Experimental Validation of a Systems Architecting Framework for Objectives Definition in a Concurrent Engineering Environment

Alessandro Aliakbargolkar*(1), Edward F. Crawley(2)

(1) (2) *Massachusetts Institute of Technology*
77 Massachusetts Avenue, Cambridge MA – United States
email: (1) golkar@mit.edu (2) crawley@mit.edu

ABSTRACT

The definition of stakeholder objectives is a critical task within the pre-phase A of space missions and programs. Effective systems architecting requires comprehensive analysis of the value proposition of intended engineering developments. Oftentimes, objectives are ambiguous or unclear, or defined without support of systematic analysis – being rather driven by non-stated intents. Ambiguity in objectives stems from the uniqueness of each mission and the emergence of multiple and often conflicting views on objectives between stakeholder groups: said tensions need to be represented and analyzed in concurrent design processes. This paper presents the experimental validation of a Delphi-structured systems architecting framework that has been developed at MIT to support system architects in the analysis of stakeholder needs towards the development of systems architectures. The proposed multidisciplinary framework allows engineers to assess the impact of ambiguity in a systems architecture within a concurrent engineering session. The framework supports Program Managers and Principal Investigators in reducing ambiguity in their objectives, and consequently identifying architectures with optimal and resilient programmatic trade-offs, stakeholder objectives and engineering performance. The experimental validation study involved 45 graduate students in a concealed retrospective analysis of the business case of satellite constellations for commercial comunications. The study compares decision-making outcomes obtained through conventional group decision-making with results obtained using the proposed Delphi-based systems architecting framework. The results of the study show statistically signicant differences in the quality of resulting decision-making, thereby strengthening the case for the novel approach. The method is proposed for implementation in concurrent engineering centers and user workshops aimed at defining stakeholder objectives for new space missions.

I. INTRODUCTION

Concurrent design infrastructures are increasingly being used in Europe and in the United States to conduct feasibility studies of new missions, identify efficient mission architectures, and allow mission scoping in reduced time frames compared to conventional "over the fence" design practice [1]. When new space missions are defined, concurrent design infrastructures are used in the definition of the overall objectives the mission is supposed to meet, hence defining how the mission delivers value to its stakeholders. Within a concurrent engineering study, system architects define the value delivery process of a mission identifying and interviewing stakeholders, trying to understand their needs and encoding them into mission objectives [2]. Tradespace exploration is then performed to assess potential design solutions meeting, and characterize them using performance, cost, risk, and other evaluation metrics of interest. Typical concurrent design studies perform stakeholder analysis in the form of requirements analysis, where mission stakeholders and design engineers meet in plenary meetings to scope the mission and define overall mission requirements [3].

While plenary meetings and direct stakeholder interactions are the norm in industrial practice, several issues might occur in the process:

1) **Peer pressure effects:** Direct, unstructured interactions are subject to peer pressure, where some stakeholders prevail in the discussion and drive requirements off-target from actual perceived stakeholder needs;

* **Now Assistant Professor, Skolkovo Institute of Science and Technology, Moscow (Russia) – golkar@skolkovotech.ru**

2) **Requirements ambiguity:** Stakeholders elicit their needs ambiguosly, where ambiguity is driven by uncertainties on future budget, political instability in mission support, low technology maturity, lack of mission heritage, requirements creep (the *wants* versus the *needs*) and/or dissenting stakeholder opinions;

3) **Adverse negotiation externalities:** Stakeholders negotiate contrasting requirements, such as perceived science or performance needs of a mission, and the need to keep the mission on budget and on schedule. Negotiations might lead to sub-optimal solutions when stakeholders cannot afford to consider all available options either for lack of time or resources to invest;

4) **Hidden agendas:** Hidden agendas carried on by stakeholders might threaten a cost-efficient definition of mission requirements (*intended* requirements versus *stated* requirements);

5) **Lack of decision acceptance:** consensus on needs or requirements is not reached among stakeholders, hindering acceptance – thereby funding and political support – for further development of the mission concept.

