

Figl, K. (2009). Usability-Fragebögen im Vergleich. Tagungsband Mensch & Computer. Oldenbourg: Berlin, 143-152.

# ISONORM 9241/10 und Isometrics: Usability-Fragebögen im Vergleich

Kathrin Figl

Institut für Wirtschaftsinformatik und Neue Medien, WU Wirtschaftsuniversität Wien

## **Zusammenfassung**

Eine der wichtigsten Methoden zur Evaluation von Software-Usability ist der Einsatz von Benutzerfragebögen. Diese Studie widmet sich dem Vergleich der zwei im deutschsprachigen Raum verbreitetsten Fragebögen, Isonorm und Isometrics, die beide dasselbe Ziel verfolgen, Software gemäß der ISO Norm 9241/10 zu evaluieren. Im Rahmen eines experimentellen Designs beurteilten 50 Versuchspersonen die Usability der Softwarepakete SPSS bzw. STATISTICA mit Hilfe der beiden Fragebögen. Die Bewertungen anhand der beiden Fragebögen lieferten dieselben Ergebnisse hinsichtlich verschiedener Usability Aspekte und korrelierten positiv miteinander. Die Ergebnisse der Fragebogenbeurteilung deuten darauf hin, dass es sowohl Personen gibt, die den Isonorm-Fragebogen präferieren, wie solche, die den Isometrics Fragebogen angenehmer und schneller zu beantworten finden.

## 1 Einleitung

Usability-Fragebögen kommt im Rahmen der Evaluation von Software eine große Bedeutung zu. In dieser Studie werden zwei Fragebögen verglichen - Isonorm und Isometrics, die beide die ISO Norm 9241/10 operationalisieren. Der Vergleich dieser beiden Fragebögen ist deshalb von besonderem Interesse, da sie einerseits in diesem Gebiet die beiden populärsten Fragebögen in deutscher Sprache sind, und weil andererseits zwar einige Untersuchungen hinsichtlich testtheoretischer Gütekriterien vorhanden sind, jedoch noch keine, in denen beide Fragebögen gemeinsam eingesetzt und verglichen wurden. Der Artikel ist wie folgt aufgebaut: Nach einer allgemeinen Einführung in Usability mit Schwerpunkt auf der 9241/110 wird ein Überblick über Usability-Fragebögen gegeben und die empirische Untersuchung dargestellt.

## 2 Usability und die ISO Norm 9241-110

Usability, zu Deutsch auch Gebrauchstauglichkeit, beschäftigt sich mit Grundsätzen ergonomischer Softwaregestaltung und hat sich im deutschen Sprachraum auch gegenüber dem enger gefassten Begriff der „Benutzerfreundlichkeit“ durchgesetzt. In der ISO-Norm 9241/11 wird Usability dadurch definiert, wie „effektiv, effizient und zufrieden stellend“ ein Benutzer seine Ziele erreichen kann indem er ein technisches Objekt zur Lösung von Aufgaben verwendet (International Organization for Standardization 1998).

Die ISO Norm 9241/110 (International Organization for Standardization 2006) behandelt die ergonomische Gestaltung von interaktiven Systemen und sie formuliert 7 Prinzipien für die Beschreibung, das Design und die Evaluation der Dialoggestaltung von Software: Aufgabenangemessenheit, Selbstbeschreibungsfähigkeit, Steuerbarkeit, Erwartungskonformität, Fehlerrobustheit, Individualisierbarkeit, Lernförderlichkeit. Sie ersetzte 1996 die ISO Norm 9241-10 (International Organization for Standardization 1996), die sich inhaltlich stark auf Software im Bürokontext fokussiert hatte. Die Prinzipien sind als voneinander abhängig zu verstehen, und manchmal wird es notwendig sein, Vorteile eines Grundsatzes gegenüber einem anderen abzuwägen. Die Auswahl dieser 7 Grundprinzipien entstand auf Grund faktorenanalytischer Analysen von Fragebogendaten, in denen Benutzer die Relevanz von Systemeigenschaften für Qualität von interaktiven Systemen einschätzten (Dzida et al. 1978, 180-195). Der Hintergedanke war, dass die von Benutzern erlebte Qualität interaktiver Software aus statistisch nicht überlappenden Kategorien, d.h. aus Faktoren besteht, da Benutzer sie in unterschiedlicher Art erleben.

