2019 NDIA GROUND VEHICLE SYSTEMS ENGINEERING AND TECHNOLOGY SYMPOSIUM
SYSTEMS ENGINEERING TECHNICAL SESSION
AUGUST 13-15, 2019 - NOVI, MICHIGAN

UTILIZING A STRUCTURED DECISION PROCESS FOR U.S. GOVERNMENT SOURCE SELECTIONS: CASE OF A COMBAT VEHICLE ENGINE

Gregory D. Hartman, PhD[1]

[1]US Army Combat Capabilities Development Command, Ground Vehicle Systems Center, Detroit Arsenal, MI

ABSTRACT

U.S. Government procurement spending exceeds $500B annually. A request for proposal is one of the more common forms of solicitation, and source selection (SS) is the process for evaluating proposals submitted by contractors. The U.S. Department of Defense and the Army promulgate manuals and supplements that direct the SS process within those organizations. Those publications identify "trade-offs" as a preferred method for conducting a SS, and encourage the use of this process "when it may be in the best interest of the Government to consider award to other than the lowest-price offeror." Under this process, cost and non-cost factors are evaluated and the contract is awarded to the offeror proposing the combination of factors that represents the best value based on the evaluation criteria. This case study will describe how a trade-off, or structured decision, process was used to support a U.S. Army SS by thoroughly evaluating multiple vendors and their proposals of a major subsystem for a major defense acquisition program. The purpose of this case study is not to focus solely on how to accomplish a trade-off, or execute a SS, but rather to share lessons-learned about how to address special situations encountered during a SS trade-off.

Citation: Hartman, G., "Utilizing a Structured Decision Process for U.S. Government Source Selections: Case of a Combat Vehicle Engine", In *Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium* (GVSETS), NDIA, Novi, MI, Aug. 13-15, 2019.

1. INTRODUCTION

U.S. Government (USG) procurement spending exceeds $500 billion annually, with purchases ranging from major weapon systems to janitorial services. Thousands of contract solicitations are posted annually on the U.S. Government's Federal Business Opportunities website; 79,985 solicitations were posted in 2014 alone.[1] One of several methods that the government may use to solicit proposals and award contracts is negotiated procurement. Contracting by negotiation permits the government to award a contract based on factors other than price. In a negotiated procurement, the

DISTRIBUTION A. Approved for public release; distribution unlimited. OPSEC #2443.

government uses full and open competition and issues a Request for Proposal (RFP); one or more contractors (suppliers) submit a proposal in response to the RFP. The agency that issued the RFP can engage in discussions with offerors about their proposals. The agency awards a contract to the offeror that provides the best value, taking into account cost and all other factors specified in the RFP, such as achieving technical requirements, quality, and past performance. An RFP is one of the more common forms of solicitation, and source selection (SS) is the process for evaluating proposals submitted by contractors. As posted on the U.S. Government Accountability Office (GAO) website, "If a party interested in a government contract believes that an agency has violated procurement law or regulation in a solicitation for goods or services, or in the award of a contract, it may file a bid protest" with the GAO. The GAO, an independent, nonpartisan agency that works for the U.S. Congress, is tasked with investigating how the federal government spends taxpayer dollars. One of its specific duties is issuing legal decisions and opinions pertaining to solicitation bid protest rulings. Based on the GAO Bid Protest Annual Report to Congress for Fiscal Year 2016, there were more protest cases filed for that year than any year since 2001. GAO's review revealed that "the most prevalent reasons for sustaining protests during the 2016 fiscal year were: (1) unreasonable technical evaluation; (2) unreasonable past performance evaluation; (3) unreasonable cost or price evaluation; and (4) flawed selection decision."

The U.S. Department of Defense (DoD) and the Army promulgate manuals and supplements that direct the SS process within those organizations. Those publications identify "trade-offs" as a preferred method for conducting a SS, as emphasized by the following excerpt: "On most acquisitions, the trade-off process will be most effective and will result in the best value to the Government. Use this process when it is in the Government's best interest to consider award to other than the lowest price offeror. Under this process, you evaluate both cost (or price) and non-cost factors and award the contract to the offeror proposing the combination of factors that represents the best value based on the evaluation criteria. Inherent in this process is the necessity to make trade-offs considering the non-cost strengths and weaknesses, risks, and the cost (or price) offered in each proposal."[2]

The above paragraphs emphasize the importance of ensuring that a sound and traceable evaluation process is followed for a DoD SS, and that the trade-off process is recognized as a highly-regarded solution. This case study will describe how a trade-off, or structured decision (SD) process was used to augment a specific U.S. Army SS by thoroughly and soundly evaluating multiple vendors and their proposals of a major subsystem for a major defense acquisition program. The purpose of this case study is not to focus solely on how to accomplish a SD, or how a SS is executed, but to share lessons-learned regarding how to address special situations that may be encountered while carrying out a SD. In particular, this case study will explore the following: 1) how to assign relative importance weighting when an upper-level constraint has been imposed, 2) how to establish a criterion rating scale that avoids any alternative receiving zero utility with respect to the criterion, and 3) how to mask the identity of vendors and respective proposal alternatives when necessary. Additionally, this case study includes key characters who were involved in the SS, though their names have been disguised.

