

Die Priorisierung von Problemhinweisen in der software-ergonomischen Qualitätssicherung

Marc Hassenzahl

Technische Hochschule,
Darmstadt

Jochen Prümper

Fachhochschule für Technik
und Wirtschaft, Berlin

Uta Sailer

Ludwig-Maximilians-
Universität, München

Zusammenfassung

Im Rahmen software-ergonomischer Qualitätssicherung sind die Ressourcen für Gestaltung beschränkt. Um die effiziente Bearbeitung zu gewährleisten, wird die Priorisierung von Problembereichen eines Softwaresystems nötig. Es werden zwei Arten der Priorisierung von Problemhinweisen, nämlich *Priorisierung nach Fehlerkosten* und *Priorisierung nach Auftretensstabilität*, vorgestellt und empirisch analysiert. Hierbei steht der Aspekt der mentalen Beanspruchung im Zentrum der Überlegungen. Prinzipiell eignen sich beide Methoden zur Priorisierung und Auswahl der wichtigsten Problembereiche eines Softwaresystems. Ob man sich für die eine oder andere Methode entscheiden soll, hängt jedoch von verschiedenen Randbedingungen, wie Auftretenshäufigkeit und zeit-ökonomischen Überlegungen ab. In jedem Fall kann eine der hier vorgeschlagenen Priorisierungsmethoden den Prozeß der software-ergonomischen Qualitätssicherung effizienter gestalten und damit im Ergebnis die mentale Beanspruchung der Benutzer reduzieren.

1 Einleitung

Unserer Erfahrung nach ist es nicht schwer, im Laufe eines Softwareentwicklungsprozesses von Projektverantwortlichen, Endanwendern oder Entwicklern eine Erklärung zu bekommen, welche Mängel im Rahmen software-ergonomischer Qualitätssicherung am dringlichsten bearbeitet werden müssen. Leider stimmen diese Prioritäten in den seltensten Fällen miteinander überein: Entwickler denken stark in technischen Kategorien von Funktionalität und Realisierung, Benutzer in Kategorien der Gebrauchstauglichkeit und Projektverantwortliche nicht zuletzt in Kategorien der Verträglichkeit der Gestaltungsmaßnahmen mit dem Budget.

Damit stellt sich die Frage, wie Problembereiche einer Software priorisiert werden können, so daß sie auf der einen Seite möglichst unabhängig von den unterschiedlichen Wünschen und Sichtweisen einzelner Gruppen sind, es auf der anderen Seite aber zu einem Konsens der beteiligten Gruppen über die Rangfolge der Bearbeitung kommt. Dies ist besonders dann wichtig, wenn Benutzerpartizipation das Mittel der Erarbeitung von Lösungsvorschlägen für Problembereiche im Rahmen einer formativen Evaluation [5,15] ist.

Egal welche Art und Weise der Partizipation man bevorzugt, immer wird es eine Beschränkung der Ressourcen geben. Eine bedeutende Rolle spielen dabei Einschränkungen, die durch den - häufig sehr engen - zeitlichen und finanziellen Rahmen, in dem die Gestaltung beendet sein soll und durch die beschränkte Verfügbarkeit von beteiligten Benutzern entstehen [9].

Der erste Schritt zur Auswahl besonders schwerwiegender Problemursachen ist das Aufstellen einer Rangfolge von Problemhinweisen. Dies hat den Vorteil, daß ein beliebiger Prozentsatz der Problembereiche eines Softwaresystems systematisch abgearbeitet, d.h. umgestaltet werden kann. Dieser Prozentsatz muß zwar an die jeweiligen Ressourcen angepaßt werden; man ist sich aber immer sicher, wichtige Probleme zuerst bearbeitet zu haben.

Oft spielt bei der Auswahl der zu bearbeitenden Probleme die Antizipation der durch die Umgestaltung entstehenden Kosten eine große Rolle. Es werden also die Probleme ausgewählt, von denen man annimmt, daß sie leicht und damit kostengünstig zu beheben sind. In dieser Strategie spiegelt sich der Trugschluß wieder, daß die Art des Problems die Umgestaltungs-kosten festlegt. Dem ist nicht so. Vielmehr ist es der konkrete Gestaltungsvorschlag des software-ergonomischen Experten, der die Umgestaltungs-kosten bestimmt. Es liegt also in der Verantwortung des Gestalters, eine finanziell vertretbare, kreative und ggf. unvollständige Lösung eines Problems zu erarbeiten, anstatt das ganze Problem fallen zu lassen.

Was aber sind die wichtigsten Probleme?

Zwei Möglichkeiten der Priorisierung sollen im folgenden diskutiert und empirisch untersucht werden: *Priorisierung nach Fehlerkosten* und *Priorisierung nach Auftretensstabilität*.

2 Die Priorisierung von Problemhinweisen

Bevor diese beiden Methoden der Priorisierung vorgestellt werden, soll zunächst das Konzept der „Fehlereinheit“ als Gegenstand der Priorisierung eingeführt werden, da die in der Fehlereinheit enthaltenen Daten die empirische Grundlage der Priorisierung bilden.

