The Emerald Research Register for this journal is available at
**www.emeraldinsight.com/researchregister**

The current issue and full text archive of this journal is available at
**www.emeraldinsight.com/1463-5771.htm**

# Quest 123: a benchmarking system for technology assessment at NASA

Madjid Tavana
*La Salle University, Philadelphia, Pennsylvania, USA*

***Keywords*** *Benchmarking, Decision making, Analytical hierarchy process, Project management*

***Abstract*** *Technology assessment is a difficult task at the Mission Control Center (MCC). The difficulty is inherent in the unavailability of structured information and exacerbated by the lack of a systematic assessment process. New technology deployment to the MCC requires testing and certification in three labs: Quest 1, 2, and 3. The Mission Control Center Systems (MCCS) architecture team, a multidisciplinary group of MCC experts and scientists is chartered to redefine the next generation of MCCS by developing a systematic process to assess and certify new technologies. Quest 123 is a benchmarking tool that was successfully implemented at the Johnson Space Center to assess and certify new technology initiatives for each lab before final deployment to the MCC. Quest 123 integrates the analytic hierarchy process with an additive multi-criteria decision-making model into a dynamic benchmarking framework.*

## 1. Introduction

The concept of benchmarking has become synonymous with achieving successful performance. The original meaning of the word benchmark refers to a metric unit on a scale of measurement. A benchmark is the standard of excellence against which one measures and compares actual performance (Jarrar and Zairi, 2001). While a number of definitions for benchmarking can be found in the literature, they all essentially share the same theme. Benchmarking is the process of continuously improving performance by adapting outstanding practices and processes found inside or outside the organization (Cox *et al.*, 1997; Francis and Holloway, 2002; Jarrar and Zairi, 2001; Spendolini, 1992). Some benefits from benchmarking include establishing goals, meeting requirements, measuring performance, and becoming competitive (Camp, 1989; Drew, 1997). There are many typologies that have been developed to describe or classify the characteristics of benchmarking. The most commonly cited is Camp's (1995) distinction between internal, competitive, functional, and strategic benchmarking. Internal benchmarking studies the best practices in an organization, competitive benchmarking deals with the best competitors in an industry, functional benchmarking investigates companies with similar processes in the same function outside one's industry, and strategic benchmarking studies the best global business practices. Czuchry *et al.* (1995), Dorsch and Yasin (1998) and Jackson *et al.* (1994) provide a good survey of benchmarking literature.

Benchmarking: An International Journal
Vol. 11 No. 4, 2004
pp. 370-384

1463-5771
DOI 10.1108/14635770410546773

---

This research was supported by NASA grant number NAG 9-1526. The author is grateful to Jack Knight, Steven Gonzales, Anthony Bruins, Mission Control Center Systems architecture team, and Quest Labs, all from Johnson Space Center, and his colleagues Dennis Kennedy and Anne Marie Smith for their assistance with this research project.

The fundamental problem in any benchmarking is finding meaningful data to use as a benchmark (Landeghem and Persoons, 2001). Most benchmarking models focus on the "hard" rather than "soft" data (Cassell *et al.*, 2001). Sometimes softer measures are ignored because of the difficulty in measuring them. Reliance on hard measures alone often undermines the strategies the company must pursue to survive in the long-term. Today's smart organizations are searching for ways to integrate soft measures into their regular performance evaluation (Geanuracos and Meiklejohn, 1993). Most current tools such as "spider" or "radar" diagrams and the "Z" chart for gap analysis focus on presenting the hard data in some graphical form (Camp, 1995). Statistical methods utilized in benchmarking such as regression and various descriptive statistics are also concerned with hard data (Blumberg, 1994; Schefcyzk, 1993). One approach which is used mostly for softer data is the analytic hierarchy process (AHP). AHP has been a very popular technique for determining weights in multi-criteria decision-making (MCDM) problems (Shim, 1989; Zahedi, 1986). The advantage of AHP is its capability to elicit judgments and scale them uniquely using a procedure that measures the consistency of these scale values (Saaty, 1989). AHP utilizes the perceptual and subjective values of decision-makers in a weighted scoring technique to analyze various benchmarks (Sarkis, 2001). AHP has been successfully utilized for benchmarking in process performance (Frei and Harker, 1999), logistics performance (Korpela and Tuominen, 1996), strategic performance (Lu *et al.*, 1994; Partovi, 2001), quality performance (Min *et al.*, 2002), and project management performance (Prasanta, 2002).

Although benchmarking is widely used in the public and private sectors, public organizations and government agencies typically face unique operational issues and strategic environments that differ from private organizations because they are not required to maximize and increase profitability (Dorsch and Yasin, 1998). Nevertheless, they are broadly similar in organizational goals and objectives, and performance assessment is considered to be of equal importance in both sectors. Internal benchmarking is largely used at NASA to accelerate organizational transformation and institutionalize continual improvement in mission operations, aviation, information technology, and quality assurance. This paper presents Quest 123, an internal benchmarking tool that integrates AHP with an additive MCDM model into a dynamic benchmarking framework for technology assessment at NASA.