## 3 Usability-Fragebögen

Viele Aspekte der Usability sind am besten zu erheben, indem man Benutzer befragt (Nielsen 1997, 110). Usability-Frageböge sind neben mündlichen Befragungen in Form von Interviews zu den meinungsbasierten Evaluationsmethoden zu zählen. Zur Beurteilung von Fragebögen können Gütekriterien wie Objektivität, Reliabilität und Validität herangezogen werden. In den folgenden Abschnitten werden die deutschsprachigen Usability-Fragebögen Isonorm 9241/10 und Isometrics vorgestellt, die sich beide an der Isonorm 9241/10 orientieren. Es existieren zahlreiche weitere Usability Fragebögen, auf die hier nur kurz eingegangen werden kann. ErgoNorm (Dzida et al. 2000) ist ein weiterer deutscher Fragebogen, in dem die 7 Gestaltungsgrundsätze von der Isonorm 9241/10 operationalisiert wurden. Im Gegensatz zu Isometrics und Isonorm 9241/10 werden in diesem Fragebogen die Gestaltungsgrundsätze den Probanden jeweils kurz erklärt. Außerdem ist das Antwortformat dichotom – Fragen werden mit „Ja“ und „Nein“ beantwortet, und Nein-Antworten sollen schriftlich erklärt werden. Auf weitere Fragebögen sei hingewiesen, z.B. auf EU-CON II (Stary et al. 1997), den Questionnaire for User Interface Satisfaction - QUIS (Chin et al. 1988), das Software Usability Measurement Inventory - SUMI (Kirakowski & Corbett 1993) oder auf den Fragebogen zur Bewertung von Software von IfADo (Institut für Arbeitsphysiologie (IfADo) 1995).

### 3.1 Isonorm 9241/10

Der Fragebogen Isonorm 9241/10 wurde 1993 veröffentlicht (Prümper & Anft 1993). Die Skalen des Fragebogens stellen eine Operationalisierung der 7 Kriterien der gleichnamigen Isonorm dar. Der Fragebogen wurde entwickelt, um die Normkonformität nach Isonorm 9241/10 überprüfen oder um Prototypen bei iterativer Softwareentwicklung evaluieren zu können (Prümper 1997, 253). Der Fragebogen ist für das Ausfüllen durch Benutzer vorgesehen, und die Durchführung dauert ca. 10 Minuten. Für den Isonorm Fragebogen liegen allgemeine Normwerte für Softwaresysteme vor, die auf der Beurteilung von 41 Programmen basieren (Prümper 1997). Insgesamt besitzt der Fragebogen 35 Items, wobei zu jeder der 7 Skalen fünf Items gehören, die aus bipolaren Aussagen bestehen. Das Antwortformat (siehe Abbildung 1) stellt eine 7-stufigen Ratingskala dar, mit der von sehr negativ ("---") bis sehr positiv ("+++") die bipolaren Aussagen beurteilt werden.

#### Aufgabenangemessenheit

Unterstützt die Software die Erledigung Ihrer Arbeitsaufgaben, ohne Sie als Benutzer unnötig zu belasten?

Die Software...	---	--	-	~/+	+	++	+++	
ist kompliziert zu bedienen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ist unkompliziert zu bedienen.

Abbildung 1: Beispielitem aus dem Fragebogen Isonorm 9241/10

Bezüglich der Testgütekriterien liegen bereits Untersuchungen mit 1264 Benutzer und 178 Softwareprogrammen vor (Prümper 1997, 253). Hinsichtlich der Reliabilitätsbestimmung zeigte die Wiederholungsmethode, dass der Isonorm 9241/10 Fragebogen eine hohe Retest-Reliabilität aufweist (von  $r_{t1/t2} = 0.59$  bei der Skala Lernförderlichkeit bis zu  $r_{t1/t2} = 0.68$  bei der Skala Fehlertoleranz). Die Konsistenzanalyse bescheinigte eine gute innere Konsistenz der einzelnen Skalen (Cronbach's alpha von 0.81 bis 0.89). Des Weiteren wurden die Konstruktvalidität sowie die innere und äußere kriterienbezogene Validität überprüft. Die Konstruktvalidität wurde überprüft, indem man Beurteilungen jeweils zweier Versionen von MS-Word für DOS und MS-Excel für Windows verglich. Es zeigte sich, dass anhand des Fragebogens Isonorm 9241/10 Veränderungen zwischen verschiedenen Versionen einer Software gemessen werden können. In den Untersuchungen zur inneren kriterienbezogenen Validität wurden Bewertungen der Fragebögen Isonorm 9241/10, QUIS (Chin et al. 1988) und BBD (Spinas 1987) korreliert, die zeigten, dass der Fragebogen Isonorm 9241/10 zu vergleichbaren Ergebnissen führte. Die äußere kriterienbezogene Validität wurde mit Hilfe von Expertenbefragungen nach EVADIS 2 bestätigt. Für den Fragebogen wurde auch eine Normierung anhand von 42 Softwareprogrammen erstellt, wobei die Skala Individualisierbarkeit die niedrigste, die Skala Steuerbarkeit die höchste Beurteilung erhielt. Der Fragebogen ist neben der Papierversion auch als Onlineversion erhältlich, welche dieselben Ergebnisse liefert (Richter 1999).