Concurrent design facilities embody significant yet unexploited potential in including all stakeholders in efficient definition of mission requirements – stakeholders including policymakers, investors, managers, and so forth. So far, concurrent design was focused on engineering stakeholders: the designers and technical managers of a mission. In this paper we argue that concurrent design can be effectively extended to include all actors of a space mission in the design process, using a structured approach. Academic researchers have proposed structured approaches for stakeholder analysis in systems architecting [4] and tradespace exploration [5], that are suitable for concurrent engineering environments. However, these methods do not consider explicitly the abovementioned issues of peer pressure effects, requirements ambiguity, negotiation externalities, hidden agendas, and lack of decision acceptance. These aspects are traditionally left to social science research, while this paper argue that systems engineers – and in particular system architects – need to be aware of those and proactively include them in design iterations.

This paper address this idea presenting a Delphi-Based Systems Architecting Framework (DB-SAF) for systems architecting under ambiguous stakeholder requirements [2], suitable for implementation in a concurrent engineering environment. The present paper covers only distinctive features of DB-SAF – a more comprehensive description can be found in previous publications [6]. The main focus of this paper is instead to demonstrate that DB-SAF is effective in reducing peer pressure, mitigating requirements ambiguity and negotiation externalities, uncovering potential hidden agendas and improving decision acceptance. This goal has been achieved by developing an experimental validation study, conducted with 45 graduate students at MIT, on a retrospective concealed case analysis of the first generation of the Iridium satellite constellation[1]. DB-SAF is a structured approach for integrated requirements elicitation and computational tradespace exploration, with stakeholders and design engineers involved together in design iteration loops. The remainder of this paper is structured as follows. Section II provides an overview of the main features of the DB-SAF approach. Section III describes the experimental validation study that has been developed to demonstrate the effectiveness of DB-SAF, including discussion of experimental hypotheses, the description of the study, and presentation of results including their statistical validation. Section IV summarizes conclusions from the experimental study and defines avenues of future research.

II. DELPHI-BASED SYSTEMS ARCHITECTING FRAMEWORK (DB-SAF) APPROACH

This section describes the Delphi-Based Systems Architecting Framework (DB-SAF), and discusses its proposed implementation in a concurrent engineering environment. Further discussion of DB-SAF is available in previous publications [2, 6]. An overview of DB-SAF is shown in Figure 1. A detailed description of DB-SAF can be found in [2], while a DB-SAF implementation in a human spaceflight case study can be found in [6]. The key in the DB-SAF approach is the structured negotiation among stakeholders. Once all stakeholders are interviewed, opinion distributions are derived for all panels, and organized in boxplots as shown in Figure 2.

[1] Disclaimer Statement: This study was created for educational and academic research purposes only. In no way do the statements made in this study express official positions of the Massachusetts Institute of Technology.

