

***D-Side: A Facility and Workforce Planning Group Multi-criteria Decision Support System
for Johnson Space Center***

Final Report
Faculty Fellowship Program - 2004
Johnson Space Center

Prepared by: Madjid Tavana, Ph.D.
Academic Rank: Professor
University & Department: La Salle University
Management Department
Philadelphia, PA 19141

NASA/JSC
Directorate: Mission Operations Directorate
Division: Advanced Operations Development Division
Branch: Operations, Research, & Strategic Development Branch

JSC Colleague: Anthony C. Bruins
Date Submitted: July 28, 2004
Contract Number: NAG 9-1526 and NNJ04JF93A

ABSTRACT

“To understand and protect our home planet, to explore the universe and search for life, and to inspire the next generation of explorers” is NASA’s mission. The Systems Management Office at Johnson Space Center (JSC) is searching for methods to effectively manage the Center’s resources to meet NASA’s mission. *D-Side* is a group multi-criteria decision support system (GMDSS) developed to support facility decisions at JSC. *D-Side* uses a series of sequential and structured processes to plot facilities in a three-dimensional (3-D) graph on the basis of each facility’s alignment with NASA’s mission and goals, the extent to which other facilities are dependent on the facility, and the dollar value of capital investments that have been postponed at the facility relative to the facility’s replacement value. A similarity factor rank orders facilities based on their Euclidean distance from Ideal and Nadir points. These similarity factors are then used to allocate capital improvement resources across facilities. We also present a parallel model that can be used to support decisions concerning allocation of human resources investments across workforce units. Finally, we present results from a pilot study where 12 experienced facility managers from NASA used *D-Side* and the organization’s current approach to rank order and allocate funds for capital improvement across 20 facilities. Users evaluated *D-Side* favorably in terms of ease of use, the quality of the decision-making process, decision quality, and overall value-added. Their evaluations of *D-Side* were significantly more favorable than their evaluations of the current approach.

Keywords: NASA, Multi-Criteria Decision Making, Decision Support System, AHP, Euclidean Distance, 3-D Modeling, Facility Planning, Workforce Planning.

1. INTRODUCTION

Over the past decade, NASA's budget has been the target of economizing. Like many other organizations throughout the public and private sectors, one of the most pressing problems facing the Johnson Space Center (JSC) is deciding how to allocate its resources in the face of fiscal constraints. The Systems Management Office at JSC is responsible for developing methods to optimize JSC's facilities and workforce. To do so, the Systems Management Office seeks to allocate increasingly limited resources in a way that is aligned with NASA's mission and strategic goals while also ensuring that operations at each facility are not adversely affected by operational interruptions at other facilities. Finally, the Systems Management Office seeks to develop a resource allocation process that is perceived by stakeholders as fair and free from undesirable personal or political biases.

At many organizations, decisions about allocating resources across facilities follow an annual cycle that is aligned with the organization's fiscal budgeting cycle (Gregory and Pearce 1999). Resource allocation typically begins by considering the investments that each facility would like to make in order to meet its needs and objectives. Because resources are usually limited, these potential investments are prioritized by evaluating them against a set of criteria. Texts on finance and accounting generally recommend that such decisions be evaluated using financial metrics such as payback period, net present value, or profitability index (e.g., Garrison and Noreen 2002). However, for mission-driven agencies in the public sector (such as NASA), these financial metrics are often less helpful than they are in other (e.g., private sector) settings. Whatever the criteria used to rank order facilities, resources are then allocated over the period covered by the budget cycle on the basis of each facility's relative priority. At the end of the cycle, any new or unmet needs and objectives serve as inputs in the next annual cycle.

Currently, the Systems Management Office uses a multi-criteria decision making (MCDM) model to allocate capital improvement among its facilities. A facility review team selected by JSC leadership uses the following five criteria to assess each facility with a weighted sum method:

- Mission - How much does this asset support NASA?
- Availability - How available does this asset need to be?
- Exclusivity - Can this asset be found elsewhere?
- Potential Future Need - Can there be a critical need for this asset in the future?
- Advanced Technology Development - Does this asset contribute to cutting edge research?

The weighting for the criteria are captured with the Analytic Hierarchy Process (AHP) and Expert Choice software (Expert Choice 2004). The facility review team rates each facility on the above five criteria using a 0 (asset does not support the criterion) to 5 (asset fully supports the criterion) rating scale. Each facility's score is the weighted sum of its asset ratings across the criteria. Facilities with higher scores are considered critical and receive more funding.

As part of the workforce planning conducted annually, a staffing review team is formed by JSC leadership to determine the impact of potential changes in work based on changes in strategy and identified center goals. Budget constraints, programmatic changes, probable attritions, and development needs/interests of current staff are used to develop full-time equivalent work demand for each directorate for the fiscal year. The staffing review team then compares the work demand with actual headcount and recommends full-time equivalent numbers for each directorate. The current approach reports on current workforce supply, forecast Center

workforce requirements (demand), and assess future gaps so the Center can allocate appropriate funds to close the gap.

D-Side is a group multi-criteria decision support system (GMDSS) developed for JSC to guide funding decisions for its facilities and workforce. Group decision support systems (GDSS) are interactive, computer-based systems that help a group of decision-makers solve problems and make choices. A group decision support system should support not only group interactions but also structure decision making processes with modeling tools (DeSanctis and Gallupe 1987). Group decision making is a complex human activity and involves assessments of multiple conflicting criteria and various decision alternatives. *D-side* embeds a multi-criteria decision making model (MCDM) into its group decision support system.

Over the last several decades, a philosophy and a body of intuitive and analytical MCDM models have been developed. Schoemaker and Russo (1993) describe four general approaches to MCDM ranging from intuitive to highly analytical. These methods include intuitive judgments, rules and shortcuts, importance weighting, and value analysis. They argue that analytical methods such as importance weighting and value analysis are more complex but also more accurate than the intuitive approaches (Schoemaker and Russo 1993).

Embedding a MCDM into a GDSS requires aggregation of individual preferences into a group decision. In a loosely coupled procedure (*Parallel Coupling*), group members individually assess all alternatives based on their own preferences and reach an individual decision. At the end, individual decisions are synthesized to arrive at a group decision. In contrast, in a tightly coupled procedure (*Sequential Coupling*), group members collectively assess all alternatives based on group preferences and arrive at a group decision. According to Arrow's Impossibility Theorem (Arrow 1963), no group decision making method is perfect. This theorem is regarded as a very important work in modern social choice theory and shows that a group decision outcome can never satisfy every group member. Cao et al. (2003) show that parallel coupling tends to make group members more satisfied with process outcome while the sequential coupling tends to produce better results with respect to the decision quality and decision confidence. *D-Side* supports both parallel and sequential coupling. However, the facility and workforce planning model described in this paper will focus on sequential coupling for simplicity. Results from previous research show that groups using a GDSS outperform groups not using a GDSS or using a GDSS without a problem-modeling tool (Barkhi 2002, Fjermestad and Hiltz 2001, Lam 1997).

