Spring 2019

An Investigation of General Criteria for Assessing Space Flight Systems of Diverse Mission Concept Designs

Cindy L. Daniels
Old Dominion University, cldaniels020@icloud.com

AN INVESTIGATION OF GENERAL CRITERIA FOR ASSESSING SPACE FLIGHT SYSTEMS OF DIVERSE MISSION CONCEPT DESIGNS

By

Cindy L. Daniels
B.S. Mathematics August 1981, Northern Michigan University
M.S. Information Systems May 1999, The George Washington University
Master of Engineering Management January 2000, The George Washington University

A Dissertation Prospectus Submitted to the Faculty of Old Dominion University in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

ENGINEERING MANAGEMENT

OLD DOMINION UNIVERSITY
2019

Approved by:

Resit Unal (Director)

C. Ariel Pinto (Member)

Pilar Pazos (Member)

Kaitlynn Castelle (Member)

ABSTRACT

AN INVESTIGATION OF GENERAL CRITERIA FOR ASSESSING SPACE FLIGHT SYSTEMS OF DIVERSE MISSION CONCEPT DESIGNS

Cindy L. Daniels

Old Dominion University 2019
Director: Dr. Resit Unal

The purpose of this research is to investigate the general criteria for assessing the technical implementation risk factors of proposed space science missions at the mission concept stage. According to the National Aeronautics and Space Administration (NASA) Space Flight Program and Project Management Handbook (NASA, 2012), the mission concept review objectives are "To evaluate the feasibility of the proposed mission concept(s) and its fulfillment of the program's needs and objectives. To determine whether the maturity of the concept and associated planning are sufficient to begin Phase A" (p.33). Experts previously defined two technical risk factors, to assess aspects of the space flight systems and mission design and operations of proposed mission concepts. Criteria were developed to address these two technical risk factors, which are comprised of 23 criteria. The space flight systems factor was assumed to be addressed by 16 criteria, while seven criteria were assumed to address the mission design and operations factor.

The criteria were developed by experts approximately 20 years ago, and no research has previously been conducted to determine whether all 23 of the evaluation criteria are necessary for assessing the implementation risk of proposed space flight systems and mission design and operations for proposed mission concepts. NASA uses these 23 criteria to conduct expert peer reviews to assess the implementation risk of over 500

unique space science mission concept proposals. An expert peer review process is used because the proposed concepts lack the detailed design information necessary for a quantifiable assessment of risk. The result of the expert peer review of each proposal is a set of ratings with a paragraph explaining the rationale for each rating, based on the 23 criteria.

This research used 356 records from past assessments of proposed mission concepts that have been assessed using a five-level qualitative rating scale. A research approach which utilizes exploratory factor analysis and past records to analyze the ratings of the 23 criteria was used. Factor analysis was used to determine if the current factor structure was valid, whether all criteria had substantial loadings on the current factor, and whether all current criteria were necessary. Factor analysis was also used to determine if any of the criteria measured the same construct. This research used a discriminatory power scale to code criteria scores for factor analysis and to identify the criteria of significance to decision makers. This research identified criteria that could be eliminated or could be combined with other criteria. A result of this research is a reduced set of criteria for assessing space flight systems and for mission design and operations that can be accomplished by an expert peer review panel for a diverse set of space mission concepts. A refined set of criteria could result in a less expensive and quicker evaluation process. This can enable decision makers on early assessments of space flight systems to make decisions more efficiently by allowing them to focus only on the most important criteria. This refined set of criteria contributes to the literature on the qualitative risk assessment of space flight systems and mission design and operations. This research is supported by the existing body of literature in using factor analysis to refine a measurement instrument. Using factor analysis to evaluate criteria for spaceflight

systems contributes another application of the use of factor analysis, beyond its historical use in psychology, education, and healthcare (Williams, Onsman, & Brown, 2010). This research provides a method that engineering managers can use to analyze and to refine a qualitative measurement instrument for assessment by a group of experts. This method could be useful in assessments that require a broad scope of required expertise.

ACKNOWLEDGMENTS

I would like to thank the following individuals for their contribution to my efforts in completing this dissertation:

My dissertation committee: Dr. Resit Unal, Dr. Ariel Pinto, Dr. Pilar Pazos, and Dr. Kaitlynn Castelle. I would particularly like to thank Dr. Unal for his patience and his encouragement in keeping me focused, both to continue this research and to finish this dissertation. I am grateful to Dr. Castelle for her helpful discussions on factor analysis and for her feedback on my research. I have enjoyed all of my graduate coursework at ODU. Dr. Pazos's class on research methods was instrumental in identifying a method and an approach that could be used to conduct this research. Dr. Pinto's classes on risk analysis provided a context for this research on risk assessments in spaceflight.

I would like to thank my husband Charlie, who has been patient and understanding of the long hours (and years) necessary to complete this work. I am fortunate to have someone in my life who understood the sacrifices in time, activities, and events that had to be made in order to complete my dissertation.

I would also like to thank Dr. Waldo Rodriguez and Victor Lucas at NASA who were kind enough to step up and act for me during periods when I needed to be away from work for extended times, in order to complete this research.

I am grateful to all of you who helped me in myriad ways. I could not have completed this research without your support and guidance.

LIST OF ACRONYMS AND ABBREVIATIONS

AO = Announcement of Opportunity

CBR = Criteria Based Review

FA = Factor Analysis

HVHDP = High or Very High Discriminatory Power

KMO = Kaiser-Meyer-Olkin

MSA = Measure of Sampling Adequacy

NASA= National Aeronautics and Space Administration

PCA = Principal Component Analysis

TMC = Technical, Management, and Cost

SMD = Science Mission Directorate

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Background

This research investigated the current general criteria for assessing the technical implementation risk factors of proposed space science missions at the mission concept stage. The National Aeronautics and Space Administration (NASA) conducts expert peer reviews of proposed space science missions at this stage. According to the NASA Space Flight Program and Project Management Handbook (2012), the mission concept review objectives are “To evaluate the feasibility of the proposed mission concept(s) and its fulfillment of the program's needs and objectives. To determine whether the maturity of the concept and associated planning are sufficient to begin Phase A” (p. 33). The mission concept proposals, reviewed with an expert peer review process, include earth observing missions, planetary science missions, heliophysics missions, and astrophysics missions. The mission concepts received for each proposal evaluation are defined by the space science discipline and by a cost cap specified for each separate program competition. NASA has defined programs that have specific space science goals that are competed on a regular basis. The competed programs for space science mission concepts include the New Frontiers program, the Discovery program, the Mars program, the Explorers program, and the Earth Venture program. The New Frontiers, Discovery, and Mars programs are space science planetary missions. The Explorers program includes astrophysics and heliophysics space science missions. The Earth Venture program includes earth science missions. The scope of these competitions is limited by a specified

cost cap which varies by program. A summary of programs, science disciplines, and cost caps is provided in Appendix D.

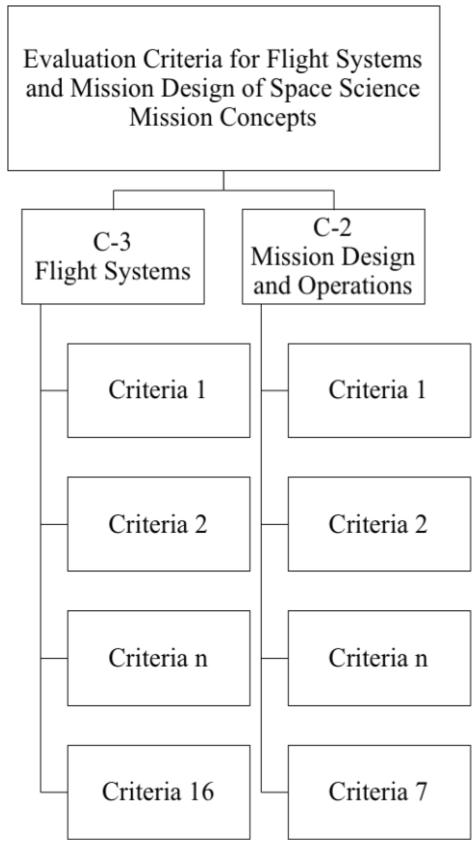
NASA solicits mission proposals for space science missions via a public Announcement of Opportunity (AO). Proposals submitted in response to a NASA AO must be developed at the proposer's expense. Proposals to an AO are considered to be in the concept development stage (NASA, 2012). An expert peer review process is used because the proposed concepts lack the detailed design information to enable a quantifiable assessment of risk.

NASA uses an expert peer review process to assess the technical implementation risk of all space science mission concept proposals received in response to an AO. The expert peer review assesses each proposed mission concept against a standard set of criteria for technical implementation risk. The result of each expert peer review of a mission concept proposal is a set of ratings and text based on standard evaluation criteria. The rating is accompanied by a paragraph providing the rationale for the rating. The NASA standard AO template (NASA, 2014a) defines five factors that comprise the "Technical, Management, and Cost (TMC) Feasibility of the Proposed Mission Implementation, Including Cost Risk" (p. 61-63). The full text of the five TMC factors is shown in Appendix A. A review panel of engineers with relevant experience assesses these five factors. This panel will be referred to in the remainder of this paper as "the expert peer panel." A separate science peer panel uses a different standard set of factors and evaluation criteria to evaluate the science merit and science implementation merit of the proposed mission concept. The ratings of the criteria of both review panels are considered by the selecting official when making selections for further maturation of the mission concept. The process of assessing proposals in response to an AO is referred to

as "a step one evaluation." The selection official usually selects several mission concepts for a phase A study, which is also referred to as "step two." NASA funds the proposal team of each selected mission concept to conduct a phase A study to further develop the mission concept. At the end of the phase A study, NASA uses an expanded set of evaluation criteria to assess each phase A study before making a final selection of one or more missions that will proceed to development and flight.

This research focuses on two factors and on the associated evaluation criteria used by the expert peer panel to assess the unfunded mission concept proposals in step one. Over the last 20 years, a standard set of factors and associated evaluation criteria has been used by expert peer panels to evaluate mission concepts. These expert peer panels have evaluated each mission concept against 23 criteria which were intended to measure two of the five technical implementation factors. This research is limited to the two technical implementation factors that address the flight systems and the mission design. The NASA standard AO defines the two factors as the following (NASA, 2014a): "C-2. Adequacy and robustness of the mission design and plan for mission operations … C-3. Adequacy and robustness of the flight systems" (p. 61). Factor C-2 will be referred to as *mission design* for the remainder of this paper. Factor C-3 will be referred to as *flight systems* for the remainder of this paper. The criterion number and the definition for each criterion are shown in Appendix B. Seven criteria are defined to measure aspects of the mission design factor, and 16 criteria are defined to measure aspects of the flight systems factor. This research is limited to the 23 criteria that are assumed to address the flight systems and the mission design factors. Figure 1, below, shows a summary of the flight systems and mission design evaluation factors and their associated evaluation criteria.

Figure 1: Flight Systems and Mission Design Criteria



Problem Statement

NASA currently utilizes two separate evaluation factors to evaluate flight systems and mission design of space science mission concepts. There is a total of 23 criteria that are currently assumed to measure these two factors, but no study, prior to this research, had been carried out to determine if all of the 23 criteria currently employed by the expert peer panels are necessary to assess the flight systems and mission design factors. The current two factor structure assumes that 16 criteria are necessary to assess the flight

system factor and that seven criteria are necessary to assess mission design factor. This research explored the 23 criteria and the relationship to the factor that they are intended to assess, in order to determine the validity of the current factor structure.

Each of the 23 evaluation criteria are rated on a five-level scale by the expert peer review. Due to the large number of criteria, it is an expensive and time-consuming process to staff an expert peer review panel, a panel with sufficient expertise to cover all of the criteria and to conduct a review of each criterion for each mission concept proposal. The large number of criteria creates a barrier to proposers, due to the large amount of information that must be prepared, in order to address all 23 criteria. This research used exploratory factor analysis to test the validity of the current factor structure. A discriminatory power scale was used to code the data and rank the criteria for the analysis. Exploratory factor analysis was conducted on the full data set and on subsets, using criteria with high and very high discriminatory power.

The Delimitations

This study was limited to analyzing the 23 evaluation criteria that are used to assess the factors of flight systems and mission design as defined in the NASA Announcement of Opportunity Standard PI-led Mission AO (NASA, 2014a). These two factors and their associated criteria are listed in Appendix B. AOs may have additional criteria above the standard 23 criteria for flight systems and mission design. However, this research is limited to the standard 23 criteria.

Significance of Problem

The number of criteria used to evaluate proposed mission concepts determines the time and the cost for proposers to prepare a proposal and for expert review panels to evaluate proposals for space science mission concepts. Proposers make significant

investments in time and money to develop proposals that address the 23 criteria for flight systems and mission design. Consequently, the proposing community has asked NASA to simplify and to reduce the information that they must provide for a mission concept proposal. NASA's time and cost to conduct an expert peer panel review of space science mission concepts is driven by the number of criteria and the number of mission concept proposals reviewed. Reducing the time that it takes for an expert peer review panel to evaluate proposals could reduce the length of time from proposal submission to selection. Reducing the time from proposal submission to selection could provide more time for the development of the spaceflight systems. NASA has excess launch capabilities on the launch vehicle for an already selected primary space science mission. The excess launch capability could be utilized by a secondary payload, if it can launch when the primary mission is ready. An example of this is the NASA Small Innovative Mission for Planetary Exploration (SIMPLEx) Announcement of Opportunity (NASA, 2018) that solicits small complete missions to fly on an already selected primary spacecraft mission. The selection of a secondary payload to launch with a primary mission requires that the secondary mission must meet the launch date of the primary mission. Reducing the time from proposal submission to selection will provide more development time for the secondary payload to meet the primary mission launch date.

Each of the criteria must be addressed in each proposed mission concept and represents a design constraint to the mission concept proposal. If the research results in a reduction in the number of criteria, the constraints on the mission concept are reduced. From a system engineering perspective, a reduction in the number of criteria to only those necessary would reduce over-constraining a mission concept design at this very early stage of design.

The time and cost to develop a space flight proposal at the mission concept stage also acts as a barrier to potential proposers. If the cost to propose a mission is reduced, more organizations may submit proposals, and organizations that already participate may decide to submit multiple proposals. More proposals will increase the number of new space science investigation ideas, which represent an additional choice to NASA for selecting the best scientific investigation(s).

CHAPTER 2

LITERATURE REVIEW

The literature review will address topics relevant to this research, among them space flight mission concept design criteria, qualitative and quantitative risk assessment based on individual and group expert judgment, expert judgement processes for individual and groups, decision making under uncertainty, and exploratory factory analysis applications.

Griffin and French (2004) described the process for defining design criteria for a space flight mission: “The basic goals and constraints of a given space mission will generally be defined by the user or customer for the resulting system. Such goals will usually be expressed in terms of the target and activity” (p. 7). An example of a target and activity provided by Griffin and French (2004) is “Deploy a spacecraft in a geosynchronous communication satellite capable of carrying 24 transponders" (p. 7). Pisacane (2005) states “The pre-phase A- advanced studies - product is a set of mission goals and one or more concepts that can satisfy the goals” (p. 11-12). Once a space flight mission is specified, in terms of the destination of the spacecraft and the activity to be conducted, the system engineering process proceeds to specify design criteria for spacecraft flight systems. Griffin and French (2004) state “the type of mission to be flown and the performance requirements that are imposed define the spacecraft design that results” (p. 17).

The NASA competed missions covered under this research do not define a specific space science mission with one target destination and specified instruments for scientific measurements. A NASA space science mission AO only defines a general area

of space science, a cost cap for the selected mission, a start date for spacecraft development, a launch date, and an end of mission date. For example, the NASA Discovery 2014 Announcement of Opportunity (NASA, 2014b) states “The NASA Science Mission Directorate (SMD) is addressing this strategic goal through Strategic Objective 1.5: Ascertain the content, origin, and evolution of the solar system and the potential for life elsewhere … Investigations may target any body in the Solar System except for the Earth and Sun, in order to advance the objectives … ” (p. 2-3). The NASA Discovery 2010 Announcement of Opportunity (NASA 2010) also states “Investigations may target any body in the Solar System, including Mars and Earth‘s Moon” (p. 3). The NASA New Frontiers 2009 Announcement of Opportunity (NASA, 2009) provides guidance, as follows, to proposers: “Proposals shall describe a science investigation that addresses a preponderance of the science objectives for one out of any of the eight mission concepts” (p. 4). The competed mission programs considered under this research also include the Explorers program and the Earth Venture program, which allow a broad scope of missions to be proposed by only limiting the mission to a specified space science discipline. For instance, the NASA Astrophysics Explorer Program 2016 MIDEX Announcement of Opportunity (NASA, 2016a) states “The goal of NASA’s Explorers Program is to provide frequent flight opportunities for high quality, high value, focused astrophysics science investigations that can be accomplished under a not-to-exceed cost cap and that can be developed relatively quickly, generally in 36 months or less, and executed on-orbit in less than three years“ (p. 3).

The NASA Earth Venture Mission 2 (EVM 2) Announcement of Opportunity (NASA, 2015b) also references broad goals as opposed to a specific target. The NASA strategic plan (NASA, 2014c) states that NASA strategic goal 2 is to “Advance

understanding of Earth and develop technologies to improve the quality of life on our home planet" (p. 25). The NASA Earth Venture Mission 2 (EVM 2) Announcement of Opportunity (2015b) references NASA strategic goal 2 and states “The NASA Science Mission Directorate (SMD) is addressing this strategic goal by pursuing the Earth Science Goals” (p. 1-2).

Since the NASA Announcements of Opportunity to request proposals for space science investigations do not specify one target and an activity of the mission, a diverse set of mission concepts to be evaluated are received. NASA has developed flight systems and mission design evaluation criteria that are not specific to a target and activity. The NASA flight systems and mission design evaluation criteria are defined in the NASA Standard AO for PI-Led Missions (NASA, 2014a); these are shown in Appendix A in paragraph form and in Appendix B in a list by criteria number. The criteria for both factors have been defined by experts in those areas. The flight system factor is addressed by a set of 16 criteria and the mission design factor is addressed by a set of seven criteria. These 23 general criteria for flight systems and mission design are used by expert peer review panels, using expert judgement, to assess proposed mission concepts. These are general criteria that could be applied to the flight systems and mission design of any spaceflight mission.

Pisacane (2005) identifies products that are developed to respond to a NASA AO, which include “Develop top-level requirements, Develop subsystem-level requirements….Identify system and subsystem characteristics….“ (p. 13). The descriptions of the products reflect the lack of maturity of the design of proposed flight systems and mission design in the proposals assessed by the expert review panel.

Once a mission is selected for further study or development, many analytical methods are available to optimize specific parts of the flight systems. Some examples include an analytical hierarchy process that was used to decide between six concepts for lunar surface power based on seven evaluation criteria (Matthews, Coomes, & Khan, 1994), and a genetic algorithm that was used for a trade-off of low earth orbit spacecraft power supply system (Mohamed, Amer, Mostafa, & Mahmoud, 2016). These analytical methods are useful for optimizing a flight system for a specific mission. Since NASA receives proposals for missions that do not have a common target and a specific mission, these comparative analytical methods are not appropriate. In addition, each of these analytical methods requires specific information that is not available for pre-Phase A proposals. Morgan (2014), stated:

> Although such analytical strategies can provide valuable insight, they can never hope to include all relevant factors. In such situations, the community of applied decision analysis has long used quantitative expert judgments in the form of subjective probability distributions that have been elicited from relevant expert … Expert elicitation can make a valuable contribution to informed decision making". (p. 7176)

According to Goossens, Cooke, Hale, and Rodić-Wiersma (2008), "Expert judgement has always played a large role in science and engineering. Increasingly, expert judgement is recognized as just another type of scientific data, and methods are developed for treating it as such" (p. 236). Expert judgment has been used for quantitative assessments in spaceflight applications. In a study by Monroe, Lepsch, and Unal (2002), individual experts were surveyed to determine "the uncertainty associated with weight estimating relationships for a launch vehicle design study" (p. 1). Another

study developed a methodology to elicit expert judgment, to aggregate the data, and to determine uncertainty distributions in support of decision analysis in high technology systems design (Chytka, Conway, & Unal, 2006).

NASA uses expert judgement to assess the 23 criteria for flight systems and mission design for pre-Phase A space science mission proposals. The expert peer review panel conducts a qualitative assessment, since the criteria are qualitative in nature. For instance, criterion 7 (see Appendix B) is "an assessment of the proposer's understanding of the processes required to accomplish development." The scope of flight systems and mission design includes many topics which are addressed by the criteria. However, the criteria themselves cover large topic areas that consider many elements. Topics covered by each criterion are shown in Appendix C Table C.1. A qualitative assessment is used by NASA, since the amount of quantitative information on each of the criteria is limited in a pre-Phase A proposal.

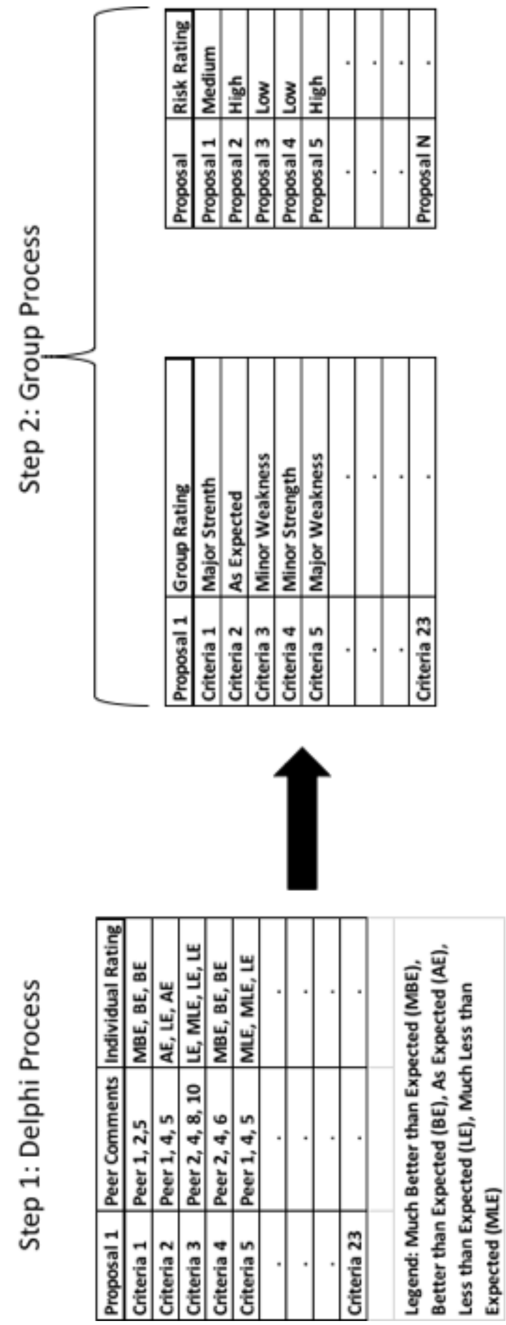
The expert review process for pre-Phase A proposals includes both an individual review using a modified Delphi process and a group discussion process. In the health field, when rigorous controlled studies based on evidence are not available to provide diagnostic criteria, "formal group consensus methods have been developed to organize subjective judgements and to synthesize them with the available evidence" (Nair, Aggarwal, & Khanna, 2011, p. 95). Nair et al. (2011) describe four consensus techniques, which include the Delphi method, Nominal Group Technique, the RAND/UCLA Appropriateness Method, and the National Institute of Health's (NIH) consensus development conference. Nair et al. (2011) identified a key advantage of the Delphi approach, which is that "each participant expresses their opinion freely and impersonally" (p. 98). Jones and Hunter (1995) found that a feature of the Delphi

process is that it avoids issues of dominance by using a questionnaire for participants to answer in private. The first step of the NASA expert peer review process is a Delphi process. Each expert rates the proposal against the criteria individually, on a five-level scale, and provides comments to substantiate the rating. The individual review is done only once in the NASA process. However, in some descriptions of the Delphi process, the participants go through several rating rounds of a revised questionnaire (Jones & Hunter, 1995). After the NASA expert individual rating process is complete, a group process is used to review and refine the ratings and rationale. According to Clemen and Winkler (2006), "the fundamental principle that underlies the use of multiple experts is that a set of experts can provide more information than a single expert" (p. 188). Nair et al. (2011) found the NIH Consensus Development Conference has two advantages. They stated that the first advantage is that a "mix of practicing physicians, researchers, consumers and others … come together and jointly evaluate an existing technology" (Nair et al., 2011, p. 102). Another advantage is that an unbiased panel was used (Nair et al., 2011). The NASA group discussion process has some similarities; among them are that there are a mix of experts in the different areas to be evaluated, and that the participants in the group discussion are unbiased. NASA conducts a rigorous institutional and personal conflict-of-interest review screening of the experts before they can participate in the group discussion which reviews all of the individual comments and ratings of the criteria.

Several rounds of group discussion and revisions occur in the process under study, in order to reach general agreement on the ratings of each criterion on a five-level scale. The ratings are substantiated by a paragraph detailing the rationale for the rating.

The result of the group discussion is the expert peer review panel ratings that are the subject of this research. The expert peer review panel process is shown in Figure 2.

Figure 2 Two Step Expert Peer Review Process



Similar two step processes to gather and refine expert opinions are described in the literature. For instance, Nair et al. (2011) states "In practice, a combination of 2 formal consensus methods or their modifications can be used in a 2-step process, where one method is used for item generation or some initial consensus and then the other method is used for final consensus" (p. 97).

Floyd and Widaman (1995) state that exploratory factor analysis has two purposes. One use is "to identify a set of more general latent variables, or factors, that explain the covariances among the measured variables. In theory, these latent variables are the underlying causes of the measured variable" (p. 286-287). In addition, Floyd and Widaman (1995) defined a second use, as follows:

> The second and related use of exploratory factor analysis is for data reduction, in which a set of measured variables is to be combined into summary indices. The goal is to discover optimal weightings of the measured variables so that a large set of related variables can be reduced to a smaller set of general summary scores that have maximal variability and reliability. (p. 286-287)

The records of the expert peer review panels were analyzed using exploratory factor analysis for both purposes described by Floyd and Widaman (1995). Exploratory factor analysis was applied to the ratings of the 23 criteria on the expert review panel records in order to identify the latent variable(s) that explain the covariances of the measured variables (the 23 rated criteria). The exploratory factor analysis of the expert peer panel records can also be used for reducing the set of measured variables (in this case, the 23 criteria) to a smaller set of criteria.

Tabachnick and Fidell (2013) stated that "In exploratory FA, one seeks to describe and summarize data by grouping together variables that are correlated. The

variables themselves may or may not have been chosen with potential underlying processes in mind" (p. 614).

There are many examples of using exploratory factor analysis to identify factors that explain covariances among variables and for data reduction. Johnson and Stevens (2001) used exploratory factor analysis on an existing school environment instrument which included 56 measured variables and eight factors. Exploratory factor analysis was conducted on a random sample of half of the completed surveys; this resulted in a reduction of 13 measured variables on the survey and a reduction in one factor (Johnson & Stevens, 2001).

Factor analysis is a commonly used technique in the fields of social science, psychology, health, and medicine. For example, a Google Scholar search using the term "exploratory factor analysis social science" returned about 2,710,000 results. A Google Scholar search using the term "exploratory factor analysis psychology" returned about 1,950,000 results. A Google Scholar search of the term "exploratory factor analysis health" returned about 1,710,000 results. A Google Scholar search of the term "exploratory factor analysis medicine" returned about 1,650,000 results.

This research proposes to use exploratory factor analysis in the field of engineering managment. The expert peer review panel is composed of engineers. The criteria assessed by the expert peer review panel includes evaluation of hardware and software proposed for spaceflight systems and the engineering development processes. A Google Scholar search of the term "exploratory factor analysis and design criteria resulted in 258 results. Hsu (2012), in a study of criteria for blog design, used exploratory factor analysis. Hsu (2012) conducted a literature review, and expert interviews were used to develop the measurement instrument, which included 23 criteria.

After the exploratory factor analysis, 23 criteria were retained and five factors were identified. This in an example of software design application of exploratory factor analysis. Tran and Molenaar (2014) used exploratory factor analysis in a study of risk factors in the design-build project delivery method used for selection for highway design and construction projects (2014). They developed an initial list of 39 risk items from previous research and workshops (Tran & Molenaar, 2014). The exploratory factor analysis resulted in retaining 23 of the risk items grouped into seven risk factors (Tran & Molenaar, 2014). This is an example of using exploratory factor analysis in an engineering application. Tran and Molenaar's measurement instument used an ordinal Likert scale with explanations for the rating (Tran & Molenaar, 2014). Tzeng, Chiang, and Li (2007) used exploratory factor analysis in a study to develop criteria for e-learning programs and stated that "When the evaluation criteria in real complex problems are too large to determine the dependent or independent relation with others, using factor analysis can verify independent factors" (p. 1030).

This research explored the current two factor structure for flight systems and mission design and assessed whether the current criteria load on the assumed factor of either flight systems or mission design and if all criteria are necessary. Each criterion was coded by the discriminatory power rating that identifies the significance of the rating. Alternative factor structures were assessed for generalization and for subgroup analysis. This research recommends a reduced set of general evaluation criteria to assess flight systems and mission design of a wide range of space flight mission at the mission concept stage. The criteria will not be restricted to a specific mission and target, since the research is based on unique missions with different targets and goals. The research will contribute to the current practice of expert peer panel reviews of spaceflight systems and

mission design by providing a revised set of criteria to evaluate flight systems and mission design. A substantial reduction in the current 23 critieria would reduce the time and money spent on the expert review, and it will also reduce the proposers' cost and their time to produce a proposal. Time and money freed up from this process can then be utilized to provide more development time for selected missions. Also, an efficient expert review panel will allow NASA to be able to quickly take advantage of secondary launch capacity. A reduction in the number of criteria and the time to evaluate proposals would also lower the barrier to propose a mission concept. This could result in additional proposals and additional new scientific ideas to consider for selecting the next space science mission.

Analysis of records on which decisions were made in the past to select space science missions is an example of how the analysis of data on which past decisions were made can contribute to better, more efficient decision making.

Factor analysis has been used extensively in the fields of social science, psychology, health, and medicine. This research adds to the literature on exploratory analysis by demonstrating that exploratory factor analysis can be used in a spaceflight engineering application.

This research also demonstrates a method for analyzing the product of a group decision process. This research adds to the literature of conducting research on group decision making.

The results of this research identify criteria with high discriminatory power to use as general evaluation criteria to assess the flight systems and the mission design of a wide range of space flight mission. The set of criteria could be used by analysts to evaluate a broad range of opportunities in commercial space. The use of an expert peer panel in

conjunction with the refined set of criteria would provide a very cost-efficient method to use in assessing a large number of diverse commercial space opportunities at the mission concept stage.

Below is a summary of the literature reviewed in Table 1. The shaded area represents the research to be addressed in this paper.

Table 1: Literature Review

		Space Flight Mission Concept		Analysis Technique	Expert Judgment					Engineering Application		Analysis
Author	Space Flight Application	Mission Specific Design Criteria	Non-Mission Specific Design Evaluation Criteria	(Non-EFA/CFA)	Individual	Group	Quan.	Qual.	Non Engineering Application	Hardware	Software	EFA
Griffin & French 2004	X	X										
Pisacane 2005	X	X										
Matthews, Coomes, & Khan 1994	X	X		X								
Mohamed, Amer, Mostafa, & Mohmoud 2016	X	X		X								
Monroe, Lepsch, & Unal, 2002	X			X	X		X			X		
Chytka, Conway, & Unal, 2006	X			X	X		X			X		
Nair et al., 2011					X	X		X	X			
Johnson and Stevens 2001					X			X	X			X
Hsu 2012					X			X			X	X
Tran & Molenaar 2014					X			X		X		X
Daniels Research	X		X		X	X		X		X	X	X

Legend: Quantitative (Quan.), Qualitative (Qual.), Exploratory Factor Analysis (EFA)

CHAPTER 3

RESEARCH METHODOLOGY

Research Question

The primary research question is: Are the 23 criteria currently used to assess space flight systems and mission design at the mission concept stage all necessary? This question will be addressed by analyzing past records of expert peer reviews that have assessed space flight systems and mission design using a standard set of 23 criteria.

Hypotheses

The criteria numbers below correspond to the same criteria numbers used in Appendix B.

The following hypotheses will be tested as a result of this research:

H0a: All 23 criteria are necessary to assess flight systems and mission design.

H0b: Criteria N is necessary to assess flight systems and mission design where N = 1 through 23.

H0c: Criteria 1 through 16 measure the flight systems factor.

H0d: Criteria 17 through 23 measure the mission design factor.

H0e: There are only two factors, which are the flight systems factor and the mission design factor.