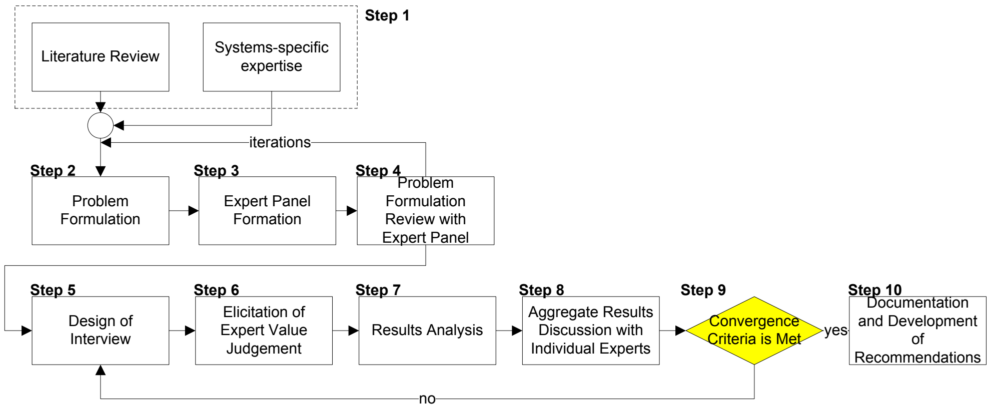

Figure 1 Delphi-Based Systems Architecting Framework (DB-SAF) Approach Overview

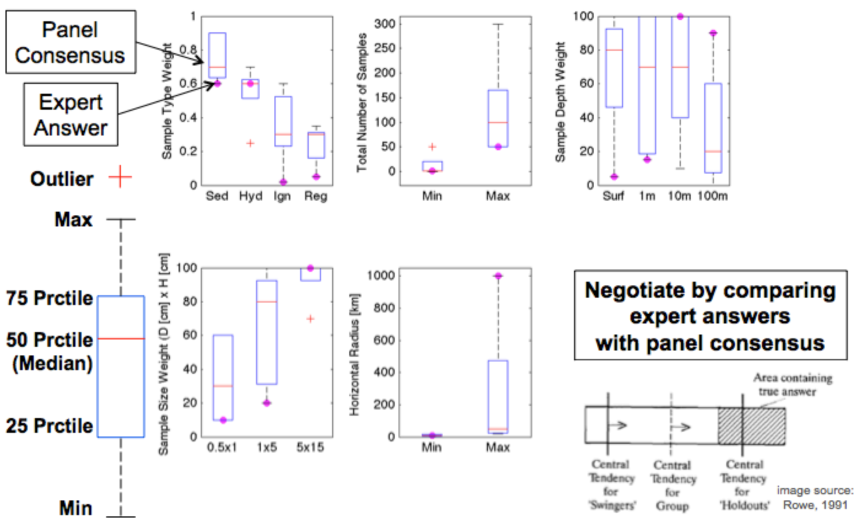

Figure 2 Delphi-Based Systems Architecting Framework - Negotiation Approach

Boxplots in the figure are to be interpreted as follows. The red line is the median score for each requirement attribute, while blue box edges define the 25th and 75th percentile of answers, respectively. Black whiskers represent maximum and minimum scores, while red crosses identify outliers. Results are then overlaid to individual scores provided by each expert, represented by pink dots. Iteration interviews are then conducted with experts to revise answers provided in the previous round in light of the opinion distributions. Experts are free to either change their answer or leave it unmodified – comments and motivations are recorded for each expert to provide rationale. Said rationale is then exchanged in anonymous form between experts in successive iterations. The goal is to reduce opinion spreads for each requirement attribute to obtain consensus, or identify areas of open debate – or potential hidden agendas that can be uncovered by post-analysis of expert comments. When convergence is called in iterations, median values are used to define value functions that can be used in tradespace analysis along with traditional performance, cost, and risk evaluation metrics. Tradespace analysis is not discussed in this paper; the reader is referred to [2] for further discussion of DB-SAF.

As described in this paper, DB-SAF is readily implementable in a concurrent engineering environment to support early phases of requirements definition and mission formulation. The question arises on whether DB-SAF improves the outcomes of said processes, when compared to the consolidated unstructured practice. Section III describes an experimental case study that has been conducted at MIT, which results suggest the proposed tool is effectively able to improve decision-making – and identifies inherent limitations in the process.

III. EXPERIMENTAL VALIDATION: IRIDIUM CASE STUDY

This section describes the validation case study that has been developed to evaluate DB-SAF and compare its performance with conventional Group Decision-Making (GDM). The purpose of the validation is to investigate the hypothesis that DB-SAF yields to better decision-making than what could be obtained with a Group Decision-Making (GDM) process. The validation consists in an analog case study based on a historical venture that is widely regarded as a systems architecting failure. The analog case is provided as an input to an experimental study. For the purposes of this paper the focus is on validating the requirements elicitation part of DB-SAF. The effectiveness of underlying systems architecting model for tradespace exploration has already been discussed ubiquitously in the literature by benchmark with existing systems (see [7] as an example). The goal is to verify whether reduction of ambiguity, achievement of compromise, elicitation of open debate, and identification of hidden agendas occur more effectively using DB-SAF rather than a GDM expert elicitation process.