## 3.2 Isometrics

In dem Fragebogen Isometrics sind ebenfalls die 7 Gestaltungsgrundsätze von der Isonorm 9241/10 operationalisiert (Hamborg 2002). Die Gestaltungsgrundsätze werden mit jeweils 7 bis 12 Items erhoben. Insgesamt besteht der Fragebogen aus 75 Items, die auf einer 5-stufigen Ratingskala beantwortet werden. Es besteht auch die Möglichkeit „Keine Angabe“ anzukreuzen, wie in Abbildung 2 gesehen werden kann. Es existiert eine Kurzversion des Fragebogen Isometrics<sup>s</sup> (short) für summative und eine Langversion Isometrics<sup>L</sup> (long) für formative Evaluationen. Diese beiden Versionen unterscheiden sich darin, dass in der Langversion zusätzlich zu jedem Item ein Wichtigkeitsrating gegeben werden kann. Außerdem sollen die Probanden in der Langform ihre Antworten begründen, indem sie Beispiele für Mängel notieren.

		stimme nicht	stimme wenig	stimme müßelmaßig	stimme ziemlich	stimme sehr	
	<b>Aufgabenangemessenheit</b>						
11	Die Software zwingt mich, überflüssige Arbeitsschritte durchzuführen.	1	2	3	4	5	Keine Angabe

Abbildung 2: Beispielitem aus dem Fragebogen Isometrics

Die Entwicklung des Fragebogens zeichnete sich dadurch aus, dass im Sinne von inhaltlicher Validität versucht wurde, die Items möglichst theoretisch fundiert auszuwählen (Gediga et al. 1999; Hamborg et al. 1996). Zuerst wurden 651 Items aus anderen Usability-Fragebögen und Checklisten gesammelt. Dann wurden redundante oder zu kompliziert formulierte Items entfernt und von sechs Usability-Experten den 7 Grundsätzen der Isonorm 9241/10 zugeordnet (Interrater-Reliabilität: Kappa = 0.75). Danach wurden die Itemformulierungen hinsichtlich Verständlichkeit und die Subskalen nach testtheoretischen Grundlagen optimiert. Die Reliabilität der Skalen des Isometrics Verfahrens sind für verschiedene Softwaresysteme vergleichbar (Hamborg et al. 1996; Willumeit et al. 1995, 8). Dies wurde anhand der Evaluationen mehrerer Programme (z.B. LaTeX, MS Word) erhoben. Die Ergebnisse wurden hinsichtlich der statistischen Qualität der Skalen als zumindest zufrieden stellend beschrieben (Gediga & Hamborg 1999). Für eine Validierung wurden die Skalenmittelwerte verschiedener Programme miteinander verglichen, und die Bewertungsunterschiede entsprachen den Erwartungen; z.B. wurde die Selbstbeschreibungsfähigkeit bei einem Programm ohne graphischer Benutzeroberfläche niedriger als bei anderen beurteilt. Zum Isometrics Verfahren liegen Normtabellen zu mehreren bekannten Softwareprogrammen vor, z.B. LaTeX, MS Word, WinWord, SAP/Term. Hinsichtlich der Effektivität von Isometrics<sup>L</sup> im Vergleich mit den Evaluationsmethoden Lautes Denken, Videokonfrontation und Heuristische Evaluation zeigt sich, dass in Isometrics<sup>L</sup> 3-5 mal so viele Anmerkungen über Probleme der Software als bei den anderen gemacht wurden (Hamborg 2002, 310). In einer anderen Studie wurde der Einsatz von Isometrics als Gruppen- und Einzelversion für formative und summative Evaluation verglichen (Hamborg et al. 1999, 14). Es zeigte sich, dass bei der summativen Evaluation die Bewertungsergebnisse in der Gruppenversion weniger streuen und strenger sind als in der Einzelversion. Bei der formativen Evaluation lieferte die Einzelversion mehr Informationen bzw. Anmerkungen als die Gruppenversion.

