Distributed Collaborative Team Effectiveness: Measurement and Process Improvement

R. Wheeler, J. Hihn & B. Wilkinson
Jet Propulsion Laboratory/California Institute of Technology

ABSTRACT: This paper describes a measurement methodology developed for assessing the readiness, and identifying opportunities for improving the effectiveness, of distributed collaborative design team teams preparing to conduct a concurrent design session. In addition, the methodology can be extended to evaluate the impact of multiple new technology introductions on the effectiveness of the same design teams.

1 INRODUCTION

[1]The state of the art in engineering design is the use of collaborative engineering teams and more recently the use of distributive collaborative teams. NASA's Jet Propulsion Laboratory (JPL) has used a collaborative engineering team (Team X) to develop and evaluate deep space mission conceptual designs since 1995. In the past few years, JPL has been exploring the effectiveness of distributive design sessions with NASA's Goddard Space Flight Center (GSFC) and Johnson Space Center as well as the European Space Agency. The impact on JPL's ability to generate conceptual designs and mission proposals has been extensive. Tasks that used to require 6 to 12 months and as much as a million dollars now are completed in days to weeks and cost about $50K to complete. The approach has proven so successful that it has been expanded into the development of collaborative engineering teams for the design of instruments, ground systems and mission formulation teams. However, as these teams start to reach maturity and as problems occur in the early attempts with distributive teaming, questions start arising as to how to improve the effectiveness of these teams.

Improving the effectiveness of engineering teams requires improving the effectiveness of team members as well as how they interact or collaborate as a team. What is required is to systematically measure the impact of an intervention on the team's working processes as well as its current state. Unfortunately, few design teams attempt to measure their performance, thinking that it is impossible or too difficult.

This paper will discuss how our collaboration effectiveness measurement methodology was developed, as well as the general principles behind the design of the methodology and how it can be customized and applied to different collaborative applications. We will conclude with a brief discussion of some actual results arising from use of the methodology.

2 MEASUREMENT METHODOLGY

2.1 *Approach*

The objective of the design effort was to develop a measurement methodology that could be used to:

- Evaluate two or more team's "readiness" to conduct a distributed, concurrent design session;
- Prioritize opportunities for improvement in the collaborative design environment; and
- Assess the impact of the application of new technologies within the context of the collaborative design environment.

In addition, the measurement methodology had to enable relatively rapid scoring, grouping and displaying of results. The measurement instrument is a structured questionnaire designed to be applied to multiple design session phases and to enable multiple evaluator's displays of results. When implemented as an on-line survey capability, knowledgeable evaluators could complete it quickly and obtain immediate summaries and comparisons of their responses with other participants or to other time periods.

[1] The work described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology and Goddard Space Flight Center, National Aeronautics and Space Administration. JPL is under contract with the National Aeronautics and Space Administration.

Depending upon the actual implementation approach utilized, this rapid data collection methodology may sacrifice collecting the detailed information necessary for pinpointing specific process improvement opportunities. In such cases, it is assumed that once a problem area is identified utilizing the questionnaire, further investigations will be undertaken by the collaborative process managers to isolate the specific details of the problem and initiate actions to resolve or mitigate them. Alternatively, process improvement details can be collected on the questionnaire from the evaluators, but that will slow down the data collection process.

2.2 *Measurement Instrument Structure Overview*

The methodology described in this paper utilizes a structured measurement instrument that can be used to evaluate the preparation, the concurrent design session effectiveness[2], and, finally, the completion and close-out of the design session products. To comprehensively evaluate a particular design session's effectiveness, multiple attributes or dimensions must be assessed. The instrument described illustrates these multiple dimensions, based on concurrent space mission design sessions conducted by teams at JPL and GSFC. However, as discussed below, all components of the methodology, including all components of the questionnaire, rating scales and the rating approaches, can and should be customized for application to different collaborative environments and applications[3].

