CAR-TR-768 March 1995
CS-TR-3463

Assessing Users' Subjective Satisfaction with the Information System for Youth Services (ISYS)

Laura Slaughter†, Kent L. Norman†, Ben Shneiderman*

Human-Computer Interaction Laboratory
†Department of Psychology
*Department of Computer Science
University of Maryland, College Park, MD 20742-3255
ben@cs.umd.edu

Abstract

In this investigation, the Questionnaire for User Interaction Satisfaction (QUIS 5.5), a tool for assessing users' subjective satisfaction with specific aspects of the human/computer interface was used to assess the strengths and weaknesses of the Information System for Youth Services (ISYS). ISYS is used by over 600 employees of the Maryland State Department of Juvenile Services (DJS) as a tracking device for juvenile offenders. Ratings and comments were collected from 254 DJS employees who use ISYS. The overall mean rating across all questions was 5.1 on a one to nine scale. The ten highest and lowest rated questions were identified. The QUIS allowed us to isolate subgroups which were compared with mean ratings from four measures of specific interface factors. The comments obtained from users provided suggestions, complaints and endorsements of the system.

The Information System for Youth Services (ISYS) is an on-line real-time processing system programmed in IDMS/R, a relational database program, and runs on an IBM mainframe computer. It is used by the State of Maryland's Department of Juvenile Services (DJS) as a tracking system for juvenile offenders. This system currently includes information for approximately 50,000 juveniles. The ISYS system is used at facilities throughout the state to enter and access the data needed by the Department's employees.

The Questionnaire for User Interaction Satisfaction (QUIS) is a measurement tool designed to assess a computer user's subjective satisfaction with the human-computer interface and has proven reliability and validity across many types of interfaces (Chin, Diehl, & Norman, 1988). It was developed at the Human-Computer Interaction Laboratory (HCIL), University of Maryland, College Park and is currently licensed to over 76 sites both in academia and industry. The QUIS contains a demographic questionnaire, a measure of overall system satisfaction, and four specific interface factors (screen factors, terminology and system information, learning factors, and system capabilities). The QUIS was modified for use with this project. Several questions were added and questions not pertaining to ISYS were omitted.

The Department of Juvenile Services contracted the Human Computer Interaction Laboratory (HCIL) to evaluate the ISYS user interface. Although ISYS has automated many of the DJS employees' tasks, it has many shortcomings which prevent it from becoming a useful tool for the employee. Because of these difficulties in using the system, employees often do not enter data about youths in a timely fashion. This makes the data in the system unreliable for those who must access and use this data for their particular job function. In our evaluation of this interface, we used the Questionnaire for User Interaction Satisfaction (QUIS) to obtain user satisfaction ratings for ISYS. We examined these ratings in order to determine what areas of the system needed improvement, which aspects were satisfactory and to suggest possible reasons why these aspects were rated as higher or lower.

The Department of Juvenile Services consists of three major employee divisions, Administrative, Field Services and Residential Services. The employees within these divisions use the system for a variety of job tasks. For example, employees in Field Services, which consists mainly of Juvenile Counselors, enter data about the youths. While some of the information Field Services employees place in the system may be useful to them, the majority of the data is accessed by others such as Administrative employees who must gather information for statistical or financial purposes. Because of these variations in system use, we predicted differences in satisfaction measures between the employee divisions. Questions were added to the demographic portion of the QUIS in order to identify these subgroups.

A separate section for the QUIS was created to answer questions about ISYS training. The additional ISYS training section was requested by the Department of Juvenile Services to look at the amount and quality of the training given to employees. For the most part, their goal was to document the inadequacy of training.

Many of the results from our analyses of the QUIS ratings and comments made by users are presented in this report. It was our goal in this project to make recommendations to the Department of Juvenile Services for improvements to ISYS. For the most part, these recommendations were made after many hours of observation and informal interviews with employees. The QUIS results from this investigation were partly used to confirm the findings made from the observations and give a scientific backing to the recommendations.

