

(This paper has been published in: R. Baud, M. Fieschi, P. Le Beux, P. Ruch (eds): The New Navigators: from Professionals to Patients. Proceedings of Medical Informatics Europe (MIE 2003), 4 – 7 May 2003, St. Malo, France. Studies in Health Technology and Informatics. Volume 95. Amsterdam, Berlin, Oxford: IOS Press. ISBN 1 58603 347 6. Pages 643-648.

Evaluation of user acceptance of information systems in health care – the value of questionnaires -

Elske Ammenwerth^a, Frieda Kaiser^a, Immanuel Wilhelmy^b, Stefan Höfer^c

^a*University for Health Informatics and Technology Tyrol (UMIT), 6020 Innsbruck, Austria*

^b*TILAK IT Department, 6020 Innsbruck, Austria*

^c*University of Innsbruck, 6020 Innsbruck, Austria*

Abstract

The use of modern information technology (IT) offers tremendous opportunities such as reducing clinical errors and supporting health care professionals in providing care. Evaluation of user satisfaction is often seen as a surrogate for the success of an information systems.

We will present the evaluation of a report writing system at the Innsbruck University Medical Center based on a standardized, validated psychometric questionnaire. The results show high reliability and validity of the questionnaire. They also show some interesting differences in user satisfaction between departments, due to differences in working processes and preconditions.

Psychometric questionnaires can be seen as a reliable and valid method to measure certain psychological constructs. Their development requires, however, methodological rigour and sufficient time. Psychometric questionnaires allow only a limited interaction between researcher and user, their results may be very dependant on the time of measurement, and their interpretation often needs external knowledge. Those limitations have to be taken into account when preparing evaluation studies.

Keywords:

Medical informatics, evaluation studies, user satisfaction, questionnaires, psychometrics

1. Introduction

Information technology (IT) is emerging more and more in health care. It is evident that the use of modern information technology offers tremendous opportunities, but there are also hazards associated with information technology in health care: modern information systems are costly, their failures may cause negative effects on patients and staff, and possibly, when insufficiently designed, they may result in spending more time with the computer than with the patient. Therefore, a rigorous evaluation of IT in health care is recommended and of great importance for decision makers and users of future information systems [1].

Evaluation can be defined as the decisive assessment of defined objects, based on a set of criteria, in order to solve a given problem [2].

There have been many attempts to describe the factors which can illuminate whether an information system can be seen as a success or not. For example, DeLone [3] describes six major dimensions of success of management information systems: system quality, information quality, use, *user satisfaction*, individual impact, and organizational impact. User satisfaction in his model is closely intertwined with use, and both are the precondition for the effects of a system. Other authors also stress the importance of user satisfaction resp. user acceptance evaluation. For example, Goodhue [4] states that user evaluation is often used as a surrogate for the success of a management information system, because objective criteria for system success are difficult to be found.

A lot of research has been done to built theories of user acceptance and user satisfaction (we will use both terms in this paper). A well known model is the technology acceptance model (TAM) by Davis [5], correlating perceived usefulness, perceived ease of use, attitude towards using, and actual system use. Another model by Ohmann [6] distinguishes between system-independent and system-dependent factors for user satisfaction, the latter e.g. characterized by user satisfaction with the content, the interface and the organization. The Task-Technology-Fit model by Goodhue [4] indicates that user evaluations depend on the fit between user, technology, and task.

In general, user acceptance seems to reflect whether a system adequately fits the characteristics of the users (e.g. computer knowledge) and the characteristics of the task (e.g. report writing) which is to be performed. Thus, user acceptance can be seen as an adequate indicator whether an information system really supports users in their clinical working processes. When this support is the aim of an information system, then user acceptance can even be seen as an adequate indicator for the overall system's success.

The evaluation of user acceptance is mostly conducted using standardized psychometric questionnaires in order to quantitatively measure the construct "user acceptance". Psychometric analyses deal with the measurement of human characteristics. To assure that these instruments fulfill required quality standards, they have to be rigorously developed and validated, comprising an iterative proceeding of design, pilot application, validation, and re-design [7], [8]. The validation of psychometric questionnaires is of great importance in order to show whether they really measure what they are intended to measure (validity), and whether they do this in an objective and reliable way. Besides validity, reliability and objectivity, feasibility, economic aspects and comparability are other requirements for questionnaires [9].