## 2. Benchmarking at NASA

Throughout the history of human spaceflight, the Mission Control Center Systems (MCCS) at the Johnson Space Center (JSC) has been a model for mission planning, communications, command, and control architectures. The mission operations directorate (MOD), overseeing the Mission Control Center (MCC) and MCCS, must ensure that the overall system performance meets current and planned needs while looking for innovative ways to curtail operational costs and continually address evolving operational scenarios. The directorate must also enforce the highest return on investment on the funding it receives annually. This vision provides a basis for the long-term, as well as day-to-day, decision-making that ultimately impacts requirements, design change, and budget plans. The MCCS architecture team, a multidisciplinary group of MOD experts and scientists is chartered to redefine the next generation of MCCS by developing integrated systems design architecture.

The original MCCS was designed as nonintegrated pieces of a whole, independently supporting the larger goal of human spaceflight operations. While more advanced computing capabilities have allowed the MCCS to morph from completely independent functioning systems into a distributed design architecture, the technological advances of the last several years have allowed for the potential implementation of a true integrated systems design architecture. The MCCS, which has always served as the nerve center of US human spaceflight operations, has evolved from a centralized mainframe computer architecture in which all aspects of mission planning, communications, command, and control were performed predominantly from multiple buildings located at JSC to a distributed architecture with multiple remote facilities located around the world.

The current MCCS is a functionally robust set of distributed systems primarily supporting the space shuttle program (SSP) and the International Space Station (ISS) mission operations. The MCCS also performs the following functions: real-time data (telemetry and trajectory) monitoring and analysis; real-time command; near real-time data storage, retrieval and analysis; space-to-ground and ground-to-ground voice, video, data, and mail distribution; as well as real-time and near real-time planning and simulations.

Forged around the uniquely complex and demanding requirements of human spaceflight, the MCCS has developed and evolved within the limits of technological capabilities of the time. The dynamic and continually evolving environment in which MCCS functions has demanded that the over-arching structure of the systems – the systems architecture – continues to evolve as well.

As a result of dwindling funding for the foreseeable future, the MCC must focus on ways to reduce costs while still maintaining and even expanding its capability to support the SSP and an increasingly larger and more complex ISS. As part of an earlier MCC mission computing strategy study, the MCC adopted the following set of goals that the MCCS architecture team uses as major evaluation factors in its early proof-of-concept work toward the vision.

- *Design for change*. Commercial standards compliant, standard interfaces, and simplified configuration management processes for migration.
- *Design for flexibility*. Acommodate future manned spaceflight program requirements and operations concepts without significant impacts to existing architecture design.
- *Design for connectivity*. Information transfer between international partners, MCC users (flight control), payload community, joint space operations with government, industry, and academia collaborations using the Internet where practical.
- *Design for access*. Security commonality/simplification – robust, but simple security for user concurrent access to multiple security levels that does not constrain system functionality.
- *Design for cost reduction*. Commodification (a move toward an infrastructure that can be supplied by many different vendors), consolidation, operational concept modifications, and re-engineer.
- *Design for ease of use and effectiveness*. Intuitive graphical user interface layouts that supports data visualization, and intelligent systems; minimal steps for task completion, effective decision-making and job productivity.

- *Design from a systems perspective*. Develop and test new systems and process improvements that address all issues and concerns relative to the overall systems as a whole entity.

Deployment of new technology initiatives to the MCC is a long and difficult task requiring a four-phase certification process in three separate labs: Quest 1, 2, and 3. Approximately about 5 percent of all technology initiatives reviewed are certified in all labs and deployed to the MCC. A technology initiative is initially considered for possible adoption and testing during phase I of the testing and certification process. Once adopted, the initiative is put through stand alone testing in Quest 1 during phase II of the testing process. Upon certification in Quest 1, the initiative is put through additional systems integration testing in Quest 2 during phase III of the process. Once certified in Quest 2, the initiative is ready for human and systems integration testing in Quest 3 during phase IV of the process. Finally, initiatives certified in Quest 3 are deployed to the MCC.

Prior to implementation of Quest 123, the MCCS architecture team used a simple and unstructured naive approach to assess and certify new technology initiatives in each lab. Using this naive approach, MCCS architecture team members assigned a subjective numerical score between 0 and 100 to each technology initiative. The average score of all team members was used to certify or deny certification to a technology initiative. No formal assessment guidelines were provided to the MCCS architecture team. Once the mean score was calculated, the team used this score to decide if the technology has adequate merits to be certified. The MCC management had expressed their concerns about the lack of a systematic framework and the potential for inconsistency in technology assessment. Quest 123 was created in response to this concern.

## 3. The model

Quest 123 uses a filtering system to certify technology initiatives in each lab. In order to develop this system, we analyzed 112 initiatives reviewed by the MCCS architecture team over the last 2 years. The median score for all 112 initiatives reviewed during the initial screening in phase I was 52 while the mean score was 68 (Table I). Of the 48 initiatives approved for follow-up testing in Quest 1 in phase II 45 had an overall score higher than the median while only 32 scored higher than the mean. Since the means were larger than the medians, the distribution was highly skewed. As a result, the medians provided a better measure of central tendency and approximation for the number of initiatives approved in all assessment phases.