2. SITUATION

During 2002 – 2009, the Future Combat System (FCS) acquisition program was the centerpiece of the U.S. Army's transformation to a lightweight, rapidly-deployable and lethal network-centric force. Manned Ground Vehicles (MGVs) were the cornerstone of the Army's FCS Brigade Combat Team (FBCT), as they were to host at least 75 percent of the FBCT lethality and ground sensor capability [1]. Boeing and Science Applications International Corporation (SAIC) were the Lead System Integrators (LSIs), and General Dynamic Land Systems (GDLS) and United Defense Limited Partnership (UDLP; name later changed to BAE Systems Land and Armaments[3]) were the MGV integrators. One major subsystem to be integrated was the engine; one common engine was to be selected for all eight planned MGV variants. In 2004, GDLS worked with the USG, LSI, and UDLP, to prepare and release an RFP to initiate the MGV engine selection. Prior to release of the RFP, market research had been conducted to determine what vendors and technologies existed that could potentially meet the engine requirements. The ensuing process to identify the most preferred engine proposal is described below.

3. THE STRUCTURED DECISION PROCESS OVERVIEW AND HOW IT AUGMENTS SOURCE SELECTION

GDLS employs a rigorous SD process that allows for complex or strategic trade-off decisions to be made at the subsystem and system level by utilizing a cross-functional team facilitated by a decision and risk analysis (DRA) team member. Although this process description is focused on military vehicle applications, it is flexible enough to be applied to almost any subsystem, system or system-of-systems (SoS),[4] such as FCS. Decisions at the subsystem level can include selection of software packages, weapons, equipment, and resources; system-level decisions can include a communications network, fleet of vehicles or watercraft, and port (air, land, sea) security concepts. The SD process can also be applied to decisions that are less complex or of less strategic importance. Generally, these decisions do not require formal facilitation and can be conducted entirely by the decision maker (DM). For these informal decisions, the DRA Team offers an in-house-developed Desktop Decision Tool that guides the DM through identifying (1) the purpose of the decision, (2) needs or requirements that must be met, (3) evaluation criteria, and (4) solution alternatives. For formal decisions, as illustrated in this case study, the same process is facilitated by an experienced and trained DRA team member who utilizes commercial-off-the-shelf software (Logical Decisions for Windows (LDW)), to document and provide assessment results.

Appropriate subject matter experts (SMEs) and stakeholders participate in the entire process. Usually, key customer representatives are identified as stakeholders and may include the DM. Once requirements are verified and understood, appropriate evaluation criteria based on those requirements are defined. Technology parameters and subsystems are then described and options for each subsystem identified. SMEs rate the subsystem options with respect to each evaluation criterion. Based on the evaluation criteria rating data, relative importance weights (RIWs), and single-attribute utility functions (SUFs), LDW delivers a ranking of the alternatives based on utility, and facilitates sensitivity analysis with respect to the RIWs, enabling the stakeholders to decide on a preferred solution.

The following steps outline the SD process that was employed for the MGV Engine SS:

1. Establish the SD framework,
2. Refine evaluation criteria and SUFs,
3. Characterize proposal alternatives,
4. Establish RIWs for evaluation criteria, and
5. Select the preferred alternative.

3.1. Establish the SD Framework: Identify Members and Evaluation Criteria, and Establish Importance Hierarchy

For this SS, the GDLS Program Manager was the SS Authority (SSA) and the SD owner. The SS evaluation was separated into four categories: technical, cost, management, and risk. A SS Evaluation Board was established by assigning a panel for each category other than risk; each of the risk subcategories was assigned to the respective category panel to assess functional risk per each alternative: technical risk to technical panel, cost risk to cost panel, and schedule risk to management panel. The risk assessment from each panel would be integrated into the risk goal in LDW. Figure 1 displays the organizational chart for the SS. Furthermore, each panel had a DRA facilitator who was well-versed in operating LDW; I was the DRA facilitator for the Management Panel.

For ground vehicles, the SD process usually begins with reviewing customer-identified driving requirements to identify and define evaluation criteria that assess the performance and burdens of the subsystems. Burden criteria include, but are not limited to: size (internal and external volumes); weight; required power; costs (developmental, production, and life-cycle); and risk (cost, schedule, and performance). In most analyses, data collection requires a significant portion of the SD schedule; therefore, resource limitations and data availability must be seriously considered when identifying evaluation criteria. Generally, natural data is preferred (e.g., miles per gallon, rounds per minute, pounds, dollars). However, when natural data is not readily available, constructed criteria can be used (e.g., Preferred / Acceptable / Unacceptable; Excellent / Above Average / Average / Below Average / Poor; Pass / Fail), but it must be defined in sufficient detail to minimize imprecision and subjectivity. For this SS the evaluation areas, elements, and criteria were identified and briefly described by USG, then incorporated into Section M (Evaluation Factor for Award) of the GDLS RFP prior to its release.