This report will not center on a comparison of observational data and QUIS data. The goal in this paper is to document the use of the QUIS in the evaluation of a system. We used ratings from the questionnaire to identify the strengths and weaknesses of the system. User's comments were used to acknowledge specific problems of the system. In our discussion, we will show how the results from the QUIS allowed us to give a more in depth look at the system and infer why specific aspects of the system rated as they did.

METHODS

Respondents

Department of Juvenile Services employees who completed the QUIS were selected from a list of 669 ISYS accounts.

Participants were chosen so that stratified by division and function, there would be a 90% response rate and a 90% confidence that at least 30 respondents would be sampled for each employee function. There were a total of 309 QUIS forms administered to DJS employees. Of this number 291 returned the QUIS. The total number of employees who completed the QUIS and stated that they use ISYS was 254. Only ratings from actual users (254) were used in the analyses of the QUIS.

We used frequency data from the demographic and training sections to characterize the population of respondents. Question 1.1 asks "Do you use a personal computer or a terminal?". Thirty said that they use a personal computer, 111 stated that they use a terminal, 78 said they used both, and 35 did not respond. Question 1.4 asks "What do you use ISYS for?, Do you enter or lookup data?". Thirty-eight said that they enter data, 58 said that they lookup data, 137 said that they did both, 2 said neither, and 19 did not respond. Question 8.3 asks "How many days of training did you receive?" It asks for days of initial and on-going training. Table 1 shows the frequency data for this question.

Days of Training	Count (Initial)	Count (On-Going)
0-2	126	88
3-4	41	1
5-6	4	0
7-8	0	0
9-10	2	1
11-12	0	0
13-14	1	0
15-16	0	0
17-18	0	1
18-19	1	0

Table 1. Frequency data for question 8.3.

Instrumentation

The QUIS

The Questionnaire for Interaction Satisfaction (QUIS) is arranged in a hierarchical format and contains: (1) a demographic questionnaire, (2) six scales that measure overall reaction ratings of the system, and (3) four measures of specific interface factors: screen factors, terminology and system feedback, learning factors and system capabilities. Each of the four specific interface factors has a main component question followed by related subcomponent questions. Each item is rated on a scale from 1 to 9 with positive adjectives anchoring the right end and negative on the left. In addition, "not applicable" is listed as a choice. Additional space which allows the rater to make comments is also included within the questionnaire. The comment space is headed by a statement that prompts the rater to comment on each of the specific interface factors. We made several changes to the standard format of the QUIS for this project. We have listed them in the following paragraphs and we also state the motive behind each change.

Modification of QUIS for ISYS

In Part 1, the demographic questionnaire, the additions of classification, division, employee location, and employee function were added to the list of identifying information. This information was used to determine the user's functions and job-related tasks to be completed when using the ISYS system. It was also used to identify subgroups. For question 1.1 in this section, the choices of personal computer or terminal were typed directly on the questionnaire. DJS employees completing the QUIS circled PC, Terminal or both. Also, the name of software (ISYS) was written in the answer line for question 1.2. These additions were made so that the questions would be easier for the employees to answer and to make it clear to them exactly what they would be rating. A new question, 1.5 was added to Part 1. The question was " What do you use ISYS for?" We included this at DJS's request for the purpose of learning the employee's function and use of ISYS in greater detail. Also, Question 1.6 was added, "Do you use the menu or command line to access the screen you need?" This was used to determine the user's proficiency at using ISYS.

There was one addition to Part 3 "Overall User Reactions". Question 3.7 was added to determine the amount of helpfulness or hindrance the ISYS system provides the user when completing job tasks. This new question included the adjectives "hindrance" and "helpful" as the anchors to the rating scale.

In part 4 "Screens", Question 4.2.1., "use of reverse video" was changed to "use of bolding" because ISYS uses bolding instead of reverse video. On question 4.3, "were the screen layouts helpful" was changed to "were the screen layouts helpful in completing tasks?" to make the question more specific. A new question, 4.3.1 was added "type of information on screen" with "irrelevant" and "relevant" as the anchors to the rating scale. This question was added to ascertain if relevant information was displayed on the screen. Also, the question 4.3.4 was added "format of information and task to be completed" with adjectives "hindrance" and "helpful" as anchors to the rating scale. This question was created to determine if the format of screen identifies and communicates tasks.