Research Method

A deductive approach was used to address the research problem. An Ex Post Facto design approach was used to analyze past records of expert peer panel ratings of the standard 23 criteria used to assess flight systems and mission design for space science mission concepts. Leedy and Ormrod (1993) describe ex post facto designs as follows:

> Ex post facto designs (the term *ex post facto* literally means “after the fact”) provide an alternative means by which a researcher can investigate the extent to which specific independent variables (a virus, a modified curriculum, a history of family violence, or a personality trait) may possibly affect the dependent variable(s) of interest. Although experimentation is not feasible, the researcher identifies events that have already occurred or conditions that are already present and then collects data to investigation a possible relationship between these factors and subsequent characteristics or behaviors. (p. 238-239)

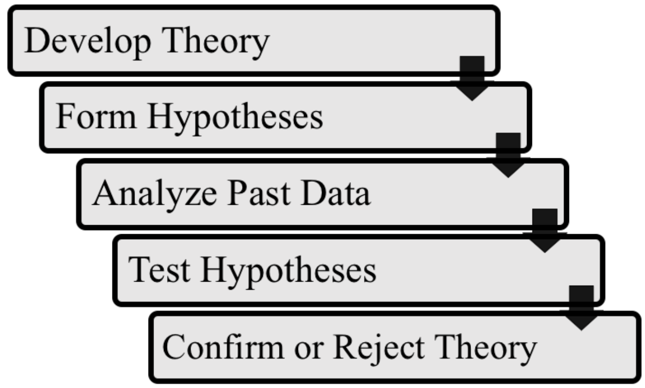
Leedy and Ormrod (1993) further describe ex post facto designs as follows:

> Although an ex post facto study lacks the control element – and so does not allow us to draw definite conclusions about cause and effect – it is nevertheless a legitimate research method that pursues truth and seeks a solution of a problem through the analysis of data. Science has no difficulty with such a methodology. Medicine uses it widely in its research activities. (p. 239)

The research method is based on a postpositive world view which uses the scientific method as the accepted approach, and according to Creswell (2018), “Thus, in the scientific method - the accepted approach to research by postpositivists - a researcher begins with a theory, collects data that either supports or refutes the theory, and then makes necessary revisions and conducts additional tests” (p. 5).

The research method is shown below in Figure 3.

Figure 3: Research Method



To accomplish the research, records from past expert peer review panels were analyzed. Each criterion was evaluated and rated by the expert peer panel and was documented on a Criteria Based Review (CBR) form. Each of the criteria was rated by the expert peer panel on a five-level adjectival scale and was coded for this research as discussed below. Exploratory factor analysis was used to test the hypotheses. Tabachnick and Fidell (2013) stated, “If the researcher had generated hypotheses regarding both the number and nature of the factors expected of graduate students, comparisons between the hypothesized factors and the factor solution provide a test of the hypotheses” (p. 616).

The number of factors to extract can be specified *a priori*, and according to Hair, Anderson, Babin, and Black (2010), this is “useful when testing a theory or hypothesis about the number of factors to be extracted. It also can be justified in attempting to replicate another researcher’s work and extract the same number of factors that was previously found” (p. 109). Burnett and Dart (1997) agreed that exploratory factor analysis can be used to test an existing instrument factor structure by stating that “Conventional factor analytic techniques are also used to validate the factor structure of

existing instruments using samples which differ in characteristics from the original scale development sample" (p. 126). This research approach was used by Tsai, Pietrzak, Southwick, and Harpaz-Rotem (2011) to conduct a study "aimed to extend previous research by exploring (1) the factor structure of two of the most commonly used screening measures of PTSD and depression" (p. 311). They "hypothesized that an exploratory factor analysis would yield a factor solution characterized by both specific and non-specific symptoms of PTSD" (Tsai et al., 2011, p. 311). To test this hypothesis, Tsai et al. stated (2011) "a principal components factor analysis was conducted that specified *a priori* two factors" p. 311. Conway and Huffcutt (2003) conducted a review and evaluation of exploratory factor analysis and stated, "EFA can also be applied to existing instruments to assess dimensionality … Another hypothesis-testing example is when EFA is conducted under different conditions to see if the number of factors changes" (p. 149).

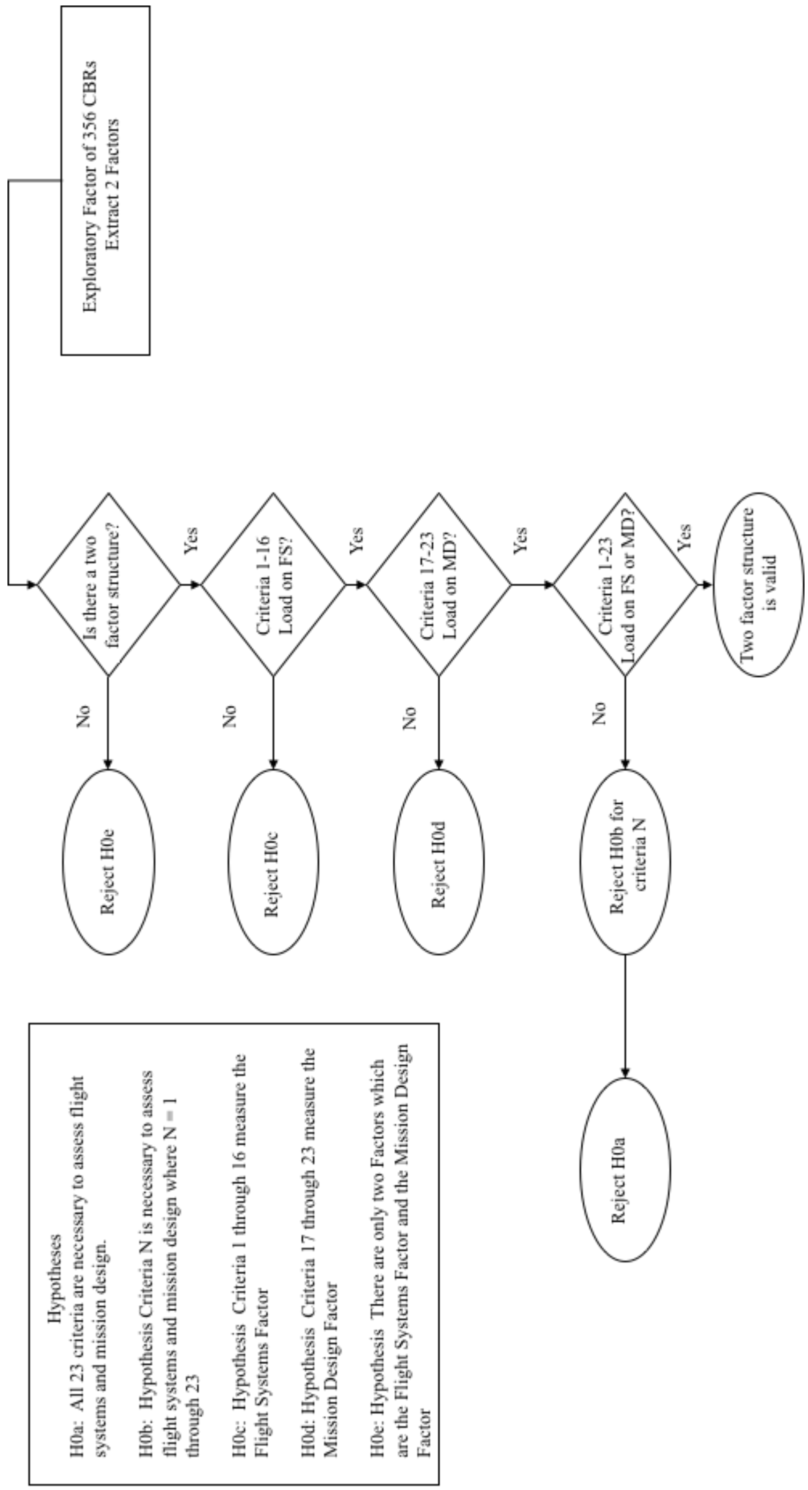
The number of factors to be extracted can be determined in a number of ways. Field (2013) stated that using the Kaiser method to determine the number of factors "appears to be accurate when the number of variables in the analysis is less than 30 and the resulting communalities (after extraction) are all greater than 0.7, or when the sample size exceeds 250 and the average communality is greater than or equal to 0.6" (p. 877). However, according to Costello and Osborne (2005), retaining all factors with eigenvalues greater than 1.0 is "among the least accurate methods for selecting the number of factors to retain" (Velicer & Jackson, 1990).

In addition, Burnett and Dart (1997) stated that the number of factors to be extracted when validating an existing instrument should be based on the existing instrument:.

> It should be noted that the practice of using eigenvalues greater than one or a Scree test to determine the number of factors to be extracted should not be undertaken when validating existing instruments. The number of factors to be extracted should be set to the number of substantive scales which form the instrument rather than making modifications to theory via deleting or reallocating items on the basis of sample specific results. (p. 129)

According to Burnet and Dart (1997), "Items were removed if they did not link with their hypothesized co-item, or if they did not fulfil a .4 minimum loading criteria and a nondual loading criteria" (p. 128). This research approach was used by Barbeite and Weiss (2004) in investigating the validity of a computer self-efficacy scale. Exploratory factor analysis was conducted to determine if the number of factors for each scale measured in the current sample was consistent with those found in previous studies. Consistency in the number of factors would indicate a constant conceptual domain (p. 5). The hypotheses of this research will be tested with exploratory factor analysis with the two-factor structure that is hypothesized. The resulting two factor structure will be examined against the hypotheses. The criteria loadings on each factor will be reviewed to determine if the criteria load on the expected factor. The minimum loading of the criteria on the factor was also examined. A summary of the hypotheses testing process is shown in Figure 4 below.

Figure 4 Hypothesis Testing Process



Data Analysis Techniques

The 23 criteria act as a survey instrument for consistently evaluating each proposed mission concept. For the purposes of this research, the CBR forms are considered to be the survey responses for each proposed mission concept. Techniques used to refine surveys will be used to analyze the criteria ratings in order to investigate whether all 23 criteria should be retained. Fricker, Kulzy, and Appleget (2012) stated, “Factor analysis is a method for identifying latent traits from question-level survey data. It is useful in survey analysis whenever the phenomenon of interest is complex and not directly measurable via a single question” (p. 30). The source of data for this research comprised 356 records of CBRs for space science mission concepts that have been completed. The current 23 criteria were developed by a group of experts on flight systems and mission design. No analysis has been done to substantiate these 23 criteria. Other fields have similar problems with surveys that have been developed and used for an extended time without analysis to validate the survey instrument. An example is a study of an existing School Level Environment Questionnaire (SLEQ) where exploratory factor analysis was used to validate the existing instrument (Johnson & Stevens, 2001). This study started with eight scales with seven items each, for a total of 56 items. After exploratory factor analysis, 13 items and one scale were dropped, which resulted in a total of 43 items and seven scales.

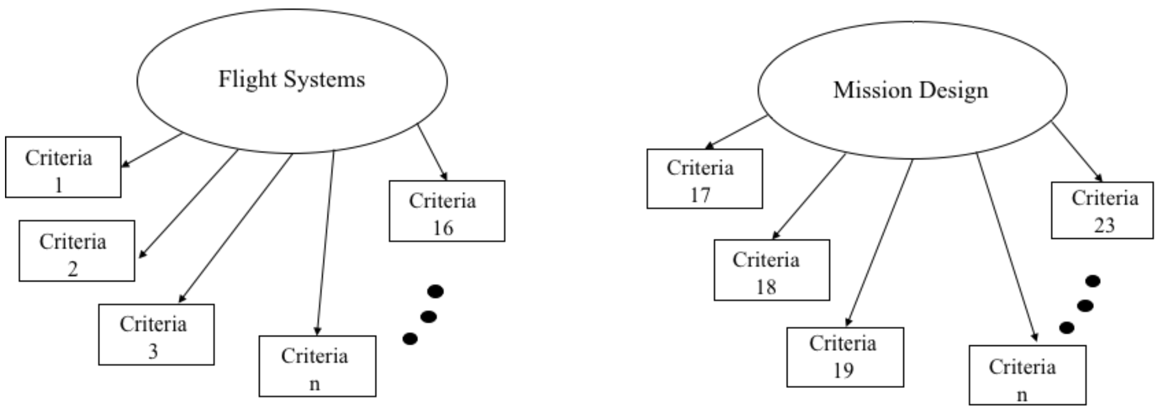
A review of the literature has indicated that exploratory factor analysis can be used to determine if the current 23 criteria used in past mission concept proposal evaluations are all necessary and whether any latent variable exists which could result in a new list of factors for a more efficient evaluation (Kline, 1994). This research will use

exploratory factor analysis to evaluate the loadings of the criteria on the assumed factor structure, based on past records.

There are 16 criteria assumed to measure the flight systems factor, as is documented in Appendix B. There are seven criteria assumed to measure mission design. This implies the factor structure shown below in Figure 5. The 16 criteria on the left load only on the flight systems factor. The seven criteria on the right load only on the mission design factor. These criteria were developed by experts, and they have not been previously analyzed to determine to what extent the 16 criteria measure the flight systems factor or the extent to which the seven criteria measure mission design. According to Byrne (2016), "Exploratory factor analysis (EFA) is designed for the situation where links between the observed and latent variables are unknown or uncertain" (p.6). Nunnally and Bernstein (1994) described factor analysis as being useful at identifying relationships among a set of observed variables and then, through data reduction, reducing the number of variables to a smaller set of variables into factors that have common characteristics. Gorsuch (2015) stated, "Factor analysis allows one to analyze numerous variables at a time, to unravel relationships among variables to factors, and to stress parsimonious solutions" (p. 9).

Exploratory factor analysis was conducted on 356 records since the links between the criteria and the factor, for both flight systems or mission design, is currently uncertain. Hypotheses can be tested based on the results of exploratory factory analysis of a two-factor structure.

Figure 5 Current Factor Structure



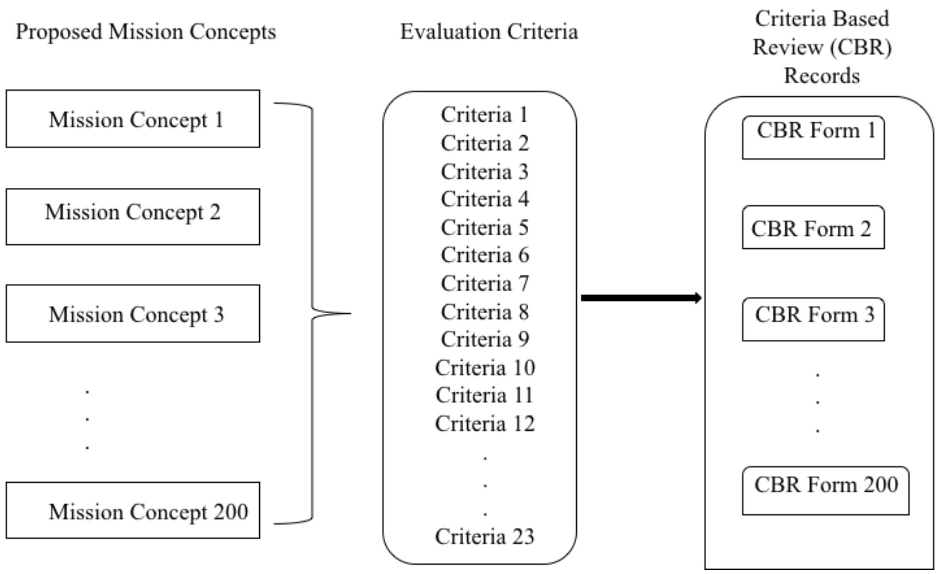
Records for Analysis

There are over 500 mission concepts that were rated by an expert peer review, based on the 23 criteria. The most recent 356 expert peer reviews were used for this research. The NASA 2016 Medium Explorer (MIDEX) Full Missions Evaluation Plan (NASA, 2016a) refers to the form that documents the expert review panel ratings of the criteria as “Form C.” Form C will be referred to as the “Criteria Based Review (CBR) form” in this paper. The result of each expert peer review is a CBR form for each of the mission concepts. The CBR form provides a rating of each criterion and a qualitative description (also called “a finding”) which provides the rationale for each strength or weakness. The qualitative description is a paragraph describing the strength or weakness in detail for each criterion. The adjectival rating is either a Major Strength, a Major Weakness, a Minor Weakness, or a Minor Strength. This is a paradigmatic corroboration as described by Saldana (2016), who discussed a type of mixed data transformation (p. 26). According to Saldana (2016),

> Paradigmatic corroboration occurs when the quantitative results of a data set do not simply harmonize or complement the qualitative but corroborate it. In other words, the quantitative analytic results "jive" with or appear to correspond with the qualitative analytic outcomes … Paradigmatic corroboration provides the analyst a "reality check" of his or her analytic work. It also provides two sets of lenses to examine the data for a multidimensional and more trustworthy account. (p. 26-27)

In examining the CBR material, it is important to note that one group of approximately eight to 10 people have developed the qualitative description of the strength or weakness and also came to a general agreement on the rating to be assigned the grade of Major Strength, Major Weakness, Minor Weakness, or Minor Strength. Although the grade definitions appear to be ordered categories, the group that is determining the grade uses general agreement to reach their decision. Individuals in the group can have some variation in the degree or the strength of their agreement on the rating, but a measure of individual strength of agreement was not documented. The fact that a group reached consensus on the grade based on a qualitative description of the issue means that the rating of the criteria corresponds to the qualitative text. The qualitative text can provide interpretation to the meaning of the quantitative analysis of the exploratory factor analysis on the 23 criteria. This process is shown in Figure 6, below.

Figure 6: Criteria Based Review Process



Criteria Ratings

Each expert review panel assesses each of the 23 criteria for each proposed mission concept and writes a paragraph or a "finding" that describes the mission concept's strength or weakness. Adjectival ratings of a major strength, a minor strength, a major weakness, a minor weakness, or as expected are assigned by the expert peer review panel, by general agreement. The standard definitions of the criteria ratings are posted publicly for each mission concept review and are the same for each mission concept evaluation. For New Frontiers 4 Announcement of Opportunity, Lucas (2017) defined the five ratings as follows:

> Major Strength: A facet of the implementation response that is judged to be well above expectations and can substantially contribute to the ability of the project to meet its technical requirements on schedule and within cost.

Minor Strength: A strength that is worthy of note and can be brought to the attention of proposers during debriefings, but is not a discriminator in the assessment of risk of a step one or pre-Phase A proposal.

Major Weakness: A deficiency or set of deficiencies taken together that are judged to substantially weaken the project's ability to meet its technical objectives on schedule and within cost.

Minor Weakness: A weakness that is sufficiently worrisome to note and can be brought to the attention of proposers during debriefings, but is not a discriminator in the assessment of risk of a step one or pre-Phase A proposal.

*Note: Findings that are considered "as expected" are not documented on the CBR form (p. 19)

Coding of CBR Forms

The purpose of the research is to analyze the CBR records and to determine which criteria will help a decision maker discriminate a proposed mission concept as either significant for rejection (negative) or selection (positive) at the mission concept stage. A discriminatory power scale is used to code the CBR data for analysis.

The discriminatory power scale reflects the importance of the adjectival rating assigned by the expert peer review panel. The importance of the ratings is determined by the definitions of the ratings and by how the information is subsequently used by the expert peer review and decision makers who are reviewing the CBR. After all of the criteria have been rated, the expert peer review panel polls each member on one of three overall risk ratings for the mission concept. The three possible risk ratings are low risk, medium risk or high risk. The definition of each risk rating is standard across all of the expert peer review panels and are stated in publicly posted evaluation plans. As defined

in the Small Explorers (SMEX) and Missions of Opportunity Preproposal Conference Technical, Management, and Cost (TMC) Evaluation (Daniels, 2007), the definitions of low risk, medium risk, and high risk are:

> Low Risk: There are no problems in the proposal that cannot be normally solved within the time and cost proposed. Problems are not of sufficient magnitude to doubt the Proposer's capability to accomplish the investigation.
>
> Medium Risk: Problems have been identified, but are considered within the proposal team's capabilities to correct with good management and application of effective engineering resources. Mission design may be complex and resources tight.
>
> High Risk: Problems are of sufficient magnitude such that failure is highly probable. (p.21)

In addition, the instructions in the NASA 2014 Astrophysics Small Explorer (SMEX) evaluation plan (NASA, 2015a) state that "Only Major findings are considered in the risk rating" (p.49). This means that only the major findings, of either a major weakness or a major strength, have discriminatory power in impacting the risk rating assigned to each proposed mission concept. This is consistent across all of the expert peer review panels. Consequently, major weaknesses and major strengths are more important than minor weaknesses and minor strengths. Also, since the risk ratings definitions focus on problems, weaknesses are more important than strengths. The risk rating definitions move higher (more negative) as more weaknesses are identified. Consequently, criteria described in the qualitative text with a corresponding major weakness will be coded as a five on the discriminatory power scale. Criteria described in

the qualitative text with a corresponding major strength will be coded as a four on the discriminatory power scale.

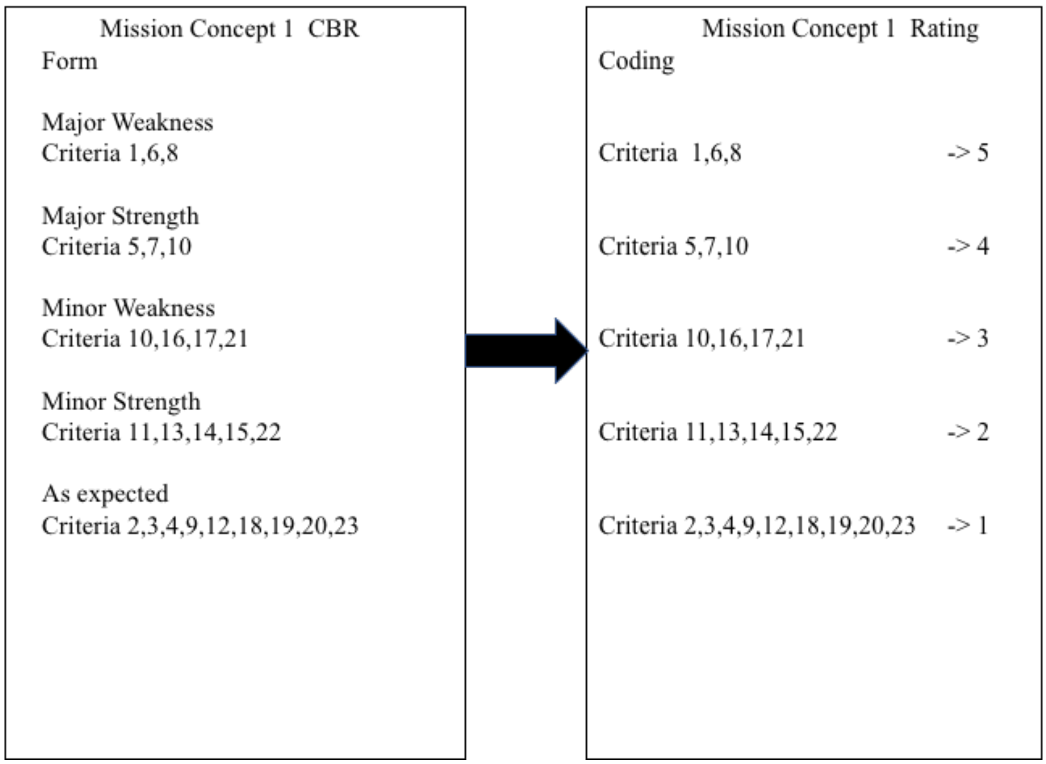
Minor weaknesses are important to proposals selected for a phase A study, since minor weaknesses and strengths are considered in the risk rating for the phase A study. A compete briefing of all of the major and minor strengths and weaknesses is provided to each proposer, to allow them to resolve the weaknesses in the phase A study or in their next proposal to an AO. Since weaknesses are more important for a proposer to resolve in the phase A study or for their next proposal, a criterion identified in the qualitative text of a minor weakness will be coded as a three in the discriminatory power scale. Criteria identified in the qualitative text as a minor strength will be coded as a two. Criteria that are not documented on the CBR form are “As Expected” and will be coded as one. Table 2 below summarizes coding of the CBR data on the Discriminatory Power Scale for analysis.

Table 2: Coding of CBR Data

Discriminatory Power Scale	Discriminatory Power Score
VERY HIGH - Very high discriminatory power (very negative)	5
HIGH - High discriminatory power (very positive)	4
MEDIUM - Medium discriminatory power (negative)	3
LOW - Low discriminatory power (positive)	2
NONE - No discriminatory power (neutral)	1

All criteria will be assigned one coding based on each CBR form. Figure 7 below provides an example of coding one CBR Form.

Figure 7: Example of Coding a CBR Form



Resolving Data Quality Issues

The definitions for criteria 1 through 23 are shown in Appendix B. The criteria are broad in scope and, consequently, there are several topics for a criterion that may be documented in the CBR that may be the focus of the qualitative text that is provided with the rating. Table C.1 lists key words and topics associated with each criterion. This table was used to ensure consistency in determining which criteria each qualitative statement represented. Each CBR record was reviewed twice, using Table C.1 as a check, to ensure that each topic discussed in a qualitative statement was associated consistently with the

same criteria. The topics associated with the criteria would be useful in interpreting the results of the analysis.

A second possible issue in determining the criteria of a qualitative text statement is that several topics may be mentioned in the qualitative text paragraph. Several approaches were used to address this issue. First, a summary bold statement, which usually summarizes the significant issue of the paragraph so the correct criteria can be identified, was provided at the beginning of the qualitative text of a major strength or major weakness. However, if the bold summary statement contained multiple topics, the topic listed first was used to determine the criteria.

The expert peer review panel is trained to identify the most significant topics first. A bold summary statement is not provided at the beginning of the qualitative statement for minor strengths or minor weaknesses. The expert peer review is trained to focus on describing the most important topic first in the qualitative statement. So, for a minor weakness or a minor strength, the first topic described is used to determine the criteria to be coded on the discriminatory power scale.

Another possible problem, due to broad criteria definitions, is the possibility that a positive aspect of a criterion could be documented in the same CBR form as a negative aspect of the same criterion. The coding scheme summarized in Table 3 below was used to resolve conflicting ratings of criteria in the same CBR form. The order of importance of the discriminatory power of the rating was used to resolve any conflict. If there is a major weakness in a CBR form, and any other conflicting qualitative text of the same criterion appears, then the CBR is coded as a five on the discriminatory power scale, since major weaknesses have the most discriminatory power. If a major strength qualitative text conflicts with a statement of lower discriminatory power, then the CBR

form is coded as a four on the discriminatory power scale. Table 3 below summarizes how conflicting ratings within a CBR form are resolved.

Table 3: Resolving Conflicting Ratings

		Criteria N Rating 1			
		MW	mW	mS	MS
Croteria N Rating 2	MW	MW	MW	MW	MW
	mW	MW	mW	mW	MS
	mS	MW	mW	mS	MS
	MS	MW	MS	MS	MS

MW is a Major Weakness
mW is a minor Weakness
mS is a minor Strength
MS is a Major Strength

The space science program CBR records used for this research were from the programs listed in Appendix D. The most recent records were reviewed, and this review resulted in a total of 356 records coded on the discriminatory power scale for analysis. Of the records reviewed, four were dropped because pages were missing from the CBR form.

Sample Size

Several references address the sample size required to conduct exploratory factor analysis, based either on the number of variables or the number of factors. A minimum of five participants per variable for this analysis was recommended by Munro (2005). Since this research has 23 variables, then the minimum required number of CBR records

would be 115. A minimum of five to 20 per factor was recommended by Suhr (2006). Since there are two factors in the assumed factor structure, then the minimum number of CBR records would be from 10 to 40. A total of 356 CBR records were reviewed and were coded on the discriminatory power scale. The entire 356 data points could be used for exploratory factor analysis.

Normality

Fabrigar, Wegener, MacCallum, and Strahan (1999) recommend reviewing the data for normality and "If nonnormality is severe (e.g., skew > 2; kurtosis >7) … one might wish to use a principal factors procedure" (p.283). Fabrigar et al. (1999) recommend using principal component analysis, since it has the advantage of no distributional assumptions. A review of the Kolmogorov-Smirnov and the Shapiro-Wilk tests of normality is shown in Table F.35. Both tests indicate that all of the criteria fail tests of normality. Steinskog, Tjøstheim, and Kvamstø, (2007) criticized the Kolmogorov-Smirnov test, stating that "the test usually leads to systematic drastic errors" (p. 1). However, Steinskog et al. (2007) recommended the Shapiro-Wilk test as a good alternative. A review of the data histograms and the Q-Q plots (Figures F.1 through F.33) indicate that some data are normally distributed and most are not normally distributed. The highlighted cells in the descriptive statistics Table F.34 indicate criteria where nonnormality is severe (e.g., skew > 2; kurtosis >7), as described by Fabrigar et al. (1999). In regard to multivariate normality, Hair et al. (2010) stated

> As data deviate more from the assumption of multivariate normality, then the ratio of respondents to parameters needs to increase. A generally accepted ratio to minimize problems with deviations from normality is 15 respondents for each parameter estimated in the model. Although some estimation procedures are

> specifically designed to deal with nonnormally distributed data, the researcher is always encouraged to provide sufficient sample size to allow for the sampling error's impact to be minimized, especially for nonnormal data. (p. 643)

In addition, Hair et al. (2010) stated that

> larger sample sizes r*educe* the detrimental effects of nonnormality. In small samples of 50 or fewer observations and especially if the sample size is less than 30 or so, significant departures from normality can have a substantial impact on the results. For sample sizes of 200 or more, however, these same effects may be negligible. (p. 72)

For this research, there were 23 criteria. The sample size of 356 exceeded the generally accepted ratio to minimize problems due to nonnormality, since 15 records for each 23 criteria would be a total of 345. The 356 samples were also above the recommended number of 200 samples, so the detrimental effects of nonnormality should be negligible. Field (2013) stated "the best advice is that if your sample is large then don't worry too much about normality at all" (p.184). According to Hair et al. (2010),

> From a statistical standpoint, departures from normality, homoscedasticity, and linearity apply only to the extent that they diminish the observed correlations. Only normality is necessary if a statistical test is applied to the significance of the factors, but these tests are rarely used. In fact, some degree of multicollinearity is desirable, because the objective is to identify interrelated sets of variables. (p. 103)

Tabachnick and Fidell (2013) stated that

> As long as PCA and FA are used descriptively as convenient ways to summarize the relationships in a large set of observed variables, assumptions regarding the

> distributions of variables are not in force. If variables are normally distributed, the solution is enhanced. To the extent that normality fails, the solution is degraded but may still be worth worthwhile. However, multivariate normality is assumed when statistical inference is used to determine the number of factors. Multivariate normality is the assumption that all variables, and all linear combinations of variables, are normally distributed. Although tests of multivariate normality are overly sensitive, normality among single variables is assessed by skewness and kurtosis. If a variable has substantial skewness and kurtosis, variable transformation is considered. (p. 618)

More than half of the criteria distributions were nonnormally distributed and had substantial skewness and kurtosis. Due to the broad nature of the criteria, transforming the variables would make interpretation challenging. This research followed the recommendation of Fabrigar et al. (1999) and used principal component analysis (PCA) since it has no distributional assumptions. See Appendix F for details on criteria distributions.

Principal Component Analysis

This research used PCA for extraction, since the data set had a large percent of nonnormally distributed data. It is important to understand that PCA is different than factor analysis. According to Henson and Roberts (2006), "PCA is intended to simply summarize many variables into fewer components, and the latent constructs (i.e., factors) are not the focus of the analysis. In addition, Tabachnick and Fidell (2013) stated

> FA produces *factors*, while PCA produces *components* … the difference between PCA and FA is in the variance that is analyzed. In PCA, all the variances in the observed variables are analyzed. In FA, only shared variance is analyzed;

> attempts are made to estimate and eliminate variance due to error and variance that is unique to each variable… Components are simply aggregates of correlated variables. In that sense, the variables "cause" – or produce – the component. There is no underlying theory about which variables should be associated with which factors; they are simply empirically associated. It is understood that any labels applied to derived components are merely convenient descriptions of the combination of variables associated with them, and do not necessarily reflect some underlying process. (p. 614-615)

Validity and Reliability

Reliability and validity will be addressed for the exploratory factor analysis. Tavakol and Dennick (2011) noted that "Validity is concerned with the extent to which an instrument measures what it is intended to measure" (p.1). Content or face validity, according to Hair et al. (2010), "is an assessment of the correspondence of the variables to be included in a summated scale and its conceptual definition" (p. 125). Hair et al. (2010) described using experts as one way to accomplish this subjective assessment of the correspondence of the items or the measured variables to the concept (p. 125). Face validity, on the current measurement instrument under study, was addressed when it was developed by experts. Tavakol, Mohagheghi, and Dennick (2008) stated, "Reliability is concerned with the ability of an instrument to measure consistently" (p.1). According to Hair et al. (2010),

> A second and more commonly used measure of reliability is internal consistency, which applies to the consistency among the variables in a summated scale. The rational for internal consistency is that the individual items or indicators of the

> scale should all be measuring the same construct and thus be highly intercorrelated. (p. 125)

Hair et al. (2010) described two other measures to assess internal consistency: "the item-to-total correlation (the correlation of the item to the summated scale score) and the inter-item correlation (the correlation among items)" (p.125). MacCallum, Roznowski, Mar, and Reith (1994) recommended that the item-to-total correlations exceed .50 and that the inter-item correlations exceed .30. For this research, Cronbach's alpha, the item-to-total correlation, and the inter-item correlations were assessed for factors with multiple criteria.