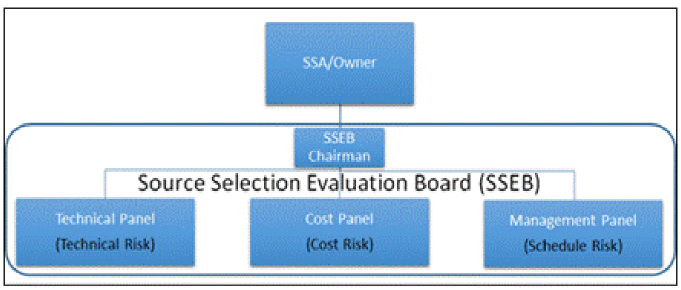

Figure 1: MGV Engine Source Selection Organization.

Although RIWs will be specifically addressed later in the process, this is an appropriate time to draw attention to the fact that the areas of evaluation identified in a previous paragraph are listed in order of importance. In the same context that the evaluation areas, elements, and criteria are identified in Section M of the RFP, it also specifies the importance hierarchy between and within the evaluation areas as determined by USG. At the highest level, the Technical area is worth slightly more than Risk, which is slightly more than Cost, and Cost is significantly more important than Management; and Technical, Risk, and Management, when added together, are significantly more important that Cost. Figure 2 offers one array that meets these parameters.

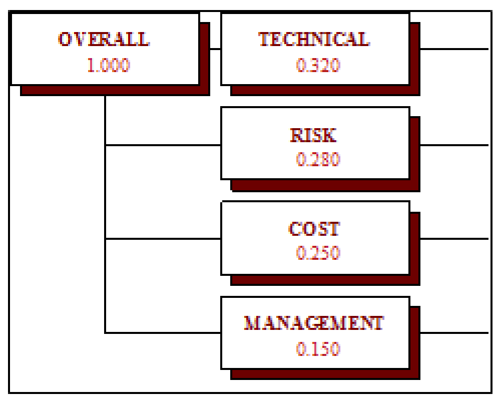

Figure 2: Overall Level-2 Goals Hierarchy with Relative Importance Weights

Within the Technical area, there are five evaluation elements: deployability, sustainability, agility & versatility, survivability, and responsiveness, each with 1-5 evaluation criteria. The elements are listed in order of priority, with the exception that deployability and sustainability are of equal importance. Within the Risk area, each of the three evaluation elements, technical, cost, and schedule, are equally important. The Cost area also consists of three evaluation elements: operating & support (O&S) cost, which is more important than the most probable development cost, which is of equal importance with unit production cost. The Management area has three evaluation elements, each with 2-5 evaluation criteria: program execution, which is equally important with past performance, which is more important than Earned Value Management System (EVMS). Exhibit 1 displays the level-3 goals hierarchy for the just-identified evaluation areas and elements. Furthermore, it presents one combination of RIWs that meets the importance relationship constraints established in the RFP's Section M and will be used later as a starting point for allocating the evaluation criteria RIWs.[5]

3.2. Refine Evaluation Criteria and Single-Criterion Utility Functions

Since I was a member of the Management Panel, that area will be the main focus for the remainder of this case study. The Management Team Lead and the subcontract administrator were both from GDLS. The remainder of the Management Panel consisted of at least one person from the USG, LSI, and UDLP. Per the SS Plan, it had been determined that the week before the proposals were due would be an opportune time for each panel to review their respective evaluation information contained in Section M of the RFP and determine if any refinement was warranted for the evaluation criteria. For natural data criteria, which is mainly what the other two panels would expect to review, it is not likely that much refinement would be needed. However, for the Management area, all evaluation criteria had to be constructed, using a three-point scale. Fortunately, the time provided an opportunity for panel members to discuss and reach consensus on all evaluation criteria. An example that reflects how Management criteria are structured is the Problem Reporting System criterion:

0 = No formal problem reporting system in place,
1 = Problem reporting system in place, but not adequately staffed or supported by management, or
2 = Fully-developed and documented closed-loop problem reporting system with adequate staffing.