With this in mind, an evaluation was conducted on user acceptance of an electronic report writing system at the Innsbruck University Medical Center. This evaluation was mainly based on questionnaires. Based on our results, we will discuss the value and shortcomings of questionnaires to be used in evaluation studies of health information systems.

2. Evaluation of user acceptance of an electronic report writing system at the Innsbruck University Medical Center

2.1 Goal of the Study

The goal of the study was to evaluate the positive and negative effects of an electronic report writing system at the Innsbruck University Medical Center (based on Cerner HNA Millennium®), seen from the view of users. The study was conducted in those three departments which worked with the system at the time of the study: neurology, internal medicine and surgical transplantation. A sub-goal was to assure validity and reliability of the used questionnaires. The functionality of this system comprised creating, correcting/editing and retrieving discharge reports as well as short medical reports. Three main user groups were identified for this study: junior physicians, senior physicians, and clerical assistants.

2.2 Study Design

The study was a prospective, descriptive, quantitative study. Through a cross-sectional study design, a representative sample from each of the different user groups and from each of the three departments were questioned via a written validated questionnaire. We were able to reuse a questionnaire presented by Boy et al [10] which had been used in at least two questionnaire studies. Only slight modifications have to be done (e.g. exchanging the name of the system). The application of the questionnaires in [10] had shown a satisfactory reliability coefficient Cronbach Alpha of 0.84. The survey was carried out in February 2002 and the sample size was obtained from the users known in the system. Our questionnaire consisted of 17 main psychometric questions on user acceptance (e.g. “The system is easy to use”; “I can profit from the system in my work”), supplemented by additional question on how often they use different functions of the system, on how the overall user acceptance is, and on demographic data. Finally, two free-text questions on positive and negative aspects of the systems were added to allow for unexpected observations.

2.3 Execution of the study

After a small-scale pre-test with three users to test the clarity of the already validated questionnaire, it was distributed to 90 users. 59 of them finally responded to the 2-page questionnaire (14 junior physicians, 26 senior physicians, 19 clerical assistants), making up a return rate of 66%.

2.4 Study Results

The outcome of the study showed medium to high level of user satisfaction with the computer-based report writing system. Figure 1 displays the mean values to general satisfaction (by department, and by occupational group), based on the main psychometric part of the questionnaires. The complete results can be found in [11].

The results show some interesting differences between the departments. While in the neurology department, the clerical assistants are much more satisfied with the system than

both physician user groups, the situation is the opposite in the other two departments. After analysis of the answers (including the free text answers), some explanations could be found such as differences in work processes and differences in the preconditions before the introduction of the IT system. For example, in the internal medicine department, the organization of report writing was already very well organized before the new system was introduced. The new way to do it now put more work on the physicians, e.g. they now needed to correct the reports themselves, therefore, the physicians criticized the higher amount of work for them in this department. The results of this study will be used to better prepare the introduction in other departments (e.g. better organizational preparation of changed workflow).

