
Die Evaluation von Software auf Grundlage des Entwurfs zur internationalen Ergonomie-Norm ISO 9241 Teil 10 als Beitrag zur partizipativen Systemgestaltung - ein Fallbeispiel ¹

Jochen Prümper, Michael Anft
DATA TRAIN GmbH, Berlin

Zusammenfassung

In dem vorliegenden Beitrag wird ein Verfahren zur Beurteilung von Software auf Grundlage des Entwurfs zur internationalen Ergonomie-Norm ISO 9241/10 vorgestellt. Neben der Bewertung von bestehender Software soll es sowohl zur Beurteilung von Prototypen im Rahmen eines iterativen Systemdesigns eingesetzt werden können, als auch als Grundlage für die Moderation mit Benutzergruppen dienen, wenn mit ihnen Designanforderungen erarbeitet werden. Der erste Einsatz fand im Rahmen der Neuentwicklung eines integrierten Softwarepakets für mittelständische Verlage statt. Das Ergebnis der Benutzergruppensitzung war eine Fülle von konkreten Hinweisen auf Schwachstellen der bestehenden Lösung.

1 Einleitung

Software-ergonomische Gestaltungsbemühungen sollen einen effizienteren System-einsatz, Belastungsminimierung beim Benutzer und geringere Schulungskosten durch leichtere Erlernbarkeit bewirken (Bullinger [3]). Auf diesem Weg bedarf es geeigneter Evaluationsverfahren mit deren Hilfe Schwachstellenanalysen bei bestehenden Systemen durchgeführt werden können bzw. beim Prototyping eine schrittweise Verfeinerung ermöglicht werden kann. Nur auf Basis einer zuverlässigen Ist-Analyse können gezielte Verbesserungsmaßnahmen ansetzen.

In den letzten Jahren wurde eine Reihe von Verfahren zur Beurteilung von Dialog-schnittstellen entwickelt (siehe z.B. Baitsch, Katz, Spinus und Ulich [1], Clegg et al. [5], Oppermann, Murchner, Reiterer & Koch [29], Ravden & Johnson [33],

¹ Diese Arbeit entstand im Rahmen des Projektes "Verlag 2000: Eine benutzerfreundliche integrierte Lösung für die mittelständische Verlags- und Druckereibranche unter Berücksichtigung von zu verbessernden Arbeitsbedingungen für die Beschäftigten." Dieses Projekt wird gefördert vom Bundesministerium für Forschung und Technologie, Förderschwerpunkt "Arbeit und Technik" (Förderkennzeichen: 01 HK 601 8). Wir danken Michael Frese, Klaus Hartmannsgruber, Ruth Orthey, Eckhard Ruch, Dankmar Scheuchl und Ferdinand Solzbacher für ihre Mitarbeit.

Shneiderman [34], Spinas [35]). Eine gute Übersicht liefern Hampe-Neteler und Rödiger [17].

EVADIS II (Oppermann et al. [29]) bietet Experten die Möglichkeit, eine Software entlang einer Reihe von Prüffragen an einer für den Benutzer typischen Arbeitsaufgabe (Standardaufgabe) zu evaluieren. Die 12 Kriterien von EVADIS II (z.B. "Verfügbarkeit", "Nützlichkeit", "Selbstbeschreibungsfähigkeit" oder "Erlernbarkeit") beziehen sich auf die Ein-/Ausgabe-, Dialog-, Werkzeug- und technische Organisationsschnittstelle des IFIP-Benutzerschnittstellenmodells (Dzida [8]). EVADIS II bietet dem Experten eine sehr detaillierte Evaluationsmöglichkeit und berücksichtigt dabei auch allgemeine Kriterien menschengerechter Arbeit. Die Nachteile von EVADIS II sind allerdings u.a. in der sehr aufwendigen Durchführung und den mangelnden Einsatzmöglichkeiten bei Sachbearbeitungsaufgaben zu sehen (vgl. Hampe-Neteler & Rödiger [17]). Darüber hinaus eignet sich EVADIS II nicht für eine Bewertung durch Endbenutzer.

Bei dem QUIS (Questionnaire for User Interface Satisfaction) (Norman & Shneiderman [28], Shneiderman [34]) beurteilen Benutzer auf Basis von Adjektivpaaren Bildschirmgestaltung, Systemmeldungen, Erlernbarkeit sowie allgemeine Systemaspekte, wie Antwortzeiten oder Zuverlässigkeit. Der QUIS ist in einer Lang- und Kurzform verfügbar und stellt ein robust einsetzbares Verfahren dar, für den selbst Reliabilitäts- und Validitätsergebnisse vorliegen (vgl. Chin, Diehl & Norman [4]). Kritisieren kann man allerdings die fehlende theoretische Fundierung (Kinder, [22]) und den Umstand, daß die verwendete Operationalisierung häufig einen deutlichen Handlungsbezug vermissen läßt und dafür eher Eindrucksurteile im Vordergrund stehen.

