

ANALYSE KOGNITIVER
BENUTZERMODELLE FÜR DIE
EVALUATION VON
MENSCH-MASCHINE-SYSTEMEN

vorgelegt von
Diplom-Informatiker
Jeronimo A. B. D. Dzaack
aus Düsseldorf

von der Fakultät V
Verkehrs- und Maschinensysteme
der Technischen Universität Berlin

im DFG-Graduiertenkolleg
Prospektive Gestaltung von Mensch-Technik-Interaktion
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
— Dr.-Ing. —

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Manfred Thüring
Gutachter: Prof. Dr.-Ing. Leon Urbas
Gutachter: Prof. Dr.-Ing. Matthias Rötting

Tag der wissenschaftlichen Aussprache: 22. Februar 2008

Berlin 2008
D 83

ZUSAMMENFASSUNG

Die Entwicklung und Gestaltung von Schnittstellen für Mensch-Maschine-Systeme erfordern, neben der Kenntnis der Aufgabe und der Arbeitsumgebung, auch eine Vorstellung über die kognitiven Anforderungen an zukünftige Benutzer. Ein Grund dafür liegt in der zunehmenden Informationsdichte und dem ansteigenden Automatisierungsgrad von technischen Systemen. Daraus ergibt sich die Frage, welche Art von Schnittstelle die kognitiven Prozesse bei der Interaktion am besten unterstützt und wie diese prospektiv gestaltet werden kann. Die vorliegende Arbeit leistet einen Beitrag, um kognitive Aspekte bei fertigungs- und regelbasierten Tätigkeiten in der Mensch-Maschine-Interaktion in den Entwicklungsprozess zu integrieren.

Für die Analyse und Bewertung von Schnittstellen bietet sich in frühen Systementwicklungsphasen die kognitive Benutzermodellierung an, die eine computerbasierte Simulation des zukünftigen Benutzerverhaltens auf Basis formaler Spezifikationen erlaubt. Die Analyse und Interpretation der Simulationsdaten ermöglicht das Aufdecken von Fehlern in der Schnittstellengestaltung und dem Interaktionsdesign. Die kognitive Modellierung wird aber aufgrund des hohen Aufwands und fehlender Werkzeuge für die Modellentwicklung und Analyse von Simulationsdaten nur selten angewendet.

In dieser Arbeit wird überprüft, in welcher Art und Weise Simulationsdaten kognitiver Benutzermodelle für die Bewertung von Mensch-Maschine-Systemen in frühen Phasen der Systementwicklung analysiert und angewendet werden können. Dabei liegt der Fokus auf simulierten Blickbewegungsdaten der kognitiven Architektur ACT-R (*atomic components of thought – rational*) und empirisch erhobenen Blickbewegungsdaten für fertigungs- und regelbasierte Tätigkeiten. Die Blickbewegungsdaten stellen für die Bewertung von Mensch-Maschine-Schnittstellen einen wichtigen Faktor dar, um beispielsweise die visuelle Suche und die räumliche Gestaltung zu untersuchen.

Für die effiziente und effektive Verarbeitung und die vergleichende Analyse der simulierten und empirisch erhobenen Daten wurde das Werkzeug SimTrA (*Simulation Trace Analyzer*) entwickelt. SimTrA gibt dem Entwickler ein Methodenrepertoire an die Hand, um Schnittstellen bereits in frühen Systementwicklungsphasen formal und computerbasiert auf Basis von Simulationsdaten kognitiver Benutzermodelle zu bewerten. Die Arbeit stellt das Konzept, die Realisierung und die Anwendung von SimTrA in einer experimentellen Evaluationsstudie vor.

Die Ergebnisse der Evaluationsstudie belegen, dass kognitive Benutzermodelle in frühen Systementwicklungsphasen angewendet werden können und unter bestimmten Bedingungen mit empirischen Daten vergleichbare Aussagen ermöglichen. Des Weiteren konnte in der Arbeit gezeigt werden, dass die Analyse von Blickbewegungsdaten kognitiver Benutzermodelle durch SimTrA ermöglicht wird und die Bewertung von Mensch-Maschine-Systemen unterstützt. SimTrA stellt somit eine wertvolle Erweiterung für die kognitive Benutzermodellierung und für die prospektive Bewertung von Mensch-Maschine-Systemen dar.

ABSTRACT

Creating and designing interfaces for human-machine systems demands not only an understanding of the task and its environment, but also an appreciation of the capabilities and cognitive demands of the users. This becomes more and more important because of the growing information density and automation of human-machine systems. Regarding the environment, an important issue is what kind of interface best supports cognitive processes and how can an interface be designed to best match the future user requirements. The work presented here makes a contribution to integrate cognitive aspects of skill- and rule-based activities during human-machine interaction into the development process of human-machine systems.

In early stages of the development process cognitive modeling can be applied to evaluate human-machine interfaces. Using formal cognitive user models to simulate the future user behavior and requirements allows us to analyze the behavior on a detailed level. This helps to detect errors in the interaction design of interfaces and gives indications about the cognitive demands of the future users. But cognitive modeling is seldom used because of the high effort and a lack of tools for the development and analysis of cognitive user models.

In this dissertation the question is raised, in what way cognitive model data can be analyzed and used to evaluate human-machine systems in early design stages. The focus is directed to eye movement data simulated by cognitive user models in ACT-R (*atomic components of thought – rational*) and human eye movement data. In questions of usability, eye movement studies play an important role and are a reliable way to evaluate visual search and the spatial design of interfaces among other things

To process both the cognitive model data and the empirical data and to compare the outcomes in an effective and efficient manner the tool SimTrA (*Simulation Trace Analyzer*) was developed. The tool provides a repertoire of methods to evaluate human-machine interfaces in a formal and computerized way based on cognitive model data. The concept and the implementation of SimTrA and its application in an evaluation study are shown in this dissertation.

The results of the evaluation study show that cognitive user models in general can be applied in early stages of the development process of human-machine interfaces and under certain conditions provide statements comparable with empirical data. The study further indicates that the analysis of simulated eye movement data with SimTrA is possible and supports the evaluation of human-machine systems. Hence SimTrA is a useful tool for cognitive modeling and for prototype evaluation in early design phases of human-machine systems.

DANKSAGUNGEN

Ich bedanke mich bei allen, die mich im Rahmen dieser Doktorarbeit unterstützt haben. Mein besonderer Dank gilt Herrn Prof. Dr.-Ing. Leon Urbas für die freundliche und vertrauensvolle Betreuung und seine wertvolle Unterstützung. Ebenfalls gilt mein Dank Herrn Prof. Dr.-Ing. Matthias Röttig für die Übernahme des Korreferats und seine kritischen Anmerkungen und wertvollen Anregungen.

Ich danke der DFG (Deutsche Forschungsgemeinschaft), die diese Doktorarbeit durch ein Stipendium im Graduiertenkolleg prometei ermöglicht hat. Mein Dank gilt der Arbeitsgruppe MoDyS und dem ZMMS, die durch einen regen Austausch und durch die Schaffung von sehr guten Arbeits- und Forschungsbedingungen erheblich zum Gelingen dieser Arbeit beigetragen haben.

Die Arbeit entstand in einem interdisziplinären Rahmen der Arbeitsgruppe MoDyS und des Forschungsschwerpunktes 7 des Graduiertenkollegs prometei. Besonders möchte ich mich bei Sandra Trösterer bedanken, die mir in allen psychologischen Fragestellungen mit Rat und Tat zur Seite stand. Mein Dank gilt auch Marcus Heinath, der sich mit mir zusammen der kognitiven Benutzermodellierung gestellt hat und immer die nötige Ruhe besaß, mich auf den Boden der Tatsachen zurückzuholen. Ebenfalls möchte ich mich bei Jürgen Kiefer und bei Nele Pape bedanken, die mir viele Antworten auf kognitionspsychologische Fragen geben konnten.

Besonders möchte ich mich bei meinen Kollegen des Graduiertenkollegs prometei bedanken. Die Möglichkeiten zum interdisziplinären Austausch haben mir in vielen Situationen geholfen, neue Sichtweisen und Wege zu finden und auszuprobieren. Ich danke allen für den Austausch und die anregenden Gespräche, besonders Carmen Bruder, Monica de Filippis, Charlotte Glaser und Anne Klostermann für die wertvollen Diskussionen, die mir immer sehr viel Spaß gemacht haben.

Mein Dank gebührt ebenfalls der Vielzahl von Diplomanden, Studenten, Praktikanten und Versuchspersonen, die durch ihre Einsatzbereitschaft dazu beigetragen haben, dass das in dieser Arbeit vorgestellte Werkzeug implementiert, getestet und angewendet wurde. Im Besonderen danke ich dafür Jeanette Andrack, Daniel Döhring, Robert Lischke und Robert Seybt.

Für die gründliche Durchsicht und die vielen Anregungen und Hinweise danke ich Charlotte Glaser, Marcus Heinath, Anne Klostermann, Nele Pape, Sandra Trösterer und Ruth Wurster.

Ich danke meiner Familie und meinen Freunden, die durch die geduldige Unterstützung und durch das Bereitstellen von Freiräumen wesentlich zum Gelingen dieser Arbeit beigetragen haben.

