

Aspekte effizienten Usability Engineerings

Aspects of Efficient Usability Engineering

Ralf Nieschulz, Björn Schuller, Michael Geiger, Technische Universität München
Robert Neuss, Usaneers GmbH, München

In diesem Artikel werden einige Methoden vorgestellt, die eine effiziente Durchführung von Usability-Studien ermöglichen. Sie wurden schrittweise im Rahmen laufender Untersuchungen und konkreter Fragestellungen u. a. aus den Bereichen Gestenerkennung, Benutzeradaption und Multimodalität am Lehrstuhl für Mensch-Maschine-Kommunikation der TU München entwickelt. Dabei wurde insbesondere auf eine allgemeine Anwendbarkeit geachtet. Weitgehende Automatisierung sowie der Einsatz von Sprachsteuerung vereinfachen die Durchführung von Usability-Untersuchungen und führen zu aussagekräftigen und nachvollziehbaren Ergebnissen.

This article introduces several methods which allow an efficient execution of usability studies. They have been developed gradually within the framework of ongoing examinations and specific questions concerning gesture recognition, user adaptation and multi-modality at the Institute of Human-Machine Communication of the Technical University of Munich. In the development process certain stress was laid on a general applicability. Extensive automation, as well as the use of speech control facilitate the execution of usability surveys and lead to meaningful and comprehensible results.

1 Einführung

Eine Tendenz, welche sich bei Geräten der Informations- und Kommunikationstechnik auf Grund neuer technischer Möglichkeiten und steigender Anforderungen abzeichnet, ist die zunehmende Integration einer Vielzahl neuer Funktionen. Dies führt zu komplexeren Benutzungsschnittstellen, deren Handhabung für den Anwender immer schwieriger wird. Usability-Aspekte treten daher mehr und mehr in den Vordergrund. Um die sensorischen Kanäle des Menschen optimal auszunutzen, werden neben der etablierten haptischen Eingabe und visuellen Ausgabe zunehmend auch natürliche Dialogformen wie Sprache und Gestik eingesetzt. Offensichtlich birgt die Multimodalität – hier verstanden als die Verwendung mehrerer Kommunikationskanäle – ein großes Potenzial zur Leistungssteigerung und Vereinfachung der Gerätebedienung. Verglichen mit monomodalen Anwendungen erhöht sich jedoch neben dem Realisierungsaufwand in jedem Fall der Aufwand für Usability.

In der Literatur werden verschiedene Ansätze multimodaler Systeme beschrieben [1]. Ein möglicher An-

satz besteht darin, den individuellen Interaktionspräferenzen der Benutzer durch ein redundantes Angebot an Ein- und Ausgabemodalitäten entgegen zu kommen. Vorteile sind unter anderem ein erleichteter Erstkontakt aufgrund mehrerer Eingabealternativen, sowie die Möglichkeit, bei Langzeitnutzung die unterschiedlichen Stärken verschiedener Modalitäten zur Lösung einer Aufgabe besonders effizient zu kombinieren.

Wie Arbeiten am Lehrstuhl für Mensch-Maschine-Kommunikation zeigen, gestaltet sich der Entwurf eines entsprechend aufgabenorientierten Systems als vielschichtig [2], da mit neuen Modalitäten Probleme hinsichtlich technischer Realisierung und Gebrauchstauglichkeit auftauchen. Dies relativiert die angestrebten Vorteile. Beispielsweise erschweren Randbedingungen wie Fehlerkennungen, unterschiedliche Verarbeitungsgeschwindigkeiten von Erkennungssystemen oder die automatische Segmentierung von Benutzeraktionen die Interpretation der Eingaben und somit die Bedienung. Gleichzeitig gilt es, Humanfaktoren zu berücksichtigen, welche sich im Interaktionsverhalten des Benutzers manifestieren: Modalitätenwechsel, das Ignorieren von Eingabeauf-



forderungen, Planänderungen, Pausen, Interaktion mit menschlichen Partnern, Gewohnheiten, Ablenkung etc. Um sinnvolle Bedienkonzepte erarbeiten zu können, werden deswegen aufwändige Vorüberlegungen angestellt und Wizard-of-Oz-Versuche¹ (WoO) durchgeführt, mit dem Ziel, das menschliche Verhalten und typische Bediensituationen zu kategorisieren.

Aus solchen Voruntersuchungen ergeben sich zahlreiche Vereinfachungsmöglichkeiten. Für viele Anwendungen ist es beispielsweise nicht zwingend notwendig, dass Eingabekanäle jederzeit gleichzeitig kombiniert werden können (*parallel ergänzende Multimodalität*), sondern es reicht oft aus, dass Modalitäten beliebig gewechselt werden können (*seriell redundante Multimodalität*). Durch derartige Vereinfachungen der Randbedingungen von Testszenerien ist es mit vertretbarem Aufwand möglich, reelle Prototypen aufzubauen und durch kontinuierliche Benutzertests zu optimieren und weiter zu entwickeln. Dieser Artikel soll nun ausgewählte Aspekte innovativen Usability Engineerings erläutern, welche zur Datenerhebung und Prototypenentwicklung eingesetzt werden.