2.1 „Fehlereinheiten“ als Gegenstand der Priorisierung

Personen haben oft ganz unterschiedliche Probleme im Umgang mit Softwaresystemen. Zur systematischen Erfassung derartiger Probleme hat sich in den letzten Jahren die „handlungsorientierte Fehlertaxonomie“ [16] bewährt, da mit ihr Fehlerereignisse systematisch klassifiziert und verbessert werden können [6]. Hier werden *Fehlerereignisse*, die in Folge eines „Mismatches“, d.h. einer Nicht-Passung von *Mensch* und *Computersystem* entstehen, als Nutzungsprobleme bezeichnet und nach Schritten im Handlungsprozeß und nach Ebenen der Handlungsregulation aufgeschlüsselt.

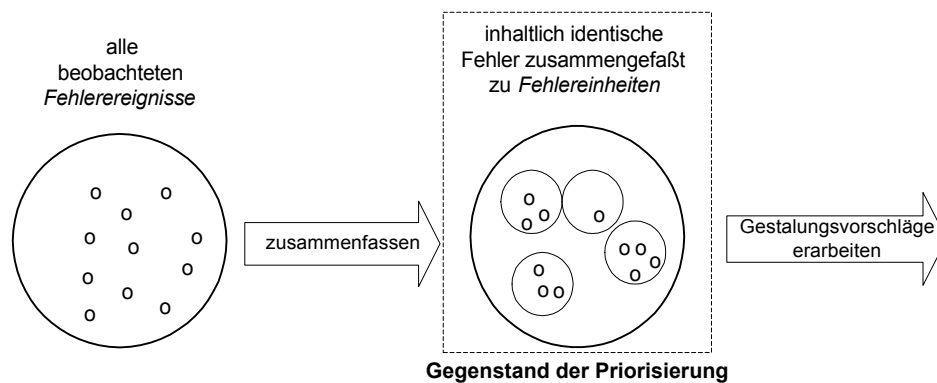


Abb. 1: Zusammenhang zwischen Fehlerereignissen und Fehlereinheiten

Nun sind viele Probleme im Umgang mit einem Softwaresystem keine singulären Ereignisse. Dies bedeutet, daß inhaltlich identische Probleme zusammengefaßt werden können. Diese zusammengefaßten Fehlerereignisse können als *Fehlereinheiten* bezeichnet werden. Abbildung 1 zeigt schematisch dieses Vorgehen. Beobachtete, inhaltlich identische *Fehlerereignisse* werden auf der theoretischen Grundlage der handlungsorientierten Fehlertaxonomie [16] zu einer *Fehlereinheit* zusammengefaßt. Die Fehlertaxonomie mit ihren *Fehlerkategorien* liefert den Fehlerbegriff und ermöglicht durch die Zuordnung der *Fehlereinheit* zu einer Fehlerkategorie eine erste allgemeine Interpretation des zugrundeliegenden Problems. Gegenstand der Analyse und Gestaltung bleibt allerdings im ganzen Prozeß die *Fehlereinheit*.

Eine Fehlereinheit besteht aus einer *inhaltlichen Beschreibung*, der Zuordnung zu einer *Fehlerkategorie*, der *mittlere Fehlerbewältigungszeit* und der *Auftretenshäufigkeit* (vgl. Abb. 2). Die *inhaltliche Beschreibung* der Fehlereinheit hält den unmittelbaren Aufgabenzusammenhang und die Funktion des Systems fest, auf die sich der Fehler bezieht. Dies macht den Fehler für andere Personen nachvollziehbar und gibt Anhaltspunkte für konkrete Veränderungen.

Die Zuordnung der Fehlereinheit zu einer *Fehlerkategorie* gibt erste Hinweise, welche Prozesse am Zustandekommen der zugrundeliegenden Fehlerereignisse beteiligt sind und welche grundsätzlichen Gestaltungsmaßnahmen zu ergreifen sind.

Fehlerbewältigungszeit und *Auftretenshäufigkeit* sind Kennwerte, die Fehlereinheiten miteinander vergleichbar machen und eine Einschätzung der Konsequenzen einer Fehlereinheit erlauben. Die Fehlerbewältigungszeit ist die Zeit, die zur Korrektur eines Fehlers benötigt wird. Die Auftretenshäufigkeit ist die Anzahl der inhaltlich identischen Fehler, die zu einer Fehlereinheit gebündelt werden.

<i>Inhaltliche Beschreibung:</i>		
Der Benutzer möchte Antragsblätter aufrufen. Statt des Menüpunkts "Formantrag" wählt er "Haushalt".		
<i>Fehlerkategorie:</i>	<i>mittlere Fehlerbewältigungszeit:</i>	<i>Auftretenshäufigkeit:</i>
Wissensfehler	6 sec	2

Abb. 2: Bestandteile einer Fehlereinheit

2.2 Methoden der Priorisierung von Problemen

Nachdem *Fehlereinheiten* als Hinweise auf Problembereiche eines Softwaresystems vorliegen, stellt sich die Frage, wie diese Fehlereinheiten zur Bearbeitung priorisiert werden können. Im folgenden werden zwei Methoden vorgestellt: erstens nach *Fehlerkosten*, zweitens nach *Auftretensstabilität*.