The instrument structure consists of **areas, attributes** and **attribute elements**. These are discussed in detail in the next section. The **areas** identify the major categories of importance to enabling an effective design session. For example, key preparation categories can include major categories such as defining customer objectives, evaluating the distributed collaborative infrastructure, or planning the design process. Key close-out categories can include issues of collaboration effectiveness and follow-up in addition to an evaluation of the process preparation effectiveness after the session has occurred. Each area consists of two or more **attributes**. The attributes are the key measurement categories for each area--they are the level at which the evaluators assign a score. Finally, each attribute contains **attribute elements.** The attribute elements are questions relevant to the attribute. They serve to remind the evaluator of the dimensions of the attribute.

[2] In the context of this paper, "improvements in effectiveness" is used in its broadest sense; i.e., it can refer to improvements in the quality of products produced by the collaborative team and/or improvements in the time taken to produce those products.

[3] Anyone who is interested in obtaining the JPL forms can contact R. Wheeler. The forms will need to be customized to your organizational environment.

2.3 *Populating the Questionnaire*

The areas, attributes and attribute elements illustrated in this paper were derived from systematic observations, and participant debriefings, of actual concurrent mission design sessions conducted by teams at JPL and GSFC[4]. Both organizations regularly conduct concurrent design sessions lasting about a week. The design teams are similar in structure, consisting of ten or more mission and spacecraft design disciplines, each assigned to a particular workstation in a concurrent design facility. Both organizations utilize a facilitator to guide and manage the concurrent design session in their respective facilities.

Two collaborative design sessions utilizing both teams were observed, each lasting about a week. They involved mission and spacecraft design proposals being jointly developed by the two organizations. Experienced observers were stationed at both facilities during the design sessions. Detailed observations describing the interactions among collaborators and their supporting infrastructure were recorded in real time by the observers. In addition, they interviewed design session participants when they had available time concerning their impressions of the collaboration effectiveness and supporting technology infrastructure. Several weeks after the design sessions were complete, the same observers interviewed key participants, including the facilitators and the customers, asking them to comment in retrospect about the design session effectiveness: what worked, what didn't work and what needed improvement.

The results of the design session observations and interviews were grouped into the categories described in the measurement instrument structure described above. The specific areas and attributes identified in this case study are illustrated in Figure 1. Five areas[5] -- defining customer objectives, getting technical requirements, planning the distributed collaborative design process, evaluating the distributed collaborative infrastructure and getting ready for the distributed collaborative design session are used for a pre-session evaluation to assess team readiness. These areas are used again in the post-session evaluation to assess preparation effectiveness. In addition, five other areas--initial briefings, issues of leadership, team dynamics, process dynamics and follow up are added to assess design session and close-out effectiveness.

Figure 2 illustrates examples of the attribute element questions. The element questions are mainly used to remind the evaluator of likely dimensions of the attribute. When he or she rates the attribute, he

[4] Special thanks to Ellen Herring and John Martin of GSFC for participating in the data collection and reduction process described in this example.

[5] Only three areas are illustrated in figure 1.

may want to note what element of the was considered the key "driver" behind the attribute rating.

Area	Attribute
A1 Defining Customer Objectives	Purpose of collaboration Definition of design session objectives Competing objectives of multiple customers
A2 Getting Technical Requirements	Requirements negotiation Documentation of technical requirements Pre-existing designs Technology needs
A3 Planning the Distributed Collaborative Design Process	Cultural evaluation Roles and responsibilities Design session process Customer Availability
B1 Initial Briefings	Pre-session agreements and information Customer (Technical) Briefing
B2 Issues of Leadership	Design Session Management Team Management
B3 Team Dynamics	Team Unity
B4 Process Dynamics	Evolution of the group process Benefits of collaboration
B5 Follow up	Achievement of session objectives

Figure 1: Some Key Areas and Attributes

2.4 *Scoring*

The measurement instrument is exercised by utilizing two or more evaluators to assign a score to each attribute identified in the pre- or post-design session instrument. The selection of evaluators and the type of rating scale depend upon the desired final use of the results. In the simpliest case, evaluators simply might want to understand the readiness of participants for a distributed design session. Application of the measurement methodology might utilize local and distal team leaders to rate each of the attribute elements using a three-point scale (say, high, medium or low readiness). Another application might be a pre- and post session comparison to rate the impact of some change introduced into the design process. The scorers in that case might be the team members. The most sophisticated use of this methodology would be for selecting and evaluating the introduction of multiple technology and process upgrades into the design team environment. Application of this approach would likely require the use of multiple evaluators trained in the application of the measurement instrument. It would also require more discriminating rating scales of at least five-points to ten-points, as illustrated in the hypothetical example in Figure 3.