Validation Hypotheses

The validation is designed to test the following hypotheses:

Hypothesis 1: The average perceived effectiveness of DB-SAF in reducing adverse peer pressure effects is greater than the average effectiveness achieved by GDM among team members of a systems architecting team.
Hypothesis 2: The average perceived effectiveness of DB-SAF in identifying and reducing ambiguity is greater than the average effectiveness achieved by GDM among team members of a systems architecting team.
Hypothesis 3: The average perceived effectiveness of DB-SAF in improving Pareto-efficient compromises is greater than the average effectiveness achieved by GDM among team members of a systems architecting team.
Hypothesis 4: The average perceived effectiveness of DB-SAF in eliciting open debate is greater than the average effectiveness achieved by GDM among team members of a systems architecting team.
Hypothesis 5: The average perceived effectiveness of DB-SAF in uncovering hidden agendas is greater than the average effectiveness achieved by GDM among team members of a systems architecting team.
Hypothesis 6: The average effectiveness of DB-SAF in improving decision acceptance among group members is greater than the average decision acceptance achieved by GDM.

Case Study Introduction

Iridium is a constellation of satellites that was designed to enable space-based personal communications. Iridium was initially built and developed by Motorola, leading company in the telecommunications industry. After nurturing the project for a decade and raising capital for a total of $6 billion dollars between debt financing, shareholders and Motorola's own investment, Iridium "solved a problem that very few customers needed solved" [8]. In 1999 Iridium filed for Chapter 11 bankruptcy, becoming one of the top 20 bankruptcies in the U.S. . Researcher identified the main reasons of Iridium's collapse to the following elements in unforeseen increase of capability of land-based cellular infrastructure, inherent engineering limitations, schedule delays and poor partner support [9].

Case Study Validation Question

The Iridium project consisted of two stages: a research and development stage (1987-1996), and a manufacturing and operations stage (1996-1999) [8]. This study analyzes Iridium in 1996, when the corporate board made the decision of transitioning from one phase to another. The following question is formulated:

Could ambiguities in Iridium's value proposition be identified and characterized in 1996 by supporting decision-makers with a Delphi-based systems architecting approach?

An analog case study has then been developed. In this study, a system architecting team is called to "make or break" the case of Suborbital Spacelines LLC, a fictional aerospace venture willing to launch a point-to-point suborbital transportation service. The following section introduces this study, and shows how the case has been developed using Iridium's data in a concealed context.

Analog Case Study – Suborbital Spacelines LLC

Suborbital Spacelines LLC is a new joint venture effort in the aerospace industry to deliver a commercial point-to-point passenger transportation service. Suborbital has conceived and designed a revolutionary concept for a spaceplane able to carry 100 passengers from London to New York in less than one hour. System engineers, business analysts and policy experts at Suborbital worked together to ensure commercial viability of their concept. Following several internal systems architecting studies and commissioning external consulting companies to conduct market surveys and market sizing estimates, Suborbital managers decided to form a steering committee to analyze the Company's business case and take a final decision to authorize the transition from detailed design to manufacturing spaceplanes and operating them. The steering committee includes three panels for Business, Technology and Policy matters. Each panel is staffed by junior and senior experts, in addition to panel leads. Suborbital's Chief Executive Officer will chair the panel and will hold the final decision on Suborbital's transition to manufacturing and operations.