Quest 123 is a MCDM that utilizes AHP to develop an overall score for each technology initiative under consideration. The project evaluation and selection

| Assessment phases | Number of initiatives reviewed | Median | Mean | Number of initiatives approved | Number of initiatives with an overall score greater than median | Number of initiatives with an overall score greater than mean |
|---|---|---|---|---|---|---|
| Phase I | 112 | 52 | 68 | 48 | 45 | 32 |
| Phase II | 48 | 55 | 71 | 25 | 23 | 14 |
| Phase III | 25 | 58 | 74 | 9 | 8 | 6 |
| Phase IV | 9 | 61 | 77 | 5 | 5 | 3 |

**Table I.** Historical cases and their characteristics methods proposed in the literature can be classified into two categories: compensatory and non-compensatory models (Oral *et al.*, 2001). Compensatory models reduce a multidimensional evaluation to a single-dimensional one through an aggregating "value" function, thereby establishing trade-offs between criteria. Multi-objective mathematical programming, multi-attribute utility theory, and AHP can be cited as examples for this category. Non-compensatory models such as Ordinal Ranking and ELECTRE trade-offs between criteria are restricted. The proposed method in this study belongs to the first group and it is similar to the multi-criteria analysis methods of Oral *et al.* (2001) and Tavana (2003). To formulate the model algebraically, let us assume: $w_i$ as the importance weight of the *ith* criterion $(i = 1, 2, \ldots, m)$, $z_i^{jkf}$ the preference rating of the *kth* technology initiative for the *ith* criterion in the *fth* phase of the evaluation process by the *jth* decision-maker $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n;\ k = 1, 2, \ldots, p;$ and $f = 1, 2, \ldots, q)$, $S^{kf}$ the overall score of the *kth* technology initiative in the *fth* phase of the evaluation process $(k = 1, 2, \ldots, p;$ and $f = 1, 2, \ldots, q)$, and $\tilde{S}^f$ the median overall score for the *fth* phase of the evaluation process $(f = 1, 2, \ldots, q)$ where $0 \leq w_i \leq 1$; $0 \leq z_i^{jkf} \leq 10$; and $0 \leq S^{kj} \leq 10$.

Assuming that $a^{jkf}$ is the alternative representing the *jth* decision-maker and the *kth* technology initiative in the *fth* phase of the evaluation process and $Z = \{z_i^{jkj}\}$, represents the preference rating of the *kth* technology initiative for the *ith* criterion in the *fth* phase of the evaluation process by the *jth* decision-maker:

$$\begin{array}{ccccc}
w_i & a^{111} & a^{211} & \ldots & a^{npq} \\
w_1 & z_1^{111} & z_1^{211} & \ldots & z_1^{npq} \\
w_2 & z_2^{111} & z_1^{211} & \ldots & z_1^{npq} \\
\vdots & \vdots & \vdots & \ldots & \vdots \\
w_m & z_m^{111} & z_m^{211} & \ldots & z_m^{npq}
\end{array}$$

$$S^{kf} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} w_i z_i^{jkf}}{n}, \quad (i = 1, 2, \ldots, m;\ \text{and}\ j = 1, 2, \ldots, n).$$

$$S^{kf} \geq \tilde{S}^f \quad \text{(Grant certification)}$$

$$S^{kf} < \tilde{S}^f \quad \text{(Deny certification)}$$

The MCCS architecture team identified $i$ criteria to be used as major evaluation criteria when early proof-of-concept work was under consideration. The importance weight of each criterion $(w_i)$ was captured and measured with AHP using the questionnaire presented in Appendix 1. The MCCS architecture team was asked to provide their subjective assessment of each pairwise comparison. Assuming that in an MCCS architecture team member's mind, $c_1, c_2, \ldots, c_i$ are the $i$th criteria that contribute to a technology initiative's success. The team member's goal is to assess the relative importance of these factors.

Saaty's AHP (Forman and Gass, 2001; Saaty and Vargas, 1998) is a method of deriving a set of weights to be associated with each of the $i$th criteria. Initially, the team member is asked to compare each possible pair $c_j$, $c_k$ of criteria, and provide quantified judgments on which one of the criteria is more important and by how much. These judgments are represented by an $i \times i$ matrix:

$$A = (a_{jk}) \quad (j, k = 1, 2, \ldots, i)$$

If $c_j$ is judged to be of equal importance as $c_k$, then $a_{jk} = 1$; if $c_j$ is judged to be more important than $c_k$, then $a_{jk} > 1$; and if $c_j$ is judged to be less important than $c_k$, then $a_{jk} < 1$

$$a_{jk} = 1/a_{kj}, \quad a_{jk} \neq 0$$

Thus, the matrix $A$ is a reciprocal matrix, that is the entry $a_{jk}$ is the inverse of the entry $a_{kj}$. $a_{jk}$ reflects the relative importance of $c_j$ compared with criterion $c_k$. For example, $a_{12} = 1.25$ indicates that $c_1$ is 1.25 times as important as $c_2$.

