https://ntrs.nasa.gov/search.jsp?R=20190002710 2020-05-09T20:20:34+00:00Z

Survey and Assessment of Crew Performance Evaluation Methods Applicable to Human Spacecraft Design

Christine Fanchiang, University of Colorado, Boulder, CO, USA (650) 302-2692, christine.fanchiang@colorado.edu
Jessica J. Marquez, NASA Ames Research Center, Moffett Field, CA, USA, (650) 604-6364, jesssica.j.marquez@nasa.gov
Brian F. Gore, NASA Ames Research Center, Moffett Field, CA, USA, (650) 604-2542, brian.f.gore@nasa.gov
David Klaus, University of Colorado, Boulder, CO, USA, (303) 492-3525, klaus@colorado.edu

Abstract— **Space is an unforgiving environment where the actions of the crew play a critical role in their health and safety. Given the limited number of crewmembers typically onboard a spacecraft and the multitude of complex systems they must operate, the performance of each individual is of paramount importance. Spacecraft habitat layout and operations are two main drivers affecting crew performance efficiency. Having the capability to analyze and compare crew performance across various spacecraft configurations can help identify improvements early in the conceptual design process where changes are less costly to implement, ultimately reducing overall project costs and improving long-term operations of the system. Currently, there are few comprehensive methods readily available for evaluating crew integration within a spacecraft in the conceptual design phase. In order to address this shortcoming, the goal of this work was to analyze various specialized evaluation methods found in analogous industries that have potential application to human spacecraft design. A survey of more than 400 human performance evaluation methods was completed. Over twenty different attributes were identified for each method and a variety of analyses were conducted to characterize and evaluate their potential use for assessing human spacecraft design options. The analysis revealed a particular deficiency of quantitative evaluation methods that are applicable early in the systems engineering design phase. It also identified five existing methods that could be supplemented to achieve the needs of an early design evaluation method. Additional discussion describes potential issues that must be overcome when developing a method specific for use in human spacecraft evaluations.**

TABLE OF CONTENTS

1. INTRODUCTION 1
2. BACKGROUND 2
3. METHODOLOGY 4
4. RESULTS 4
6. CONCLUSION 9

1. INTRODUCTION

Designing a human spacecraft is a complex engineering process of balancing performance and safety while minimizing mass, volume, power, and cost. The tight engineering and budget constraints as well as the extreme environmental conditions of space drive the need for critical and thorough evaluations and analyses of the spacecraft design prior to flight. The need to protect and appropriately utilize the crew demands even more stringent requirements for safety and operability of the system.

Crew members provide a number of functions that enhance the capabilities of the spacecraft, including but not limited to conducting complex science objectives, executing contingency plans in the event of failures, and being test subjects themselves for human research in space. As such, maintaining adequate levels of crew performance is critical when analyzing spacecraft design. Various design choices ranging from controls and display designs to cabin architecture and layout all have an impact on crew performance.

Human factor engineers and practitioners are well-aware of these design challenges, and over the years have helped publish a number of standards and guidelines for spacecraft designers [1-6]. Even with the helpful documentation, evaluating whether the design appropriately accommodates and utilizes the crew is still a challenge for a number of reasons. A National Research Council Report (2007) [7] lists three general categories as contributing to inadequate human integration into complex systems: 1) poor research to practice translation 2) lack of effective methods and tools, and 3) poor documentation of human task performance. Consequently, there is often little advocacy and inadequate resource allocation for human systems integration (HSI) in organizations.

NASA has also noted that there is a lack of effective methods and tools available to quantify human factor measures for spacecraft design. As reported in NASA's Human Research Roadmap, one of the gaps points to a lack of: "[m]ethodologies and metrics for integrated vehicle/system level evaluations leveraging multiple, complementary tools/methods such as digital modeling, [human-in-the-loop] HITL evaluations, and population analysis." [8]

The combination of these factors leads to a growing interest in finding and defining effective tools or methods for understanding how humans can be effectively integrated into a complex system. Because of this apparent dearth of existing crew performance metrics for spacecraft design, the goal of this paper is to review and assess the few that have been developed for spacecraft, as well as investigate methods from analogous industries and determine whether

they might be useful to adapt for the evolving human spacecraft industry.

2. BACKGROUND

A. A Growing Human Spaceflight Industry

As prospective missions are being developed for human spaceflight, a variety of new providers and users have joined in the industry. While historically human spaceflight has been in the realm of government control, the recent years have seen a growing interest in commercialization of space through tourism [9]. The mission concepts range from a few minutes of suborbital experience to a one-way journey to Mars, while the anticipated users range from veteran pilots to retired grandmothers. The diversity of participants in this new industry is important to consider when ensuring safety and accommodation requirements for the spacecraft.

In recognition of the changing landscape of human spaceflight, the Federal Aviation Administration (FAA) and NASA have been proactively developing recommendations and standards for evaluation and analysis of spacecraft prior to flight. NASA has established their own Human-Rating Requirements (NPR 8705.2B) [10] that ensures compliance by their commercial crew providers, while the FAA has more recently published a list of recommended practices for supporting good human systems integration [11].