Area	Attribute	Attribute Elements
A1. Defining Customer Objectives	**A1.1 Purpose of collaboration**	Is there a clear reason for the collaborative effort? Will the session be a design of a "new" system or review of an existing system?
	A1.2 Definition of design session objectives	Have the distributed team leaders/facilitators met with the sponsor to determine the objectives of the collaborative design session? Do the team leaders agree on the objectives of the session? Have the team leaders documented agreed-upon objectives of the design session?
A4. Evaluating the Distributed Collaborative Infrastructure	**A4.1 Procedures for communicating between teams**	Have the communications procedures been worked out? Have speaking and identity protocols been developed and have the teams been trained in their use? Are there mechanisms and procedures for electronically contacting individuals separately from the group?
	A4.4 Data/Information Electronic Exchange	Can all information be shared and reviewed by the group in real time? How will electronically generated information be exchanged and displayed in real time? ... Are groupware and procedures identified that enable "screen" sharing between distributed sites?...
	A4.5 Shared tools	Is there common software that team members can use for collaboratively developing their designs and products? If different tools are used, are there agreed-to standards for exchanging resulting products in real time? Is there a tool unique to one team that will be used in the design session?

Figure 2: Measurement Instrument Structure

3 APPLICATIONS [6]

Regardless of the scale used, the results can be rank ordered by attributes, evaluators, gaps between the evaluators or by time periods. In all the examples illustrated below, each attribute is rated using two ten point rating scales (see Figure 3). The first scale is always attribute importance. The second scale measures the outcome parameter of interest, for example, readiness (or capability) or collaboration effectiveness. A ten point rating scale is rarely used in formal surveys because human subjects cannot reliably discriminate ten distinct intervals. However, in this case, because relative comparisons rather than absolute outcomes are the parameters of interest, a ten point scale is used to help "spread out" the outcome scores. The reader should be cautioned that a five to seven point scale should be used if absolute scores are more important than relative comparisons.

[6] All of the scoring examples provided in this paper are hypothetical.

Area	Attribute	Attribute Elements	Attribute Importance	Attribute Readiness	Opportunity for Process Improvement	Process Improvement Opportunity Score	Readiness Score
			Importance Scale 0 1 2 3 4 5 6 7 8 9 10 (Low Importance → High Importance)	Capability Scale 0 1 2 3 4 5 6 7 8 9 10 (Low Capability → High Capability)	Maximum Capability minus Existing Capability	Opportunity for Improvement X Importance	Capability X Importance
A1. Defining Customer Objectives	A1.1 Purpose of collaboration	Is there a clear reason for the collaborative effort? Will the session be a design of a "new" system or review of an existing system?	7	5	5	35	35
	A1.2 Definition of design session objectives	Have the distributed team leaders/facilitators met with the sponsor to determine the objectives of the collaborative design session? Do the team leaders agree on the objectives of the session? Have the team leaders documented agreed-upon objectives of the design session?	8	8	2	16	64
A4. Evaluating the Distributed Collaborative Infrastructure	A4.1 Procedures for communicating between teams	Have the communications procedures been worked out? Have speaking and identity protocols been developed and have the teams been trained in their use? Are there mechanisms and procedures for electronically contacting individuals separately from the group?	4	6	4	16	24
	A4.4 Data/Information Electronic Exchange	Can all information be shared and reviewed by the group in real time? How will electronically generated information be exchanged and displayed in real time? ... Are groupware and procedures identified that enable "screen" sharing between distributed sites?	9	10	0	0	90
	A4.5 Shared tools	Is there common software that team members can use for collaboratively developing their designs and products? If different tools are used, are there agreed-to standards for exchanging resulting products in real time? Is there a tool unique to one team that will be used in the design session?	10	4	6	60	36