Then the vector $w$ representing the relative weights of each of the $i$th criteria can be obtained by computing the normalized eigenvector corresponding to the maximum eigenvalue of matrix $A$. An eigenvalue of $A$ is defined as $\lambda$ which satisfies the following matrix equation:

$$Aw = \lambda w$$

where $\lambda$ is a constant, called the eigenvalue, associated with the given eigenvector $w$. Saaty (1977, 1980, 1983, 1990, 1994) has shown that the best estimate of $w$ is the one associated with the maximum eigenvalue ($\lambda_{\max}$) of the matrix $A$. Since the sum of the weights should be equal to 1.00, the normalized eigenvector is used. Saaty's algorithm for obtaining this $w$ is incorporated in the software Expert Choice (2000).

One of the advantages of AHP is that it encourages team 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, $\lambda_{\max}$, should equal $i$, the number of criteria that are compared. In general, the responses are not perfectly consistent, and $\lambda_{\max}$ is greater than $n$. The larger the $\lambda_{\max}$, the greater is the degree of inconsistency. Saaty defines the consistency index (CI) as $(\lambda_{\max} - i)/(i - 1)$, and provides the following random index (RI) table for matrices of order 3-10.

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|
| RI | 0.58 | 0.90 | 1.12 | 1.32 | 1.41 | 1.45 | 1.49 | 1.51 |

This RI is 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 team member is made aware that his or her pairwise comparisons are logically inconsistent, and he or she is encouraged to revise them.

The responses were processed with Expert Choice (2000), and those with inconsistency ratios greater than 0.10 were asked to revisit their judgments as it is suggested by Satty. The mean importance weights were calculated for the MCCS architecture team after necessary adjustments were made to inconsistent responses. Each MCCS architecture team member was presented with his/her individual score along with the group mean weights. IT members were given the opportunity to revisit their judgments and make revisions to their pairwise comparison scores based on this feedback. Some MCCS architecture team members took advantage of this opportunity and revised their judgments in the second round. The mean importance weights for the first and second round are presented in Table II. As it is shown, the second round results were slightly different from the first round results.

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. The geometric aggregation rule is used to avoid the controversies associated with rank reversal (Dyer, 1990a, b; Harker and Vargas, 1990; Saaty, 1990; Tavana, 2002).

## 4. A technology certification decision scenario at NASA

Quest 123 is a four-phase filtering process utilized at NASA to certify technology initiatives in three test labs before final deployment to the MCC. Figure 1 shows a graphical presentation of Quest 123. The process begins with an initial screening in phase I ($f = 1$). All $j$ MCCS architecture team members individually assess the $k$th technology initiative under consideration by providing a performance rating $\left(z_i^{jk1}\right)$ between 0 and 10 for each of the $i$th evaluation criteria adopted by MCC. All individual performance ratings in phase I are combined with the previously established performance weights of criteria ($w_i$) using the model presented earlier to calculate an overall score ($S^{k1}$) for the $k$th initiative in phase I. This score is compared with the median score ($\tilde{S}^1$) of the first phase, retrieved from Quest database. If $S^{k1} \geq \tilde{S}^1$, the initiative is selected for phase II testing in Quest 1 lab. If $S^{k1} < \tilde{S}^1$, the initiative is rejected for further testing. The Quest database is updated with all technology initiative scores, whether the initiative is accepted or rejected. The inclusion of new entries into the Quest database results in recalculation of the median score.

A Quest database was initially created with 112 records for phase I ($\tilde{S}^1 = 5.64$), 48 records for phase II ($\tilde{S}^2 = 6.48$), 25 records for phase III ($\tilde{S}^3 = 7.36$), and nine records for phase IV ($\tilde{S}^4 = 8.14$). The first technology initiative considered by Quest 123 was the 113th initiative. The MCCS architecture team members individually evaluated this

| Evaluation factor | Round 1 | Round 2 |
|---|---|---|
| Ease of use and effectiveness | 0.296 | 0.293 |
| Systems perspective | 0.184 | 0.191 |
| Flexibility | 0.140 | 0.132 |
| Access | 0.086 | 0.083 |
| Connectivity | 0.091 | 0.091 |
| Change | 0.088 | 0.084 |
| Cost reduction | 0.116 | 0.126 |
| Inconsistency ratio | 0.081 | 0.068 |