## 4 Empirische Studie

Hinsichtlich des Vergleichs der Usability-Fragebögen Isometrics und Isonorm sind zentrale Fragen, ob Software in beiden Fragebögen in ähnlicher Weise beurteilt wird, d.h. die beiden Fragebögen vergleichbare Evaluationsresultate liefern und ob Benutzer das Ausfüllen der Fragebögen unterschiedlich erleben. Weiters ist es von Interesse zu sehen, ob die einzelnen Skalen gute Reliabilitäten aufweisen, und ob die Skalen voneinander unabhängig sind oder starke Interkorrelationen aufweisen.

### 4.3 Versuchsdesign

Für die Untersuchung wurde ein experimentelles Design gewählt; Versuchspersonen sollten ein neu in einem Usability-Test kennengelerntes Softwareprodukt anhand der beiden Fragebögen Isonorm und Isometrics beurteilen. Um möglichst allgemeine Aussagen zum Vergleich der Fragebögen treffen zu können, wurden zwei inhaltlich ähnliche Softwareprodukte gewählt: SPSS und STATISTICA von der Firma Statsoft, die sich beide auf die Analyse sozialwissenschaftlicher Daten konzentrieren. Die Wahl von Statistik-Programmen schien insbesondere deshalb sinnvoll, weil sich ein Großteil der Käufer vor allem an den Kriterien Usability und Bearbeitungsgeschwindigkeit orientiert (McCullough, 2000, p. 350). In einem Zwei-Versuchsgruppendedesign lernten Teilnehmer der Versuchsgruppe 1 das Statistikprogramm SPSS (Version 12.0), Teilnehmer der Versuchsgruppe 2 das Statistikprogramm STATISTICA (Version 7.1) kennen und beurteilten im Anschluss die Usability dieses Programms anhand der Fragebögen Isonorm und Isometrics (siehe Abbildung 3). Die Versuchsgruppen waren weiters in Subversuchsgruppen je nach Reihenfolge der Vorlage der Fragebögen Isonorm und Isometrics eingeteilt.

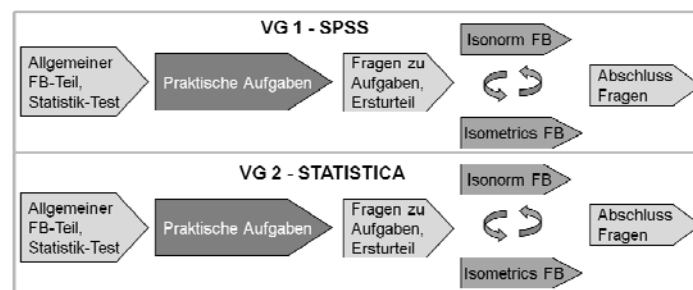


Abbildung 3: Überblick über das Versuchsdesign

### 4.4 Usability-Test

Der Usability-Test wurde auf zwei Notebooks durchgeführt um eine einheitliche Umgebung zu garantieren. Anhand von 14 Aufgaben mit derselben Datendatei mit 10 Datensätzen und je 10 Variablen lernten die Versuchspersonen die beiden Programme SPSS bzw. STATISTICA kennen. Es wurde darauf geachtet, dass die Versuchspersonen ein möglichst umfas-

sendes Bild der Software erlangen und sowohl Datenmanipulation, deskriptive Statistiken, statistische Analysemodule, Graphikerstellung und die Hilfe kennen lernen konnten. Auch wurden absichtlich drei Eingabefehler in der Angabe eingebaut, damit die Versuchspersonen die Fehlertoleranz des Programms testen und Fehlermeldungen kennen lernen konnten. Die Aufgaben wurden in drei verschiedenen Schwierigkeitsgraden gestellt. Am Anfang gab es Aufgaben mit vollständiger Anleitung, in der jeder Mausklick beschrieben wurde. Dann gab es Aufgaben, in denen zuvor erklärte Verfahren bzw. Graphiken selbstständig ohne Anleitung mit anderen Variablen ausgeführt werden sollten. Abschließend sollten völlig neue Prozeduren ohne Erklärung selbstexplorierend gelöst werden. Der Usability-Test wurde mit einem Fragebogen begleitet, der allgemeine Fragen, einen kurzen Statistik-Test, Anleitungen für den praktischen Teil in SPSS bzw. STATISTICA, eine allgemeine Einschätzung des Programms, den Usability-Fragebögen Isonorm und Isometrics sowie Fragen zum Vergleich dieser beiden Usability-Fragebögen inkludierte. Die Fragebögen Isonorm und Isometrics wurden im Original übernommen, mit der Ausnahme, dass die Antwortmöglichkeit „Keine Angabe“ beim Isometrics Fragebogen weggelassen wurde, um Vergleichbarkeit mit dem Isonorm Fragebogen zu gewährleisten. Die Skala Individualisierbarkeit, für deren Beantwortung Langzeiterfahrung mit der Software nötig wäre, wurde bei beiden Fragebögen weggelassen. Ergänzt wurden alle Items der beiden Fragebögen um jeweils ein zusätzliches Item, in dem auch der eigene Kenntnisstand zum Beurteilen des Items auf einer 5-stufigen Skala eingeschätzt werden sollte.