D-Side, a group multi-criteria decision support system (GMDSS), offers several advantages. First, *D-side* has a strong strategic focus; resource allocation decisions are guided by explicit consideration of each facility's contribution to the attainment of the organization's mission and strategic goals. Second, *D-side* employs the Analytic Hierarchy Process and Expert Choice software to determine the relative importance (i.e., weight) of each of the organization's strategic goals. Facilities that are closely aligned with the most important strategic goals are given higher priority than facilities that are closely aligned with less important strategic goals. Third, *D-side* also considers interdependencies among facilities; resource allocation decisions are influenced by the extent to which operational interruptions at each facility might adversely affect operations at other facilities. This is accomplished through a modified version of the Mission Dependency Index (developed by the Naval Facilities Engineering Service Center). Fourth, a core metric in the facility management industry, the dollar value of capital investments that have been postponed at the facility relative to the facility's replacement value, is used to compare facilities. Fifth, the *D-side* process produces a plot of the facilities in three-dimensional space,

thereby allowing decision makers to view simultaneously the rank ordering of the facilities, the relative standing of each facility on each of the underlying criteria, and facilities that are similar (i.e., that cluster together) in terms of the underlying criteria. That is, *D-side* incorporates recommendations concerning the visual display of quantitative data by enabling users to view simultaneously a large amount of multivariate and comparative information from many different perspectives (Tuft 1990, 2001). Sixth, *D-side* incorporates several tools from the industrial and organizational psychology research literature to enhance the accuracy of judgments and minimize the likelihood that judgments will be distorted by self-serving or other political biases. Seventh, on the basis of the criteria described above (i.e., alignment with strategic goals, interdependence of facilities, and relative value of postponed capital investments) *D-side* provides clear guidance about the proportion of available resources that should be allocated to each facility. Eighth, *D-Side* employs a shared methodology for both facility and workforce planning models, thereby enabling information sharing between the two models. Ninth, *D-Side* offers a set of integrated decision-support features including visualization and animation, consolidation and aggregation, modeling, what-if, and goal-seek analysis.

Section 2 of this paper presents a detailed explanation of the facilities model. Section 3 provides an overview of the workforce model. Section 4 presents guidelines for the rating process. Section 5 illustrates a hypothetical example. Section 6 describes the results of a pilot study of *D-Side* and presents users' evaluations of *D-Side* versus the current approach in terms of ease of use, the decision-making process, decision quality, and overall value-added. Section 7 presents some concluding comments and managerial implications.

2. THE FACILITIES MODEL

Two panels of expert judges are selected by JSC leadership to participate in this process: the directorate representative panel (*D-Panel*) and the facility management panel (*F-Panel*). The *D-Panel* members are Directorate representatives selected from Center Operations, Engineering, Life Sciences, Mission Operations, Space Shuttle, International Space Station, and Systems Management Office. The *F-Panel* is comprised of the facility managers selected from various Directorates by JSC leadership. *D-Side* uses a series of intuitive and analytical methods to plot each facility in a three-dimensional (3-D) graph based on its Euclidean distance from the Ideal and Nadir points. The overall methodology is depicted in Figure 1.

Insert Figure 1 Here

Conceptually, the methodology is identical for facilities and workforce units, working through three distinct but related phases. The eight-step procedure described below is utilized to systematically evaluate each facility:

1.a. Center Priority Weights: Center Priority Weights are used to determine the Facility Priority Index (P_i) for each facility under review at JSC. P_i attempts to quantitatively score facilities based on how well they support NASA's mission and goals. To determine facility alignment with NASA's mission, each existing facility is scored against a set of strategic goals. This requires the scoring of over 250 separate facilities against 10 goals each planning period. These goals along with their respective missions identified in the 2003 NASA Strategic Plan (National Aeronautics and Space Administration 2003) are presented in Figure 2.

Insert Figure 2 Here

The *D-Panel* initially identifies the relative importance of each of the four missions (W_j ; $j=1, \dots, 4$). Then, for each mission, the *D-Panel* identifies the relative importance of each goal associated with the mission (W_{jk} ; $j=1, \dots, 4$ and $k=1, \dots, l_j$ where l_j is the number of goals associated with mission j). Next, the *F-Panel* collectively scores each facility against each goal in Phase 2 (Step 2.a).

The relative importance (i.e., weighting) for the missions and goals is captured through a series of pairwise comparisons using the AHP and Expert Choice software in a GMDSS. The *D-Panel* members are asked to provide their subjective assessment of each pairwise comparison. Saaty's AHP (Saaty and Vargas 1998, Forman and Gass 2001) uses these pairwise comparisons to derive a weight for each mission and goal.

Initially, the *D-Panel* members are asked to compare each possible pair of missions by providing their judgments about which mission is more important and by how much. Because there are four missions, this requires each *D-Panel* member to make six pairwise comparisons [$n(n-1)/2$]. Then, within each mission, the *D-Panel* members are asked to compare each possible pair of strategic goals (g_i and g_j) and provide judgments about which goal is more important and by how much. For example, three strategic goals are associated with the mission 'Understand and protect our home planet'; weighting these three goals requires three pairwise comparisons. Assuming that the *D-Panel* member is evaluating q goals, AHP quantifies these judgments and represents them in a $q \times q$ matrix:

$$M = (m_{ij}) \quad (i, j = 1, 2, \dots, q)$$

If g_i is judged to be of equal importance as g_j , then $m_{ij} = 1$

If g_i is judged to be more important than g_j , then $m_{ij} > 1$

If g_i is judged to be less important than g_j , then $m_{ij} < 1$

$$m_{ij} = 1/m_{ji} \quad m_{ij} \neq 0$$

Because the entry m_{ij} is the inverse of the entry m_{ji} , the matrix M is a reciprocal matrix. m_{ij} reflects the relative importance of goal g_i compared with goal g_j . For example, $m_{12} = 1.25$ indicates that g_1 is 1.25 times as important as g_2 .

Then, the vector v representing the relative weights of each of the q goals can be found by computing the normalized eigenvector corresponding to the maximum eigenvalue of matrix M . An eigenvalue of M is defined as λ which satisfies the following matrix equation:

$$Mv = \lambda v$$

where λ is a constant, called the eigenvalue, associated with the given eigenvector v . Saaty (1977, 1983, 1989, 1990, 1994) has shown that the best estimate of v is the one associated with the maximum eigenvalue (λ_{max}) of the matrix M . Because the sum of the weights should be equal to 1.00, the normalized eigenvector is used. Saaty's algorithm for obtaining this v is incorporated in Expert Choice software used in this study.

The relative importance of each mission is found through a series of pairwise comparisons among missions. The relative weights of the missions (W_j ; $j=1, \dots, 4$) are the v values computed using the Matrix M , where m_{ij} is the relative importance of mission i compared with mission j . Similarly, the relative importance of each strategic goal for each mission (W_{jk}) is computed.