Using the currently available recommendations and requirements is one of the first steps to build a foundation for identifying effective tools and methods for spacecraft design analysis. It is important to understand how human spacecraft have historically been evaluated, and whether lessons learned are extensible to the new era of spacecraft designs.

B. Defining Tools, Methods, and Framework

A number of methods and tools have previously been developed to help evaluate and assess various aspects of designs and user interfaces. But before determining the best method to use for spacecraft design, it is necessary to define three central terms used throughout this paper; tools, methods, and frameworks. A tool is a device to measure, collect and analyze data; a method is a technique to measure data that can integrate a number of tools to establish a process; and a framework provides the foundation for how the data is construed. To improve readability, this paper uses the word '*methodologies'* to bundle these three terms together.

C. Current Methodologies for Human Spacecraft Design

Historically, NASA has developed and relied on the systems engineering process as a systematic, requirements-based methodology to ensure compliance of the design. The systems engineering process is broken down in the NASA Systems Engineering Handbook (2007) [12] into six main phases.

Project Phase A: Concept and Technology Development: determines feasibility and desirability of suggested system and establishes an initial baseline compatibility with NASA's strategic plans.

Project Phase B: Preliminary Design and Technology Completion: define project in enough detail to establish an initial baseline capable of meeting mission needs.

Project Phase C: Final Design and Fabrication: complete detailed design of the system (and its associated subsystems, including its operations systems), fabricate hardware, and code software.

Project Phase D: System Assembly, Integration and Test, Launch: assemble integrate the products and create the system, meanwhile developing confidence that it will be able to meet the system requirements, conduct launch and prepare for operations.

Project Phase E: Operations and Sustainment: conduct the mission and meet the initially identified need and maintain support for that need.

Project Phase F: Closeout: implement the systems decommissioning/disposal plan developed in Phase C and analyzed returned data and samples [12].

Each of the project phases correspond to different activities that occur throughout the system design lifecycle. Several tools and methods have been developed to help designers integrate crewmembers at different stages of the systems engineering design process. They can be classified into four main categories as aggregated in this study. Figure 1 illustrates when each tool or method is applied across the different systems engineering process. Phase F is not included here as it does not typically influence vehicle design impacts on crew performance.

These four methods have been used frequently on historical space missions, but have a number of limitations that drive the need for improvements in the methods as is described in the following paragraphs.

Phase A- Historical data on spacecraft design impacts to crew performance is often in the form of anecdotal reports and are limited to a unique and limited set of users. To-date there have been about eleven different space vehicles ever flown with humans aboard [13] and close to 500 people who have been in space [14]. The data gathered from these missions while useful are limited and particular to a unique user population of highly skilled, educated, trained, and healthy individuals. Therefore, it can be difficult to extrapolate and generalize the data to future users that do not fit the same criteria.

Phase B - Lessons learned from analogous industries can play a large role in defining many basic human factor requirements from anthropometric constraints to sensory limitations. Many accident investigation reports from the

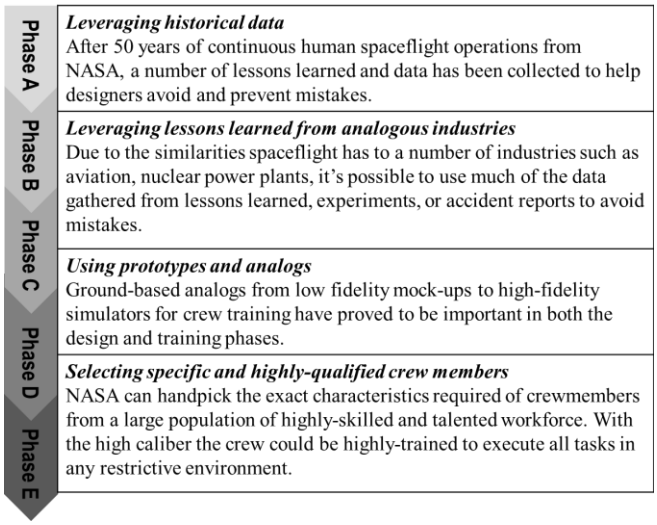


Figure 1. Current methodologies for evaluating human performance in a system mapped to systems engineering phases.

aviation industry have revealed critical human and systems integration issues. This is a wealthy source of data, but is limited to ground based interactions. The microgravity, radiation, and lighting environments of space are unique and create a number of effects that are detrimental to the human and can affect performance in a variety of ways that are not fully understood.

Phase C and D – Prototype and analog environments can provide relevant and insightful research, but high fidelity simulations can be expensive to produce and also come late in the process. There have been attempts to use ground-based analogs as indicators for future spaceflight conditions [15], but again are limited by the realism due to limited crew selection bias or non-similar operational dynamics (i.e. large polar expeditions vs small astronaut crew). Alternatively cheap and quick prototypes can be very useful early in the concept design, but consequently does not provide enough detail for requirement verification.