2 Versuchsumgebung

Ein Usability-Experiment gestaltet sich üblicherweise wie folgt: Eine Versuchsperson wird mit der Bedienung des zu testenden Geräts konfrontiert und dabei von einem Versuchsleiter beobachtet. Hauptziel ist es, Aussagen über Aspekte wie Benutzerfreundlichkeit oder Akzeptanz zu treffen. Zu den Grundaufgaben eines Versuchsleiters (siehe auch Bild 1) gehören die Ablaufsteuerung und die Überwachung des Usability-Experiments. Er führt die Versuchsperson durch ein Testszenario, indem er anwendungstypische Bedienungsaufgaben stellt (z. B. im Kontext Telefonanwendung: „Rufen Sie Herrn Zobl an.“). Dabei beobachtet er das Benutzerverhalten und macht sich Notizen zum Bedienverhalten der Versuchsperson und zu Auffälligkeiten. Auf diese Aufzeichnungen wird bei der anschließenden Videoauswertung zurückgegriffen, wobei interessante Situationen detaillierter untersucht werden. Darüber hinaus muss der Versuchsleiter mit der Testperson interagieren – z. B. als Telefongesprächspartner. Handelt es sich zudem um eine WoO-Studie, d. h. der Versuchsleiter simuliert das Systemverhalten bzw. Teilkomponenten des Systems, so werden die Grenzen der Machbarkeit selbst für geübte Experimentatoren schnell erreicht. Einer-

¹ In Wizard-of-Oz-Versuchen beobachtet ein Versuchsleiter, Wizard genannt, eine Versuchsperson und simuliert das Verhalten eines technischen Systems. In unseren Szenarien steht oft ein reales System zur Verfügung und der Wizard ersetzt Eingabegeräte wie einen Gestenerkennung oder ein sprachverstehendes System.



Bild 1: Versuchsleiter im Kontrollraum.

seits muss die Versuchsperson permanent beobachtet werden, um alle Benutzeraktionen zu erfassen, andererseits gilt es möglichst unmittelbar auf diese Aktionen zu reagieren – üblicherweise durch haptische Eingaben. Zu lange Reaktionszeiten führen zu abweichendem Benutzerverhalten (z. B. Blickabwendungen) und beeinflussen die Studie somit in unerwünschter Weise. Bei WoO-Experimenten zur gestischen Bedienung² ist die permanente visuelle Beobachtung der Benutzeraktionen von besonders großer Bedeutung. Es hat sich gezeigt, dass es einer Einzelperson kaum möglich ist, gleichzeitig zu beobachten, in angemessener Zeit mit haptischen Eingaben die entsprechenden Systemreaktionen auszulösen und zusätzlich den Ablauf des Experiments zu steuern.

Dieser Sachverhalt legt es nahe, die anfallenden Aufgaben (Stellen der Bedienungsaufgaben, Systemsteuerung durch WoO, Beobachten des Benutzerverhaltens etc.) auf mehrere Versuchsleiter zu verteilen. Insbesondere bei Usability-Untersuchungen zur multimodalen Bedienung ergäbe sich somit ein hoher Personalaufwand. Die nachfolgend beschriebenen Methoden (siehe auch [4]) erlauben die weitgehende Automatisierung der Versuchsumgebung, wobei auch die Bedienoberfläche des Versuchsleiters an benutzerfreundliche Kriterien angepasst wird – sozusagen Usability für Usability-Untersuchungen. Ziel ist es, die Durchführung einem einzelnen Versuchsleiter zu ermöglichen und dabei sicherzustellen, dass alle Einzelversuche unter nahezu identischen Rahmenbedingungen ablaufen und somit gut vergleichbar sind.

Das System ist aus einzelnen Modulen aufgebaut (Bild 2). Sie kommunizieren via Netzwerk (TCP/IP) miteinander und können somit auf verteilten Rechnern laufen. Eine Anpassung an die konkrete Frage-

² Unter gestischer Bedienung wird hier die berührungslose Steuerung von Geräten mittels Hand- oder Kopfbewegungen verstanden [3].

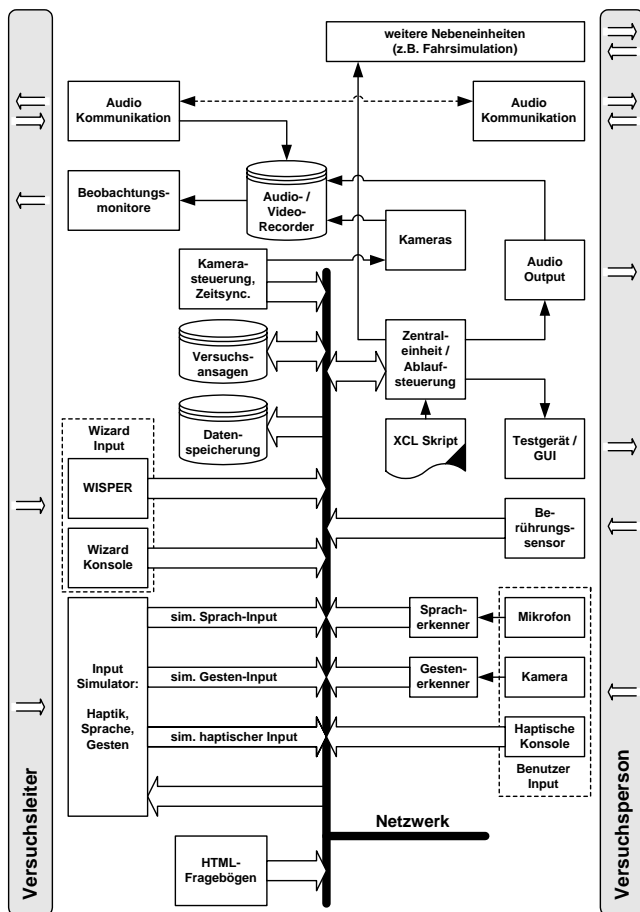


Bild 2: Systemübersicht.

stellung sowie Erweiterungen des Systems gestalten sich daher unkompliziert. Im Folgenden wird auf einige dieser Module ausführlicher eingegangen.