One of the advantages of AHP is that it encourages panel members to be consistent in their pairwise comparisons. Saaty suggests a measure of consistency for the pairwise comparisons. When the judgments are perfectly consistent, the maximum eigenvalue, λ_{max} , should equal q , the number of goals that are compared. In general, the responses are not

perfectly consistent, and λ_{max} is greater than q . The larger the λ_{max} , the greater is the degree of inconsistency. Saaty defines the consistency index (CI) as $(\lambda_{max} - q) / (q - 1)$, and provides a random index (RI) for matrices of order 3 to 10 based on a simulation of a large number of randomly generated weights. Saaty recommends the calculation of a consistency ratio (CR), which is the ratio of CI to the RI for the same order matrix. A CR of 0.10 or less is considered acceptable. When the CR is unacceptable, the panel member is made aware that his or her pairwise comparisons are logically inconsistent and is encouraged to revise them.

The responses are processed with Expert Choice and panel members with inconsistency ratios greater than 0.10 are asked to reconsider their judgments as suggested by Saaty. The mean importance weights are calculated after the necessary adjustments are made to inconsistent responses. Each panel member is presented with his or her individual score along with the *D-Panel* group mean weights. Panel members are given the opportunity to revisit their judgments and make revisions to their pairwise comparison scores based on this feedback.

There has been some criticism of AHP in the operations research literature. Harker and Vargas (1987) show that AHP does have an axiomatic foundation, the cardinal measurement of preferences is fully represented by the eigenvector method, and the principles of hierarchical composition and rank reversal are valid. On the other hand, Dyer (1990a) has questioned the theoretical basis underlying AHP and argues that it can lead to preference reversals based on the alternative set being analyzed. In response, Saaty (1990) explains how rank reversal is a positive feature when new reference points are introduced. We use the geometric aggregation rule to avoid the controversies associated with rank reversal (Dyer 1990a, Saaty 1990, Harker and Vargas 1990, and Dyer 1990b).

The result of this step is a set of weights representing the relative importance of the four missions and 10 strategic goals.

1.b. Facility Backlogs and Replacement Values: The Facilities Engineering Division at JSC regularly monitors its facilities and keeps data concerning deferred maintenance and current replacement value of its facilities. An official *Deferred Maintenance Parametric Estimating Guide* is used to perform this continuous assessment. The data collected by the Facilities Engineering Division is used to determine the Facility Condition Index (C_i) of each facility in Phase 2 (Step 2.c).

2.a. Facility Priority Index (P_i): Facility Priority Index attempts to quantitatively score facilities based on how well they support NASA's mission and strategic goals. *F-Panel* members rate each facility against each of the strategic goals identified in the 2003 NASA Strategic Plan (NASA 2003) using a rating scale between 0 and 100 (S_{ijk}). A simple additive weighting model is used to compute a P_i for facility i by multiplying the relative importance weight of each mission (W_j) by the relative importance weight of each goal (W_{jk}) and the facility score (S_{ijk}):

$$P_i = \sum_{j=1}^4 \sum_{k=1}^{l_j} W_j W_{jk} S_{ijk} \quad (\text{where } l_j \text{ is the number of goals associated with mission } j) \quad (1)$$

The P_i calculation yields a score between 0 and 100. Facilities with large P_i are highly aligned with the mission and goals while facilities with small P_i are less closely aligned with NASA's mission and goals.

2.b. Facility Dependency Index (D_i): The purpose of the Facility Dependency Index is to reduce the risk of consequences associated with not funding or under-funding facilities. D_i is a modified version of the Mission Dependency Index developed by the Naval Facilities Engineering Service Center (NFESC) to describe the relative importance of infrastructure in

terms of mission criticality (Antelman and Miller 2002). The Facility Dependency Index is an operational risk management metric that reflects the extent to which other facilities are dependent on the focal facility. It does this by evaluating responses to two questions that together yield a score from 0 to 100, with 100 being the highest score. The first question focuses on the impact that an interruption of operations at one facility would have on other facilities. The second question focuses on the difficulty of relocating the function or service provided by the facility. *F-Panel* members respond to the following two questions for each facility:

Question 1: How long could the operations at this facility be stopped without adverse impact on other facilities? The possible responses are:

- None : Any interruption will immediately have an adverse impact on other facilities
- Very Brief: A few hours
- Brief : One day
- Very Short: Two days
- Short : Three to seven days
- Long: One to two weeks
- Very Long : More than two weeks

Question 2: If the facility was not functional, could its operations be relocated to another facility or a temporary facility? The possible responses are:

- Impossible: Alternate facility is not available.
- Very Difficult: Alternate facilities exist, but utilization would require an extraordinary effort with respect to person-hours and money and the job may be compromised.
- Difficult: Alternate facilities exist, but utilization would require a moderate effort with respect to person-hours and money, however, the job would not be compromised.
- Possible: Alternate facilities are readily available with little effort.
- Very Possible: Alternate facilities are readily available with no effort.

The responses to Questions 1 and 2 are referred to Table 1 to determine the Facility Dependency Index of the facility. Facilities with high D_i scores should be viewed as ‘critical.’ Other facilities are adversely affected by work interruptions at such facilities and the operations of such facilities cannot be easily relocated.

Insert Table 1 Here

2.c. Facility Condition Index (C_i): The Facility Condition Index, first cited in 1991, has long been a core metric in the facility management industry to provide a comparison between different facilities (Rush 1991). C_i measures the relative condition of a single facility and is defined as the facility backlog (B_i) divided by the facility replacement value (V_i) times 100 ($C_i = B_i/V_i \times 100$). Facility backlog is the dollar value of capital investments that have been postponed due to lack of funding or other reasons. The facility replacement value shows how much it would cost to replace a facility if it had to be built from scratch. The C_i calculation yields a score between 0 and 100. Facilities with lower C_i 's are in better condition while facilities with higher C_i 's are in worse condition.

While C_i is a useful index, relying only on C_i to allocate resources can be misleading because what is broken should not always be considered a top priority for funding (e.g., when a facility is not closely linked to missions/goals and when the facility's operations have little impact on other facilities). Our model therefore combines three indices to address this concern. Each facility is represented as a bubble in the 3-D model. The diameter of the bubbles represents the amount of investment needed to bring the facility to the desired condition (Facility Backlog).

3.a. Similarity Factors (Z_i): The Ideal point represents a hypothetical facility that would have received the maximum score (i.e., 100) on each of the evaluation criteria (Facility Priority, Facility Dependency, and Facility Condition). In contrast, the Nadir point represents a hypothetical facility that would have received the minimum score (i.e., 0) on each of the evaluation criteria. Our model relies on the premise that the facility that is closest to the Ideal and furthest from the Nadir is preferred to other facilities and is a better candidate for funding. (Note that the facility that is closest to the Ideal would be, by definition, furthest from the Nadir.) That is, the rank ordering of facilities from the Ideal will be inversely (and highly) correlated with the rank ordering of facilities from the Nadir. A graphical view of this model is presented in Figure 3.