Phase D and E - Crew member selection and training continues to be a highly involved process. But as the industry evolves to include new populations of users a number of concerns have been brought up. In anticipation of the larger range of user age and health, the FAA has supported research in characterizing health and performance concerns of future spaceflight participants who may not be as healthy or fit as the typical NASA astronaut [16]. The expanded user population brings additional complication to ensuring quality spacecraft designs that can account for all ranges of human health and performance.

These methods lack flexibility, are limited by resources, and are constrained by the scarcity of spaceflight data. These limitations have driven the development of more quantitative and objective approaches that attempt to address those shortcomings. More powerful and capable computer hardware coupled with improved algorithms and tools have made computational modeling more accessible and approachable for analyzing spacecraft systems [17].

One methodology that has become more prevalent in the human spaceflight industry is the application of risk-based assessments for human error. As described by Chandler et al. (2006) [18] in a NASA Technical Report, NASA's adoption of Human Reliability Analysis (HRA) methods has been driven in part by the large number of mishaps attributed to human error in the industry. While incorporating human error into overall system reliability establishes early consideration and planning for human performance impacts, there are still limitations to this method. No HRA methods (with the exception of the Human Factors Process Failure Mode and Effects Analysis (HF PFMEA)) have been specifically designed for an aerospace application; most methods have been developed for nuclear power plant operations. Three specific limitations that fall out from this have been identified in the Technical Report as: "1) coverage of [performance shaping factors] PSF's and task characteristics unique to space missions; 2) applicability of the underlying data, [human error probability] HEP estimates, and PSF weights to space environment; and 3) significant differences in human action time scales between nuclear plant operation and space missions" [18].

Another more computationally intensive methodology that is gaining traction in human spaceflight is computer-aided modeling and simulation of spacecraft design. A number of the computational tools have been adapted from the aviation industry and applied in specific simulations for spacecraft design. An example of a well-known computational model is the Man-machine Integration Design and Analysis Systems (MIDAS), where it has been used previously for analysis of task and workload applications on space shuttle orbiter glove box design. [17]. The use of these human performance models (HPM) can help reduce design cost and improve efficiency, but still have many challenges to overcome. A major setback to human performance models is their complexity and difficulty for the spacecraft designer to use [19].

Given the limitations of the methods available for evaluating human spacecraft design, an alternative source of existing methodologies can be gleaned from industries that have similarities to human spaceflight. The authors chose to look across these analogous industries, and determine if they have tools, methods, or frameworks that could be adaptable to human spaceflight. Ultimately, the goal of this work is two-fold: 1) to aggregate a number of methodologies for evaluating human system performance and 2) to establish a formal analysis that identifies what methodologies are appropriate to utilize for the growing human spaceflight industry.

3. METHODOLOGY

To establish a formal survey and analysis of the various human performance evaluation methodologies, three main steps were involved:

Step 1: Aggregate as many methodologies from aerospace and analogous industries.

Step 2: Determine important characteristics specific to the spacecraft design process.

Step 3: Establish a formal evaluation criteria for analyzing the crew performance evaluation methodologies.

This section describes each of these steps and how over 400 methodologies were aggregated and reviewed. It also describes how the various characteristics were identified and deemed important for the human spacecraft design process.

A. Aggregating Tools, Methods, and Frameworks

The initial step in aggregating the various methodologies was identifying existing ones in the human spaceflight industry. This was done through a literature search for human performance models specific for spaceflight. Key word choices for the literature search included: human and crew performance models, evaluation tools, spacecraft design process, and human factor design tools.

Due to the evolution occurring in the spaceflight industry, it was important to look across a number of other analogous industries to determine if there were existing methodologies that would be appropriate or applicable for human spacecraft design.

In order to find these other methodologies, it was critical to identify what is considered an analogous industry. A number of unique attributes of human spaceflight were defined and listed in the first column of Table 1. The authors then identified various industries that possessed similar attributes. Five industries stood out for their likeness to human spaceflight: commercial aviation, naval operations, industrial engineering, nuclear power plant, and robotics and automation. A second literature review then looked specifically at the methodologies used that pertain to measuring or understanding human performance. Through this process, over 400 methodologies were collected and reviewed.

B. Characterization of Spacecraft Design Process

For each tool, method, or framework, twenty-two characteristics were recorded for comparison and analysis. The characteristics ranged from general classifications (e.g. models or simulation, performance assessment) to the required resources needed for using the methodology. The characteristics were selected because they were deemed by the authors as potential indicators that could be useful for spaceflight.

C. Establish Evaluation Criteria

A formalized evaluation process was established to help identify what criteria were most important for a spacecraft design methodology. The four criteria of interest were: early design phase, quantitative output, state driven, and flexible architecture. Each of these criteria could be defined by filtering the twenty-two characteristics identified from Step 2. A more detailed description of each of the four criteria is captured in the results section.

From the collection method to characterization and establishment of the evaluation criteria, a final list of methodologies was identified as potential candidates for future development.