3 XCL – Programmierung von Versuchsabläufen

Um das Design von Versuchsszenarien zu beschleunigen und übersichtlich zu gestalten, wurde die Versuchsbeschreibungssprache XCL (eXperiment Control Language; siehe Bild 3) eingeführt. Mit Hilfe dieser Interpretersprache können Versuchsabläufe weitgehend automatisiert werden. XCL verfügt über Schleifen und Kontrollstrukturen, sowie über Kommentierungs- und Segmentierungsfunktionen. Der Programmcode wird in einer zentralen Steuereinheit abgearbeitet, welche aufgabenbedingte Nebeneinheiten und unter Umständen das zu untersuchende Gerät selbst kontrolliert, sowie die Ergebnisdatenströme einheitlich bündelt.

Die Nebeneinheiten steuern einerseits Aufgaben wie die Ansage der Versuchsaufgaben oder das Abfragen von Fragebögen, andererseits Ablenkungseffekte oder Nebenaufgaben zur Bedienung. Letztere können etwa im Fahrzeug- oder Luftfahrtbereich

```

<S> ::= <CMD_SEQ>
<CMD_SEQ> ::= <CMD_LINE> | <CMD_LINE><CMD_SEQ>
<CMD_LINE> ::= <TASK_DEF> | <TU_CMD> | <TS_CMD> | <GUI_CMD> | <WAIT_CMD> | <JUMP_CMD> | <REACT_CMD> | <WIZMSG_CMD>
<TASK_DEF> ::= task <TASK_ID> <TASK_DESC>
<TASK_ID> ::= <chars>
<TASK_DESC> ::= <chars>
<TU_CMD> ::= sim <tu_command> | <TU_ID> <tu_command>
<TS_CMD> ::= app <ts_command> | <TS_ID> <ts_command>
<GUI_CMD> ::= gui <gui_command> | gui <GUI_ID> <gui_command>
<WAIT_CMD> ::= wait <digits>
<WIZMSG_CMD> ::= wizmsg <chars>
<REACT_CMD> ::= <WIZWAIT_SCR> | <WIZWAIT_CMD>
<WIZWAIT_SCR> ::= wizwait_script | wizwait_script <rcscript> | ...
<WIZWAIT_CMD> ::= wizwait_cmd | wizwait_cmd <INTERN_CMD> | ...
<INTERN_CMD> ::= <WAIT_CMD> | <TU_CMD> | <TS_CMD> | <GUI_CMD> | <WIZMSG_CMD>
<JUMP_CMD> ::= <SCRIPT_CMD>
<JUMP_CMD> ::= skip_task | repeat_task
<SCRIPT_CMD> ::= script <rcscript>
<WIZMSG_CMD> ::= wizmsg <chars>
    
```

Bild 3: XCL-Auszug in Backus-Naur-Form.

eine Simulation der – aus Sicht des späteren Anwenders – Hauptaufgabe (z. B. Steuerung eines Fahr- oder Flugzeugs) darstellen. Durch die zentrale Ansteuerung solcher Nebenaufgaben wird eine Synchronisation mit der Versuchsaufgabe gewährleistet. Im Beispiel „Fahrzeugsimulation“ kann so etwa zu einer geforderten Bedienungsaufgabe zeitgleich eine extreme Fahrsituation simuliert werden. Eine Benutzerinteraktion kann somit gezielt in Stress- oder sicherheitsrelevanten Situationen erprobt und die kognitive Last des Probanden in Bediensituationen beeinflusst werden. Gleichzeitig werden für alle Probanden konstante Umgebungs- und Ablaufbedingungen gewährleistet.

3.1 Vorhersehbarkeit von Versuchsabläufen

Bei vielen Versuchen ist die Abfolge der Bedienaktionen vorhersehbar. Unter der Annahme eines erwartungskonformen Verhaltens der Versuchspersonen kann durch eine Programmierung vorab nahezu der gesamte Ablauf automatisiert werden. Der Versuchsleiter gibt dem System in einem solchen Szenario nur noch den Zeitpunkt einer Probandenaktion an. Das System führt die zu diesem Zeitpunkt erwartete Aktion des Testgeräts aus und fährt mit



dem Ablauf fort. Solche Versuche können etwa beim Sammeln von Benutzeräußerungen oder Gesten für das Training stochastischer Sprach- oder Gestenerkennungssysteme zum Einsatz kommen. Hier werden der Versuchsperson in der Regel klare Einzelaufgaben gestellt, bei denen vor allem die gewählte Formulierung oder Geste von Interesse ist, das Bedienziel aber a priori bekannt ist. Eine weitere Motivation multimodaler Untersuchungen ist die Erhebung der Modalitätenpräferenz eines Anwenders bezüglich einer bestimmten Intention. Auch hier ist in der Regel ein Ablaufplan vorherbestimmt.