**Table II.**
First and second round importance weights

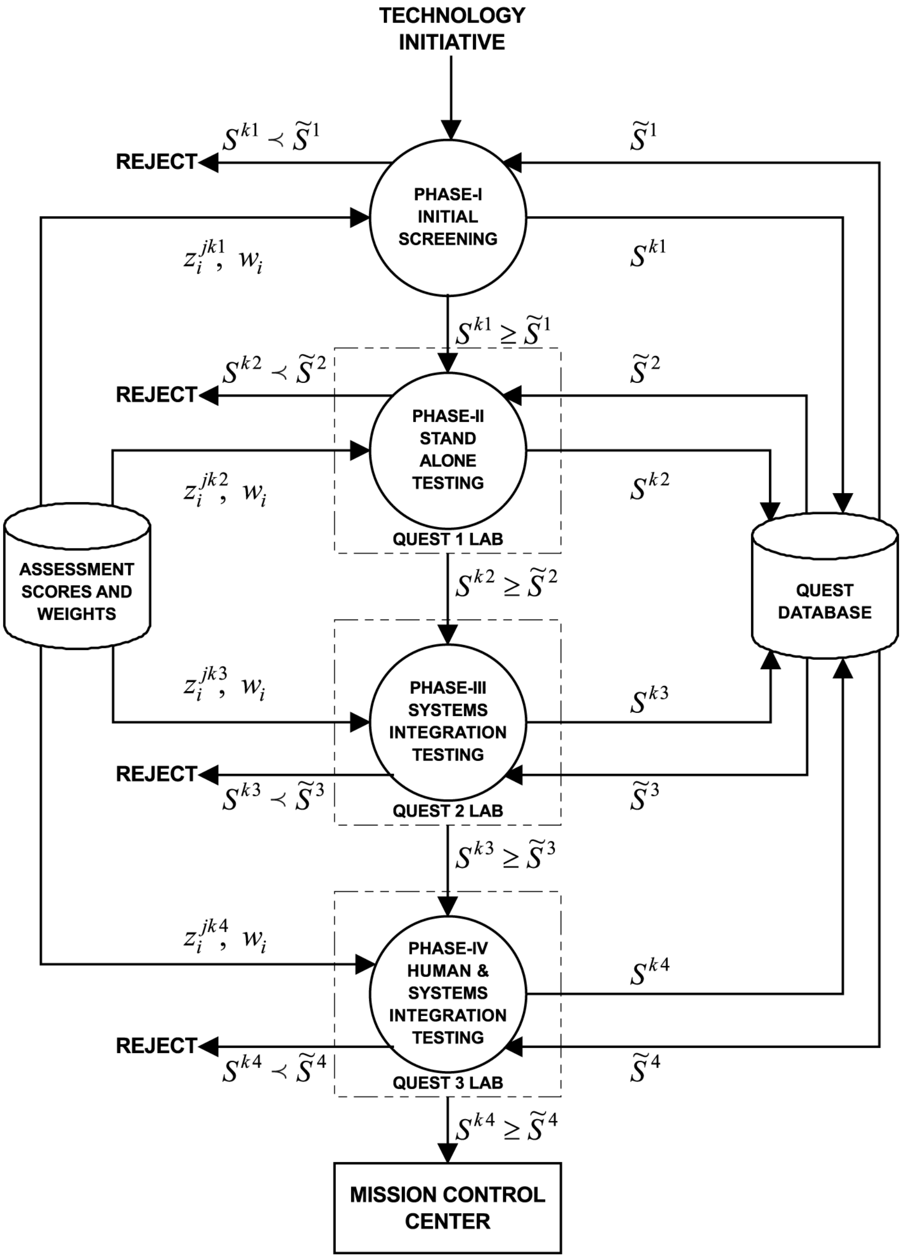

**Figure 1.**
New technology assessment process

technology initiative and assigned a numerical preference rating between 0 and 10 for each of the seven evaluation criteria adopted by MCC. The combination of individual performance ratings with the overall performance weights of criteria using the model presented earlier resulted in an overall score for each team member. The median score of the individual team member scores were used as the overall score for the 113th initiative in phase I ($S^{113,1} = 6.76$). Since $S^{113,1} \geq \tilde{S}^1$, the 113th technology initiative was approved for adoption and stand alone testing in Quest 1. As a result, the number of cases in phase I was increased by 1-113 and a new median was recalculated for this phase ($\tilde{S}^1 = 5.67$).

The MCCS architecture team members individually performed the stand alone testing on this initiative during phase II in Quest 1. Again, a numerical preference rating between 0 and 10 for each of the seven evaluation criteria adopted by MCC was given to this initiative for stand alone testing. Aggregation of the individual scores produced a median score of 6.94 for this initiative in phase II ($S^{113,2} = 6.94$). Since $S^{113,2} \geq \tilde{S}^2$, the 113th technology initiative was approved for adoption and systems integration testing in Quest 2. As a result, the number of cases in phase II was increased by 1 to 49 and a new median was recalculated for this phase ($\tilde{S}^2 = 6.54$).

The MCCS architecture team members individually performed systems integration testing on this initiative during phase III in Quest 2. Again, a numerical preference rating between 0 and 10 for each of the seven evaluation criteria adopted by MCC was given to this initiative for systems integration testing. Aggregation of the individual scores produced a median score of 6.86 for this initiative in phase III ($S^{113,3} = 6.86$). Since $S^{113,3} < \tilde{S}^3$, the 113th technology initiative was denied certification and removed from the process and further testing. It should be noted that deletion of this initiative does not change the number of cases and the median of phase III. All technology initiatives are similarly assessed by Quest 123.

The process describes here is an iterative one whereby the median is used as a filter to separate the higher ranked initiatives from the lower ranked initiatives that were evaluated in the Quest labs. Initially, a certain number of initiatives qualify based on the overall median and are sent to Quest 1 in phase I. The scores are re-evaluated in Quest 1 during phase II and those ranked above the median are allowed to proceed to Quest 2 for phase III testing. Similar process is repeated for all four phases in all three Quest labs until the technology initiative is certified to be deployed to the MCC.

## 5. Evaluation of Quest 123

Prior to implementation of Quest 123, the MCCS architecture team used a simple scoring method described earlier to evaluate new technology initiatives at MCC. One to the lack of a systematic process and inadequate thorough analysis, MCC management and the team were concerned about the potential for inconsistency and poor decisions. The team decided to replace the current simple scoring method with a more complete, structured, and analytical process.