According to Pett, Lackey, and Sullivan (2003), the internal consistency aspects of reliability of the structure can be addressed in exploratory factor analysis stage in several ways. Individual items that have low correlations (< .3) with other items should be dropped (Pett et al., 2003). Following the recommendations of Hair et al. (2010), items that do not have a significant loading (< .3) on a factor should be deleted after considering whether the communality of the item indicates that the item is of little interest (Hair, Anderson, Tatham, & Black, 1995). The consistency of a scale can be addressed by Cronbach's coefficient alpha (α) (Cronbach, 1951), which measures the portion of total variance in a scale attributed to a common source (DeVellis, 1991). According to Cohen (1977), values above .70 are considered acceptable, values above .80 are good, and values above .90 are excellent. Generally, the lower limit for this reliability measure is $\leq$ .7. However, Hair et al. (2010) noted that it could be as low as .60 for exploratory factor analysis (p. 125). In regard to Cronbach's alpha or coefficient alpha, Tavakol and Dennick (2011) stated:

> If a test has more than one concept or construct, it may not make sense to report alpha for the test as a whole as the larger number of questions will inevitable inflate the value of alpha. In principle therefore, alpha should be calculated for each of the concepts rather than for the entire test or scale. (p. 54)

Field (2013) agreed and stated that “if your questionnaire has subscales, α should be applied separately to these subscales” (p. 709). The consistency of the factor solutions in this research were addressed for factors with multiple criteria. Cronbach’s coefficient alpha (α) (Cronbach, 1951) was assessed for each factor with multiple criteria. Tavakol and Dennick (2011) stated that “the reliability of an instrument is closely associated with its validity. An instrument cannot be valid unless it is reliable. However, the reliability of an instrument does not depend on its validity” (p.53).

CHAPTER 4

RESULTS AND DISCUSSION

Analysis of Current Two Factor Structure

Analysis was conducted to test the current two factor structure against the hypotheses. To test the hypothesis that there are only two factors for the current 23 criteria, two factors were specified for extraction and the results were reviewed to test the hypothesis. Table 4 below shows the measures of sampling adequacy (MSA) and Bartlett's test of sphericity for the two-factor solution.

Table 4 Full Set Measure of Sampling Adequacy

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.652
Bartlett's Test of Sphericity	Approx. Chi-Square	620.302
	df	253
	Sig.	0.000

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) was 0.652, which is greater than the minimum required of .5. Kaiser (1974) provided an interpretation of KMO statistics that is shown in Table 5 below.

Table 5: Interpretation of the KMO Statistics

KMO Statistic	Interpretation
in the .90's	marvelous
in the .80's	meritorious
in the .70's	middling
in the .60's	mediocre
in the .50's	miserable
below .5	unacceptable

Bartlett's test of sphericity shown in Table 4 was less than 0.05, which indicated that there were relationships among the criterion (Pett et al, (2003). However, Pedhazur and Schmelkin (1991) cautioned that

> Bartlett's sphericity test is affected by sample size. When N is large, as it should be in factor analytic studies … the null hypothesis will almost always be rejected. … rejection of the null hypothesis should not be construed as evidence that the correlation matrix is appropriate for FA. (p. 599-600)

After these preliminary checks, Hutcheson and Sofroniou (1999) recommended reviewing the individual variables for MSA; these are shown in the diagonal of the anti-image correlation matrix in Table 6. Following Kaiser's (1974) recommendations, values below .5 were unacceptable. Criterion 8 is .487, criterion 10 is .457, and criterion 12 is .486. Consequently, criteria 8, 10, and 12 all failed the test of individual sampling adequacy.

Table 6 Individual MSA and Partial Correlations

Anti-image Correlation

	Criteria_1	Criteria_2	Criteria_3	Criteria_4	Criteria_5	Criteria_6	Criteria_7	Criteria_8	Criteria_9	Criteria_10	Criteria_11	Criteria_12	Criteria_13	Criteria_14	Criteria_15	Criteria_16	Criteria_17	Criteria_18	Criteria_19	Criteria_20	Criteria_21	Criteria_22	Criteria_23
Criteria_1	.626[a]	0.068	0.021	0.032	-0.044	-0.024	-0.093	-0.036	-0.078	0.068	-0.011	0.087	-0.032	-0.05	-0.132	0.021	-0.079	0.057	-0.024	0.047	-0.066	-0.112	0.078
Criteria_2	0.068	.562[a]	-0.053	-0.076	-0.047	-0.147	-0.012	-0.069	0.034	0.071	0.038	0.163	0.063	0.045	-0.064	-0.059	-0.053	-0.09	-0.093	-0.12	-0.018	-0.045	0.017
Criteria_3	0.021	-0.053	.655[a]	-0.062	-0.01	0.027	-0.056	0.103	0.053	-0.097	-0.083	-0.012	-0.063	-0.077	-0.055	0.028	0.02	-0.012	-0.015	-0.023	-0.023	0.06	0.022
Criteria_4	0.032	-0.076	-0.062	.722[a]	-0.066	0.095	-0.099	-0.064	-0.074	0.031	-0.001	-0.083	-0.04	-0.005	-0.098	0.074	-0.033	-0.065	0.016	-0.14	0.044	-0.116	-0.051
Criteria_5	-0.044	-0.047	-0.01	-0.066	.732[a]	-0.019	-0.176	-0.007	-0.043	0.012	0.029	0.07	-0.065	-0.054	0.071	0.179	-0.033	0.03	0.077	-0.098	-0.099	-0.101	0.029
Criteria_6	-0.024	-0.147	0.027	0.095	-0.019	.658[a]	-0.111	0.067	0.013	-0.019	0.011	-0.051	-0.114	-0.135	-0.071	0.085	0.103	-0.004	0.031	-0.005	-0.142	-0.129	-0.133
Criteria_7	-0.093	-0.012	-0.056	-0.099	-0.176	-0.111	.765[a]	0.065	0.019	-0.091	-0.015	0.053	-0.159	-0.087	-0.033	0.043	-0.086	-0.063	0.031	-0.09	0.042	0.009	-0.036
Criteria_8	-0.036	-0.069	0.103	-0.064	-0.007	0.067	0.065	.487[a]	0.017	0.014	-0.011	-0.053	0.032	-0.091	-0.031	-0.039	0.053	0.016	0.022	0.021	0.011	0.006	-0.038
Criteria_10	-0.078	0.034	0.053	-0.074	-0.043	0.013	0.019	0.017	.528[a]	0.017	0.037	0.036	0.012	0.099	-0.143	-0.247	-0.047	-0.015	-0.084	-0.063	-0.013	-0.123	-0.075
Criteria_11	0.068	0.071	-0.097	0.031	0.012	-0.019	-0.091	0.014	0.017	.457[a]	0.058	0.007	0.037	0.091	-0.042	0.039	-0.008	0.069	-0.045	-0.039	-0.005	-0.041	0.014
Criteria_12	-0.011	0.038	-0.083	-0.001	0.029	0.011	-0.015	-0.011	0.037	0.058	.690[a]	0.021	-0.013	-0.156	-0.03	0.062	-0.002	-0.009	-0.069	-0.052	-0.042	-0.185	-0.07
Criteria_13	0.087	0.163	-0.012	-0.083	0.07	-0.051	0.053	-0.053	0.036	0.007	0.021	.486[a]	0.028	-0.013	0.005	0.05	-0.053	-0.146	0.101	0.011	-0.25	0.01	-0.014
Criteria_14	-0.032	0.063	-0.063	-0.04	-0.065	-0.114	-0.159	0.032	0.012	0.037	-0.013	0.028	.721[a]	0.013	0.035	-0.078	-0.008	-0.08	0.091	-0.079	-0.108	-0.038	-0.012
Criteria_16	-0.05	0.045	-0.077	-0.005	-0.054	-0.135	-0.087	-0.091	0.099	0.091	-0.156	-0.013	0.013	.709[a]	-0.109	0.003	0.068	-0.052	-0.008	-0.028	-0.092	-0.106	0.026
Criteria_17	-0.132	-0.064	-0.055	-0.098	0.071	-0.071	-0.033	-0.031	-0.143	-0.042	-0.03	0.005	0.035	-0.109	.613[a]	0.01	0.038	-0.081	0.015	-0.064	0.036	0.04	0.102
Criteria_19	0.021	-0.059	0.028	0.074	0.179	0.085	0.043	-0.039	-0.247	0.039	0.062	0.05	-0.078	0.003	0.01	.572[a]	0.099	-0.036	0.048	0.046	-0.027	-0.006	-0.044
Criteria_20	-0.079	-0.053	0.02	-0.033	-0.033	0.103	-0.086	0.053	-0.047	-0.008	-0.002	-0.053	-0.008	0.068	0.038	0.099	.601[a]	-0.048	0.049	0.001	-0.081	-0.09	-0.019
Criteria_21	0.057	-0.09	-0.012	-0.065	0.03	-0.004	-0.063	0.016	-0.015	0.069	-0.009	-0.146	-0.08	-0.052	-0.081	-0.036	-0.048	.646[a]	-0.05	0.039	0.073	-0.127	-0.057
Criteria_22	-0.024	-0.093	-0.015	0.016	0.077	0.031	0.031	0.022	-0.084	-0.045	-0.069	0.101	0.091	-0.008	0.015	0.048	0.049	-0.05	.528[a]	-0.027	-0.011	0.001	-0.005
Criteria_23	0.047	-0.12	-0.023	-0.14	-0.098	-0.005	-0.09	0.021	-0.063	-0.039	-0.052	0.011	-0.079	-0.028	-0.064	0.046	0.001	0.039	-0.027	.771[a]	0.039	-0.046	0.007
Criteria_15	-0.066	-0.018	-0.023	0.044	-0.099	-0.142	0.042	0.011	-0.013	-0.005	-0.042	-0.25	-0.108	-0.092	0.036	-0.027	-0.081	0.073	-0.011	0.039	.531[a]	0.085	0.023
Criteria_9	-0.112	-0.045	0.06	-0.116	-0.101	-0.129	0.009	0.006	-0.123	-0.041	-0.185	0.01	-0.038	-0.106	0.04	-0.006	-0.09	-0.127	0.001	-0.046	0.085	.715[a]	-0.048
Criteria_18	0.078	0.017	0.022	-0.051	0.029	-0.133	-0.036	-0.038	-0.075	0.014	-0.07	-0.014	-0.012	0.026	0.102	-0.044	-0.019	-0.057	-0.005	0.007	0.023	-0.048	.554[a]

According to Hair et al. (2010),

> The correlations among the variables can also be analyzed by computing the partial correlations among variables. A partial correlation is the correlation that is unexplained when the effects of other variables are taken into account. If "true" factors exist in the data, the partial correlation should be small, because the variables can be explained by the variables loading on the factors. If the partial correlations are high, indicating no underlaying factors, then factor analysis is inappropriate….and a rule of thumb would be to consider partial correlations above .7 as high. (103-104)

The off diagonal values in the anti-image correlation matrix shown in Table 6 are the negatives of the partial correlations. There are no values in the off diagonal in Table 6 that exceed .7. This check of partial correlations passed this check of correlations among the variables.

The correlation matrix was also inspected for factorability. Hair et al. (2010) recommended a substantial number of correlations over .30. The correlation in this data set did not meet that standard. However, the null hypothesis, that states that there are no correlations in the data, was rejected on overage 7.7 times for each criterion, based on the one tailed test of significance. The full data set of 356 records was considered appropriate for factor analysis, based on the review of partial correlations and on the test of significance of the correlation matrix.

According to Hair et al. (2010),

> the communality of each criteria represents the amount of variance accounted for by the factor solution for each variable… a researcher may specify that at least one-half of the variance of each variable must be taken into account. Using this

guideline, the researcher would identify all variables with communalities less than .5 as not having sufficient explanation. (p. 119)

Using .5 as a guideline for communalities, a review of the communalities shown in Table 7 for the two factor solution revealed that none of the criteria have communalities over .5. Consequently, the two factor structure did not represent even half the variance of each variable.

Table 7 Communalities

	Initial	Extraction
Criteria_1	1	0.089
Criteria_2	1	0.204
Criteria_3	1	0.084
Criteria_4	1	0.249
Criteria_5	1	0.286
Criteria_6	1	0.246
Criteria_7	1	0.356
Criteria_8	1	0.024
Criteria_10	1	0.311
Criteria_11	1	0.01
Criteria_12	1	0.141
Criteria_13	1	0.194
Criteria_14	1	0.216
Criteria_16	1	0.235
Criteria_17	1	0.173
Criteria_19	1	0.228
Criteria_20	1	0.061
Criteria_21	1	0.134
Criteria_22	1	0.125
Criteria_23	1	0.241
Criteria_15	1	0.268
Criteria_9	1	0.345
Criteria_18	1	0.031

Extraction Method: Principal Component Analysis.

Hair et al. (2010) provided guidance on expectations for the total variance explained in various applications as follows:

> No absolute threshold has been adopted for all applications. However, in the natural sciences the factoring procedure usually should not be stopped until the extracted factors account for at least 95 percent of the variance or until the last factor accounts for only a small portion (less than 5%). In contrast, in the social sciences, where information is often less precise, it is not uncommon to consider a solution that accounts for 60 percent of the total variance (and in some instances even less) as satisfactory. (p. 109)

This research used 60 percent as the guideline for an acceptable measure of total variance explained. The total variance extracted was approximately 18% for the two-factor solution; this was unacceptably low. Table 8 shows that the total variance explained by the two factors was only 18.5%.

Table 8 Total Variance Explained

Initial Eigenvalues			
Component	Total	% of Variance	Cumulative %
1	2.647	11.509	11.509
2	1.604	6.975	18.484
3	1.417	6.159	24.643
4	1.286	5.591	30.234
5	1.231	5.353	35.587
6	1.166	5.07	40.657
7	1.143	4.969	45.626
8	1.095	4.761	50.386
9	1.03	4.478	54.864
10	1.004	4.365	59.229
11	0.949	4.126	63.355
12	0.879	3.821	67.176
13	0.848	3.688	70.863
14	0.836	3.635	74.498
15	0.786	3.418	77.916
16	0.742	3.226	81.141
17	0.701	3.047	84.188
18	0.691	3.004	87.192
19	0.665	2.891	90.083
20	0.636	2.765	92.848
21	0.574	2.497	95.345
22	0.542	2.355	97.699
23	0.529	2.301	100

Extraction Method: Principal Component Analysis.

Based on this analysis, the current two factor solution with the existing set of 23 criteria did not pass the tests on MSA of individual variables. Communalities were too low, and the two-factor solution did not explain a substantial amount of the variance in the 23 criteria. Consequently, hypothesis H0e, which states there are only two factors which are the flight systems factor and the mission design factor hypothesis, was rejected.

Hair et al. (1995) recommended reviewing the rotated factor matrix and deleting items with weak loadings of less than $|.3|$ from the factor solution. Criteria loadings on

the two factors are shown in Table 9, where weak loadings below the absolute value of .3 were suppressed.

Table 9

Two Factor Rotated Component Matrix

Coefficients Below Absolute Value of .3 Suppressed

Rotated Component Matrix[a]		
	Component	
	1	2
Criteria_7	0.587	
Criteria_9	0.547	
Criteria_5	0.502	
Criteria_23	0.459	
Criteria_16	0.458	
Criteria_4	0.455	
Criteria_6	0.45	
Criteria_14	0.42	
Criteria_12	0.375	
Criteria_21	0.322	
Criteria_17	0.306	
Criteria_1		
Criteria_3		
Criteria_20		
Criteria_18		
Criteria_10		0.543
Criteria_15		-0.491
Criteria_13		-0.441
Criteria_19		0.41
Criteria_2		0.354
Criteria_22		0.353
Criteria_8		
Criteria_11		

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

The current two factor solution assumes that criteria 1 through 16 will address the flight systems factor and load on the same factor. A review of Table 9 shows that criteria 1 through 16 did not load on the same factor. Criteria 4, 5, 6, 7, 9, 12, 14, and 16 loaded on the first factor, but criteria 2, 10, 13, and 15 loaded on the second factor. The remaining factors had weak loadings, and loadings below the absolute value of .3 were

suppressed. Criteria loadings are shown in Table 10, with loadings below the absolute value of .1 suppressed.

Table 10

Two Factor Rotated Component Matrix

Coefficients Below Absolute Value of .1 Suppressed

Rotated Component Matrix[a]		
	Component	
	1	2
Criteria_7	0.587	-0.109
Criteria_9	0.547	0.213
Criteria_5	0.502	-0.185
Criteria_23	0.459	0.173
Criteria_16	0.458	-0.16
Criteria_4	0.455	0.204
Criteria_6	0.45	-0.207
Criteria_14	0.42	-0.197
Criteria_12	0.375	
Criteria_21	0.322	0.174
Criteria_17	0.306	0.282
Criteria_1	0.292	
Criteria_3	0.243	-0.158
Criteria_20	0.234	
Criteria_18	0.143	0.102
Criteria_10	0.127	0.543
Criteria_15	0.166	-0.491
Criteria_13		-0.441
Criteria_19	-0.244	0.41
Criteria_2	0.281	0.354
Criteria_22		0.353
Criteria_8		0.15
Criteria_11		

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

Criteria 13 still showed no loading on factor 1 at this very weak level of loading. Hypothesis H0c, which states that criteria 1 through 16 measure the flight systems factor, was rejected, since some of criteria 1 through 16 (specifically, criterion 12 and criterion 13) did not load on the same factor.

The current two factor solution assumes that criteria 17 through 23 will address the mission design factor and will load on the same factor. In Table 9, criteria 17, 21, and 23 loaded on the first factor, while 19 and 22 loaded on the second factor. Some criteria were not shown as loading on either factor, due to weak loading levels (below .3). As shown in Table 10, with loadings below .1 suppressed, criterion 20 loaded only on the first factor; criterion 22 loaded on the second factor. The matrix for the two-factor solution in Tables 9 and 10 demonstrates that criteria 20 and 22 did not load on the same factor. Tables 9 and 10 demonstrate that criteria 17 through criteria 23 did not load on the same factor. Consequently, hypothesis H0d, which states that criteria 17 through 23 measure the mission design factor, was rejected, since these criteria did not all load on the same factor. Tables 9 and 10 show the use of the varimax rotation method. Other rotation methods, including direct oblimin, quartimax, equamax, and promax, were also used to extract two factors and the results were reviewed. These other rotation methods all had the same result of the 23 criteria not loading on the expected factor.

The reliability of the criteria composing the extracted two factors was reviewed by assessing three measures of internal consistency. Criteria 1 through 16 were assessed using these three measures to determine their reliability in measuring flight systems. Hair et al. (2010) recommended that Cronbach's alpha as low as .6 could be acceptable for factor analysis. Cronbach's alpha for the flight systems criteria shown in Table 11 was 0.524, which was too low.

Table 11 Cronbach's Alpha for Flight Systems

Reliability Statistics Criteria 1 - 16		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.524	0.499	16

Table 12 shows Cronbach's alpha if item deleted for the flight systems criteria 1 through 16. A review of criteria 1 through 16 shows that if any one of them were deleted, Cronbach's alpha would still be under .6 and would be unacceptably low. The item-to-total correlations are also shown in Table 12, and were below the minimum of .5 recommended by Hair et al. (2010).

Table 12 Cronbach's Alpha if Item Deleted for the Flight Systems Factor

Item-Total Statistics - Flight Systems					
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Criteria_1	26.33	31.675	0.161	0.076	0.514
Criteria_2	27.09	32.107	0.113	0.085	0.527
Criteria_3	26.68	31.88	0.129	0.07	0.523
Criteria_4	26.85	30.542	0.265	0.118	0.488
Criteria_5	27.25	29.758	0.303	0.139	0.478
Criteria_6	28.08	31.349	0.28	0.158	0.488
Criteria_7	27.26	28.963	0.372	0.201	0.46
Criteria_8	28.47	35.54	-0.026	0.041	0.532
Criteria_9	28.11	29.904	0.352	0.188	0.469
Criteria_10	28.06	33.608	0.047	0.056	0.538
Criteria_11	28.47	35.484	-0.028	0.039	0.534
Criteria_12	28.38	33.126	0.206	0.107	0.505
Criteria_13	28.34	35.408	-0.045	0.114	0.543
Criteria_14	27.75	30.722	0.249	0.115	0.492
Criteria_15	28.58	35.416	0.156	0.131	0.523
Criteria_16	28.12	31.37	0.274	0.16	0.489

The inter-item correlations are shown in Table 13. No criteria met the minimum of .30 recommended by Hair et al. (2010).

Table 13 Inter-Item Correlation Matrix – Flight Systems

Inter-Item Correlation Matrix																
	Criteria 1	Criteria 2	Criteria 3	Criteria 4	Criteria 5	Criteria 6	Criteria 7	Criteria 8	Criteria 9	Criteria 10	Criteria 11	Criteria 12	Criteria 13	Criteria 14	Criteria 15	Criteria 16
Criteria_1	1	-0.021	-0.006	0.024	0.115	0.079	0.144	0.02	0.158	0.104	-0.056	0.069	-0.079	0.081	0.072	0.118
Criteria_2	-0.021	1	0.059	0.127	0.089	0.143	0.083	0.047	0.109	0.031	-0.052	0.013	-0.15	0	-0.02	0.026
Criteria_3	-0.006	0.059	1	0.093	0.064	0.039	0.128	-0.101	-0.002	-0.065	0.095	0.106	0.011	0.095	0.043	0.113
Criteria_4	0.024	0.127	0.093	1	0.161	0.008	0.192	0.052	0.202	0.103	-0.02	0.079	0.056	0.104	-0.032	0.082
Criteria_5	0.115	0.089	0.064	0.161	1	0.135	0.289	-0.023	0.188	0.018	0.004	0.06	-0.046	0.167	0.11	0.138
Criteria_6	0.079	0.143	0.039	0.008	0.135	1	0.205	-0.057	0.193	-0.013	0.005	0.089	0.058	0.188	0.18	0.223
Criteria_7	0.144	0.083	0.128	0.192	0.289	0.205	1	-0.081	0.156	0.006	0.076	0.096	-0.048	0.252	0.028	0.175
Criteria_8	0.02	0.047	-0.101	0.052	-0.023	-0.057	-0.081	1	0.002	0.001	-0.048	0.01	0.043	-0.058	-0.013	0.071
Criteria_9	0.158	0.109	-0.002	0.202	0.188	0.193	0.156	0.002	1	0.152	-0.007	0.237	-0.019	0.127	-0.022	0.191
Criteria_10	0.104	0.031	-0.065	0.103	0.018	-0.013	0.006	0.001	0.152	1	-0.024	-0.018	-0.062	0.016	-0.024	-0.07
Criteria_11	-0.056	-0.052	0.095	-0.02	0.004	0.005	0.076	-0.048	-0.007	-0.024	1	-0.054	-0.013	-0.025	-0.011	-0.082
Criteria_12	0.069	0.013	0.106	0.079	0.06	0.089	0.096	0.01	0.237	-0.018	-0.054	1	-0.002	0.068	0.053	0.233
Criteria_13	-0.079	-0.15	0.011	0.056	-0.046	0.058	-0.048	0.043	-0.019	-0.062	-0.013	-0.002	1	0.016	0.248	0.046
Criteria_14	0.081	0	0.095	0.104	0.167	0.188	0.252	-0.058	0.127	0.016	-0.025	0.068	0.016	1	0.137	0.093
Criteria_15	0.072	-0.02	0.043	-0.032	0.11	0.18	0.028	-0.013	-0.022	-0.024	-0.011	0.053	0.248	0.137	1	0.13
Criteria_16	0.118	0.026	0.113	0.082	0.138	0.223	0.175	0.071	0.191	-0.07	-0.082	0.233	0.046	0.093	0.13	1

Criteria 1 through 16 failed three internal consistency measures of reliability. This demonstrated that the CBR measurement instrument, which contains criteria 1 through 16, did not measure consistently (Tavakol & Dennick, 2011). In regard to validity, Tavakol and Dennick (2011) stated that:

> Validity is concerned with the extent to which an instrument measures what it is intended to measure. Reliability is concerned with the ability of an instrument to measure consistently. It should be noted that the reliability of an instrument is closely associated with its validity. An instrument cannot be valid unless it is reliable. (p. 53)

Since the flight systems factor failed three measures of reliability, then the flight systems instrument was not reliable. Since the flight systems factor was not a reliable measure, then the flight systems factor was not valid.

Hypothesis H0c, which states that criteria 1 through 16 measure the flight systems factor, was rejected, since criterion 1 through 16 failed three internal consistency measures of reliability; the measurement of those criteria was not reliable or valid for measuring flight systems.

The criteria 17 through 23 were assessed, using three internal consistency measures of reliability for the mission design factor. Cronbach's alpha for the mission design criteria, shown in Table 14, was 0.073, and did not meet the minimum of .6.

Table 14 Cronbach's Alpha for Mission Design

Reliability Statistics Criteria 17-23		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.073	0.104	7

Table 15 shows Cronbach's alpha if item deleted for criteria 17 through 23. A review of the seven criteria shows that if any one of them were deleted, Cronbach's alpha would still be under .6, and would be unacceptably low.

Table 15 Cronbach's Alpha if Item Deleted for the Mission Design Factor

	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Criteria_17	12.08	5.924	0.07	0.034	0.01
Criteria_18	13.61	8.137	0.031	0.019	0.069
Criteria_19	13.13	7.761	-0.073	0.022	0.141
Criteria_20	11.88	6.313	-0.025	0.023	0.134
Criteria_21	11.11	6.693	0.146	0.03	-.034[a]
Criteria_22	12.76	7.154	0.001	0.005	0.086
Criteria_23	13.32	7.277	0.085	0.026	0.025

a. The value is negative due to a negative average covariance among items. This violates reliability model assumptions. You may want to check item codings.

The item-to-total correlations are shown in Table 15. They did not meet the minimum of .5. In addition, the negative value of Cronbach's alpha if item deleted for criteria 21 violated reliability model assumptions. The inter-item correlations are shown in Table 16. No criteria met the minimum of .30.

Table 16 Inter-Item Correlation Matrix – Mission Design

Inter-Item Correlation Matrix							
	Criteria 17	Criteria 18	Criteria 19	Criteria 20	Criteria 21	Criteria 22	Criteria 23
Criteria_17	1	-0.073	-0.003	-0.017	0.109	0.02	0.121
Criteria_18	-0.073	1	0.048	0.024	0.093	0.003	0.019
Criteria_19	-0.003	0.048	1	-0.108	0.022	-0.005	-0.088
Criteria_20	-0.017	0.024	-0.108	1	0.074	-0.049	0.048
Criteria_21	0.109	0.093	0.022	0.074	1	0.031	0.042
Criteria_22	0.02	0.003	-0.005	-0.049	0.031	1	0.029
Criteria_23	0.121	0.019	-0.088	0.048	0.042	0.029	1

The mission design criteria 17 through 23 failed three internal consistency measures of reliability. This demonstrated that the mission design factors which contained criteria on17 through 23 did not measure consistently and were not reliable (Tavakol & Dennick, 2011). Since an instrument cannot be valid unless it is reliable (Tavakol & Dennick, 2011), the measurement instrument for the mission design, which were criteria 17 through 23, were not reliable and not valid. Hypothesis H0d, which states that criteria 17 through 23 measure the mission design factor, was rejected, since criteria 17 through 23 failed three internal consistency measures of reliability, and the measurement of those criteria was not valid for measuring mission design.

Hair et al. (2010) provided guidelines for identifying significant factor loadings based on sample size. Based on their guidelines, for a sample size of 350, a factor loading of .30 is significant. The sample size of 356 was used to generate the two-factor solution shown in Table 9. Based on the sample size of 356, it was appropriate to remove criteria that have loadings of less than the absolute value of .3, since they would not be significant. Table 9 shows that criteria 1, 3, 8, 11, 18, and 20 did not have loadings significant enough to load on either of the factors in the two-factor solution. Based on the two factor solution, these criteria were not necessary to assess flight systems and mission design. Since criteria 1, 3, 8, 11, 18 and 20 did not load on either factor,

hypothesis H0b, which states that criteria N is necessary to assess flight systems and mission design where N-1 through 23, was rejected for criteria 1, 3, 8, 11, 18, and 20. In addition, hypothesis H0a, which states that all 23 criteria are necessary to assess flight systems and mission design, was rejected, since criteria 1, 3, 8, 11, 18, and 20 were not necessary since they did not load on either of the two factors. A summary of the MSA and internal consistency measures for the current two factor solution is shown in Appendix E Table E.1.

High or Very High Discriminatory Power (HVHDP) Criteria

Since the current two factor solution was rejected, exploratory factor analysis was used to investigate alternative factor structures. Exploratory factor analysis is useful early in the process of defining a measurment instrument. According to Fabrigar and Wegener (2012),

> We use factor analysis when we want to know how many constructs a set of measured variables is assessing and what these contructs might be, but we are not yet at a point at which we want to test specific hypotheses about how the constructs might be causally related … One of the primary uses of factor analysis is in helping to identify the key constructs needed to account for a particular area of inquiry. Frequently, in the early stages of an area of research, the basic constructs making up the domain of interest have yet to be definitively identified … Fortunately, factor analysis provides a clear method for testing the dimensionality of a set of items and detemining which items appropriately belong together as part of the same scale or subscale. (p. 20-21)

Burston, Eley, Parker, and Tuckett (2017) wanted a valid instrument to measure moral distress within the Australian residential and community care environment. No

instruments had been specifically designed for the aged care environment, so a similar instrument was chosen from the acute care environment. Burston et al. (2017) described the process to develop a valid instrument from a closely related instrument as follows:

> Normality was assessed using histograms and measures of shape (kurtosis and skew). Statistical testing of data to determine internal consistency [reliability] and construct validity of the amended instrument was undertaken. Reliability of the instrument was assessed using Cronbach's alpha, with a score of 0.70 used to determine reliability. Construct validity was determined using principal axis factoring with orthogonal rotation of extracted factors by varimax rotation [Kaiser normalisation]. (p. 4)

This research will follow this approach in investigating an alternative measurement instrument for flight systems and mission design.

Criteria with high or very high discriminatory power (HVHDP) ratings were investigated, since a goal of this research was data reduction and since HVHDP criteria were of primary interest to retain. Exploratory factor analysis was conducted to determine if there was a factor structure for HVHDP criteria. Exploratory factor analysis and measures of reliability were used to investigate potential data reduction among the criteria with HVHDP ratings. Table 17 provides descriptive statistics of each criterion, ranked from high to low, based on the mean of the discriminatory power ratings.

Table 17 Criteria Ranked by HVHDP Mean Score

	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation
Criteria_21	356	4	1	5	1260	3.54	0.944
Criteria_1	356	4	1	5	1160	3.26	1.278
Criteria_3	356	4	1	5	1035	2.91	1.341
Criteria_20	356	4	1	5	984	2.76	1.473
Criteria_4	356	4	1	5	973	2.73	1.224
Criteria_17	356	4	1	5	915	2.57	1.379
Criteria_2	356	4	1	5	888	2.49	1.34
Criteria_5	356	4	1	5	833	2.34	1.28
Criteria_7	356	4	1	5	830	2.33	1.264
Criteria_22	356	4	1	5	673	1.89	1.065
Criteria_14	356	4	1	5	653	1.83	1.228
Criteria_10	356	4	1	5	545	1.53	1.173
Criteria_19	356	4	1	5	541	1.52	0.962
Criteria_6	356	4	1	5	538	1.51	1.028
Criteria_9	356	4	1	5	527	1.48	1.144
Criteria_16	356	4	1	5	524	1.47	1.036
Criteria_23	356	4	1	5	472	1.33	0.805
Criteria_13	356	4	1	5	444	1.25	0.805
Criteria_12	356	4	1	5	431	1.21	0.789
Criteria_11	356	4	1	5	397	1.12	0.572
Criteria_8	356	4	1	5	397	1.12	0.481
Criteria_18	356	3	1	4	369	1.04	0.321
Criteria_15	356	2	1	3	358	1.01	0.106

The review of the means and standard deviations in Table 17 reveals that several criteria had low mean scores and very low standard deviation, which implies that some criteria had a low number of HVHDP ratings. Table 18 lists the total number of high and very high discriminatory power ratings that each criterion received in the sample of 356 mission concept proposals.

Table 18 Criteria Ranked by Number of High or Very High Discriminatory Power (HVHDP) Ratings

Criteria	4	5	No. of High and Very High DP Ratings	% of Sample
21	97	68	165	46%
20	67	59	126	35%
1	30	92	122	34%
3	42	62	104	29%
17	60	42	102	29%
2	61	32	93	26%
4	59	33	92	26%
5	17	32	49	14%
7	15	32	47	13%
10	6	26	32	9%
9	7	24	31	9%
14	1	29	30	8%
13	17	4	21	6%
16	6	15	21	6%
6	3	14	17	5%
19	2	10	12	3%
22	5	7	12	3%
12	0	11	11	3%
23	6	5	11	3%
11	0	5	5	1%
18	4	0	4	1%
8	0	1	1	0%
15	0	0	0	0%

Criteria 16 is highlighted in Table 18; this indicates that criterion and those above it had at least 20 High or Very High discriminatory power ratings. Criteria 23 is highlighted; this indicates that criterion and those above had at least 10 High or Very High discriminatory power ratings. Exploratory factor analysis was conducted on the group of criteria that had 20 or more HVHDP ratings and on the group of criteria that had 10 or more HVHDP ratings.