2.2.1 Priorisierung nach Fehlerkosten

Zur Priorisierung kann man Fehlereinheiten im Hinblick auf ihre Kosten betrachten. Dabei lassen sich *ökonomische* Kosten eines Problems für ein Unternehmen (Institution) und *psychische* Kosten für den einzelnen Benutzer unterscheiden [16]. So wurde in einer Studie gefunden, daß 10% der gesamten Computerarbeitszeit auf Fehlerbewältigung entfällt. Diese Fehlerbewältigungszeit steht wiederum in signifikantem Zusammenhang zu negativen emotionalen Reaktionen [1,2]. Betrachtet man Computersysteme als Werkzeuge zur effizienten und effektiven Bearbeitung von Arbeitsaufgaben, ist dies bedenklich. Im Gegensatz zum Computer als Spielzeug muß der Computer als Werkzeug aus dem Aufmerksamkeitsbereich des Benutzers zurücktreten [8]. Er darf den Benutzer bei der Erledigung seiner Arbeitsaufgabe nicht behindern.

Die *Priorisierung nach Fehlerkosten* geht davon aus, daß die Probleme zuerst behoben werden müssen, die gleichzeitig hohe ökonomische und hohe psychische Kosten verursachen. Diese Fehlerkosten setzen sich zusammen aus der Häufigkeit des Auftretens des Fehlerereignisses und der Zeit, die benötigt wird, um das Fehlerereignis zu bewältigen. Mit anderen Worten: Fehlereinheiten kommt dann eine hohe Priorität zu, wenn sowohl die Auftretenshäufigkeit als auch die Bewältigungszeit hoch ist. Abbildung 3 veranschaulicht diesen Gedanken. Fehlereinheiten lassen sich in eine Rangfolge bringen, indem die Bewältigungszeiten aller Fehlerereignisse einer Fehlereinheit summiert werden.

Aufretenshäufigkeit	Bewältigungszeit	
	niedrig	hoch
hoch	mittlere Priorität	hohe Priorität
niedrig	niedrige Priorität	mittlere Priorität

Abb. 3: Priorität in Abhängigkeit von Auftretenshäufigkeit und Bewältigungszeit

2.2.2 Priorisierung nach Auftretensstabilität

In der handlungstheoretischen Fehlertaxonomie werden Nutzungsprobleme als „Mismatch“ definiert. Das bedeutet, daß eine eindeutige Schuldzuweisungen in Richtung Computer oder Benutzer nicht möglich ist [12]. Aus praktischen Erwägungen bietet es sich jedoch an, den einen oder anderen Pol dieses „Mismatches“ stärker zu berücksichtigen. Unterläuft mehreren Personen derselbe Fehler, dann ist es sinnvoll, die Software zu verändern; passiert es hingegen, daß eine einzelne Person Fehler macht, die anderen Personen nicht unterlaufen, dann sind entsprechende Qualifizierungsmaßnahmen sinnvoll [4].

Die Priorisierung nach Auftretensstabilität geht davon aus, daß „systembedingte“ Fehlereinheiten mit einer höheren Priorität bearbeitet werden sollen, da hier die Chance, durch Veränderungen des Systems ein verringertes Auftreten des Problems bewirken zu können, höher eingeschätzt werden muß.

3 Methode

3.1 Versuchspersonen

Im Rahmen einer Studie zur partizipativen Evaluation und Gestaltung einer Bürosoftware wurden 12 Personen, sieben Frauen und fünf Männer untersucht. Das durchschnittliche Alter der Teilnehmer betrug 34 Jahre. Sie waren im Mittel neun Jahre in ihrem Beruf tätig, davon im Mittel 5 Jahre und 1 Monat in dem Tätigkeitsbereich, den die Funktionalität des Softwaresystems berührt. Die Teilnehmer arbeiteten im Durchschnitt seit rund 86 Monaten mit Computern und absolvierten im Schnitt 5½ Trainingstage für die untersuchte Software. Die Frage „Wie gut beherrschen Sie die getestete Software?“ beantworteten die befragten Personen auf einer Skala von 1 (sehr schlecht) bis 7 (sehr gut) im Durchschnitt mit 2,8. Mit Computern arbeiteten die Personen im Mittel rund 17 Stunden pro Woche. Insgesamt arbeiteten die Befragten durchschnittlich mit 3,7 verschiedenen Softwareprogrammen. Die getestete Software war zum Zeitpunkt der Untersuchung noch nicht am Arbeitsplatz eingeführt.

3.2 Abhängige Variablen

Abhängige Variablen sind die *Aufgabenbearbeitungszeit*, die *Fehlereinheiten* und die *mentale Beanspruchung*. Die Aufgabenbearbeitungszeit muß registriert werden, da nur so der *Anteil der Fehlerbewältigungszeit an der Aufgabenbearbeitungszeit* bestimmt werden kann. Dies ist nötig, da individuelle Unterschiede in der Bearbeitungsgeschwindigkeit sich auch in längeren Fehlerbewältigungszeiten niederschlagen.

3.2.1 Aufgabenbearbeitungszeit

Die *Aufgabenbearbeitungszeit* ist die Zeit, die jeder Teilnehmer zur Erledigung einer Aufgabe benötigt. Eine maximale Aufgabenbearbeitungszeit wurde nicht vorgegeben. Vielmehr lag es in der Verantwortung der Teilnehmer, die Aufgabe dann zu beenden, wenn sie der Meinung waren, entweder das gewünschte Arbeitsziel erreicht zu haben oder aber keine Möglichkeit der Aufgabenlösung mehr zu sehen.