In part 5 "Terminology and ISYS System Administration" there were several changes made. The ISYS system has both terms and codes. Question 5.1 was changed to "use of terms and codes throughout system". Questions 5.1.3 "task codes" and 5.1.4 "computer codes" were also added. There are two types of error messages given in the ISYS system. Question 5.6 was changed to "system and data entry error messages". Questions 5.6.3 "data entry error messages clarify the problem" and 5.6.4 " phrasing of data entry error messages" were also included. To include the codes used in the ISYS system, question 6.3.2 "remembering specific rules about entering codes" was added.

Finally, Part 8 "training" was added to the questionnaire at DJS request to include questions about training.

Design and Procedures

The distribution and collection of the questionnaires was monitored by the Department of Juvenile Services. Two introductory letters were included at the beginning of the questionnaire to explain the purpose of and give instructions for the QUIS. One was an introductory letter from our research team at the University of Maryland and the second was a letter written by DJS administrators of the QUIS.

The questionnaires were given to the employees by their supervisors. The supervisors also received a letter describing how their employees should complete the questionnaire. Employees completed the questionnaire on their own and were required to return the QUIS to the DJS administrators within one week.

RESULTS

Profile of Ratings for ISYS

The first approach to analysis of QUIS results is to calculate the mean and standard deviation for each item. To provide an overview of the ratings for ISYS, the profile of mean ratings for each main component question and their 95% confidence interval is shown in Figure 1.

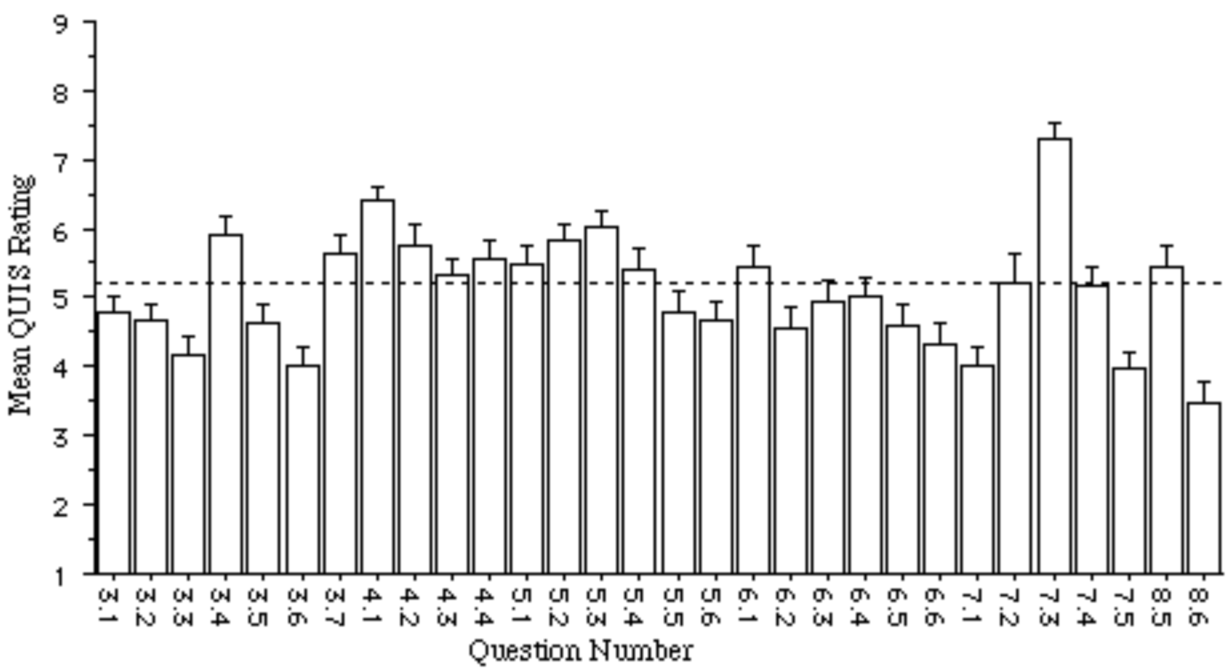

Figure 1. Mean QUIS ratings for each main component question. The dashed line indicates the position of the overall mean rating (5.1) of ISYS.