Outliers

Several criteria were considered to be outliers, due to the very low number of HVHDP ratings or because the criteria only applied to a subgroup of the total missions. The criteria excluded from the HVHDP analysis include criteria 8, 9, 11, 15, and 18. Criterion 9 had over 20 discriminatory power ratings that were high or very high. However, criterion 9 is defined in Appendix B as an assessment of the adequacy of plans for entry descent and landing. This criterion only applies to the subgroup of the sample that are planetary missions. Hair et al. (2010) stated it was inappropriate to apply factor analysis to a sample with two subgroups for a set of items known to differ for the two subgroups. Since criterion 9 was not applicable to both planetary and non-planetary missions, criterion 9 was removed from consideration in the factor analysis. Criterion 8, which was an assessment of the adequacy of the plans for launch operations, received the third lowest point total as shown in Table 18, and received only one HVHDP rating, as shown in Table 19. Criterion 15 was an assessment of the maturity and technical readiness of the operations systems. As shown in Table 19, this criterion received no HVHDP ratings. Criterion 18 was an assessment of the overall mission architecture. There are only four HVHDP ratings for criteria 18, which was used predominately for complex planetary missions. Three of the four HVHDP for criterion 18 were on planetary missions, so this criterion was excluded due to the very low number of HVHDP ratings and because it predominately applied to the planetary subgroup of missions. Criterion 11 was an assessment of the plans for advanced engineering developments and was rarely used, since only five HVHDP ratings were received out of a sample of 356. Criteria 8, 9, 11, 15, and 18 were outliers, as discussed above, and were excluded from the factor analysis on HVHDP criteria.

Criteria with 20 or More HVHDP Ratings

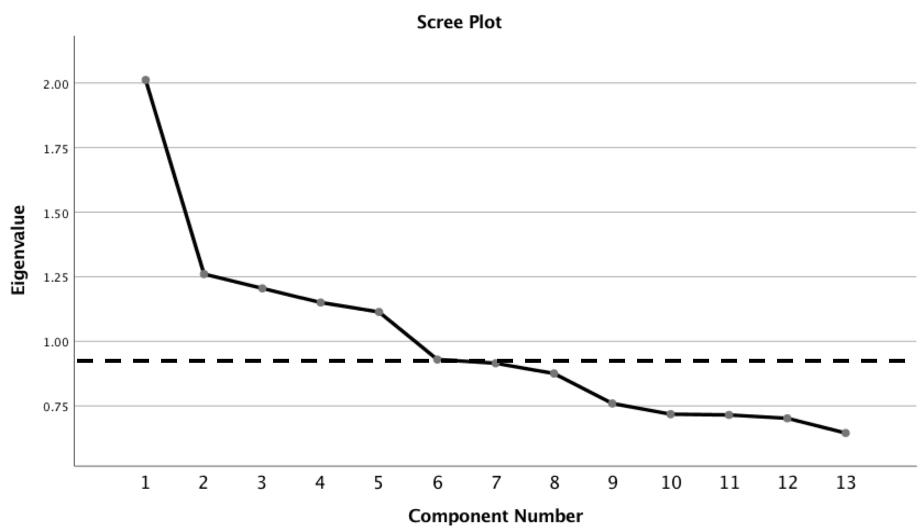
The criteria that had 20 or more HVHDP ratings and were not excluded for other reasons (criteria 9) are: 1, 2, 3, 4, 5, 7, 10, 13, 14, 16, 17, 20, and 21. This data set was referred to as 20 HVHDP. Exploratory factor analysis was performed to identify a factor structure for the 20 HVHDP criteria. The unit of analysis is the criteria. This type of factor analysis, where the unit of analysis is the variable, is referred to as R factor analysis (Hair et al., 2010). Hair et al. (2010) stated that "factor analysis can also be used to achieve data reduction in two ways….the purpose is to retain the nature and character of the original variable, but reduce their number to simplify the subsequent multivariate analysis" (p. 98). Principal component analysis was used as the extraction method since, according to Hair et al. (2010), "Component factor analysis is most appropriate when: Data reduction is a primary concern, focusing on the minimum number of factors needed to account for the maximum portion of total variance represented in the original set of variables" (p.107). Osborne (2015) stated that "PCA computes the analysis without regard to the underlying latent structure of the variables, using all the variance in the manifest variables" (p. 1). The criteria for extracting factors for this research were that the factor solution would have at least 60% of the variance explained. Hair et al. (2010) stated that "As with other aspects of multivariate models, parsimony is important. The notable exception is when factor analysis is used strictly for data reduction and a set level of variance to be extracted is specified" (p. 111). In addition, Pett et al. (2003) indicated that the original goals of the factor analysis should also be considered when deciding how many factors to extract. For this research, a larger than usual number of factors was extracted, in order to retain criteria in the factor solution that would have a large number

of HVHDP scores. A substantial number of the HVHDP criteria were needed to cover the scope of a spaceflight mission.

Principal component analysis was used, since many of the criteria were nonnormally distributed. As discussed previously, Fabrigar et al. (1999) recommended the use of principal component analysis, since it has the advantage of no distributional assumptions.

A scree plot is shown in Figure 8 for the criteria with 20 HVHDP ratings. Hair et al. (2010) stated “The point at which the curve first begins to straighten out is consider to indicate the maximum number of factors to extract” (p.110). This resulted in a maximum of six factors to be extracted, as shown by the dash line in Figure 8.

Figure 8 Scree Plot for Criteria with 20 HVHDP



The six-factor solution is shown in Table 19. It includes nine of the original 13 criteria 20 HVHDP criteria. The factor structure had a simple structure that is defined by Santos and Clegg (1999) as “All input factors loading on to a specific construct should exhibit one-way moderate to high loading (coefficient of .40 or greater) and very low complementary loading (ideally approaching zero) on other constructs” (p. 3).

Table 19 Six Factor Solution for 20 HVHDP Criteria

Rotated Component Matrix[a]

	Component					
	1	2	3	4	5	6
Criteria_5	0.722					
Criteria_7	0.688					
Criteria_14	0.621					
Criteria_1		0.761				
Criteria_17		0.732				
Criteria_21			0.882			
Criteria_2				0.907		
Criteria_3					0.945	
Criteria_16						0.901

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 8 iterations.

Bartlett's test of sphericity shown in Table 20 was less than 0.05, which indicated, according to Yong and Pearce (2013), that there were relationships among the criteria.

Table 20 KMO and Bartlett's Test for Six Factor of 20 HVHDP

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.646
Bartlett's Test of Sphericity	Approx. Chi-Square	140.596
	df	36
	Sig.	0.000

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA), shown in Table 20, was 0.646, which was greater than the minimum of .5 that was required. Individual criteria passed checks for MSA, as shown in the diagonal of the anti-image correlation matrix in Table 21.

Table 21 Anti-image Correlation Matrix for Six Factor of 20 HVHDP

Anti-image Correlation	Criteria_1	Criteria_2	Criteria_3	Criteria_5	Criteria_7	Criteria_14	Criteria_16	Criteria_17	Criteria_21
Criteria_1	.617[a]	0.045	0.044	-0.074	-0.092	-0.044	-0.075	-0.145	0.051
Criteria_2	0.045	.573[a]	-0.04	-0.08	-0.048	0.037	0.014	-0.092	-0.099
Criteria_3	0.044	-0.04	.682[a]	-0.017	-0.083	-0.063	-0.082	-0.061	0.001
Criteria_5	-0.074	-0.08	-0.017	.641[a]	-0.233	-0.096	-0.088	0.056	0.03
Criteria_7	-0.092	-0.048	-0.083	-0.233	.660[a]	-0.191	-0.097	-0.056	-0.072
Criteria_14	-0.044	0.037	-0.063	-0.096	-0.191	.670[a]	-0.028	0.028	-0.091
Criteria_16	-0.075	0.014	-0.082	-0.088	-0.097	-0.028	.706[a]	-0.117	-0.076
Criteria_17	-0.145	-0.092	-0.061	0.056	-0.056	0.028	-0.117	.587[a]	-0.085
Criteria_21	0.051	-0.099	0.001	0.03	-0.072	-0.091	-0.076	-0.085	.614[a]

a. Measures of Sampling Adequacy(MSA)

The criteria for the six-factor solution passed the test for communality, as shown in Table 22.

Table 22 Communalities for Six Factor Solution of 20 HVHDP

	Initial	Extraction
Criteria_1	1	0.741
Criteria_2	1	0.856
Criteria_3	1	0.913
Criteria_5	1	0.696
Criteria_7	1	0.536
Criteria_14	1	0.718
Criteria_16	1	0.874
Criteria_17	1	0.701
Criteria_21	1	0.812

Extraction Method: Principal Component Analysis.

The total variance, shown in Table 23, was 76.1% , which was acceptable.

Table 23 Total Variance Explained for Six Factor Solution of 20 HVHDP

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.825	20.278	20.278	1.825	20.278	20.278	1.521	16.899	16.899
2	1.145	12.72	32.998	1.145	12.72	32.998	1.158	12.863	29.762
3	1.087	12.079	45.077	1.087	12.079	45.077	1.083	12.034	41.795
4	0.974	10.826	55.903	0.974	10.826	55.903	1.05	11.662	53.458
5	0.956	10.621	66.524	0.956	10.621	66.524	1.034	11.492	64.95
6	0.859	9.541	76.065	0.859	9.541	76.065	1	11.115	76.065
7	0.747	8.299	84.364						
8	0.738	8.201	92.565						
9	0.669	7.435	100						

Extraction Method: Principal Component Analysis.

This six-factor solution had only two factors that had more than one criterion. Factor one was composed of criteria 5, 7, and 14. Factor two was composed of criteria 1 and 17. The reliability measure Cronbach's alpha, as shown in Table 24, was 0.481 and did not meet the minimum of .6.

Table 24 Cronbach's Alpha for Factor One of 20 HVHDP

Cronbach's Alpha	N of Items
0.481	3

Cronbach's alpha, if item deleted, shown in Table 25, was under the minimum of .6 for all criteria. All of the criteria failed the item-to-total correlation minimum of .5, as shown in Table 25.

Table 25 Item-Total and Cronbach's Alpha if Item Deleted for Factor One of Six Factor 20 HVHDP

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Criteria_5	4.17	3.885	0.289	0.093	0.402
Criteria_7	4.17	3.671	0.354	0.126	0.286
Criteria_14	4.67	4.171	0.26	0.073	0.448

In addition, all of the criteria failed the inter-item correlation minimum of .3, as shown in Table 26.

Table 26 Inter-Item Correlation Matrix Factor One of 20 HVHDP

	Criteria_5	Criteria_7	Criteria_14
Criteria_5	1	0.289	0.167
Criteria_7	0.289	1	0.252
Criteria_14	0.167	0.252	1

Since factor one failed these internal consistency measures of reliability, factor one was not reliable and not valid.

The reliability measures were also reviewed for factor two, which was composed of criteria 1 and 17. The reliability measure Cronbach's alpha was 0.267 for factor two and did not meet the minimum of .6, as shown in Table 27.

Table 27 Cronbach's Alpha of Factor Two for Six Factor Solution of 20 HVHDP

Cronbach's Alpha	N of Items
0.267	2

Cronbach's alpha if item deleted was not calculated, since there were only two criteria. All of the criteria failed the item-to-total correlation minimum of .5, as shown in Table 28.

Table 28 Item-Total of Factor two for Six Factor Solution of 20 HVHDP

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Criteria_14	2.57	1.902	0.154	0.024	
Criteria_17	3.26	1.634	0.154	0.024	

In addition, all of the criteria failed the inter-item correlation minimum of .3. A summary of the MSA and internal consistency measures for the six factor solution of the 20 HVHDP is shown in Appendix E Table E.2. In summary, the 20 HVHDP six factor solution failed all three reliability tests of internal consistency. According to Hair et al. (2010), internal consistency is a commonly used measure of reliability "which applies to the consistency among variables in a summated scale. The rationale for internal consistency is that the individual items or indicators of the scale should all be measuring the same construct and thus be highly intercorrelated" (p.125). Since factor two failed these internal consistency measures of reliability, factor two was not a reliable measure of the factor.

Factor Structure Stability of 20 HVHDP

The factor structure of an existing scale can be assessed by using another sample to test the instrument, to perform factor analysis, and to compare the two resulting factor structures. Walsh, Seldomridge, and Badros (2007) stated the following:

> The purpose of the study was to re-examine the stability of the factor structure of the California Critical Thinking Disposition Inventory using principal components factor analytic procedures. The question the researchers sought to answer was whether the structure of the four factors reported in the studies by Walsh and Hardy (1997) and Kakai (2003) showed stability upon re-examination. (p. 145)

Hair et al. (2010) recommends a similar process for assessing the factor structure stability, as follows:

> "If sample size permits, the researcher may wish to randomly split the sample into two subsets and estimate the factor models for each subset. Comparison of

the two resulting factor matrices will provide an assessment of the robustness of the solution across the sample. (p. 122)

An assessment of the stability of a solution across the sample was tested by creating two random sets, A and B, with each having 178 samples. Suhr (2006) recommended a minimum of five to 20 samples per factor. Assuming a six factor solution, the sample size for set A and B was adequate to analyze factor structure stability. Exploratory factor analysis was conducted, which resulted in a six factor solution for both sets. The six factor solution for set A is shown in Table 29.

Table 29 Set A Six Factor Solution of 20 HVHDP Criteria

Rotated Component Matrix[a]

	Component					
	1	2	3	4	5	6
Criteria_3	0.861					
Criteria_16	0.612					
Criteria_14		0.908				
Criteria_7		0.626				
Criteria_1			0.849			
Criteria_17			0.525	-0.499	0.422	
Criteria_5				0.824		
Criteria_2					0.903	
Criteria_21						0.971

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 11 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for set A, shown in Table 30, was 0.575, which was greater than the minimum .5 required. Bartlett's test of sphericity, shown in Table 30, was less than 0.05, which indicated, according to Yong and Pearce (2013), that there were relationships among the criteria. In addition, the total variance extracted was 67.8%, which was above the minimum of 60%.

Table 30 KMO and Bartlett’s Test for Set A of 20 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.575
Bartlett's Test of Sphericity	Approx. Chi-Square	91.72
	df	36
	Sig.	0.000

The factor structure for set B with 20 HVHDP Criteria is shown in Table 31.

Table 31 Set B Six Factor Solution of 20 HVHDP

	Component					
	1	2	3	4	5	6
Criteria_5	0.818					
Criteria_7	0.767					
Criteria_16		0.76				
Criteria_17		0.727	0.458			
Criteria_10			0.871			
Criteria_1				0.98		
Criteria_2					0.992	
Criteria_3						0.973

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 8 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for set B, shown in Table 32, was 0.596, which is greater than the .5 minimum required. Bartlett’s test of sphericity, shown in Table , was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 67.8%, which was above the minimum of 82.3%.

Table 32 KMO and Bartlett's Test for Set B of 20 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.596
Bartlett's Test of Sphericity	Approx. Chi-Square	55.099
	df	28
	Sig.	0.002

The entire factor solution of random set A and B were not the same, leading to the conclusion that the overall factor structure of either set A or set B for 20 HVHDP was not generalizable across the full sample. However, criterion 2 was a single criteria factor in set A and B and in the full data set six factor solution for the 20 HVHDP criteria. Criterion 2 was stable across the full data set and two random samples, using the 20 HVHDP criteria.

A summary of MSA and internal consistency measures set A and B for the 20 HVHDP criteria is shown in Appendix E Table E.3.

Criteria with 10 or More HVHDP Ratings

A larger set of criteria that have ten or more HVHDP ratings is identified in Table 18. Criteria 23 and those listed above criteria 23 in Table 18 all have at least ten HVHDP ratings. Exploratory factor analysis was conducted on this larger group of criteria that have ten or more HVHDP ratings. This set included criteria 1 through 7, 10, 12, 13, 14, 16, 17, 19, and 20 through 23. Exploratory factor analysis was used to identify potential data reduction among the criteria with 10 HVHDP ratings. The seven factor solution for the 10 HVHDP criteria is shown in Table 33.

Table 33 Seven Factor Solution for 10 HVHDP

	Component						
	1	2	3	4	5	6	7
Criteria_4	0.665						
Criteria_23	0.638						
Criteria_3	0.465						
Criteria_12		0.703					
Criteria_16		0.668					
Criteria_19			-0.735				
Criteria_5			0.607				
Criteria_2				0.702			
Criteria_6				0.636			
Criteria_14					0.878		
Criteria_1						0.836	
Criteria_17						0.568	
Criteria_20							0.727
Criteria_21							0.631

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 17 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for the seven factor 10 HVHDP, shown in Table 34, was 0.650 which is greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 34, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 62.3%, which was above the minimum of 60 %. Factors 1, 2, 3, 4, 6, and 7 did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the seven factor solution of the 10 HVHDP is shown in Appendix E Table E.4.

Table 34 KMO and Bartlett's Test for Seven Factor 10 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.65
Bartlett's Test of Sphericity	Approx. Chi-Square	258.938
	df	91
	Sig.	0.000

Hair et al. (2010) stated,

> Validation of any factor analysis result is essential, particularly when attempting to define underlying structure among the variables. Optimally, we would always follow our use of factor analysis with some form of confirmatory factor analysis, such as structural equation modeling … but this type of follow-up is often not feasible. We must look to other means, such as split sample analysis or applications. (p. 139)

According to Hair et al. (2010), if the two split samples provide similar results, then "we can be reasonably assured that the results are stable within our sample" (p. 141).

In order to test the robustness of a solution across the sample, two random sets, A and B, were created, with each having 178 samples. Suhr (2006) recommended a minimum of five to 20 samples per factor. The sample size for set A and B was adequate to analyze a seven factor structure. Exploratory factor analysis was conducted, which resulted in a seven factor solution for both sets. The seven factor solution for set A is shown in Table 35.

Table 35 Set A Seven Factor Solution for 10 HVHDP

	Component						
	1	2	3	4	5	6	7
Criteria_4	0.702						
Criteria_23	0.649						
Criteria_21	0.559	0.436					
Criteria_12		0.782					
Criteria_16		0.678					
Criteria_5			0.715				
Criteria_19			-0.608		0.501		
Criteria_2				0.769			
Criteria_6				0.613			
Criteria_14					0.859		
Criteria_1						0.763	
Criteria_17						0.675	
Criteria_20							0.933

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 22 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for set A 10 HVHDP, shown in Table 36, was 0.618, which was greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 36, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 67.8%, which is above the minimum of 60 %. Factors one through six did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the seven factor solution of the set A 10 HVHDP is shown in Appendix E Table E.5.

Table 36 KMO and Barlett's Test for Set A 10 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.618
Bartlett's Test of Sphericity	Approx. Chi-Square	153.918
	df	78
	Sig.	0.000

The factor structure for set B is shown in Table 37.

Table 37 Set B Seven Factor Solution for 10 HVHDP

	Component						
	1	2	3	4	5	6	7
Criteria_5	0.734						
Criteria_19	-0.708						
Criteria_2		0.819					
Criteria_23		0.684					
Criteria_21			0.895				
Criteria_14	0.464		0.469				
Criteria_1				0.782			
Criteria_17				0.735			
Criteria_6					0.736		
Criteria_4					-0.641		
Criteria_3						0.918	
Criteria_12							0.914

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 11 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for set B 10 HVHDP, shown in Table 38, was 0.607, which is greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 38, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 71.6%, which was above the minimum of 60 %. Factors one through five did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the seven factor solution of the set B 10 HVHDP is shown in Appendix E Table E.5.

Table 38 KMO and Bartletts's Test for Set B 10 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.607
Bartlett's Test of Sphericity	Approx. Chi-Square	124.458
	df	66
	Sig.	0.000

Set A and set B had two similar factors. Factor six in set A had criteria 1 and 17, which was the same as factor four in set B. Factor three in set A has two criteria, 5 and 19, which were the same as two of the three criteria in factor one in set B. Criteria 14 also loaded on factor one in set B, but it had a low cross-loading. In addition, the full data set factor for 10 HVHDP had factor three with criteria 5 and 19, and factor six with criteria 1 and 17. It is reasonable to conclude that a factor with criteria 1 and 17 and a factor with criteria 5 and 19 would be stable within the 10 HVHDP data set. However, the entire factor structure of set A or set B was not stable across the HVHDP data set. A summary of MSA and internal consistency measures for set A and B using the 10 HVHDP criteria is shown in Appendix E Table E.5.

Subgroup Analysis

Since the analysis of the two random samples (set A and set B) from the full data set did not support factor structure stability for the 20 HVHDP or for the full 10 HVHDP data sets, subgroups within the full sample were compared to determine if they had significantly different factor structures that could be the causing factor structure instability in the full set of data. Tabachnick and Fidell (2013) stated that

> First, samples that are known to be different with respect to some criterion (e.g., socioeconomic status) may also have different factors. Examination of group differences is often quite revealing. Second, underlying factor structure may shift in time for the same subjects with learning or with experience in an experimental setting and these differences may also be quite revealing. Pooling results from diverse groups in FA may obscure differences rather than illuminate them. On the other hand, if different samples do produce the same factors, pooling them is desirable because of increase in sample size. (p. 617)

Hair et al. (2010) agreed that “whenever differing groups are expected in the sample, separate factor analyses should be performed, and the results should be compared to identify differences not reflected in the results of the combined sample” (p. 103). Hair et al. (2010) further stated that

> Variables that are better discriminators between the subgroups of the sample will load on the later factors, many times those not selected by the criteria discussed previously. When the objective is to identify factors that discriminate among the subgroups of a sample, the researcher should extract additional factors beyond those indicated by the methods just discussed and examine the additional factors’ ability to discriminate among the groups. (p.110-111)

Two subgroups were created from the full data set, and factor analysis was conducted on the subgroup of planetary records and the factor structure was compared to the factor structure of the subgroup of non-planetary records. A second subgroup analysis was conducted, to determine if the factor structure changed over time. The data set was split in half by time, and the factor structure of the newer records was compared to the factor structure of the older records.

Planetary Versus Non-Planetary

The full data set contained planetary missions and non-planetary missions. There were 182 records that were planetary and 174 that were non-planetary missions. Factor analysis was performed on the planetary data set and on the non-planetary data set and the factor structure was compared. The factor structure for the planetary data using the 20 HVHDP criteria is shown in Table 39.

Table 39 Planetary Factor Solution of 20 HVHDP Criteria

	Component				
	1	2	3	4	5
Criteria_14	0.759				
Criteria_7	0.702				
Criteria_3		0.814			
Criteria_20		-0.605			
Criteria_16			0.777		
Criteria_10			-0.616		
Criteria_21				0.902	
Criteria_5					0.911

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 9 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for Planetary 20 HVHDP, shown in Table 40, was 0.586, which was greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 40, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 71%, which was above the minimum of 60%. Factors 1 through 3 did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the five factor solution of the Planetary 20 HVHDP is shown in Appendix E Table E.6.

Table 40 KMO and Bartlett's Test for Planetary 20 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.586
Bartlett's Test of Sphericity	Approx. Chi-Square	44.199
	df	28
	Sig.	0.027

The factor structure for the non-planetary data is shown in Table 41.

Table 41 Non-Planetary Factor Solution of 20 HVHDP Criteria

	Component				
	1	2	3	4	5
Criteria_7	0.76				
Criteria_14	0.716				
Criteria_5	0.658				
Criteria_17		0.775			
Criteria_10		0.73			
Criteria_13			0.891		
Criteria_1				0.783	
Criteria_2			-0.411	-0.647	
Criteria_3					0.927

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for Non-Planetary 20 HVHDP, shown in Table 42, was 0.565, which is greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 42, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 68%, which was above the minimum of 60%. Factors one through four did not meet reliability measures of internal consistency, with the exception that the inter-item correlation between C5 and C7 was above .3. A summary of MSA and

internal consistency measures for the five factor solution of the Non-Planetary 20 HVHDP is shown in Appendix E Table E.6.

Table 42 KMO and Bartlett's Test for Non-planetary 20 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.565
Bartlett's Test of Sphericity	Approx. Chi-Square	79.091
	df	36
	Sig.	0.000

The planetary and non-planetary 20 HVHDP factor structures had criteria 3, 5, 7, 10, and 14 in common. In addition, the planetary, non-planetary, and full set 20 HVHDP factor solutions had criteria 3, 5, 7, and 14 in common. These four criteria were consistently found to contribute significantly to the percentage of variance explained. In addition, the factor composed of 5, 7, and 14 was in the full set factor solution for 20 HVHDP and in the non-planetary factor solution for 20 HVHDP. In addition, a factor composed of criteria 7 and 14 was in the planetary factor solution for 20 HVHDP. For all three cases, the factor is the first in the factor solutions which represents the largest eigenvalue. A factor with criteria 7 and 14 was generalizable across the full set, planetary and non-planetary 20 HVHDP. However, three out of eight criteria of the planetary factor solution were different from any criteria in the non-planetary solution. Also, four of the nine criteria in the non-planetary solution were different from any criteria in the planetary solution. This indicates that the planetary and non-planetary data were significantly different, in which criteria were important and in the factor structure. This significant difference between the planetary and non-planetary factor solutions and

the criteria composing the factor solutions could be one of the causes of factor instability in the full data set.

Factor analysis was also performed on the planetary data set and the non-planetary data set for the 10 HVHDP criteria. The factor structure for the planetary data using the 10 HVHDP criteria is shown in Table 43.

Table 43 Planetary Factor Solution of 10 HVHDP Criteria

	Component					
	1	2	3	4	5	6
Criteria_6	0.868					
Criteria_14	0.624					
Criteria_23		0.69				
Criteria_10		0.679				
Criteria_17			0.842			
Criteria_12				0.895		
Criteria_16				0.477		
Criteria_3					0.845	
Criteria_20					-0.478	
Criteria_1						0.908

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 24 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for Planetary 10 HVHDP, shown in Table 44, was 0.553, which was greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 44, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 71.2%, which was above the minimum of 60%. Factors one, two, four, and five did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the six factor solution of the Planetary 10 HVHDP is shown in Appendix E Table E.7.

Table 44 KMO and Bartlett's Test for Planetary 10 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.553
Bartlett's Test of Sphericity	Approx. Chi-Square	76.655
	df	45
	Sig.	0.002

The non-planetary factor structure is shown in Table 45.

Table 45 Non-Planetary Factor Solution for 10 HVHDP Criteria

	Component					
	1	2	3	4	5	6
Criteria_19	0.812					
Criteria_10	0.811					
Criteria_5		0.909				
Criteria_7		0.605	0.423			
Criteria_4			0.828			
Criteria_23			0.662	0.435		
Criteria_16				0.919		
Criteria_17					0.946	
Criteria_3						0.981

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for non-planetary 10 HVHDP, shown in Table 46, was 0.666, which was greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 46, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 80.5%, which was above the minimum of 60 %. Factors one, two, three, and four did not meet all reliability measures of internal consistency. However, inter-item correlation was passed for factor one with criteria 10 and 19, for factor two for with criteria 5 and 7, for factor three with the pair of criteria 4 and 7, and

for the pair of criteria 7 and 23. A summary of MSA and internal consistency measures for the six factor solution of the non-planetary 10 HVHDP is shown in Appendix E Table E.7.

Table 46 KMO and Bartlett's Test for Non-Planetary 10 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.666
Bartlett's Test of Sphericity	Approx. Chi-Square	157.993
	df	36
	Sig.	0.000

The planetary structure and non-planetary structure for 10 HVHDP each have criteria 17 as a single criteria factor. However, criteria 17 was not a single criteria factor in the full data set 10 HVHDP factoring. The remaining factor structure of the planetary and the non-planetary data sets were not similar. Five of the ten criteria of the planetary factor solution did not appear in the non-planetary factor solution. Also, five of the nine criteria in the non-planetary factor solution did not appear in the planetary factor solution for 10 HVHDP. Since the two subgroups had five out of six factors that were different, and approximately half of their respective criteria were different, the differences in the planetary and non-planetary data may be the cause of factor instability in the full set factor structure. A summary of the MSA and internal consistency measures for the planetary and non-planetary data sets using the 10 HVHDP criteria is shown in Table E.7.

The two subgroups of planetary missions and non-planetary missions were tested with the 20 HVHDP criteria and the 10 HVHDP criteria. The criteria composing the factor solutions and the factors were not generally similar. Since the planetary and non-

planetary subgroups did not have similar factor structures, they may be the source of the factor structure instability in the full data set analysis.

Recent CBR Records Versus Older CBR records

Since the analysis of the two random samples (set A and set B) from the full data set did not support full factor structure stability for the 20 HVHDP or for the 10 HVHDP data sets, subgroups within the full sample were compared, in order to determine if they had significantly different factor structures that would change over time. The full data set was composed of recent CBR records (2009-2018) and older CBR records (1998-2007). There were 165 recent records from 2009-2018 and there were 191 older CBR records from 1998-2007. Factor analysis was performed on the recent record set and on the older record set, and the factor structure compared. The factor structure for the recent records (2009-2018) using the 20 HVHDP criteria is shown in Table 47.

Table 47 2009-2018 Factor Solution of 20 HVHDP

	Component					
	1	2	3	4	5	6
Criteria_10	0.772					
Criteria_4	0.591					
Criteria_17	0.525					
Criteria_20		0.827				
Criteria_5		0.714				
Criteria_14			0.75			
Criteria_21			0.712			
Criteria_2				0.872		
Criteria_16					0.92	
Criteria_1						0.853

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 18 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for recent data 20 HVHDP, shown in Table 48, was 0.589, which was greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 48, was less than

0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 71.3%, which was above the minimum of 60%. Factors one, two, and three did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the six factor solution of the non-planetary 10 HVHDP is shown in Appendix E Table E.8.

Table 48 KMO and Bartlett's Test for 2009-2018 20 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.589
Bartlett's Test of Sphericity	Approx. Chi-Square	74.918
	df	45
	Sig.	0.003

The factor structure for the older data (1998-2007) is shown in Table 49. The only similarity is that criterion 2 is a single criterion factor in both the recent and older data sets. Generally, the factor structure of the recent and the older records is not similar.

Table 49 1998-2007 Factor Solution for 20 HVHDP

	Component					
	1	2	3	4	5	6
Criteria_5	0.721					
Criteria_7	0.667					
Criteria_14	0.595					
Criteria_10		0.724				
Criteria_20		0.65				
Criteria_17			0.841			
Criteria_16			0.497			
Criteria_3				0.867		
Criteria_2					0.901	
Criteria_21						0.934

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 9 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for older data (1998-2007) 20 HVHDP, shown in Table 50, was 0.600, which is greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 50, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 71%, which was above the minimum of 60%. Factors one, two, and three did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the six factor solution of the 1998-2007 20 HVHDP is shown in Appendix E Table E.8.

Table 50 KMO and Bartlett's Test for 1998-2007 20 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.6
Bartlett's Test of Sphericity	Approx. Chi-Square	85.586
	df	45
	Sig.	0.000

The factor structure for the recent data (2009-2018) was generally not similar to the factor structure for the older data (1998-2007). Since the two subgroups had different factor structures, the two subgroups may be one of the causes of factor instability in the full set factor structure. This indicates that the ratings on the 20 HVHDP criteria may be changing over time. However, criterion 2 was a single factor in both the recent and the older factor structure. Criterion 2 was also a single factor in the full 20 HVHDP data set, the set A 20 HVHDP, and the set B 20 HVHDP. Criterion 2, as a single factor, was generalizable across the 20 HVHDP data set.

Inter-item Correlation Difference Between Planetary and Non-Planetary

Criteria 5 and 7 had high inter-item correlation (>.3) for the non-planetary 10 HVHDP and 20 HVHDP factors solutions. However, when criteria 5 and 7 appeared in a factor in other data sets, including the full set 20 HVHDP, random set B 20 HVHDP, and the oldest data set 1998- 2007, the inter-item correlation was low (<.3). A factor with criteria 5 and 7 never appeared in the planetary data set factor solutions. This indicates that there is a significant difference in the data between non-planetary data sets and the full data set, random set B and the oldest data set. These inter-item correlation differences between non-planetary and other data sets are summarized in Table E.9

Recent Planetary Records

The most recent 91 records of the planetary data were analyzed to determine if the most recent planetary records had a valid factor structure. The new ranking of the HVHDP criteria for the recent planetary data is shown in Table 51. Factor analysis was conducted on the criteria with four or more high or very high discriminatory power (HVHDP) rankings. This is referred to as the Planetary 4 HVHDP criteria (1-5, 7, 9, 10, 13, 17, 20, 21).