Zeigt sich der Verlauf eines Versuchs ansatzweise abschätzbar, können gezielt Verzweigungspunkte im Versuchsskript gesetzt werden. Der Versuchsleiter kann sich so auf die Wahl aus wenigen erwarteten Probandenaktionen beschränken, ohne die volle Gerätefunktionalität bedienen zu müssen. Beispielsweise könnte ein Benutzer in einem Teilziel im Kontext Telefonbedienung die Aufgabe haben, Frau Koch anzurufen. Als zu erwartende Möglichkeiten bieten sich etwa „Rufen Sie Frau Koch an!“ oder „Wählen Sie die Nummer 0043 ...“ Zu diesem Zeitpunkt des Versuchs kennt das System das Bedienziel (Frau Koch anrufen) und stellt dem Versuchsleiter diese Funktion samt diverser Parameter als Shortcut bereit.

Weicht der Plan einer Versuchsperson von dem Angenommenen ab, da diese etwa das System erforschen will, das Ziel missverstanden hat, oder nicht mit der Bedienung zurecht kommt und so einen eventuellen Mangel des Konzepts aufdeckt, steht dem Versuchsleiter trotzdem die volle Systemfunktionalität für einen manuellen Eingriff zur Verfügung. Dies dient der Aufrechterhaltung der Simulation bis zur Rückführung zum eigentlichen Ablaufplan. Gelingt dies, kann der Versuch bis auf die abweichende Sequenz als reproduzierbar angesehen werden.

3.2 Kommentierungsfunktionalität

Ein Versuch wird bei der Beschreibung in XCL in Einzelziele zerlegt. Die dadurch entstehenden Etappen können direkt angesprungen, übersprungen oder wiederholt werden. Hierzu verfügt der Versuchsleiter über eine grafische Bedienkonsole. Hier werden die aktuelle Versuchsphase und Anweisungen oder Kommentare angezeigt. Die hierzu bereit gestellten speziellen Kommentierungsfunktionen erlauben zusätzlich, das Lesen der Skripte einfacher zu gestalten. Generell ist das direkte Ablegen des XCL-Skripts in einer Versuchsdatenbank möglich. Das Skript dient so als exakte Versuchsbeschreibung, wodurch eine weitere Dokumentation entfällt.

4 WISPER – Sprachgesteuerte Versuchskontrolle

4.1 Ablaufsteuerung

Nach der Ausarbeitung eines Versuchsszenarios werden die darin gestellten Bedienungsaufgaben als vorgefertigte Sprachsequenzen abgespeichert. Die Aufgaben müssen vom Versuchsleiter während des Versuchs nicht vorgelesen, sondern lediglich abgerufen werden. Durch die abgespeicherten Sprachsequenzen ist gewährleistet, dass die Aufgaben bei verschiedenen Versuchspersonen stets identisch in deutlich verständlicher Weise formuliert sind, wodurch die Ergebnisse später optimal vergleichbar sind. Die Aufgabenblöcke können durch Sprachkommandos wie „nächste Aufgabe“ oder „Aufgabe wiederholen“ sequenziell gesteuert werden. Das speziell für Usability-Experimente entwickelte Spracherkennungssystem WISPER (*WIZard SPEech Recognition*), das durch einen Fuß-taster aktiviert wird (Push-to-Talk), übergibt die erkannten Kommandos an die Ablaufsteuerung, welche daraufhin das Vorlesen der aktuellen Aufgabe startet. Die Sprachbedienung ermöglicht dem Versuchsleiter die ständige Beobachtung der Versuchsperson per Monitor ohne Blickabwendungen.

4.2 Simulation von Erkennungssystemen

Eine weitere wichtige Aufgabe von WISPER im Rahmen einer WoO-Studie ist die Simulation eines automatischen Erkennungssystems, etwa zur Sprach- oder Gestenerkennung. Der Versuchsleiter übersetzt beobachtete (bzw. gehörte) Aktionen der Versuchsperson in Sprachkommandos, welche von WISPER an die Ablaufsteuerung der zu evaluierenden Applikation weitergeleitet werden. Für die Applikation macht es keinen Unterschied, ob der Input von einem real implementierten Erkennungssystem oder von WISPER kommt.

Ein Beispielszenario: Eine Versuchsperson führt eine Geste aus, indem sie mit einer Handbewegung das Abheben eines Telefonhörers imitiert. Der WoO beobachtet dies auf seinem Monitor, betätigt den Fuß-taster um die WISPER-Spracherkennung zu aktivieren und spricht „abheben“. WISPER erkennt das Sprachkommando und veranlasst den Input-Simulator (siehe Bild 2), einen entsprechenden Systembefehl an die Zentraleinheit zu senden, welche daraufhin einen Telefonanruf initiiert.

Diese Methodik mag auf den ersten Blick etwas umständlich erscheinen, bietet jedoch neben der stark verringerten Ablenkung des Versuchsleiters weitere Vorteile. Durch den stark begrenzten Wortschatz treten bei WISPER praktisch keine Fehlerkennungen auf, die das Benutzerverhalten auf unerwünschte und unerwartete Weise beeinflussen könnten. Manchmal

ist jedoch gerade die Reaktion eines Benutzers auf Systemfehler von Interesse; in diesem Fall lassen sich simulierte Fehlerkennungen gezielt und reproduzierbar einsetzen.

Als vorteilhaft erweisen sich zudem die verglichen mit haptischer Versuchssteuerung deutlich kürzeren Gesamterkennungszeiten (verstrichene Zeit zwischen vollständig abgeschlossener Benutzeraktion und Systemreaktion; typische Werte: ca. 300 ms).

