
N A S A

N A S A

**AN OVERVIEW OF THE EVALUATION
PLAN FOR PC/MISI - PC-BASED
MULTIPLE INFORMATION SYSTEM INTERFACE**

Bee Lee Lim and Philip P. Hall

**The University of Southwestern Louisiana
Center for Advanced Computer Studies
Lafayette, Louisiana**

April 24, 1985

AN OVERVIEW OF THE EVALUATION
PLAN FOR PC/MISI - PC-BASED
MULTIPLE INFORMATION SYSTEM INTERFACE

ABSTRACT

This document represents an initial evaluation plan for the personal computer multiple information system interface (PC/MISI) project.

The document is intended to be used as a blueprint for the evaluation of this system and each objective of the design project is discussed along with the evaluation parameters and methodology to be used in the evaluation of the implementation's achievement of those objectives.

The potential of the system for research activities related to more general aspects of information retrieval is also discussed.

TABLE OF CONTENTS

I.	INTRODUCTION	4
II.	GENERAL OBJECTIVES	5
III.	GENERAL EVALUATION METHODOLOGY	6
IV.	USER CATEGORIZATION.	8
V.	SPECIFIC OBJECTIVES IN EVALUATING PC/MISI.	9
VI.	EVALUATION OF DESIGN OBJECTIVES.	13
	6.1 Evaluation Parameters and Methodology	13
	6.2 Miscellaneous Parameters.	23
	6.3 Evaluating Impact of User Experience.	23
	REFERENCES	26

I. INTRODUCTION

The objective of this document is to describe a plan for the evaluation of the Personal Computer-Based Multiple Information System Interface (PC/MISI) System currently being developed at the University of Southwestern Louisiana. The design plan from which the system is being developed has been described in a previous document [Hall, 84]. The system is intended to provide a much easier framework which casual users can utilize to access remote information sources. A common interface is provided through which users can access multiple systems. The objectives of this project, as described in the design plan, include the development of a system which can be used for research activities into various problems associated with providing access to information stored in IS&R systems. The evaluation activities which are presented in the remainder of this document include a plan for the evaluation of the effectiveness of the system itself in improving the ability of casual users to access IS&R systems as well as a plan for the utilization of the builtin evaluation mechanisms in more generalized research activities.

II. GENERAL OBJECTIVES

The following general objectives are intended to present an overview of the goals of the evaluation project.

- (1) Develop an evaluation methodology for the PC/MISI system.
- (2) Identify relevant performance indexes to be used in the evaluation.
- (3) Construct profiles of user behavior including characterization of PC/MISI system usage, host system usage, user error and user experience factors.
- (4) Construct measures and predictors of user success and user satisfaction with system usage.

III. GENERAL EVALUATION METHODOLOGY

This section provides a framework for the evaluation activities to be conducted. This methodology will be used in the development of experiments to evaluate specific aspects of the PC/MISI system as described in Chapter V.

- (1) Determine the evaluation objectives.
- (2) Determine the specific parameters to be monitored initially based upon the overall objectives.
- (3) Design and implement the monitoring facility into the system.
- (4) Design and implement the data analysis tools to be used in analyzing the monitored data (the statistical package is to be purchased).
- (5) Design and conduct the monitoring experiment to collect the data to be analyzed.
- (6) After the experiment has been completed, perform the data analysis (which will include data validation), making evaluations and drawing appropriate conclusions.
- (7) Identify monitor improvements and enhancements as implied by the results of the analysis (add new parameters that were

found to be necessary delete parameters that were found to be unnecessary, etc.).

- (8) Identify system improvements and enhancements as implied by the results of the analysis.

IV. USER CATEGORIZATION

Since relative merits of different interface levels of PC/MISI are to be measured, different user groups are needed in order to conduct an experiment which is meaningful. The user groups can be categorized as follows:

- (1) Users with no previous experience in computing, interactive terminals, or information system at all, i.e., totally naive casual users.
- (2) Users with some experience in computing and interactive terminals, but with no previous experience in utilizing information systems.
- (3) Users with experience in computing, interactive terminals and information systems.
- (4) Users with knowledge of specific subject areas.

V. SPECIFIC OBJECTIVES IN EVALUATING PC/MISI

The specific evaluation objectives correspond closely to the design objectives described in the original design plan [Hall, 84]. The purpose of the evaluation of these objectives is to determine if the system implementation has, indeed, achieved the objectives for which it was developed.

- (1) To evaluate the ease of access to mutiple information systems to both casual and experienced users.