Die (englisch-sprachige) Checkliste von Ravden und Johnson [33] eignet sich sowohl für End-Benutzer als auch für Experten, die mit dem Entwurf, der Entwicklung und der Evaluation von Mensch-Computer-Schnittstellen befaßt sind. Diese Checkliste enthält 11 Themenblöcke - wie z.B. "Konsistenz", "Kompatibilität", "Informative Rückmeldungen", "Flexibilität und Kontrolle" oder "Fehlervermeidung- und korrektur" - bei denen Fragen danach beantwortet werden, wie häufig ein beschriebenes Ereignis auftritt. Hampe-Neteler & Rödiger [17] kritisieren bei diesem Verfahren, daß, ungeachtet einer annehmbaren theoretischen Fundierung, Aussagen über testtheoretische Güte kaum zu machen sind.

Zusammenfassend kann man sagen, daß trotz dieser Verfahrensentwicklungen noch stets ein Defizit an wissenschaftlich fundierten und gleichzeitig praxisgerechten Evaluationsverfahren besteht (vgl. von Benda [2], Kirakowski & Corbett [23], Nielsen [27], Shneiderman [34]). Kirakowski & Corbett [23] und Shneiderman [34] kritisieren bei der Entwicklung von Software-Evaluationsverfahren die fehlende

Berücksichtigung von Gütekriterien wie Reliabilität und Validität und Nielsen [27] unterstreicht insbesondere das Defizit an kostengünstigen, robust einsetzbaren Verfahren, was zur Folge hat, daß im kommerziellen Bereich Methoden wie Benutzerpartizipation, Evaluation von Dialogschnittstellen oder iteratives Design erst gar nicht zum Einsatz kommen (Gould & Lewis [14]).

Zudem verwenden die diversen Evaluationsinstrumente einen bunten Strauß an mehr oder weniger einsichtigen oder zumindest allgemein anerkannten Bewertungskriterien; ein Umstand, der dem Praktiker die Entscheidung für oder gegen den Einsatz eines Verfahrens sicherlich nicht erleichtert.

1988 wurde nun, nach Vorarbeiten von Dzida und Itzfeldt [9], mit der DIN 66 234/8 [6] der Versuch unternommen, zumindest auf nationaler Ebene eine Einigung hinsichtlich ergonomischer Dialoggestaltungsprinzipien herzustellen. Obwohl diese Norm nicht unwiderrspochen blieb (siehe z.B. Moll & Ulich [25]) und ihr teilweise "leicht begriffliche Unschärfe" (Frese & Brodbeck [12], S. 113) vorgeworfen wurde, so gibt sie doch immerhin dem Software-Entwickler eine einheitliche Nomenklatur in die Hand, die ihn davon befreit "undifferenziert über *Benutzerfreundlichkeit* zu reden" (Dzida [7], S. 443).

2 Der Normentwurf ISO 9241/10

Im September 1991 wurde von der International Organisation for Standardisation (ISO) ein Entwurf für die ergonomischen Grundsätze zur Gestaltung von Dialogschnittstellen vorgestellt (ISO 9241/10 [19]).

Mit dem Norm-Entwurf ISO 9241/10 (1991, 1992) - der sich zu großen Teilen an die DIN 66 234/8 (1988) anlehnt - wurde erstmals der Vorschlag für eine internationale Standardisierung für Dialogschnittstellen unterbreitet. Er enthält die sieben Grundsätze: Aufgabenangemessenheit, Selbstbeschreibungsfähigkeit, Steuerbarkeit, Erwartungskonformität, Fehlerrobustheit, Individualisierbarkeit und Erlernbarkeit.

DIN 66234/8 wurde damit um die zwei Kriterien Individualisierbarkeit (der Grad, indem der Benutzer den Dialog den eigenen Anforderungen und Bedürfnissen entsprechend anpassen kann) und Erlernbarkeit (der Grad, indem der Dialog den Lernprozeß des Benutzers so unterstützt, daß der Lernaufwand möglichst gering ist) erweitert.

3 Eine Operationalisierung des Normentwurfs ISO 9241/10

Um zu analysieren, inwieweit ein Softwaresystem Grundsätzen entspricht, wie sie in dem Entwurf zur ISO 9241/10 formuliert wurden, müssen diese über ein Beurtei-

lungsverfahren operationalisiert werden - "sie sind interpretations- und konkretisierungsbedürftig" (Hampe-Neteler und Rödiger [17], S. 60).

Bei der Operationalisierung der ISO 9241/10 war es unser Bestreben, ein effizient einsetzbares, robustes Verfahren zu entwickeln, das erste Hinweise auf Schwachstellen und somit Verbesserungspotential bezüglich der Benutzerfreundlichkeit von Softwaresystemen gibt. Es soll sowohl zur Beurteilung von bereits etablierter Software eingesetzt werden, als auch zur ersten Beurteilung von Prototypen im Rahmen eines iterativen Systemdesigns. Darüber hinaus soll es eine geeignete Grundlage für die Moderation mit Benutzergruppen sein, wenn mit ihnen konkrete Designforderungen erarbeitet werden.