3.2.2 Fehlereinheiten

Fehlereinheiten lassen sich über Auftretenshäufigkeit und Fehlerbewältigungszeit charakterisieren. *Auftretenshäufigkeit* ist die Anzahl der inhaltlich identischen Fehlerereignisse, die zu einer Fehlereinheit gebündelt werden. *Fehlerbewältigungszeit* ist definiert als die Zeitspanne vom Beginn der fehlerhaften Aktion, bis zu dem Punkt, an dem eine neue Aktion ausgeführt wird. Diese Aktion kann dann entweder direkt der Fehlerbewältigung dienen, dann wird sie mitgemessen, oder aber wiederum ein neues Fehlerereignis sein, dann beginnt eine neue Messung.

3.2.3 Mentale Beanspruchung

Zur Annäherung an die psychische Belastung, im Sinne von subjektiv aversiv erlebter Beanspruchung [13], kam die BSMA-Skala [3,17] zum Einsatz. Diese eindimensionale Skala ermöglicht die Einschätzung der mentalen Beanspruchung von 0 bis 220. Neben den Zahlen der Skala erleichtern verbale Anker wie „kaum anstrengend“ bei 20 oder „außerordentlich anstrengend“ bei 204 die Einschätzung. Ein Vorteil dieses Verfahrens ist seine leichte Anwendbarkeit und schnelle Beantwortbarkeit. Dies ermöglicht, Beanspruchungsmessungen mehrmals während einer Testphase durchzuführen.

3.3 Versuchsablauf

Jeder Teilnehmer hatte in Einzelsitzungen nacheinander jeweils fünf Testaufgaben unterschiedlicher Komplexität in einer randomisierten Reihenfolge zu lösen. Bei der Aufgabenkonstruktion wurde hohen Wert auf Repräsentativität und unterschiedliche Schwierigkeitsgrade gelegt. Jede Aufgabe sprach andere Aufgabenbereiche der Tätigkeit und damit auch andere Funktionalitätsbereiche des Softwaresystems an.

Die Aufgabenbearbeitungszeit wurde mit Ergonom 1.0 [14] aufgezeichnet; Aktionen zur Aufgabenlösung sowie die verbalen Äußerungen des Benutzers mit ScreenCam 2.0 [7], einer Protokollsoftware, die die Aufzeichnung von Bildschirmaktionen als eine Art „Film“ ermöglicht, der unabhängig von dem Softwaresystem abgespielt werden kann.

Die Beurteilung der Beanspruchung mittels BSMA wurde direkt nach jeder Aufgabenbearbeitung vorgenommen, d.h. im Laufe der Untersuchung wurden von jedem Teilnehmer fünf Beurteilungen abgegeben.

4 Ergebnisse

4.1 Fehleranalyse

Von 60 Untersuchungseinheiten (12 Teilnehmer mal 5 Testaufgaben) wurden 47 Einheiten (78%) analysiert. 13 Untersuchungseinheiten konnten aufgrund von Systemabstürzen, Netzwerkproblemen oder ähnlichem bei der Aufgabenbearbeitung nicht protokolliert werden.

Insgesamt wurden 353 Fehlerereignisse registriert. Faßt man diese zusammen, ergeben sich 117 verschiedene Fehlereinheiten. Diese Fehlereinheiten bündeln im Mittel rund 3 Fehlerereignisse (Auftretenshäufigkeit) mit einer mittleren Fehlerbewältigungszeit von 10,7 Sekunden.

Abbildung 4 zeigt alle 117 Fehlereinheiten mit ihren Kennwerten „mittlere Fehlerbewältigungszeit“ und „Auftretenshäufigkeit“. Es wird deutlich, daß der größte Anteil der Probleme relativ niedrige Auftretenshäufigkeiten und relativ kurze Fehlerbewältigungszeiten aufweist. Nur wenige Fehlereinheiten haben hohe Bewältigungszeiten oder hohe Auftretenshäufigkeiten

und es finden sich keine, die hohe Werte auf beiden Kriterien erreichen. Die beiden Kennwerte korrelieren nicht miteinander.

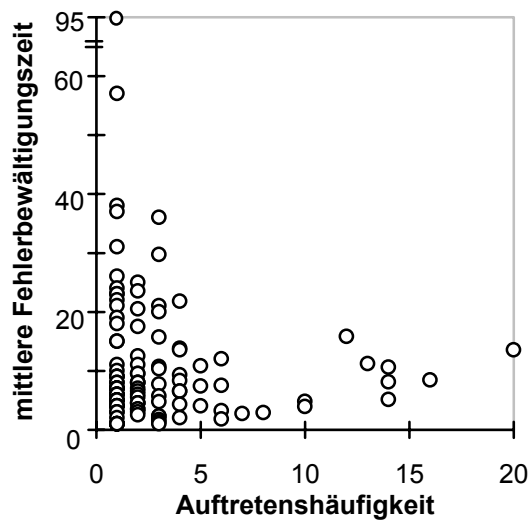


Abb. 4: Die Kombination mittlerer Fehlerbewältigungszeit und Auftretenshäufigkeit bei Fehlereinheiten

4.2 Ergebnisse zur Priorisierung nach Kosten

Die Priorisierung nach Kosten macht nur dann Sinn, wenn (1) der Anteil der Fehlerbewältigungszeit an der Aufgabenbearbeitungszeit so groß ist, daß überhaupt von *ökonomischen* Kosten im Sinne einer Beeinträchtigung der Produktivität ausgegangen werden kann und, wenn (2) die angenommenen *psychischen* Kosten der Probleme über die Fehlerbewältigungszeit eingeschätzt werden können.