## 4.5 Stichprobe

Die Stichprobe bestand aus 50 Versuchspersonen, davon 50% weiblich bzw. männlich. Die Personen waren von 21 bis 34 Jahre alt und hatten zumindest Matura/Abitur als höchste abgeschlossene Schulbildung.

## 4.6 Ergebnisse des Usability-Tests

Von den praktischen Aufgaben, die dazu dienten, die Statistikprogramme kennen zu lernen, wurden durchschnittlich 15.8 der insgesamt 17 Aufgaben sowohl in der SPSS- als auch in der STATISTICA-Versuchsbedingung gelöst. Für die praktischen Beispiele haben die Versuchspersonen im Durchschnitt 26 Minuten verwendet, für das Ausfüllen der Fragebögen benötigten die meisten Versuchspersonen ca. eine halbe Stunde.

Zur Prüfung ob die beide Programme unterschiedlich in den Fragebögen Isonorm und Isometrics beurteilt wurden zweifaktorielle multivariate Varianzanalyse mit Messwiederholung ( $df_{\text{Hypothese}} = 6$ ,  $df_{\text{Fehler}} = 39$ ) gerechnet, mit den Fragebögen als Innersubjektfaktor und der Versuchsbedingung der unterschiedlichen Statistikprogramme als Zwischensubjektfaktor. Der Innersubjektfaktor „Fragebogen“ (Isonorm vs. Isometrics) hatte keinen Einfluss auf diese Beurteilung der Usability ( $F=1.21$ ,  $p=0.32$ ,  $\eta^2=0.16$ ). Das bedeutet, die beiden Fragebögen liefern dieselben Ergebnisse. Es zeigt sich weiters, dass der Zwischensubjektfaktor „Statistikprogramm“ (SPSS vs. STATISTICA) einen signifikanten Einfluss auf die Bewertung der Usability hat ( $F=3.40$ ,  $p=0.009$ ,  $\eta^2=0.34$ ), d.h. eines der beiden Softwareprodukte – SPSS - wurde unterschiedlich bewertet (siehe Abbildung 4 und Abbildung 5). In der Dimen-

sion Steuerbarkeit erreichte SPSS in beiden Fragebögen bessere Werte, in der Dimension Selbstbeschreibungsfähigkeit erhielt es nur im Fragebogen Isometrics eine bessere Beurteilung. Dies bedeutet, dass die Menüführung in SPSS einfacher als in STATISTICA ist, da es einfacher ist, Befehle abzubrechen, und dass man sich leichter in Dokumenten bewegen und zwischen Menüs wechseln kann.

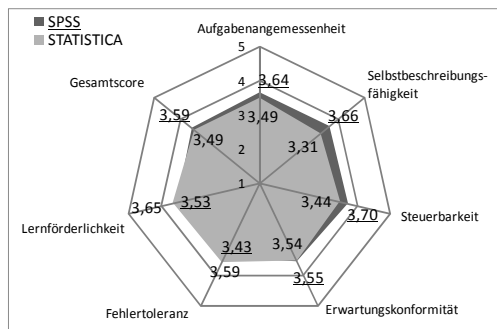


Abbildung 4: Beurteilung der Usability im Isometrics-Fragebogen (Mittelwerte)

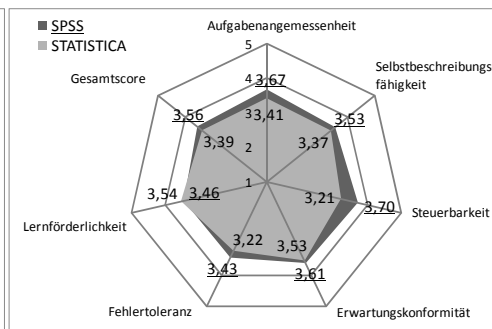


Abbildung 5: Beurteilung der Usability im Isonorm-Fragebogen (Mittelwerte)