Die Fragebogenmethode erschien uns für diese Zwecke die am besten geeignete, da ihr Vorteil in einer wenig aufwendigen und leichten Durchführung liegt, die in allen Phasen einer Systementwicklung eingesetzt werden kann.

In Anlehnung an handlungstheoretische Überlegungen (Hacker [15], Volpert [38]) bemühten wir uns bei der Itemformulierung um möglichst tätigkeitsnahe Beschreibungen.

Den "Granulationsgrad" (Rauterberg [32], S. 11) der Beschreibungssprache versuchten wir so zu wählen, daß die spezifischen Eigenschaften verschiedener Oberflächentypen hinreichend genau differenziert werden können, und er gleichzeitig auf möglichst viele Oberflächentypen einheitlich anwendbar ist. Mit diesem mittleren Granulationsgrad soll sich das Verfahren für die Evaluation einer möglichst große Bandbreite von Softwaresystemen anbieten.

Um dem Anspruch eines ökonomisch einsetzbaren Verfahrens gerecht zu werden, wurde der Umfang des Fragebogens so zugeschnitten, daß jeder einzelne der sieben Grundsätze von ISO 9241/10 über fünf Items operationalisiert wurde. Insgesamt umfaßt das Verfahren also 35 Items.

Bei der Fragebogenkonstruktion entschieden wir uns für ein sieben-stufiges, bipolares Frageformat, mit einem Antwortschema von - - - bis +++ (codiert: 1 - 7). Von der Gegenüberstellung dieser positiven und negativen Ausprägungen versprechen wir uns, daß die Beurteiler angeregt werden, differenzierter über den Sachverhalt nachzudenken, und sie nicht von vorneherein durch Vorgabe einer Richtung beeinflußt werden (Tränkle [36]).

Da das Beurteilungsinstrument sich eng an dem Normentwurf ISO 9241/10 orientiert, taufen wir es auf den Namen ISONORM 9241/10.

Die Software ...

	---	--	-	-/+	+	++	+++	
bietet keine Möglichkeit, die Arbeit an jedem Punkt zu unterbrechen und dort später ohne Verluste wieder weiterzumachen.								bietet die Möglichkeit, die Arbeit an jedem Punkt zu unterbrechen und dort später ohne Verluste wieder weiterzumachen.
erzwingt eine unnötig starre Einhaltung von Bearbeitungsschritten.								erzwingt keine unnötig starre Einhaltung von Bearbeitungsschritten.
ermöglicht keinen leichten Wechsel zwischen einzelnen Menüs oder Masken.								ermöglicht einen leichten Wechsel zwischen einzelnen Menüs oder Masken.
ist so gestaltet, daß der Benutzer nicht beeinflussen kann, wie und welche Informationen am Bildschirm dargeboten werden.								ist so gestaltet, daß der Benutzer beeinflussen kann, wie und welche Informationen am Bildschirm dargeboten werden.
erzwingt unnötige Unterbrechungen der Arbeit.								erzwingt keine unnötigen Unterbrechungen der Arbeit.

Abb.1: Operationalisierung des Grundsatzes Steuerbarkeit mit Mittelwertsprofil des Adreßmoduls einer Verlagssoftware (N=42)

Abb. 1 liefert aus dem Verfahren ISONORM 9241/10 ein Beispiel für den Grundsatz "Steuerbarkeit" (mit den Ergebnissen einer Beurteilung von 42 Benutzern des Adreß-Moduls einer Verlagssoftware).

4 Einsatz des ISONORM 9241/10 im Rahmen partizipativer Systemgestaltung

Der erste Einsatz des Verfahrens ISONORM 9241/10 findet im Rahmen eines Projektes statt, bei dem es um die Entwicklung eines integrierten Softwarepakets für klein- und mittelständische Verlage geht.

Unter dem Stichwort "partizipative, iterative Software-Entwicklung" (Floyd [11], Hacker, Müller-Holz auf der Heide & Aschersleben [16], Keil-Slawik [21], Ulich [37]) wird die Softwareentwicklung in diesem Projekt in enger Zusammenarbeit mit den potentiellen Endbenutzern durchgeführt.

Im Vordergrund steht die Form der aktiven Partizipation (Heilmann [18]), bei der den Benutzern die Möglichkeit gegeben wird, den Systemgestaltungsprozeß aktiv mitzubeeinflussen.