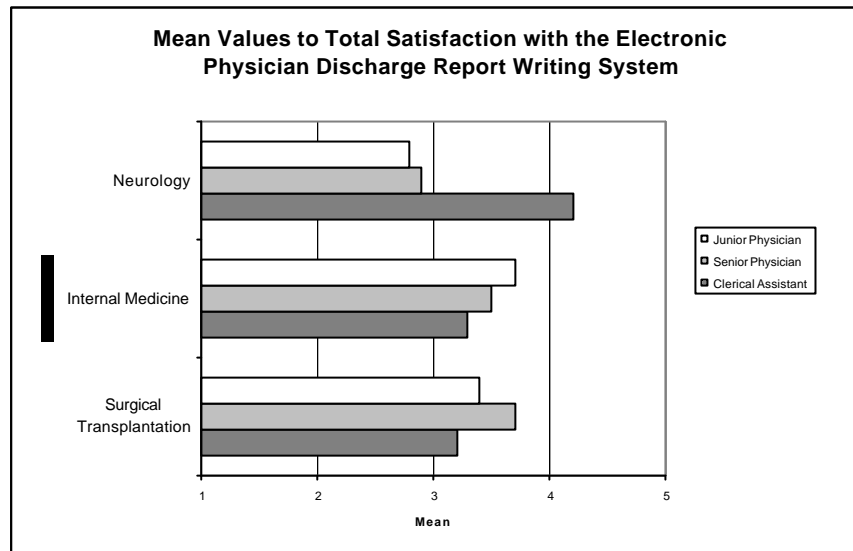


Figure 1: Mean values to total satisfaction in the three departments, based on 17 psychometric questions, by occupational group. Scale: 1 – very dissatisfied, 2 – fairly dissatisfied, 3 – neither satisfied, nor dissatisfied, 4 – fairly satisfied, 5 – very satisfied.

2.5 Validation of the questionnaire

The analysis of reliability of the used questionnaire showed satisfactory high results. The reliability index Cronbach alpha was 0.93 for the 17 psychometric questions. This shows a high internal consistency of the questions, meaning that all questions “point in the same direction”, thus showing a high reliability. A detailed analysis of each question showed a high selectivity index ($> 0,5$) for 15 of the 17 questions. For example, the question “The system supports research” showed a low selectivity index, as many users (e.g. clerical assistants) do not work in the research area. Thus, this question does not contribute much to their overall user acceptance. The questionnaires should now be modified to indicate that this question should only be used for special user groups.

For the analysis of validity, we chose to empirically check the criterion-related validity which consists of the correlation between the test result and an external criterion related to the construct. As external criterion in absence of a gold standard we chose the one question on the overall user satisfaction with the system (“How satisfied are you overall with the

system?") as we assumed, that people who are overall satisfied, should be consistent with their answers to the individual question. Each of the 15 remaining psychometric questions (after exclusion those two with low selectivity index) as well as the mean of the psychometric questions were positively correlated with the overall satisfaction with the system (correlation coefficient r between 0.42 and 0.78, $p < 0.001$ for each correlation). The psychometric questions are thus good correlated to the overall satisfaction, pointing to an acceptable validity of the questions.

The free text comments by 48 of the 59 users indicated to some aspects that were not covered by the psychometric questions. For example, the personal communication between clerical assistants and physicians as well as junior and senior physicians seemed to be negatively affected by the system. The system replaced the direct contact during the necessary correction cycle by asynchronous electronic communication. Some users thus criticized the reduced personal communication. This aspect could be added to an updates version of the questionnaire.

3. Discussion and Conclusion

One important application of questionnaires is to get an insight into what people think. In this case, questionnaires try to quantitatively measure a so-called "psychological construct" (such as user satisfaction). When rigorous methods for the development and validation of such psychometric questionnaires are followed, then they can measure a construct with high reliability. These questionnaires are typically based on standardized and well-structured closed questions, allowing for an exact and quantitative measurement.

It is often useful to add, however, some open-ended questions to a psychometric questionnaire. This allows the researcher to detect aspects which have not been taken into account during the development of the questionnaire. For example, in our case, changes in communication patterns seem to bother many users, but this was not part of the questionnaire. Thus, open-ended questions help to give hints for a future redesign and validity improvement of a questionnaire. The often used reliability measure Cronbach Alpha only informs us if all questions point in the same direction – not whether they completely cover the construct which the researcher is interested in.

The validity of questionnaires is usually more difficult to assess than its reliability. Typical validity approaches are the correlation with an external factor (criterion-based validity), the agreement with a theoretical prediction (construct-based validity), or subjective assessment of the validity of the questions by external experts (content validity). For criterion-based validity, it is difficult to prove that this external factor itself is valid. We chose the self-reported overall satisfaction as external factor, however, this is clearly not an optimal choice, as it is a subjective factor, and also part of the same questionnaire. The often used correlation with external usage patterns (e.g. high satisfaction and intensive use are correlation), as done e.g. in the theory of Davis [5], should however not be applied, when the use of our system is mostly mandatory (as in our case). Content validity, finally, is surely the weakest (but easiest) approach. We applied it by discussing the content and adequacy of the questionnaires with different users and with persons responsible for the IT.