The primary purpose of the SUF, referred to as a Common Unit in LDW, is to convert all units of measure (e.g., miles per gallon, rounds per minute, pounds, dollars) into one common unit of measure (utility between 0 and 1), to allow all criteria to be equally compared against each other. The default SUF is a straight line between the two data-range end-points. Admittedly, the Management evaluation criteria SUFs are quite simple: the utility for 2 = 1.0, for 0 = 0.0, and the utility for 1 is somewhere in-between, depending on the criterion. It was not very difficult to lead the panel members to consensus on whether the mid-point utility should be a 0.68, 0.7, or 0.8. The important task was ensuring that for each criterion a solid rationale for the chosen mid-point utility was recorded. Exhibit 2 displays a few notional examples of SUFs that are more thought-provoking.

3.3. Characterize Proposal Alternatives

For this SS, more than four vendors submitted a total of 17 engine proposals. Once proposals were received and vetted by Chris Aileron to ensure that they complied with RFP requirements, necessary information was shared with the appropriate panel. Although vendor proposals typically include extensive amounts of information and data addressing the RFP requirements, members of the Panels conducted independent data collection and analysis to augment the offeror's, or vendor's information, as well as gauge confidence in that information. Management Panel analysis focused on establishing an objective perspective on a vendor's past performance and current ability to manage a development program, as well as their ability to identify and manage schedule risk. This allowed each vendor's proposal to be evaluated objectively against each criterion and receive a rating of 2, 1, or 0.

To assess the risk associated with each proposal, GDLS applied a structured risk management (RM) process that aligns with guidance from the Department of Defense Acquisition University (DAU) and the Project Management Institute (PMI). For each proposal alternative, potential risks were identified and assessed with respect to likelihood of occurrence, and consequence concerning cost, schedule and performance. Risk can be a measure of low confidence in some particular data (e.g., a vendor's claimed performance or costs for its proposal), as well as the uncertainty in whether a supplier's proposed solution will meet one or more of the customer's requirements. The integration of SD and RM provides increased customer satisfaction through the streamlined identification, assessment, and mitigation of risks associated with design alternatives. Once a particular alternative is selected, the risks for that alternative have already been identified and provide a point of departure for the respective project's RM activities. Overall, it is expected that each proposal will have appropriate technical risks identified and assessed, and that the respective proposed schedule and costs will account for resources necessary to mitigate those risks. For this SS, risk was measured against a five-point scale, ranging from Low with a utility of 1.0, to High, with zero utility.

3.4. Allocate RIWs and Select the Preferred Alternative

Once all proposals had been rated with respect to the Management evaluation criteria, the data were entered into LDW. A meeting was then held with the stakeholders to review the SUFs and rating, and determine a RIW for each evaluation criterion. Ordinarily, when assigning RIWs the tendency is to have them strictly replicate the significance of the criteria in the overall requirements structure; the importance hierarchy imposed by the RFP's Section M somewhat encouraged this approach. Another method is to determine RIWs strictly on the respective criterion's range of rating data, i.e., the greater the range, the more importance allocated. This aligns with the practice of eliminating non-discriminating criteria, those where all alternatives are rated equally against a given criterion; the assigned RIW of the cancelled criterion is then proportionally allocated to the remaining criteria. The objective is to gain consensus from all stakeholders that they can accept the RIW allocation. In this case it was a bit of a challenge with the USG-imposed importance constraints. Once explanations were shared and discussed, arriving at the final RIWs did not turn out to be terribly difficult. Beginning with the strawman RIWs provided in Exhibit 1, for each evaluation criterion, I elicited from the stakeholders an RIW that reflected their relative importance preference, yet remained within the imposed constraint. Exhibit 3 displays the level-4 goals hierarchy of the Management area, including evaluation criteria with parent elements and respective RIWs. Again, the presented RIWs represent one possible combination that remains within the boundaries of the imposed importance relationship.

An interesting development transpired early in the alternative characterization effort. When I reiterated to the panel members that a rating of 0 assigned for any criterion would result in zero utility for the associated alternative, some were not receptive to that scheme. To relieve the members' concern, all Management evaluation criteria were modified by adding two scale end-points of -1 and 3, and setting the 0 point utility to 0.1; 2, 1, and 0 remained as the only permissible ratings to be assigned. Figure 3 displays an example of a revised Management evaluation criterion SUF.

LDW produces several informative output graphics that facilitate decision-making and report generation. For this SD, the outputs were presented live to the stakeholders to encourage discussion and gain stakeholder buy-in with the SD results. The first one shown was the stacked-bar ranking of the alternatives (see Figure 4[6]). I explained to the stakeholders that for each alternative the multi-colored stacked-bar ranking reflects the level of utility for each evaluation criterion, which is the product of the criterion rating for that alternative, multiplied by the RIW of that criterion. Furthermore, the criteria displayed at the bottom of the diagram are listed in decreasing order of RIW from left to right, top to bottom. "Note the groupings of alternatives," I pointed out to the stakeholders. "This grouping is indicative of the number of vendors. Also, in addition to seeing that the first two groups are roughly equal in overall utility, we can quickly determine that the first group demonstrates higher utility than the second group for Product Delivery, while the second group has more utility in Master Schedule and Transition to Production, compared to the first." After a moderate amount of discussion, the stakeholders expressed their satisfaction with what was presented, and we moved on.