Figure 3: Scoring Methodology for Evaluating Design Session Preparation

Figure 4 illustrates an example of applying the methodology to preparing a "readiness assessment" for an upcoming concurrent design session using distributed design teams. In this case appropriate team or process managers complete the questionnaire, scoring each attribute on *importance* and *readiness* of their design team for the upcoming design session. The final score for each attribute is obtained by multiplying the *importance* score by the *readiness* (or capability) score. In the example illustrated, all of the evaluators scores are averaged, although gaps can be identified by displaying individual evaluators scores. By agreeing in advance to an arbitrary cut-off point (say, 30 is selected as the "minimum readiness cutoff" level), process manag

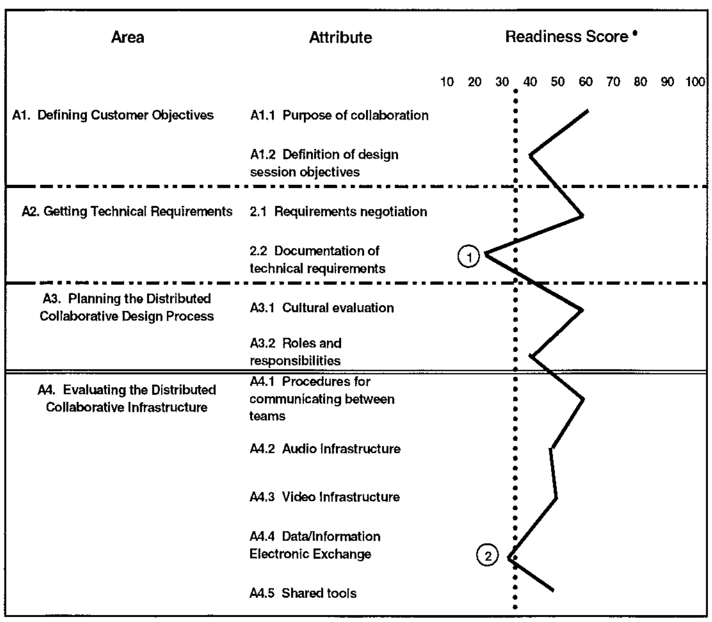

Figure 4: Readiness Assessment Results

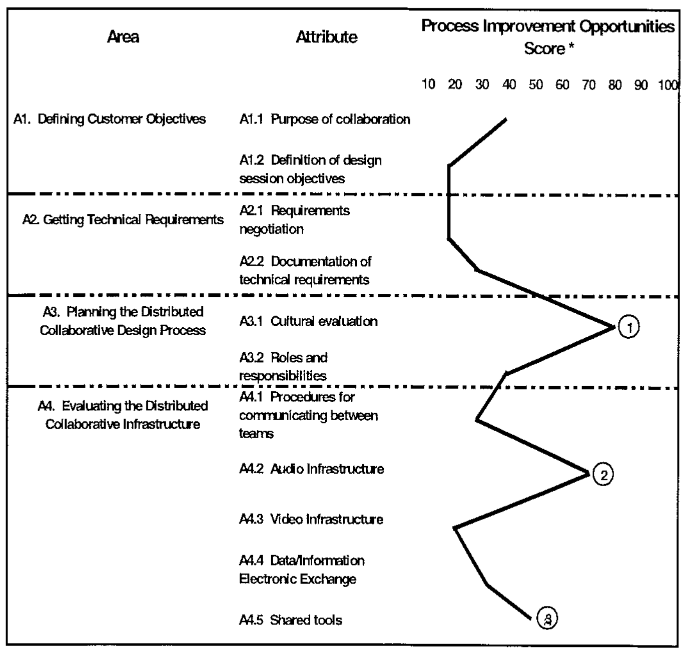

Figure 5: Process Improvement Opportunities

ers can identify for discussion and possible mitigation in advance of the design session specific areas the teams are not ready to effectively collaborate.