INHALTSVERZEICHNIS

1	EINLEITUNG	1
1.1	Problem- und Zielstellung	2
1.2	Aufbau der Arbeit	3
2	THEORETISCHE GRUNDLAGEN	5
2.1	Mensch-Maschine-Systeme	5
2.1.1	Entwicklung von Mensch-Maschine-Systemen . .	6
2.1.2	Bewertung von Mensch-Maschine-Systemen . . .	7
2.1.3	Diskussion	10
2.2	Simulation kognitiver Prozesse	12
2.2.1	Kognitive Architekturen	13
2.2.2	Kognitive Modellierung	19
2.3	Analyse simulierter kognitiver Prozesse	22
2.3.1	Methoden zur Analyse von kognitiven Simulati- onsexperimenten	23
2.3.2	Auswahl einer Evaluationsmethode	24
2.4	Blickbewegungen	25
2.4.1	Blickbewegungen als Indikatoren für kognitive Prozesse	25
2.4.2	Parameter der Blickbewegung	27
2.4.3	Methoden der Blickbewegungserfassung	28
2.4.4	Simulation von Blickbewegungen	31
2.5	Ableiten der Fragestellung	33
3	SIMULATION TRACE ANALYZER	37
3.1	Voruntersuchung	37
3.1.1	Untersuchungsaufgabe	37
3.1.2	Untersuchungsgruppen	39
3.1.3	Untersuchungsdesign	39
3.1.4	Untersuchungsablauf	40
3.1.5	Geräte	40
3.1.6	Untersuchungsauswertung	40
3.1.7	Ergebnisse und Diskussion	41
3.2	Anforderungsanalyse	43
3.2.1	Datenintegration	43
3.2.2	Datenverarbeitung	43
3.2.3	Analyse- und Vergleichsmethoden	44
3.2.4	Datensichtung	45
3.3	Konzept	45
3.3.1	Prozess der Datenverarbeitung	47
3.3.2	Aufbereitung der Rohdaten	48
3.3.3	Analyse der Daten	49
3.3.4	Bereinigung der Daten	51
3.3.5	Vergleich der Daten	52
3.4	Realisierung des Werkzeugs SimTrA	52
3.4.1	Allgemeine Angaben	52
3.4.2	Aufbereitung der Rohdaten	53
3.4.3	Analyse der Daten	57
3.4.4	Bereinigung der Daten	60
3.4.5	Vergleich der Daten	61

3.5	Verifikation des Werkzeugs SimTrA	63
4	EXPERIMENTELLE STUDIEN	67
4.1	Experimentelle Hauptstudie	67
4.1.1	Aufgabe	67
4.1.2	Versuchsgruppen	68
4.1.3	Versuchsdesign	72
4.1.4	Versuchsablauf	73
4.1.5	Geräte	73
4.1.6	Versuchsauswertung	74
4.1.7	Experimentelle Hypothesen	75
4.1.8	Ergebnisse	76
4.1.9	Diskussion	87
4.2	Weitere experimentelle Studien	90
4.2.1	Blocks World	91
4.2.2	Vernetztes Fahren	92
4.2.3	Hierarchical Task Analysis Mapper	93
5	DISKUSSION	97
5.1	Die Bewertung von Mensch-Maschine-Systemen durch kognitive Benutzermodelle	97
5.2	Das Werkzeug SimTrA	98
5.3	Mögliche Anwendungsgebiete von SimTrA	100
6	ZUSAMMENFASSUNG UND AUSBLICK	103
6.1	Zusammenfassung	103
6.2	Ausblick	105
	LITERATURVERZEICHNIS	107

ABBILDUNGSVERZEICHNIS

Abbildung 1	Lebenszyklus eines Mensch-Maschine-Systems (nach Kraiss 1995). Zugeordnet sind die jeweiligen Methoden zur Evaluation der Benutzungsschnittstelle.	7
Abbildung 2	Die Informationsorganisation in ACT-R. Die Informationen in den Buffern, die mit den Modulen assoziiert sind, werden durch Produktionsregeln angefragt und geändert (nach Anderson et al., 2004).	18
Abbildung 3	Phasen der Modellkonstruktion (nach Wallach 1998, S. 40).	19
Abbildung 4	Relation zwischen natürlichem System S und künstlichem System S'. α und β stellen eine beobachtbare Abbildung der Input- und Outputrelation dar. Die Relation von S zu S' ist nur theoriegeleitet auf Grundlage von α und β zu interpretieren (nach Zinterhof 1987, S. 105ff.).	20
Abbildung 5	Schematische Darstellung der Empirie (links) und des Simulationsexperiments (rechts). P: Pegel, V: Regelventil, H: Heizung, GUI: Graphical User Interface, AGI: ACT-R Graphical Interface.	38
Abbildung 6	Relative Häufigkeiten der 6 wichtigsten lokalen Fixationspfade für Modell und Versuchspersonen mit Standardabweichung. H: Heizung, V: Regelventil, P: Pegel.	42
Abbildung 7	Schematische Darstellung der Komponenten von SimTrA.	46
Abbildung 8	Die Verarbeitungsschritte von SimTrA: Dargestellt ist der Informationsfluss und die jeweils abgelegten Dateien bei der Verarbeitung von drei Datensätzen. at: Tabelle der AOIs, et: Tabelle der Elemente der kognitiven Architektur, lt: Tabelle der ausgewählten Versuchspersonen, st: Tabelle der ausgewählten Zeitintervalle, tt: Tabelle der Blickbewegungsdaten, VPN: Versuchspersonen.	47
Abbildung 9	1: Komponentenmodell von SimTrA. 2: Die wichtigsten Dateien, die angelegt und abgerufen werden. 3: Verzeichnisstruktur von SimTrA.	53
Abbildung 10	Benutzungsoberfläche von SimTrA für die Datenaufbereitung.	54
Abbildung 11	Benutzungsoberfläche von SimTrA für die Auswahl von Analysemethoden.	57
Abbildung 12	Benutzungsoberfläche von SimTrA für die Bereinigung von Datensätzen.	60
Abbildung 13	Benutzungsoberfläche von SimTrA für den Vergleich von Datengruppen.	61
Abbildung 14	Übersicht der Testdaten.	64

Abbildung 15	Ansicht der Versuchsumgebung für die Kontroll- und Steuerungsaufgabe (H1202: Heizung, S1202: Regelventil).	68
Abbildung 16	Schematische Darstellung des internen Ablaufzykluses des Top-down-orientierten kognitiven Modells.	70
Abbildung 17	Schematische Darstellung des Bottom-up-orientierten kognitiven Modells.	71
Abbildung 18	Drei Schnittstellen für die Untersuchung mit den für die Analyse festgelegten AOIs (Szenario 1, Szenario 2 und Szenario 3).	72
Abbildung 19	Ansicht des Versuchsaufbaus für die Versuchspersonen.	74
Abbildung 20	Übersicht der betrachteten Faktoren für die Bewertung der Performanz der drei Versuchsgruppen. Dargestellt sind der Mittelwert und die Standardabweichung für jeden betrachteten Faktor für die drei Szenarien und zusammengefasst über die drei Szenarien.	76
Abbildung 21	Fixationspfade der funktional zusammengefassten AOIs für die drei Versuchsgruppen und die drei Szenarien mit Standardabweichungen.	81
Abbildung 22	Statistische Analyse der Übergangshäufigkeiten der drei Szenarien für die drei Versuchsgruppen. Dargestellt sind die H_c -Werte mit einer Standardabweichung für die funktional zusammengefassten AOIs (links) und die 3 bzw. 5 AOIs (rechts).	82
Abbildung 23	Die lokalen Fixationspfade des Bottom-up-Modells größer als 3 % und die zugeordneten lokalen Fixationspfade des Top-down-Modells für die drei untersuchten Szenarien.	84
Abbildung 24	Kumulierte Fixationsdauern für die drei Versuchsgruppen in den drei untersuchten Szenarien.	85
Abbildung 25	Mittlere Übergangsdichte für die drei Versuchsgruppen und die drei Szenarien. Grundlage für den Vergleich ist ein 10 x 10 Gitter, das über das Blickfeld gelegt wurde.	86
Abbildung 26	Die Blickverteilung von zwei unterschiedlichen Versuchspersonen während der Interaktion mit der Mensch-Maschine-Schnittstelle in Szenario 3.	88
Abbildung 27	Die Aufgabenumgebung von Blocks World (nach Gray et al. 2004) mit Blickbewegungsdaten einer Versuchsperson auf das <i>Target Window</i> (rechts). Ansicht aus SimTrA der fixierten AOIs und der nachfolgend platzierten farbigen Blöcke für eine Versuchsperson (links).	91
Abbildung 28	Ansicht der Blickbewegungsdaten einer Versuchsperson in SimTrA mit zwei AOIs (für Fahraufgabe und Nebenaufgabe).	92

