

## Test IT: ISONORM 9241/10

Jochen Prümper

### 1 Introduction

„ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs), Part 10: Dialogue Principles“ (1995) is an official international standard and describes seven general ergonomic principles, which are independent of any specific dialogue technique; i.e. they are presented without reference to situations of use, applications, environments, or technology. The seven principles are as follows:

Table 1: The Dialog Principles of ISO 9241/10 (ISO 9241-10 1995, p. 5ff.)

Dialog Principle	Description
• Suitability for the task	A dialog is suitable for a task when it supports the user in the effective and efficient completion of the task.
• Self-descriptiveness	A dialog is self-descriptive when each dialog step is immediately comprehensible through feedback from the system or is explained to the user on request.
• Controllability	A dialog is controllable when the user is able to initiate and control the direction and pace of the interaction until the point at which the goal has been met.
• Conformity with user expectations	A dialog conforms with user expectations when it is consistent and corresponds to the user characteristics, such as task knowledge, education, experience, and to commonly accepted conventions.
• Error tolerance	A dialog is error tolerant if despite evident errors in input, the intended result may be achieved with either no or minimal corrective action by the user.
• Suitability for individualization	A dialog is capable of individualization when the interface software can be modified to suit the task needs, individual preferences, and skills of the user.
• Suitability for learning	A dialog is suitable for learning when it supports and guides the user in learning to use the system.

In order to analyse whether a software-system meets the dialog principles of ISO 9241-10, those principles must be characterized through a valid and reliable evaluation instrument. In this paper the software evaluation instrument ISONORM 9241/10 is presented; some reliability and validity results are discussed.

## 2 The questionnaire „ISONORM 9241/10“

The questionnaire „ISONORM 9241/10“ was designed as an evaluation instrument that is economical to use. Therefore, each of the seven principles was operationalized by five items only. The questionnaire has a seven-tier, bi-polar question format. The answers range from „- - -“ to „+ + +“ (coded: 1-7). Filling out the questionnaire takes approximately 10 minutes. Figure 1 shows a sample item referring to the principle „suitability for the task“.

Figure 1: Sample item from ISONORM 9241/10

<i>The software ...</i>	---	--	-	-+	+	++	+++	<i>The software ...</i>
requires unnecessary inputs.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	does not require unnecessary inputs.

## 3 Reliability

In the following the results with regard to Cronbach's alpha and the re-test reliability are presented.

### 3.1 Subjects

1265 users have up to now used ISONORM 9241/10 in the evaluation of software. The average age of the subjects was 34.5 years, 51.8% were female, 48.2% male. The users evaluated 178 different software programs. Their general computer experience was on average 77 months and their experience with the evaluated software 25 months. The question „How well do you know the evaluated software?“, was answered on a seven-point scale ranging from „very bad“ (1) to „very good“ (7) on average with 5.1.

### 3.2 Cronbach's alpha

As can be seen from table 1, the values for Cronbach's alpha with regard to the seven scales of ISONORM 9241/10 are satisfactory (scale means range between 4.4 and 5.3; standard deviations between 1.0 and 1.6).

### 3.3 Re-Test Reliability

To determine the re-test reliability, 49 users from the sample were asked at two times of measurement (on average after a period of 6.7 months) to evaluate the software they most frequently used. Total re-test reliability amounted to  $r = .77$  ( $p < .001$ ,  $N = 49$ ). Taking into consideration the possible factors which could reduce the re-test reliability (e.g. expertise changing over time might have an influence on how the software is evaluated), the ISONORM 9241/10 evaluations can be assessed as being stable over time (see tab. 2).

Table 2: Reliability of ISONORM 9241/10  
(<sup>⊙</sup> N between 1208 and 1251; <sup>⊠</sup> N= 49, \*  $p < .001$ )

Principle	alpha <sup>⊙</sup>	Re-Test <sup>⊠</sup>
• Suitability for the task	.81	.67*
• Self-descriptiveness	.86	.62*
• Controllability	.84	.64*
• Conformity with user expectations	.84	.60*
• Error tolerance	.87	.68*
• Suitability for individualization	.89	.63*
• Suitability for learning	.83	.59*

## 4 Validity

A first validity study was conducted by Prümper (1993). In this study it was shown that the user-friendliness of systems with a GUI was judged to be significantly better than those systems without a GUI, across all seven principles of ISO 9241-10. A second study includes correlating ISONORM 9241/10 with other software-evaluation instruments. For this purpose, two user-oriented questionnaires and one expert-evaluation inventory were used. The results will be presented in the following.