Figure 5 illustrates an example of applying the methodology to identifying process improvement opportunities prior to a concurrent design session. The scoring and scorers are identical to that discussed above. However, this time the final score is calculated by multiplying the importance score by the opportunity for process improvement score. This score is merely the inverse of the readiness score, that is, the maximum score that can be obtained for readiness (in this case, 10) minus the actual score for that attribute. The idea is that the further away a particular attribute is from it's optimal capability, the greater the opportunity to improve it. Of course, the score has to be weighted by its relative importance so that the most important opportunities with the highest yield improvement opportunities are identified.

Evaluation of the design session effectiveness employ pre and post scoring of the concurrent design session. Figure 6 illustrates an example of the scoring methodology for evaluating design session effectiveness. Process preparation effectiveness is scored before and after the design session; collaboration effective effectiveness is scored after the design session.

Previously Scored

Not Scored

Area	Effectiveness Attribute	Attribute Elements	Readiness Score	Attribute Importance	Attribute Collaboration Effectiveness	Opportunity for Program Improvement	Effectiveness Score	Program Improvement Opportunity Score
			Capability X Importance	Importance Scale 0 1 2 3 4 5 6 7 8 9 10	Effectiveness Scale 0 1 2 3 4 5 6 7 8 9 10	Maximum Effectiveness minus Existing Effectiveness	Effectiveness X Importance	Opportunity for Improvement X Importance
B1.1 Defining Customer Objectives	B1.1.1 Purpose of collaboration	Was there a clear reason for the collaborative effort? Was the session be a design of a "new" system or review of an existing system?	35	7	8	2	56	14
B1.4 Evaluating the Distributed Collaborative Infrastructure	B14.1 Procedures for communicating between teams	Have the communications procedures been worked out? Have speaking and identity protocols been developed and have the teams been trained in their use? Are there mechanisms and procedures for electronically contacting individuals separately from the group?	24	4	6	4	24	16
	B1.4.4 Data/Information Electronic Exchange	Can all information be shared and reviewed by the group in real time? How will electronically generated information be exchanged and displayed in real time? ... Are groupware and procedures identified that enable "screen" sharing between distributed sites?	90	10	6	4	60	40
B2: Collaborative Effectiveness								
B2.1 Initial Briefings	B2.1.1 Pre-session agreements and information	Were all pre-session agreements and information clearly conveyed to all design team members before or at the beginning of the design session?		6	5	5	7	30
	B2.1.2 Customer (Technical) Briefing	Did the sponsor brief both teams at the beginning of the collaborative design session about the technical objectives and issues? Were completing objectives clearly identified and prioritized? Were the reasons for collaboration clearly explained to the design teams?		5	3	7	3	35

Figure 6: Scoring Methodology for Evaluating Desing Session Effectiveness

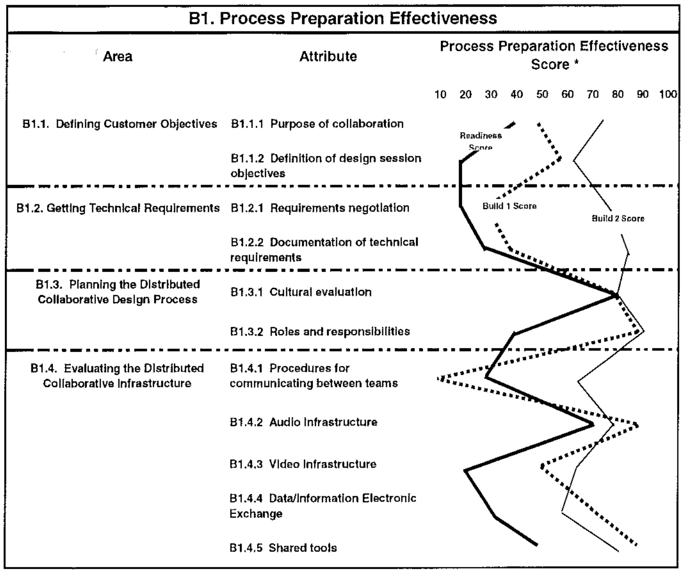

Figure 7: Preparation Improvements over Time

Figures 7 and 8 illustrate an application of this methodology to multiple technology builds upgrades. Readiness is assessed prior to the first build and thereafter after each build or upgrade delivery. Design session effectiveness is assessed after a design session is conducted utilizing a new upgrade.