(should allow user to simply choose the information system he/she is interested in (the system will perform communication and other related procedures)).

- (2) To evaluate the modularity of the system.

(i.e., the ability to expand to include more remote systems when necessary).

- (3) To evaluate the system documentation.

(how the user manuals and other documentation facilitate the ability of users to learn and utilize PC/MISI).

- (4) To evaluate the capabilities of the system to provide multilevel interaction to the remote systems.

(users with different levels of expertise should be able to interact with the host system according to their own level of expertise; smooth transition from one level to another as the user's level of expertise changes).

- (5) To evaluate if user orientation is helpful

(The user will be kept informed of his location within the system, time and date, and possibly other information. Evaluation activities will focus on determining whether or not this actually improves the user's ability to interact with the system and/or the user's impression of the system).

- (6) To evaluate the ability of the system to utilize users' knowledge.

(should provide users with "advice" on how to develop search strategies to best utilize their specific subject knowledge).

- (7) To evaluate the capabilities of downloading information.

(should provide simple and efficient procedure to store information from remote systems, edit it, print it, sort it, etc.).

- (8) To evaluate the batch processing capabilities.

(should enable users to store an entire search in a batch file and then have the entire sequence executed at the remote location with no need for further user input).

- (9) To evaluate the error handling capabilities.

(should provide interpretation of remote system messages and additional information and assistance where required).

- (10) To evaluate the display capabilities.

(should provide graphical and screen management capabilities with maximum portability).

- (11) To evaluate the relative merits of different interface levels.

(know how useful each interface is to the system users).

- (12) To evaluate the utilization of the remote systems.

(know the frequency of invocations of different remote systems, amount of information retrieved, time required to retrieve specific information, etc.).

- (13) To evaluate the response times of some of the operations incorporated.

(downloading efficiency, search time, etc.).

(14) To evaluate the relative usage of the operations incorporated.

(how frequently one operation is utilized relative to the others, etc.).

(15) To measure user success/satisfaction.

(does the user get what he wants? is he satisfied with what he has done?)

VI. EVALUATION OF DESIGN OBJECTIVES

6.1 Evaluation Parameters and Methodology

This section will discuss each of the specific design objectives and will describe the specific data which will be collected concerning each of these. The utilization of this information in evaluation will be described and the methods of collecting the information and varying different conditions to obtain comparative data will be outlined.

Objective 1: ease of access to multiple systems

Corresponding data measures: users ratings, user comments, error rates

One easy way ease of access can be evaluated (in a subjective manner) is to have a post-usage questionnaire which will include the users' ratings and comments. Then with this information, statistical analysis can be performed to satisfy the objective. Another way is to carry out an experiment which consists of two sessions. In the 1st session, users are required to access the information systems they are interested in based on the standard procedure (i.e., dialing through modem, entering userid and password again and again,

etc.). In the 2nd session, the users can merely choose the information systems they want and PC/MISI will do the rest! Then the error rates can be compared after the experiment. (Designer's and implementers' subjective evaluations (what the actual process to access a system is may be considered too.)

Objective 2: modularity of the system

Corresponding data measures: interface administrator's ratings/comments, error rates, processing time

Since only the interface administrator has the primary responsibility for the addition and maintenance of the host system files, the only way to evaluate the modularity is according to his ratings/comments, the processing time taken and the number of errors made during the expansion and/or addition process.

Objective 3: system documentation

Corresponding data measures: users ratings, users comments

Again, post-usage questionnaires can be provided to obtain the necessary information to do the statistical analysis about the system manuals and other related documentation.

Also experiments can be conducted in which one group of users is required to read and learn the manual before accessing the system and their performance in subsequent system usage is compared to users who access the system without access to explanatory material.

Objective 4: multilevel interaction capabilities

Corresponding data measures: users ratings, users comments, number of requests in using the multilevel selection

Users' ratings and comments will describe how they feel about this multilevel selection and also their feelings toward the smoothness of the transition from one level to another. Statistical analysis can be performed on the above information and also on the amount of time spent at each level by users as they become more familiar with the system and the error rates at each level.

Objective 5: user orientation

Corresponding data measure: users ratings, users comments

Only the system users can tell if the system is user oriented. Thus, we can collect users ratings and comments and

can vary the amount and type of information provided to determine if the user's perceptions of the system and ability to interact with the system are affected.