Quest 123 was evaluated by 13 MCCS architecture team members. The questionnaire, presented in Appendix 2, was used to compare the post-quest 123 with the system currently used by the MCCS architecture team. Team members were asked to compare the two systems using the following seven evaluation factors adopted by the team.

1. *Flexibility*. The number of criteria, technology initiatives, decision-makers, or phases of the process at MCC is not predetermined. The new system should be flexible enough to be able to handle these changes.
2. *Simplicity*. The new system needs to be theoretically sound and at the same time easy to understand by users and management.

(3) *Analytical*. The new system has to be analytical and capable of evaluating multi-criteria problems.

(4) *Completeness*. The new system should have facilities to collect and integrate a wide range of information concerning change, flexibility, connectivity, access, cost reduction, ease of use, effectiveness, and systems perspective.

(5) *Structured*. The new system should have manageable steps with a facility to integrate the results from each step.

(6) *Decision process quality*. The new system should improve decision process quality.

(7) *Decision quality*. The new system should improve the decision quality.

The results shown in Figure 2 show that post-quest 123 scores are higher on all of the assessment dimensions. Decision quality with a median score of seven followed by decision process quality and structured with a median score of six were strongly favored in the post-quest 123 over the pre-quest 123 system. Completeness with a score of five and analytical with a score of three for post-quest 123 were perceived moderately better than the pre-quest approach. Finally, simplicity and flexibility with median scores of two and one for post-quest 123 were perceived slightly better than pre-quest approach by the MCCS architecture team members.

Next, the Wilcoxon signed ranks test of medians was performed. All the medians were statistically different from zero at $\alpha = 0.05$. A Kruskal-Wallis test also shows that all the medians were significantly different at $\alpha = 0.05$ (Table III). These results indicate that the post-quest system was preferred to the pre-quest system on all the criteria.

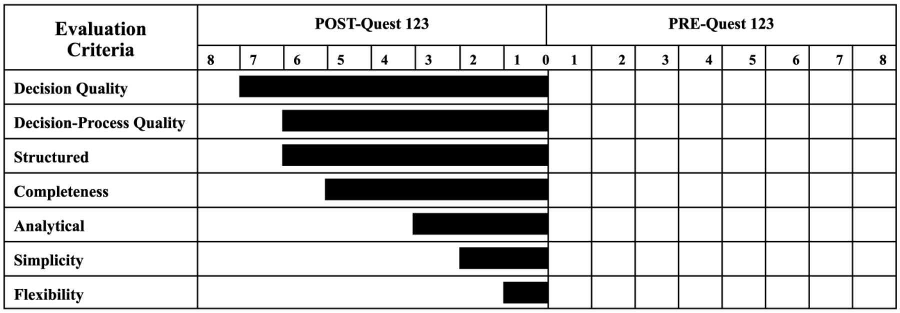

**Figure 2.**
MCCS architecture team median scores

| Evaluation factor | Median | Significant ($\alpha = 0.05$) |
|---|---|---|
| Decision quality | 7 | Yes |
| Decision process quality | 6 | Yes |
| Structured | 6 | Yes |
| Completeness | 5 | Yes |
| Analytical | 3 | Yes |
| Simplicity | 2 | Yes |
| Flexibility | 1 | Yes |

**Table III.**
Post-quest 123 test of medians

## 6. Conclusion

Advances in computer technology and data availability have made decision making more complex than ever. Quest 123 uses AHP and an additive MCDM model to reduce these complexities by systematically structuring the technology assessment problem into manageable and transparent steps at MCC. The earlier assessment model used by MCC was highly intuitive and unstructured. While intuition is still favored by practicing managers, it may be dangerously unreliable when used to solve repetitive problems because of the potential for inconsistency. Quest 123 has several attractive features that address some of the limitations inherent in the earlier naive approach used by MCC.

It should be noted that using a structured framework like Quest 123 does not imply a deterministic approach to technology assessment. While Quest 123 enables MCCS architecture team members to crystallize their thoughts and organize their judgments, it should be used very carefully. Team member judgment is an integral component of Quest 123; therefore, the effectiveness of the model relies heavily on their cognitive abilities. Quest 123 utilizes intuitive methods such as AHP and MCDM for subjective estimation of weights and scores because they cannot be obtained by empirical analysis. While these judgments often mirror a team member's beliefs, they should be used with caution. As with any decision calculus model, the practicing managers must be aware of the limitations of subjective estimates. When empirical analysis is feasible and makes economic sense, it should be utilized to improve these estimates (Lodish, 1982).

On the other hand, Quest 123 only addresses some of the problems inherent in technology certification decisions. Quantification of all evaluation criteria is a difficult task. Storing information generated during a session for use in future sessions along with information on actual performance of a technology initiative can facilitate the process of learning from past mistakes. This may be done by interfacing Quest 123 with knowledge-based systems for fine tuning weights and subjective scores, and for inferring in a non-algorithmic way, for example by inductive reasoning. Neural networks also offer a potential for developing knowledge-based models in this area. Quest 123 presents a step toward developing a more comprehensive and structured framework in benchmarking.