## 4.7 Vergleich der Fragebögen Isometrics und Isonorm

### 4.7.1 Reliabilitäts- und Itemanalyse

Reliabilitätsanalysen zur Überprüfung, ob die Items einer Skala dieselbe Dimension messen, wurden für die Skalen der Fragebögen Isonorm und Isometrics jeweils getrennt für die Beurteilung der beiden Softwareprodukte gerechnet. Die Reliabilitäten können als durchwegs zufriedenstellend bewertet werden (Cronbach's Alpha von 0.54 - 0.91 bei allen eingesetzten Skalen). Zusätzliche Kennwerte wie das Cronbach's Alpha, für den Fall, dass ein Item weggelassen würde, zeigten, dass durch Ausschluss der Items „Das System lässt sich nur in einer starr vorgegebenen Weise bedienen“, „Die Software erschwert meine Aufgabenbearbeitung durch eine uneinheitliche Gestaltung“, „Bei der Arbeit mit der Software kann es passieren, dass auch kleine Fehler schwerwiegende Folgen nach sich ziehen“ des Fragebogens Isometrics die Reliabilität der dazugehörigen Skalen etwas steigen würde. Dieses Ergebnis ist wahrscheinlich dadurch begründet, dass diese Items im Vergleich zu den anderen negativ formuliert waren.

### 4.7.2 Zusammenhänge der Skalen untereinander

Um herauszufinden, ob die Skalen der Fragebögen zwischen den Skalen Zusammenhänge bestehen. Da diese Frage unabhängig von dem beurteilten Programm von Interesse ist, wurden partielle Korrelationen mit „SPSS\_STATISTICA“ als Kontrollvariable gerechnet. Dabei ergab sich, dass im Fragebogen Isonorm alle Interkorrelationen der 5 verwendeten Skalen stark und signifikant sind ( $r=0.51-0.77$ ,  $df=45$ ). Hinsichtlich des Fragebogens Isometrics ergab sich ein ähnliches Bild: beinahe alle Korrelationen sind stark und signifikant, ( $r=0.34-$

0.73,  $df=45$ ). Allerdings bildet hier die Skala Fehlertoleranz eine Ausnahme, da deren Beurteilung keine bis niedrige Zusammenhänge mit den Bewertungen der anderen Skalen zeigt ( $r=0.09-0.34$ ). Eine Interpretation der hohen Interkorrelationen kann in der Weise gegeben werden, dass die Beurteilung der einzelnen Aspekte der Usability stark von einer Gesamteinschätzung abhängt, da angenommen werden kann, dass Versuchspersonen, die ein Programm grundsätzlich als gebrauchstauglich einschätzen, es in allen Dimensionen gut beurteilen. Dieser Effekt könnte auch durch das Versuchsdesign verstärkt worden sein, in dem die Versuchspersonen die Programme neu kennen zu lernen hatten. Es könnte durchaus sein, dass erfahrene Benutzer ein System differenzierter in Bezug auf die einzelnen Dimensionen beurteilen würden.

#### 4.7.3 Zusammenhang zwischen den Beurteilungen in den beiden Fragebögen

Hinsichtlich des Zusammenhangs der Einschätzungen in den beiden Fragebögen, wurden Korrelationen für jede einzelne der 6 Skalen sowie für den Gesamtscore berechnet. Wie Tabelle 1 zeigt, sind sämtliche Korrelationen positiv, die meisten sogar mittel bis hoch. Versuchspersonen schätzen somit die Dimensionen der Usability in beiden Fragebögen in ähnlicher Weise ein. Auch die partielle Korrelation mit der Kontrollvariable „SPSS\_STATISTICA“ erbringt dieselben Ergebnisse.