Table 51 Recent Planetary Records HVHDP Ranking

Criteria	4s	5s	Total	% of recent Planetary records
Criteria 21	29	24	53	58%
Criteria 20	13	21	34	37%
Criteria 1	7	26	33	36%
Criteria 17	19	11	30	33%
Criteria 3	14	11	25	27%
Criteria 2	18	6	24	26%
Criteria 4	13	9	22	24%
Criteria 10	6	9	15	16%
Criteria 9	4	9	13	14%
Criteria 5	4	8	12	13%
Criteria 7	0	9	9	10%
Criteria 13	6	1	7	8%
Criteria 14	1	4	5	5%
Criteria 16	0	4	4	4%
Criteria 23	3	1	4	4%
Criteria 6	0	2	2	2%
Criteria 11	0	2	2	2%
Criteria 12	0	2	2	2%
Criteria 22	0	2	2	2%
Criteria 18	1	0	1	1%
Criteria 8	0	0	0	0%
Criteria 15	0	0	0	0%
Criteria 19	0	0	0	0%

The factor structure for the recent planetary records is shown in Table 52.

Table 52 Factor Structure for Recent Planetary Data and Planetary 4 HVHDP

	Component					
	1	2	3	4	5	6
Criteria_4	0.831					
Criteria_9	0.63					
Criteria_7		0.773				
Criteria_14		0.688				
Criteria_1			0.796			
Criteria_21			-0.525			
Criteria_5				0.884		
Criteria_10			0.47	0.498		
Criteria_23					0.9	
Criteria_20						0.926

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 8 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) recent planetary and planetary 4 HVHDP criteria, shown in Table 53, were 0.670, which is greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 53, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 71%, which was above the minimum of 60%. Factors one, two, and three did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the six factor solution of the recent planetary data and planetary 4 HVHDP criteria is shown in Appendix E Table E.10

Table 53 KMO and Bartlett's Test for Recent Planetary and Planetary 4 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.670
Bartlett's Test of Sphericity	Approx. Chi-Square	66.925
	df	45
	Sig.	0.019

This six factor solution passed measures of KMO, Bartlett's test of sphericity, individual MSA, and the total variance explained. However, factors one, two, three, and four failed measures of internal consistency and were not valid. A summary of MSA and internal consistency measures for factors 1 through 4 is shown in Appendix E Table E.10.

Recent Non-planetary Records

The most recent 90 records of the non-planetary data were analyzed to determine if the most recent non-planetary records had a valid factor structure. The new ranking of the HVHDP criteria for the recent non-planetary data is shown in Table 54. Factor analysis was conducted on the criteria with four or more high or very high discriminatory power (HVHDP) rankings. This is referred to as the non-planetary 4 HVHDP criteria (1-5, 10, 13, 14, 17, 19, 20, 21, 22).

Table 54 Recent Non-planetary Records HVHDP Ranking

Criteria	4s	5s	Total	% of Recent Non-planetary Records
Criteria 21	27	7	34	38%
Criteria 20	22	6	28	31%
Criteria 1	4	15	19	21%
Criteria 2	9	8	17	19%
Criteria 17	5	12	17	19%
Criteria 4	15	1	16	18%
Criteria 3	3	12	15	17%
Criteria 10	0	13	13	14%
Criteria 19	1	6	7	8%
Criteria 5	2	3	5	6%
Criteria 14	0	5	5	6%
Criteria 22	2	3	5	6%
Criteria 13	4	0	4	4%
Criteria 7	1	1	2	2%
Criteria 8	0	1	1	1%
Criteria 11	0	1	1	1%
Criteria 16	0	1	1	1%
Criteria 18	1	0	1	1%
Criteria 6	0	0	0	0%
Criteria 9	0	0	0	0%
Criteria 12	0	0	0	0%
Criteria 15	0	0	0	0%
Criteria 23	0	0	0	0%

The factor structure for the recent non-planetary records using the Non-planetary 4 HVHDP criteria is shown in Table 55.

Table 55 Factor Structure for Recent Non-Planetary and Non-Planetary 4 HVHDP

	Component					
	1	2	3	4	5	6
Criteria_10	0.892					
Criteria_19	0.679					
Criteria_1		0.871				
Criteria_17		0.675				
Criteria_2			0.951			
Criteria_20				0.913		
Criteria_13					0.982	
Criteria_14						0.994

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) recent planetary and planetary 4 HVHDP, shown in Table 56, was 0.637, which is greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 56, was less than 0.05 which indicated that there were relationships among the criteria. In addition, the total variance extracted was 67.2%, which is above the minimum of 60%. Factors one and two did not meet reliability measures of internal consistency. A summary of MSA and internal consistency measures for the six factor solution of the recent non-planetary data using non-planetary 4 HVHDP is shown in Appendix E Table E.11

Table 56 KMO and Bartlett's Test for Recent Non-planetary 4 HVHDP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.637
Bartlett's Test of Sphericity	Approx. Chi-Square	46.998
	df	28
	Sig.	0.014

2016 Non-Planetary Data

The recent non-planetary data set was divided further, and a data set was created with non-planetary data with records from 2014 to 2016. There were 51 data points in the 2016 non-planetary data set. This data was factored, using the full set of 23 original criteria.

The factor solution is shown below in Table 57.

Table 57 Factor Structure for 2016 Non-Planetary Data with 23 Criteria

Rotated Component Matrix[a]			
	Component		
	1	2	3
Criteria_11	0.904		
Criteria_5	0.835		
Criteria_23		0.917	
Criteria_6		0.816	
Criteria_20			0.945
Extraction Method: Principal Component Analysis.			
Rotation Method: Varimax with Kaiser Normalization.			
a. Rotation converged in 4 iterations.			

The Kaiser-Meyer-Olkin (KMO) overall Measure of Sampling Adequacy (MSA) for the 2016 non-planetary records using the original 23 criteria, shown in Table 58, was 0.543, which was greater than the minimum required of .5. Bartlett's test of sphericity, shown in Table 58, was less than 0.05, which indicated that there were relationships among the criteria. In addition, the total variance extracted was 83%, which was above the minimum of 60%.

Table 58 KMO and Bartlett's Test for 2016 Non-planetary 23 Criteria

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.543
Bartlett's Test of Sphericity	Approx. Chi-Square	44.4
	df	10
	Sig.	0.000

Factor one was composed of criteria 5 and 11. Factor one met reliability measures of internal consistency. Cronbach's alpha for factor one is shown in Table 59. Cronbach's alpha was .652, which was above the minimum of .6.

Table 59 Cronbach's Alpha for Factor One 2016 Non-Planetary

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.652	0.706	2

The inter-item correlations for criteria 5 and 7 that compose factor one are shown in Table 60. The inter-item correlations were .545, which was above the minimum of .3.

Table 60 Factor One Inter-Item Correlations 2016 Non-Planetary

Inter-Item Correlation Matrix		
	Criteria_5	Criteria_11
Criteria_5	1	0.545
Criteria_11	0.545	1

The corrected item to total correlation for criteria 5 and 11 is shown in Table 61. The corrected item to total correlation was .545, which was above the required minimum of .5.

Table 61 Factor One Item to Total Correlation 2016 Non-Planetary

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Criteria_5	1.08	0.314	0.545	0.298	
Criteria_11	1.47	0.854	0.545	0.298	

Factor two was composed of criteria 6 and 23. Factor two met reliability measures of internal consistency. Cronbach's alpha for factor two is shown in Table 62. Cronbach's alpha was .645, which was above the minimum of .6.

Table 62 Cronbach's Alpha for Factor Two 2016 Non-Planetary

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.645	0.704	2

The inter-item correlations for criteria 6 and 23 that compose factor two are shown in Table 63. The inter-item correlations were .543, which was above the minimum of .3.

Table 63 Factor Two Inter-Item Correlations 2016 Non-Planetary

Inter-Item Correlation Matrix		
	Criteria_6	Criteria_23
Criteria_6	1	0.543
Criteria_23	0.543	1

The corrected item to total correlation for criteria 6 and 23 is shown in Table 64. The corrected item to total correlation was .543, which was above the required minimum of .5.

Table 64 Factor Two Item to Total Correlation 2016 Non-Planetary

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Criteria_6	1.06	0.056	0.543	0.295	
Criteria_23	1.14	0.161	0.543	0.295	

Factors one and two passed the measures of internal consistency and were reliable measures.

The 2016 non-planetary factor solution was reliable. A summary of MSA and internal consistency measures for the two factor solution of the 2016 non-planetary data using the original 23 criteria is shown in Appendix E Table E.12.

The criteria composing this three factor solution for the 2016 non-planetary were used to extract three factors in the larger recent non-planetary data set, in order to determine if the factor structure would be generalizable across the larger recent non-planetary data set. However, no factor solution was found, using criteria 5, 6, 11, 20, 23 in the larger non-planetary data set. The 2016 data set included data from three different

AO program competitions. The three expert peer reviews of those proposals were held in 2014, 2015, and 2016. The recent non-planetary data set included two additional expert peer reviews, held in 2011. The 2016 non-planetary factor solution was found to be reliable, but was not generalizable to the earlier non-planetary data set. This indicates that the factor structure changed, over time, for the non-planetary data from 2011 to 2016. This could reflect changes in the experience of the expert peer panel members or changes in the spaceflight discipline, over time.

The recent non-planetary 4 HVHDP was composed of 13 criteria (1-5, 10, 13, 14, 17, 19, 20, 21, 22). A reliable factor solution was not found using these criteria. The 2016 non-planetary factor solution with reliable measures included criteria 5, 6, 11, 20, 23. Criteria 6, 11, and 23 were not in the recent non-planetary 4 HVHDP criteria set. Table 54 lists the number of High and Very High Discriminatory Power (HVHDP) ratings that each criterion received for the non-planetary data set. Criteria 6 and 23 received no HVHDP ratings, and criteria 11 received only one.

CHAPTER 5

CONCLUSIONS

Test of Two Factor Structure Hypotheses

Two factors were extracted from the full data set to test the five hypotheses about the existing two factor structure and criteria. The extracted two factors did not pass individual tests of MSA, the communality was too low for all of the criteria, and the total variance explained by the two factor solution was unacceptably low, at 18%. Both of the extracted factors failed three reliability tests of internal consistency. Criteria 1, 3, 8, and 11 were expected to load on the flight systems factor. However, these criteria did not load on either of the extracted factors. Criteria 18 and 20 were expected to load on the mission design factor, but these criteria did not load on either of the extracted factors. Therefore, Hypothesis H0e, which states that there are only two factors which are the flight systems and the mission design factor, was rejected.

The criteria loading on the two extracted factors was reviewed, and two criteria (12, 13), which were expected to load on the flight systems factor, were loaded instead on two different factors. Since criteria 12 and 13 load on different factors, one of them was not loading on the flight systems factor. In addition, criteria 1 through 16 failed three internal consistency measures of reliability, which indicated that the criteria did not consistently measure the factor. Hypothesis H0c, which states that criteria 1 through 16 measure the flight systems factor, was rejected.

Criteria 17 through 23, which were expected to load only on the mission design factor, loaded instead on both of the two extracted factors, with criteria 17, 21, and 23 loading on the first factor, while criteria 19 and 22 loaded on the second factor. In

addition, criteria 17 through 23 failed three internal consistency measures of reliability which demonstrated that criteria 17 through 23 did not consistently measure the factor. Hypothesis H0d, which states that Criteria 17 through 23 measure the mission design factor, was rejected, since several of those criteria loaded on two different factors in the two factor solution. In addition, all of the criteria 17 through 23 failed reliability measures of internal consistency.

A review of the significant loadings, shown in Table 9, indicated that criteria 1, 3, 8, 11, 18, and 20 did not have loadings significant enough to load on either of the factors in the two factor solution. Since criteria 1, 3, 8, 11, 18, and 20 did not load on either factor, hypothesis H0b, which states that criteria N is necessary to assess flight systems and mission design where N = 1 through 23, was rejected for criteria 1, 3, 8, 11, 18, and 20. In addition, hypothesis H0a, which states that all 23 criteria are necessary to assess flight systems and mission design, was rejected, since criteria 1, 3, 8, 11, 18, and 20 were not necessary, since they did not load on either of the two factors. The hypotheses and the results are summarized in Table 65, below.

Table 65 Hypotheses Results

Hypotheses		
Hypothesis	Rejected/Not Rejected	Rationale for Rejection
H0a: Hypothesis All 23 criteria are necessary to assess flight systems and mission design.	Rejected	All 23 criteria were not needed since criteria 1, 3, 8, 11, 18, and 20 did not load on either factor.
H0b: Hypothesis Criteria N is necessary to assess flight systems and mission design where N = 1 through 23.	Rejected	Criteria 1, 3, 8, 11, 18, 20 were not necessary to assess flight systems and mission design for the current two factor structure since they did not load at a significant level on either factor.
H0c: Hypothesis Criteria 1 through 16 measure the Flight Systems Factor.	Rejected	1) Criteria 4, 5, 6, 7, 9, 12, 14, and 16 loaded on the first factor but criteria, 2, 10, 13, 15 loaded on the second factor. They did not load on the same factor for the two factor solution. 2) Criteria 1 - 16 failed three internal consistency measures of reliability for flight systems. 3) The measurement instrument was not valid since it was not reliable.
H0d: Hypothesis Criteria 17 through 23 measure the Mission Design and Operations Factor.	Rejected	1) Criteria 17, 21, 23 loaded on the first factor, but 19 and 22 loaded on the second factor. They did not load on the same factor for the two factor solution. 2) Criteria 17 - 23 failed three internal consistency measures of reliability for the mission design and operations factor. 3) The measurement instrument was not valid since it was not reliable.
H0e: Hypothesis There are only two Factors which are the Flight Systems Factor and the Mission Design Factor.	Rejected	1) Criteria 8, 11, 13 failed individual tests of MSA. 2) No criteria met guideline of .5 for communalities. 3) Total variance explained for original two factor solution was too low at 18%. 60% variance explained was the minimum acceptable. 4) Both of the extracted factors failed three reliability tests of internal consistency. 5) Criteria 1, 3, 8, 11, 18, and 20 did not load on either extracted factor.

Investigation of Alternative Factor Structures

The criteria were ranked by the number of High or Very High Discriminatory Power (HVHDP) ratings that each criterion received for the full sample of 356 CBR records. Outliers were identified and were excluded from the factor analysis. Factor analysis was performed on a set of criteria that had 20 or more HVHDP (20 HVHDP) ratings and on a set that had 10 or more HVHDP (10 HVHDP) ratings. Both 20 HVHDP and 10 HVHDP sets initially resulted in alternative factor structures that passed Bartlett's test of sphericity, Kaiser-Meyer-Olkin's (KMO) overall measure of sampling adequacy (MSA), and an individual criteria check for MSA, and that showed communalities that were acceptable. However, both alternative factor structures failed reliability measures of internal consistency, indicating that factors with multiple criteria were not consistently measuring the factor. Since the criteria in factors with multiple criteria were not consistently measuring the same construct, this provided justification to retain each of those criteria as a separate measure in the final set of recommended criteria.

Testing Generalizability of a Factor Structure Across the Sample

An assessment of the robustness of a solution across the sample was tested by creating two random sets, labeled A and B. Factor analysis was conducted on set A and on set B, using the 20 HVHDP criteria. The overall factor structure of set A and set B were compared and were not similar, leading to the conclusion that there was no evidence that the factor structure of either set A or set B for the 20 HVHDP was generalizable across the full sample. However, criterion 2 was a single criteria factor in both sets A and B and in the full set six factor solution for the 20 HVHDP criteria. Criterion 2 was stable across the full data set and the two random samples using the 20 HVHDP criteria.

Factor analysis was also conducted on set A and set B, using the 10 HVHDP criteria. There were two factors identified that were stable across the 10 HVHDP data set. One factor included criteria 1 and 17, and the other factor included criteria 5 and 19. However, the entire factor structure of set A or set B was not stable across the 10 HVHDP data set. Since the full factor structures for 20 HVHDP and 10 HVHDP for the two random sets A and B were not generalizable, subgroups within the full data set were investigated as a source of the factor structure instability.

Subgroup Analysis of Planetary and Non-Planetary

Since the analysis of the two random samples (set A and set B) from the full data set did not support factor structure stability for the 20 HVHDP or 10 HVHDP analysis, subgroups within the full sample were factored and compared, in order to determine if subgroups had significantly different factor structures that could be the cause of factor structure instability in the full set of data. The two subgroups of planetary missions and non-planetary missions were tested with the 20 HVHDP criteria. One factor with criteria 7 and 14 was generalizable across the planetary and non-planetary 20 HVHDP. However, three out of eight criteria of the planetary factor solution were different from any criteria in the non-planetary solution. Also, four of the nine criteria in the non-planetary solution were different from any criteria in the planetary solution. This indicates that the planetary and non-planetary data were significantly different for the 20 HVHDP criteria. The two subgroups of planetary missions and non-planetary missions were also tested with the 10 HVHDP criteria.

This significant difference between the planetary and non-planetary factor solutions and the criteria composing the factor solutions could be one of the causes of factor instability in the full data set.

The two subgroups of planetary missions and non-planetary missions were also tested with the 10 HVHDP criteria. Only one factor was common in a six factor solution, and approximately 40 percent of the criteria were different. The two subgroups of planetary missions and non-planetary missions were tested with the 20 HVHDP criteria and the 10 HVHDP criteria. The criteria composing the factor solutions and the factors were generally not similar. Since the planetary and non-planetary subgroups did not have similar factor structures, they may be the source of the factor structure instability in the full data set analysis.

Subgroup Analysis of Newer and Older CBR records

This research investigated whether there were significant changes in the data set over time. The full data set was split into two subgroups. One subgroup contained the most recent records from 2009-2018 and the second subgroup contained the older data from 1998-2007. Factor analysis was performed on the recent record set and on the older record set using the 20 HVHDP criteria, and the factor structure compared. The factor structure for the recent data (2009-2018) was generally not similar to the factor structure for the older data (1998-2007). Since the two subgroups had different factor structures, the two subgroups may be one of the causes of factor instability in the full set factor structure. This indicates that the ratings on the 20 HVHDP criteria may be changing over time. It was noted that criterion 2 was a single criterion factor across the two different time period factor solutions, across the full 20 HVHDP, set A, and set B factor solutions. Criterion two, as a single factor structure, is generalizable across the 20 HVHDP data set.

Analysis of Recent Planetary and Recent Non-Planetary Data

The subgroup analysis indicated there were differences both between planetary data and non-planetary data, and between newer versus the older records. A set of the most recent planetary data was analyzed, in order to investigate whether a recent data set would result in a valid factor structure for either planetary or non-planetary data. A new set of HVHDP criteria was developed from the recent planetary data and was used in factor analysis of the recent planetary data. This resulted in a factor structures that passed Bartlett's test of sphericity, Kaiser-Meyer-Olkin's (KMO) overall measure of sampling adequacy (MSA), and an individual criteria check for MSA, and that showed communalities that were acceptable. However, the factor structure for recent planetary data failed reliability measures of internal consistency, indicating that factors with multiple criteria were not consistently measuring the same construct. Since the criteria in the factor with multiple criteria were not consistently measuring the same construct, this provided justification to retain each of those criteria as a separate measure in the final set of recommended criteria.

A set of the most recent non-planetary data was also developed, in order to determine if the most recent non-planetary data would result in a valid factor structure. This resulted in a factor structures that passed Bartlett's test of sphericity, Kaiser-Meyer-Olkin's (KMO) overall measure of sampling adequacy (MSA), and an individual criteria check for MSA, and that showed communalities that were acceptable. However, the factor structure for recent non-planetary data failed reliability measures of internal consistency, indicating that factor structure was not reliable. Since the criteria in factors with multiple criteria were not reliable, then according to Tavakol and Dennick (2011),

the instrument could not be valid. This provided justification to retain each of those criteria as a separate measure in the final set of recommended criteria.

Analysis of 2016 Non-Planetary Data

The 2016 non-planetary data set was composed of the non-planetary data from 2014-2016 (2016 non-planetary). The 2016 non-planetary data was factored, using the original 23 criteria. A three factor solution was found that was reliable for the 2016 non-planetary data set. This three factor solution was tested with the larger recent non-planetary data set, but it was not generalizable for that larger data set with older records. This indicated that the factor structure changed over time for the non-planetary data from 2011 to 2016. This could reflect changes in the experience level of expert peer panel members or changes in the spaceflight discipline, over time.

Criteria Recommended for Deletion

The factor structures either indicated one criterion per factor or several criteria per factor. Most of the multi criteria factors found in the factor solutions consistently failed internal consistency measures of reliability, indicating that the multi criteria factor constructs were not reliable or valid. This was a valuable finding, since it substantiated that those criteria were unique criteria, and since it provides justification to continue to use those criteria separately.

The exception was that the 2016 multi criteria factors were found to be reliable. The 2016 non-planetary factor solution with reliable measures included criteria 5, 6, 11, 20, 23.

Table 54 lists the number of High and Very High Discriminatory Power (HVHDP) rating each criterion received for the non-planetary data set. Criteria 6 and 23 received no HVHDP ratings, and criteria 11 received only one. Although the 2016 non-

planetary solution was reliable, the criteria 6, 11, and 23 had none or one HVHDP and, consequently, were not recommended for inclusion in the recommended set of criteria.

The 20 HVHDP criteria were analyzed for how frequently there were included in one of the factor solutions in this paper. The percent of the time each of the 20 HVHDP criteria were included in the 12 different factor solutions for the 20 HVHDP or 10 HVHDP was calculated. 12 of the 13 20 HVHDP criteria were included in the 12 different factors solutions 50% of the time or more. Criteria 13 was included in the 12 different factor solutions only 13% of the time. However, criteria 13 was similar to criteria 19, and it was recommended that those two criteria be combined. Criteria that were outliers and that contributed very infrequently to the high and very high discriminatory power ratings were identified. These were recommended for deletion.

Criteria that frequently identified very negative findings (very high discriminatory power) or very positive finding (high discriminatory power) were identified and ranked. A subset of the criteria with the most frequent HVHDP ratings could be considered as a new smaller set to use in the mission concept phase, which could result in less time and money for proposers to prepare mission concepts and less time for expert peer reviews to review mission concept proposals. Based on this analysis, nine of the original 23 criteria were recommended for deletion. The 13 criteria to be retained were the 20 HVHDP criteria, which include criteria 1, 2, 3, 4, 5, 7, 10, 13, 14, 16, 17, 20, and 21. The summary of recommendations, with rationale, is presented in Table 66.

Table 66 Summary of Criteria Recommendations

Criteria	4	5	No. of High and Very High DP Ratings	% of Sample	Included in Factor Solutions	Recommend for Final Set	Notes
21	97	68	165	46%	67%	Yes	
20	67	59	126	35%	50%	Yes	
1	30	92	122	34%	83%	Yes	
3	42	62	104	29%	83%	Yes	
17	60	42	102	29%	92%	Yes	
2	61	32	93	26%	75%	Yes	
4	59	33	92	26%	50%	Yes	
5	17	32	49	14%	92%	Yes	
7	15	32	47	13%	58%	Yes	
10	6	26	32	9%	58%	Yes	
9	7	24	31	9%	Outlier	No	Only applies to Planetary
14	1	29	30	8%	83%	Yes	
13	17	4	21	6%	17%	Yes	Similar to criteria 19. Combine criteria 13 and 19 into one criteria
16	6	15	21	6%	83%	Yes	Combine 23 with 16
6	3	14	17	5%		No	This could be addressed by compliance with a requirments document and the assess in Phase A
19	2	10	12	3%		No	Combine with 13
22	5	7	12	3%		No	Assess this criteria in Phase A
12	0	11	11	3%		No	Low number of HVHDP.
23	6	5	11	3%		No	Low number of HVHDP. Can be assessed in Phase A
11	0	5	5	1%		No	Only five HVHDP ratings.
18	4	0	4	1%		No	Only 4 HVHDP ratings for criteria.
8	0	1	1	0%		No	Outlier: Received the third lowest point total of HVHDP. This should assessed in Phase A
15	0	0	0	0%		No	This criterion received no HVHDP ratings.

Factor Structure

This research produced 12 factor structures for the full data set for 20 HVHDP criteria and 10 HVHDP criteria. Although the factor solution passed measures of Bartlett's test of sphericity, Kaiser-Meyer-Olkin's (KMO) overall measure of sampling adequacy (MSA), individual MSA, and the total variance explained, the multi-criteria factors did not pass tests of internal consistency. Ford, MacCallum, and Tait (1986) stated that "Another problem with interpretation is that even when the factors appear to be clear and unambiguous, the factor structure may be unreliable because of sampling variability" (p. 96). In order to investigate the cause of failing internal consistency measures, subgroups were analyzed. The full data set was split by planetary and nonplanetary missions, and factor solutions were generated for each subgroup. The factor solutions for the planetary and nonplanetary data sets passed measures of Bartlett's test of sphericity, Kaiser-Meyer-Olkin's (KMO) overall measure of sampling adequacy (MSA), individual MSA, and the total variance explained. However, measures of internal consistency were not valid.

A second subgroup was analyzed, in order to determine if variations in the ratings of the criteria over time were a cause of factor structure instability. The full data set was split by newer (2009-2018) and older (1998-2007) records. Both the newer and the older data factor solutions passed measures for Bartlett's test of sphericity, Kaiser-Meyer-Olkin's (KMO) overall measure of sampling adequacy (MSA), sufficient communalities, and total variance extracted. However, both the newer and older data failed internal consistency measures.

Recent planetary data and recent non-planetary data were analyzed, since the factor solutions appeared to vary over time and by planetary or non-planetary records.

The factor solutions to the recent planetary and recent nonplanetary data sets passed measures of Bartlett's test of sphericity, Kaiser-Meyer-Olkin's (KMO) overall measure of sampling adequacy (MSA), individual MSA, and the total variance explained. However, measures of internal consistency were not reliable.

One possible explanation of the factor structure unreliability of the recent planetary and recent non-planetary data is that the ratings on criteria, whether planetary or non-planetary, continued to change over time. The recent planetary data set included data from 2006 forward. The recent non-planetary data set included data from 2011 forward. Each of these date ranges covers significant time periods where the ratings on criteria may be changing, due to changes in expectations based on events in space science mission development or due to changes in experience level of personnel serving on the expert review panels.

The recent planetary data set included 91 records and contained data from 2006 forward, and the recent non-planetary data set included 90 records and contained data from 2011 forward. Further analysis of data within the last five years is desirable. De Winter, Dodou, and Wieringa (2009) stated "Exploratory factor analysis (EFA) is generally regarded as a technique for large sample sizes (N), with N = 50 as a reasonable absolute minimum." (p. 147). According to Mundfrom, Shaw, and Ke (2005), "Suggested minimums for sample size include from 3 to 20 times the number of variables and absolute ranges from 100 to over 1,000" (p. 159). The number of records for the most recent five years in planetary is only 39, and there are 51 non-planetary records. Since there were 51 non-planetary records from 2014 to 2016, a non-planetary data set was created, named 2016 planetary. A three factor solution was found that was reliable

for the 2016 non-planetary data set, using the full 23 original criteria. However, the solution was not generalizable to a larger time period for non-planetary records.

No factor structure was recommended for the 13 remaining criteria, since none of the factor structures analyzed were reliable or valid, with the exception of the 2016 non-planetary three factor solution. The 2016 non-planetary three factor solution was not applicable to planetary missions, and the five criteria did not credibly address the scope of a spaceflight mission. In addition, three of the five criteria that comprise the three factor solution for the 2016 non-planetary data had 5% or less high or very high discriminatory power ratings and were not part of the recommended reduced set of criteria.

Planetary Dominance in HVHDP

The number of HVHDP ratings from planetary data was compared to the total number of HVHDP ratings for planetary and non-planetary data. This is shown in Table 67. Planetary missions are much more complex and challenging than non-planetary missions. This is reflected in the fact that the total of planetary HVHDP composes 65% of the total HVHDP.

Table 67 Planetary HVHDP ratings as a % of Total HVHDP

Criteria	Total HVHDP	Planetary	Planetary % of Total HVHDP
21	165	102	62%
20	126	70	56%
1	122	79	65%
3	104	66	63%
17	102	74	73%
2	93	58	62%
4	92	52	57%
5	49	29	59%
7	47	33	70%
10	32	17	53%
9	31	31	100%
14	30	22	73%
13	21	15	71%
16	21	18	86%
6	17	14	82%
19	12	4	33%
22	12	6	50%
12	11	11	100%
23	11	11	100%
11	5	4	80%
18	4	3	75%
8	1	0	0%
15	0	0	0%
Total	1108	719	65%

Research Contributions

The results of this research contribute to the current practice of expert peer reviews of spaceflight systems and mission design by identifying a reduced set of 13 criteria to assess flight systems and mission design for a wide range of space flight missions at the mission concept stage. The NASA Space Flight Program and Project Management Requirements require that concept studies demonstrate feasibility (NASA, 2012). This research recommends that 13 specific criteria be addressed, in order to establish the feasibility of a pre-Phase A concept study.

This research contributes to the knowledge of the feasibility criteria that must be addressed as part of the systems engineering process at the pre-Phase A mission concept stage. This is the only research on the criteria based on the past records of over 300 expert peer review panels that assessed pre-Phase A proposals to NASA space science competed mission programs. This research contributes to the knowledge of the feasibility criteria that must be addressed as part of the systems engineering process at the pre-Phase A mission concept stage; it identifies specific criteria that have high or very high discriminatory power in the overall rating of the mission concept by the expert peer panel and that are considered by decision makers at the mission concept stage.

The reduction of the current 23 criteria to 13 criteria will reduce the time and money spent on the expert peer review. It will also reduce the proposers' cost and the time needed to produce a proposal. Time and money freed up from this process can be utilized to provide more development time for selected missions. The reduction in the number of criteria and in the time needed to evaluate proposals would also lower the

barrier to propose a mission concept. This may result in additional proposals and in new scientific ideas to consider for selecting the next space science mission.

Since the research was based on a wide range of proposed spaceflight missions with no common target or goal, the 13 recommended criteria are *general* criteria that can apply to a wide range of proposed government and commercial spaceflight missions. The use of an expert peer panel in conjunction with the refined set of 13 general criteria will provide a cost-efficient method to assess a large number of diverse commercial space opportunities at the mission concept stage.

The risk criteria assessed by the expert peer panel are broad in scope and require the work of a group of engineers with different specializations within the scope of spacecraft systems and missions design and operations. The result is a qualitative rating and text that substantiates the rating on each criterion. This research demonstrates a method to analyze a qualitative product of a group expert judgment process. This method could be used to validate a measurement instrument or to refine a measurement instrument used by a group of experts. In particular, this research provides a method that engineering managers can use to analyze and to refine a qualitative measurement instrument for an assessment by a group of experts. This could be useful in assessments that require a broad scope of required expertise.

Analyzing records on which decisions were made in the past to select space science missions is an example of how analysis of data on which past decisions were made can contribute to better and more efficient decision making. There are applications in the government and in the commercial sector of analyzing past records of qualitative assessments to improve decision making in the future. This research contributes to the literature on the analysis of qualitative assessments by groups using expert judgment.

Factor analysis has been used extensively in the fields of social science, psychology, health, and medicine. This research adds to the literature on exploratory factor analysis by demonstrating that these methods can be used in a spaceflight engineering application. This research demonstrates how split set anaysis can be used to identify factors and criteria that are generalizable across a data set. This research also demonstrates how exploratory factors analysis can be used to identify subgroups within the full data set.

Study Limitations and Delimitations

This research uses records of expert engineering peer reviews of space science mission concepts at the pre-Phase A time period, which is very early in the development cycle. This research may not be applicable to assess spaceflight missions later in the development cycle.

Extension of Research

This research evaluates the criteria for two of the five factors of the TMC Feasibility of the Proposed Mission Implementation, Including Cost Risk (NASA, 2014a). All five factors are shown in Appendix A. This research could be extended to assess the criteria of the other three factors.

This research was conducted using the criteria for mission concept proposals in response to an AO. Selected mission concepts were funded by NASA to conduct a phase A study. Expanded criteria were used to assess the phase A study. Similar research could be conducted to analyze whether the expanded set of criteria used to evaluate space science mission Phase A studies are all necessary.

In addition, further analysis can be done to quantify the cost risk of the very high discriminatory power criteria that are recommended to be deleted. For each Very High

rating on a criterion, there is a paragraph that describes the major weakness which includes a likelihood and a consequence cost risk statement. The cost risk statements can be analyzed across the sample for the criteria to be deleted and the consequence of dropping the criteria recommended for deletion can be quantified.

REFERENCES

Barbeite, F. G., & Weiss, E. M. (2004). Computer self-efficacy and anxiety scales for an Internet sample: testing measurement equivalence of existing measures and development of new scales. *Computers in Human Behavior, 20*(1), 1-15. doi:10.1016/S0747-5632(03)00049-9

Burnett, P. C., & Dart, C. D. (1997). Conventional versus confirmatory factor analysis: methods for validating the structure of existing scales. *Journal of Research and Development n Education, 30*(2), 126 - 132.

Burston, A., Eley, R., Parker, D., & Tuckett, A. (2017). Validation of an instrument to measure moral distress within the Australian residential and community care environments. *International Journal of Older People Nursing, 12*(2). doi:10.1111/opn.12144

Byrne, B. M. (2016). *Structural equation modeling with Amos* (Third Edition ed.). 711 Third Avenue, New York, NY 10017: Routledge.