Dazu wurde die Aufgabenbearbeitungszeit, die Fehlerbewältigungszeit, die Fehlerhäufigkeit und die Beanspruchung pro Teilnehmer errechnet, indem die Werte der von dem jeweiligen Teilnehmer einzeln bearbeiteten Untersuchungseinheiten gemittelt wurden.

4.2.1 Ökonomische Kosten: Anteil der Fehler- an der Aufgabenbearbeitungszeit

Betrachtet man die Aufgabenbearbeitungszeit gemittelt über die Teilnehmer ($N = 12$), so benötigten diese durchschnittlich 327 Sekunden für die Bearbeitung einer Untersuchungseinheit ($s = 307$, $min = 65$, $max = 1264$). Die mittlere Fehlerbewältigungszeit pro Untersuchungseinheit betrug 102 Sekunden ($s = 151$, $min = 2$, $max = 574$). Dies entspricht einem Anteil der mittleren Fehlerbewältigungszeit an der Aufgabenbearbeitungszeit von 31,2% und weist auf eine massive Beeinträchtigung der Produktivität hin.

4.2.2 Psychische Kosten: Zusammenhang von Fehlerbearbeitungszeit und Beanspruchung

Es ergibt sich über die Teilnehmer eine mittlere Beanspruchung von 49 laut BSMA-Skala ($s = 26$, $min = 20$, $max = 120$). Dieser Skalenwert deutet auf eine eher schwache mentale Beanspruchung hin. Er liegt 9 Skalenpunkte über dem verbalen Anker „etwas anstrengend“ und 25 Skalenpunkte unter „einigermaßen anstrengend“.

Mentale Beanspruchung und Fehlerbewältigungszeit korrelieren signifikant ($r = .88$, $p < .01$, $N = 12$). Personen, die länger zur Bewältigung ihrer Fehler benötigen, fühlen sich mental höher beansprucht.

Die Korrelation von mentaler Beanspruchung und prozentualem Anteil der Fehlerbewältigungszeit an der Aufgabenbearbeitungszeit beträgt $r = .65$ ($p < .05$, $N = 12$). Das bedeutet, daß Personen, die einen höheren Anteil der Aufgabenbearbeitungszeit mit Fehlerbewältigung ver-

bringen, eine höhere Beanspruchung empfinden. Fehlerbewältigungszeit kann als Indikator für psychische Kosten betrachtet werden.

4.3 Ergebnisse zur Priorisierung nach Auftretensstabilität

Als Indikator für die Auftretensstabilität einer Fehlereinheit kann die Anzahl der unterschiedlichen Benutzer herangezogen werden, bei denen dieselbe Fehlereinheit beobachtet werden kann. Abbildung 5 zeigt diese Häufigkeiten. Von 117 verschiedenen Fehlereinheiten konnten 80 bei jeweils nur einer Person beobachtet werden. Dies sind idiosynkratische Fehlereinheiten, die den „Benutzer“-Pol des „Benutzer-Software“-Kontinuums markieren. Betrachtet man das andere Ende des Kontinuums, so finden sich keine Fehlereinheiten, die bei allen Personen gleichermaßen beobachtet wurden. Die höchste Übereinstimmung ist dieselbe Fehlereinheit bei 10 von 12 Benutzern. Dieser Fehler kann als interindividuell stabil bezeichnet werden.

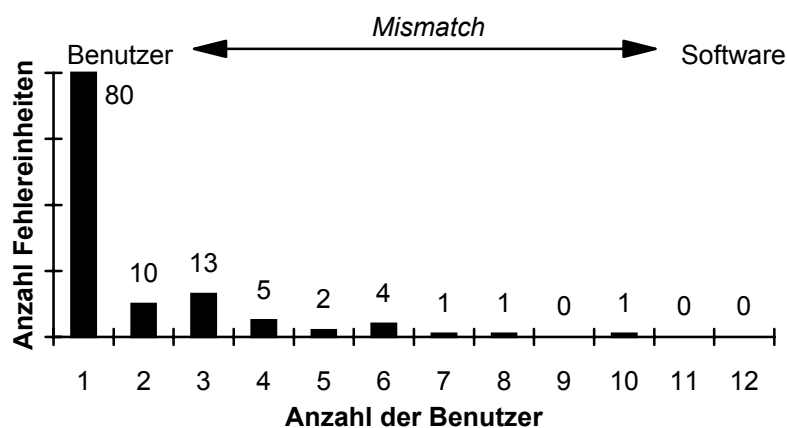


Abb. 5: Anzahl der Fehlereinheiten, die bei Benutzern gleichzeitig beobachtet wurden

4.4 Vergleich der Priorisierungen

Insgesamt ergeben sich aufgrund der hier vorgestellten Überlegungen zur Priorisierung zwei Möglichkeiten der Auswahl von Problemhinweisen, die im folgenden miteinander verglichen werden sollen: (1) Priorisierung nach Fehlerkosten und (2) Priorisierung nach Auftretensstabilität.