Tabelle 1: Pearson Korrelationen zwischen den Beurteilungen im Isonorm und Isometrics Fragebogen

Dimension	Gesamt	p	Partielle Korrelation (SPSS_STATISTICA als Kontrollvariable)	p
Aufgabenangemessenheit	<b>0.45***</b>	<b>0.001</b>	<b>0.44**</b>	<b>0.002</b>
Selbstbeschreibungsfähigkeit	<b>0.51***</b>	<b>0.000</b>	<b>0.50***</b>	<b>0.000</b>
Steuerbarkeit	<b>0.37**</b>	<b>0.010</b>	<b>0.32**</b>	<b>0.028</b>
Erwartungskonformität	<b>0.29**</b>	<b>0.048</b>	<b>0.29<sup>+</sup></b>	<b>0.051</b>
Fehlertoleranz	<b>0.24<sup>+</sup></b>	<b>0.098</b>	<b>0.28<sup>+</sup></b>	<b>0.056</b>
Lernförderlichkeit	<b>0.66***</b>	<b>0.000</b>	<b>0.66***</b>	<b>0.000</b>

#### 4.7.4 Fragebogenbeurteilung der Versuchspersonen

Die Versuchspersonen beurteilten die beiden Fragebögen anhand von drei abschließenden, vergleichenden Fragen, deren Ergebnisse in Abbildung 6 zu sehen sind.

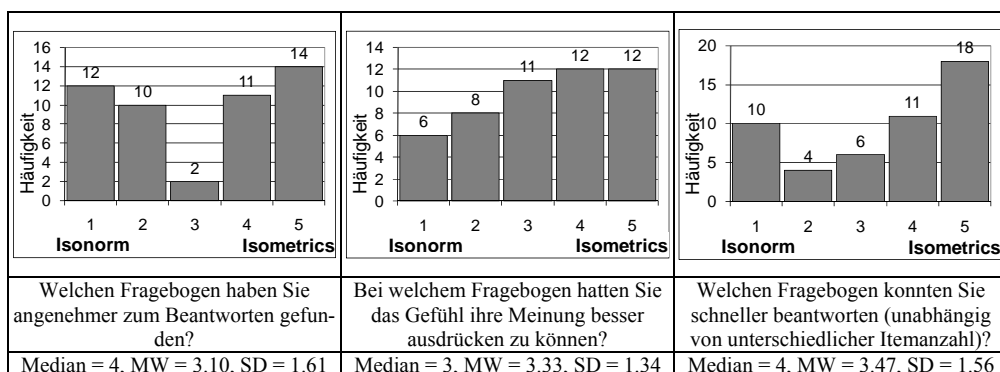


Abbildung 6: Beurteilung der Fragebögen Isonorm und Isometrics (N = 49)



Bezüglich der Fragen, ob die Versuchspersonen einen Fragebogen angenehmer bzw. schneller zum Beantworten gefunden haben, scheint es, dass wahrscheinlich auf Grund persönlicher Vorlieben für Antwortformate je ein Teil entweder Isonorm oder Isometrics bevorzugt hat. Bei der Frage, ob es in einem Fragebogen einfacher sei, seine Meinung besser auszudrücken als im anderen, hat es auch einige (11) indifferente Antworten gegeben. Die Frage, ob nun einer der beiden Fragebögen bei den Versuchspersonen „beliebter“ war, kann nicht eindeutig beantwortet werden. Die Mittelwerte zeigen eine leichte Präferenz für den Fragebogen Isonorm, die Antwortverteilung lässt aber auch erkennen, dass es Versuchspersonen gegeben hat, denen der Isometrics Fragebogen mehr „liegt“.

## 5 Diskussion

Diese Studie widmete sich dem Vergleich zweier weitverbreiteten, deutschsprachigen Usability Fragebögen – Isonorm und Isometrics, die beide die Isonorm 9241/110 operationalisieren. In dem eingesetzten Untersuchungsdesign gab es zwei Versuchsgruppen, wobei die eine das Statistik-Softwarepaket SPSS, die andere STATISTICA ca. 25 Minuten lang anhand von praktischen Aufgaben kennen lernte. Danach beurteilten die Versuchspersonen das jeweilige Programm anhand der Usability Fragebögen Isonorm und Isometrics. Es wurden sechs der insgesamt sieben verfügbaren Skalen eingesetzt: Aufgabenangemessenheit, Selbstbeschreibungsfähigkeit, Steuerbarkeit, Erwartungskonformität, Fehlertoleranz und Lernförderlichkeit. Bezüglich beider Fragebögen konnten zufriedenstellende Reliabilitäten sowie hohe Interkorrelationen der Skalen festgestellt werden. Die Bewertungen im Isonorm- und Isometrics-Fragebogen korrelierten positiv miteinander. Die Fragebogenbeurteilung deutet darauf hin, dass es sowohl Versuchspersonen gibt, die den Isonorm-Fragebogen präferieren, wie solche, die den Isometrics Fragebogen angenehmer und schneller zu beantworten finden. Zusammenfassend kann gesagt werden, dass die beiden verwendeten Fragebögen sehr gut für Usability-Untersuchungen geeignet sind, und dass sie grundsätzlich dieselben Ergebnisse für Bewertung und Vergleich von Softwareprodukten liefern.