Figure 3: Example of a Revised SUF for Management Evaluation Criteria

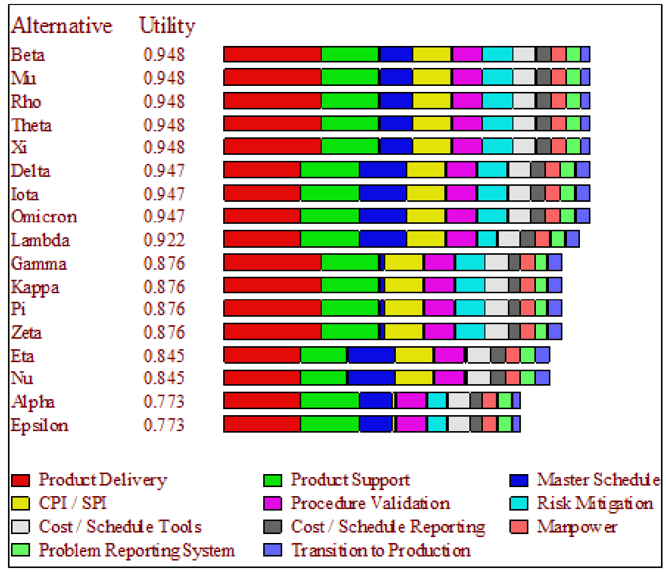

Figure 4: Management Panel Stacked-Bar Ranking of Alternatives

Next were the sensitivity graphs for each criterion with respect to RIW. Each graph shows the extent that the criterion's RIW can be increased or decreased until it produces a change in the ranking of the alternatives. Figure 5 reveals that the Master Schedule criterion is highly sensitive, since an increase of less than one percentage point to its RIW (12.1 to 13.0) will produce a change to the ranking of the top two groups of alternatives. I addressed the panel about this situation: "Fortunately, this is the only criterion that shows considerable sensitivity. At this point, we can have you, the stakeholders, revisit the ratings with respect to that criterion and decide if any revision is necessary, or change the RIW, or leave it as is." After revisiting the assigned ratings, the stakeholders decided to not make any changes and reexamine for sensitivity after the data from all panels were combined.

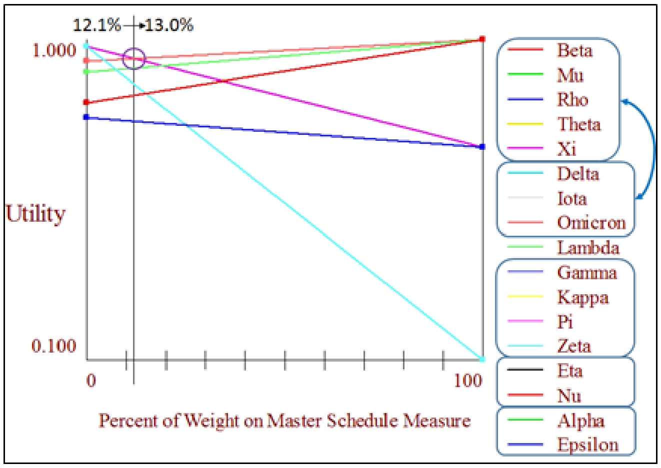

Figure 5: RIW Sensitivity Graph for Master Schedule Criterion

Another LDW output presented to the stakeholders was the direct comparison chart of any two alternatives to identify their strengths and weaknesses. Exhibit 4 reveals that alternative Beta is superior to Delta in Product Delivery by 0.053 utility, whereas Delta is superior to Beta in Master Schedule and Transition to Production by 0.052 utility. This aligns with the stacked-bar ranking observations mentioned earlier (Figure 4). These three capabilities combined provided incredibly valuable insight for the stakeholders and DM in selecting their preferred alternative.

At this point the other panels had completed the same steps, and all data were combined into one LDW model for final review with the stakeholders and SSA. However, before that information could be revealed to the stakeholders, the names of the alternatives had to be masked so only a limited number of individuals would know which costs were linked to specific vendors. This is the reason why the alternatives shared in this case are named after the Greek alphabet. A decipher key that mapped the original alternative names to the Greek names was established and retained in a secure location; only individuals with a need-to-know were allowed access. This practice can be beneficial even during routine SD circumstances, since it helps counter bias that stakeholders may have with regard to alternatives and respective vendors.

After an extended period of questions and discussion, the stakeholders and SSA expressed that they were quite satisfied with the quality of information that had been presented, and believed that it was more than adequate to identify the superior alternative. Furthermore, they were confident that the rigor and objectivity of the SD would stand up to any scrutiny in the case of a protest.