Insert Figure 3 Here

Assuming:

- n = Number of facilities under consideration.
- P_i = Facility Priority Index of the i -th Facility ($i=1, \dots, n; 0 \leq P_i \leq 100$)
- D_i = Facility Dependency Index of the i -th Facility ($i=1, \dots, n; 0 \leq D_i \leq 100$)
- C_i = Facility Condition Index of the i -th Facility ($i=1, \dots, n; 0 \leq C_i \leq 100$)
- d_i^* = Euclidean Distance of the i -th Facility from the Ideal Point ($i=1, \dots, n$)
- d_i^- = Euclidean Distance of the i -th Facility from the Nadir Point ($i=1, \dots, n$)
- Z_i = Similarity Factor of the i -th Facility to Ideal ($i=1, \dots, n; 0 \leq Z_i \leq 1$)

Our objective is to *Maximize* $Z_i = \frac{d_i^-}{d_i^* + d_i^-}$ (2)

where:

$$d_i^* = \sqrt{(P_i - 100)^2 + (D_i - 100)^2 + (C_i - 100)^2} \quad (3)$$

$$d_i^- = \sqrt{P_i^2 + D_i^2 + C_i^2} \quad (4)$$

3.b. Sensitivity Analysis: The process as proposed here presumes equal weights for P_i , D_i , and C_i . However, the process could be easily modified by attaching different weights to P_i , D_i , and C_i . A sensitivity analysis could be performed to evaluate the sensitivity of the results to the weights associated with P_i , D_i , and C_i . Assuming that \hat{w}_P , \hat{w}_D , and \hat{w}_C are the relative importance weights associated with P_i , D_i , and C_i , the Euclidean distance of the i -th Facility from Ideal (\hat{d}_i^*) and the Euclidean distance of the i -th Facility from Nadir (\hat{d}_i^-) can be found as:

$$\hat{d}_i^* = \sqrt{\hat{w}_P (P_i - 100)^2 + \hat{w}_D (D_i - 100)^2 + \hat{w}_C (C_i - 100)^2} \quad (5)$$

$$\hat{d}_i^- = \sqrt{\hat{w}_P P_i^2 + \hat{w}_D D_i^2 + \hat{w}_C C_i^2} \quad (6)$$

Consider two facilities, A (P_A, D_A, C_A) and B (P_B, D_B, C_B). Equations (5) and (6) can be used to calculate \hat{d}_A^* , \hat{d}_A^- , \hat{d}_B^* , and \hat{d}_B^- . We can find the critical value of one weight assuming that another one is fixed. For example, assuming \hat{w}_D is fixed, since $\hat{w}_P + \hat{w}_D + \hat{w}_C = 1$, both \hat{w}_P and \hat{w}_C will change with sensitivity analysis. We can find the critical weight associated with the P_i for the Nadir point (\hat{w}_P^-). The critical weight is the weight that causes rank reversal when $\hat{d}_A^- = \hat{d}_B^-$:

$$\hat{w}_P^- = \frac{\hat{w}_D (D_B^2 - D_A^2 + C_A^2 - C_B^2) + C_B^2 - C_A^2}{C_B^2 - C_A^2 + P_A^2 - P_B^2} \quad (7)$$

If A is preferred to B when $\hat{w}_P < \hat{w}_P^-$ then B is preferred to A when $\hat{w}_P > \hat{w}_P^-$

If A is preferred to B when $\hat{w}_P > \hat{w}_P^-$ then B is preferred to A when $\hat{w}_P < \hat{w}_P^-$

A and B are equally preferred if $\hat{w}_P = \hat{w}_P^-$

We can also find the critical weight associated with the P_i for the Ideal point (\hat{w}_P^*) assuming

$$\hat{d}_A^- = \hat{d}_B^-:$$

$$\hat{w}_P^* = \frac{\hat{w}_D [(D_B - 100)^2 - (D_A - 100)^2 + (C_A - 100)^2 - (C_B - 100)^2] + (C_B - 100)^2 + (C_A - 100)^2}{(C_B - 100)^2 - (C_A - 100)^2 + (P_A - 100)^2 + (P_B - 100)^2} \quad (8)$$

If A is preferred to B when $\hat{w}_P < \hat{w}_P^*$ then B is preferred to A when $\hat{w}_P > \hat{w}_P^*$

If A is preferred to B when $\hat{w}_P > \hat{w}_P^*$ then B is preferred to A when $\hat{w}_P < \hat{w}_P^*$

A and B are equally preferred if $\hat{w}_P = \hat{w}_P^*$

There are several other approaches to sensitivity analysis. For example, one might consider performing sensitivity analysis on different weights such as the Center Priority Weights or different scores such as the facility scores.

3.c. Resource Allocation: In the final step of the process, Similarity Factors (Z_i) are used to allocate capital improvement funds to each facility. Facilities with higher P_i , C_i , and D_i are more closely aligned with goals, more critical, and more in need of improvement. Therefore, a larger percentage of their backlog should be funded. Initially, we normalize Z_i 's to determine an

allocation factor for each facility ($\bar{Z}_i = Z_i / \sum_{i=1}^n Z_i$). Next, \bar{Z}_i 's are multiplied by the backlogs to

calculate the tentative capital improvement budget of each facility ($A_i = \bar{Z}_i \cdot B_i$). A_i 's are then

summed to calculate a total tentative capital improvement for the Center ($\bar{A} = \sum_{i=1}^n A_i$). Systems

Management Office uses \bar{A} as a benchmark for capital improvement. Once the actual capital

improvement budget for the Center (\bar{A}) is determined, all A_i 's are adjusted downward (if $\bar{A} > \bar{A}$)

or upward (if $\bar{A} < \bar{A}$) by $(\bar{A} - \bar{A}) / \bar{A}$ percent.

3. THE WORKFORCE MODEL

The workforce model is conceptually identical to the facilities model and uses three indices to plot each workforce unit in a 3-D graph. A workforce management panel (W -Panel), which is selected from Center Operations, Engineering, Life Sciences, Mission Operations, Orbital Space Plane, Space Shuttle, International Space Station, and the Systems Management Office, is formed in addition to the D -Panel to participate in this process. Similarity factors are used to rank order workforce units based on their Euclidean distance from the Ideal and Nadir points. This model provides a systematic and structured approach to measure workforce needs and priorities in the short and long-term. 'Workforce units' or functionally-oriented job types are used as the unit of measure for this model. The workforce units are the civil service employees and support service contractors who share work with civil servants, but not the product-oriented

contractor workforce. The workforce units for three JSC Directorates: Mission Operations, Engineering, and Space and Life Sciences are presented in Table 2 for demonstration purposes.

Insert Table 2 Here

Directorates provide workforce backlogs and workforce value data. Workforce backlog refers to the gap between anticipated and optimal human resources investments (including human resources investments that have been postponed due to lack of funding). In order to determine this figure, the Directorate must first determine the optimal civil service level during the upcoming budget period. The optimal civil service level refers to the salary, benefit, and training costs associated with the number of civil service employees (and support service contractors who share work with civil servants) that will be required to meet forecasted workload expectations during the upcoming budget period. Each Directorate must also determine the anticipated civil service level during the upcoming budget period. When identifying the anticipated civil service level, attention should be paid to attrition (retirements and other turnover) levels, the quantity of hires currently allowed into the Directorate, and the typical movement rates among workforce units. If the optimal level of human resources investments is greater than the anticipated level, then additional investments in human resources are required during the upcoming budget period. If the optimal level of human resources investments is less than the anticipated level (e.g., because forecasts indicate that workload expectations are decreasing for the workforce unit), then no additional investments in human resources are required during the upcoming budget period (and the level of human resources investments in the workforce unit might even need to be reduced).