Chytka, T. M., Conway, B. A., & Unal, R. (2006, July). An expert judgment approach for addressing uncertainty in high technology system design. *Portland International Conference On Management of Engineering and Technology*. 1: 444-449. DOI: 10.1109/PICMET.2006.296590

Clemen, R., & Winkler, R. (2006). Combining probability distributions from experts in risk analysis. *Risk Analysis, 19*(2), 187-203. doi:10.1111/j.1539-6924.1999.tb00399.x

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (revised edition ed.). New York: Academic Press.

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods, 6*(2), 147-168. doi:10.1177/1094428103251541

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 9.

Creswell, J. W. (2018). *Research design qualitative, quantitative, and mixed methods pproaches* (Fifth Edition ed.). 2455 Teller Road Thousand Oaks, California 91320: SAGE Publications Inc.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Daniels, C. L. (2007, November). *Technical, management, and cost (TMC) evaluation.* Presented at the Small Explorers (SMEX) and Missions of Opportunity Preproposal Conference. Retrieved from http://explorers.larc.nasa.gov/PDF_FILES/DANIELS-EVALUATION.pdf

de Winter, J. C. F., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research, 44*(2), 147-181. doi:10.1080/00273170902794206

DeVellis, R. F. (1991). *Scale development: Theory and Applications*. Newbury Park, CA: Sage.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299. doi:doi:10.1037/1082-989X.4.3.272

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory Factor Analysis*. 198 Madison Avenue, New York, New York 10016: Oxford University Press Inc.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. 2455 Teller Road Thousand Oaks, CA 91320: SAGE Publications Inc.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 286-299.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: a critical review and analysis. Personnel Psychology, 39(2), 291-314. doi:10.1111/j.1744-6570.1986.tb00583.x

Fricker, R. D. J., Kulzy, W. W., & Appleget, J. A. (2012). From data to informaition: using factor analysis with survey data. *Phalanx*, *45*:4, 30-34. http://hdl.handle.net/10945/38733

Goossens, L. H. J., Cooke, R. M., Hale, A. R., & Rodić-Wiersma, L. (2008). Fifteen years of expert judgement at TUDelft. *Safety Science, 46*(2), 234-244. doi:http://dx.doi.org/10.1016/j.ssci.2007.03.002

Gorsuch, R. L. (2015). *Factor analysis: classic edition*. 711 Third Avenue, New York, NY 10017: Routledge.

Griffin, M. D., & French, J. R. (2004). *Space vehicle design*: American Institute of Aeronautics and Astronautics, Inc., Reston Virginia.

Hair, J. F., Anderson, R. E., Tatham, R. L. & Black., W. C. (1995). *Multivariate data analysis and readings* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Hair, J. F., Anderson, R. E., Babin, J. B., & Black, W. C. (2010). *Multivariate data analysis* (7 ed.). Upper Saddle River, NJ 07458: Prentice Hall.

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educational and Psychological Measurement, 66*(3), 393-416. doi:10.1177/0013164405282485

Hsu, C. (2012). Evaluation criteria for blog design and analysis of causal relationships using factor analysis and DEMATEL. *Expert Systems with Applications, 39*(1), 187-193. doi:https://doi.org/10.1016/j.eswa.2011.07.006

Hutcheson, G., & Sofroniou, J. (1999). *The multivariate social scientist*. Thousand Oaks, CA 91320: Sage Publications.

Johnson, B., & Stevens, J. J. (2001). Exploratory and confirmatory factor analysis of the school level environment questionnaire (SLEQ). *Learning Environments Research, 4*(3), 325. doi:10.1023/a:1014486821714

Jones, J., & Hunter, D. (1995). Consensus methods for medical and health services research. *BMJ : British Medical Journal, 311*(7001), 376-380.

Kaiser, J. F. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31-36.

Kakai, H. (2003). Re-examining the factor structure of the California critical thinking disposition inventory. *Perceptual and Motor Skills, 96*, 435-483.

Kline, P. (1994). *An Easy Guide to Factor Analysis*. New York, NY: Routledge.

Leedy, P. D., & Ormrod, J. E. (1993). *Practical research planning and design*. Upper Saddle River, New Jersey 07458: Pearson Education Inc.

Lucas, V. (2017). *Technical, management, and cost (TMC) evaluation overview.* Presented at the New Frontiers 4 Announcement of Opportunity Preproposal Conference. Retrieved from Retrieved from https://newfrontiers.larc.nasa.gov/PDF_FILES/06-Lucas-TMCEvaluation.pdf

MacCallum, R. C., Roznowski, M., Mar, C. M., & Reith, J. V. (1994). Alternative Strategies for Cross-Validation of Covariance Structure Models AU. *Multivariate Behavioral Research, 29*(1), 1-32. doi:10.1207/s15327906mbr2901_1

Matthews, R. B., Coomes, E. P., & Khan, E. U. (1994). Hierarchical analysis of options for lunar-surface power. *Journal of Propulsion and Power, 10*(3), 425-431. doi:10.2514/3.23752

Mohamed, A. M., Amer, F. E. H., Mostafa, R. M., & Mahmoud, A. A. H. (2016, March). *Trade-off analysis of low earth orbit spacecraft power supply system by genetic algorithm.* Paper presented at the 2016 IEEE Aerospace Conference. doi:10.1109/aero.2016.7500851

Monroe, R., Lepsch, R., & Unal, R. (2002). Using expert judgement methodology to address uncertainty in launch vehicle weight estimates. In 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization: American Institute of Aeronautics and Astronautics.

Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences of the United States of America, 111*(20), 7176-7184. doi:10.1073/pnas.1319946111

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum Sample Size Recommendations for Conducting Factor Analyses. *International Journal of Testing, 5*(2), 159-168. doi:10.1207/s15327574ijt0502_4

Nair, R., Aggarwal, R., & Khanna, D. (2011). methods of formal consensus in classification/diagnostic criteria and guideline development. *Seminars in Arthritis and Rheumatism, 41*(2), 95-105. doi:https://doi.org/10.1016/j.semarthrit.2010.12.001

NASA. (2009). *Announcement of Opportunity New Frontiers 2009*. Retrieved from https://nspires.nasaprs.com/external/viewrepositorydocument/cmdocumentid=182696/solicitationId=%7BCFE00B4A-2447-7325-7004-7C3B04ED5738%7D/viewSolicitationDocument=1/New_Frontiers_AO_admin1.pdf

NASA. (2010). *Announcement of Opportunity Discovery 2010*. Retrieved from https://nspires.nasaprs.com/external/viewrepositorydocument/cmdocumentid=234098/solicitationId=%7B70D505D2-9B11-5907-029A-89AD44684EBD%7D/viewSolicitationDocument=1/DSC10%20Amendment%203.pdf

NASA. (2012). *NASA Space Flight Program and Project Management Requirements NPR 7120.5E*. Retrieved from https://nodis3.gsfc.nasa.gov/npg_img/N_PR_7120_005E_/N_PR_7120_005E_.pdf

NASA. (2014a). *Annoucement of Opportunity Standard PI-Led Mission AO*. Retrieved from https://soma.larc.nasa.gov/standardao/pdf_files/StandardAOTemplate140613.pdf

NASA. (2014b). *Announcement of Opportunity Discovery 2014*. Retrieved from https://nspires.nasaprs.com/external/viewrepositorydocument/cmdocumentid=438340/solicitationId=%7BFE7B4C63-873D-63C1-4D15-1D46E2FEA949%7D/viewSolicitationDocument=1/discovery-2014-amend1.pdf

NASA. (2014c). *Strategic Plan*. Retrieved from https://www.nasa.gov/sites/default/files/files/FY2014_NASA_SP_508c.pdf

NASA. (2015a). *Announcement of Opportunity 2014 Astrophysics Small Explorer (SMEX) Full Missions Evaluation Plan*. Retrieved from Retrieved from https://explorers.larc.nasa.gov/APSMEX/SMEX/pdf_files/Astro-SMEX-2015-01-14-V05-public-version.pdf

NASA. (2015b). *Announcement of Opportunity Earth Venture Mission - 2*. Retrieved from https://nspires.nasaprs.com/external/solicitations/summary.do?method=init&solId=%7B570984B5-0AF8-0145-6E4A-9B9C25A338CC%7D&path=open

NASA. (2016a). *Announcement of Opportunity Astrophysics Explorers Program 2016 Medium Explorer (MIDEX)*. Retrieved from https://nspires.nasaprs.com/external/viewrepositorydocument/cmdocumentid=535096/solicitationId=%7B8FB8BF8F-8248-8FB1-DDC5-54E7AE878E5A%7D/viewSolicitationDocument=1/2016%20Astro%20MIDEX%20AO_Amend2.pdf

NASA. (2018). *Third Stand Alone Missions of Opporuntiy Notice (SALMON-3) Program Element Appendix (PEA J) Small Innovative Missions for Planetary Exploration (SIMPLEx)*. Retrieved from https://nspires.nasaprs.com/external/viewrepositorydocument/cmdocumentid=624244/solicitationId=%7B523A5AF4-07C1-1FA4-0734-9C0253A172DF%7D/viewSolicitationDocument=1/PEA%20J%20SIMPLEx_amend3.pdf

Nunnally, J. C., Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.

Osborne, J. W. (2015). What is rotating in exploratory factor analysis. *Practical Assessment, Research & Evaluation, 20*(2), 1-7.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. 270 Madison Avenue New York, NY 10016: Psychology Press Taylor and Francis Group, LLC.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis*. Thousand Oaks, California 91320: Sage Publications.

Pisacane, V. L. (2005). *Fundamentals of Space Systems*. 198 Madison Avenue, New York, New York 10016: Oxford University Press, Inc.

Saldana, J. (2016). *The coding manual for qualitative researchers* (3rd Ed.). 2455 Teller Road Thousand Oaks, California 91320: Sage Publications Inc.

Santos, J. R. A., & Clegg, M. D. (1999). Factor analysis adds new dimension to extension surveys. *Journal of Extension, 37*(5), 5RIB6.

Steinskog, D., Tjøstheim, D., & Kvamstø, N. (2007). A cautionary note on the use of the Kolmogorov-Smirnov test for normality. *Monthly Weather Review, 135*(3), 1151-1154,1156-1157. doi:10.1175/MWR3326.1

Suhr, D. (2006). *Exploratory or Confirmatory Factor Analysis*. Paper presented at the SAS Users Group International Conference.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. One Lake Street, Upper Saddle River, New Jersey 0758: Pearson.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education, 2*, 53-55. doi:10.5116/ijme.4dfb.8dfd

Tavakol, M., Mohagheghi, M. A., & Dennick, R. (2008). Assessing the skills of surgical residents using simulation. *Journal of Surgical Education, 65*(2), 77-83.

Tran, D. Q., & Molenaar, K. R. (2014). Impact of risk on design-build selection for highway design and construction projects. *Journal of Management in Engineering, 30*(2), 153-162. doi:doi:10.1061/(ASCE)ME.1943-5479.0000210

Tsai, J., Pietrzak, R. H., Southwick, S. M., & Harpaz-Rotem, I. (2011). Examining the dimensionality of combat-related posttraumatic stress and depressive symptoms in treatment-seeking OEF/OIF/OND veterans. *Journal of affective disorders, 135*(1-3), 310. doi:10.1016/j.jad.2011.06.057

Tzeng, G. H., Chiang, C. H., & Li, C. W. (2007). Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL. *Expert Systems with Applications, 32*(4), 1028-1044. doi:https://doi.org/10.1016/j.eswa.2006.02.004

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis - some further observations. *Multivariate Behavioral Research, 25 (1)*, 89-95.

Walsh, C. M., Hardy, R. C. (1997). Factor structure stability of the Califiornia critial thinking disposition inventory across sex and various student majors. *Perceptual and Motor Skills, 85*, 1211-1228.

Walsh, C. M., Seldomridge, L. A., & Badros, K. K. (2007). California critical thinking disposition inventory: further factor analytic examination. *Perceptual and Motor Skills, 104*(1), 141-151. doi:10.2466/pms.104.1.141-151

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *2010, 8*(3). doi:10.33151/ajp.8.3.93

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology, 9*(2), 79-94.

APPENDIX A

Standard Evaluation Criteria

Standard Announcement of Opportunity Template dated June 13, 2014
https://soma.larc.nasa.gov/standardao/pdf_files/StandardAOTemplate140613.pdf

7.2.4 TMC Feasibility of the Proposed Mission Implementation, Including Cost Risk

The technical and management approaches of all submitted investigations will be evaluated to assess the likelihood that they can be successfully implemented as proposed, including an assessment of the likelihood of their completion within the proposed cost and schedule. The factors for feasibility of mission implementation include the following:

• Factor C-1. Adequacy and robustness of the instrument implementation plan. The maturity and technical readiness of the instrument complement will be assessed, as will the ability of the instruments to meet mission requirements. This factor includes an assessment of the instrument design, accommodation, interface, heritage, and technology readiness. This factor includes an assessment of the instrument hardware and software designs, heritage, and margins. This factor includes an assessment of the proposer's understanding of the processes, products, and activities required to accomplish development and integration of the instrument complement. This factor also includes adequacy of the plans for instrument systems engineering and for dealing with environmental concerns. This factor includes an assessment of plans for the development and use of new instrument technology, plans for advanced engineering developments, and the adequacy of backup plans to mature systems within the proposed cost and schedule when systems having a TRL less than 6 are proposed.

• Factor C-2. Adequacy and robustness of the mission design and plan for mission operations. This factor includes an assessment of the overall mission design and mission architecture, the spacecraft design and design margins (including margins for launch mass, delta-V, and propellant), the concept for mission operations (including communication, navigation/tracking/trajectory analysis, and ground systems and facilities), and the plans for launch services. This factor includes mission resiliency – the flexibility to recover from problems during both development and operations – including the technical resource reserves and margins, system and subsystem redundancy, and reductions and other changes that can be implemented without impact to the Baseline Science Mission.

• Factor C-3. Adequacy and robustness of the flight systems. This factor includes an assessment of the flight hardware and software designs, heritage, and margins. This factor includes an assessment of the proposer's understanding of the processes, products, and activities required to accomplish development and integration of all elements (flight systems, ground and data systems, etc.). This factor includes an assessment of the adequacy of the plans for spacecraft systems engineering, qualification, verification, mission assurance, launch operations, and entry/descent/landing. This factor includes the

plans for the development and use of new technology, plans for advanced engineering developments, and the adequacy of backup plans to ensure success of the mission when systems having a TRL less than 6 are proposed. The maturity and technical readiness of the spacecraft, subsystems, and operations systems will be assessed. The adequacy of the plan to mature systems within the proposed cost and schedule, the robustness of those plans, including recognition of risks and mitigation plans for retiring those risks, and the likelihood of success in developing any new technologies will be assessed.

• Factor C-4. Adequacy and robustness of the management approach and schedule, including the capability of the management team. This factor includes: the adequacy of the proposed organizational structure and WBS; the management approach including project level systems engineering; the roles, qualifications, and experience of the PI, PM, other named Key Management Team members, and implementing organization, mission management team, and known partners; the commitment, spaceflight experience, and relevant performance of the PI, PM, other named Key Management Team members, and implementing organization, mission management team, and known partners against the needs of the investigation; the commitments of partners and contributors; and the team's understanding of the scope of work covering all elements of the mission, including contributions. Also evaluated under this factor is the adequacy of the proposed risk management approach, including any risk mitigation plans for new technologies, any long-lead items, and the adequacy and availability of any required manufacturing, test, or other facilities. The approach to any proposed descoping of mission capabilities will be assessed against the proposed Baseline Science Mission. The plans for managing the risk of contributed critical goods and services will be assessed, including the plans for any international participation, the commitment of partners and contributors, as documented in Letters of Commitment, and the technical adequacy of contingency plans, where they exist, for coping with the failure of a proposed cooperative arrangement or contribution. This factor also includes assessment of proposal elements such as the relationship of the work to the project schedule, the project element interdependencies, the associated schedule margins, and an assessment of the likelihood of launching by the proposed launch date. Also evaluated under this factor are the proposed project and schedule management tools to be used on the project along with the subcontracting plan, including small and small disadvantaged businesses.

• Factor C-5. Adequacy and robustness of the cost plan, including cost feasibility and cost risk. This factor includes proposal elements such as cost, cost risk, cost realism, and cost completeness including assessment of the basis of estimate, the adequacy of the approach, the methods and rationale used to develop the estimated cost, the discussion of cost risks, the allocation of cost reserves by phase, and the team's understanding of the scope of work (covering all elements of the mission, including contributions). Proposals will be evaluated for the adequacy of the cost reserves and whether proposals with inadequate cost reserves demonstrate a thorough understanding of the cost risks. This factor also includes an assessment of the proposed cost relative to estimates generated using parametric models and analogies. Also evaluated under this factor are the proposed cost management tools to be used on the project.

APPENDIX B

Standard Evaluation Criteria for Flight Systems and Mission Design and Operations

Factor C-3 Flight Systems: Adequacy and robustness of the flight systems

Criteria 1: This factor includes an assessment of the flight hardware and software designs

Criteria 2: This factor includes an assessment of the flight hardware and software heritage

Criteria 3: This factor includes an assessment of the flight hardware and software margins

Criteria 4: This factor includes an assessment of the proposer's understanding of the processes, products, and activities required to accomplish development and integration of all elements (flight systems, ground and data systems, etc.)

Criteria 5: This factor includes an assessment of the adequacy of the plans for spacecraft systems engineering

Criteria 6: This factor includes an assessment of the adequacy of the plans for spacecraft qualification and verification.

Criteria 7: This factor includes an assessment of the adequacy of the plans for spacecraft mission assurance

Criteria 8: This factor includes an assessment of the adequacy of the plans for launch operations

Criteria 9: This factor includes an assessment of the adequacy of the plans for entry/descent/landing

Criteria 10: This factor includes the plans for the development and use of new technology

Criteria 11: This factor includes the plans for advanced engineering developments

Criteria 12: This factor includes the adequacy of backup plans to ensure success of the mission when systems having a TRL less than 6 are proposed

Criteria 13: The maturity and technical readiness of the spacecraft will be assessed

Criteria 14: The maturity and technical readiness of the subsystems will be assessed

Criteria 15: The maturity and technical readiness of the operations systems will be assessed

Criteria 16: The adequacy of the plan to mature systems within the proposed cost and schedule, the robustness of those plans, including recognition of risks and mitigation plans for retiring those risks, and the likelihood of success in developing any new technologies will be assessed

Factor C-2 Mission Design and Operations: Adequacy and robustness of the mission design and plan for mission operation

Criteria 17: This factor includes an assessment of the overall mission design

Criteria 18: This factor includes an assessment of the overall mission architecture

Criteria 19: This factor includes an assessment of the spacecraft design

Criteria 20: This factor includes an assessment of the design margins (including margins for launch mass, delta-V, and propellant)

Criteria 21: This factor includes an assessment of the concept for mission operations (including communication, navigation/tracking/trajectory analysis, and ground systems and facilities)

Criteria 22: This factor includes an assessment of the plans for launch services

Criteria 23: This factor includes mission resiliency – the flexibility to recover from problems during both development and operations – including the technical resource reserves and margins, system and subsystem redundancy, and reductions and other changes that can be implemented without impact to the Baseline Science Mission

APPENDIX C

Criteria Topics

Table C1: Topics in Criteria

	Criteria Key Words	Supporting Topics
Criteria 1	Fight H/W and S/W Designs	Lack of information on design. Missing information to drive design. Conflicting information on the design.
Criteria 2	Fight H/W and S/W Heritage	
Criteria 3	Fight H/W and S/W Margins	
Criteria 4	Understanding of: Activities to accomplish development and integraton of all elements	Development plans. AI&T. ATLO. Integration and Test. Sparing Philosophy. Environmental testing.
Criteria 5	Spacecraft Systems Engineering	Phase A trade study plans. Analysis not complete/incomplete to drive design. Requirements not specified or are conflicting. MEL. Inconsistancy in text and drawings. Requirements flow down and traceability.
Criteria 6	Spacecraft Qualification and verification	Meet specifications and required performance. Testing plans for Qual and verification. Test as you fly. NASA IV&V for S/W. Acceptance testing.
Criteria 7	Plans for Spacecraft Mission Assurance Plans	Spacecraft reliability. Redundancy. Single string. Grade of parts. Dust Issues and mitiations. Plans for contamination control. Magnetic Cleanliness plans. Parts radiation hardening. Radiation shielding. Parts burn-in. Fault tolerance. Critical events coverage. Meet orbital debris requirments at end of life. Orbital debris. Critical events req.
Criteria 8	Adequacy of plans for Lauch Operations	
Criteria 9	Plans for Entry, Descent and Landing (EDL)	Including parachutes, probes and Sample Return Capsule (SRC)
Criteria 10	Plans for the development and use of NEW technology	Disagree with TRL stated in proposal. TRL not described or defined.
Criteria 11	Advanced engineering development	
Criteria 12	Adequacy of backup plans for systems below TRL 6.	Overstating TRL level. Panel assesses it below 6.
Criteria 13	Maturity and tech readiness of the spacecraft.	
Criteria 14	Maturity and tech readiness of the subsystems.	ie. New comet sampling system, FSW, etc.
Criteria 15	Maturity and tech readiness of the operations systems.	If words "maturity and technical readiness" were used than put in here.
Criteria 16	The adequacy of the plan to mature systems within the proposed cost and schedule, the robustness of those plans, including recognition of risks and mitigation plans for retiring those risks, and the likelihood of success in developing any new technologies will be assessed	Back up plans. Descope options.
Criteria 17	Overall mission design	Trajectory design is mature or underdeveloped. Including proximity operations. Touch and Go (TAG) Balloons, Airplanes.
Criteria 18	Overall mission architecture	Trajectory, comm, other mission assets. Robust margins.
Criteria 19	Assessment of the spacecraft design	To accomplish the mission design and operations. i.e. onboard data storage to make ops/telecom DL more flexible and resilient to missed passes.
Criteria 20	Assessment of the overall design margins (including margins for launch mass, delta-V, and propellant)	Include schedule margin.Launch schedule uncertainty or flexibility. Launch date flexibility or constraints.
Criteria 21	Concept for mission operations (including communication, navigation/tracking/trajectory analysis, and ground systems and facilities)	Insufficient information on items such as deployment trajectories. Communication ground and link margin issues. Operations team readiness and experience. Include impact of launch dispersions on operations and mission planning.
Criteria 22	Plans for launch services	Compatible with proposed Launch service.
Criteria 23	This factor includes mission resiliency – the flexibility to recover from problems during both development and operations – including the technical resource reserves and margins, system and subsystem redundancy, and reductions and other changes that can be implemented without impact to the Baseline Science Mission	"Development" overlaps with C16. I put development problems in C16 and operations problems in C23. However, also considered where finding was in Form C (in Flight Systems or MD&Os section) to guide focus of finding.

APPENDIX D

Space Science Mission Programs

Mission Program	Science Disciple	Cost Cap
Discovery Program 1998	Planetary Science	$299 Million in FY 1999 dollars
Discovery Program 2000	Planetary Science	$299 Million in FY 2001 dollars
Discovery Program 2004	Planetary Science	$360 Million in FY 2004 dollars
Discovery Program 2006	Planetary Science	$425 Million in FY 2006 dollars
Discovery Program 2010	Planetary Science	$425 Million in FY 2010 dollars
Discovery Program 2014	Planetary Science	$450 Million in FY 2015 dollars
Mars Scout 2002	Planetary Science	$325 Million in FY 2003 dollars
Mars Scout 2006	Planetary Science	$475 Million in FY 2006 dollars
New Frontiers 2003	Planetary Science	$700 Million in FY 2003 dollars
New Frontiers 2009	Planetary Science	$650 Million in FY 2009 dollars
New Frontiers 2017	Planetary Science	$850 Million in FY 2015 dollars
Medium Explorer (MIDEX) 2001	Astrophysics and Heliophysics	$180 Million in FY 2002 dollars
Small Explorer (SMEX) 2003	Astrophysics and Heliophysics	$120 Million in FY 2003 dollars
Explorers 2007	Astrophysics and Heliophysics	$105 Million in FY 2008 dollars
Explorers 2011	Astrophysics and Heliophysics	$200 Million in FY 2011 dollars
Astrophysics Small Explorers 2014	Astrophysics	$175 Million in FY 2015 dollars
Astrophysics Medium Explorers 2016	Astrophysics	$250 Million in FY 2017 dollars
Heliophysics Small Explorers 2016	Heliophysics	$165 Million in FY 2017 dollars
Earth Venture Mission 1 2011	Earth Science	$150 Million in FY 2014 dollars,
Earth Venture Mission 2 2015	Earth Science	$166 Million in FY 2018 dollars

APPENDIX E

Data Analysis

Table E.1 MSA and Internal Consistency for Current Two Factor Structure

	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
FULL SET CURRENT 2 FACTOR STRUCTURE									
Current 2 Factor Structure	Pass .652	Pass 0.000	Criteria 8, 11, and 13 Fail this test since their values are below .5.	All criteria fail	Fail 18.5%				
Flight Systems Factor: criteria 1 - 16						Fail 0.524	All Fail	All Fail	All Fail
Mission Design Factor: criteria 17-23						Fail 0.073	All Fail	All Fail	Fail

Table E.2

MSA and Internal Consistency for 6 Factor 20 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indivial MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less	Item-to-total correlation must be	Inter-item correlation must be >.30
FULL SET 20 HVHDP									
20 HVHDP Criteria 6 Factor	Pass .646	Pass 0.000	Pass	Pass	Pass 76.1%				
Factor 1: criteria 5, 7, 14						Fail .481	All Fail	All Fail	All Fail C5 and C7 .289
Factor 2: criteria 1, 17						Fail .267	NA	Fail .154	Fail .154

Table E.3

MSA and Internal Consistency for Set A and B of 20 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
RANDOM SET A VS. B 20 HVHDP									
20 HVHDP Random Set A 6 Factor									
Set A 6 Factor	Pass .575	Pass 0.000	Pass	Pass	Pass 78.3%				
Factor 1: criteria 3, 16						Fail .307	NA	Fail .182	Fail .183
Factor 2: criteria 14, 7						Fail .441	Fail	Fail .283	Fail .283
Factor 3: criteria 1, 17						Fail .229	NA	Fail .129	Fail .129
Factor 4: 17, 5						Fail -.067	NA	Fail -.032	Fail -.032
Factor 5: criteria: 17, 2						Fail .201	NA	Fail .112	Fail .112
20 HVHDP Random Set B 6 Factor	Pass .596	Pass 0.000	Pass	Pass	Pass 82.3%				
Factor 1: 5, 7						Fail .446	NA	Fail .288	Fail .288
Factor 2: 16, 17						Fail .220	NA	Fail .140	Fail .140
Factor 3: 17, 10						Fail .267	NA	Fail .155	Fail .155

Table E.4

MSA and Internal Consistency for Seven Factor 10 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
FULL SET 10 HVHDP									
10 HVHDP Criteria 7 Factor	Pass .650	Pass 0.000	Pass	Pass	Pass 62.36%				
Factor 1: criteria 3, 4, 23						Fail .289	All Fail	All Fail	All Fail
Factor 2: criteria 12, 16						Fail .367	NA	Fail .233	Fail .233
Factor 3: criteria 5, 19						Fail -.519	NA	Fail -.215	Fail -.215
Factor 4: criteria 2, 6						Fail .243	NA	Fail .143	Fail .143
Factor 6: criteria 1, 17						Fail .267	NA	Fail .154	Fail .154
Factor 7: 20, 21						Fail .125	NA	Fail .074	Fail .074

Table E.5

MSA and Internal Consistency for Set A and B of 10 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
RANDOM SET A VS. B 10 HVHDP									
10 HVHDP Random Set A 7 Factor	Pass .618	Pass 0.000	Pass	Pass	Pass 67.83%				
Factor 1: criteria 4, 23, 21						Fail .381	All Fail	All Fail	All Fail
Factor 2: criteria: 12, 16, 21						Fail .391	All Fail	All Fail	All Fail
Factor 3: criteria: 5, 19						Fail -.539	NA	Fail -.222	Fail -.222
Factor 4: criteria: 2, 6						Fail .261	NA	Fail .160	Fail .160
Factor 5: criteria 14, 19						Fail 0.299	NA	Fail .183	Fail .183
Factor 6: criteria: 1, 17						Fail .229	NA	Fail .129	Fail .129
10 HVHDP RANDOM Set B 7 Factor	Pass .607	Pass 0.000	Pass	Pass	Pass 71.578				
Factor 1: criteria 5, 14, 19						Fail -.025	Fail	Fail	Fail
Factor 2: criteria 2, 23						Fail .352	NA	Fail .241	Fail .241
Factor 3: 14, 21						Fail .342	NA	Fail .213	Fail .213
Factor 4: criteria 1, 17						Fail .304	NA	Fail .180	Fail .180
Factor 5: criteria 4, 6						Fail -.166	NA	Fail -.077	Fail -.077

Table E.6

MSA and Internal Consistency for Planetary and Non-Planetary using 20 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
Planetary VS Non-Planetary 20 HVHDP									
Planetary 20 HVHDP	Pass .586	Pass 0.027	Pass	Pass	71.00%				
Factor 1: 7, 14						Fail .363	NA	Fail .222	Fail .222
Factor 2: 3, 20						Fail -.300	NA	Fail -.131	Fail -.131
Factor 3: 10, 16						Fail -.182	NA	Fail -.084	Fail -.084
Non Planetary 20 HVHDP	Pass 0.565	Pass 0.000	Pass	Pass	68%				
Factor 1: 5, 7, 14						Fail .528	Fail	Fail	**.373 for C5 and C7, others Fail**
Factor 2: 10, 17						Fail .416	NA	Fail .264	Fail .264
Factor 3: 2, 13						Fail -.227	NA	Fail -.133	Fail -.133
Factor 4: 1, 2						Fail -.136	NA	Fail -.064	Fail -.064

Table E.7

MSA and Internal Consistency for Planetary and Non-Planetary using 10 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	alpha if item deleted. If less than .6 fail	correlation must be >.50	Inter-item correlation must be >.30
Planetary VS Non-Planetary 10 HVHDP									
Planetary 10 HVHDP	Pass .553	Pass 0.002	Pass	Pass	Pass 71.161%				
Factor 1: 6, 14						Fail .387	NA	Fail .241	Fail .241
Factor 2: 10, 23						Fail .193	NA	Fail .109	Fail .109
Factor 4: 12, 16						Fail .345	NA	Fail .211	Fail .211
Factor 5: 3, 20						Fail -.300	NA	Fail -.131	Fail -.131
Non Planetary 10 HVHDP	Pass .666	Pass 0.000	Pass	Pass	Pass 80.5%				
Factor 1: 10, 19						Fail .566	NA	Fail .397	Pass .397
Factor 2: 5, 7						Fail .543	NA	Fail .373	Pass .373
Factor 3: 4, 7, 23						Fail .518	Fail	Fail	Pass .303 for C4 and C7. Pass .350 for C7 and C23. Fail C4 and C23
Factor 4: 16, 23						Fail .300	NA	Fail .179	Fail .179

Table E.8

MSA and Internal Consistency for Newer and Older Records using 20 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indivual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
Newest CBR VS. Oldest 20 HVHDP									
New 2009-2018 20 HVHDP	0.589	0.003	Pass	Pass	71.30%				
Factor 1: 4, 10, 17						Fail .368	All Fail	All Fail	All Fail
Factor 2: 5, 20						Fail .402	NA	Fail .253	Fail .253
Factor 3: 14, 21						Fail .251	NA	Fail .145	Fail .145
Old 1998-2007	Pass .600	Pass 0.000	Pass	Pass	Pass 71%				
Factor 1: 5, 7, 14						Fail .465	All Fail	All Fail	All Fail. .225 C5 and C7
Factor 2: 10, 20						Fail .184	NA	Fail .134	Fail .134
Factor 3: 16, 17						Fail .289	NA	Fail .170	Fail .170

Table E.9

Inter Item Correlation Difference between Planetary and Non Planetary

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
FULL SET 20 HVHDP									
20 HVHDP Criteria 6 Factor	Pass .646	Pass 0.000	Pass	Pass	Pass 76.1%				
Factor 1: criteria 5, 7, 14						Fail .481	All Fail	All Fail	All Fail C5 and C7 .289
20 HVHDP Random Set B 6 Factor	Pass .596	Pass 0.000	Pass	Pass	Pass 82.3%				
Factor 1: 5, 7						Fail .446	NA	Fail .288	Fail .288
Non Planetary 20 HVHDP	Pass 0.565	Pass 0.000	Pass	Pass	68%				
Factor 1: 5, 7, 14						Fail .528	Fail	Fail	**.373 for C5 and C7, others Fail**
Non Planetary 10 HVHDP	Pass .666	Pass 0.000	Pass	Pass	Pass 80.5%				
Factor 2: 5, 7						Fail .543	NA	Fail .373	**Pass .373**
Old 1998-2007	Pass .600	Pass 0.000	Pass	Pass	Pass 71%				
Factor 1: 5, 7, 14						Fail .465	All Fail	All Fail	All Fail .225 C5 and C7