4. SUMMARY

This case study described in permissible detail how a SD process was applied to augment and enhance a USG SS. Specifically, it enhanced the FCS MGV Engine SS by methodically and soundly evaluating multiple vendors and their 17 proposals to select a solution that achieved the requirements for a common engine for all eight MGV variants. The primary intent of this case study was to share lessons-learned regarding how to address unique situations encountered during the connected SS, specifically:

- assigning RIWs within imposed importance hierarchy constraints,
- using an modified rating scale to avoid an alternative receiving no utility with respect to a constructed criterion, and
- masking each vendor / alternative name (i.e., by way of the Greek alphabet) to comply with access restrictions; this practice can also help counter bias that a stakeholder may harbor toward a particular vendor or alternative.

Now that you have finished an initial read-through of the case study, I suggest that you review the below questions and consider them as you conduct your second, more-detailed review.

- What are some additional reasons why you would not want an alternative to receive zero utility for a given criterion?
- What are some additional reasons why the identities of the vendors and proposals would need to be masked?

- What are other weighting options that align with the RFP?
- Would you prefer to modify the weighting schema prescribed in the RFP, and if so, why?
- Are there any metrics missing that should be included?
- What are potential risks in this structured decision?
- What risks could be assumed?
- What risks need mitigation?
- What would a risk management plan look like for this structured decision?

1. REFERENCES

[1]Feickert, A. "The Army's Future Combat System (FCS): Background and Issues for Congress." CRS Report for Congress, 2006.

ENDNOTES

[1] Annual Review of Government Contracting, 2015 Edition. National Contract Management Association. Bloomberg Government, 6.

[2] Army Source Selection Manual dated Feb 26, 2009.

[3] https://www.defenseindustrydaily.com/bae-closes-united-defense-lp-acquisition-0755/

[4] System-of-systems is defined as a system where the sum of the whole is greater than the sum of the individual parts (or systems); the parts are integrated and interdependent, and may or may not be members of a common domain.

[5] Due to USG and GDLS Sensitive Information restrictions, the RIWs shown in this case do not reflect the actual source selection results. Furthermore, the SUFs and ratings of Management evaluation criteria have been altered and do not reflect the actual results of the source selection.

[6] The alternative names shown are based on the Greek alphabet, though at this stage of the actual SS, the names still reflected the model number of the proposed engine. In order to mask vendor identity, the alternative names were changed to those shown when the data from all panels were consolidated and presented to the stakeholders and SSA for final review.