Alternative definitions of workload backlog could also be considered. For example, workforce backlog could also refer to the extent to which the workforce unit lags competitors in compensation (and hence experiences higher than expected rates of turnover and related human resources problems) or insufficient investment in training and development for current employees (e.g., in-house training, external education).

The most straightforward definition of workforce value is the current value of the workforce unit's compensation costs (including direct compensation and benefits). However, if we want to parallel the 'replacement value' concept from the facilities model, workforce value should be defined as the cost to replace the workforce unit if it had to be hired from scratch. In addition to compensation costs, this would include an estimate of the recruiting and selection costs that would be involved in replacement (likely to be much higher for some workforce units than others due to talent scarcity in some disciplines) and the orientation and training costs that would be required to make the new employees fully productive (also likely to be much higher for some workforce units than others). The Workforce Condition Index is workforce backlog divided by workforce value times 100, yielding a score between 0 and 100.

Next, the Workforce Priority Index is calculated. This index quantitatively scores workforce units based on how well they support NASA's mission and goals. The *W-Panel* collectively rates the workforce units against each of the goals identified in the 2003 NASA Strategic Plan (NASA 2003). An additive weighting model is used to compute this index for each workforce unit by multiplying the Center Priority Weights developed earlier by the *D-Panel* and the workforce unit scores developed by the *W-Panel*. This calculation yields a score between 0 and 100 for the Workforce Priority Index.

Next, the Workforce Dependency Index is calculated for each workforce unit. Workforce Dependency Index represents the consequences associated with not funding or under-funding workforce units. Similar to the facility model, this index determines workforce dependency

based on responses from two questions. The *W-Panel* members are asked to respond to the following two questions for each workforce unit:

Question 1: How long could the operations supported by this workforce unit be stopped without adverse impact on other workforce units?

Question 2: If the workforce unit was not functional, could its operations be reassigned to another workforce unit or a temporary workforce unit?

The responses to both questions are referred to Table 2 to determine the Workforce Dependency Index of the workforce unit (similar to the facility model). Workforce units with high dependency should be viewed as 'critical' units.

Finally, the three indices (Workforce Priority, Workforce Dependency, and Workforce Condition) are used to calculate the similarity factors and plot each workforce unit as a bubble in the 3-D graph. The workforce unit that is closest to the Ideal and at the same time furthest from its Nadir would receive higher priority in allocating human resources investments than other workforce units. The similarity factors are used to allocate funds for human resources for each workforce unit.

4. GUIDELINES FOR THE RATING PROCESS

Because the final ranking of facilities (and workforce units) will depend heavily on the ratings provided by *F-Panel* (and *W-Panel*) members, it is important that ratings be perceived as reasonably accurate and fair. If the rating process is viewed by stakeholders as biased, inaccurate, or contaminated by self-serving motives, then resource allocation decisions will be viewed as unfair.

Research in the field of industrial and organizational psychology has distinguished between the fairness of outcomes (e.g., the resources allocated to each facility) and procedural fairness (Gilliland and Langdon 1998). Procedural fairness refers to perceptions about the fairness of the process used to make the decision and is strongly influenced by factors such as the relevance of criteria, opportunities for input to the decision making process, and the consistency with which the process and standards are applied (e.g., across facilities). In research on performance ratings, procedural fairness has been shown to be positively related to acceptance of performance evaluations, trust in the supervisor (rater), satisfaction with the evaluation process, motivation to improve performance, and organizational commitment. Also, supervisors appear to be less likely to distort or manipulate their ratings after steps are taken to build fairness into the process (Taylor et al. 1995). One consistent finding has been that fair procedures can make up for negative outcomes. That is, when the process is perceived to be fair and the basis for decisions is thoroughly and adequately communicated, individuals who receive negative outcomes are much more likely to perceive those outcomes as fair (Gilliland and Langdon 1998). Thus, procedural fairness is likely to be especially important to those facilities receiving negative outcomes (i.e., less than desired resource allocations).

Gilliland and Langdon (1998) suggest a number of steps that can be taken to enhance the perceived fairness of employee evaluation systems; these steps have direct parallels for evaluating facilities. For example, it is important to communicate consistently with all stakeholders about the development of the resource allocation process (including the relevance of the criteria that will be used to rank order facilities) and to alleviate any concerns about the new process through two-way communication forums (e.g., meetings, conferences). It is also important that facility managers are given opportunities to provide input, the evaluation process is standardized (i.e., the same standards are applied consistently to all facilities), multiple raters

are used to evaluate each facility (to minimize biases of individual raters), raters are knowledgeable about the operations at facilities they are asked to evaluate, and administrative decisions (i.e., resource allocations) are closely linked to the rating process. It would also be desirable to allow facilities an opportunity to rebut and request a review of ratings that they perceive to be unfair (thereby creating a sense of due process).

In developing the rating processes described here, we note two types of rating errors that can occur. Some rating errors are unintentional. For example, two raters who agree about a facility's contribution to a specific strategic goal might assign different ratings to the facility because they interpret the rating scale differently. However, some rating errors are intentional and reflect self-serving or political motives. In this case, raters may have the ability to make accurate ratings but they are unwilling to do so. In sum, raters can play organizational games and distort their ratings to achieve organizational or personal goals (Kozlowski et al. 1998). Kozlowski et al. (1998) have noted that politics and associated rating distortions are more likely when (a) there is a direct link between the ratings and desired organizational rewards (as is the case in resource allocation decisions), (b) there is a lack of surveillance of rater behavior, and (c) there is a widespread perception that others will distort their ratings (e.g., that others will provide inflated ratings concerning their own facilities). Kozlowski et al. (1998) describe several actions that organizations can take to minimize the role of politics in ratings; each has a clear implication for evaluating facilities. These recommendations include having top management serve as role models by providing fair evaluations and discouraging political game playing, allowing stakeholders to suggest potential improvements to the system itself, ensuring that evaluation criteria are widely viewed as relevant, training raters, using multiple raters for each facility, and making raters accountable for their evaluations (e.g., having to explain the reasons for their evaluations). It is noteworthy that the recommendations to reduce politics in ratings closely parallel the recommendations for enhancing procedural fairness.