Abbildung 29	Ansicht der Versuchsumgebung mit Blickbewegungen einer Versuchsperson bei der Interaktion mit der Abwasseraufbereitungsanlage in SimTrA. Eingezeichnet sind die AOIs für die 4 verschiedenen Teilprozesse der Aufgabe: Anlieferung, Homogenisierung, Trennung und Produktlager. . . .	94
Abbildung 30	Auszug aus den Analysen der Blickbewegungsdaten der Untersuchung mit einem automatisiert erstellten kognitivem Modell in HTAmap und empirischen Daten. Dargestellt sind die ungerichteten Transitionen (größer als 4 % für die empirischen Daten).	95

TABELLENVERZEICHNIS

Tabelle 1	Vergleich der drei kognitiven Architekturen ACT-R, Soar und EPIC (nach Byrne 2003).	16
Tabelle 2	Vorteile und Nachteile der kognitiven Modellbildung.	21
Tabelle 3	Parametersystematik der Verwendbarkeit von Parametern der Augen- und Blickbewegungen im Rahmen von arbeitswissenschaftlichen Aufgabenanalysen (Rötting 2001, S. 187ff.).	29
Tabelle 4	Performanz der Steuerungs- und Kontrollaufgabe der zwei Gruppen (kognitives Modell und Empirie) für die zwei Szenarien. Dargestellt sind die Mittelwerte für die Anzahl der Schaltungen der Heizung und des Regelventils und des Stabilisierungsfaktors.	41
Tabelle 5	Von SimTrA berechnete Werte der Analyse für die Testdaten im Bezug zu den manuell errechneten Werten.	65
Tabelle 6	Regeln für die Steuerung und Kontrolle der untersuchten Aufgabe. Der Pegel der Flüssigkeit im Tank ist in fünf Abschnitte eingeteilt: über der oberen Grenze, hoch, optimal, niedrig und unter der unteren Grenze. Die Heizung kann an- oder ausgeschaltet sein. Das Regelventil kann geöffnet oder geschlossen sein.	69
Tabelle 7	Performanz der Menschen für die drei Szenarien. Dargestellt sind die Mittelwerte und die Standardabweichung der betrachteten Faktoren.	77
Tabelle 8	Performanz des Bottom-up-Modells und des Top-down-Modells für die drei Szenarien. Dargestellt sind die Mittelwerte und die Standardabweichung der betrachteten Faktoren.	77

Tabelle 9	Durchschnittliche Performanz der drei Versuchsgruppen für die zusammengefassten Szenarien. Dargestellt sind die Mittelwerte und die Standardabweichung der betrachteten Faktoren.	78
Tabelle 10	Anzahl der Versuchspersonen und Mittelwerte und Standardabweichung der allgemeinen Informationen der Blickbewegungsdaten für jede Versuchsgruppe und die drei Szenarien.	80

EINLEITUNG

In der Entwicklung von Mensch-Maschine-Systemen stellen die Gestaltung, die Analyse und die Optimierung der Benutzungsoberflächen und der zugrunde liegenden technischen Prozesse eine zentrale Aufgabe dar. Sie wäre ohne Unterstützungswerkzeuge kaum zu lösen. In vielen Entwicklungsbereichen werden schon seit langem Modelle und Simulationen eingesetzt (wie z. B. Strömungsmodelle, anthropometrische Modelle), die Prozesse und Wechselwirkungen im Anwendungskontext nachbilden und abstrahieren. Das Ziel ist eine adäquate Unterstützung der Entwicklungs- und Entscheidungsprozesse.

Die erhöhte Automatisierung und die Integration komplexer technischer Prozesse in komplexe Mensch-Maschine-Systeme hat eine Verlagerung der Handlungsebene der menschlichen Tätigkeiten zur Folge (Rasmussen 1982). Es ist eine Verschiebung von rein fertigkeitsbasierten sensumotorischen Tätigkeiten hin zu überwiegend wissensbasierten und kognitiven Tätigkeiten festzustellen. Der Mensch stellt dementsprechend in modernen Mensch-Maschine-Systemen einen Anwender dar, dessen primäre Aufgabe es ist, Prozesse zu überwachen und gegebenenfalls basierend auf Kausalwissen und mentalen Modellen regulierend einzugreifen.

„In Streßsituationen, wenn die Automatisierung von Prozessen den Menschen nicht entlasten kann, ist die wissensbasierte Verarbeitung jedoch im allgemeinen zu langsam. Hier ist der Rückgriff auf schneller ausführbare Handlungsebenen (Regel- und/oder Gewohnheitsebene) notwendig“ (Urbas 1999, S. 5). Das heißt, dass die fertigkeits- und regelbasierten Tätigkeiten bei der Gestaltung von Mensch-Maschine-Systemen neben der Automatisierung von technischen Prozessen einen wichtigen Aspekt ausmachen. Diese Tätigkeiten beruhen auf kognitiven Prozessen des Menschen und werden durch die Arbeitsumgebung und die Aufgabe bestimmt.

Moderne Mensch-Maschine-Systeme stellen eine Vielzahl von graphischen Anzeigen und Interaktionselementen zu Verfügung, um die Aufgabenbearbeitung zu unterstützen. Diese sind vornehmlich durch visuelle Informationsaufnahme und -verarbeitung sowie durch die Koordination von Aufgaben charakterisiert. Somit stellen moderne Mensch-Maschine-Systeme erhöhte Anforderungen an die menschliche Kognition. Daraus ergibt sich die Notwendigkeit, die Gebrauchstauglichkeit von Mensch-Maschine-Systemen hinsichtlich der Möglichkeiten und Grenzen der menschlichen Kognition in den verschiedenen Phasen des Entwicklungsprozesses von Mensch-Maschine-Systemen zu bewerten und die Schnittstellengestaltung an diese anzupassen.

Neben den etablierten Modellierungs- und Simulationsmethoden, die Eingang in den Entwicklungsprozess von Mensch-Maschine-Systemen gefunden haben, stellt die Methode der kognitiven Modellbildung eine Möglichkeit dar, die menschliche Kognition bei fertigkeits- und regelbasierten Tätigkeiten abzubilden und je nach Anwendungsfall zu abstrahieren. Kognitive Modellierungsmethoden basieren auf kogni-

tionspsychologischen Theorien und integrieren das Wissen über Grenzen und Möglichkeiten der menschlichen Kognition. Anhand kognitiver Modelle kann das menschliche Verhalten simuliert und analysiert werden. Im Kontext von Mensch-Maschine-Interaktion sind kognitive Modelle in der Lage, das Interaktionsverhalten von Benutzern formal und computerbasiert zu simulieren. Die Analyse und Interpretation der Simulationsdaten, die bei der Interaktion kognitiver Modelle mit Mensch-Maschine-Systemen generiert werden, ermöglichen das Aufdecken von Fehlern in der Schnittstellengestaltung und im Interaktionsdesign auf Basis von kognitiven Theorien (z. B. zu Wahrnehmung, Denken, Motorik, Lernen). Somit können kognitive Modellierungsmethoden in der Systementwicklung eingesetzt werden, um das Interaktionsdesign und die Schnittstellengestaltung zu bewerten und an die Analysen und Interpretationen anzupassen. Der Einsatz kognitiver Benutzermodelle in dem Entwicklungsprozess von Mensch-Maschine-Systemen erlaubt die Anpassung von Schnittstellen an die menschlichen Grenzen und Möglichkeiten, die im Rahmen von fertigkeit- und regelbasierten Tätigkeiten auftreten.

Besonders in frühen Phasen der Systementwicklung können die formalen und quantitativen Computermodelle eingesetzt werden, um alternative Gestaltungsvarianten kostengünstig und einfach zu bewerten. Die Methode der kognitiven Modellierung ermöglicht daher die Einbeziehung von Erkenntnissen über die menschliche Kognition von Beginn der Entwicklung an.

1.1 PROBLEM- UND ZIELSTELLUNG

In einer Interviewstudie im Rahmen der Forschung zu dieser Arbeit konnte gezeigt werden, dass das Potential der kognitiven Modellierung in Forschung und Entwicklung für die Bewertung von Mensch-Maschine-Systemen bekannt ist und vereinzelt eingesetzt wird (Urbas, Dubrowsky, Dzaack & Heinath 2005). Es besteht der Wunsch, kognitive Benutzermodelle für die Bewertung von Mensch-Maschine-Systemen im Entwicklungsprozess von Mensch-Maschine-Systemen einzusetzen. Es zeigte sich aber, dass die Methode aufgrund fehlender Beispiele und Werkzeuge für Modellentwicklung und Analyse der Simulationsdaten nur selten angewendet wird.

Der Prozess der Modellbildung, Simulation und Analyse der Simulationsdaten kann als zusammenhängende Prozesskette betrachtet werden (Urbas, Schulze-Kissing & Leuchter 2005). In jedem dieser Schritte liegen Hürden und Schwierigkeiten, die den Einsatz von kognitiven Benutzermodellen behindern können. Dazu zählen die komplexe Modellentwicklung, die Anbindung von kognitiven Benutzermodellen an reale oder virtuelle Mensch-Maschine-Schnittstellen und die Analyse der simulierten Verhaltensdaten der kognitiven Modelle. Für die Übertragung der kognitiven Modellierung in den realen Anwendungskontext gilt es, diese Hürden zu überwinden. Hier wird die Analyse der Simulationsdaten kritisch betrachtet und deren Einsatz für die Bewertung von Mensch-Maschine-Systemen in prospektiven Phasen der Systemgestaltung überprüft und prototypisch umgesetzt.