Validation Protocol

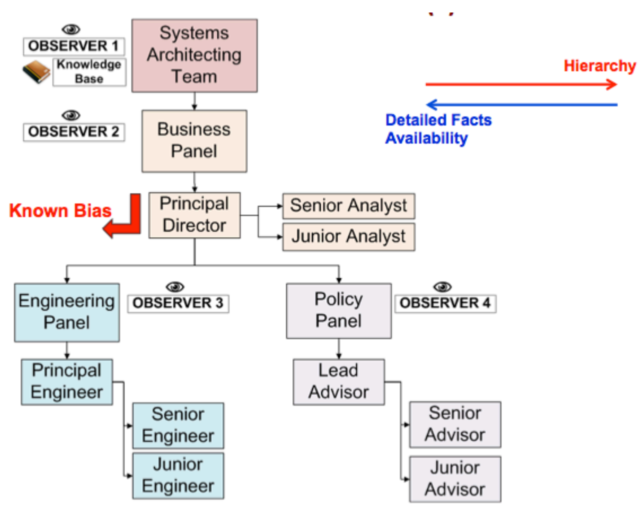

Figure 3 Validation Protocol

A validation protocol (Figure 3) has been developed and applied to the analog case study. In an experiment session, two independent teams have been supplied identical data and specific questions to answer. Team A has been instructed to approach the problem according to a GDM process, while Team B has been engaged in a structured process using DB-SAF.

Validation Inputs

A fact database to inform Suborbital's case study has been developed as set of inputs for the validation. An excerpt from the facts database is shown in Figure 4. Data included only information that was available at the time of the decisions in the historical case. Each fact is documented with a reference source from the literature. Facts have been characterized as either positive or negative biases, and as pertinent to business, engineering and policy. Facts have been evenly spread across said categories.

Fact	Literature Source	Iridium Satellite Constellation	Suborbital Spacelines LLC
Leading-edge technical feat to fulfill a need of an uncertain customer base	(Kerzner 2009)	First time satellite constellation is conceived, developed and operated. First time no-delay "anytime, anywhere" voice communications available to market. However, as the system is the first of its kind, there is no analog example to estimate a customer base for its utilization.	First time suborbital flight service is conceived, developed and operated. First time <1hr transoceanic flights available to market. However, as the system is the first of its kind, there is no analog example to estimate a customer base for its foreseen utilization.

Figure 4 Excerpt from facts database (source: [2])

Fact lists reproduced team work need across team panels by supplying to participants a subset of facts that were relevant to their domain of expertise only (i.e. expert domain-relevant facts). Peer pressure was simulated in Team A by giving influence authority to people with higher hierarchial positions, while instructing participants at lower organizational levels that hierarchy should have been respected during the GDM process. Team B was peer pressure free by virtue of anonimity of the DB-SAF process. Ambiguity has been introduced "ambiguating" fact sheets, by fuzzifying numbers, randomly concealing pieces of information, and so forth. Hidden agendas were introduced by providing specific additional instructions only to a small number of members within teams, intentionally pushing teams towards suboptimal choices that would favor individual on team-wide. The number of hidden agendas-related facts was directly proportional to the hierarchical level of each participant. This principle sets an adverse condition that biases the team towards the selection of the "wrong" answer, that determined the failure in the historical case..

Validation Process

Participants in both GDM and DB-SAF sessions were briefed with an introductory presentation with a review of the scopes of the study and with general guidelines to follow through the simulation. The study was presented to participants as a systems architecting role play. Participants were told to participate in a 1 hour study to make a go / no go decision on transitioning Suborbital from its development to its operational phase. Every participant was given a general information package with simulation and specific role information with technical, business or policy decisions according to the criteria stated above. DB-SAF participants were given 3 iteration rounds to elicit information and negotiate among teams. Following the iteration rounds, the Principal Director was given the overall information and asked to make a decision. In the GDM setting, the Principal Director led a 1 hour group discussion by asking team members to perform panel-specific iterations and therefore report to the overall committee. At the end of the session the Principal Director was called to make a decision. Participants were called to fill the evaluation survey that formed the basis for statistical analysis. It is important to note that the study was not meant to assess the group as the unit of analysis; rather the focus was on individual participants as units of analysis.