Zu diesem Zweck kommen unter der Anleitung von Arbeitswissenschaftlern Anwender und Entwickler in regelmäßigen Abständen zusammen, um miteinander die einzelnen Designphasen abzustimmen. Bei diesen Treffen werden den späteren Benutzern Prototypen der Verlagssoftware vorgestellt. Damit haben die Anwender die Möglichkeit, den Softwareentwicklungsprozeß frühzeitig zu beeinflussen, und das Programm ihren individuellen Bedürfnissen entsprechend anzupassen. Dies verhindert, daß die Software sozusagen vom Schreibtisch aus entwickelt wird, und ein Produkt entsteht, das erst nach umfangreichen und kostspieligen Anpassungen einsetzbar ist.

Allerdings hat in den meisten Fällen lediglich ein ausgewählter Kreis von Endbenutzern die Möglichkeit, an diesen Benutzergruppentreffen teilzunehmen. Deshalb sollten die restlichen von der Softwarelösung betroffenen Personen zumindest den Status der passiven Partizipation bekommen. Obwohl diese Benutzer keine unmittelbare Einflußnahme auf die Gestaltung der Software haben, können ihre Anforderungen so dennoch Berücksichtigung finden.

Um beide Partizipationsformen in der vorliegenden Softwareentwicklung zu verwirklichen, beurteilten im Rahmen der Definition des Teilmoduls "Adresse" 42 Benutzer aus drei Verlagen diese anhand des Verfahrens ISONORM 9241/10, mit dem sie bislang arbeiteten (passive Partizipation). Aus diesen Verlagen nahmen sieben Benutzer an dem ersten Benutzergruppentreffen teil, bei dem es darum ging, bestehende Schwachstellen der existierenden Lösung zu erarbeiten (aktive Partizipation).

Zu Beginn der Sitzung wurden der Benutzergruppe die Ergebnisse der Erhebung als Mittelwerte auf Itemebene präsentiert. Mittels der Technik der Kartenfrage (Klebert, Schrader & Straub [24]) wurden zu jedem Item Beispiele erfragt. Die Äußerungen wurden im Rahmen einer Gruppendiskussion aufgearbeitet und präzisiert. Dabei stand den Benutzern die von ihnen beurteilte Software zur Verfügung, sodaß sie die Möglichkeit hatten, ihre Beurteilungen anhand konkreter Situationen zu erläutern. Abb. 2 liefert einen Überblick über das Vorgehen.

1. Erläuterung der sieben Grundsätze der ISO 9241/10
2. Präsentation der Profile der einzelnen ISONORM 9241/10-Grundsätze
3. Illustration der Ergebnisse durch die Benutzer anhand von Beispielen
4. Sammeln, Besprechen und Sortieren der Karten im Plenum
5. Erarbeiten von Verbesserungsvorschlägen

Abb. 2: Ablauf des Vorgehens bei der visualisierten Diskussion des ISONORM 9241/10 im Rahmen partizipativer Systemgestaltung

Dieser Ablauf wurde, entsprechend der Anzahl der Grundsätze der ISONORM 9241/10, siebenmal durchlaufen. Der Zeitraum umfaßte sieben Stunden.

Nicht immer war es den Benutzern möglich, zu jeder einzelnen Operationalisierung der jeweiligen Grundsätze spontan ein Beispiel zu generieren. In diesen Fällen erläuterte der Moderator in Anlehnung an die "critical incident technique" (Flanagan [10]), die Einzelitems mit guten und schlechten Beispielen aus anderen Programmen oder er forderte die Benutzer auf, kleinere Arbeitsaufgaben mit dem System zu bearbeiten. Dennoch trat die Schwierigkeit auf, daß die Benutzer die einzelnen Grundsätze der ISO 9241/10 nicht immer eindeutig voneinander zu unterscheiden wußten. So kamen z.B. bei dem Grundsatz "Aufgabenangemessenheit" eine Reihe von Nennungen, die aus konzeptioneller Sicht den Grundsätzen "Steuerbarkeit" oder "Erwartungskonformität" zuzurechnen wären. Unter Umständen wird an dieser Stelle eine ungenügende theoretische Ableitung der sieben Grundsätze der ISO 9241/10 deutlich, die damit einen breiten Interpretationsraum provoziert und somit in ihrer Operationalisierung zu Überschneidungen führen muß (besonders deutlich wurde dies bei den beiden Grundsätzen "Steuerbarkeit" und "Individualisierbarkeit"). Dies hatte zur Folge, daß zu den Grundsätzen, die am Anfang der Moderation präsentiert wurden, von den Benutzern mehr Beispiele erarbeitet wurden, als zu denen, die am Ende der Moderation zur Sprache kamen. Dennoch war das Ergebnis letztendlich eine Fülle von konkreten Hinweisen auf Schwachstellen, die das Ergebnis der ISONORM 9241/10-Fragebogenerhebung sinnvoll bereicherten und erweiterten.