5 Effiziente Datenerfassung und -auswertung

5.1 Automatische und synchronisierte Datenerfassung

Während einer Usability-Studie fällt eine beträchtliche Menge an Datenmaterial an: Systemzustände, Benutzeraktionen, aktuelle Aufgabe inkl. Versuchsblock, Steueraktionen etc.. Nicht immer ist bei der Planung und Durchführung eines Experiments bereits im Detail klar, welche Daten bei der späteren Auswertung notwendig oder nützlich sein werden. Um

nicht später relevante Informationen aus dem üblicherweise mitgeschnittenen Videomaterial gewinnen zu müssen, was erfahrungsgemäß mühselig und zeitaufwändig ist, ist es wichtig, möglichst viel Information bereits automatisch während der Versuchsdurchführung zu erfassen. Eine weitere Funktion der Ablaufsteuerung ist daher, alle automatisch erfassbaren Daten mit Zeitstempel abzuspeichern. Der Versuchsablauf wird zusätzlich konventionell auf Videoband aufgenommen (meist aus mehreren Perspektiven; Bild-im-Bild-Aufzeichnung). Durch einen im Videobild eingblendeten synchronisierten Timecode lässt sich der Aufwand für eine eventuell nötige Sichtung des Videomaterials weiter verringern. Es wird ein einfacher Vergleich bestimmter Datenblöcke mit der entsprechenden Videosequenz ermöglicht (Bild 4).

5.2 Erfassung von Bedienzeiten

Da bei Usability-Studien die Bedienbarkeit von Geräten im weitesten Sinne im Mittelpunkt steht, sind Reaktions- und Ausführungsdauern der Versuchspersonen als Maß für die Effizienz der Bedienung von besonderer Bedeutung (Bild 5). Diese Zeiten lassen

NR.	DATE	TIME	EVENT	DEVICE	STATE	MENU	TASK
...
180	12. Okt 00	15:23:38	reading_task_start	Main_menu	select_device	1	15
181	12. Okt 00	15:23:41	reading_task_end	main_menu	select_device	1	15
182	12. Okt 00	15:23:41	sensor_release_start	main_menu	select_device	1	15
183	12. Okt 00	15:23:41	recognition_start	main_menu	select_device	1	15
184	12. Okt 00	15:23:42	release_time 1155 ms	main_menu	select_device	1	15
185	12. Okt 00	15:23:42	gesture 'enter' 684 ms	main_menu	select_device	1	15
186	12. Okt 00	15:23:44	sensor_release_start	phone	no_call	0	15
187	12. Okt 00	15:23:44	recognition_start	phone	no_call	0	15
188	12. Okt 00	15:23:45	release_time 1168 ms	phone	no_call	0	15
189	12. Okt 00	15:23:45	gesture 'phone_up' 573 ms	phone	initiate_call	0	15
190	12. Okt 00	15:23:54	wizard_accept_call	phone	active_call	0	15
191	12. Okt 00	15:24:12	wiz_end_call	phone	no_call	0	15
...



Bild 4: Zeitsynchrone Datenerfassung: Auszug aus einem Logfile (links), zugehöriges Videobild (rechts).

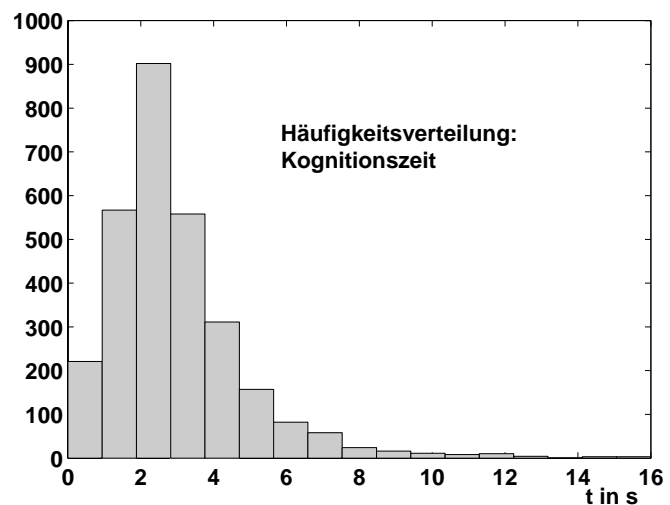
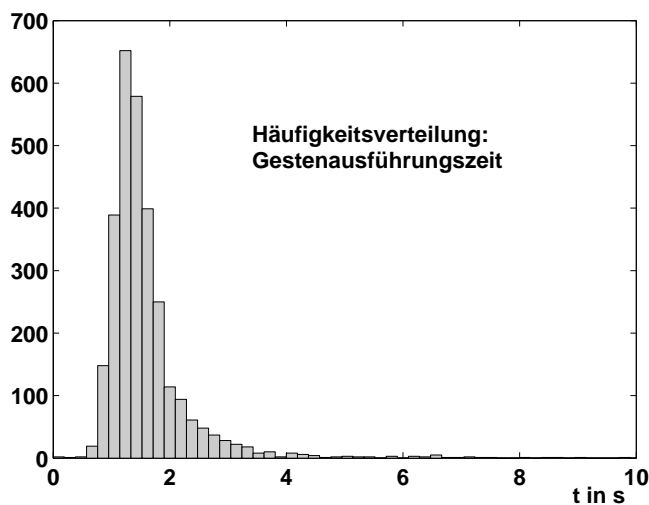


Bild 5: Beispiele gemessener Bedienzeiten aus einer Untersuchung zur Gestensteuerung im Fahrzeug [5].