Results Analysis

Figure 5 shows raw aggregate results from DB-SAF and GDM sessions. The histograms show the distribution, while the vertical dashed lines show the median values for both the DB-SAF and GDM subsets. Medians have been preferred to means being more suitable for the analysis of small samples and are more robust to the presence of outliers. A preliminary analysis shows that DB-SAF median results are greater than GDM median response in mitigating peer pressure (hypothesis 1), reducing ambiguity (hypothesis 2), achieving better compromise (hypothesis 3) and better decision acceptance (hypothesis 6). DB-SAF scored less than GDM in eliciting open debates (hypothesis 4), and achieved a tie in perceived ability of uncovering hidden agendas (hypothesis 5). In the study sessions that have been conducted, all GDM sessions decided for Suborbital's transition to operations. All DB-SAF sessions decided to delay the decision to gather more data on the case. While being an interesting result, further experimentation would be required to test its statistical significance. In this case, the unit of analysis is the group rather than the individual. Such result should be considered as a tendency and not as statistical significant. Still, it is an observation of interest as it proves how final consensus (as measured by acceptance of the final decision by individual experts) does not always lead to optimal results. These preliminary results seem to suggest that a structured tool such as DB-SAF achieves its goal more effectively than an unstructured tool such as conventional GDM.

Nevertheless, DB-SAF encountered limitations that surfaced during the study. The main drawback that has been encountered in DB-SAF is that of constraining open debate by forcing experts to share their knowledge through a structured approach with no inter-expert communication. Anonimity allows mitigation of peer pressure and availability of confirmatory and disconfirmatory results enable reduction of ambiguity and achievement of compromise. The benefits of open communication are lost in DB-SAF. While the analysis of raw data provided insights on DB-SAF, a more robust approach is required to assess claims more confidently. We use statistical significance as a tool to do so. The question of interest is whether differences in median results between DB-SAF and GDM related scores are statistically significant, and if so, to what level of confidence. Two-way significance tests are conducted in this paper.