4. RESULTS

The collection of 400 human systems integration evaluation methods were categorized based on specific attributes required for spacecraft design.

A. Methodology Classification

The first analysis classified the various types of methods into one of nine categories as shown in the Figure 2.

Table 1. Comparison table of industries analogous to human spaceflight.

Unique Human Spaceflight Attributes	Commercial Aviation	Naval Operations	Industrial Engineering	Nuclear Power Plant	Robotics/ Automation
1 Small cramped living/working quarters	X	X	X	X	
2 Limited people	X	X	X	X	X
3 People stay or get exchanged a few times	X		X	X	X
4 Highly selected people	X	X		X	
5 Dangerous outside environment	X	X			
6 Live/work inside 99% of time		X			
7 Highly-esteemed work	X			X	
8 Highly-technical work	X	X	X	X	X
9 Lots of training required for crew	X	X	X	X	
10 Environment fully controllable	X	X			

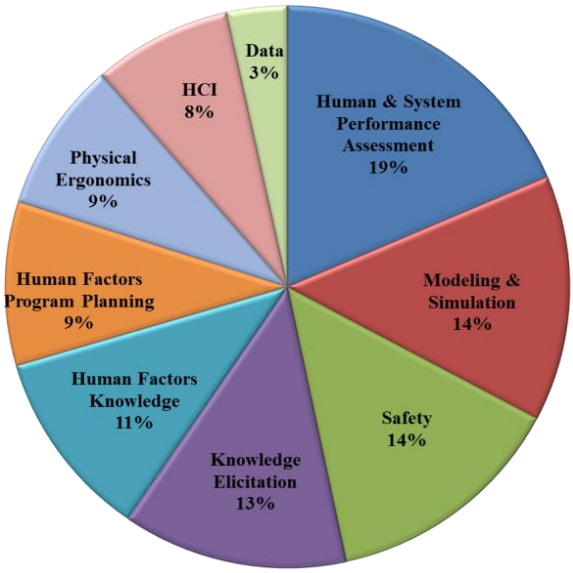


Figure 2. Breakdown of methodology classification.

A definition for each of the categories is described below in order of the most methodologies per category.

Human & System Performance Assessment: These types of methods include cognitive testing, function allocation, generic performance measurements, physiological tools, secondary task evaluations, situation awareness, stress, and workload specific tools.

Modeling and Simulation: These methods leverage any type of model or simulation that has been created to analyze human performance.

Safety: These methods are pulled from literature for evaluating accident investigations, human errors, human reliability, and risk assessments.

Knowledge Elicitation: These methods are used during usability tests in which the information can be elicited from the user in a variety of ways. These include cognitive task analysis, interviews, observation, questionnaires, and task analysis.

Human Factors Knowledge: This category captures any documentation that has been generated to help guide designers, including databases, guidelines, glossaries, standards, and technical reports.

Human Factors Program Planning: This category holds any methods and tools that are useful for managing a large and complex human system including cost/risk benefit analyses, data item descriptions, decision-making tools, economic cost analyses, and project management methods.

Physical Ergonomics: This set of methods is specific to any ergonomic or anthropometric models meant to help the designer appropriately size and place items for ease of use with minimal strain and injury. They include empirical models, postural analysis tools, various software and standards specific to anthropometric data.

Human Computer Interaction: This category is for methods that are specific to the human computer interface. It includes information regarding design as well as methods for analysis of good computer interactions.

Data: This category is for any type of methods and tools that can help sort through various types of data sets and organize or analyze them specific methods include data mining tools and data analysis tools.

The results from the classification of the 400 methodologies show a diverse spread across each methodology, with the Human & System Performance Assessments representing a majority of the grouping.

B. Quantitative vs Qualitative Methodologies

Methodologies were divided by their output type showing whether it has quantitative or qualitative outputs. Quantitative outputs are those that have a number or value associated with the evaluation, while qualitative outputs were descriptors of the system often more subjective in nature. It should be noted that in certain methodologies, the operator or observer could define a quantitative output but it is still subjective in nature. For example, an operator's rating of workload is subjective, but can be ranked on a numerical scale. In such a case, the methodology was classified as having a quantitative output but with a subjective rating. Table 2 shows an even split between quantitative and qualitative methodologies.

Table 2. Number of quantitative vs qualitative methodologies

Output Type	# of Methodologies
Quantitative (QN)	201
Qualitative (QL)	202

The equal division between the output types could be an indicator of the multidisciplinary field in which there is a mix between the more descriptive social sciences and the more data driven engineering disciplines. To gather more insight regarding what methodologies tended to be more qualitative or quantitative, Table 3 divides the output types across the different methodology classifications. The darker green colors indicate which classification tended towards which output type.

Classifications that are strongly quantitative include modeling and simulation, data tools, and safety, while the more qualitative outputs appear from the knowledge elicitation tools or human factors knowledge tools. The results reflect the general intuition in which models and simulations produce numerical values based on data or algorithm driven results while the knowledge based tools

often are in the form of questionnaires or observers recording their own insights.

Table 3. Output type across classification.