Diese Verknüpfung von Fragebogenmethode und visualisierter Diskussion (Peschke & Wittstock [30]) hat sich in unserem Fall bewährt. Die interpretationsbedürftigen, nackten Fragebogenergebnisse konnten sinnvoll ergänzt und aufbereitet werden.

Zur Illustration seien exemplarisch für die einzelnen Grundsätze jeweils eine Operationalisierung und ein Ergebnis der Kartenabfrage dargestellt (vgl. Abb. 3).

Zusammenfassend kann man sagen (vgl. Abb. 4), daß das von uns untersuchte Adreßmodul in der Bewertung der Benutzer lediglich in den drei Grundsätzen

Grundsatz	Item	Verletzung des Grundsatzes
Aufgabenangemessenheit	Die Software ist schlecht auf die Anforderungen der Arbeit zugeschnitten.	Die Adreß-Maske enthält überflüssige Felder, die nie bzw. nur selten ausgefüllt werden müssen.
Selbstbeschreibungsfähigkeit	Die Software liefert in unzureichendem Maße Informationen darüber, welche Eingaben zulässig oder nötig sind.	Die Adreß-Maske trifft keine Unterscheidung zwischen Feldern, bei denen eine Eingabe erfolgen muß oder die lediglich optional ausgefüllt werden können.
Steuerbarkeit	Die Software erzwingt eine unnötig starre Einhaltung von Bearbeitungsschritten.	Bei der Adreß-Neuanlage wird zwingend in die Erfassungsmaske für Konditionsvereinbarungen verzweigt, obwohl meist keine Konditionen benötigt werden.
Erwartungskonformität	Die Software erschwert die Orientierung durch eine uneinheitliche Gestaltung.	In der Adreß-Maske sind die Funktionstasten anders belegt als zum Beispiel in den Bereichen Auftrags erfassung, Buchhaltung oder Abonnement.
Fehlerrobustheit	Die Software gibt keine konkreten Hinweise zur Fehlerbehebung.	Wenn eine Adresse unvollständig eingegeben wird, dann springt der Positionsanzeiger in das erste Feld der Maske und es erscheint die Fehlermeldung "Eingabe falsch oder fehlend". Auf welches Feld sich die Fehlermeldung bezieht, ist nicht ersichtlich.
Individualisierbarkeit	Die Software läßt sich von dem Benutzer schwer erweitern, wenn für ihn neue Aufgaben entstehen.	Werden neue Kodierungen (z.B. Branchencodes, Hobbies, Fachgruppen) benötigt, so kann der Benutzer diese Ergänzung nicht selbst vornehmen.
Erlernbarkeit	Die Software ist schlecht ohne fremde Hilfe oder Handbuch erlernbar.	Die Einarbeitung ist mühsam, weil keine Bedienerhilfe existiert.

Abb. 3: Die sieben Grundsätze der ISO 9241 Teil 10, Beispielimens für die Operationalisierung und exemplarische Ergebnisse aus der Benutzergruppensitzung

Aufgabenangemessenheit, Erwartungskonformität und Erlernbarkeit die +-Marke erreichte. Selbstbeschreibungsfähigkeit, Steuerbarkeit und Fehlerrobustheit wurden noch ungünstiger bewertet; sie ragen nur knapp in den positiven Bereich hinein. Die Individualisierbarkeit des Adreßmoduls schnitt in der Bewertung der Benutzer am schlechtesten ab; sie bewegte sich im negativen Bereich.

Bezüglich der Aufgabenangemessenheit läßt sich festhalten, daß das Adreßmodul nicht alle Funktionen bietet, um die anfallenden Aufgaben zu bewältigen, und daß überflüssige Eingaben getätigt werden müssen. Die mangelnde Selbstbeschreibungsfähigkeit ist in erster Linie auf eine unzureichende Hilfsfunktion und auf fehlende, kontextspezifische Erklärungen zurückzuführen. Einer der kritischeren Punkte ist auch der Grundsatz der Steuerbarkeit. Hier wird insbesondere bemängelt, daß die

