed				
Margin of error (survey)				
Size specified				
Size per group specified				
Sample assignment				
Random assignment to groups				
Treatments randomly assigned				
Group equivalency demonstrated				
Not appropriate				
D. Instrumentation Issues				
Description given				
Validity information given				
Reliability information given				
None given (archived data)				
E. Type of Design				
Quantitative				
Preexperimental				
True experimental				
Quasiexperimental				
Ex post facto (causal-comparative)				
Correlational				
Survey				
Content Analysis				
Other				
Unknown/Unclear				
Qualitative				
Case study				
Narrative				
Grounded theory				
Phenomenology				
Ethnographic				
Content Analysis				
Other				
Unknown/Unclear				
F. Internal Validity Threats Discussed by Author (mark all that apply)				
Quantitative				
History				
Maturation				
Mortality (attrition)				
Selection bias				
Testing				
Regression				
Instrumentation				
Experimenter effect				
Subject effect				
Diffusion				
Location				
No threats discussed				

Qualitative				
Credibility				
Transferability				
Dependability				
Confirmability				
No threats discussed				
G. External Validity Discussion				
Population generalizability				
Ecological generalizability				
None given				
H. Data Analysis Approaches / Statistical Measures Reported				
Descriptive statistics (*M*, *Mdn*, *range*, *SD*, *freq.*, *%*)				
Inferential statistics				
t test (independent samples)				
t test (repeated measures/matched pairs)				
ANOVA (oneway)				
ANOVA (factorial)				
ANOVA (repeated measures/matched pairs)				
Covariance				
Correlation				
Regression				
Chi-square				
Not applicable (e.g., survey methodology)				
None given				
Statistical measures reported				
Confidence intervals				
Effect size				
Correlation coefficient				
Regression coefficient				
Power				
Other (specify in comments section)				
None given				
I. Conclusions				
Reported as truth (overgeneralization)				
Reported as trivia (undergeneralization)				
Reported as findings				
Reported as plausible explanations				
Reported as implications				
None given				
J. Recommendations				
Given with respect to future research				
Given with respect to practice				
None given				
K. Limitations and Delimitations				
Adequate				
Marginal				
None given				

Appendix B

IRB Documents

Florida Institute of Technology
Institutional Review Board

Notice of Exempt Review Status

From: Florida Tech Institutional Review Board
FWA00014339, Exp. 4/11/2017, IRB00001690

To: Safak Aktemur

Date: February 12, 2015

IRB Number: 15-023

Study Title: A review of aviation research: A content analysis of articles published in the Collegiate Aviation Review, 2007-2014

Dear Researcher:

Your research protocol was reviewed and approved by the IRB Chairperson. Per federal regulations, 45 CFR 46.101, your study has been determined to be minimal risk for human subjects and exempt from 45 CFR46 federal regulations and further IRB review or renewal unless you change the protocol or add the use of participant identifiers. This study is approved for one year from the above date. If data collection continues past this date, a Continuing Review Form must be submitted.

All data, which may include signed consent form documents, must be retained in a locked file cabinet for a minimum of three years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained on a password-protected computer if electronic information is used. Access to data is limited to authorized individuals listed as key study personnel.

The category for which exempt status has been determined for this protocol is as follows:

4) Research involving the **collection or study of existing data, documents, records, or specimens** if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, indirectly or through identifiers linked to the subjects.

Institutional Review Board Office
Dr. Lisa Steelman, Chair IRB
School of Psychology
(p) 674-8104
lsteelma@fit.edu
http://www.fit.edu/research/committees/irb/index.html

Florida Institute of Technology

RESEARCH INVOLVING HUMAN SUBJECTS
Exempt Application

This form shall be used if there is **minimal risk** to human subjects and one or more of the conditions below apply. If there is more than minimal risk associated with the research (none of the conditions below apply) or if the research utilizes a special population (children, prisoners, institutionalized individuals, etc.), please use the full application form found on the IRB website.

You should consult the university's document "Principles, Policy, and Applicability for Research Involving Human Subjects" prior to completion of this form. Copies may be obtained from the Office of Sponsored Programs and on the IRB website.

Name:	Safak Aktemur
Date:	February 4, 2015
Academic Unit:	College of Aeronautics
Email:	Saktemur2013@my.fit.edu
Title of Project:	A Review of Aviation Research: A Content Analysis of Articles Published in the *Collegiate Aviation Review*, 2007–2014

☒ 1) Research conducted in established or commonly accepted educational settings, involving **normal educational practices**, such as:
   a. research on regular and special education instruction strategies, or
   b. research on the effectiveness of or the comparison among instruction techniques, curricula, or classroom management methods.

☐ 2) Research involving the use of **educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior** unless:
   a. the subjects can be identified, directly or through identifiers linked to the subjects and
   b. any disclosure of subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.
   Note: This exemption does not apply to survey procedures or interviews involving minors.

☐ 3) Research involving the use of educational tests, survey or interview procedures, or observation of **public behavior** if:
   a. the subjects are elected or appointed public officials or candidates for public office or
   b. the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.

☒ 4) Research involving the **collection or study of existing data, documents, records, or specimens** if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, indirectly or through identifiers linked to the subjects.

☐ 5) Research and demonstration projects that are conducted by or subject to the approval of Department or Agency heads and that are designed to study, evaluate, or otherwise examine:
   a. **public benefit or service programs**,
   b. procedures for obtaining benefits or services under those programs,
   c. possible changes in or alternatives to those programs or procedures, or
   d. possible changes in methods or levels of payment for benefits or services under those programs.