Methodology Classification	Qualitative Output (% Total)	Quantitative Output (% Total)
Human & System Performance Assessment	7.4	11.2
Modeling & Simulation	0.0	14.1
Safety	2.0	11.9
Knowledge Elicitation	8.7	4.2
Human Factors Knowledge	10.9	0.0
Human Factors Program Planning	7.7	1.7
Physical Ergonomics	5.5	3.2
Human Computer Interaction	7.9	0.0
Data	0.0	3.5

C. Methodology Application in Systems Engineering Process

Methodologies were analyzed across the systems engineering process. Figure 3 shows the spread of methodologies in the different phases and where they are most often used.

From Figure 3 it can be seen that the majority of the methods and tools are used later in the systems engineering lifecycle. This can be interpreted as a factor of hardware design in which it is much easier to obtain direct user feedback and evaluation from an already designed and built system.

The result of the analysis illustrates the lack of methodologies early in the systems engineering phase. In addition, it is interesting to note that 60% of the methodologies available in the earlier design phases (Phase A through C) produce qualitative outputs.

D. Unique Attributes for Spacecraft Design

To determine which methodologies could potentially be applicable for human spacecraft design, a number of pertinent attributes were identified from a mix of literature review on the future of human spacecraft design as anticipated by FAA and NASA as well as insights on the human spacecraft design process itself. The following attributes that would be valuable for a potential human spacecraft design evaluation methodology are listed below with a brief rationale.

1) *Early Design Phase:* Due to the high costs of spaceflight, having an evaluative methodology early in the process will help to reduce downstream costs. The analysis shows a lack of methodologies in this early design phase.

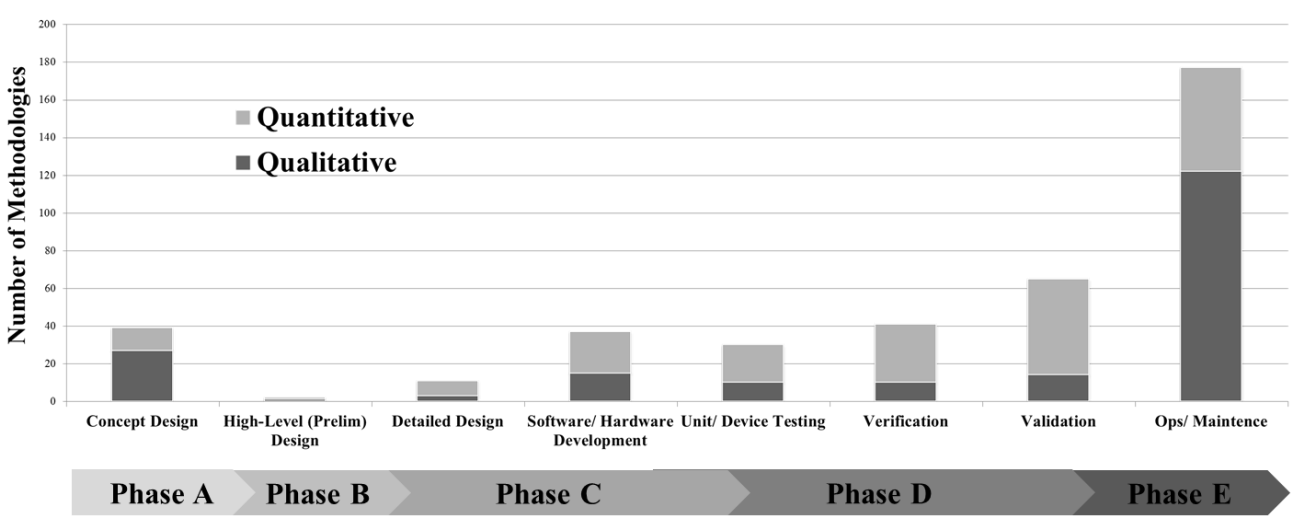
2) *Quantitative Output:* Having a quantified value for crew performance is necessary to make objective comparisons across various designs similar to what is done with metrics like mass, volume or power.

3) *State-Driven:* With the variety of upcoming human space missions whether for commercial or government, the model should be able to handle a range of mission types, human occupants, and task requirements.

A state-driven methodology is one that can be updated with a number of different mission profiles, users, and various tasks that could be performed. The methodologies that have this type of state-driven capability appear in the models and simulation type classification.

4) *Flexible Architecture*: Since human spaceflight is fairly young, there are unknown unknowns that are still to be discovered, therefore having a flexible architecture is crucial in adapting to new information that is learned with each new mission.

Figure 3. Methodology breakdown across systems engineering process.



A methodology that has a flexible architecture is one that can adapt to new information or poorly understood information.

5) *Comprehensive:* Each subsystem can be well-optimized for their required task or need, but when they are integrated into a full system, it is important to understand how each subsystem might impact others. The total system performance needs to be considered as an aggregation of subsystems which must perform well as a whole.