Das Ziel ist die Entwicklung eines Werkzeugs für die Analyse und den Vergleich von simulierten und empirischen Leistungsdaten, die bei der Interaktion von Menschen und kognitiven Modellen mit Mensch-Maschine-Systemen erfasst werden können. Die Anwendbarkeit kognitiver Modellierungsmethoden und des prototypischen Werkzeuges wird an einem Anwendungsbeispiel gezeigt. In diesem Zusammenhang soll geprüft werden, inwieweit die Analyse von Simulationsdaten heutiger kognitiver Modelle die Bewertung von Mensch-Maschine-Systemen ermöglicht und auf welcher Ebene die kognitiven Simulationsdaten betrachtet werden können. Des Weiteren sollen das Potential und die Grenzen kognitiver Modellierungsmethoden aufgezeigt und betrachtet werden.

1.2 AUFBAU DER ARBEIT

In Kapitel 2 werden die theoretischen Grundlagen der Arbeit dargestellt. Die Evaluation von dynamischen und komplexen Mensch-Maschine-Systemen, kognitive Modellierungsmethoden und Parameter für die Blickbewegungsmessung werden diskutiert. Zusammenfassend wird die Anwendung kognitiver Modellierungsmethoden für die Evaluation von Mensch-Maschine-Systemen gezeigt. Abschließend wird die der Arbeit zugrunde liegende Fragestellung abgeleitet.

Darauf aufbauend wird in Kapitel 3 die konzeptionelle Planung und die praktische Umsetzung des *Simulation Trace Analyzer* (SimTrA) für die Analyse und den Vergleich von simulierten und empirisch erfassten Blickbewegungsdaten dargestellt.

In Kapitel 4 wird die Praktikabilität von SimTrA beispielhaft an einer experimentellen Studie zur Schnittstellengestaltung von Mensch-Maschine-Systemen unter Verwendung von simulierten und empirisch erhobenen Blickbewegungsdaten gezeigt. Das Untersuchungsdesign der Studie wird eingeführt und anschließend werden die Ergebnisse diskutiert. Das Kapitel schließt mit einer Beschreibung von weiteren Studien, in denen SimTrA für die Analyse von Blickbewegungsdaten eingesetzt wurde.

Kapitel 5 diskutiert das entwickelte Werkzeug SimTrA im Kontext der Bewertung von Mensch-Maschine-Systemen in frühen Phasen der Systementwicklung.

Den Abschluss bildet eine Zusammenfassung der Arbeit und ein Ausblick auf die kognitive Modellbildung im Kontext der Bewertung von Mensch-Maschine-Systemen.

In diesem Kapitel werden die theoretischen Grundlagen für die Arbeit gelegt. Die Evaluation von dynamischen und komplexen Mensch-Maschine-Systemen und die Phasen der Systementwicklung werden diskutiert. Dabei liegt der Fokus auf der prospektiven Evaluation von Benutzungsoberflächen von Mensch-Maschine-Systemen. Anschließend werden die in den einzelnen Lebensphasen solcher Systeme anwendbaren Klassen von Methoden für die Evaluation der Benutzungsoberfläche vorgestellt. Im Folgenden wird die Simulation kognitiver Prozesse unter Verwendung kognitiver Architekturen eingeführt und deren Vorteile und Nachteile dargestellt. Dabei wird die Analyse simulierter kognitiver Prozesse besonders betrachtet und der aktuelle Stand der Forschung aufgezeigt. Die für diese Arbeit und die Bewertung von Mensch-Maschine-Systemen grundlegende Methode der Blickbewegungsanalyse wird eingeführt. Eine Übertragung auf das Gebiet kognitiver Simulationsexperimente wird gezeigt. Zusammenfassend wird der Gebrauch simulierter kognitiver Prozesse für die Evaluation von Mensch-Maschine-Systemen dargelegt. Abschließend wird die der Arbeit zugrunde liegende Fragestellung abgeleitet.

2.1 MENSCH-MASCHINE-SYSTEME

Im Allgemeinen werden technische Systeme, zu deren Funktion das Zusammenwirken mindestens eines Menschen mit einem technischen System notwendig ist, als Mensch-Maschine-Systeme bezeichnet (z. B. Kraftfahrzeuge, Navigationssysteme, fertigungstechnische Anlagen). Nach Timpe und Kolrep (2002) bezeichnet der Begriff Mensch-Maschine-Systeme „eine zweckmäßige Abstraktion des zielgerichteten Zusammenwirkens von Personen mit technischen Systemen zur Erfüllung eines fremd- oder selbstgestellten Auftrags“ (S. 12). Eine solche Abstraktion erlaubt die Betrachtung des zielgerichteten Informationsaustausches zwischen Mensch und Maschine in unterschiedlichen Situationen. Dabei wird die Struktur des Mensch-Maschine-Systems als rückgekoppeltes System verstanden. Der Mensch (oder eine Gruppe von Menschen) kann auf der Grundlage der organisationalen Verankerung, seiner Zielstellung, des Auftrags und der wahrgenommenen Rückmeldungen des technischen Systems Entscheidungen treffen und das System steuern.

Technische Systeme sind zumeist komplexe Systeme, die in ihrer Dynamik durch Nichtlinearität und Emergenz gekennzeichnet sind (Wickens & Holland 2000). Das heißt, dass komplexe Systeme nicht zu vereinfachen sind, ein in sich vielschichtiges System bleiben und dadurch das Handeln des Menschen beeinflussen.

Dörner (1995) stellt folgende Merkmale heraus, die den Menschen beim Treffen einer Entscheidung in komplexen Handlungssituationen beeinflussen und behindern können: Komplexität, Dynamik, Intranspa-

renz, Vernetztheit und Unvollständigkeit oder Falschheit der Kenntnisse über das System. Die fünf Merkmale können auf die Interaktion von Menschen mit technischen Systemen übertragen werden. Technische Systeme bestehen zumeist aus vielen vernetzten Subsystemen, die über Schnittstellen miteinander verbunden oder gekoppelt sind. Die Systeme sind intransparent, das heißt nicht alle Informationen über den Prozess sind immer einsehbar. Weiterhin ist das Systemverhalten nicht stetig, da das System sich aufgrund der Eigendynamik selbstständig weiterentwickelt. Schließlich hat ein Anwender zumeist keine vollständige Kenntnis über die Gesamtheit der Systemeigenschaften. Bei genügend großen Mensch-Maschine-Systemen sind die fünf Merkmale erfüllt, so dass von komplexen und dynamischen Mensch-Maschine-Systemen gesprochen werden kann.

Das Verhalten eines Mensch-Maschine-Systems wird durch die Informationsverarbeitung des Menschen, beruhend auf sozialen, psychologischen und biologischen Gesetzmäßigkeiten, und durch technische Prozesse, beruhend auf physikalischen Gesetzen, bestimmt (Klix 1966). Die menschliche Informationsverarbeitung wird im Besonderen durch die dargebotenen Informationen des technischen Systems und die kognitiven Prozesse des Menschen moderiert. Bei der Gestaltung von Mensch-Maschine-Systemen müssen daher beide Aspekte berücksichtigt werden. Die Benutzungsaspekte technischer Systeme werden aber heute zumeist durch die technischen Parameter eines Systems determiniert und bilden technische Strukturen ab (z. B. Anzeigen mit vielen Variablen zur Prozesssteuerung). Um eine menschengerechte Interaktion zu ermöglichen, stellt die Berücksichtigung menschlicher kognitiver Prozesse und deren Beschränkungen im Zusammenhang mit der Aufgabe und der Aufgabenumgebung eine wichtige Betrachtungsweise bei der Entwicklung von Mensch-Maschine-Systemen dar (Byrne 2007).

2.1.1 *Entwicklung von Mensch-Maschine-Systemen*

In der System- und Softwaretechnik wird der Werdegang eines Systems in einzelne voneinander abgrenzbare Lebensphasen unterteilt, die in Phasenmodellen dargestellt werden können (Kraiss 1995; Timpe & Kolrep 2002). Ein Phasenmodell stellt ein idealtypisches Vorgehen dar und beschreibt daher nicht notwendigerweise den in der Realität zu findenden Prozessablauf. Die Modelle definieren einen festgelegten organisatorischen Rahmen und erlauben die Integration wissenschaftlicher Erkenntnisse für die wirtschaftliche Herstellung und den wirtschaftlichen Einsatz technischer Systeme (Balzert 1998; Pomberger & Blaschek 1996). Dabei werden die Reihenfolge der Arbeitsabläufe, die jeweils durchzuführenden Aktivitäten, die Teilergebnisse und die Fertigstellungskriterien definiert. Die Ergebnisse jeder einzelnen Phase bilden die Grundlage der darauf folgenden Phase (Kaindl, Lutz & Trippold 1998). Somit bestimmt die Qualität der Teilergebnisse jeder einzelnen Phase die Qualität eines fertigen Systems. Es existiert eine Reihe von unterschiedlichen Modellformen für die Entwicklung von technischen Systemen, zum Beispiel das Spiralmodell, das V-Modell und das Prototypenmodell. Für einen Überblick siehe Balzert (1998) oder Kaindl et al. (1998).