Objective 6: utilize users' knowledge

Corresponding data measures: users ratings, users comments, search success

Experiments can be performed with two different groups of users. The 1st group consists of users with knowledge of a specific subject area (e.g., geology) and the 2nd group consists of users with experience in retrieval of information from IS&R systems. Then, these two groups of users can be asked to perform certain usage assignments within a specified time. After the experiment, the success of both these groups can be evaluated and an analysis made to determine the variance in the information retrieved. (Notice that in order to measure the success, we have to assume that the objective of these users is to get the correct answer for the usage assignment and that there is no reformulation of the objective during the process).

Objective 7: downloading information capabilities

Corresponding data measures: users ratings, users comments (see also Objectives 13 and 14 for the downloading efficiency, frequency counts, etc.)

Users' ratings and comments will be used to determine how easy it is to download an accession, edit it, print it, sort it, merge it, etc. Note that the rate of downloading accessions is a very important factor. This will be treated in Objective 13.

Objective 8: batch processing capabilities

Corresponding data measures: users ratings, users comments, error counts, processing time, (see also Objectives 13 and 14)

Users ratings and comments will reflect their feelings about performing searches in a batch file manner. Processing time will enable us to compare the relative efficiency and convenience of using a batch file and using a sequence of single commands and interacting with the host system each time. The error counts will enable us to analyze the error frequency using both of the above methods. Here, the error handling capabilities may be tested also. PC/MISI will correct the syntax of a command before entering it into the batch file. When performing the correction, the guidance

provided may be more user oriented than what the remote system would have given.

Objective 9: error handling capabilities

Corresponding data measures: users ratings, users comments, number of retries (if information is collected in a time-ordered sequence), designer/implementer's comments, error context, corresponding frequencies

Users' ratings and comments will allow us to know how the error handling capabilities are helping the users to correct errors committed (from the user point of view). The number of retries (if available) will give us the number of retries necessary to correct an error (from a statistical point of view). The designer/implementer's comments, on the other hand, will describe exactly what kind of error handling capabilities are incorporated into the system. Error context will enable us to know which specific types of errors are being made (misspelling of keywords, invalid system command, etc), complexity of the attempted operation and the types of operations which are most error prone, thus providing appropriate implications for user language re-design, for documentation improvement, and so on).

Objective 10: display capabilities

Corresponding data measures: users ratings, users comments, implementer's comments

Experiments will be designed with different types of displays: color highlighted error messages vs. black and white, display of information at different locations on the screen, windows, etc. Users' ratings and comments will be used in conjunction with more objective measurements to determine the usefulness of the window systems, light pen and mouse capabilities incorporated into PC/MISI. Measurements can be made of retrieval efficiency using different combinations of these capabilities.

Objective 11: relative merits of different interface levels

Corresponding data measures: session time, number of errors, users ratings, users comments, number of accessions retrieved, quality of accessions retrieved

Here, experiments can be carried out on a group of users using different interface levels. Then, the session time (the time they start using the system through the time of completion for a fixed task) may be compared. Also, with this fixed task performed via different levels, the number of

errors committed can be collected during the session. Retrieval success based on number of accessions retrieved and quality of accessions retrieved versus time may also be analyzed to satisfy the objective. The users feelings of the relative merits may be found in the ratings and comments. The combination of information retrieved by these measures should provide a means of determining which levels are best suited to which types of users.

Objective 12: utilization of remote systems

Corresponding data measures: frequencies of the invocations of remote systems, session times using the remote systems

These frequency counts and session time counts will allow us to determine the amount of utilization of different remote systems. Appropriate actions such as the removal of some very under-utilized remote system can then be taken. Experiments can be conducted in which users are provided with some general information concerning the information available in different systems and then be allowed to choose the system from which to extract specified information.

Objective 13: response times

Corresponding data measures: response time of each operation

The response time will allow us to evaluate the efficiency of each operation. User ratings correlated with response times can be used to evaluate the impact of differing response times on the user's perception of the system.

Objective 14: operation counts

Corresponding data measures: frequency count of each operation

Less frequently referenced operation may be moved to secondary menu, put into different command table, etc.

Objective 15: user success/satisfaction

Corresponding data measures: users ratings, users comments, search success

In order to measure user success/satisfaction, we need to know the objectives of the user. Thus, the easiest way to measure is to gather information from users' ratings and comments. Also, we can measure the user success as described in Objective 6; that is, to assign the user a certain task (e.g., usage assignment) and examine the results after he has

completed the task (assuming that the objective is to get the task completed). The user satisfaction is not an easy measure because, even though the user may not complete his task, he might still be very satisfied (he learned something!) with what he has done. Therefore, no obvious single measure for user satisfaction is available, but attempts will be made to extrapolate user satisfaction from success and the validity of these measures can be determined by comparison with subjective information obtained from user ratings and comments.