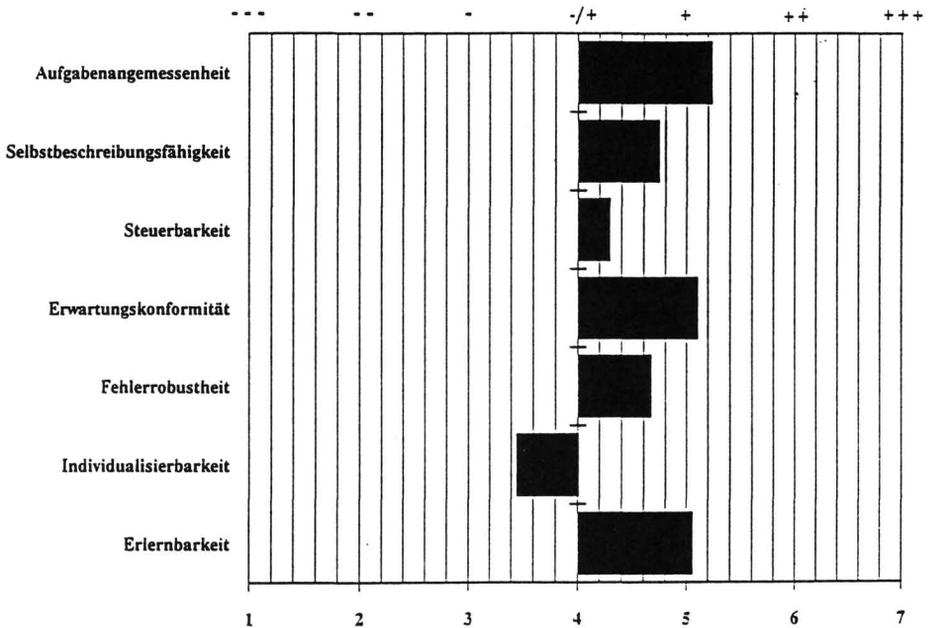


Abb. 4: Die Ergebnisse der Benutzerbefragung (N=42)

Software nur ungenügend die Möglichkeit bietet, die Arbeit zu unterbrechen, und ohne Verluste später wieder weiterzumachen. Darüber hinaus kann der Benutzer kaum beeinflussen, wie und welche Informationen am Bildschirm dargeboten werden. Im Rahmen der Erwartungskonformität wird kritisiert, daß die Software zu uneinheitlich gestaltet ist, und daß dadurch die Orientierung erschwert wird, sowie der Umstand, daß sich Erfahrungen aus den einen Programmteilen nicht immer auf andere Programmteile übertragen lassen. Im Kontext Fehlerrobustheit schlägt besonders negativ zu Buche, daß die Software keine konkreten Hinweise zur Fehlerbehebung gibt, und daß sie nicht immer unmittelbar über fehlerhafte Eingaben informiert. In der besonders ungünstigen Bewertung der Individualisierbarkeit spiegelt sich wieder, daß sich das Adreßmodul von den Benutzern nur schwer erweitern läßt, wenn für sie neue Aufgaben entstehen, und daß es sich schlecht an ihre persönliche, individuelle Art der Arbeitserledigung anpassen läßt. Die Erlernbarkeit erleidet Einbußen, weil die Software nicht dazu ermutigt, auch einmal neue Funktionen auszuprobieren. Zudem steht keine Bedienerhilfe zur Verfügung, sodaß die Software nur schwer ohne fremde Hilfe erlernbar ist.

Deutlich wurde bei allen Grundsätzen, daß bei dem untersuchten Adreßmodul von den Benutzern noch viele Verbesserungen gefordert werden.

5 Zusammenfassung und Ausblick

In dem vorliegenden Beitrag wurde ein Verfahren zur Beurteilung von Software auf Grundlage des Entwurfs zur internationalen Ergonomie-Norm ISO 9241/10 als Beitrag zur partizipativen Systemgestaltung vorgestellt. Das Verfahren ISONORM 9241/10 erwies sich im Rahmen einer Software-Entwicklung, die in enger Zusammenarbeit mit Endbenutzern durchgeführt wird, als ein praktikables Mittel zu einer ersten Schwachstellenanalyse. Im Fall der vorgestellten Benutzergruppensitzung lag seine besondere Qualität darin, daß es einen definierten roten Faden bot, anhand dessen Benutzer und Softwareentwickler das zur Bearbeitung anstehende Produkt diskutieren konnten. Allerdings ersetzt das von uns durchgeführte Prozedere keine Tiefenanalysen, wie sie etwa durch beobachtungsbasierte Fehleranalysen (Frese & Zapf [13], Zapf, Brodbeck & Prümper [39]), computerunterstützte Logfileprotokolle (Müller-Holz auf der Heide, Aschersleben, Hacker & Bartsch [26]) oder detaillierte Leitfäden (Oppermann et al. [29]) möglich sind. Der Einsatzschwerpunkt lag bei der vorliegenden Untersuchung darin, besonders am Anfang des Softwareentwicklungsprozesses Endbenutzer und Entwickler miteinander ins Gespräch zu bringen und dabei grundlegende Designprobleme zu erörtern.

Bislang können wir noch nicht beurteilen, wie sich die Qualität der von uns untersuchten Software im Vergleich zu anderen Lösungen verhält. Unser Streben geht deshalb dahin, für eine repräsentative Menge von Softwareprodukten und eine repräsentative Benutzerstichprobe Beurteilungen zu gewinnen, um typische Profilverläufe von diversen Branchen- und Standardlösungen, Systemoberflächen, Betriebssystemen etc. zu erhalten. Erste Ergebnisse, bei dem ein Vergleich zwischen Großrechner- und PC-Anwendungen durchgeführt wurde (Prümper & Anft [31]) sprechen dafür, daß der ISONORM 9241/10 deutliche Unterschiede zwischen verschiedenen Systemen aufzuzeigen vermag.