Ein einfaches Phasenmodell ist in Abbildung 1 dargestellt (Kraiss 1995). Hauptbestandteile dieses Phasenmodells sind die Entwicklung, die Realisierung und die Nutzung eines technischen Systems. Der Entwicklungsprozess beginnt mit der Definition der Funktionen, die das zu entwickelnde System abbilden soll, basierend auf einer konzeptionellen Anforderungsanalyse. Daraus werden die Spezifikationen für den Systementwurf abgeleitet und detaillierte Anforderungen für die (Teil-)Systeme festgelegt. In der Realisierungsphase werden die (Teil-)Systeme technisch implementiert und zu einem Gesamtsystem integriert. Die abschließende Nutzungsphase stellt den ordnungsgemäßen Gebrauch bis zur Außerbetriebnahme des technischen Systems dar.

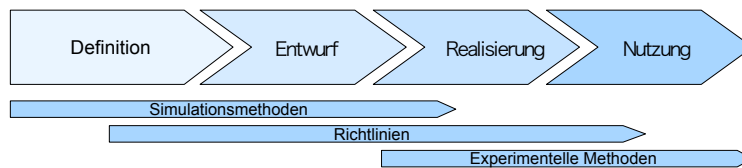


Abbildung 1: Lebenszyklus eines Mensch-Maschine-Systems (nach Kraiss 1995). Zugeordnet sind die jeweiligen Methoden zur Evaluation der Benutzungsschnittstelle.

2.1.2 Bewertung von Mensch-Maschine-Systemen

Die gebrauchstaugliche Gestaltung der Benutzungsoberfläche und der dargebotenen Organisationseinheiten der Komponente Maschine stellt in der Systementwicklung von Mensch-Maschine-Systemen einen wichtigen Bestandteil dar. In den einzelnen Phasen bieten sich unterschiedliche Methoden für die Evaluation von Mensch-Maschine-Systemen an, wobei technische, ökonomische, ökologische und ergonomische Kriterien berücksichtigt werden müssen. Die ergonomischen Kriterien beziehen sich in einem Mensch-Maschine-System auf die unmittelbare Interaktion von Mensch und Maschine, das heißt auf die für die Bedienung des technischen Systems gestalteten Elemente (die Benutzungsschnittstelle oder die Mensch-Maschine-Schnittstelle). Bei der Gestaltung der technischen Komponente stellen die Berücksichtigung der menschlichen Leistungsfähigkeiten und -grenzen, die Optimierung der Beanspruchung und die sichere und verlässliche Handhabung wichtige Faktoren dar (Baggen & Hemmerling 2002). In dynamischen Mensch-Maschine-Systemen sind somit die Prozesse der Informationsaufnahme, -verarbeitung und -abgabe bei der Interaktion von Mensch und Maschine besonders zu unterstützen.

Bei der benutzerfreundlichen Gestaltung von technischen Systemen sind im Wesentlichen drei Aspekte zu beachten. Nach DIN EN ISO 9241-11 (Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten – Teil 11: Anforderungen an die Gebrauchstauglichkeit; DIN EN ISO 9241-11 1998) sind dies die Effektivität, die Effizienz und die Zufriedenstellung durch die Nutzung eines technischen Systems.

Die Effektivität beschreibt die Vollständigkeit und Genauigkeit, mit der ein Benutzer ein bestimmtes (Teil-)Ziel erreicht. Die Effizienz stellt das Verhältnis der Effektivität zu dem zur Zielerreichung benötigten Aufwand dar. Die Zufriedenstellung beschreibt die Freiheit von Beeinträchtigungen und die Akzeptanz der Benutzer. Ausgehend von diesen Kriterien können Maße für die Benutzungsfreundlichkeit ermittelt werden. Gebräuchliche Maße für die Effektivität sind zum Beispiel Fehlerraten und die Anzahl gelernter Funktionen. Geeignete Maße für die Effizienz stellen Ausführungszeiten, Lernzeiten und die Anzahl der Dialogschritte dar. Die Zufriedenheit kann durch die Erhebung der Akzeptanz oder der Benutzungshäufigkeit gemessen werden (Marrenbach 2001; DIN EN ISO 9241-11 1998).

Für die Evaluation der Benutzerfreundlichkeit eines Systems bieten sich Simulationsmethoden, Richtlinien und experimentelle Methoden an (Freed, Shafto & Remington 1998). In Abbildung 1 sind diese Methoden den einzelnen Lebensphasen eines Mensch-Maschine-Systems zugeordnet.

Simulationsmethoden

Simulationsmethoden basieren auf formalen Modellen des Benutzers beziehungsweise dessen Verhalten und stellen eine quantitative Beschreibung des Handlungsablaufs dar. Diese Methoden sind besonders für die Abbildung von fertigungs- und regelbasiertem Verhalten geeignet. Die mit formalen Ansätzen beschriebenen idealtypischen Modelle simulieren für die Bedienung von technischen Systemen objektive Leistungsdaten, wie beispielsweise Ausführungszeiten, Fehlerraten und die Anzahl von Interaktionen bei der Bedienung. Die Ermittlung von Schwachstellen in der Gestaltung von Mensch-Maschine-Systemen und die Evaluation verschiedener Gestaltungsalternativen werden somit ermöglicht. Anhand der simulierten Leistungsdaten kann die Effizienz und die Effektivität eines Systems bewertet werden. Aufgrund der Nutzung von Systemspezifikationen und Annahmen über das Verhalten des Menschen als Grundlage für die Entwicklung von formalen Simulationsmodellen sind diese bereits in frühen Systementwicklungsphasen einsetzbar. Aspekte wie Individualität, Adaptivität und Multitasking, die wichtige Eigenschaften des menschlichen Verhaltens sind, können mit heutigen quantitativen Methoden nur schwer abgebildet werden (Rasmussen 1983). Ansätze für die Integration von individuellem Verhalten zum Beispiel durch unterschiedliche Strategien oder Wissensbasen (Taatgen 2001) und Multitasking werden heute erprobt (Salvucci 2005) und sind exemplarisch durch Auswahlprozesse oder Zufallskomponenten in formalen Modellierungsmethoden implementiert.

Es existiert eine Vielzahl von Methoden zur Simulation von menschlichem Benutzerverhalten, die in zwei Klassen eingeteilt werden können (Salvucci & Lee 2003).

HIGH-LEVEL-ANSÄTZE High-level-Ansätze beschreiben das Verhalten auf einer aggregierten Ebene und definieren Interaktionen des Benutzers mit einem technischen System als statische Sequenz menschlicher Handlungen. High-level-Ansätze sind geeignet, Fehler und Schwierigkeiten in der Gestaltung festzustellen, die auf sequenziellem Verhal-

ten beruhen (z. B. Handeln im Störfall, Abarbeiten einer Handlungsanweisung). Beispiele für diese Methoden sind GOMS (*Goals, Operators, Methods and Selection Rules*; Card, Moran & Newell 1983) und TAG (*Task Action Grammar*; Payne & Green 1986).

LOW-LEVEL-ANSÄTZE Low-level-Ansätze sind im Gegensatz dazu geeignet, komplexe Phänomene (wie z. B. Entscheiden, Mehrfachaufgaben und Lernen) zu beschreiben und abzubilden. Sie beschreiben das menschliche Verhalten auf einer atomaren Ebene und erlauben eine genauere Betrachtung der menschlichen Kognition als High-level-Ansätze. Beispiele dafür sind ACT-R (*atomic components of thought – rational*; Anderson et al. 2004), EPIC (*Executive-Process/Interactive Control*; Kieras & Polson 1988) und Soar (*State, Operator and Result*; Newell 1990).

Richtlinien

In Richtlinien wird Expertenwissen erfasst und Entwicklern und Konstrukteuren von technischen Systemen bereitgestellt. Das Expertenwissen beruht zumeist auf subjektiven Erfahrungen der Experten und wird in verdichteter Form zusammengefasst. Daraus resultieren allgemein gehaltene Richtlinien, die für die Anwendung auf einen konkreten Fall eine Bandbreite von Interpretationen zulassen. Beispiele sind die Richtlinien für die ergonomische Gestaltung von Bürotätigkeiten mit Bildschirmgeräten (DIN EN ISO 9241-11 1998) und die Richtlinien für die barrierefreie Gestaltung von multimedialen Angeboten (BITV 2002). Auf Grundlage der Richtlinien können kriterienorientierte Prüfsysteme entwickelt werden, die eine subjektive Bewertung des Systemzustandes ermöglichen (z. B. IsoMetrics; Willumeit, Gediga & Hamborg 1996). Eine Gefahr in der Anwendung von Richtlinien besteht darin, dass ein technisches System durch die Entwickler an spezielle Richtlinien und deren Prüfmethoden angepasst wird und somit für die Bearbeitung von einigen wenigen Arbeiten optimal gestaltet wurde und nicht immer generalisierbar sind (Greif & Grediga 1988). Richtlinien können in der Systementwicklung durchgängig eingesetzt werden und unterstützen die Bewertung der Effizienz, Effektivität und durch heuristische Aussagen vereinzelt auch die der Benutzerzufriedenheit (Henninger 2001).