A comprehensive methodology integrates different aspects of the human spacecraft interactions. The purpose of the evaluation is to provide a high-level concept design comparisons, it must encompass all aspects of how the spacecraft design might influence the crew performance. Therefore such a methodology should include the quality of the habitation environment as well as the controllability of the spacecraft. It would entail understanding how the space is properly used as both a workspace and living area. To evaluate methodologies for "comprehensiveness" required a deeper understanding of each methodology and what information and foundational knowledge it contained.

From the first three filtering mechanisms (early design phase, quantitative, and state-driven), sixteen methodologies remained and are listed below in Table 4. Filtering for the last two attributes (flexible and comprehensive) required a more in-depth and subjective review of each methodology. After reviewing each of the remaining methodologies, five stood out as adaptable as a human spaceflight evaluation methodology, and have been identified on Table 4 with a thick dashed line.

Each of five methods is described with more detail in the following sections with the recommended extensible aspects that would be needed to create a more standardized method of evaluating human spacecraft design.

1. Integrated Performance Modeling Environment (IPME)

IPME is a Unix-based integrated environment of simulation and modeling tools for answering questions about systems that rely on human performance to succeed. It combines both the top-down approach of task network modeling with bottom-up approaches for simulating human behavior. It is interoperable with other models and external simulations. IPME has enhanced usability through a user friendly graphical user interface, and provides a full-featured discrete event simulation environment built on Micro-Saint modeling software. The following are specific functions it can perform:

• Allows the user to select from two different workload models

• Supports performance shaping function (PSF) approaches and built-in micro models

Table 4. Filtered methodologies for relevant human spaceflight attributes.

#	Method	Detailed Classification
1	Micro Saint (in MIDAS)	Task Network Model
2	IPME	Task Network Model
3	WinCrew	Task Network Model
4	IMPRINT	Task Network Model
5	ACT-R	Cognitive Models
6	MIDAS	Cognitive Models
7	D-OMAR	Cognitive Models
8	SAMPLE	Cognitive Models
9	Soar	Cognitive Models
10	GLEAN	Cognitive Models
11	Epic	Cognitive Models
12	APEX	Cognitive Models
13	CREAM	Error Models
14	Bayesian Belief Nets	Error Models
15	PARADyM	Error Models
16	Neural Networks	Error Models

• Uses built-in micro model functions that represent basic human actions and behaviors such as the rate at which text is read or the time to reach or move a motor control [20].

Getting access requires a license. . There is are also five day training courses available from time to time in Boulder, CO. This seems to be a flexible and extensible platform, but requires more contact with the company for access to software.

2. Improved Performance Research Integration Tool (IMPRINT)

IMPRINT is an embedded discrete event task network modeling language (Micro Saint) that uses task-level information to construct networks representing flow and performance time and accuracy for operations and maintenance missions. Workload profiles are generated so that crew-workload distribution and soldier-system task allocation can be examined. Manpower requirements estimates can be generated for a single system or Army-wide. Additionally, outputs can be used as a basis for estimating manpower lifecycle costs [21]. This system also uses Micro Saint as a backbone for the structure and would require a license for use.

3. Man Machine Integrated Design and Analysis System (MIDAS)

MIDAS is a dynamic, integrated human performance model (HPM) environment that facilitates design, visualization, and computational evaluation of complex human-system

concepts in simulated operational environments; symbolically represents many mechanisms that underlie and cause human behavior; it combines graphical equipment prototyping, dynamic simulation, and HPMs to reduce design cycle time, support quantitative predictions of human-system effectiveness, and improve design of crew stations and their operating procedures[22]. MIDAS consists of three components, the basic operator model, a Micro Saint Sharp task simulator, and an anthropometric model. While there is no direct cost in acquiring the MIDAS package, the anthropometric model requires specific software which would need to be purchased.

4. Cognitive Reliability and Error Analysis Model (CREAM)

CREAM is a human reliability analysis approach for probabilistic safety assessments. It is a stand-alone method that can be used for accident analysis and for larger design methodologies for more interactive systems. CREAM would allow for the designer to do particular tasks including:

"• identify tasks that require human cognition and therefore depend on cognitive reliability

• determine the conditions where cognitive reliability and ensuing risk may be reduced

• provide an appraisal of the consequences of human performance on system safety which can be used in PSA."[23]

The purpose of CREAM is to help designers evaluate how likely a particular design could induce performance errors by the operators. It provides a risk assessment to understand the error propagation throughout a system.

5. Performance and Reliability Analysis via Dynamic Modeling (PARADyM)

A model developed by Draper Laboratory to analyze human-in-the-loop performance. The model uses Matlab® and Simulink ® to model different components of a system capturing individual dynamics, failure modes, and any dependencies. One of the unique aspects of this tool is its use of Markov modeling to determine the systems reliability over a given mission timeline. Initial probabilities are set for various Markov states and are propagated to other states over the systems lifetime via specified transition rates based on the failure rates of the components [24].

The limitation for this system for adding the human-in-the-loop is similar to the CREAM model in that it lacks real data from human spaceflight mission failure rates. While error probabilities and failure rates of the human can and have been adapted from error rates found in nuclear power plant operations, it has not been thoroughly verified as a tool for spacecraft design.