Table E.10

MSA and Internal Consistency for Recent Planetary 4 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
Recent Planetary (2006 and later without Mars 2006)									
Recent Planetary 4 HVHDP	Pass .679	Pass 0.019	Pass	Pass	Pass 73.4%				
Factor 1: 4, 9						Fail .397	NA	Fail .254	Fail .254
Factor 2: 7, 14						Fail .349	NA	Fail .214	Fail .214
Factor 3: 1, 10, 21						Fail .257	All Fail	All Fail	All Fail
Factor 4: 5, 10						Fail .408	NA	Fail .259	Fail .259

Table E.11

MSA and Internal Consistency for Recent Non-planetary 4 HVHDP

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
Recent Non-planetary	Pass .637	Pass 0.14	Pass	Pass	Pass 67.2				
Factor 1: 10, 19						Fail .525	NA	Fail .361	Pass 10, 19
Factor 2: 1, 17						Fail .426	NA	Fail .271	Fail .271

Table E.12

MSA and Internal Consistency for 2016 Non-planetary

						Reliability			
	KMO >.5 OK	Bartlett's test of sphericity less than 0.05 OK	Indiviual MSA <.5 fail	Communality less the .5 fail	Total Variance over 60%?	Cronbach's alpha Fail if less than .6	Cronbach's alpha if item deleted. If less than .6 fail	Item-to-total correlation must be >.50	Inter-item correlation must be >.30
2016 Non-Planetary	Pass .543	Pass 0.000	Pass	Pass	83%				
Factor 1: 5, 11						Pass 0.652	NA	Pass 0.545	Pass 0.545
Factor 2: 6, 23						Pass 0.645	NA	Pass 0.543	Pass 0.543

APPENDIX F

DATA CHARACTERISTICS

Figure F.1

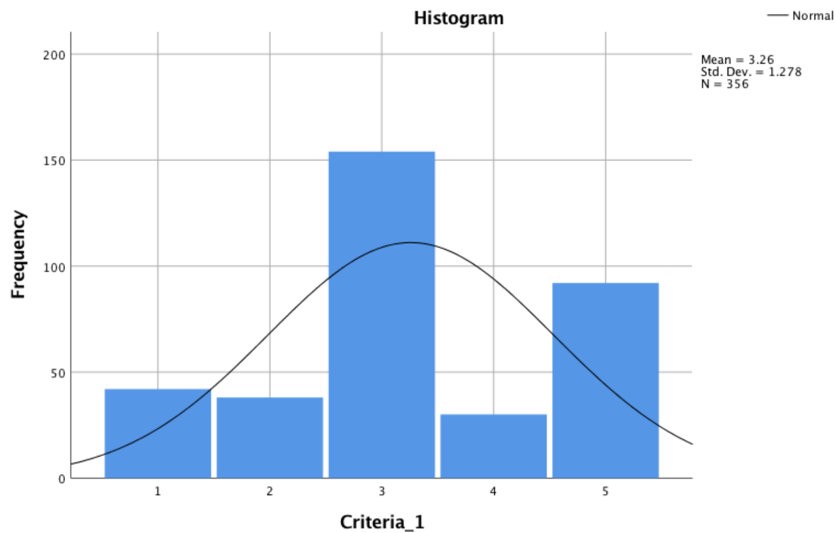


Figure F.2

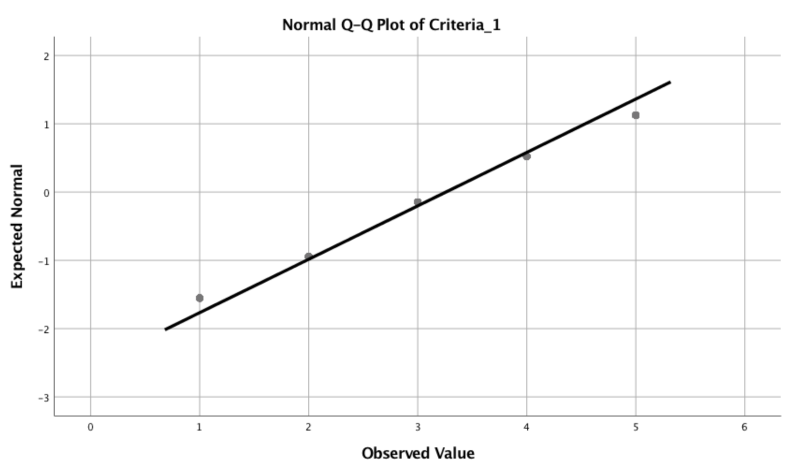


Figure F.3

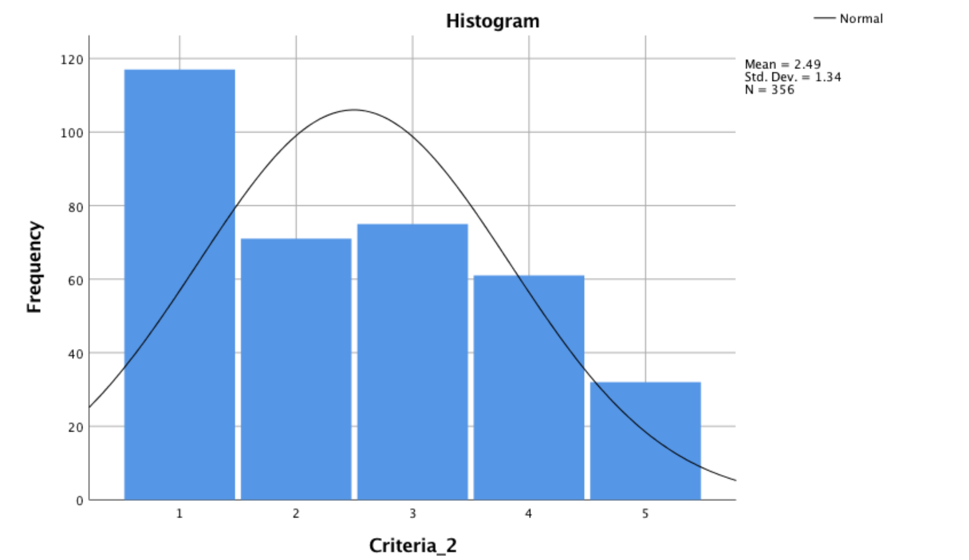


Figure F.4

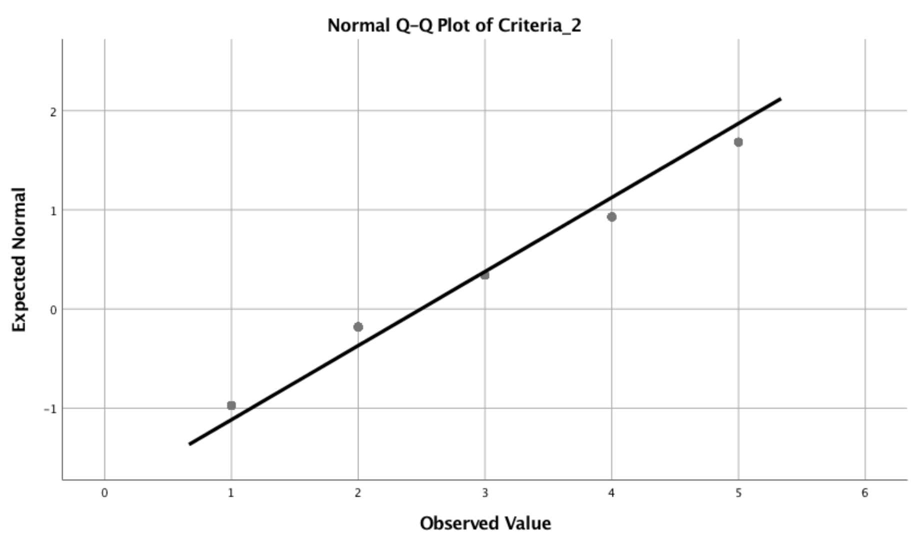


Figure F.5

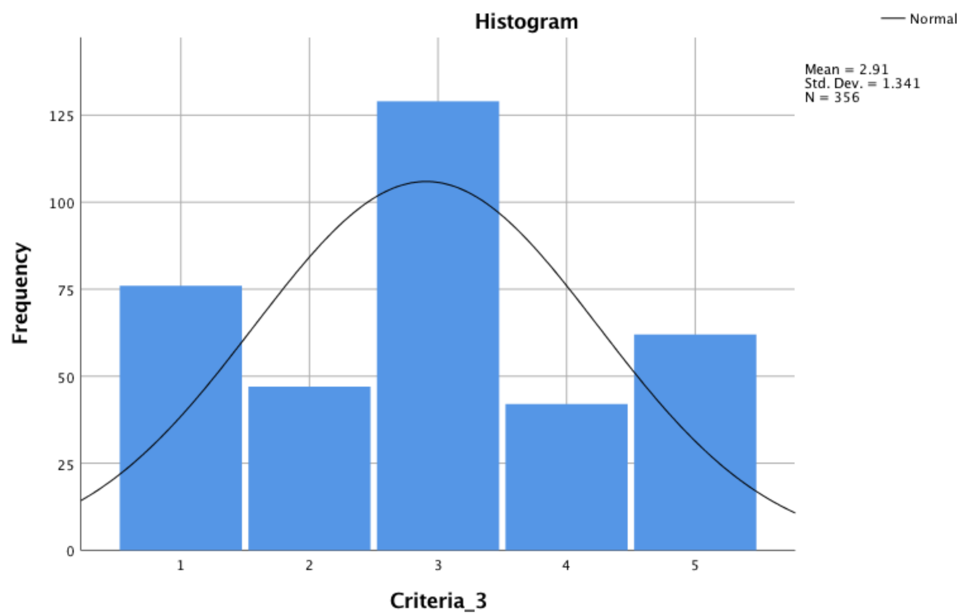


Figure F.6

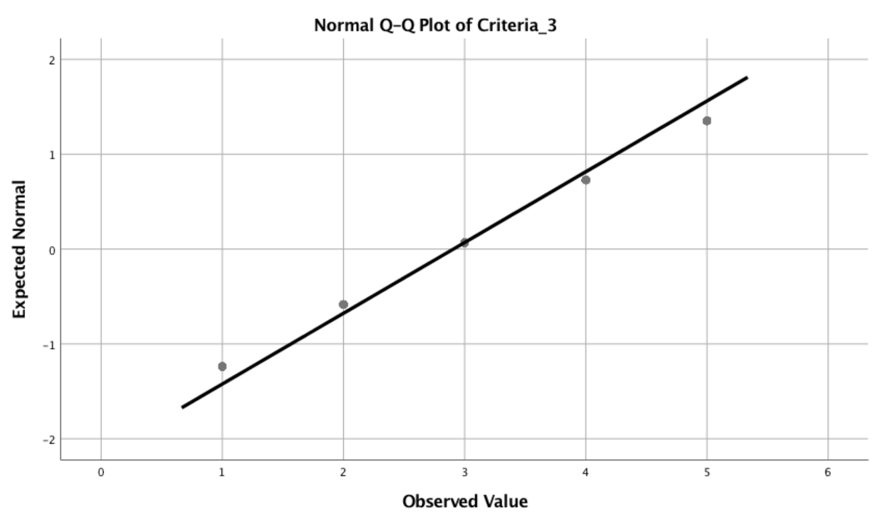


Figure F.7

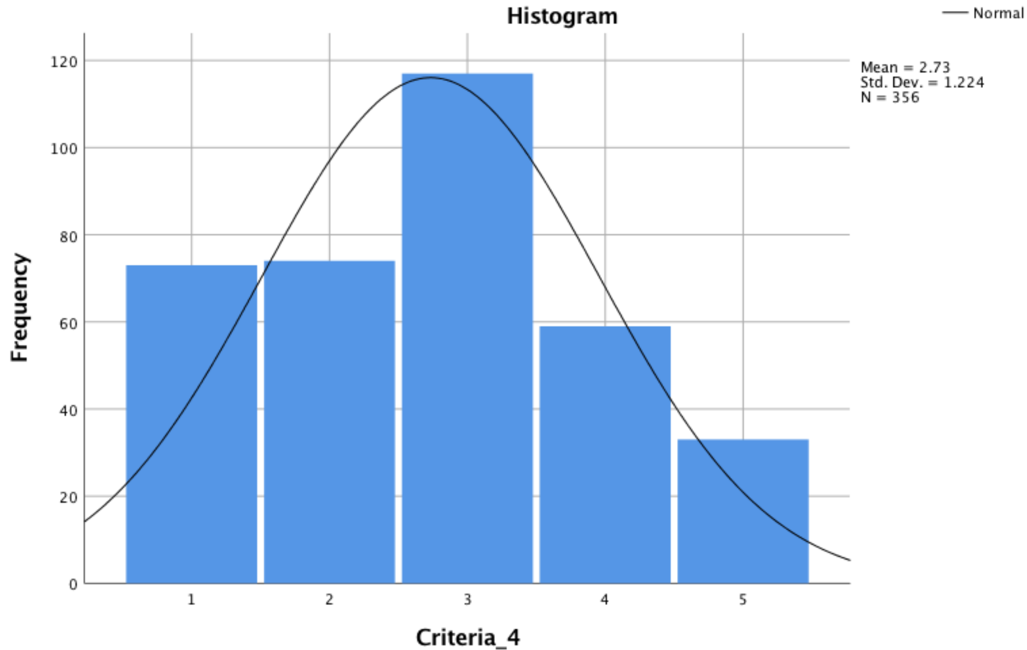


Figure F.8

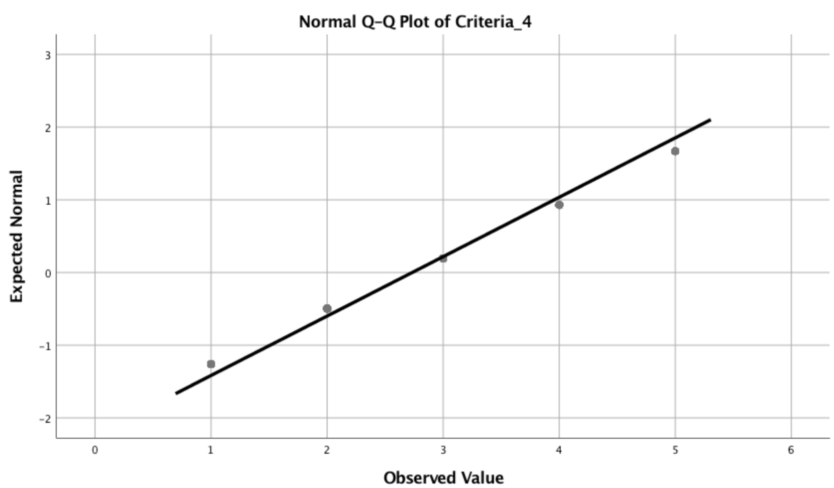


Figure F.9

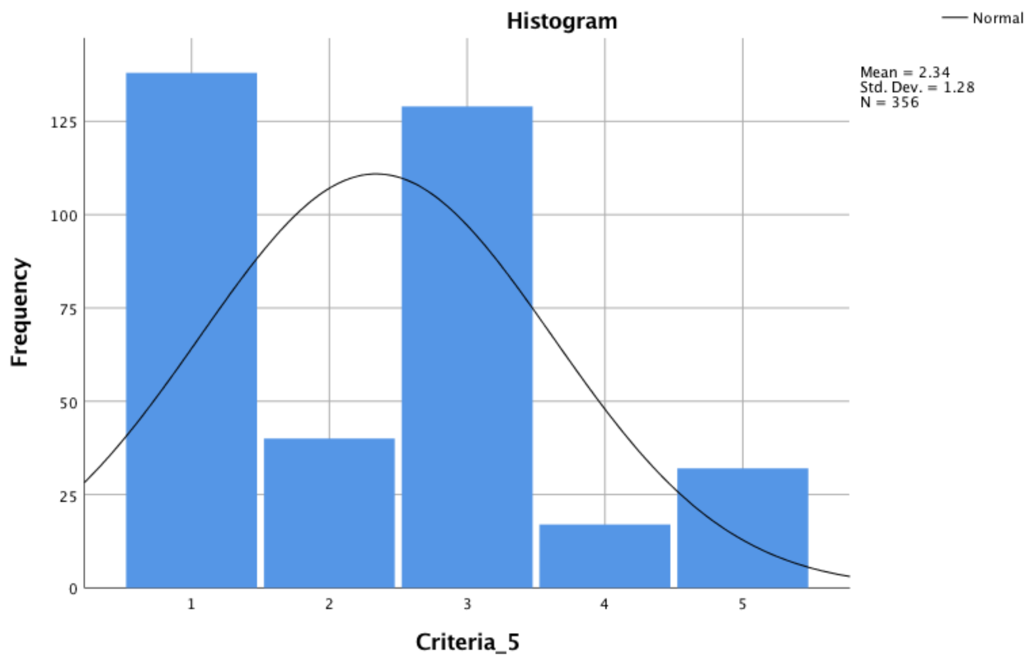


Figure F.10

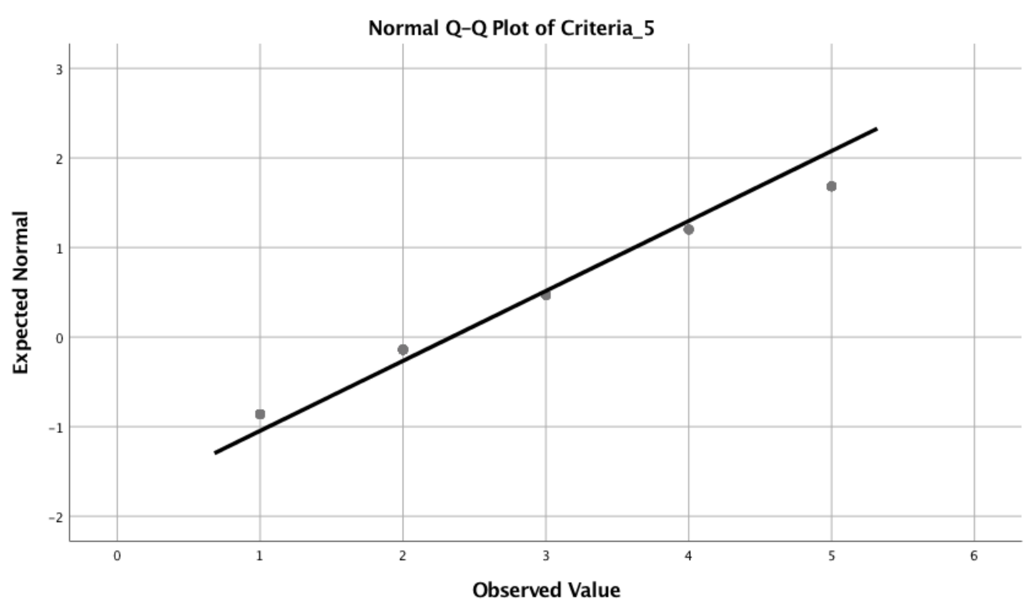


Figure F.11

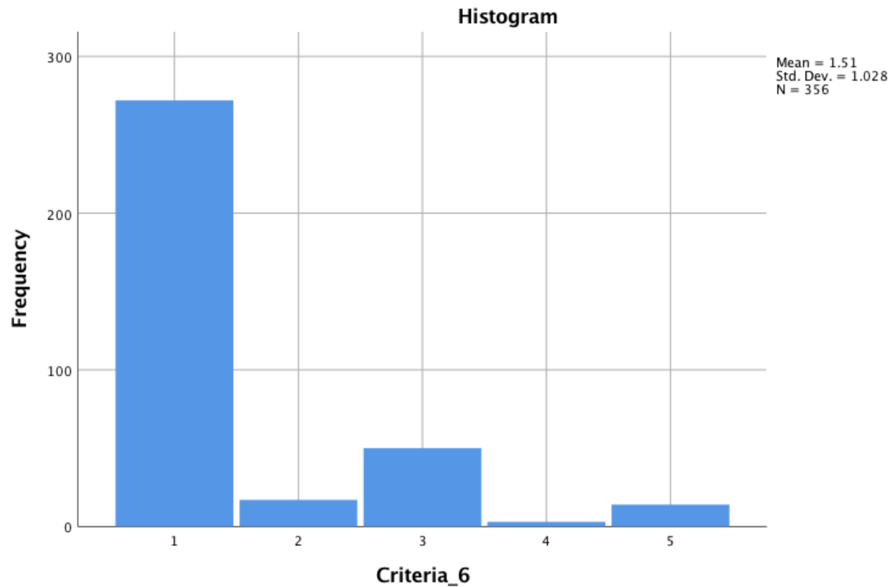

Histogram
Mean = 1.51
Std. Dev. = 1.028
N = 356
Frequency
300
200
100
0
1
2
3
4
5
Criteria_6

Figure F.12

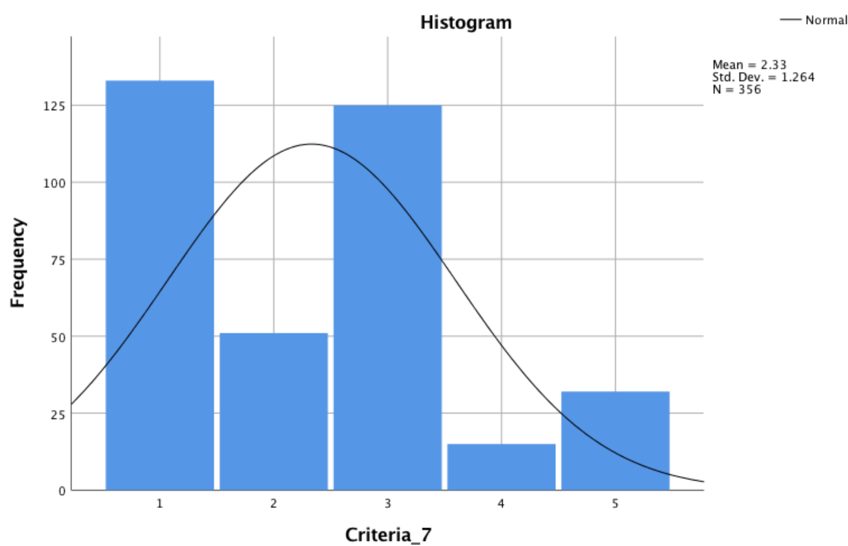

Histogram
Normal
Mean = 2.33
Std. Dev. = 1.264
N = 356
Frequency
125
100
75
50
25
0
1
2
3
4
5
Criteria_7

Figure F.13

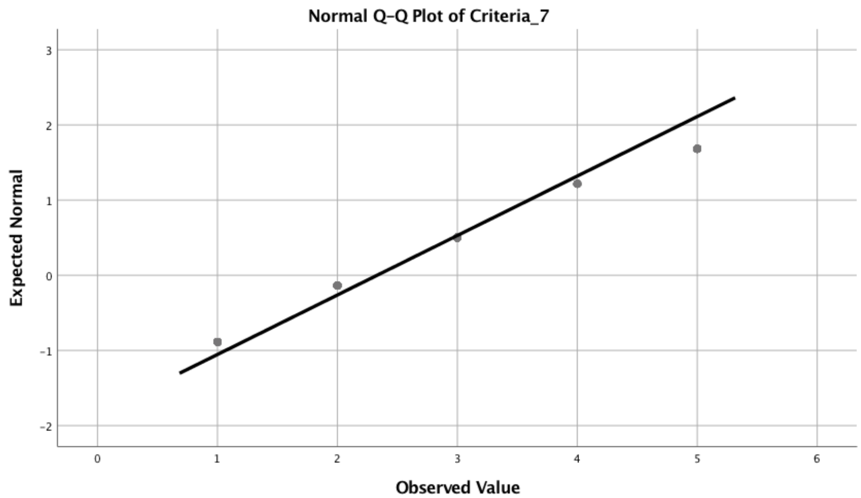


Figure F.14

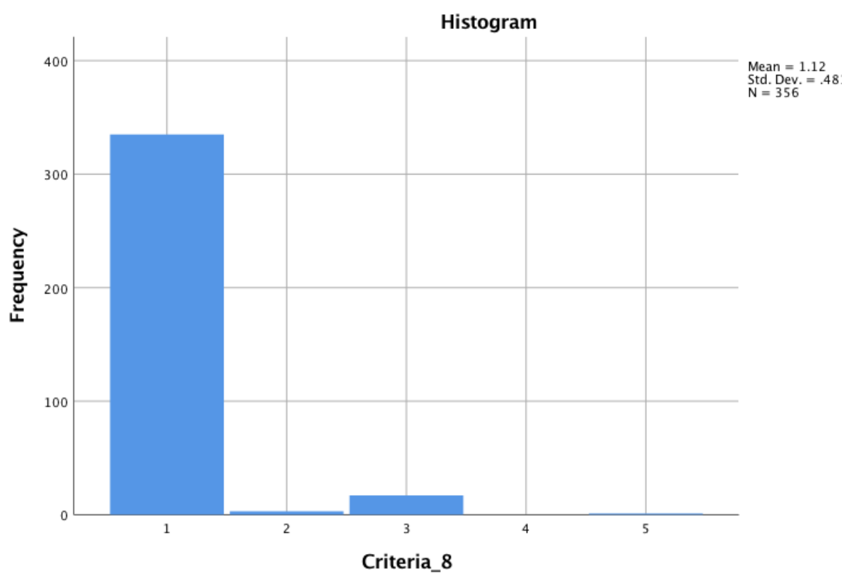


Figure F.15

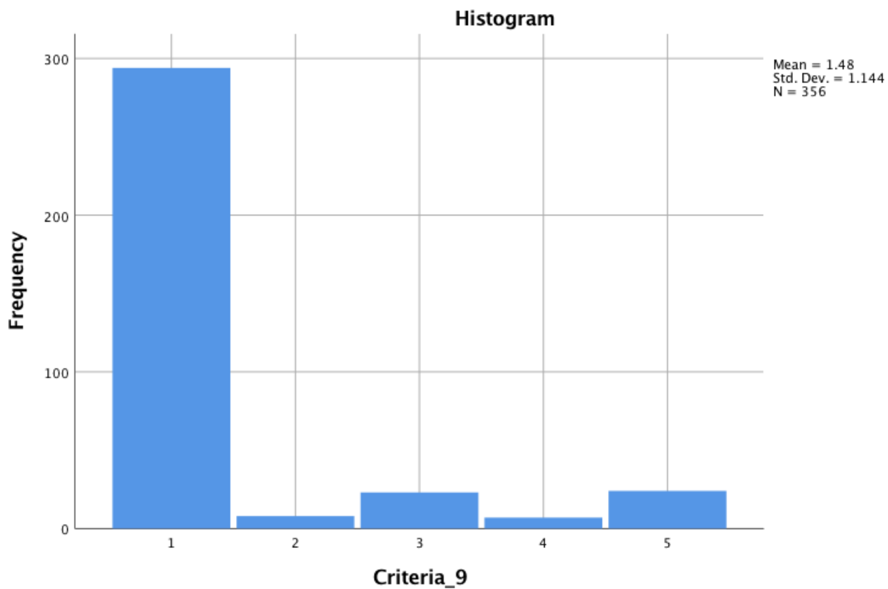


Figure F.16

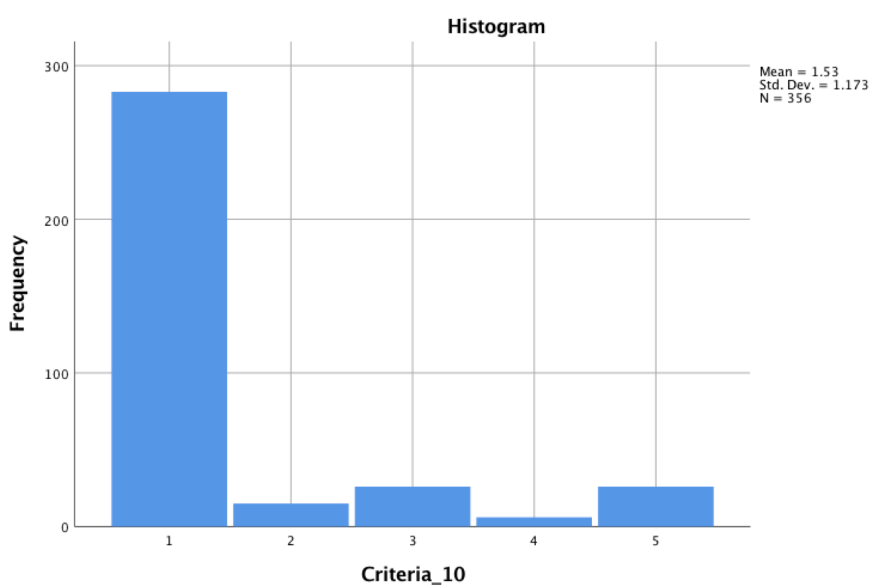


Figure F.17

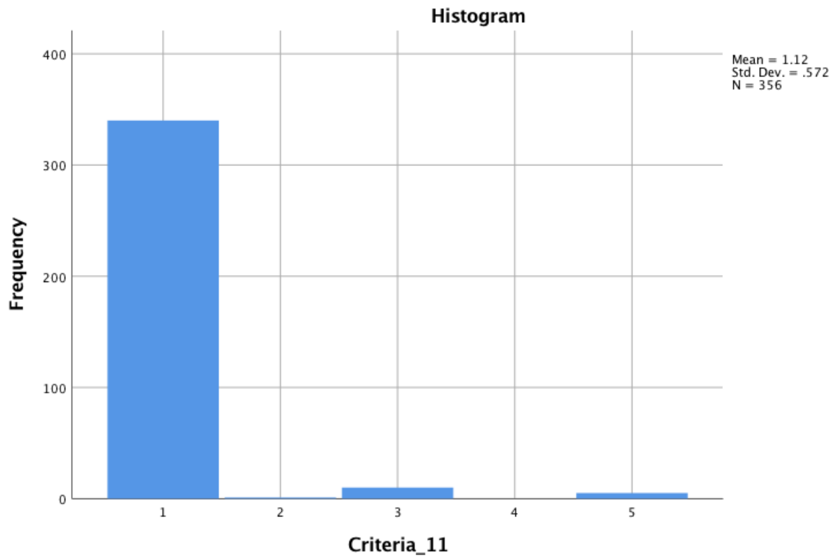

Histogram
Mean = 1.12
Std. Dev. = .572
N = 356
Frequency
400
300
200
100
0
1
2
3
4
5
Criteria_11

Figure F.18

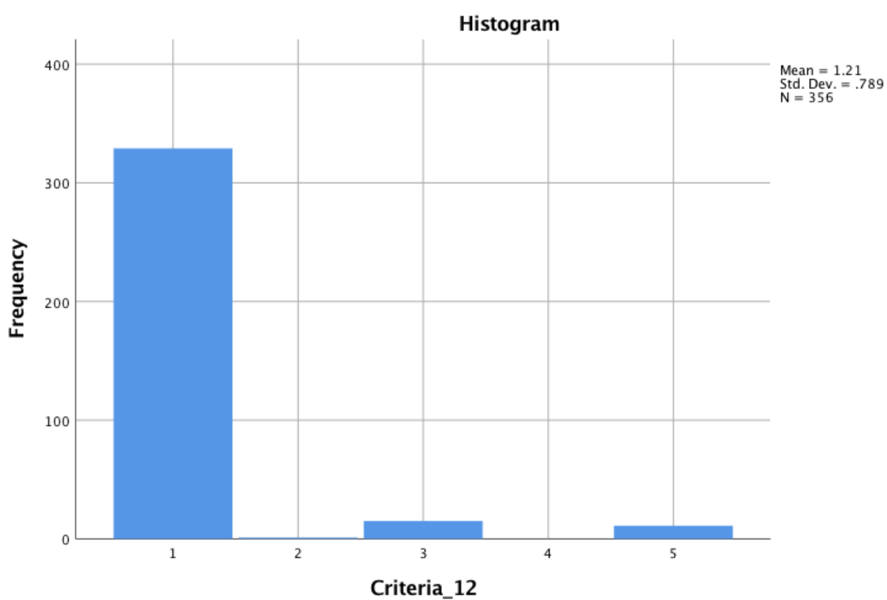

Histogram
Mean = 1.21
Std. Dev. = .789
N = 356
Frequency
400
300
200
100
0
1
2
3
4
5
Criteria_12