Bild 6: Berührungssensor in Fahrzeugumgebung.

sich automatisch mittels eines Berührungssensors erfassen. Die Versuchsperson wird dazu angehalten, diesen Sensor permanent zu berühren und nur zur Ausführung einer gestellten Aufgabe loszulassen. Um keine unerwünschten Nebeneffekte zu erzielen, ist bei der Realisierung des Sensors auf eine natürliche Haltung der Versuchsperson zu achten, was beispielsweise mit einem Lenkradsensor im Fahrzeug (Bild 6) oder einer Ablagefläche für die Hand bei der Bildschirmarbeit erreicht wird.

6 Automatische Messung und Anpassung an die Kontrollbedingung

Durch eine initial durchgeführte automatische Messung der Kontrollbedingung (Baseline-Erhebung) kann vor einem Versuch das individuelle Leistungsniveau einzelner Probanden bei der Erfüllung von Bedienungsneben- oder -hauptaufgaben gemessen werden. Ziel einer solchen Erhebung ist es, bei einer Versuchsdurchführung oder -auswertung unterschiedlichen interindividuellen Leistungsniveaus gerecht zu werden. Die Ergebnisse können nach einem Versuch in der Auswertung berücksichtigt werden, um statistische Verzerrungen bei der Analyse zu vermeiden. In speziellen Versuchen, in denen eine kognitive Grundlast der Probanden gewährleistet sein soll – etwa, um Stresssituationen zu simulieren – ist die Baseline-Erhebung eine Möglichkeit zur Regelung der Probandenlast. Diese Regelung kann durch die automatische Messung und Auswertung der Kontrollbedingung augenblicklich erfolgen. Im Beispiel einer Fahrzeugsimulation kann etwa durch Variation der Schwierigkeit der Fahrstrecke die Fähigkeit einer Versuchsperson, ein Fahrzeug zu steuern, ermittelt werden.

7 Automatische Rekonstruktion des Versuchsablaufs

Die abgespeicherten Daten erlauben eine einfache, direkte Auswertung, sowie die authentische Rekonstruktion des gesamten Versuchsablaufs durch einen automatischen *Testreplayer*. Dieser lässt sich sowohl zu Demonstrations- wie auch zu Evaluierungszwecken verwenden. Mit seiner Hilfe lassen sich aber zudem neue Verfahren und Methoden ohne die Durchführung einer neuen Versuchsreihe anhand der bereits gewonnen Versuchsdaten entwickeln und testen. Die durch ein neues Zusatzmodul aus den Versuchsdaten gewonnenen Erkenntnisse und Schlüsse, z. B. zur Bildung eines Benutzermodells zur Adaption eines Geräts an den Benutzer, lassen sich direkt im virtuellen, quasi durch den Testreplayer durchgeführten Versuch verwenden, um das geänderte bzw. verbesserte Systemverhalten zu untersuchen. Natürlich ersetzt dieses Vorgehen keine abschließende Untersuchung des so erweiterten Gesamtsystems, aber die Entwicklungszeit wird um einige Zwischenschritte verkürzt.

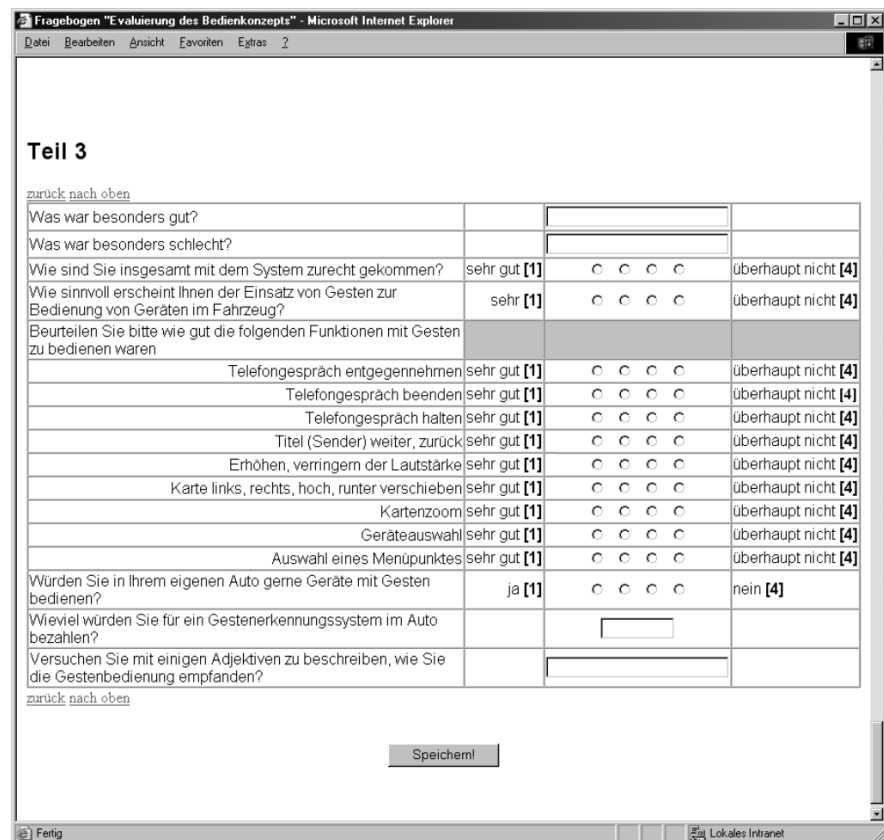
Dieses Verfahren wurde beispielsweise bei der Entwicklung eines Hilfesystems zur gestischen Bedienung im Fahrzeug [5] angewandt.