When raters are motivated to provide accurate ratings, rater training can enhance the accuracy of their ratings (Hauenstein 1998). Hauenstein (1998) reviewed empirical research in this area and described key elements in successful rater accuracy training. First raters are familiarized with the rating criteria. Second, raters are given examples of facility characteristics associated with different points (e.g., 10, 30, 50, 70, 90) on the rating scale for each criterion (e.g., characteristics of facilities that might serve as indicators concerning their degree of alignment with specific strategic objectives). Third, raters complete practice ratings. For example, they can be provided with relevant information about a real or hypothetical facility and asked to rate the facility on one or more criteria (e.g., alignment with strategic objectives or questions 1 and 2 of the Facility Dependency Index). Fourth, the distribution of ratings from all raters is then displayed, often accompanied by 'true' ratings (developed earlier based on consensus discussions among subject matter experts who are highly knowledgeable about the facility being rated and the rating criteria). Raters then discuss reasons for differences in their ratings (e.g., "What information led you to give this facility a rating of 70 in terms of its alignment with strategic goal *i* whereas others gave the facility a rating of 40?"). For example, raters might discuss what information about a facility seems most relevant to them in arriving at a rating on a specific criterion. Discussion might also focus on whether inaccurate raters were simply too lenient (or harsh) or whether they did not recognize what information about the facility was most important for evaluating the specific criterion. Raters might also be cautioned to avoid halo errors (the tendency to form an overall impression about a facility and incorrectly rate the facility as high or low on all criteria on the basis of that overall impression). For

example, a facility whose operations would be difficult to relocate (or where interruption of operations would adversely affect other facilities) is not necessarily closely aligned with many strategic goals. Fifth, the process of completing practice ratings and subsequent discussions is repeated concerning several real or hypothetical facilities.

Based on rating research and recommendations described above, we offer the following guidelines for the rating processes used to generate P_i and D_i .

1. Raters should receive rater accuracy training prior to completing ratings that will influence resource allocation decisions.
2. Rating scales should include definitions that describe the meaning of several points on the scale (e.g., 90 = This facility exists primarily to support attainment of this strategic objective).
3. Multiple raters (e.g., 5 or more) should be used. Raters should not be asked to rate facilities with which they have little or no familiarity.
4. Facility managers should always participate in discussing and rating their own facility.
5. The reliability (i.e., consistency) of raters should be tracked. If a set of i raters evaluates j facilities, the reliability of their ratings can be indexed using an intraclass coefficient (Shrout and Fleiss 1979). Using an i raters \times j facilities analysis of variance framework, $ICC = (MS_{between} - MS_{within}) / MS_{between}$. Note that ICC yields the reliability of the average of the ratings made by i raters. ICC values greater than 0.70 are considered acceptable. To estimate the reliability that could be expected from using more or fewer raters, the Spearman-Brown prophecy formula can be used: $r_{nn} = nr_{xx} / [1 + (n - 1)r_{xx}]$ Where r_{nn} is the estimated reliability based on n times as many raters as those at hand, and r_{xx} is the reliability based on the current number of raters. For example, if r_{xx} was 0.80 using 4 raters, then the estimated reliability for 2 raters (i.e., $n = 0.50$) would be $0.67 = (0.50)(0.80) / [1 + (0.50 - 1)0.80]$.
6. Raters whose ratings are consistently outliers should receive additional training or be removed from the process. When rating a single facility, outliers can be detected by displaying the distribution of ratings. Across j facilities, a rater would be considered an outlier if ICC increases when the rater's evaluations are removed from the dataset.
7. Substantial disagreements among raters should be discussed thoroughly and the facility rating on the criterion should be reached through consensus. Where only minor disagreements occur among raters, the average rating can serve as the facility rating.
8. A facilitator should guide rating sessions to ensure that the same process is applied systematically to all facilities. The facilitator can also provide training to new raters and facilitate discussions to reach consensus when there are substantial disagreements among raters. A few studies have focused on the method participants used to interact with GDSS, with emphasis on the use of human facilitators. In general, it has been shown that facilitation enhances the effectiveness of groups using GDSS (Khalifa et al. 2002).

5. HYPOTHETICAL EXAMPLE

This section illustrates the concepts of Ideal and Nadir using four hypothetical facilities at JSC. First, the Facility Priority Index (P_i) is computed for each of the four facilities. P_i Shows how closely aligned each facility is with NASA's mission and strategic goals. The first step in determining P_i is to have the *D-Panel* at JSC identify the relative importance of each mission and strategic goal by using AHP and Expert Choice software. Each panel member is presented with his or her individual weights along with the *D-Panel* group mean weights. The panel members

are given the opportunity to revise their judgments according to this feedback. Table 3 presents the *D-Panel* mean weights after two rounds of judgment.

Insert Table 3 Here

The next step in determining P_i is having the *F-Panel* rate each facility against each strategic goal using a rating scale between 0 and 100. Higher scores are given to facilities that are highly aligned with a goal while lower scores are given to facilities that are less closely aligned with the goal. Table 3 shows the scoring of the facilities along with the P_i for each facility.

Following the calculation of P_i 's, we calculate the Facility Dependency Index (D_i) of each facility. The *F-Panel* is asked to respond to the two questions used to determine the Facility Dependency Index. Table 4 shows the responses to the two questions along with the corresponding D_i for each facility. The results show that facility 3 is the most critical facility while facility 2 is the least critical facility.

Insert Table 4 Here

Next, the Facilities Engineering Division provides pertinent information concerning capital investment backlog and current replacement value of each facility. These figures are used to determine a Facility Condition Index (C_i) for each facility. As shown in Table 5, facility 3 has the largest C_i while facility 2 has the smallest C_i . Furthermore, facility 1 has the largest capital investment backlog while facility 2 has the smallest backlog.

Insert Table 5 Here

Using the mission priority, facility dependency, and facility condition indices, we calculate the similarity factor (Z_i) of each facility and rank order them in descending order. As shown in Table 6, facility 3 is most similar to the Ideal with a similarity factor of 0.64 while facility 2 is least similar to the Ideal with a similarity factor of 0.17. The proximity of facility 3 to the Ideal and facility 2 to the Nadir can also be viewed in the 3-D model presented in Figure 4(a). All four facilities are represented as bubbles in these figures. The size of the bubbles represents the backlog. Bubble 1 has the largest diameter ($B_1=\$2,800,000$) while bubble 2 has the smallest backlog ($B_2=\$1,100,000$). The user interface in *D-Side* supports visualization and animation by allowing the decision maker to rotate the 3-D model and view the facilities from different angles in the cube as it is shown in Figure 4(b). Additional 2-D views of the 3-D model given in Figure 4(c, d and e) can also provide valuable information.