Highest Rated Questions		**Lowest Rated Questions**	
Question	t-test results p<.0001	Question	t-test results p<.0001
5.5.1	t(232) = 8.180	6.6.1	t(164) = -8.917
5.1.2	t(217) = 8.222	7.1	t(237) = -9.267
7.4.1	t(222) = 9.040	5.6.1	t(230) = -10.026
7.5.2	t(200) = 9.082	6.5.3	t(196) = -10.070
5.3	t(221) = 9.687	3.6	t(220) = -10.290
4.3.1	t(236) = 12.461	7.2.2	t(234) = -10.322
4.1.1	t(233) = 14.420	7.5.1	t(229) = -11.531
4.1	t(237) = 14.439	7.5	t(229) = -12.288
4.1.2	t(223) = 16.155	6.4.1	t(236) = -14.627
7.3	t(228) = 17.482	7.2.3	t(222) = -14.911

Table 2. Highest and lowest rated items from the QUIS for ISYS.

Ten Highest and Ten Lowest Rated Questions

The second approach to the analysis of items on the QUIS is to determine the ten best and ten worst features of the system. A within-subjects approach was used to identify which items were perceived as highest and lowest relative to each subject's average rating. Deviation scores from the subject's overall mean for each of the subject's ratings were calculated. For each item of the QUIS, a one sample t-test was performed on the deviation scores. The p-values were set at p<.0001 by a Bonferonni adjustment to control for an inflated alpha error rate. These results are summarized in Table 2.

Differences in Ratings Across Subgroups

The third approach to analysis of the QUIS is to compare subgroups of respondents on mean ratings of each section of the QUIS. A one-way analysis of variance was performed on the data for employee division and mean ratings across questions for each section of the QUIS. The three divisions of employees gave significantly different mean ratings for two sections of the questionnaire. In both cases where significant differences occurred between the employee divisions, the Residential Services employees tended to rate questions higher.

There were significant differences between divisions on mean overall rating across all questions in Section 4 "Screens", F(2,240) = 5.128, p<.01. Significant differences were found between Administrative (M = 5.6) and Residential Services (M = 6.4) (Fisher's PLSD, p< .02), and between Field Services (M = 5.7) and Residential Services (M = 6.4) (Fisher's PLSD, p< .002). In this section of the QUIS, Residential Services tended to give higher ratings to ISYS than both Administrative and Field Services.

There were significant differences between divisions on mean overall rating for Section 5 "Terminology and System Information", F(2,238) = 3.706, p<.03. Significant differences were found between Field Services (M = 5.1) and Residential Services (M = 5.8) (Fisher's PLSD, p< .01). In this section, Residential Services tended to rate ISYS higher than Field Services employees.

Question 1.3 and Mean Rating for Each Section

Question 1.3 asks "On the average, how much time do you spend per week on the ISYS system?" and has four choices: 1) less than one hour, 2) one to less than four hours, 3) four to less than ten hours, and 4) over ten hours. A one-way analysis of variance was performed on the data for Question 1.3 and mean ratings across questions for each section of the QUIS. Only significant results are reported. In each case, employees with less computer hours per week tended to rate ISYS lower than their more experienced co-workers.

There were significant differences for Question 1.3 on the mean overall rating for Section 3 "Overall Reactions" (F(3, 237) = 3.137, p<.03). Significant differences were found between employees who use ISYS less than one hour per week (M = 4.3) and those who use it four to less than ten hours (M = 5.2) (Fisher's PLSD, p< .004). Significant differences were also found between employees who use it less than one hour per week (M = 4.3) and those who use it over ten hours per week (M = 5.1) (Fisher's PLSD, p< .04). Employees who use the system less than one hour per week tended to rate ISYS lower than employees who use it four to ten hours and those who use it over ten hours per week.