6.2 Miscellaneous Parameters

This section lists some other measurements which will be collected and will be used in a variety of ways to correlate different components of the evaluation data base.

(1) User's name and affiliation

(know who the users are).

(2) Date of interactive session

(how often the system is used and how the use is distributed).

(3) User ratings of the interactive session

(4) User comments on the interactive session

(5) Average session cost

(on-line time and PC/MISI local time).

(6) Output reports generation

6.3 Evaluating Impact of User Experience

The following questions will be addressed utilizing previously

mentioned data. These questions represent some very important research areas concerned with user/system interaction.

- (1) To what extent does familiarity with the PC/MISI system increase chances of success?
- (2) To what extent does familiarity with the host system increase chances of success?
- (3) To what extent does familiarity with interactive terminals increase chances of success?
- (4) To what extent does familiarity with computing increase chances of success?

This document is intended to provide a foundation for the development of future evaluation activities utilizing the PC/MISI system. Therefore a number of possible areas of evaluation activity have been described in general terms and many of the evaluation mechanisms described are subjective in nature. More detailed evaluation plans concerning specific areas of evaluation will be developed in the future utilizing much more specific and objective measurements.

REFERENCES

- [Dominick, 75]. W. D. Dominick, "System Performance Evaluation of Interactive Retrieval", in Personalized Data Base System, B. Mittman, L. Borman (editors), John Wiley & Sons Inc., 1975, Chapter 12, pp. 216-227.
- [Dominick, et.al, 78]. W. D. Dominick, J. Urban, "Application of A Generalized Evaluation Methodology for Analyzing User Interaction with the MADAM System at the University of Southwestern Louisiana", Technical Report CMPS-78-6-3, Computer Science Department, University of Southwestern Louisiana, Lafayette, Louisiana, 1978.
- [Hall, 84]. P. P. Hall, "Design Criteria for a PC-Based Common Interface to Multiple Remote Information Systems", USL/DEMS NASA/PC R&D Working Paper Series Report Number DEMS.NASA/PC R&D-9, August 13, 1984, 21p.
- [Hall, 85]. P. P. Hall, "The Design of PC/MISI, A PC-Based Common User Interface to Remote Information Storage and Retrieval Systems", Master Thesis, University of Southwestern Louisiana, Lafayette, LA, 1985, 83p.

[Lancaster, 79]. F. W. Lancaster, "Criteria by Which Information Services May be Evaluated", in Information Retrieval Systems: Characteristics, Testing and Evaluation, Chapter 8, John Wiley & Sons Inc., New York, 1979.

1. Report No. <i>11-82</i>	2. Government Accession No. <i>183587</i> 44597	3. Recipient's Catalog No.	
4. Title and Subtitle USL/NGT-19-010-900: AN OVERVIEW OF THE EVALUATION PLAN FOR PC/MISI - PC-BASED MULTIPLE INFORMATION SYSTEM INTERFACE		5. Report Date April 24, 1985 <i>DATE OVERRIDE</i>	6. Performing Organization Code
7. Author(s) BEE LEE LIM AND PHILIP P. HALL		8. Performing Organization Report No.	
9. Performing Organization Name and Address University of Southwestern Louisiana The Center for Advanced Computer Studies P.O. Box 44330 Lafayette, LA 70504-4330		10. Work Unit No.	
12. Sponsoring Agency Name and Address		11. Contract or Grant No. NGT-19-010-900	
		13. Type of Report and Period Covered FINAL; 07/01/85 - 12/31/87	
		14. Sponsoring Agency Code	
15. Supplementary Notes			
16. Abstract This Working Paper Series entry represents an initial evaluation plan for the personal computer multiple information system interface (PC/MISI) project. The document is intended to be used as a blueprint for the evaluation of this system and each objective of the design project is discussed along with the evaluation parameters and methodology to be used in the evaluation of the implementation's achievement of those objectives. The potential of the system for research activities related to more general aspects of information retrieval is also discussed. This report represents one of the 72 attachment reports to the University of Southwestern Louisiana's Final Report on NASA Grant NGT-19-010-900. Accordingly, appropriate care should be taken in using this report out of the context of the full Final Report.			
17. Key Words (Suggested by Author(s)) Microcomputer-Based Environment, Common User Interface, Multiple Remote Information Systems, PC/MISI Evaluation Plan, PC-Based Research and Development		18. Distribution Statement	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 27	22. Price*