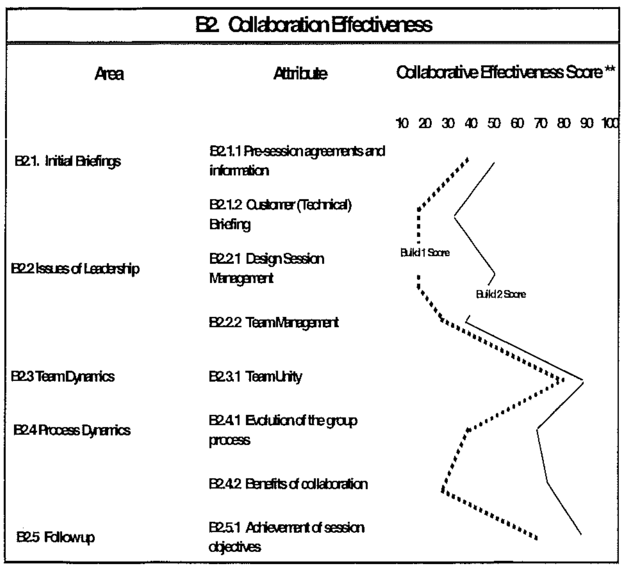

Figure 8: Design Session Effectiveness Improvements over Time

INITIAL STUDY RESULTS

During the first JPL-GSFC study, problems arose with virtually every aspect of the design session including requirements specification, audio communication, video communication, major process differences, tool differences, and basic file transfer. In spite of all of this, the study was concluded relatively successfully.

Finally, this methodology can be used to adjust program strategies as illustrated in figure 9.

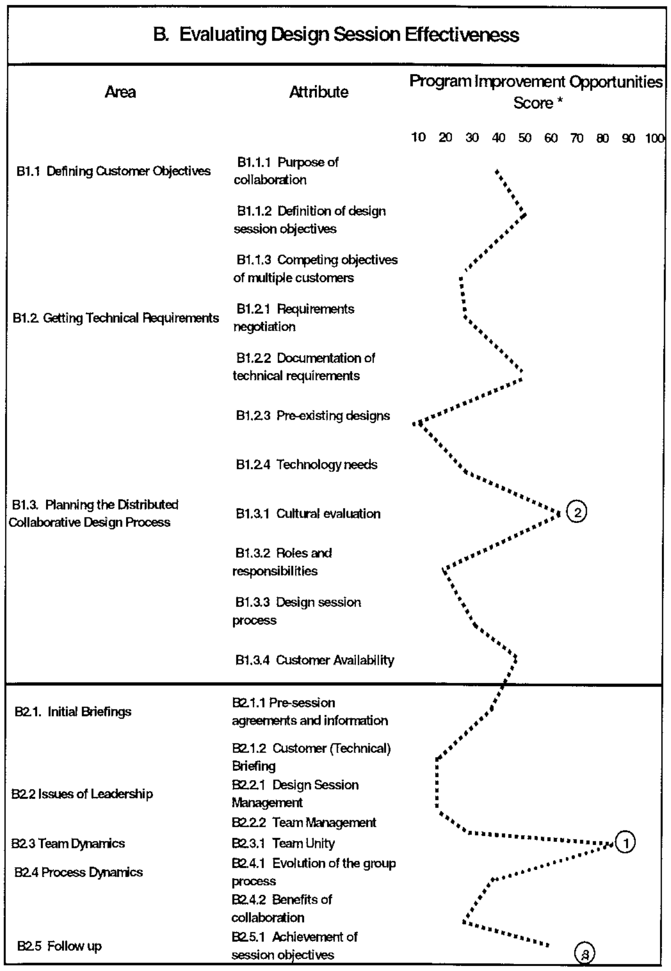

Figure 9: Identifying Priorities

Later analysis indicated that the collaborative facilities needed to be rearranged and the design process needed to be modified to better support distributive collaborative design. The main changes recommended were:

1. High quality audio communication with the ability to make presentations to both teams simultaneously and the ability to support multiple sidebars (small groups formed in real time to solve issues that crop up during the design session).
2. Clear simultaneous presentation of the products being developed with the ability to edit from both sites.
3. Well defined communication process and roles to reduce confusion.
4. Better pre-session work and clear customer requirements.