8 Elektronische Fragebögen

Nach jedem einzelnen Versuchsabschnitt erfolgt ein Interview und eine Einschätzung der Versuchsperson durch den Versuchsleiter. Die gewonnenen subjektiven Daten, d. h. die Einschätzung des zu untersuchenden Systems durch den Benutzer und die Akzeptanz, sind in einer Benutzerstudie von großer Bedeutung, die oftmals unterschätzt wird.

Der Versuchsleiter füllt gemeinsam mit der Versuchsperson einen HTML-Fragebogen aus. Er liest die Fragen jeweils kommentarlos vor und gibt die mündliche Antwort des Probanden online in den Fragebogen ein. Dies hat den Vorteil, dass die Versuchsperson sich ganz auf das Beantworten der Fragen konzentrieren und das Ausfüllen auf einem Display mitverfolgen kann (Bild 7), so dass sich Missverständnisse möglichst vermeiden bzw. sofort klären lassen. Dies animiert die Versuchspersonen zudem, sich intensiver mit der Fragestellung auseinander zu setzen, als beim anonymen Beantworten von Fragebögen. Natürlich ist bei dieser Vorgehensweise zu beachten, dass die Versuchsperson nicht in irgendeiner Form beeinflusst wird. Der Fragebogen wird schließlich vor dem Abspeichern automatisch (skriptbasiert) auf Vollständigkeit überprüft.

Die anschließende Einschätzung der Versuchsperson durch den Versuchsleiter geschieht ebenfalls mittels



Teil 3

[zurück nach oben](#)

Was war besonders gut?		<input type="text"/>	
Was war besonders schlecht?		<input type="text"/>	
Wie sind Sie insgesamt mit dem System zurecht gekommen?	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Wie sinnvoll erscheint Ihnen der Einsatz von Gesten zur Bedienung von Geräten im Fahrzeug?	sehr [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Beurteilen Sie bitte wie gut die folgenden Funktionen mit Gesten zu bedienen waren			
Telefongespräch entgegennehmen	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Telefongespräch beenden	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Telefongespräch halten	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Titel (Sender) weiter, zurück	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Erhöhen, verringern der Lautstärke	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Karte links, rechts, hoch, runter verschieben	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Kartenzoom	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Geräteauswahl	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Auswahl eines Menüpunktes	sehr gut [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	überhaupt nicht [4]
Würden Sie in Ihrem eigenen Auto gerne Geräte mit Gesten bedienen?	ja [1]	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	nein [4]
Wieviel würden Sie für ein Gestenerkennungssystem im Auto bezahlen?		<input type="text"/>	
Versuchen Sie mit einigen Adjektiven zu beschreiben, wie Sie die Gestenbedienung empfanden?		<input type="text"/>	

[zurück nach oben](#)

Fertig Lokales Intranet

Bild 7: Beispiel eines HTML-Fragebogens.

HTML-Fragebogen (nicht sichtbar für die Versuchsperson) mit Vollständigkeitsüberprüfung. Da die Ergebnisse der Interviews in elektronischer Form vorliegen, wird wiederum die nachfolgende Auswertung sowie der Vergleich mit den objektiv erfassten Daten deutlich erleichtert.

9 Zusammenfassung

Während bei Usability-Studien üblicherweise der Anwender und die Benutzerfreundlichkeit eines Geräts im Mittelpunkt stehen, wird hier zusätzlich die Schnittstelle zwischen Versuchsleiter und Experiment optimiert. Auch ein Versuchsleiter muss zur Durchführung von Untersuchungen Geräte bedienen, deren Bedienoberfläche benutzerfreundlichen Kriterien genügen sollte.

Die vorgestellten Methoden werden am Lehrstuhl für Mensch-Maschine-Kommunikation der TU München gewinnbringend eingesetzt und erweisen sich als effizient bezüglich Vorbereitung, Durchführung, Auswertung und Archivierung von Usability-Experimenten. Die Effizienz bezieht sich hierbei sowohl auf die zeitsparende Durchführung von Untersuchungen, als auch auf die Qualität des gewonnenen Datenmaterials. Insbesondere die Untersuchung komplexer Fragestellungen, wie etwa aus dem Bereich der multimodalen Mensch-Maschine-Interaktion, wären ohne die

Entwicklung obiger Verfahren nicht oder nur sehr eingeschränkt möglich.

Eine Entlastung des Versuchsleiters resultiert zu großen Teilen aus der weitgehenden Automatisierung der Versuchsumgebung. Die Skriptsprache XCL erlaubt eine einfache und schnelle Programmierung von Versuchsabläufen. Durch die sprachgesteuerte Kontrolle mittels WISPER wird dem Versuchsleiter die Durchführung der Untersuchungen erleichtert. Er kann sich besser auf die wesentlichen Aufgaben einer Usability-Studie wie das Beobachten von Benutzeraktionen konzentrieren. Auch auf die Qualität des gewonnenen Datenmaterials wirkt sich die Verwendung von WISPER positiv aus, z. B. durch geringere Systemreaktionszeiten oder durch Vermeidung von Fehlerkennungen bei der Simulation von Erkennungssystemen.