Auf Grundlage der Überlegungen zur Priorität in Abhängigkeit von Auftretenshäufigkeit und Bewältigungszeit (vgl. Abb. 3) schlagen wir vor, die 25% der Fehlereinheiten zur Bearbeitung auszuwählen, die den größten Anteil an der gesamten Fehlerbewältigungszeit ausmachen. In dem vorliegenden Fall sind dies 29 von insgesamt 117 Fehlereinheiten. Diese 25% der Fehlereinheiten sind für 69% der gesamten Fehlerbewältigungszeit verantwortlich. Könnten diese Fehlereinheiten beseitigt werden, würde sich der Anteil der Fehlerbewältigungszeit an der Aufgabebearbeitungszeit von 31,2% auf 12,3% verringern und damit auch die assoziierten ökonomischen und psychischen Kosten. Obwohl dies ein idealisierter Wert ist, der davon ausgeht, daß keine neuen Probleme durch Umgestaltung erzeugt werden, wird deutlich, daß in der Bearbeitung weniger, aber sorgfältig ausgewählter Fehlereinheiten ein hohes Potential zur Verbesserung steckt.

Eine gröbere Priorisierung kann anhand der Auftretensstabilität vorgenommen werden. Hierbei kann die Anzahl der Benutzer, die dasselbe Problem haben, als Rangfolge betrachtet werden. Im vorliegenden Fall ist unser Vorschlag, auf die Bearbeitung der idiosynkratischen Fehler zu verzichten. Die 37 der 117 Fehlereinheiten (31,6%), auf die dieses Kriterium zutrifft, sind immer noch eine akzeptable Menge an zu bearbeitenden Problemhinweisen. Auf diese Art ausgewählte Problemhinweise sind für 63,2% der gesamten Fehlerbewältigungszeit

verantwortlich. Die Beseitigung der zugrundeliegenden Probleme könnte zu einer Reduktion des Anteils der Fehlerbewältigungszeit an der Aufgabenbearbeitungszeit von 31,2% auf 14,3% führen.

Die Ergebnisse der beiden Priorisierungsmethoden sind ähnlich. Vergleicht man die Selektion aufgrund beider Priorisierungsmethoden, so findet man, daß von den 29 Problemhinweisen mit dem größten Anteil der Fehlerbewältigungszeit an der gesamten Fehlerbewältigungszeit (Priorisierung nach Fehlerkosten) 19 identisch mit den durch die zweite Methode ausgewählten sind (vgl. Abb. 6). Dies entspricht einer signifikanten Korrelation ($\Phi = .42, p < .001, N = 117$ Fehlereinheiten). Auch der Anteil der erklärten Fehlerbewältigungszeit ist bei beiden in etwa vergleichbar. Allerdings müssen bei der Priorisierung nach Auftretensstabilität 8 Problemhinweise (6,8%) mehr bearbeitet werden.

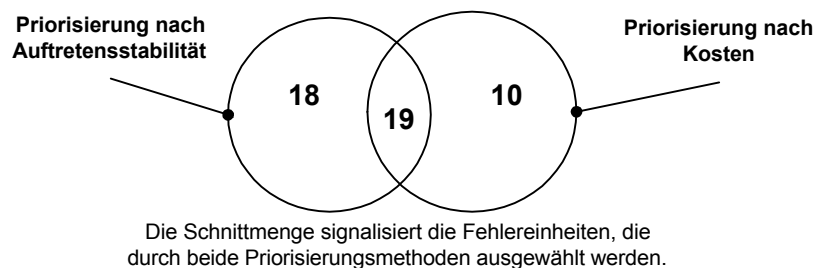


Abb. 6: Anzahl der über Priorisierungsmethoden ausgewählten Fehlereinheiten

5 Diskussion

In den Ergebnissen der Fehleranalyse dominieren Fehlereinheiten, die Fehlerereignisse mit relativ kurzen Fehlerbewältigungszeiten und geringen Auftretenshäufigkeiten zusammenfassen. Um detaillierte Hinweise auf die Schwachstellen eines Softwaresystems zu bekommen, muß mit einem hohen Detailliertheitsgrad in der Analyse gearbeitet werden. Je detaillierter die Analyse, desto weniger Fehlerereignissen werden als ähnlich klassifiziert und zu Fehlereinheiten gebündelt.

Ein Drittel der Aufgabenbearbeitungszeit wird auf die Bewältigung von Fehlern verwendet. Dieser Anteil ist hoch, selbst wenn man davon ausgeht, daß die Teilnehmer relativ ungeübt im Umgang mit dem System waren und dadurch der zeitliche Aufwand zur Behebung steigt [10, 11]. Eine Beseitigung der identifizierten Problembereiche wird zu einer deutlichen Steigerung der Produktivität des Benutzers führen.