There were significant differences for Question 1.3 on mean overall ratings for Section 5 "Terminology and System Information", ($F(3,236) = 3.914$, $p<.01$). Significant differences were found between employees who use ISYS less than one hour per week ($M = 5.0$) and those who use it four to less than ten hours per week ($M = 5.2$) (Fisher's PLSD, $p<.001$); between employees who use ISYS one to less than four hours ($M = 5.0$) and those who use it four to less than 10 hours ($M = 5.2$) (Fisher's PLSD, $p< .003$); and between those who use ISYS four to less than ten hours ($M = 5.2$) and those who use it over ten hours per week ($M = 6.0$) (Fisher's PLSD, $p<.01$). In this section of the QUIS, employees who use the system less than 10 hours per week tended to rate ISYS lower than the employees who use it over ten hours per week.

There were significant differences for Question 1.3 on mean overall rating for Section 6 "Learning", ($F(3,238) = 6.936$, $p<.001$). Significant differences were found between employees who use ISYS less than one hour per week ($M = 3.9$) and those who use it one to less than four hours per week ($M = 4.7$) (Fisher's PLSD, $p<.005$); between those who use the system less than one hour per week ($M = 3.9$) and those who use it four to less than ten hours ($M = 5.0$) (Fisher's PLSD, $p<.0002$); and between those who use the system less than one hour per week ($M = 3.9$) and those who use it over ten hours per week ($M = 5.3$) (Fisher's PLSD, $p<.0001$). In Section 6 "Learning", employees who used the system less than one hour per week tended to rate ISYS lower than employees who used ISYS from one to over ten hours per week.

Comments

The fourth approach to the analysis of the QUIS is the inspection of comments from users. This is often the most interesting and useful diagnostic analysis. The comments listed below were chosen because they represent the common complaints, suggestions and viewpoints of many other users. A brief discussion follows each comment.

Error Messages

"*Although the error codes appear on the last screen just before sign off, these codes do not explain or indicate what were the errors I made. I could make the same error over and over, and not even know what the mistake is because, although the information is displayed, it does not tell me anything, like looking at the letters of a foreign language.*" Juvenile Counselor

Complaints about the error messages were the most frequently reported, stating that the terminology used in the error messages is not easily understood. A large number of users indicated that the manuals were not very much help. One comment on error messages simply states "Confusing, Very confusing, Real confusing, Extremely confusing"- Program Director

Learning

"*I really enjoyed learning the system. Each chance that I had, I would get on ISYS, and play around with it.*" Juvenile Counselor

"*you would have to be a masochist to want to learn this system*" - Program Director

Comments made about learning the system were varied. Almost all of the users said they had more hands-on learning with little or no training. People with computer experience tend not to find learning the system as difficult as most of the new users. A Juvenile Counselor concludes "Since I was pretty much familiar with computers, I really did not have that much trouble learning and operating ISYS. But, for those who have had no computer experience, ISYS is difficult and not user friendly."

Help Messages

"*We should have a help option in the system*"-Juvenile Counselor

The lack of instruction manuals, or the ability of the available instruction manuals to clarify the problem is a common complaint. Updating manuals regularly and placing help options in the system were common suggestions. Some users asked for on-line tutorials.

Screen Layout

"*Too much text on screen, organized too illogically and not related to task*"- Administrator

Statements included suggestions for the addition and deletion of information displayed on the screen. For instance, a Program Specialist wrote that "Info on social history screen often reflect data on parents with whom youth does not reside." This person asked for additional lines on the screen to place the data for people with whom a child actually lives (e.g. grandparents). Users also made a point of stating that they often cannot access accurate data from ISYS either because it is not entered or the screens do not allow additional data which would be useful.