Experimentelle Methoden

Bei experimentellen Methoden werden anhand von ablauffähigen Versionen des zukünftigen Produktes (Prototypen) die Gestaltung und die Funktion eines Systems bewertet (Balzert 1998). Die Bewertung der Ergonomie stützt sich dabei auf objektive und subjektive Maße, die während und nach der experimentellen Bewertung erhoben werden können (DIN EN ISO 9241-11 1998). Objektive Bewertungsmethoden sind beispielsweise Handlungsprotokolle, Blickbewegungsmessungen und Videoaufzeichnungen. Subjektive Methoden beziehen sich auf die Erfahrungen der Benutzer. Beispiele dafür sind Fragebogen- oder Interviewmethoden. Das Ziel der experimentellen Methoden liegt in der Erhebung von Merkmalen und Hinweisen über das zukünftige Benutzerverhalten durch die Produktgestaltung (Weimer 1995). Für die

experimentelle Bewertung werden Prototypen benötigt, die erst in späteren Phasen der Systementwicklung zur Verfügung stehen. Neben der Erstellung von Prototypen ist die Durchführung von experimentellen Methoden zeit- und kostenintensiv (z. B. Kosten der Versuchspersonen, Experimentalplanung). Experimentelle Methoden erlauben neben der Bewertung der Effizienz und der Effektivität eines technischen Systems auch die Bewertung der Benutzerzufriedenheit.

2.1.3 Diskussion

Die zunehmende Komplexität von Mensch-Maschine-Systemen und die damit steigenden Anforderungen in der Bedienung verlangen eine entsprechende Berücksichtigung der menschlichen Komponente im Entwicklungsprozess, das heißt der kognitiven Prozesse mit deren Grenzen und Ansprüchen (vgl. Timpe, Jürgensohn & Kolrep 2002). Beispielsweise führen unzureichende Gestaltungslösungen von Mensch-Maschine-Systemen häufig zu Handlungsfehlern, die auf einer Entkopplung des Anwenders vom Prozess (*out of loop*) und damit auf dem Fehlen von Informationen beruhen oder auf Wahrnehmungsüberforderung aufgrund der Überführung von nicht visuellen Informationen in eine visuelle Darstellung in der Benutzungsoberfläche zurückzuführen sind (Hedecke 2002).

Eine Vielzahl von Problemen in der Bedienung von technischen Systemen sind auf die unterschiedlichen Sichtweisen der Entwickler und der Anwender auf ein technisches System zurückzuführen (Norman 1986). Gründe dafür sind beispielsweise unterschiedliches Systemwissen, Kausalwissen und Interaktionswissen der Entwickler und Anwender. Infolgedessen muss eine Bewertungsmethode objektive Maße zu Verfügung stellen, die schon frühzeitig als Kommunikationshilfe zwischen Entwickler und Benutzer dienen können. Besonders die prospektive Bewertung von Mensch-Maschine-Systemen ist im Gegensatz zu der retrospektiven Bewertung und der Bewertung in späteren Systemlebensphasen für das Aufdecken von Gestaltungsfehlern des zukünftigen Produktes von Bedeutung. In frühen Entwicklungsphasen können Fehler mit erheblich geringerem Aufwand behoben und Gestaltungsalternativen exploriert werden (Holzinger 2005; Reinhart, Lindemann & Heinzl 1996). Beispielsweise können Kosten für eine Nachbesserung am fertigen Produkt bis um das Zehntausendfache höher sein als in frühen Entwicklungsphasen (Freise 2003).

Folglich ergeben sich spezielle Anforderungen an die Entwicklung von Mensch-Maschine-Systemen. Zum einen muss der Faktor Mensch mit dessen kognitiven Anforderungen und Grenzen systematisch in die Bewertungsprozesse integriert werden. Zum anderen müssen die Bewertungsmethoden in frühen Phasen der Systementwicklung einsetzbar sein und objektive Bewertungen zulassen. So können die Bewertungsergebnisse von Anfang an als Kommunikationshilfe zwischen Entwickler und Benutzer genutzt werden.

Richtlinien und experimentelle Methoden werden in dem Entwicklungsprozess von Mensch-Maschine-Systemen für die Bewertung der Gebrauchstauglichkeit heute standardmäßig eingesetzt. Formale Methoden der Simulation und Modellierung von Benutzerverhalten kommen

nur selten zum Einsatz (Kindsmüller, Leuchter, Schulze-Kissing & Urbas 2004).

Richtlinien ermöglichen eine Prognose der Benutzerfreundlichkeit bereits in der Entwurfsphase von Mensch-Maschine-Systemen und stellen ein wichtiges Repertoire von Prüflisten, Heuristiken und Kriterien bereit, die eine Bewertung der Schnittstelle ermöglichen. Experimentelle Methoden beruhen auf empirisch erhobenen Daten und ermöglichen eine Bewertung der Schnittstelle auf der Basis von ablauffähigen Prototypen und können daher oft erst im Anschluss an die Entwurfsphase eingesetzt werden.

Beide Verfahren stellen keine ausreichenden Methoden bereit, um das Interaktionsdesign und die Schnittstellengestaltung im Kontext der beteiligten kognitiven Prozesse adäquat bewerten zu können. Diese sind für das menschliche Verhalten von großer Bedeutung und werden durch die Gestaltung der Arbeitsumgebung und der Aufgabe beeinflusst (Byrne 2007). Daher muss der Einfluss der Gestaltung einer Mensch-Maschine-Schnittstelle auf die menschliche Kognition bei der Entwicklung von Mensch-Maschine-Systemen verstärkt von Beginn an betrachtet werden.

In frühen Phasen der Systementwicklung können formale Simulationsmethoden eingesetzt werden. Kognitive Modellierungsmethoden können den Entwicklern frühzeitig das zukünftige Benutzerverhalten kommunizieren und erlauben eine computerbasierte Betrachtung des Benutzerverhaltens bei der Interaktion mit technischen Systemen. Im Gegensatz zu experimentellen Methoden und dem Einsatz von Richtlinien werden durch formale Methoden objektive Daten bereitgestellt.

Hinsichtlich der Anforderungen der DIN EN ISO-Norm 9241-11 können alle drei methodischen Ansätze (d. h. Simulationsmethoden, Richtlinien und experimentelle Methoden) die Effektivität und die Effizienz eines Gestaltungsansatzes messen. Die Zufriedenheit kann anhand von experimentellen Methoden und basierend auf heuristischen Maßen durch den Einsatz von Richtlinien gemessen werden.

Es ist notwendig, dass objektive Bewertungsmaße von Anfang an in den Systementwicklungsprozess integriert werden. Diese Forderung ist durch den Einsatz formaler Simulationsmethoden erfüllbar. Für den Einsatz in zivilen Bereichen, wie der Evaluation von Gebrauchselektronik, Leitsystemen oder Software, fehlt heute ein Methodenrepertoire, das die Erstellung von Benutzermodellen und die Analyse der simulierten Leistungsdaten ermöglicht (Kraiss 1995; Kindsmüller et al. 2004). Erste Schritte für eine durchgängige Unterstützung der Benutzermodellierung im Rahmen der Systementwicklung werden in Urbas, Schulze-Kissing und Leuchter (2005) beschrieben.

Aufgrund der Möglichkeit, formale Simulationsmethoden bereits in frühen Phasen der Systemgestaltung einzusetzen und der durch diese Methoden erfassbaren objektiven Daten, wird in dieser Arbeit der Ansatz der kognitiven Benutzersimulation für die Bewertung von Mensch-Maschine-Systemen betrachtet.

2.2 SIMULATION KOGNITIVER PROZESSE

Die Simulation kognitiver Prozesse bedeutet die Konstruktion von Modellen, die ausgewählte Eigenschaften der zu simulierenden mentalen Prozesse abbilden und über diese mittels angenommener Strukturen operieren (Kieras 1985). Diese Methode kann als kognitive Modellierung bezeichnet werden und stellt einen wichtigen Aspekt der Theoriebildung und Plausibilisierung von Theorien dar (Cooper 2002; Opwis 1992; Wallach 1998). Ein Modell stellt eine geeignete Abstraktion der Realität dar, die anstelle des Originals genutzt werden und als Kommunikationsgrundlage dienen kann (Strube 2001). Die Abstraktion erlaubt die Formulierung von allgemeinen Aussagen über einen Gegenstandsbereich, der eine Vielzahl von Einzelergebnissen zusammenfasst (Gigerenzer 1988, 1991; Dörner 1994). In der Psychologie werden Modelle konstruiert, um das Wissen über Strukturen und Zusammenhänge kognitiver Prozesse abbilden und erforschen zu können (Opwis 1992).

Dabei ist zu beachten, dass eine Modellrepräsentation vollständig und glaubwürdig ist (Cooper 2002). Vollständigkeit bedeutet in diesem Zusammenhang, dass durch Abstraktion keine Details verloren gehen, die einen wichtigen Einfluss auf das Verhalten des Originals besitzen. Die Glaubwürdigkeit eines Modells bezieht sich auf die akkurate Abstraktion des Originals, so dass keine unnötigen oder zusätzlichen Annahmen und Inhalte in das Modell integriert werden. Die Konstruktion kognitiver Modelle in der Kognitionspsychologie stellt dementsprechend den Versuch dar, „für ausgewählte kognitive Leistungen Symbolstrukturen (für Daten und Regeln) anzugeben und zu zeigen, dass mit eben diesen Daten und Regeln die zu erklärende kognitive Leistung erbracht werden kann“ (Tack 1995, S. 177).