Anfallende Daten, wie beispielsweise Bedienzeiten oder Kontrollbedingungen, werden weitgehend automatisch erfasst, wodurch sie in hohem Maße vergleichbar sind und nicht zuletzt die nachfolgende Auswertung vereinfacht und beschleunigt wird. Mit dem vorgestellten „Testreplayer“ lassen sich Versuchsabläufe automatisch rekonstruieren, wodurch ebenfalls die Auswertung erleichtert und die Zeit zur Anpassung und Weiterentwicklung des zu untersuchenden Systems deutlich reduziert wird. Die versuchsbegleitenden Fragebögen werden automatisch auf Vollständigkeit überprüft und liegen zur einfa-



chen Weiterverarbeitung in elektronischer Form vor.

Durch die Modularität der Versuchsumgebung wird ein hohes Maß an Flexibilität gewährleistet. Der Versuchsaufbau kann somit problemlos angepasst und erweitert werden. Ist für eine Untersuchung der Einsatz eines speziellen Moduls nicht erforderlich, können die übrigen Methoden weiterhin ohne Einbußen eingesetzt werden.

Die vorgestellten Aspekte effizienten Usability-Engineerings wurden bisher in Untersuchungen aus den Anwendungsbereichen Multimodalität, natürlichsprachliche Bedienung, Gestensteuerung, sowie Adaption von Geräten an Benutzer und Kontext eingesetzt. Sie bewährten sich sowohl bei der Entwicklung neuer, als auch in der Evaluierung bestehender Bedienoberflächen und -konzepte.

Literatur

- [1] *Sharma, R.; Pavlovic, C., und Huang, T. S.*: Toward Multimodal Human-Computer Interaction. Proc. of the IEEE, IEEE, 1998. Vol. 86, No 5, S. 853–869.
- [2] *Neuss, R.*: Usability Engineering als Ansatz zum multimodalen Mensch-Maschine Dialog. Dissertation, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, 2001.
- [3] *Zobl, M.; Geiger, M.; Bengler, K., und Lang, M.*: A Usability Study on Hand Gesture Controlled Operation of In-Car Devices. Proc. of the 9th Intl. Conf. on Human-Computer Interaction (HCI International 2001), New Orleans, Louisiana, USA, 5.–10. 8. 2001. Ed.: Lawrence Erlbaum Ass., New Jersey, 2001. „Poster Sessions: Abridged Proceedings“, S. 166–168.
- [4] *Geiger, M.; Nieschulz, R.; Zobl, M.; Neuss, R., und Lang, M.*: Methods for Facilitation of Wizard-of-Oz Studies and Data Acquisition. Proc. of the 9th Intl. Conf. on Human-Computer Interaction (HCI International 2001), New Orleans, Louisiana, USA, 5.–10. 8. 2001. Ed.: Lawrence Erlbaum Ass., New Jersey, 2001. „Poster Sessions: Abridged Proceedings“, S. 191–193.
- [5] *Nieschulz, R.; Geiger, M.; Bengler, K., und Lang, M.*: An Automatic, Adaptive Help System to Support Gestural Operation of an Automotive MMI. Proc. of the 9th Intl. Conf. on Human-Computer Interaction (HCI International 2001), New Orleans, Louisiana, USA, 5.–10. 8. 2001. Ed.: Lawrence Erlbaum Ass., New Jersey, 2001. Vol. 1: „Usability Evaluation and Interface Design“, S. 272–276.
- [6] *Balbo, S.*: Evaluation automatique des Interfaces Utilisateur: un pas vers l'automatisation. PhD thesis, University of Grenoble, 1994.



Dipl.-Ing. Ralf Nieschulz (2. v. r.), **Dipl.-Ing. Björn Schuller** (1. v. r.) und **Dipl.-Ing. Michael Geiger** (1. v. l.) sind als wiss. Assistenten bzw. Mitarbeiter am Lehrstuhl für Mensch-Maschine-Kommunikation der TU München tätig. Schwerpunkte ihrer Forschungstätigkeit sind die Benutzermodellierung zum adaptiven Mensch-Maschine-Dialog (Nieschulz), Methoden des Usability Engineerings, fehlerrobuste multimodale Mensch-Maschine-Kommunikation, Sprachverstehen und Emotionserkennung (Schuller), und die Entwicklung von Methoden zur berührungslosen Interaktion mit elektronischen Geräten im Fahrzeug (Geiger).

Adresse: Lehrstuhl für Mensch-Maschine-Kommunikation, TU München, D-80290 München,
E-Mail: {nieschulz, schuller, geiger}@ei.tum.de

Dr.-Ing. Robert Neuss (2. v. l.) fertigte seine Doktorarbeit am Lehrstuhl für Mensch-Maschine-Kommunikation der TU München [2] mit einem zwischenzeitlichen Forschungsaufenthalt am japanischen ATR in Kyoto an. Im April 2000 gründete er die Usaneers GmbH. Dieses Unternehmen entwickelt aufgabenorientierte, benutzerfreundliche sowie ansprechende Oberflächen unter Einsatz von multimodalen Technologien wie Sprach- oder Schrifterkennung mit Methoden des Usability-Engineerings.

Adresse: Usaneers GmbH, Paul-Lagarde-Str. 18, D-80686 München,
E-Mail: robert.neuss@usaneers.de