Together with interviews and observations, questionnaires are an important methods for systems analysis and system evaluation. One big disadvantage of questionnaires is that they

do not allow for a flexible interaction between researcher and user, e.g. when something is unclear or important. Thus, psychometric questionnaires may present clear quantitative results, but they are often not sufficient for a full interpretation of the findings. For example, as in our case, we saw differences in user attitude between user groups and departments. However, the questionnaire did not help to clearly answer the 'why' – what was the reason? 'Why'-questions can better be answered using interactive methods such as interviews, or, more generally spoken, qualitative evaluation methods. Thus, for the interpretation of psychometric questionnaires, further external information normally needs to be gathered, such as background information, usage patterns, or additional open-ended questions.

Another possible drawback of questionnaires is that they are often only applied at one given point in time (as in our case). Any one-point psychometric analysis of user satisfaction in a complex clinical environment should thus only be interpreted very carefully, as changes of work processes, of staff members, of organizational structures, of patient clientele, as well as software updates, training sessions etc. all may affect the results, as they tend to change the fit between user, technology and task. And, therefore, user satisfaction may also quickly change because it is an indicator for this fit.

Summarizing, psychometric questionnaires can be seen as a reliable and valid method to measure certain psychological attributes such as user satisfaction. They can give clear quantitative results. In this sense, it is a method of the objectivist evaluation tradition. When validated questionnaires are available, then their application and analysis is rather efficient (in our example, the selection, adaptation and pre-test of the questionnaires took only 23 hours). Their new development, in contrast, is very time-consuming and should only be done when no available questionnaires are available. Psychometric questionnaires, however, allow only a limited interaction between researcher and user, their results may be very dependant on the time of measurement, and their interpretation often needs external knowledge. Those limitations have to be taken into account when preparing evaluation studies.

4. References

- [1] Rigby M. Evaluation: 16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives. In: Patel V, Rogers R, Haux R, editors. Proceedings of the 10th World Congress on Medical Informatics (Medinfo 2001). Amsterdam: IOS Press; 2001. p. 1198-202.
- [2] Heinrich L. Informationsmanagement. München Wien: Oldenbourg; 1999.
- [3] DeLone W, McLean E. Information systems success: the quest for the dependent variable. *Inform Systems Res* 1992; 3(1):60-95.
- [4] Goodhue D, Klein B, March S. User evaluations of IS as surrogates for objective performance. *Information & Management* 2000; 38:87-101.
- [5] Davis F. User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *Int J Man-Machine Studies* 1993; 38:475-487.
- [6] Ohmann C, Boy O, Yang Q. A systematic approach to the assessment of user satisfaction with health care systems: constructs, models and instruments. In: Pappas C, editors. Medical Informatics Europe '97. Conference proceedings. Amsterdam: IOS Press. 43 Pt B; 1997. p. 781-5.
- [7] Kaplan R, Saccuzzo D. Psychological testing: principles applications and issues. Belmont: Brooks Cole Wadsworth; 1982.
- [8] Cronbach L. Essentials of psychological testing. New York: Harper & Row; 1970.
- [9] Lienert GA, Raatz U. Testaufbau und Testanalyse. Weinheim: Psychologie Verlags Union; 1998.

- [10] Boy O, Ohmann C, Aust B, Eich H, Koller M, Knode O, et al. Systematische Evaluierung der Anwenderzufriedenheit von Ärzten mit einem Krankenhausinformationssystem - Erste Ergebnisse. In: Hasman A, Blobel B, Dudeck J, Engelbrecht R, Gell G, Prokosch H, editors. Medical Infobahn for Europe - Proceedings of MIE2000 and GMDS 2000. Amsterdam: IOS Press; 2000. p. 518-522.
- [11] Ammenwerth E, Kaiser F. Evaluation der Benutzerzufriedenheit mit der rechnergestützten Arztbrief- und Befundschreibung, UMIT Publication Series No. 1/2002, University for Health Informatics and Technology Tyrol, Innsbruck, 2002.

5. Address for correspondence

Ass.-Prof. Dr. Elske Ammenwerth, Research Group Assessment of Health Information Systems, University for Health Informatics and Technology Tyrol (UMIT), A - 6020 Innsbruck, elske.ammenwerth@umit.at.