Exhibit 1. Overall Level-3 Goals Hierarchy with Strawman RIWs

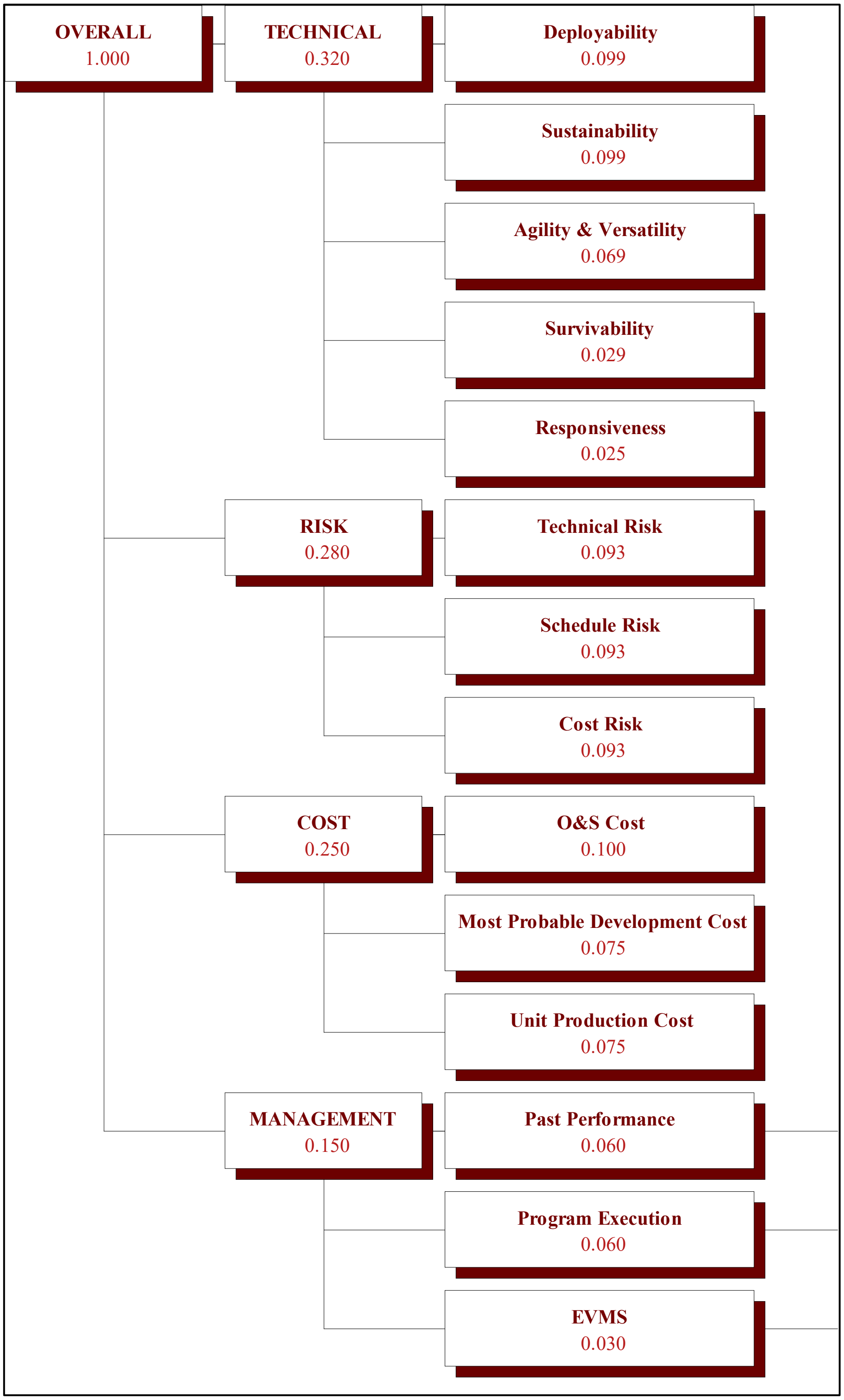

Exhibit 2. SUF Examples

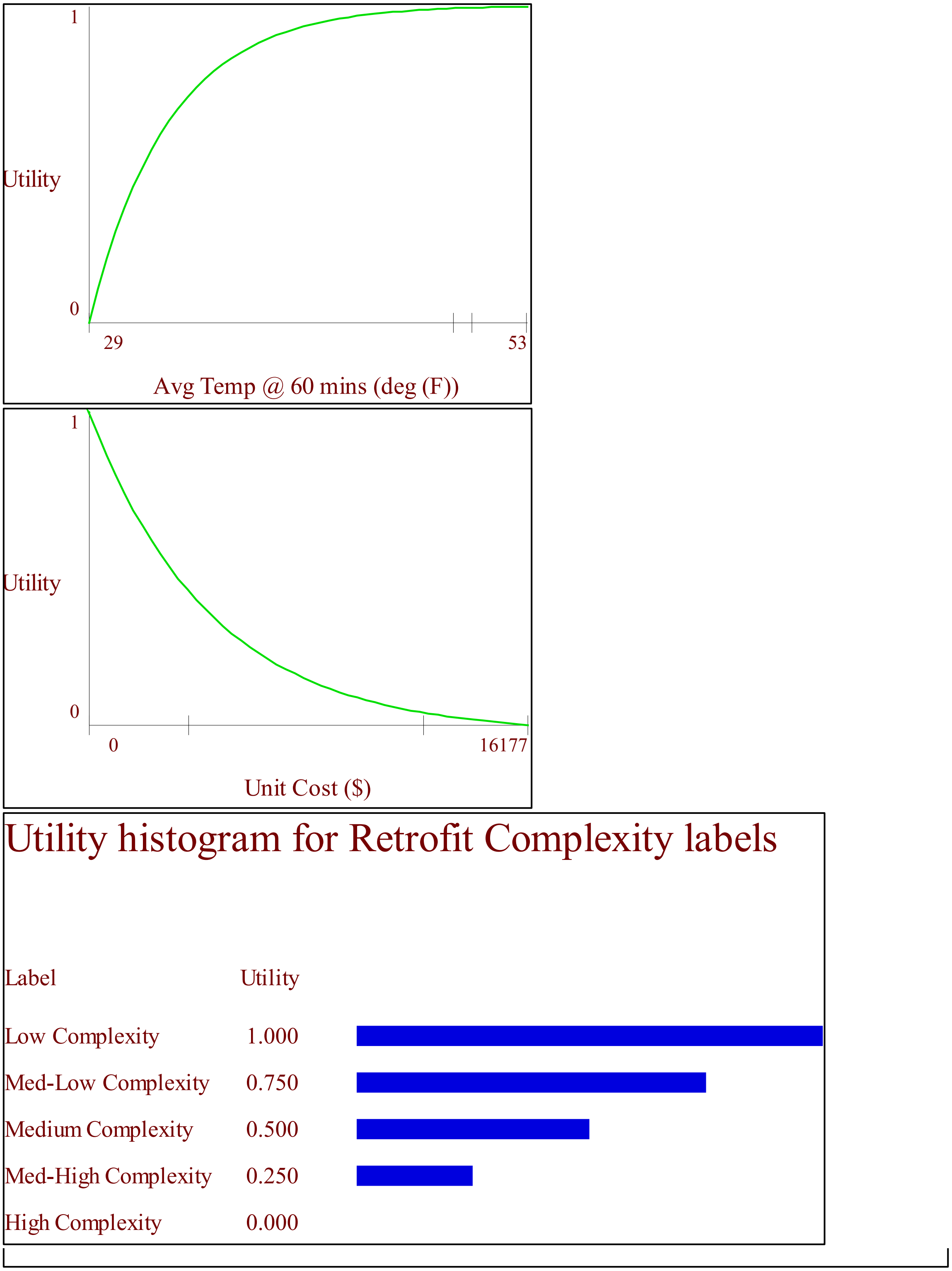