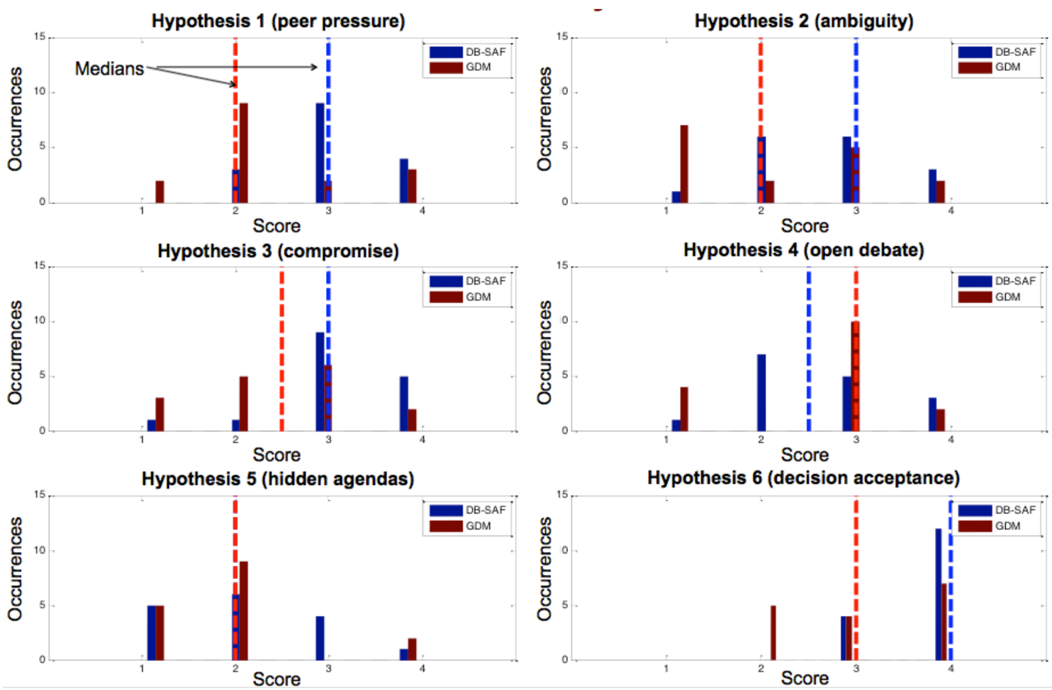

Figure 5 Validation Study Results

Two tests for statistical significance have been employed. The first one is the Wilcoxon rank sum test [10], a non-parametric test that does not require data normality. The other approach that has been employed for statistical significance testing is the bootstrap permutation test [10]. These methods are particularly useful to make statistical inference on populations with skewed distributions for which the normality assumption does not hold. Table 1 shows the results of the Wilcoxon rank sum test and for bootstrap permutation tests for the six hypothesis under scrutiny. The two populations (DB-SAF and GDM data) are compared by differences in the median. The null hypothesis is rejected for hypothesis 1 (mitigate peer pressure), hypothesis 3 and hypothesis 6 at the 95% confidence level, and hypothesis 2 at the 85% confidence level. The null hypothesis cannot be rejected for hypothesis 4 (elicit open debate) and hypothesis 5 (uncover hidden agenda) at reasonable confidence levels (above 80%). Failure in rejecting the null hypothesis for hypothesis 4 and 5 is confirmation of advantages and disadvantages of DB-SAF discussed previously, and further confirmed by participants' feedback. The results from the bootstrap permutation tests confirm the findings obtained with the preliminary inspection of results and the Wilcoxon rank sum test. Results show improved peer pressure mitigation, reduction of ambiguity, achievement of compromise and decision acceptance with DB-SAF as perceived by individual experts.

Table 1 Statistical Significance Tests Results

Wilcoxon Rank Sum Test	Confidence level	Bootstrap Permutation Tests	Confidence level
Hypothesis 1 (peer pressure)	95%	Hypothesis 1 (peer pressure)	98%
Hypothesis 2 (ambiguity)	85%	Hypothesis 2 (ambiguity)	92%
Hypothesis 3 (compromise)	95%	Hypothesis 3 (compromise)	97%
Hypothesis 4 (open debates)	20%	Hypothesis 4 (open debates)	43%
Hypothesis 5 (hidden agendas)	35%	Hypothesis 5 (hidden agendas)	58%
Hypothesis 6 (decision acceptance)	95%	Hypothesis 6 (decision acceptance)	99%

IV. CONCLUSIONS AND FUTURE WORK

This paper introduced the DB-SAF framework as a structured approach for requirements definition, and proposes it for implementation in a concurrent engineering design environment. The validity of DB-SAF is supported by the results of an experimental validation case study. Validating a decision-support tool focusing on stakeholder ambiguities is a challenge, given all the exogenous factors and behavioral phenomena that need to be considered. A concealed simulation approach has been devised to validate DB-SAF. An experimental pilot study has been designed and DB-SAF results have been compared with traditional GDM methods.

The validation uncovered strengths and limitations of the framework. DB-SAF proved being effective in mitigating peer pressure, reducing ambiguities, achieving compromises and improving decision acceptance. At the same time, the validation gave the opportunity to highlight intrinsic method challenges in eliciting open debates and uncovering hidden agendas. DB-SAF achieves beneficial results by minimizing and anonymizing inter-expert interactions. Conventional GDM, is prone to suffer from adverse effects such as peer pressure and ambiguity. On the other hand, direct expert interaction facilitates resolution of ambiguities. This tension was reflected in comparing feedback from study participants. The resulting trade-off highlights how the selection of an appropriate decision-making tool depends on the goals to be achieved by the system architects. If peer pressure, ambiguity and hidden agendas are primary concerns, then DB-SAF is probably an effective choice for decision making. If open debate and high expert interaction is preferred, GDM offers leeway to do so, at the expense of exposure to adverse effects as discussed in this chapter.