### 4.1 The User-oriented Questionnaires

The first questionnaire is a german translation of the „*QUIS*“ (long form; Shneiderman 1987) by Kinder (1991) and the second the „*BBD*“ from Spinus (1987). 31 users were requested to evaluate their software by means of ISONORM 9241/10 and the two questionnaires mentioned above. The users' general computer experience was on average 58 months, the experience with the evaluated software 18 months on average. The question „How well do you know the evaluated software?“ was answered on a seven-point scale ranging

from „very bad“ (1) to „very good“ (7) on average with 5.6. As can be seen from table 3, ISONORM 9241/10 significantly correlates with the other two user-oriented questionnaires.

Table 3: Validity of ISONORM 9241/10 (N = 31 users, \* p < .001)

	QUIS	BBD
ISONORM 9241/10	.73*	.71*

#### 4.2 The Expert-Evaluation Inventory

EVADIS II is a comprehensive evaluation system to be used by experts in the field of software ergonomics (Reiterer & Oppermann 1993). For the validity study 13 different software systems were evaluated by a specialist in software ergonomics together with an experienced user of the software. Each evaluation session lasted approximately three hours. A subsample (N=383) of the full sample was used for this analysis. Each software was evaluated by 29.5 users on average,  $N_{\min.} = 11$ ,  $N_{\max.} = 79$ . The correlation between EVADIS II and the mean ISONORM 9241/10 judgements was  $r = .59$  ( $p < .01$ ,  $N = 13$  software programs).

This shows that the user-oriented questionnaire generates results similar to the outcomes of the expert-evaluation system, but with much less effort in time and money.

## 5 Discussion

In this paper the software evaluation instrument ISONORM 9241/10 was introduced, and some reliability and validity results were presented.

The reliability (Cronbach's alpha and re-test reliability) as well as the validity yielded satisfactory results. A further advantage of ISONORM 9241/10 is that it was constructed on the basis of an international standard. Therefore, it is likely to become internationally accepted. The questionnaire ISONORM 9241/10 is currently available in German (Prümper & Anft 1997), Dutch (Prümper & Anft 1998) and English (Prümper & Anft 1999). Further translations are in the offing.

## 6 References

ISO 9241-10 (1995). *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 10: Dialogue principles*. Brussels: CEN - European Committee for Standardization.

- Kinder, A. (1991). *Testen und Bewerten von Software durch Benutzer*. Gießen: Justus-Liebig-Universität Gießen (unpublished master thesis).
- Prümper, J. (1993). Software-Evaluation based upon ISO 9241 Part 10. In T. Grechenig & M. Tscheligi (Eds.) *Human Computer Interaction* (p. 255-265). Berlin: Springer.
- Prümper, J. (1997). Der Benutzungsfragebogen ISONORM 9241/10: Ergebnisse zur Reliabilität und Validität. In: R. Liskowsky, B.M. Velichkovsky & W. Wüschmann (Eds.), *Software-Ergonomie '97 - Usability Engineering: Integration von Mensch-Computer-Interaktion und Software-Entwicklung* (p. 253-262). Stuttgart: Teubner.
- Prümper, J. & Anft, M. (1997). ISONORM 9241/10 - Beurteilung von Software auf Grundlage der Internationalen Ergonomie-Norm ISO 9241/10 (Questionnaire). In: Döbele-Martin, C., Martin, G., Richenhagen, G. & Prümper, J. (1997). *Ergonomie-Prüfer - Handlungshilfen zur ergonomischen Gestaltung der Bildschirmarbeit* (p. 103-113). Oberhausen: TBS.
- Prümper, J. & Anft, M. (1998). *ISONORM 9241/10 - Beoordeling van software op basis van de Internationale Ergonomie-Norm ISO 9241/10* (unpublished questionnaire).
- Prümper, J. & Anft, M. (1999). *ISONORM 9241/10 - Evaluation of Software based upon International Standard for Ergonomic Design ISO 9241/10* (unpublished questionnaire).
- Reiterer, H. & Oppermann, R. (1993). Evaluation of user interfaces: EVADIS II - a comprehensive evaluation approach. *Behaviour & Information Technology*, 3, 137-148.
- Shneiderman, B. (1987). *Designing the user interface*. Reading MA: Addison-Wesley.
- Spinas, P. (1987). *Arbeitspsychologische Aspekte der Benutzerfreundlichkeit von Bildschirmsystemen*. Zürich: ADAG.
- Smelcer, J.B. (1995). User errors in database query composition. *International Journal of Human-Computer Studies*, 42, 353-381.
- Zapf, D., Brodbeck, F.C., Frese, M., Peters, H. & Prümper, J. (1992). Errors in working with office computers: a first validation of a taxonomy for observed errors in a field setting. *International Journal of Human-Computer Interaction*, 4 (4), 311-339.

Quelle:

Prümper, J. (1999). Test IT: ISONORM 9241/10. In H.J. Bullinger & J. Ziegler (Eds.), *Human-Computer Interaction - Communication, Cooperation, and Application Design* (S. 1028-1032). Mahwah, New Jersey: Lawrence Erlbaum Associates.