Insert Table 6 and Figure 4 Here

D-Side empowers the decision makers with tools to perform what-if analysis, goal-seeking, and sensitivity analysis. For example, to study the sensitivity of weight to rank reversal, we can find the critical weight associated with the P_i for the Nadir point (\hat{w}_p^-) and the Ideal point (\hat{w}_p^+). Note that the default relative importance weight associated with P_i , D_i , and C_i , is 0.33 and $\hat{w}_p = \hat{w}_D = \hat{w}_C$. Given the following P_i , D_i , and C_i :

Facility 1 (40, 40, 25)

Facility 2 (20, 20, 10)

Facility 3 (80, 80, 40)

Facility 4 (30, 60, 20)

facility 3 totally dominates others while facility 2 is totally dominated by the others. Therefore, we focus our sensitivity analysis on facilities 1 and 4. Using equation (7), we find \hat{w}_p^- , the critical value of P_i as a function of D_i :

$$\hat{w}_p^- = \frac{1775 \hat{w}_D - 225}{475}$$

Considering different values for \hat{w}_D , we calculate \hat{w}_p^- . As shown in Table 7, when $\hat{w}_D = 0.20$, we find $\hat{w}_p^- = 0.27$, meaning that 0.27 is the reversal weight for P_i (compared with the default weight of 0.33 for P_i). In other words, the rank order changes when moving the weight from below 0.27 to above 0.27. $\hat{w}_p^- > 1$ or $\hat{w}_p^- < 0$ corresponds to the infeasible cases where there is no reversal and the order does not change. Similar calculations are presented in Table 7 for \hat{w}_p^* where $\hat{w}_p^* > 1$ or $\hat{w}_p^* < 0$ corresponds to infeasible cases:

$$\hat{w}_p^* = \frac{-2775 \hat{w}_D + 775}{-525}$$

Insert Table 7 Here

Normalized Similarity Factors (\bar{Z}_i) are used to determine the required funding needed for each facility (A_i). The tentative capital improvement amount suggested by *D-Side* for all 4 facilities equals \$1,769,000 (see Table 8). This amount is used by the Systems Management Office as a benchmark for capital improvement distribution. Depending on the available budget, A_i 's are adjusted downward or upward. For example, if JSC leadership approves \$690,000 (10% of total backlogs) for facility improvement, all A_i 's will be adjusted downward by $(1,769,000 - 690,000)/1,769,000$ or 61.0% (see table 8).

Insert Table 8 Here

6. PILOT STUDY

Twelve facility managers volunteered to participate in the pilot study. Their mean number of years at NASA was 18.58 (SD = 7.15). They had on average 8.83 years (SD = 3.83) of facility management experience at NASA and 6.58 years (SD = 1.78) of experience with the current facility evaluation process. Seventeen percent had a bachelor's degree, 66% had a master's degree, and 17% had a doctoral degree.

Participants were randomly assigned to one of two sessions. Both groups were asked to evaluate 20 randomly selected facilities at JSC with a total backlog of \$19,100,000. Group I participants first evaluated the facilities using the current approach (115 minutes) and then evaluated the facilities using *D-Side* (200 minutes). Group II participants first evaluated the facilities using *D-Side* (225 minutes) and then evaluated the facilities using the current approach (80 minutes). In both sessions, the survey containing the dependent variables (described below) was completed at the end of the session (i.e., after all facilities had been evaluated using both approaches).

Table 9 shows the combined results for the two groups using the current approach. The similarity factors presented in Table 10 provide an overall ranking of the facilities using *D-Side* (sorted in a descending order of Z_i). Since we had two groups working on the same problem, the average priority (P_i) and dependency (D_i) indices for the two groups were used to arrive at the final figures. Note that the C_i 's remain unchanged for both groups. The correlation between the

Overall Score from the current approach (Table 9) and Z_i from *D-Side* (Table 10) was .88. That is, the two approaches yielded similar rank orderings of the facilities. The 3-D and 2-D views of the *D-Side* results are presented in Figure 5. Using the normalized similarity factors, \$969,510 was considered as the tentative capital improvement amount for the 20 facilities. However, it was expected that JSC leadership approves 10% or \$1,910,000 towards capital improvement for these 20 facilities. Therefore, all figures were adjusted upward by a factor of $(969,510 - 1,910,000)/969,510$ or 97.0%.

Insert Figure 5 and Tables 9 and 10 Here

Dependent Variables

Our choice of dependent variables was influenced primarily by Bharati and Chaudhury's (2004) decision satisfaction model, Benbasat and Lim's (1993) meta-analysis, DeLone and McLean's (1992) usability model, Gallupe and DeSanctis's (1988) effectiveness model, and Limayem and DeSanctis's (2000) decisional guidance model. Collectively, these (and many other) studies point to the importance of ease of use, the quality of the decision process, and decision quality as key antecedents of users' satisfaction with and use of decision support systems.

We therefore created four scales: Ease of Use (7 items), Decision Process (8 items), Decision Quality (6 items), and Overall Value-Added (4 items). Where feasible, items were adapted from items used in previous studies (e.g., Aldag & Power, 1986; Bharati and Chaudhury, 2004; Davis, 1989; DeLone and McLean, 1992; Gallupe, DeSanctis & Dickson, 1988; Niederman & DeSanctis, 1995; Watson, DeSanctis, & Poole, 1988; Srinivasan, 1985). All items are presented in the Appendix. Each item was answered using a 7-point rating scale where 1 = strongly disagree, 4 = neutral, and 7 = strongly agree.

The reliability (Cronbach's alpha) of each of the scales was good. For the seven Ease of Use items, $\alpha = .74$ and $.85$ for the current approach and *D-Side*, respectively. For the eight Decision Process items, $\alpha = .77$ and $.67$ for the current approach and *D-Side*, respectively. For the six Decision Quality items, $\alpha = .88$ and $.84$ for the current approach and *D-Side*, respectively. For the four Overall Value-Added items, $\alpha = .92$ and $.91$ for the current approach and *D-Side*, respectively.

Following Campbell and Fiske (1959), Davis (1989) pointed out that items intended to measure the same construct for the same approach (convergent validity) should be more highly correlated than items intended to measure different constructs for the same approach (discriminant validity) or different constructs for different approaches. We found that the average correlation among items measuring the same construct for the same approach ($r = .39$) was higher than the average correlation among items measuring different constructs for the same approach ($r = .32$) or different constructs for different approaches ($r = -.26$).

We created a score for each participant on each scale (for each approach) by computing the mean responses to the Ease of Use, Decision Process, Decision Quality, and Overall Value-Added items.

Users' Evaluations

Descriptive statistics concerning the dependent variables are presented in Tables 11 and 12. Inspection of the means in Table 11 indicates that users' reactions to the current approach were slightly below or near the center of the 7-point rating scale whereas their reactions to *D-Side* were more favorable. Inspection of individual-level ratings indicated that 11 of the 12 facility managers evaluated *D-Side* more favorably than the current approach on all four dependent variables. Independent groups t-tests indicated that there were no significant differences on the dependent variables for users in the two sessions. The correlations in Table 12

indicate that, for each approach, correlations among the dependent variables were moderate to high. Also, users' opinions about the current approach were negatively related to their opinions about *D-Side*. That is, positive opinions about *D-Side* were associated with negative opinions about the current approach.

Insert Tables 11 and 12 Here

We used multiple regression to examine the extent to which Overall Value-Added was predicted by Ease of Use, Decision Process, and Decision Quality. For the current approach, $R^2 = .74$, $p = .01$. For *D-Side*, $R^2 = .80$, $p < .01$. Inspection of the standardized beta coefficients (and significance tests associated with them) indicated that Decision Quality was the strongest predictor of Overall Value-Added for both the current approach and *D-Side* (see also the correlations in Table 12).