Die Einschätzung der subjektiv empfundenen, mentalen Beanspruchung, die zwischen „etwas anstrengend“ und „einigermaßen anstrengend“ liegt, deutet auf eine eher schwache Beanspruchung hin. Selbst die Person, die die höchste Beanspruchung angab, konnotierte mit „ziemlich anstrengend“ einen moderaten Wert. Hier wird der artifizielle Charakter einer experimentellen Gebrauchstauglichkeitsstudie offensichtlich. Man kann davon ausgehen, daß ähnliche Probleme im Kontext der alltäglichen Aufgabenbearbeitung aufgrund ihrer Konsequenzen zu massiveren Beanspruchungen führen werden. Der Bedeutung der Erfüllung bzw. Nichterfüllung einer Aufgabe im Rahmen einer solchen Untersuchung wird ein geringerer Stellenwert als im Arbeitsalltag zugemessen.

Der signifikante Zusammenhang zwischen Fehlerbewältigungszeit bzw. dem Anteil der Fehlerbewältigungszeit an der Aufgabenbearbeitungszeit und subjektiv empfundener mentaler Belastung macht die Fehlerbewältigungszeit zu einem Indikator psychischer Fehlerkosten. Personen, die einen höheren Anteil ihrer Aufgabenbearbeitungszeit mit Fehlerbewältigung verbringen müssen, empfinden auch eine höhere Beanspruchung. Es liegt also nahe, Problemhinweise so zu priorisieren, daß die Fehlereinheiten, die den größten Anteil der Fehlerbewälti-

gungszeit ausmachen, zuerst zu bearbeiten sind. So hat die Bearbeitung weniger Fehlereinheiten eine große Auswirkung auf die Reduktion der Fehlerbewältigungszeit und damit u.U. auch auf die Reduktion der Beanspruchung.

Priorisierung nach Auftretensstabilität ist eine einfachere Methode, da die Fehlerbewältigungszeit nicht erhoben werden muß. Im Vergleich zur Priorisierung nach Kosten führt sie zu ähnlich guten Ergebnissen. Von Nachteil ist, daß im Gegensatz zur Priorisierung nach Fehlerkosten die Zahl der ausgewählten Fehlereinheiten nicht von vornherein festliegt, da die Auftretenshäufigkeit idiosynkratischer Fehlerereignisse von Situation zu Situation unterschiedlich sein kann.

Die Übereinstimmung der anhand beider Methoden ausgewählten Fehlereinheiten ist abhängig von der Konstellation der Auftretenshäufigkeiten und der Bewältigungszeiten von Fehlerereignissen. Die Priorisierung nach Auftretensstabilität unterschätzt im Gegensatz zur Priorisierung nach Kosten systematisch die Einflußkraft von Fehlerereignissen mit hohen Fehlerbewältigungszeiten und niedrigen Auftretenshäufigkeiten. In vorliegendem Fall wurde diese Kombination nicht oft beobachtet, was zu der hohen Übereinstimmung führte. Sollte es der Regelfall sein, daß es eher selten Probleme gibt, die nur vereinzelt auftreten und dann stets lange brauchen, bis sie bewältigt sind, dann ist die Methode der Priorisierung nach Auftretensstabilität der Priorisierung nach Kosten vorzuziehen: sie ist leichter durchführbar und erzielt trotzdem ähnlich gute Ergebnisse. Gibt es viele Fehlerereignisse mit hohen Fehlerbewältigungszeiten und niedrigen Auftretenshäufigkeiten, ist die Priorisierung nach Kosten die Methode der Wahl.

Im Zentrum der vorgeschlagenen Priorisierungsarten stehen Fehlerkosten, die über Fehlerbewältigungszeiten abgeschätzt werden. Dies geschieht vor dem Hintergrund, daß Beanspruchung im Arbeitsalltag zu Streß wird, und daß durch software-ergonomische Gestaltungsmaßnahmen die Beanspruchung reduziert werden kann. Deutlich wird, daß durch die Anwendung der vorgeschlagenen Priorisierungsmethoden zur Auswahl von Fehlereinheiten ein großer Teil der Fehlerbewältigungszeit - und damit der Fehlerkosten - reduziert werden könnte.

Offen ist, ob diese Kriterien zur Priorisierung von allen Personengruppen geteilt werden. Unter Umständen ziehen Personen Kriterien vor, die wenig mit objektiven Fehlerkosten, sondern eher mit ganz persönlichen Interessen, mit Vorstellungskraft, Gestaltungsspielraum oder finanziellen Aspekten zu tun haben.

Der Einsatz von Fehleranalysen und das Konzept der Fehlereinheit kann im Rahmen einer formativen Evaluation hilfreich sein. Besonders die Fehlereinheit bietet die Möglichkeit, qualitative mit quantitativen Daten zu verbinden. Aber sie darf nicht das einzige Mittel zur Analyse eines Softwaresystems sein. In einer empirischen Gebrauchstauglichkeitsstudie sind Merkmale des Systems und Merkmale der teilnehmenden Benutzer, z.B. Werkzeugwissen und Computerwissen, stark miteinander verschränkt. Es wird dabei nicht nur das System an sich getestet, sondern auch die Angemessenheit des Systems für die testende Benutzerpopulation. Verändert sich diese, werden sich auch die Ergebnisse der Untersuchungen ändern. Um diesen Einfluß zu kontrollieren, müssen auf der einen Seite Maßnahmen zur Erfassung der relevanten Benutzermerkmale ergriffen werden, auf der anderen Seite müssen beobachtete Nutzungsprobleme immer auf dem Hintergrund der Benutzerpopulation interpretiert und erst dann in Gestaltungsmaßnahmen umgesetzt werden.