Bevor derartige Messungen jedoch aussagekräftig werden können, gilt es, das beschriebene Verfahren den notwendigen Objektivitäts-, Validitäts- und Reliabilitätsuntersuchungen zu unterziehen (vgl. Kirakowski & Corbett [23], Rauterberg [32]). Entsprechende Untersuchungen sind im Gange.

6 Literatur

- [1] Baitsch, C., Katz, C., Spinas, P., & Ulich, E. (1989). Computerunterstützte Büroarbeit: Ein Leitfaden für Organisation und Gestaltung. Zürich: Verlag der Fachvereine.
- [2] Benda, H. von (1988). Neue Technologien: Mensch-Computer-Interaktion. In: F. Frey, C. Graf Hoyos & D. Stahlberg (Hrsg.), *Angewandte Psychologie: Ergebnisse und neue Perspektiven* (S. 170-186). München: Urban und Schwarzenberg.

- [3] Bullinger, H.J. (1990). Wettbewerbsvorteile durch Software-Ergonomie. In: H.J. Bullinger. (Hrsg.), *Software-Ergonomie in der Praxis: Richtlinien, Methoden und Werkzeuge für die Gestaltung interaktiver Systeme* (S. 13-23). Berlin: Springer.
- [4] Chin, J.P., Diehl, V.A., & Norman, K.L. (1988). Development of an instrument measuring user satisfaction of the human-computer-interface. CHI'88 Conference Proceedings - Human Factors in Computing Systems (S. 213-218). New York: ACM.
- [5] Clegg, C. W., Warr, P., Green, T., Monk, A., Kemp, N., Allison, G., & Landsdale, M. (1988). *People and Computers - How to evaluate your companys new technology*. Chichester: Ellis Horwood.
- [6] DIN Norm 66 234 Teil 8 (1988). *Bildschirmarbeitsplätze: Grundsätze ergonomischer Dialoggestaltung*. Berlin: Beuth.
- [7] Dzida, W. (1985). Ergonomische Normen für die Dialoggestaltung: Wem nützen die Gestaltungsgrundsätze im Entwurf DIN 66 234, Teil 8? In: H.J. Bullinger (Hrsg.), *Software-Ergonomie '85* (S. 430-443). Stuttgart: Teubner.
- [8] Dzida, W. (1988). Modellierung und Bewertung von Benutzerschnittstellen. *Software Kurier*, 1, 13-28.
- [9] Dzida, W., & Itzfeldt, W.D. (1978). User-perceived quality of interactive systems. *IEEE Transactions on Software Engineering*, SE-4 (4), 270-276.
- [10] Flanagan, J.G. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327-358.
- [11] Floyd, C. (1984). A systematic look at prototyping. In: R. Budde, K. Kuhlenkamp, L. Mathiassen, & M. Zillighoven (Eds.), *Approaches to prototyping* (S. 1-18). Berlin: Springer.
- [12] Frese, M., & Brodbeck, F. C. (1989). *Computer in Büro und Verwaltung*. Berlin: Springer.
- [13] Frese, M., & Zapf, D. (Hrsg.) (1991). *Fehler bei der Arbeit mit dem Computer: Ergebnisse von Beobachtungen und Befragungen im Bürobereich*. Bern: Huber.
- [14] Gould, J.D., und Lewis, C. (1983). Designing for usability: Key principles and what designers think. *Proceedings of the ACM, CHI'83 Boston, MA*, 50-53.
- [15] Hacker, W. (1986). *Arbeitspsychologie. Psychische Regulation von Arbeitstätigkeiten*. Bern: Huber.
- [16] Hacker, S., Müller-Holz auf der Heide, B., und Aschersleben, G. (1991). Prototyping in einem Designteam - Vorgehen und Erfahrungen bei einer Software-Entwicklung unter Benutzerbeteiligung. In: M. Frese, Chr. Kasten, C. Skarpelis, & B. Zang-Scheucher (Hrsg.), *Software für die Arbeit von morgen: Bilanz und Perspektiven anwendungsorientierter Forschung* (S. 179-189). Berlin, Heidelberg, New York: Springer.
- [17] Hampe-Neteler, W., & Rödiger, K.-H. (1992). *Software-Ergonomie - Verfahren der Evaluierung und Standards zur Entwicklung von Benutzeroberflächen*. Forschungsbericht des Studiengangs Informatik der Universität Bremen, Bericht Nr. 2/92.
- [18] Heilmann, H. (1981). *Modelle und Methoden der Benutzermitwirkung in Mensch-Computer-Systemen*. Stuttgart: Wiesbaden.
- [19] ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs). (1991). Part 10: Dialogue principles - First committee draft, September 1991.
- [20] ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs). (1992). Part 10: Dialogue principles - Second committee draft, June 1992.
- [21] Keil-Slawik, R. (1988). Integrierte Systementwicklung. In: E. Nullmeier & K.-H. Rödiger (Hrsg.), *Dialogsysteme in der Arbeitswelt* (S. 205-228). Mannheim: BI-Wissenschaftsverlag.
- [22] Kinder, A. (1991). *Testen und Bewerten von Software durch Benutzer*. Gießen: Justus-Liebig-Universität Gießen (unveröff. Diplomarbeit).