## References

Blumberg, D.F. (1994), "Strategic benchmarking of service and logistic support operations", *Journal of Business Logistics*, Vol. 15 No. 2, pp. 89-119.

Camp, R.C. (1989), *Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance*, ASQC Quality Press, Milwaukee, WI.

Camp, R.C. (1995), *Business Process Benchmarking: Finding and Implementing Best Practices*, ASQC Quality Press, Milwaukee, WI.

Cassell, C., Nadin, S. and Gray, M.O. (2001), "The use and effectiveness of benchmarking in SMEs", *Benchmarking: An International Journal*, Vol. 8 No. 3, pp. 212-22.

Cox, J.R.W., Mann, L. and Samson, D. (1997), "Benchmarking as a mixed metaphor: disentangling assumptions of competition and collaboration", *Journal of Management Studies*, Vol. 34 No. 2, pp. 285-314.

Czuchry, A.J., Yasin, M.M. and Dorsch, J.J. (1995), "A review of benchmarking literature – a proposed model for implementation", *International Journal of Materials and Product Technology*, Vol. 10 No. 1-2, pp. 27-45.

Dorsch, J.J. and Yasin, M.M. (1998), "A framework for benchmarking in the public sector: literature review and directions for future research", *International Journal of Public Sector Management*, Vol. 11 No. 2-3, pp. 91-115.

Drew, J.E. (1997), "From knowledge to action: the impact of benchmarking on organizational performance", *Long Range Planning*, Vol. 30 No. 3, pp. 427-41.

Dyer, J.S. (1990a), "Remarks on the analytic hierarchy process", *Management Science*, Vol. 36 No. 3, pp. 249-58.

Dyer, J.S. (1990b), "A clarification of 'Remarks on the analytic hierarchy process'", *Management Science*, Vol. 36 No. 3, pp. 274-5.

Expert Choice (Computer Software) (2000), Decision Support Software, Inc., McLean, VA.

Forman, E.H. and Gass, S.I. (2001), "The analytic hierarchy process – an exposition", *Operations Research*, Vol. 49 No. 4, pp. 469-86.

Francis, G. and Holloway, J. (2002), "Beyond comparisons – the role for the operational researcher in benchmarking", *Journal of the Operational Research Society*, Vol. 53, pp. 283-91.

Frei, F.X. and Harker, P.T. (1999), "Measuring aggregate process performance using AHP 1", *European Journal of Operational Research*, Vol. 116 No. 2, pp. 436-42.

Geanuracos, J. and Meiklejohn, I. (1993), *Performance Measurement: The New Agenda*, Business Intelligence Ltd, London.

Harker, P.T. and Vargas, L.G. (1987), "The theory of ratio scale estimation: Saaty's analytic hierarchy process", *Management Science*, Vol. 33, pp. 1383-403.

Harker, P.T. and Vargas, L.G. (1990), "Reply to 'Remarks on the analytic hierarchy process'", *Management Science*, Vol. 36 No. 3, pp. 269-73, by J.S. Dyer.

Jackson, A.E., Safford, R.R. and Swart, W.W. (1994), "Roadmap to current benchmarking literature", *Journal of Management in Engineering*, Vol. 10 No. 6, pp. 60-7.

Jarrar, Y.F. and Zairi, M. (2001), "Future trends in benchmarking for competitive advantage: a global survey", *Total Quality Management*, Vol. 12 No. 7-8, pp. 906-12.

Korpela, J. and Tuominen, M. (1996), "Benchmarking logistics performance with an application of the analytic hierarchy process", *IEEE Transactions on Engineering Management*, Vol. 43 No. 3, pp. 323-32.

Landeghem, R.V. and Persoons, K. (2001), "Benchmarking of logistical operations based on a causal model", *International Journal of Operations & Production Management*, Vol. 21 No. 1-2, pp. 254-67.

Lodish, L.M. (1982), "Experience with decision calculus models and decision support systems", in Schulz, R. and Zoltners, A. (Eds), *Marketing Decision Models*, North-Holland, New York, NY.

Lu, M.H., Madu, C.N., Kuei, C.H. and Winokur, D. (1994), "Integrating QFD, AHP and benchmarking in strategic marketing", *The Journal of Business and Industrial Marketing*, Vol. 9 No. 1, pp. 41-50.

Min, H., Min, H. and Chung, K. (2002), "Dynamic benchmarking of hotel service quality", *Journal of Services Marketing*, Vol. 16 No. 4, pp. 302-19.

Oral, M., Kettani, O. and Cinar, U. (2001), "Project evaluation and selection in a network of collaboration: a consensual disaggregation multi-criteria approach", *European Journal of Operational Research*, Vol. 130, pp. 332-46.

Partovi, F.Y. (2001), "An analytic model to quantify strategic service vision", *International Journal of Service Industry Management*, Vol. 12 No. 5, pp. 476-99.

Prasanta, K.D. (2002), "Benchmarking project management practices of Caribbean organizations using analytic hierarchy process", *Benchmarking: An International Journal*, Vol. 9 No. 4, pp. 326-56.

Saaty, T.L. (1977), "A scaling method for priorities in hierarchical structures", *Journal of Mathematical Psychology*, Vol. 15, pp. 234-81.