Der Einsatz von Computermodellen in der Kognitionspsychologie ist auf das Paradigma der Informationsverarbeitung im Sinne der nicht-numerischen Informationsverarbeitung zurückzuführen (Neisser 1967; Newell & Simon 1972; Anderson 1983; Johnson-Laird 1988). Kognitive Prozesse werden laut dieses Paradigmas als wissensbasierte Vorgänge der Symbolverarbeitung aufgefasst. Wissensbasierte Systeme sind durch die Verwendung einer formalen Syntax gekennzeichnet, die prozedurales Wissen (Steuerungswissen) und deklaratives Wissen (Faktenwissen) explizieren und modularisiert darstellen, das aufgrund assoziativer Zugriffe durch Kontrollkomponenten aktiviert werden kann (Laubsch 1985; Smith 1982). Neben konzeptionellen Aspekten bilden Computermodelle der Kognition ebenfalls die formalen und funktionalen Beziehungen der menschlichen Kognition ab (Habel, Kanngießer & Strube 1990). Somit werden die Überprüfung kognitionspsychologischer Theorien und die Entwicklung neuer Hypothesen auf deren Grundlage durch den Einsatz berechenbarer Modelle ermöglicht.

Die Nutzung von Computermodellen erlaubt gegenüber der Anwendung neuronaler Netze oder den rein mathematischen Ansätzen zur Theorieprüfung und -generierung beispielsweise die graphische Veranschaulichung, die Demonstration der zeitlichen Dynamik und die Produktion von Verhaltensdaten auf Grundlage von formalisierten Theorien (Schaub 1993). Im Rahmen der Simulation kognitiver Prozesse können theoretische Annahmen über den Denkhandlungsprozess durch ein lauffähiges Computermodell auf Konsistenz, Plau-

sibilität und Vollständigkeit geprüft werden. Für die Entwicklungen von Computermodellen der menschlichen Kognition werden in der Kognitionspsychologie vermehrt kognitive Architekturen eingesetzt, die kognitionspsychologische Theorien zusammenfassen und auf deren Grundlage die Simulation kognitiver Prozesse möglich ist.

2.2.1 Kognitive Architekturen

Kognitive Architekturen integrieren isolierte Modelle und Theorien der Subprozesse und Subkomponenten der menschlichen Kognition. Anstatt die Welt in eine Anzahl von kleinen Paradigmen mit eigenen Lösungen und eigener Logik zu zerlegen, bilden sie die menschliche Kognition in ihrer Gesamtheit ab. Dieser Ansatz wurde erstmals von Newell in Form des *General Problem Solver* vorgestellt (Newell & Simon 1972; Newell 1973) und später unter der Bezeichnung *Unified Theories of Cognition* zusammengefasst (Newell 1990).

Kognitive Architekturen bezeichnen „das spezifische System invarianter Strukturen und primitiver Strukturen des kognitiven Apparates“ (Wallach 1998, S. 68) und determinieren die atomaren Bausteine der Kognition, also die kognitiven Operationen, die als Primitive angesehen werden. Diese werden in Architekturmerkmalen (z. B. Strukturen und Mechanismen) und durch wechselseitige Beschränkungen repräsentiert, die als allgemeine Gesetzmäßigkeiten unabhängig von Aufgabenstellung, Wissensbeständen und Aufgabenumgebung allgemein gelten (Chandrasekaran 1990). Zusammenhängende kognitive Funktionen werden in funktionalen Modulen repräsentiert, die den Zugriff auf die internen Strukturen und Mechanismen erlauben (z. B. Module für Wahrnehmung, Kognition, Handlung; Gray 2007). Funktionale Module können Untermodule enthalten, die wiederum kognitive Zusammenhänge abbilden. Komplexere Prozesse werden anhand der zur Verfügung stehenden Konstrukte aus den Primitiven und der Interaktion der funktionalen Module zusammengesetzt (Pylyshyn 1989).

Die Kontrolle und die Kommunikation der an kognitiven Prozessen beteiligten Komponenten kann anhand eines dreiteiligen Kontrollsystems abgebildet werden. Diese ist durch festgelegten Strukturen einer kognitiven Architektur implementiert und ermöglicht eine interaktive Kommunikation über die einzelnen funktionalen Module hinweg (Gray 2007).

Der erste Kontrolltyp stellt die Kommunikation zwischen funktionalen Modulen des kognitiven Apparates dar. Das heißt die Ausgaben eines funktionalen Moduls dienen als Eingaben eines anderen funktionalen Moduls. Diese kann direkt von Modul zu Modul erfolgen oder über eine zentrale Komponente.

Die interne Kontrolle eines funktionalen Moduls wird als zweiter Kontrolltyp aufgefasst. Eingaben werden durch andere funktionale Module oder die Aufgabenumgebung bereitgestellt. Innerhalb des funktionalen Moduls werden die Eingaben verarbeitet und gegebenenfalls zu Ausgaben transformiert. Die Mechanismen können in Untermodulen, die jeweils eine Theorie implementieren, bereitgestellt werden, die in drei Untertypen eingeteilt werden können. Der erste Untertyp stellt empirisch oder heuristisch begründete Statistiken dar, wie beispielsweise

se Fitts' Gesetz zur Vorhersage für den Zeitverbrauch von geführten motorischen Bewegungen (Fitts 1954). Die zweite Gruppe repräsentiert kognitive Theorien, die zur Laufzeit des kognitiven Modells ausgeführt werden können und auf mechanischen oder algorithmischen Ausdrücken beruhen (Anderson 1987), wie beispielsweise die auf einer rationalen Analyse beruhende Aktivierungstheorie des deklarativen Gedächtnisses (Anderson & Schooler 1991). Die dritte Gruppe von Submodulen fasst Methoden zusammen, die vor der Laufzeit eines kognitiven Modells berechnet und deren Ergebnisse offline bereitgestellt werden, wie zum Beispiel Wechselkosten bei Mehrfachaufgaben oder die vorherige Berechnung von Blickbewegungszeiten.

Der dritte Kontrolltyp kombiniert aufgabenspezifisches Wissen mit den Merkmalen der kognitiven Architektur, um Strategien oder Methoden für die Aufgabenbearbeitung zu erzeugen. Kontrolle vom Typ drei stellt die Anwendung der kognitiven Architektur in einem speziellen Aufgabenszenario dar.

Anhand der auf kognitiven Architekturen basierenden kognitiven Modelle mit ihren interaktiven Kontrollstrukturen können unterschiedliche kognitionspsychologische Annahmen und verschiedene Arten der wechselseitigen Kommunikation der an der Kognition beteiligten Module rechnergestützt untersucht und quantitativ vorhergesagt werden (Gluck, Ball & Krusmark 2007).

Produktionensysteme

Die meisten kognitiven Architekturen basieren auf Produktionensystemen, die als „elaborierteste Rahmenvorstellung“ der menschlichen Kognition betrachtet werden (Opwis 1992, S. 73). Produktionensysteme bestehen aus drei Komponenten: einem Datenspeicher, einem Produktionen- oder Regelspeicher und einem Interpretierer (vgl. Opwis 1992).

Der Datenspeicher enthält das dem System bekannte deklarative Wissen (Faktenwissen) und stellt die Datenbasis eines Produktionensystems dar. Der Produktionen- oder Regelspeicher enthält das operative Wissen des Systems in Form einer Menge von Produktionen oder Regeln. Diese bestehen aus einem Bedingungsteil und einem Aktionsteil. Der Bedingungsteil enthält eine Menge von Bedingungen, die erfüllt sein müssen, damit der Aktionsteil ausgeführt werden kann. Der Interpretierer steuert den Prozess der Informationsverarbeitung durch einen mehrstufigen Zyklus: Auswerten, Konfliktlösung und Ausführung.

Das Auswerten der Bedingungsteile aller Produktionen eines Produktionensystems führt zu einer Menge ausführbarer Produktionsregeln, der Konfliktmenge. Besteht die Konfliktmenge aus mehr als einer Produktionsregel, wird der Konflikt durch die Auswahl einer Produktionsregel gelöst. Dabei können verschiedene Heuristiken zum Einsatz kommen (z. B. Reihenfolgedominanz, Spezifität, Zufallswahl; vgl. Wallach 1998). Abschließend wird der Aktionsteil der selektierten Produktionsregel ausgeführt, wodurch Manipulationen an der Datenbasis vorgenommen werden können. Die iterative Ausführung der drei Schritte wird solange fortgesetzt, bis die Konfliktmenge keine Produktionsregel enthält oder der Aktionsteil einer Produktion den Zyklus ausdrücklich beendet.