DISCUSSION

One use of the data from this study might be to compare ISYS to other systems that have been evaluated by the QUIS. If we collected reports from many QUIS evaluations, we could also create standards of what might be judged as a "good" system. One study, (Harper & Norman, 1993) rated several software packages that were used in an experimental computer classroom. Comparing ISYS with those systems, the ISYS rates below average. However, we will not continue to compare ISYS with other systems rated by QUIS for two reasons 1) ISYS is a completely different system than other systems we have data available for and we do not want to compare it with systems that are not similar and 2) although QUIS is used extensively in research at many sites in academia and industry, reports are not always published so we are limited in the number of systems obtainable for comparison. Instead of the system comparison approach, we will examine the results of our evaluation in order to understand how this system meets the demands of its users.

The overall mean rating of all sections of the QUIS was 5.1, on a 9-point scale. If we look back to the profile of results shown as Figure 1, it is obvious that the system did not rate extremely well on any question. In fact, the highest rated question , 7.3 "ISYS system tends to be... noisy-quiet", received only a mean overall rating of seven on the nine-point scale. If we were to define seven on a nine-point scale as an acceptable overall rating of a "good" system, the overall rating made by the users' of this system shows us that this system is in need of much improvement.

Five Highest Rated Questions

One of the most important examinations of our results was a close analysis of the ten highest and lowest rated questions from the QUIS. We have included a discussion for the results of the five highest rated questions as an example.

The five highest rated questions deal with "factual" information that makes it easy to understand why they rated the highest of all questions. Question 4.3.1 "Type of information on screen- irrelevant vs. relevant" had a high rating because the information on the screen relates to the employee's work. However, not all of the information listed on the screens is important to each employee. Many of them might feel that much of the data on the screens are repetitive and extraneous. This may be the reason that, although it is one of the highest rated questions, it still had an overall mean of only 6.2. Question 4.1.1 "Image of Characters- fuzzy vs. sharp", Question 4.1" Characters on the computer screen- hard to read vs. easy to read", and Question 4.1.2 " Character shapes (fonts)- barely legible vs. very legible" were rated high because the characters on the computer screen in this system are legible. The highest rated question, Question 7.3 "ISYS system tends to be-noisy vs. quiet" is not a surprise because ISYS is a quiet system. Although the questions with the highest mean ratings represent the best areas of the system, these aspects could still be improved a great deal. With the exception of question 7.3, each of the highest rated questions had a mean rating of only six on a nine-point scale.

Using the QUIS in the evaluation of ISYS helped confirm what problems and strengths existed with this system. The QUIS demographic section helped us identify subgroups to compare with mean ratings of the four measures of specific interface factors: screen factors, terminology and system feedback, learning factors and system capabilities. We were able to collect the data from the seven scales that measure overall reaction ratings of the system, and the four measures of specific interface factors in order to complete a profile analysis. From this, we were able to produce a ten best and ten worst questions analysis. The comments collected from this questionnaire provided a list of suggestions, complaints and endorsements for the system. These results from our investigation are not only helpful for the re-engineering of the system (Vanniamparampil, A., et. al., 1995), they will also provide other researchers with a detailed evaluation of a system using a satisfaction questionnaire.

ACKNOWLEDGEMENTS

Our sincere thanks to HCIL members who were involved directly or indirectly in the projects listed here, and contributed to their successful completion. We would like to thank Chris Cassett for his help with customizing the QUIS, and the distribution and collection of the questionniare from users. We also thank the users of ISYS for their thoughtful completion of the QUIS and the comments they provided. The preparation of this document was supported in part by the Maryland Department of Juvenile Services.

REFERENCES

Chin, J.P., Diehl, V.A., & Norman, K.L.(1988). Development of an instrument measuring user satisfaction of the human-computer interface. In CHI '88 Conference Proceedings: Human Factors in Computing Systems, (pp. 213-218), New York: Association for Computing Machinery.

Harper, B.D. & Norman, K.L.(1993). Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. In Proceedings of Mid Atlantic Human Factors Conference, (pp.224-228), February 23-26, 1993.

Vanniamparampil, A., Shneiderman, B., Plaisant, C. & Rose, A. (1995). User interface re-engineering: A diagnostic approach. (Technical Report) University of Maryland, Department of Computer Science.