Exhibit 3. Management Level-4 Goals Hierarchy with Acceptable RIWs

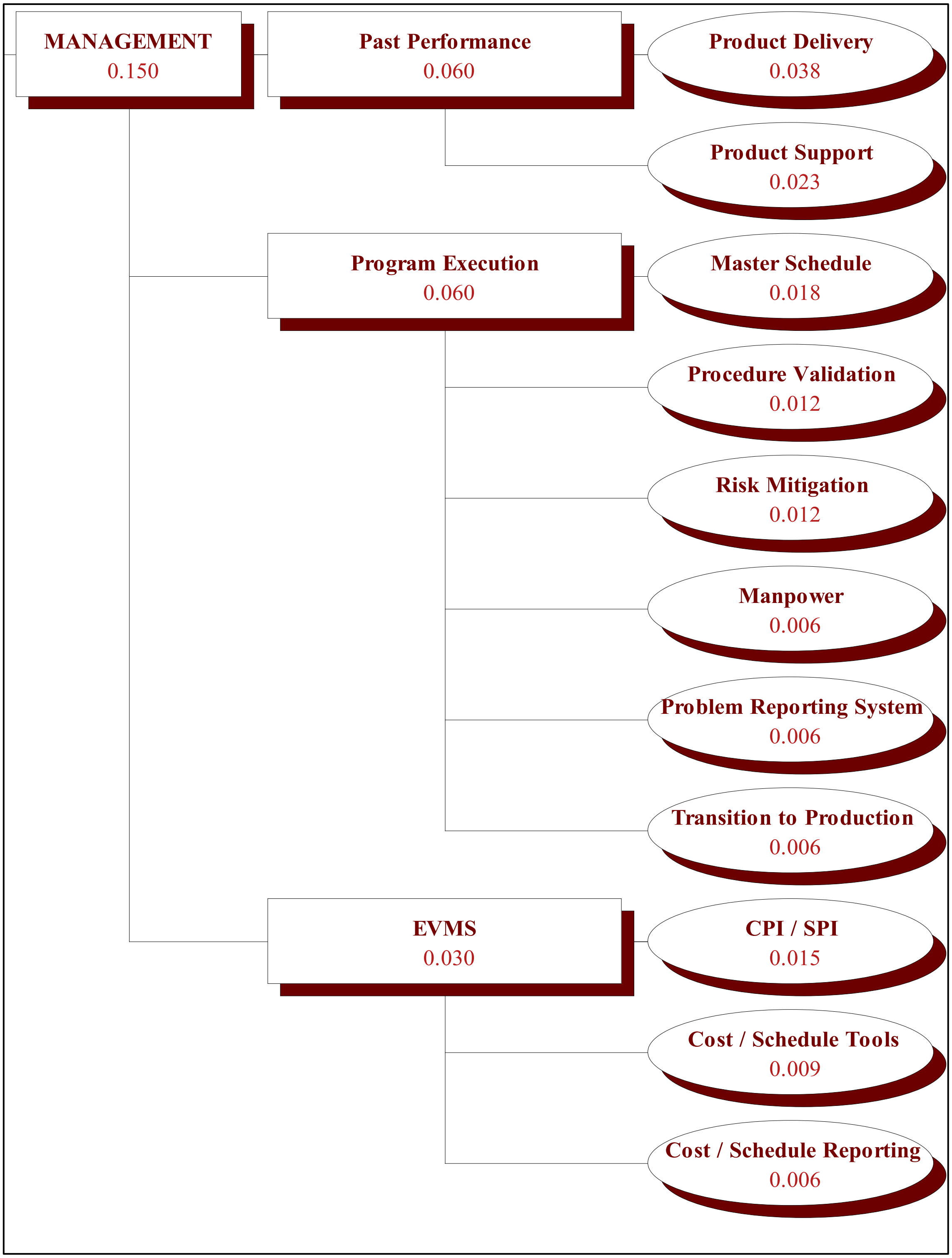

Exhibit 4. Direct Comparison between Beta and Delta Alternatives