## 6 Literaturverzeichnis

- Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of a Tool Measuring User Satisfaction of the Human-Computer Interface. SIGCHI '88 New York: ACM/SIGCHI.
- Dzida, W., Herda, S., & Itzfeldt, D. W. (1978). User-perceived quality of interactive systems. International Conference on Software Engineering, Atlanta.
- Dzida, W., Hofmann, B., Freitag, R. et al. (2000). Gebrauchstauglichkeit von Software. Ergonorm: Ein Verfahren zur Konformitätsprüfung von Software auf der Grundlage von DIN EN ISO 9241 Teile 10 und 11. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.
- Gediga, G., & Hamborg, K.-C. (1999). IsoMetrics: An usability inventory supporting summative and formative evaluation of software systems. HCI International '99, New Jersey: Lawrence Erlbaum Associates.

- Gediga, G., Hamborg, K.-C., & Düntsch, I. (1999). The IsoMetrics Usability Inventory: An operationalisation of ISO 9241-10 supporting summative and formative evaluation of software systems. *Behaviour and Information Technology*, 18, 151-164.
- Hamborg, K.-C. (2002). Gestaltungsunterstützende Evaluation von Software: Zur Effektivität und Effizienz des IsoMetricsL Verfahrens. In M. Herzeg, W. Prinz & H. Oberquelle (Hrsgs.), *Mensch & Computer 2002: Vom interaktiven Werkzeug zu kooperativen Arbeits- und Lernwelten* (S. 303-312). Stuttgart: Teubner.
- Hamborg, K.-C., Gediga, G., Döhl, M. et al. (1999). Softwareevaluation in Gruppen oder Einzelevaluation: Sehen zwei Augen mehr als vier? In U. Arend & K. Pitschke (Hrsgs.), *Softwareergonomie '99: Design von Informationswelten* (S. 97-109). Stuttgart: Teubner.
- Hamborg, K.-C., Willumeit, H., & Gediga, G. (1996). Untersuchungen zu Itemformulierungen des IsoMetrics-Verfahrens, *Osnabrücker Schriftenreihe Software-Ergonomie - 1*. Osnabrück: Universität Osnabrück.
- Institut für Arbeitsphysiologie (IfADo) (1995). Fragebogen zur Bewertung von Software: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA).
- International Organization for Standardization (1996). ISO 9241-10:1996- Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 10: Dialogue principles.
- International Organization for Standardization (1998). ISO 9241-11:1998 - Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability.
- International Organization for Standardization (2006). ISO 9241-110:2006 Ergonomics of human-system interaction -- Part 110: Dialogue principles.
- Kirakowski, J., & Corbett, M. (1993). SUMI: The Software Usability Measurement Inventory. *British Journal of Educational Technology*, 24(3), 210-212.
- Nielsen, J. (1997). Let's Ask the Users. *IEEE Software*, 14(3), 110-111.
- Prümper, J. (1997). Der Benutzungsfragebogen ISONORM 9241/10: Ergebnisse zur Reliabilität und Validität. *Software-Ergonomie '97*, Stuttgart: Reports of the German Chapter of the ACM.
- Prümper, J., & Anft, J. (1993). Die Evaluation von Software auf Grundlage des Entwurfs zur internationalen Ergonomie-Norm ISO 9241 Teil 10 als Beitrag zur partizipativen Systemgestaltung - ein Fallbeispiel. *Software-Ergonomie '93*, Stuttgart: Teubner.
- Richter, M. (1999). Online Befragung als neues Instrument zur Beurteilung der Benutzerfreundlichkeit von Software. In E.-D. Reips, B. Batanic, W. Bandilla, et al. (Hrsgs.), *Current Internet science - trends, techniques, results. Aktuelle Online Forschung - Trends, Techniken, Ergebnisse*. Zürich: Online Press.
- Spinas, P. (1987). *Arbeitspsychologische Aspekte der Benutzerfreundlichkeit von Bildschirmsystemen*. Zürich: ADAG.
- Sary, C., Riesenecker-Caba, T., Kalkhofer, M., & Flecker, J. (1997). EU-CON. ein Verfahren zur EU-konformen Software-ergonomischen Bewertung und Gestaltung von Bildschirmarbeit. ETH.
- Willumeit, H., Gediga, G., & Hamborg, K.-C. (1995). *Validation of the IsoMetrics usability inventory*: Universität Osnabrück.