IMPLEMENTATION

Based on what was learned from the initial study and the development of the measurement instrument it was realized that the methodology need to be standardized and made available via computer displays for ease and efficiency of data input and display of the results. The measurement tool and methodology described in this paper has been implemented as a web-based on-line questionnaire to assess the "readiness" of JPL and GSFC design teams to conduct a future collaborative design session[7]. Each team facilitator or system engineer can log on and complete the questionnaire assessing his team's readiness for the upcoming design session. For each rated area, the evaluator can add specific comments relevant to the rating assigned. When both questionnaires are complete, multiple displays of the results can be obtained--for example, a rank ordering of the least to the most ready attributes from the perspective of one of the evaluators or the same data rank ordered by the largest gaps between the evaluators. The value of the approach is that the leaders can get together to discuss and resolve or mitigate lack of preparation problems prior to the actual design session.

LESSONS LEARNED

To date two key lessons learned have arisen as the measurement instrument has been used to evaluate and assess actual distributive collaborative design sessions. In some cases, the team lead could not answer a number of the key preliminary questions that were originally designed for him, particularly the detailed facility questions. As a result, the data entry interface is being revised to enable partial data entry by people with specialized knowledge. For example, the facility manager/center engineer can complete only the Evaluating the Distributed Collaborative Infrastructure portion of the questionnaire. Another issue that arose is that no input was obtained from the customer in the original methodology. To address this issue, the next version of the on-line questionnaire will enable customers to rate team readiness and importance in the Defining Customer Objectives and Getting Technical Requirements sections. Introducing the customer as a scorer has caused a new problem in displaying the results. The original scoring methodology focused specifically on distributed team's readiness to conduct and design session as perceived by their team leaders. Collecting input from the customer now adds a new dimension for comparison that will require the development of new data displays. We are in the process of working out how to summarize and display the results from three rather than tow different perspectives.

The second major lesson learned was that, currently, the majority of design sessions do not involve the coordination of two distributed design teams but rather the participation of a number of individuals, or small groups, remotely connecting into a main design team session using a workstation terminal. These connections can be from multiple sites, consisting of participants representing different combinations of the customer teams, (sub) contractors, and engineering or science specialists. The availability of minimal remote site capabilities greatly simplifies the assessment as the complexity of team member interactions and possible corrective actions are greatly reduced. For example, when there is only one design team, remote participants default to that team's design process and standard products. Secondly, from a technology assessment perspective significant infrastructure investments on the part of the individual remote participants are not feasible. The focus becomes one primarily of improved data transfer to the remote participants and possible video on computer displays.

Based on our experience to date, we conclude that this assessment tool needs to be a "living tool" that is constantly reviewed and revised as the needs and objectives of distributed design session process.

GENERALIZATION AND NEW APPLICATIONS

All elements of this methodology can be generalized and applied to new applications. New areas, attributes and attribute elements can be derived from observations of new case studies utilization the approach as described above. Selection of the scale to be used in rating the attributes should depend on the use of the results as discussed in the applications section above. To evaluate the effectiveness of multiple technology upgrades, the attributes must be kept constant and the use of impartial evaluators will likely be required.

CONCLUSIONS

It took a bit less then $30,000 to observe the sessions (with travel), conduct interviews, identify and reduce the categories, attributes and define the questions. It is estimated that a new organization building on the methodology discussed in this paper and forms should be able to do an initial assessment including form modification for significantly less. For larger projects with multiple companies and maybe even international partners, the development cost could be more. The original budget for this task from NASA's Intelligent Synthesis Environment program was only $130 K (1 month additional support from each center would have been required) to implement this at six NASA centers. Hence, for a large organization with projects in the tens to hundreds of millions this is a very affordable task with potentially huge payoff.

[7] Implemented by Carmel Cortney at NASA's GSFC.