Careful analysis needs to be conducted to extend the generalization of these results. One should consider the temporal dimension in decision-making. Group exercises of knowledge elicitation of any fashion are subject to change with time and participants. The same group of experts could reach opposite conclusions on a systems architecting problem if the meeting is held in different times of the year. Consider for instance the outcomes of a systems architecting team meeting of a human spaceflight mission before and after a budget decision by the White House. Likewise, different groups of experts could reach different conclusions due to unforeseen bias unbalance among the group. This limitation is minimized by have a larger group size while intentionally incorporating multiple perspectives on the problem of interest.

In addition to method-specific challenges, one must also consider limitations in the evaluation of validation-specific results. Extreme care is required to maximize the likelihood of success in reproducing subjective factors such as peer pressure, hidden agendas and ambiguity are hard to reproduce in a controlled setting. Reliability, internal consistency and external validation need to be primary concerns. Challenges include the difficulty of the target population reproducing with high fidelity – in this case, high-level decision makers and senior experts include the challenges. Participants may not be interpreting their role correctly in a behavioral simulation setting. Observers have been used to mitigate this risk.

This validation showed that one of the most challenging elements to reproduce were hidden agendas, as individuals had limited time to digest the information and to maturate a highly biased side in the discussion.

Furthermore, cultural and national difference are exogenous factors that play important roles in subjective perception of leadership styles. What could be perceived as peer pressure, for instance, in a European meeting, might not in fact be perceived in equal manner in a US environment, and vice versa.

Lastly, exposing people to new problems requires simplification. Simplification is traded against realism in a case study. A validation case study must be of analog complexity to "real" case studies in order to maintain relevance to its scopes. Notwithstanding challenges, this validation showed strong tendencies in support of major claims in the DB-SAF framework. Its implementation in concurrent design facilities is encouraged to support stakeholder analysis and requirements definition, thus enhancing systems architecting and pre-phase A / phase A studies of new space missions.

V. REFERENCES

1. Bandecchi, M., Melton, B., and Ongaro, F. "Concurrent Engineering Applied to Space Mission Assessment and Design," *ESA Bulletin* Vol. 99, No. Journal Article, 1999.
2. Aliakbargolkar, A. "A Framework for Space Systems Architecting under Stakeholder Objectives Ambiguity," *Aeronautics and Astronautics*. Vol. Ph.D., Massachusetts Institute of Technology, Cambridge, MA, 2012.
3. Agency, E. S. "The ESA Concurrent Design Facility: Concurrent Engineering applied to space mission assessments." 2010.
4. Cameron, B. G., Crawley, E. F., Geilson, L., and Rebentisch, E. S. "Value flow mapping: Using networks to inform stakeholder analysis," *Acta Astronautica* Vol. 62, No. 4-5, 2008, pp. 324-333.
5. Ross, A. M., and Hastings, D. E. "The Tradespace Exploration Paradigm," *INCOSE 2005 International Symposium, Rochester, NY, July*. 2005.
6. Aliakbargolkar, A., Rudat, A., and Crawley, E. F. "A Delphi-Based Framework for Systems Architecting of In-Orbit Exploration Infrastructure for Human Exploration beyond Low Earth Orbit," *International Astronautical Congress*. International Astronautical Federation, Naples, Italy, 2012.
7. De Weck, O., De Neufville, R., and Chaize, M. "Staged Deployment of Communications Satellite Constellations in Low Earth Orbit," *Journal of Aerospace Computing, Information, and Communication* Vol. 1, No. 3, 2004, pp. 119-136.
8. Kerzner, H. *Project management : a systems approach to planning, scheduling, and controlling*. Hoboken, N.J.: John Wiley & Sons, 2009.
9. Finkelstein, S., and Sanford, S. H. "Learning from corporate mistakes: The rise and fall of iridium," *Organizational Dynamics* Vol. 29, No. 2, 2000, pp. 138-148.
10. Moore, D. S., McCabe, G. P., and Craig, B. *Introduction to the Practice of Statistics*: W.H. Freeman, 2007.