☐ 6) **Taste and food quality evaluation** and consumer acceptance studies if:
   a. wholesome foods without additives are consumed or
   b. food is consumed that contains food ingredients found to be safe by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture.

1. List the objectives of the proposed project.

   The purpose of this study is to determine if actual practices of aviation researchers are consistent with commonly recommended research methods and procedures, and to inform the aviation community as to the state of the art with respect to eight selected volumes of *Collegiate Aviation Review* (CAR), which is a refereed journal published by the University Aviation Alliance (UAA). The research questions that will guide the study are as follows:

   1. To what extent do aviation researchers state the purpose of their study and corresponding research questions?
   2. To what extent do aviation researchers define their target/accessible populations?
   3. What sampling strategy do aviation researchers commonly use?
   4. To what extent do aviation researchers describe their sample, including representativeness, sample size, and assignment?
   5. To what extent do aviation researchers address instrumentation issues such as validity and reliability?
   6. What research methodology/design do aviation researcher commonly use?
   7. To what extent do aviation researchers give attention to threats to internal validity?
   8. What data analysis methods do aviation researchers commonly use?
   9. In what way do aviation researchers report their conclusions?
   10. To what extent do aviation researchers state their recommendations, limitations, and delimitations?

2. Describe the research project design/methodology. Discuss how you will conduct your study, and what measurement instruments you are using. If your project will use a questionnaire or structured interview, attach. Please describe your study in *enough detail* so the IRB can identify what you are doing and why.

   The research methodology is content analysis. I will review and code all the research articles published in volumes 25–32 of CAR. These volumes are being selected to reflect the status of published research in CAR over the most recent 8-year period. A total of 92 articles in the targeted volumes will be reviewed. A purposive selection method will be used to insure a balanced coverage of the time period of interest. The data collection instrument will be a coding form my advisor and I developed. The categories on the form will correspond to the 10 research questions given above. A copy of this form is provided with this application. My advisor and I will read and code each article independently and then meet to compare our coding for intercoder reliability. If there are any areas of disagreement, we will then ask a third person to review and code the problematic article to arrive at a consensus.

3. Describe the characteristics of the subject population, including number, age, sex, and recruitment strategy (attach actual recruitment email text, recruitment flyers etc).

   The target population is all articles published in CAR and the accessible population is all articles published in CAR between 2008 and 2014.

4. Describe any potential risks to the subjects (physical, psychological, social, legal, etc.) and assess their likelihood and seriousness. Research involving children must carefully assess risks and describe the safeguards in place to minimize these risks.

   There are no known risks to any subjects because the sample will consist of published articles and the names of the authors will not be revealed.

5. Describe the procedures you will use to maintain the confidentiality and privacy of your research subjects and project data.

   All articles will be coded manually using paper copies of the coding form. For organizational purposes, the names of the article authors will be recorded on this form. All completed forms will be stored in my advisor's locked file cabinet in his office. After data analysis is completed, he and I will then shred the coded forms. No authors' names will be reported in my thesis or in any article I prepare for publication.

6. Describe your plan for informed consent (attach proposed form).

   Informed consent is not needed because no human subjects will be participating in the study. I will be reviewing and coding the content of published articles and reporting frequencies and percentages relative to the categories given on the coding form.

7. Discuss the importance of the knowledge that will result from your study and what benefits will accrue to your subjects (if any).

To date, no similar content analysis studies have been conducted with respect to aviation research published in refereed journals. As a result, this study could be the first step to determine if the actual practices of aviation researchers are consistent with commonly recommended research methods and procedures. If other content analyses of refereed aviation journals are conducted periodically, then the study will contribute to the aviation research community for the development of scholarly publications in the area. The researchers planning to publish their studies and the editors of refereed aviation journals can benefit from the findings of the study. In the light of findings of the study the editor of CAR can have a chance to criticize their standards toward the quality of the articles published in the most recent 8 years. Also, the researchers in the aviation field who endeavor to publish their research will have the opportunity to design their studies based on the findings of the proposed study.

8. Explain how your proposed study meets criteria for exemption from Institutional Review Board review (as outlined on page 1 of this form).

This study satisfies the exempt criteria for research involving human subjects because it involves the collection and study of existing documents that are publicly available to members of UAA (Paragraph 4). The proposed study also will be conducted in an established or commonly accepted educational setting involving normal educational practices (Paragraph 1).

Signature Assurances

I understand Florida Institute of Technology's policy concerning research involving human subjects and I agree:

1. to accept responsibility for the scientific and ethical conduct of this research study,
2. to obtain prior approval from the Institutional Review Board before amending or altering the research protocol or implementing changes in the approved consent form,
3. to immediately report to the IRB any serious adverse reactions and//or unanticipated effects on subjects which may occur as a result of this study,
4. to complete, on request by the IRB, a Continuation Review Form if the study exceeds its estimated duration.

STUDENT:
~~PI Signature~~ _____ Date 2/4/2015

Advisor Assurance: If primary investigator is a student
This is to certify that I have reviewed this research protocol and that I attest to the scientific merit of the study, the necessity for the use of human subjects in the study to the student's academic program, and the competency of the student to conduct the project.

Major Advisor Michael A. Gallo Date 2/4/2015

Major Advisor (print) MICHAEL GALLO

Academic Unit Head: It is the PI's responsibility to obtain this signature
This is to certify that I have reviewed this research protocol and that I attest to the scientific merit of this study and the competency of the investigator(s) to conduct the study.

Academic Unit Head _____ Date 2.5.15

FOR IRB USE ONLY

IRB Approval _____ Date 2-12-15
Name

15-023
IRB #