Saaty, T.L. (1980), *The Analytic Hierarchy Process*, McGraw-Hill, New York, NY.

Saaty, T.L. (1983), "Axiomatic foundations of the Analytic Hierarchy Process", *Management Science*, Vol. 32 No. 2, pp. 841-55.

Saaty, T.L. (1989), "Decision making, scaling, and number crunching", *Decision Sciences*, Vol. 20, pp. 404-9.

Saaty, T.L. (1990), "An exposition of the AHP in reply to the paper 'Remarks on the analytic hierarchy process'", *Management Science*, Vol. 36 No. 3, pp. 259-68.

Saaty, T.L. (1994), "How to make a decision: the analytic hierarchy process", *Interfaces*, Vol. 24 No. 6, pp. 19-43.

Saaty, T.L. and Vargas, L.G. (1998), "Diagnosis with dependent symptoms: Bayes theorem and the analytic hierarchy process", *Operations Research*, Vol. 46 No. 4, pp. 491-502.

Sarkis, J. (2001), "Benchmarking for agility", *Benchmarking: An International Journal*, Vol. 8 No. 2, pp. 88-107.

Schefcyzk, M. (1993), "Industrial benchmarking: a case study of performance analysis techniques", *International Journal of Production Economics*, Vol. 32 No. 1, pp. 1-11.

Shim, J.P. (1989), "Bibliographical research on the analytic hierarchy process (AHP)", *Socio-Economic Planning Sciences*, Vol. 23, pp. 161-7.

Spendolini, M.J. (1992), *The Benchmarking Book*, AMACOM, New York, NY.

Tavana, M. (2002), "Euclid: strategic alternative assessment matrix", *Journal of Multi-Criteria Decision Analysis*, Vol. 11 No. 2, pp. 75-96.

Tavana, M. (2003), "A multi-criteria group-decision-making model for evaluating and prioritizing advanced-technology projects at NASA", *Interfaces*, Vol. 33 No. 3, pp. 40-56.

Zahedi, F. (1986), "The analytic hierarchy process – a survey of the method and its applications", *Interfaces*, Vol. 16, pp. 96-108.

## Further reading

Eyrich, H.G. (1991), "Benchmarking to become the best of breed", *Manufacturing Systems*, Vol. 9 No. 4, pp. 40-7.

## Technology Assessment Questionnaire

**INSTRUCTIONS:** Please respond by placing an 'X' on the appropriate location to indicate your perception of the relative importance of the **technology assessment factors** for each of the following pairwise comparisons.

*Placing an 'X' at equal indicates that factors A and B are equally important*

*Placing an 'X' to the left of equal indicates that factor A is more important than factor B*

*Placing an 'X' to the right of equal indicates that factor B is more important than factor A*

| TECHNOLOGY ASSESSMENT FACTOR A | Extreme | | Very Strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very Strong | | Extreme | TECHNOLOGY ASSESSMENT FACTOR B |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | |
| **Change** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Flexibility** |
| **Change** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Connectivity** |
| **Change** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Access** |
| **Change** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Cost Reduction** |
| **Change** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Ease of Use and Effectiveness** |
| **Change** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Systems Perspective** |
| **Flexibility** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Connectivity** |
| **Flexibility** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Access** |
| **Flexibility** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Cost Reduction** |
| **Flexibility** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Ease of Use and Effectiveness** |
| **Flexibility** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Systems Perspective** |
| **Connectivity** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Access** |
| **Connectivity** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Cost Reduction** |
| **Connectivity** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Ease of Use and Effectiveness** |
| **Connectivity** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Systems Perspective** |
| **Access** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Cost Reduction** |
| **Access** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Ease of Use and Effectiveness** |
| **Access** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Systems Perspective** |
| **Cost Reduction** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Ease of Use and Effectiveness** |
| **Cost Reduction** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Systems Perspective** |
| **Ease of Use and Effectiveness** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | **Systems Perspective** |

**Figure A1.**

## Appendix 2

### POST-Quest 123 Evaluation Questionnaire

**INSTRUCTIONS:** The following criteria are adopted by the MCCS Architecture Team to evaluate Quest 123 system. Please respond by placing an 'X' on the appropriate location to indicate your preference for the POST-Quest 123 system or PRE-Quest 123 system based on Flexibility, Simplicity, Analytical Capabilities, Completeness, Structured Framework, Decision-Process Quality, and Decision Quality.

***Placing an 'X' at equal indicates***
***NO PREFERENCE between Post-Quest123 and Pre-Quest 123***

***Placing an 'X' to the left of equal indicates that POST-Quest 123 is MORE PREFERRED than the PRE-Quest 123***

***Placing an 'X' to the right of equal indicates that PRE-Quest 123 is MORE PREFERRED than POST-Quest 123***

**<<<< POST-Quest 123 Preferred>>>>** **<<<< PRE-Quest 123 Preferred >>>>**

| | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Extreme | | Very Strong | | Strong | | Moderate | | Equal | | Moderate | | Strong | | Very Strong | | Extreme |
| **Flexibility** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| **Simplicity** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| **Analytical.** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| **Completeness** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| **Structured** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| **Decision-Process Quality** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| **Decision Quality** | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

**Figure A2.**