5. DISCUSSION

There are a number of tools, methods, and frameworks available focused on human and complex system integration efforts. It is important to note that while a number of methodologies were gathered for this analysis, it does not represent a comprehensive list of every human performance evaluation methodology.

Through the process of collecting the methodologies, it became clear that there are no systematic or standard processes readily available to help find, much less, characterize and understand the methodologies that currently exist outside of intensive literature searches. In general, there is a need for a better database and collection of these various methodologies to be helpful for future designers of complex human systems.

Besides the process itself, the analysis from the collected data also provided a number of interesting insights on the type of methodologies available. The first analysis grouped methodology type into a specific classification. While useful as an exercise and as an early analysis tool, it can greatly vary depending on the definition and creation of different classifications. There may be another set of classifications that could help better distinguish the type of tools, methods, or frameworks in a more useful form such as field of application (aviation, nuclear power plant operations, or military operations.). Future analysis needs to include more meaningful ways to classify the methodologies.

The analysis also shows that there are few methodologies that exist or are used early in the concept design phase of the systems engineering process. The number of methodologies currently in use might be an indicator of both the difficultly of creating an early concept design methodology or that designers have not found them particularly useful for this early phase. Both these considerations must be addressed for adoption of future approaches.

Also evident from this work is that while there are a number of methodologies for human performance evaluation in complex systems, there still doesn't seem to be one that supports all the attributes that would be required for human spacecraft design. For example, the CREAM approach helps designers quantify error probabilities for one design, but it lacks flexibility in changing efficiently to another design without having to re-map errors and any coupling of factors.

Another option besides adapting and modifying existing methodologies is to create a new one from scratch. The creation of a new methodology can be a completely new framework, or it could be a merger of various other methodologies. The considerations for each must be thoroughly weighed, as it could be largely useless to add another methodology to the over 400 existing ones, but if it proves to be more beneficial and well-targeted to the human spacecraft designer, then the effort could be worthwhile.

6. CONCLUSION

Numerous tools, methods, and frameworks have been created to evaluate human integration into complex systems. These methods are often applied for specific uses and hard to adapt to new or different systems. Through a number of analyses on the various methodologies, five stood out as most applicable, extensible, and accessible for human spaceflight: IMPE, IMPRINT, MIDAS, CREAM, and PARADyM. While each of these methodologies has their specific limitations, a predominant concern for all of them includes the limited data regarding human performance issues in spaceflight.

As human spaceflight becomes more accessible and routine, there will be a continued push for better methodologies to ensure efficient and high-quality spacecraft designs. This work is a proactive step in developing the right tools, methods, and frameworks for future designers. Ultimately the goal of having better methodologies is to ensure the humans are protected, appropriately accommodated, and efficiently utilized by the spacecraft to achieve the best mission outcomes.

REFERENCES

[1] NASA (2014). "Human Integration Design Handbook (HIDH)", NASA/SP-2010-3407/REV1, June 5, 2014.

[2] NASA (2006), "Health and Medical Policy for Human Space Exploration" NPD-8900.5A, May 2006.

[3] NASA (2007a), "Design, Development, Test, and Evaluation (DDT&E) Considerations for Safe and Reliable Human Rated Spacecraft Systems", NASA Engineering and Safety Center Technical Report: RP-06-108 Vol 1, V.1. 1 May 2007.

[4] NASA (2007b), "NASA Space Flight Human System Standard Volume 1: Crew Health", NASA-STD-3001, March 2007.

[5] NASA (2008), "Health and Medical Requirements for Human Space Exploration", NPR-8900.1, May 2008.

[6] NASA (2011), "NASA Space Flight Human System Standard Volume 2: Human Factors, Habitability, and Environmental Health", NASA-STD-3001, January 2011.

[7] NRC (2007). *Human-systems Integration in the System Development Process*. The National Academies Press, National Research Council.

[8] NASA (2013). Evidence Report: Risk of Incompatible Vehicle/Habitat Design. Human Research Program Space Human Factors and Habitability Element. 8 March 2013.

[9] Tauri Group (2010). *Suborbital Reusable Vehicles: A 10-Year Forecast of Market Demand.*

[10] NASA (2011). "NASA Procedural Requirements (NPR) Human-Rating Requirements for Space Systems," NPR-8705.2B change 3, Nov 2011.

[11] FAA (2014). Recommended Practices for Human Space Flight Occupant Safety. Version 1.0. August 27, 2014. http://www.faa.gov/about/office_org/headquarters_offices/ast/news_announcements/media/Recommended_Practices_for_HSF_Occupant_Safety-Version1.pdf. Last accessed 1 Oct 2014.

[12] NASA (2007). NASA Systems Engineering Handbook. NASA/SP-2007-6105 Rev 1.