- [23] Kirakowski, J., & Corbett, M. (1990). *Effective methodology for the study of HCI*. Amsterdam: North Holland.
- [24] Klebert, K., Schrader, E., & Straub, W.G. (1985). *Kurzmoderation: Anwendung der Moderationsmethode in Betrieb, Schule und Hochschule, Kirche und Politik, Sozialbereich und Familie bei Besprechungen und Präsentationen*. Hamburg: Windmühle.
- [25] Moll, T., & Ulich, E. (1988). Einige methodische Fragen in der Analyse von Mensch-Computer-Interaktionen. *Zeitschrift für Arbeitswissenschaft*, 42, 70-76.
- [26] Müller-Holz auf der Heide, B., Aschersleben, G., Hacker, S., & Bartsch, T. (1991). Methoden zur empirischen Bewertung der Benutzerfreundlichkeit von Bürosoftware im Rahmen von Prototyping. In: M. Frese, Chr. Kasten, C. Skarpelis, & B. Zang-Scheucher (Hrsg.), *Software für die Arbeit von morgen: Bilanz und Perspektiven anwendungsorientierter Forschung* (S. 409-420). Berlin, Heidelberg, New York: Springer-Verlag.
- [27] Nielsen, J. (1989). Usability engineering at a discount. In G. Salvendy & M.J. Smith (Eds.), *Designing and using human-computer interfaces and knowledge based systems* (S. 394-401). Amsterdam: Elsevier.
- [28] Norman, K., & Shneiderman, B. (1989). *Questionnaire for user interface satisfaction Vers. 5.0*, Maryland, College Park, USA: University of Maryland, HCI-Lab.
- [29] Oppermann, R., Murchner, B., Reiterer, H., & Koch, M. (1992). *Software-ergonomische Evaluation. Der Leitfaden EVADIS II*. Berlin, New York: De Gruyter.
- [30] Peschke, H., & Wittstock, M. (1987). Benutzerbeteiligung im Software-Entwicklungsprozeß. In: K.P. Fähnrich (Hrsg.), *Software-Ergonomie* (S. 81-92). München: Oldenbourg.
- [31] Prümper, J., & Anft, M. (1992). *Entwicklung eines Software-Beurteilungsverfahrens auf Grundlage des Entwurfs zu ISO 9241/10. Konzept und erste Ergebnisse*. Manuskript. Data Train, Berlin.
- [32] Rauterberg, M. (1992). Läßt sich die Gebrauchstauglichkeit interaktiver Software messen? Und wenn ja, wie? *Ergonomie und Informatik*, 16, 3-18.
- [33] Ravden, S., & Johnson, G. (1989). *Evaluating usability of human-computer interfaces: a practical method*. Chichester: Ellis Horwood.
- [34] Shneiderman, B. (1992). *Designing the user interface: Strategies for effective human-computer interaction*. Reading, Massachusetts: Addison-Wesley.
- [35] Spinaz, P. (1987). *Arbeitspsychologische Aspekte der Benutzerfreundlichkeit von Bildschirmssystemen*. Zürich: ADAG.
- [36] Tränkle, U. (1983). Fragebogenkonstruktion. In: H. Feger & J. Bredenkamp (Hrsg.), *Enzyklopädie der Psychologie, Themenbereich B Methodologie und Methoden, Serie I Forschungsmethoden der Psychologie, Band 2 Datenerhebung* (S. 222-301). Göttingen: Hogrefe.
- [37] Ulich, E. (1987). Zur Frage der Individualisierung von Arbeitstätigkeiten unter besonderer Berücksichtigung der Mensch-Computer-Interaktion. *Zeitschrift für Arbeits- und Organisationspsychologie*, 31, 86-93.
- [38] Volpert, W. (1987). Psychische Regulation von Arbeitstätigkeiten. In: U. Kleinbeck & J. Rutenfranz (Hrsg.), *Enzyklopädie der Psychologie - Arbeitspsychologie* (S. 1-42). Göttingen: Hogrefe.
- [39] Zapf, D., Brodbeck, F.C., & Prümper, J. (1989). Handlungsorientierte Fehlertaxonomie in der Mensch-Computer Interaktion. *Zeitschrift für Arbeits- und Organisationspsychologie*, 33, 178-187.

Jochen Prümper und Michael Anft

DATA TRAIN GmbH, Ibsenstr. 13, O-1071 Berlin