In weiteren Arbeiten muß geprüft werden, inwiefern das Konzept der Fehlereinheit praxistauglich ist und in Verbindung mit der hier vorgeschlagenen Priorisierung in ein umfassendes Gestaltungskonzept eingepaßt werden kann. Da es immer auch die Möglichkeit gibt, Probleme durch beteiligte Personengruppen, wie Benutzer, Entwickler oder Auftraggeber, priorisieren zu lassen, müssen Untersuchungen folgen, die auch diese Methode weiter explorieren und bewerten.

Wie dem auch sei: die vorliegende Arbeit konnte zeigen, daß Priorisierung anhand empirischer Daten möglich ist und im Rahmen software-ergonomischer Qualitätssicherung einen Beitrag zur effizienten Bearbeitung von Problembereichen eines Softwaresystems zu leisten vermag.

Literaturverzeichnis

- [1] F.C. Brodbeck: Fehlerbewältigungszeiten und die Nutzung von Unterstützungsmöglichkeiten. In: M. Frese & D. Zapf (Hg.): Fehler bei der Arbeit mit dem Computer - Ergebnisse von Beobachtungen und Befragungen im Bürobereich. Huber, Bern, 1991, 80-94.
- [2] F.C. Brodbeck, D. Zapf, J. Prümper & M. Frese: Error handling in office work with computers: A field study. In: Journal of Occupational and Organizational Psychology 60(4) (1993), 303-317.
- [3] K. Eilers, F. Nachreiner & K. Hänecke: Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. In: Zeitschrift für Arbeitswissenschaft 40 (1986), 215-224.
- [4] M. Frese & D. Zapf: Fehlersystematik und Fehlerentstehung: Eine theoretische Einführung. In: M. Frese & D. Zapf (Hg.): Fehler bei der Arbeit mit dem Computer - Ergebnisse von Beobachtungen und Befragungen im Bürobereich. Huber, Bern, 1991, 14-31.
- [5] L. Herrmann, L.L. Morris & C. Taylor Fitz-Gibbon: Evaluator's handbook. Beverly Hills, Sage, 1988.
- [6] A. Kensik, J. Prümper & M. Frese: Ergonomische Gestaltung von Software auf Grundlage handlungsorientierter Fehleranalysen. In: H.D. Böcker (Hg.): Software-Ergonomie '95 - Anwendungsbereiche lernen voneinander. Stuttgart, Teubner, 1995, S. 217-232.
- [7] Lotus: ScreenCam 2.0. Lotus Development Corporation, 1995.
- [8] T.W. Malone: Toward a theory of intrinsically motivating instruction. In: Cognitive Science 4 (1981), 333-369.
- [9] L. Müller-Böling: Akzeptanz und Partizipation - Sind Systemgestalter lernfähig? In: K.T. Schröder (Hg.): Arbeit und Informationstechnik. Springer, Berlin, 1986, 153-166.
- [10] J. Prümper: Handlungsfehler und Expertise. In: M. Frese & D. Zapf (Hg.): Fehler bei der Arbeit mit dem Computer - Ergebnisse von Beobachtungen und Befragungen im Bürobereich. Huber, Bern, 1991, 118-130.
- [11] J. Prümper, D. Zapf, F.C. Brodbeck & M. Frese: Some surprising differences between novice and expert errors in computerized office work. In: Behaviour & Information Technology 11(6) (1992), 319-328.
- [12] J. Rasmussen: Human Error Data. Facts or Fiction. Riso National Laboratory, Roskilde, 1985.
- [13] I. Udris & M. Frese: Belastung, Streß, Beanspruchung und ihre Folgen. In: D. Frey, C. Graf Hoyos & D. Stahlberg (Hg.): Angewandte Psychologie. Psychologie Verlags Union, München, 1988, 428-447.
- [14] S. Willmeroth: Ergonom 1.0. Ziff Verlag GmbH, 1993.
- [15] H. Wottawa & H. Thierau: Evaluation. Huber, Bern, 1990.
- [16] D. Zapf, F.C. Brodbeck & J. Prümper : Handlungsorientierte Fehlertaxonomie in der Mensch-Computer Interaktion. In: Zeitschrift für Arbeits- und Organisationspsychologie 33 (1989), 178-187.
- [17] R. Zijlstra & L. van Doorn: The construction of a scale to measure subjective effort. Universität Groningen: Unveröffentlichter Bericht, 1985.

Adressen der Autoren

Marc Hassenzahl
TH-Darmstadt
Fachbereich Psychologie
Hillebergstraße 42
64319 Pfungstadt

hassenzahl@aol.com
hassenzahl@hrz1.hrz.th-darmstadt.de

Prof. Dr. Jochen Prümper
FHTW-Berlin
Fachgebiet Wirtschaftspsychologie
Treskowallee 8
10313 Berlin

pruemper@aol.com
pruemper@rz.fhtw-berlin.de

Uta Sailer
LMU- München
Fachbereich Psychologie
Orleansstraße 39
81667 München

saileru@aol.com
sailer@mip.paed.uni-muenchen.de