[13] Cohen, M. M. (2008), 'Testing the Celentano Curve: An Empirical Survey of Predictions for Human Spacecraft Pressurized Volume", 38th International Conference on Environmental Systems, San Francisco, CA, June 30-July 2, 2008

[14] Romero, J. J. (2007). "How Many People Have Been in Space?-asks Gail".By Joshua J Romero Posted March 26, 2007 on http://scienceline.org/2007/03/ask-romero-people_in_space/ Last Accessed 1 Oct 2014.

[15] Sandal, G. M., Leon, G. R., and Palinkas, L. (2006). *Human challenges in polar and space environments*. Rev Environ Sci Biotechnol (2006) 5:281-296.

[16] FAA (2014). The Annual Compendium of Commercial Space Transportation: 2013. February 2014.

[17] Gore, B. F., and Smith, J. D. (2006). 'Risk assessment and human performance modelling: the need for an integrated systems approach', Int. J. Human Factors Modelling and Simulation, Vol. 1, No. 1, pp.119-139.

[18] Chandler, F. T., Chang, Y. H., J., Mosleh, A., Marble, J. L., Boring, R. L., and German, D. I. (2006). Human Reliability Analysis Methods Selection Guidance for NASA. NASA/OSMA Technical Report. July 2006.

[19] Pew, R. W. (2008). More Than 50 Years of History and Accomplishments in Human Performance Model Development. Human Factors: The Journal of the Human Factors and Ergonomics Society. 2008 50:489.

[20] Fowles-Winkler, A. M. (2003). Modelling with the Integrated Performance Modelling Environment. Proceedings 15th European Simulation Symposium.

[21] U.S. Army (2010). United States Army Research Laboratory. http://www.arl.army.mil/www/default.cfm?page=445 Last accessed 20 May 2014.

[22] Gore, B. F. (2010). Man-machine integration design and analysis system (MIDAS) v5: Augmentations, motivations, and directions for aeronautics applications. In P. C. Cacciabu, M. Hjalmdahl, A. Luedtke and C. Riccioli (eds.) Human Modelling in Assisted Transportation. Heidelberg, Springer-Verlag Berlin Heidelberg.

[23] Hollnagel, E. (1998) Cognitive Reliability and Error Analysis Method (CREAM) 1998, New York: Elsevier Science.

[24] Bortolami, S. B., Duda, K. R., and Borer, N. K. (2009). Markov Analysis of Human-In-The-Loop System Performance. IEEEAC paper #1200, Version 1, updated December 29, 2009.

BIOGRAPHY

Christine Fanchiang *received her Bachelor's Degree in Aerospace Engineering from MIT and is now in the doctoral program in the Aerospace Engineering Sciences Department at the University of Colorado at Boulder with an emphasis in Bioastronautics. Her research focuses on defining an operability index for human-rating space vehicles to better understand the effects of spacecraft design on crew performance. She is currently funded through the NASA Harriet G. Jenkins Graduate Fellowship.*

Dr. David M. Klaus *received his BS in mechanical engineering from West Virginia University in 1984 and his MS and PhD in aerospace engineering sciences from the University of Colorado Boulder in 1991 and 1994, respectively. Between 1985 and 1990, he worked at the Kennedy Space Center in Florida and the Johnson Space Center in Houston as a Systems Engineer on the Space Shuttle Program. He spent 1994/95 as a postdoc at the German Institute of Aerospace Medicine (DLR) in Cologne on a Fulbright Scholarship. Dr. Klaus is currently an Associate Professor in the aerospace engineering sciences department at the University of Colorado Boulder. He is an AIAA Associate Fellow and CU President's Teaching Scholar, and has received various teaching and research awards including the ASGSB Thora W. Halstead Young Investigator's Award (2003); AIAA Rocky Mountain Section Educator of the Year Award (2004); CU Provost Faculty Achievement Award (2007); CU Boulder Faculty Assembly Excellence in Teaching Award (2007); and the CU Charles Hutchinson Memorial Teaching Award (2011).*

Dr. Jessica Marquez *received her BSE in mechanical & aerospace engineering at Princeton University and her MS and PhD in human systems engineering from MIT in 2007. Her research is in the Human System Integration Division at Ames. As a researcher, she develops tools for people who support human space exploration, including trainers, flight controllers, and astronauts. She previously supported the development and deployment of a variety of NASA planning and scheduling software tools for space operations, focusing on the International Space Station Program. Her research interests include human space exploration, human and automation-robotic integration, human-computer interactions, and Earth-analog missions.*

Dr. Brian Gore *is a Principal Investigator for San Jose State University in the Human Systems Integration Division at NASA Ames Research Center. Dr. Gore holds a PhD in Mechanical and Industrial Engineering from the University of Toronto and has 17 years of Human Factors and Ergonomics research experience in the surface transportation and aerospace domains using human-in-the-loop simulation and computational modeling methodologies. Dr. Gore holds membership in a number of professional organizations and standards committees and has taken on elected and lead roles within these organizations (HFES, SAE, IEEE, AIAA, SIGMA Xi). Dr. Gore currently leads a number of research projects aimed at increasing the fidelity of human performance models, and at predicting human-system performance in the aerospace domain.*