We examined the relationship between users' background characteristics and the dependent variables. 'Years at NASA' was positively related to opinions about the current approach (for Ease of Use, $r = .57$, $p = .05$; for Decision Process, $r = .69$, $p = .01$, for Decision Quality, $r = .82$, $p < .01$, for Overall Value-Added, $r = .65$, $p = .02$). However, 'years at NASA' was unrelated to opinions about *D-Side*. 'Years of facility management experience at NASA' was positively related to opinions about the current approach (for Decision Process, $r = .58$, $p = .05$; for Decision Quality, $r = .71$, $p < .01$). 'Years of facility management experience at NASA' was unrelated to opinions about *D-Side*. 'Years of experience with the current facility evaluation process' was unrelated to opinions about the current process or to opinions about *D-Side*. However, education level was positively related to opinions about *D-Side* (for Ease of Use, $r = .68$, $p = .02$; for Decision Quality, $r = .59$, $p < .04$, for Overall Value-Added, $r = .71$, $p = .01$) but was unrelated to opinions about the current approach. In sum, users with many years of facility management experience at NASA viewed the current approach more favorably than those with less facility management experience. (Although, as noted above, 11 of 12 users viewed *D-Side* more favorably than the current approach.) And users with higher levels of education viewed *D-Side* more favorably than those with less education.

We used paired t-tests to compare users' opinions concerning the current approach and *D-Side* on each of the four dependent variables. Significance tests indicated that users had more favorable opinions concerning *D-Side* than the current approach on each of the four dependent variables (for Ease of Use, $t = 7.75$, $df = 11$, $p < .001$; for Decision Process, $t = 8.12$, $df = 11$, $p < .001$; for Decision Quality, $t = 6.99$, $df = 11$, $p < .001$; for Overall value-Added, $t = 7.31$, $df = 11$, $p < .001$).

For each dependent variable, we calculated Cohen's d (i.e., the mean difference divided by the standard deviation of the difference scores) as a measure of the effect size. According to Cohen (1988), small, medium, and large values of d are about .20, .50, and .80 respectively. All effect sizes were very large (for Ease of Use, $d = 2.23$; for Decision Process, $d = 2.35$; for Decision Quality, $d = 2.01$; for Overall value-Added, $d = 2.12$), thereby indicating that the mean differences between the current approach and *D-Side* were both statistically and practically significant.

7. CONCLUSION AND MANAGERIAL IMPLICATIONS

The pilot study illustrated that, despite the longer amount of time required to use *D-Side* relative to the current approach (and the similar rank ordering of facilities using the two approaches), facility managers viewed *D-Side* much more favorably in terms of ease of use, decision process, decision quality, and overall value-added. *D-Side* is quite flexible in that it can easily be adapted

by other organizations, for example, by substituting other criteria for those listed here or adding additional criteria. Although this paper uses three criteria (e.g., facility priority index, facility dependency index, and facility condition index), the mathematics of *D-Side* can accommodate any number of relevant criteria and determine the distance of each facility or workforce unit from the Ideal and Nadir (although the 3-D views are naturally limited to only three criteria at a time). As illustrated by the workforce model presented here, the basic framework of *D-Side* can be readily adapted to other decision-making contexts (e.g., evaluating proposals within facilities, evaluating new business opportunities, and so on). Another advantage of *D-Side* is the ease with which it places both inherently subjective criteria (e.g., the alignment of facilities with strategic goals) and more objective criteria (e.g., facility condition index) on a common measuring scale (0 to 100). Finally, sensitivity analyses enable decision makers to understand the conditions (criterion weights) that would cause reversals in the rank ordering of facilities.

We believe that one of strength of *D-Side* is its integration of research from multiple disciplines. That is, it combines research from literature on decision-making (e.g., AHP, MCDM) and research from the industrial and organizational psychology literature (e.g. procedural justice, rater accuracy training). The use of tools from either one of these disciplines without the other is likely to be inadequate. For example, if facility priority index ratings were derived primarily or solely on the basis of ratings made by facility managers concerning their own facilities, self-serving and political biases would likely contaminate ratings and create widespread perceptions that the process was unfair despite its appearance of objectivity in terms of the underlying mathematics. At the same time, the industrial and organizational psychology literature does not offer useful guidance about (a) optimal approaches to weight criteria (e.g., AHP), (b) how changing criterion weights will affect resource allocation decisions (e.g., sensitivity analysis), or (c) how to combine subjective (e.g., ratings) and objective (e.g., financial) data in a way that offers an easily understood rationale to support resource allocation decisions.

While the focus on three criteria might be viewed by some as limiting, we believe it is actually a strength. *D-Side* forces decision-makers to carefully consider and select the most relevant criteria rather than simply brainstorming and including a potentially lengthy list of relevant criteria. *D-Side* also emphasizes the importance of combining strategic (and hence long-term) criteria along with more tactical and short-term criteria (e.g., facility backlog and the extent to which operational interruptions at each facility might adversely affect operations at other facilities).

As with any MCDM, selecting the 'right' criteria is critical. In each application of *D-Side*, one empirical question of applied interest is whether criteria are highly correlated with each other. For example, if workforce unit scores on one criterion (e.g., alignment with strategic goals) were found to be highly correlated (e.g., $r > .80$) with scores on another criterion (e.g., workforce condition index), then it could be argued that the two criteria are redundant. In this situation, one of the two criteria might be replaced with another relevant yet conceptually independent criterion. Moreover, we expect that when decision makers are allowed to include a lengthy list of criteria (an outcome deliberately constrained by *D-Side*), it is highly likely that some criteria will be highly correlated (and therefore redundant) with other criteria. When this occurs, the net effect is that whatever is being measured by the highly correlated (redundant) criteria is given additional mathematical weight in final decisions. This points to the importance of (a) initially selecting a small number of conceptually independent criteria that address both

strategic and tactical perspectives, and (b) periodically examining the correlations among criteria.

Cluster analysis can be used to identify groups of facilities that are similar on the underlying criteria. Note that these groups of facilities would be clustered together in the 3-D model. Identifying relatively homogeneous clusters of facilities can be helpful because it enables senior managers to identify resource allocation and related management actions that are likely to address the needs of all the facilities in a cluster (rather than having to make separate decisions about each facility on its own).

Using a structured, step-by-step approach like *D-Side* is not intended to imply a deterministic approach to facility and workforce planning. While *D-Side* enables decision makers to crystallize their thoughts and organize data by placing both inherently subjective criteria and more objective criteria on a common measuring scale, it should be used very carefully. As with any decision analysis model, the researchers and practicing managers must be aware of the limitations of subjective estimates. *D-Side* should not be used blindly to plug-in numbers and crank-out solutions. The effectiveness of the model relies heavily on the ability and willingness of decision makers to provide sound judgments. Potentially, decision makers could make poor judgments as they do with any approach. Such judgments can generate misleading results and ultimately poor decisions.

REFERENCES AND TABLES ARE AVAILABLE UPON REQUEST