Figure F.19

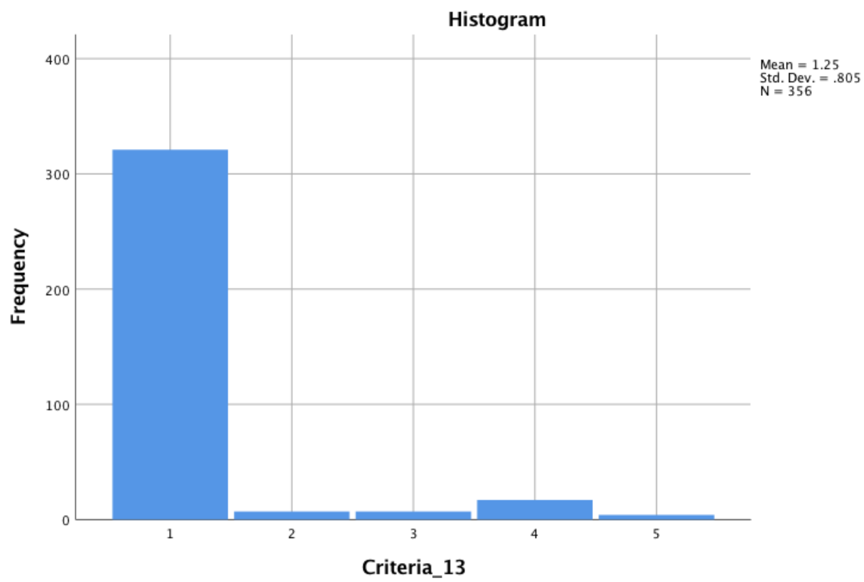


Figure F.20

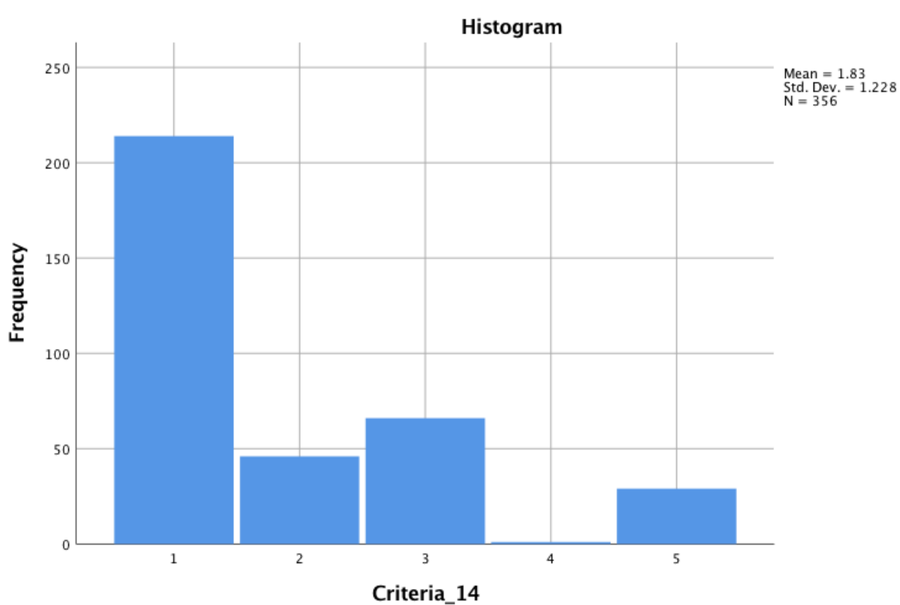


Figure F.21

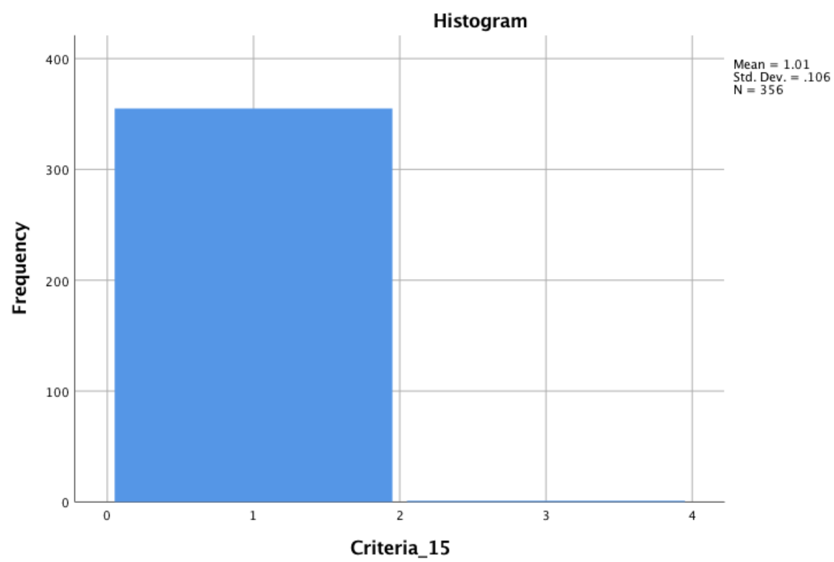


Figure F.22

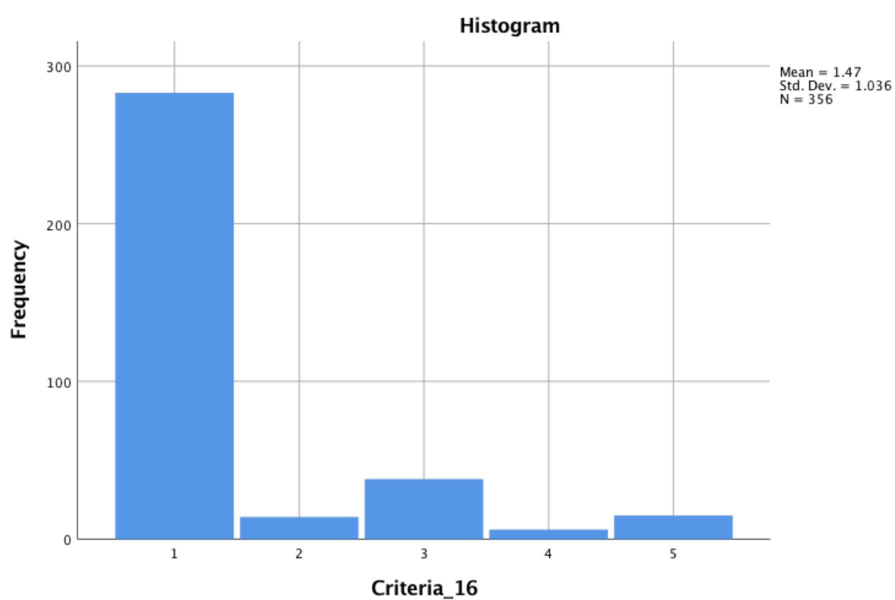


Figure F.23

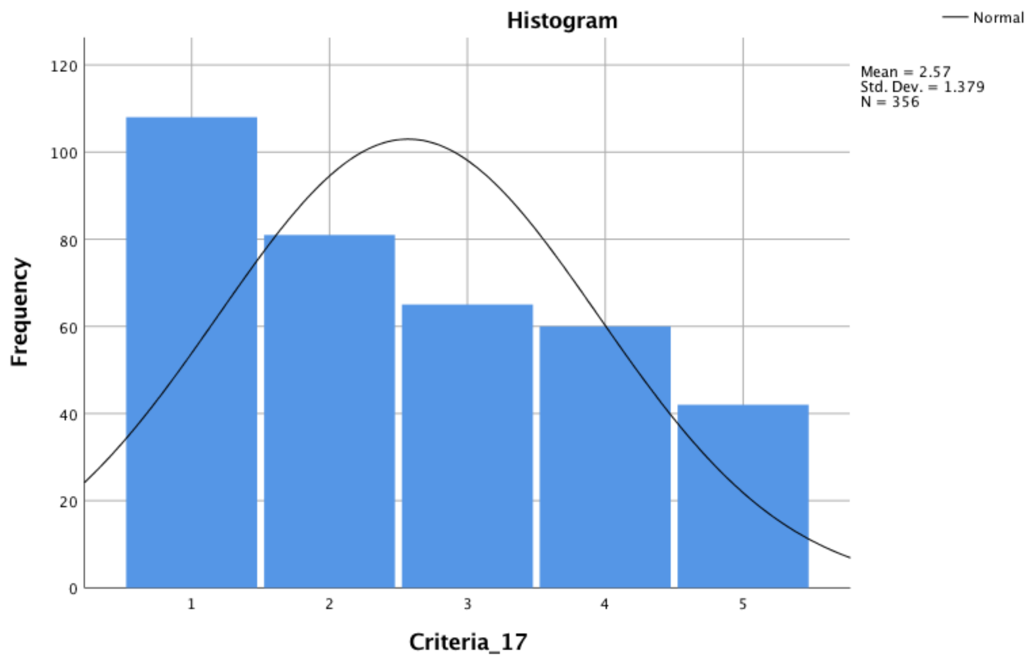


Figure F.24

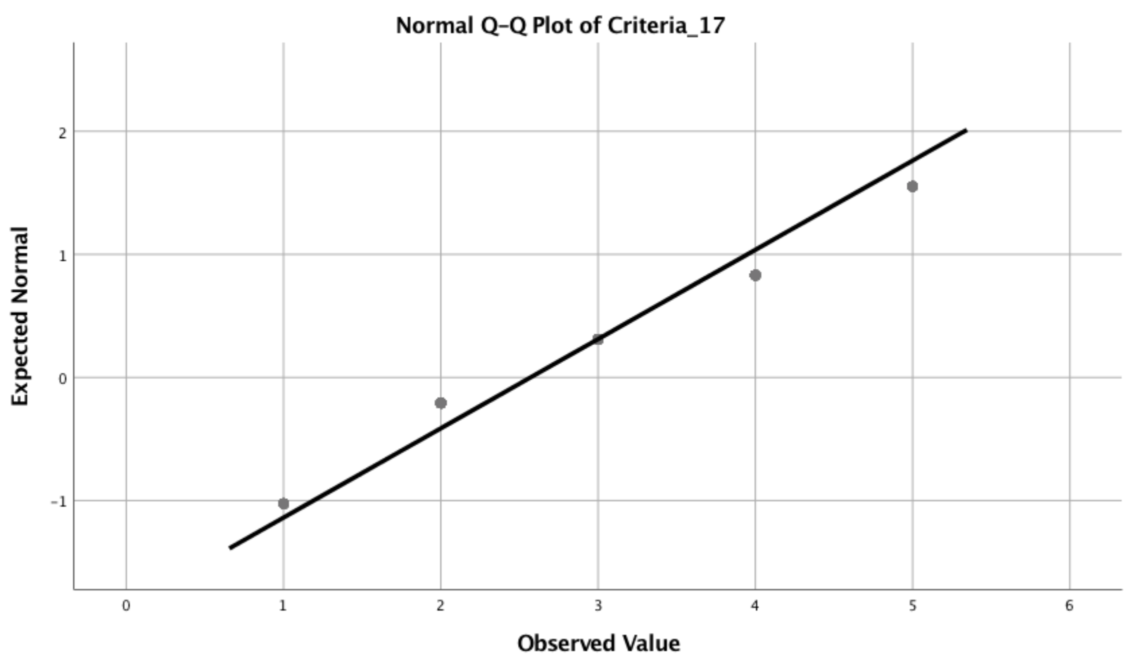


Figure F.25

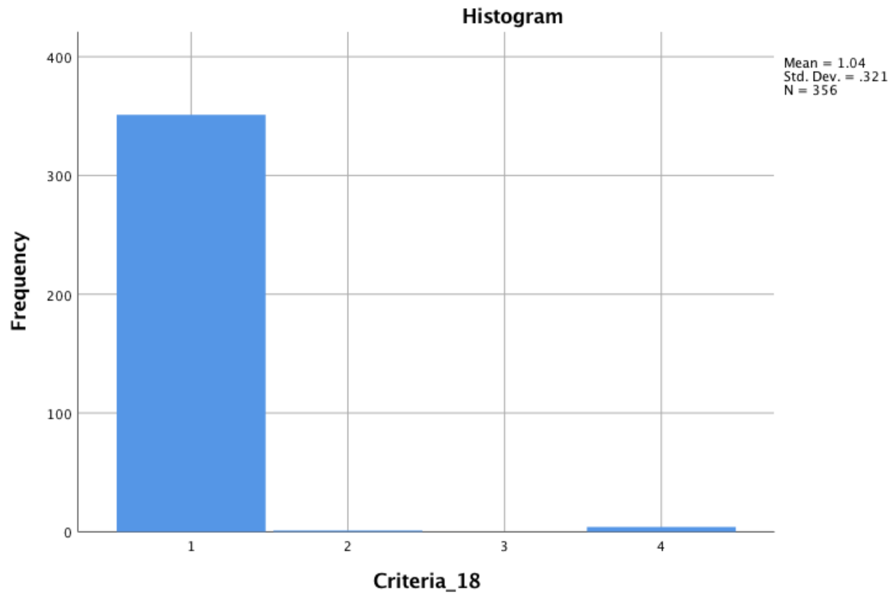

Histogram
400
300
200
100
0
Frequency
1
2
3
4
Criteria_18
Mean = 1.04
Std. Dev. = .321
N = 356

Figure F.26

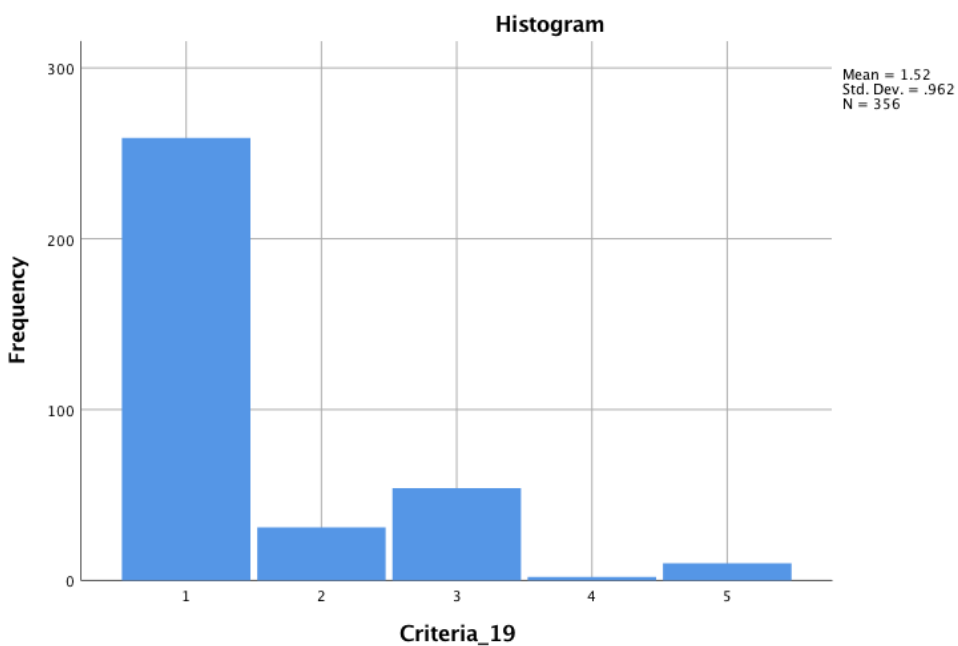

Histogram
300
200
100
0
Frequency
1
2
3
4
5
Criteria_19
Mean = 1.52
Std. Dev. = .962
N = 356

Figure F.27

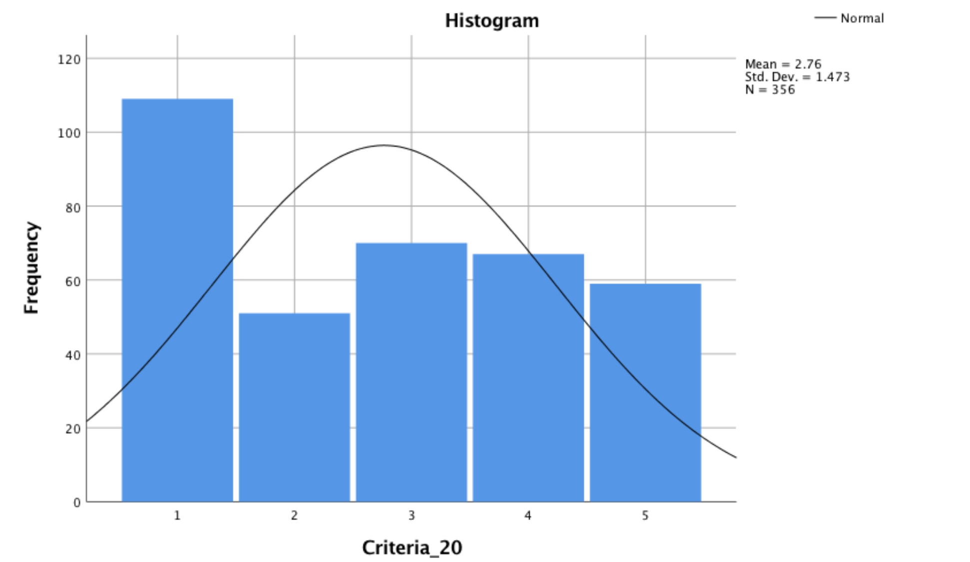


Figure F.28

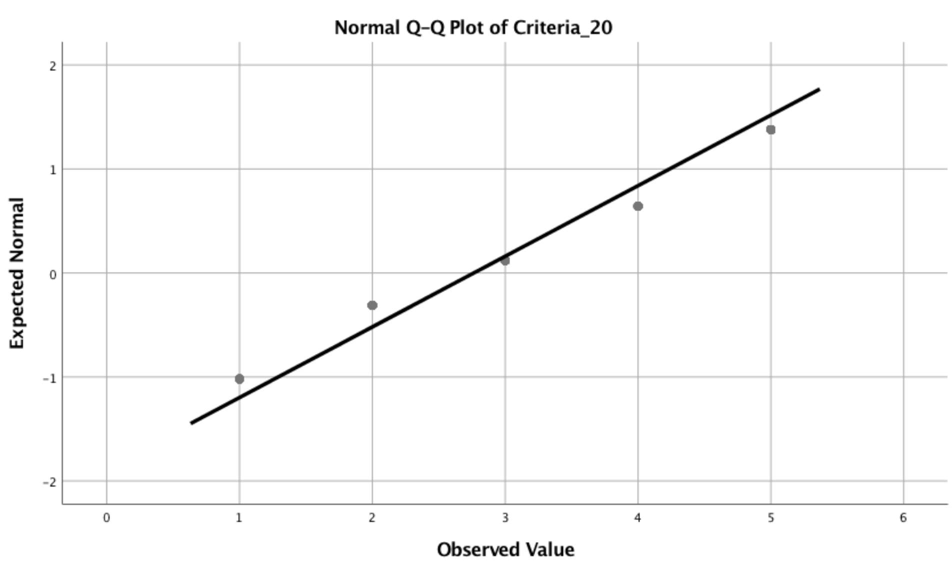


Figure F.29

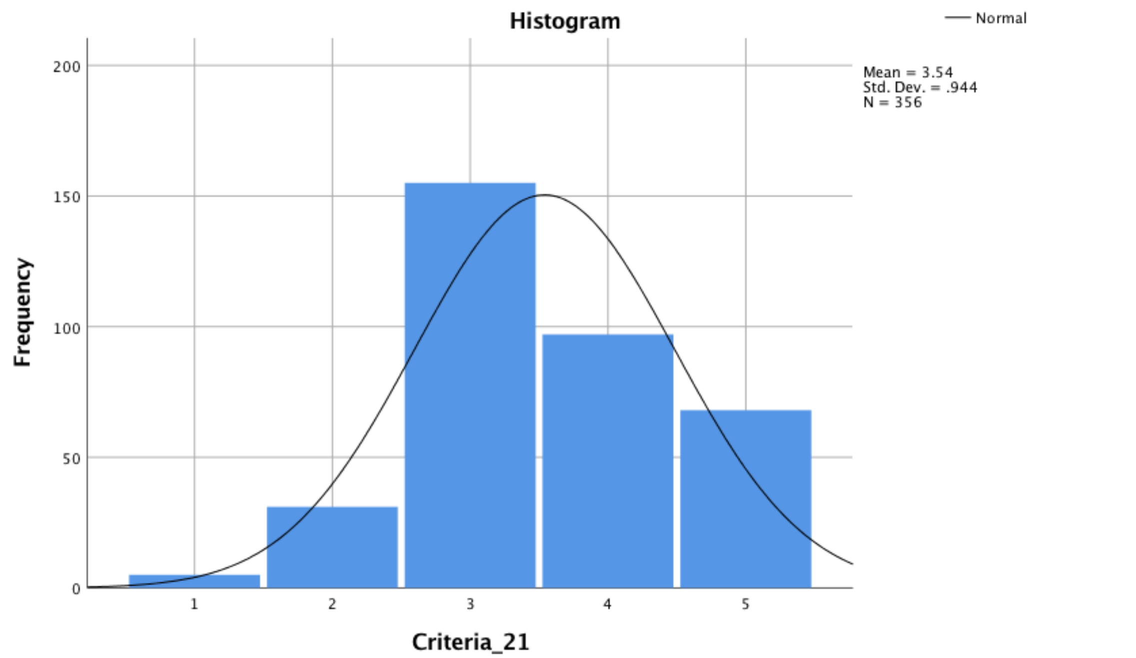


Figure F.30

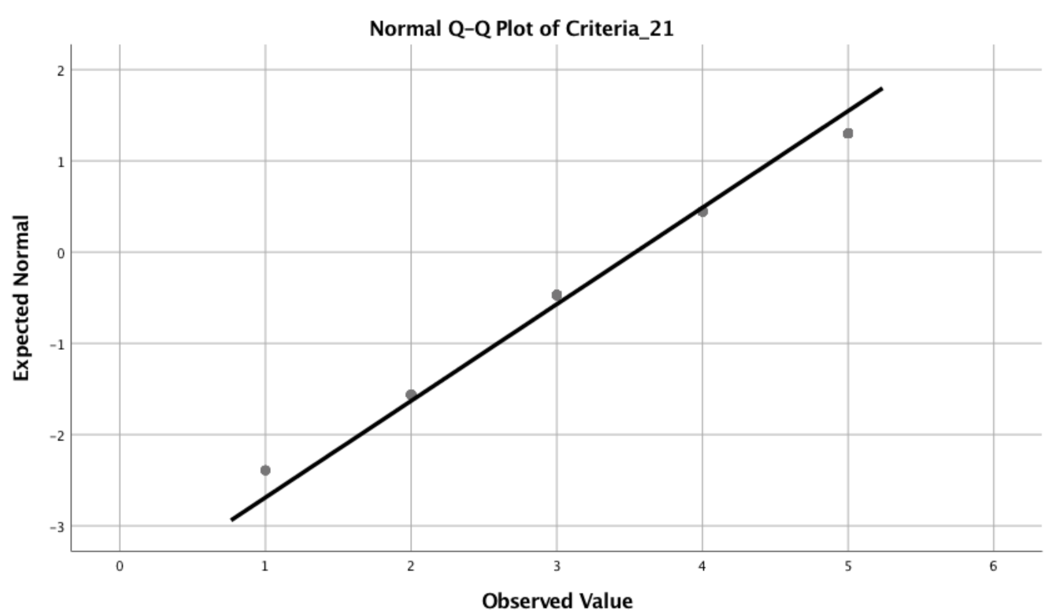


Figure F.31

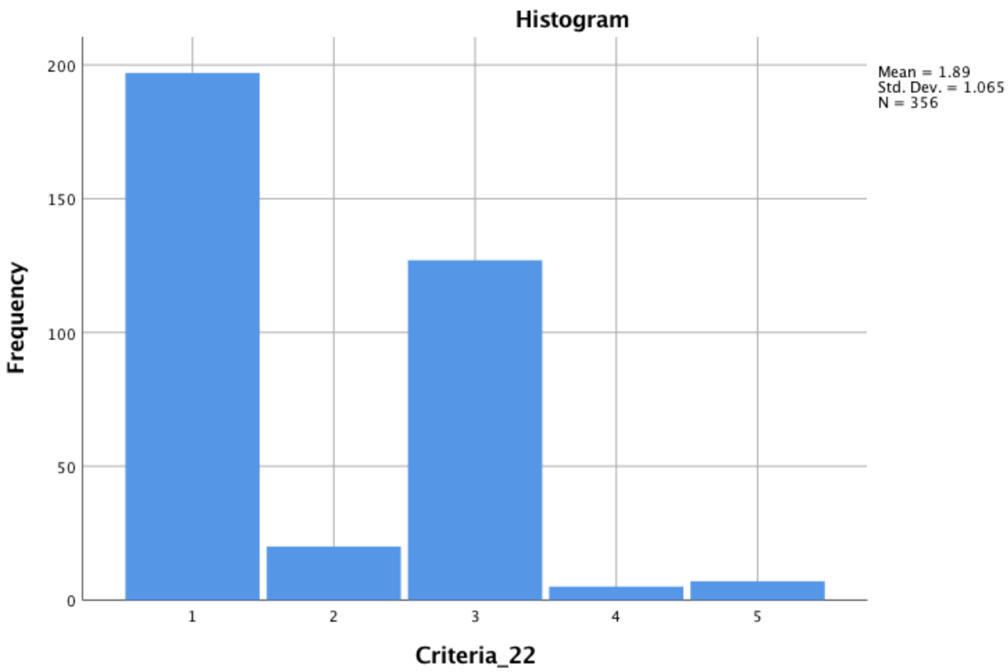

Histogram
Mean = 1.89
Std. Dev. = 1.065
N = 356
Frequency
200
150
100
50
0
1
2
3
4
5
Criteria_22

Figure F.32

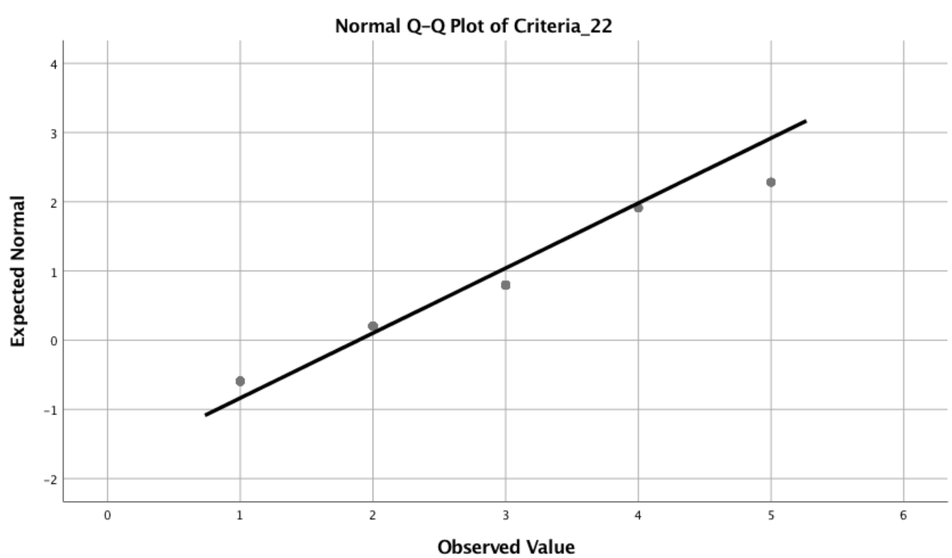

Normal Q-Q Plot of Criteria_22
Expected Normal
4
3
2
1
0
-1
-2
0
1
2
3
4
5
6
Observed Value

Figure F.33

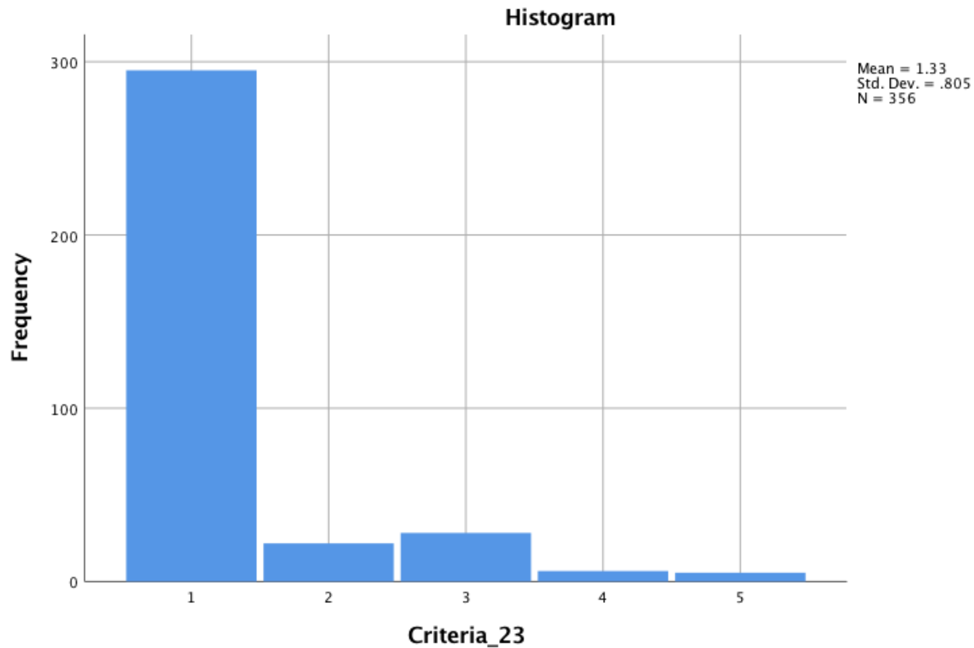

Histogram
Mean = 1.33
Std. Dev. = .805
N = 356
Frequency
300
200
100
0
1
2
3
4
5
Criteria_23

Table F.34 Descriptive Statistics

Descriptive Statistics

	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Criteria_1	356	4	1	5	1160	3.26	1.278	-0.087	0.129	-0.833	0.258
Criteria_2	356	4	1	5	888	2.49	1.34	0.38	0.129	-1.097	0.258
Criteria_3	356	4	1	5	1035	2.91	1.341	0.071	0.129	-0.997	0.258
Criteria_4	356	4	1	5	973	2.73	1.224	0.151	0.129	-0.849	0.258
Criteria_5	356	4	1	5	833	2.34	1.28	0.518	0.129	-0.678	0.258
Criteria_6	356	4	1	5	538	1.51	1.028	2.013	0.129	3.277	0.258
Criteria_7	356	4	1	5	830	2.33	1.264	0.561	0.129	-0.573	0.258
Criteria_8	356	4	1	5	397	1.12	0.481	4.418	0.129	20.666	0.258
Criteria_9	356	4	1	5	527	1.48	1.144	2.283	0.129	3.832	0.258
Criteria_10	356	4	1	5	545	1.53	1.173	2.135	0.129	3.224	0.258
Criteria_11	356	4	1	5	397	1.12	0.572	5.444	0.129	30.774	0.258
Criteria_12	356	4	1	5	431	1.21	0.789	3.927	0.129	14.852	0.258
Criteria_13	356	4	1	5	444	1.25	0.805	3.264	0.129	9.5	0.258
Criteria_14	356	4	1	5	653	1.83	1.228	1.394	0.129	0.958	0.258
Criteria_15	356	2	1	3	358	1.01	0.106	18.868	0.129	356	0.258
Criteria_16	356	4	1	5	524	1.47	1.036	2.203	0.129	3.928	0.258
Criteria_17	356	4	1	5	915	2.57	1.379	0.378	0.129	-1.139	0.258
Criteria_18	356	3	1	4	369	1.04	0.321	9.011	0.129	80.701	0.258
Criteria_19	356	4	1	5	541	1.52	0.962	1.888	0.129	3.06	0.258
Criteria_20	356	4	1	5	984	2.76	1.473	0.147	0.129	-1.38	0.258
Criteria_21	356	4	1	5	1260	3.54	0.944	0.007	0.129	-0.469	0.258
Criteria_22	356	4	1	5	673	1.89	1.065	0.684	0.129	-0.644	0.258
Criteria_23	356	4	1	5	472	1.33	0.805	2.662	0.129	6.825	0.258

Table F.35 Tests of Normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Criteria_1	0.237	356	0.000	0.866	356	0.000
Criteria_2	0.196	356	0.000	0.869	356	0.000
Criteria_3	0.182	356	0.000	0.886	356	0.000
Criteria_4	0.173	356	0.000	0.905	356	0.000
Criteria_5	0.24	356	0.000	0.829	356	0.000
Criteria_6	0.455	356	0.000	0.551	356	0.000
Criteria_7	0.228	356	0.000	0.835	356	0.000
Criteria_8	0.536	356	0.000	0.248	356	0.000
Criteria_9	0.489	356	0.000	0.467	356	0.000
Criteria_10	0.469	356	0.000	0.503	356	0.000
Criteria_11	0.535	356	0.000	0.201	356	0.000
Criteria_12	0.529	356	0.000	0.285	356	0.000
Criteria_13	0.522	356	0.000	0.337	356	0.000
Criteria_14	0.353	356	0.000	0.694	356	0.000
Criteria_15	0.518	356	0.000	0.028	356	0.000
Criteria_16	0.471	356	0.000	0.513	356	0.000
Criteria_17	0.191	356	0.000	0.871	356	0.000
Criteria_18	0.531	356	0.000	0.088	356	0.000
Criteria_19	0.433	356	0.000	0.591	356	0.000
Criteria_20	0.191	356	0.000	0.864	356	0.000
Criteria_21	0.253	356	0.000	0.876	356	0.000
Criteria_22	0.352	356	0.000	0.719	356	0.000
Criteria_23	0.486	356	0.000	0.463	356	0.000

a. Lilliefors Significance Correction

VITA

Cindy L. Daniels

Department of Engineering Management and Systems Engineering

Old Dominion University, Norfolk VA 23529

EDUCATION

Doctor of Philosophy, Engineering Management, Old Dominion University, 2014 (expected)

Master of Engineering Management, The George Washington University, Washington D.C. 2000

M.S. Information Systems, The George Washington University, Washington D.C. 1999

B.S. Mathematics, Northern Michigan University, Marquette MI 1981

WORK EXPERIENCE

Ms. Cindy Daniels has over 30 year of experience at NASA in project management and engineering management. Ms. Daniels is currently the Director of the NASA Science Office for Mission Assessments at NASA Langley Research Center.