Vorteile von Produktionensystemen sind nach Wallach (1998) der modulare und homogene Aufbau und die Abstraktheit der benutzten formalen Ausdrücke und Mechanismen. Sie ermöglichen die ziel- als auch die datengesteuerte Abbildung von Prozessen, wobei der Aufbau von Produktionensystemen analog zu gedächtnispsychologischen Strukturvorstellungen aufgebaut ist. Somit sind sie als Grundlage zur Erklärung von menschlichen Lern- und Informationsverarbeitungsprozessen geeignet.

Auswahl einer kognitiven Architektur

Die drei bekanntesten kognitiven Architekturen, die gegenwärtig aktiv weiterentwickelt und für Fragestellungen der Mensch-Maschine-Interaktion eingesetzt werden, sind ACT-R, Soar und EPIC. Ein ausführlicher Vergleich ist zu finden bei Byrne (2003). Alle drei kognitiven Architekturen basieren auf Produktionensystemen, wobei sich die Granularität und die Verknüpfungsprozesse der drei Architekturen unterscheiden. Jede Architektur hat einen eigenen Fokus und ist somit für verschiedene Fragestellungen anwendbar. In Tabelle 1 sind die wichtigsten Architekturmerkmale vergleichend dargestellt (nach Byrne 2003).

Hinsichtlich der Modellierung von Benutzerverhalten in dynamischen und komplexen Mensch-Maschine-Systemen sind mehrere Kriterien in Betracht zu ziehen, die eine Auswahl einer kognitiven Architektur bestimmen. Diese können in harte und weiche Faktoren unterteilt werden. Die harten Faktoren beziehen sich auf Strukturen und Mechanismen der kognitiven Architektur, die weichen auf die eine kognitive Architektur begleitenden Umstände, wie zum Beispiel Lernmaterialien, Anwendung und Ausrichtung.

Für die harten Faktoren ergibt sich folgendes Bild. Das Handeln in dynamischen und komplexen Systemen bedarf der Interaktion mit einer Systemumgebung. Das Lernen von neuen oder angepassten Handlungsstrukturen muss durch eine kognitive Architektur für die Simulation von Benutzerverhalten in komplexen und dynamischen Mensch-Maschine-Systemen möglich sein, um die vorhandene Wissensbasis adäquat an Veränderungen anpassen zu können. Für eine Analyse der Handlungsprozesse und -strukturen ist es notwendig, zwischen prozeduralem und deklarativem Wissen innerhalb der kognitiven Architektur zu unterscheiden, um empirische Unterschiede in der Nutzung differenzieren zu können.

Der Vergleich der drei kognitiven Architekturen zeigt, dass ACT-R hinsichtlich der harten Faktoren die Spezifikationen und Anforderungen erfüllt (siehe Tabelle 1). Soar ermöglicht keine Unterscheidung von Wissensseinheiten und bot lange keine Interaktions- und Wahrnehmungsmöglichkeiten von Arbeitsumgebungen an. Erst durch eine Erweiterung von Soar, die auf EPIC beruht, kann Soar/EPIC auf direktem Weg mit externen Arbeitsumgebungen interagieren. EPIC stellt keine Lernmechanismen zur Generierung neuer Wissensseinheiten bereit.

Hinsichtlich der weichen Faktoren stellt ACT-R ebenfalls die zweckmäßigste Wahl dar. Lernmaterialien und eine umfangreiche Dokumentation werden angeboten. Dadurch ist eine selbstständige Erweiterung

Tabelle 1: Vergleich der drei kognitiven Architekturen ACT-R, Soar und EPIC (nach Byrne 2003).

	ACT-R	Soar	EPIC
Verwendungszweck	Gedächtnis und Problemlösen	Lernen und Problemlösen	Mehrfachaufgaben und menschliche Wahrnehmung und Motorik
Zyklus	Produktionszyklus, serielles Ausführen von Produktionsregeln (Konfliktlösemechanismen)	Zweistufiger Entscheidungszyklus, serielles Ausführen von Produktionsregeln. Alle möglichen Produktionsregeln werden ausgeführt und mögliche Änderungen der Datenbasis berechnet. Anhand von Präferenzen werden die tatsächlichen Änderungen an der Datenbasis durchgeführt	Produktionszyklus, paralleles Ausführen von Produktionsregeln
Symbolische und aktivierungs-basierte Informationsverarbeitung	Hybrider Ansatz	Symbolische Informationsverarbeitung	Symbolische Informationsverarbeitung
Motorische und perzeptuelle Systeme	Ja	Nein, aber in der Verbindung von Soar mit EPIC stehen motorische und perzeptuelle Systeme bereit	Ja
Management von Zielen	Ziel-Buffer mit einem aktuellen aktiven Ziel. Das Ziel kann durch Produktionsregeln gewechselt werden	Zielspeicher mit einem aktiven Ziel. Serielle Abarbeitung des Zielspeichers ist notwendig	Parallele (beschränkte) Verfolgung von Zielen (z. B. nur ein Augenpaar)
Datenspeicher	Deklaratives und prozedurales Wissen wird getrennt gespeichert (Chunks und Produktionsregeln)	Keine Trennung zwischen deklarativem und prozeduralem Wissen. Das Wissen wird in Form von Produktionsregeln gespeichert	Deklaratives und prozedurales Wissen wird getrennt gespeichert (Chunks und Produktionsregeln)
Lernmechanismen	Ziele werden als Chunks abgespeichert. Lernen anhand von Beispielen ist möglich	Gelöste Problemstellungen werden zu neuen generalisierten Produktionsregeln zusammengefasst (chunking)	keine
Lernunterstützung	Großes Lernmaterial in Form von Dokumentation und Tutorials vorhanden, eine Sommerschule wird angeboten	Einige Lernmaterialien vorhanden, FAQ	keine
Benutzergruppe (ausschließlich der Entwickler)	Wachsende Benutzergruppe, primär in der Psychologie	Einige Benutzer, primär im Bereich der künstlichen Intelligenz	keine

und Anpassung der Architektur möglich. Die theoretischen und praktischen Grundlagen von ACT-R werden von einer wachsenden Forscher- und Entwicklergruppe kontinuierlich erweitert und um neue Aspekte ergänzt. ACT-R ist kostenfrei und in aktuellen Versionen online zu beziehen. ACT-R wird in umfangreichen Studien im zivilen Bereich eingesetzt, Soar kommt eher im militärischen Bereich zum Einsatz und EPIC wird vereinzelt in psychologischen Studien als kognitive Modellierungssprache genutzt (Byrne 2003).

Aus diesen Gründen wurde ACT-R als kognitive Architektur zur Simulation kognitiver Prozesse für die Vorhersage von Benutzerverhalten in komplexen und dynamischen Mensch-Maschine-Systemen ausgewählt. Ein weiteres Kriterium waren die vorhandenen Erfahrungen und Vorarbeiten in der Anwendung von ACT-R auf dem Gebiet der Benutzermodellierung der Arbeitsgruppe, in deren Rahmen diese Arbeit entstand. Im folgenden Abschnitt wird ACT-R kurz vorgestellt.

Atomic Components of Thought – Rational

Die kognitive Architektur ACT-R (*atomic components of thought – rational*) stellt eine formale Theorie dar, um die menschliche Kognition zusammenhängend zu beschreiben (Anderson 1993; Anderson & Lebiere 1998; Anderson et al. 2004). ACT-R ist eine hybride kognitive Architektur, wobei die symbolverarbeitende Struktur durch ein Produktionensystem und die subsymbolische Struktur durch eine Vielzahl von parallelen Prozessen, die durch mathematische Gleichungen beschreibbar sind, bestimmt wird.

Das zugrunde liegende Architekturkonzept von ACT-R besteht aus einer Menge von Modulen, die verschiedene Arten von Informationen verarbeiten. In Abbildung 2 sind die wichtigsten Module dargestellt: Das visuelle Modul ermöglicht die Identifizierung von Objekten im visuellen Feld und die Ausrichtung der visuellen Aufmerksamkeit, das manuelle Modul kontrolliert die motorischen Handlungen, das deklarative Modul ermöglicht die Abfrage von Wissensseinheiten aus dem Gedächtnis und das intentionale Modul steuert die Zielverfolgung und Intentionen. Ein zentrales Produktionensystem koordiniert das Verhalten der einzelnen Module. Es ist nicht für alle Aktivitäten der einzelnen Module sensitiv, sondern kann nur auf eine limitierte Menge an Informationen reagieren, die durch einzelne Datenspeicher, die Buffer der Module, bereitgestellt werden (Anderson et al. 2004).

In ACT-R gibt es zwei Arten von Gedächtnissen. Das deklarative Gedächtnis speichert Wissen in Form von Chunks. Sie werden in ACT-R nicht ausschließlich als Symbol interpretiert, sondern sind zusätzlich mit einem Aktivierungslevel assoziiert. Dieser Wert sagt aus, wie häufig und wann zuletzt ein Chunk abgerufen wurde (Anderson & Lebiere 1998), und ist ein Maß für die Relevanz des Chunks im augenblicklichen Kontext der Aufgabe. Je öfter beispielsweise ein Chunk abgerufen wurde, umso höher ist dessen Aktivierungslevel, und umso einfacher ist der Chunk durch das kognitive System zukünftig abzurufen.

Eine zentrale Annahme in ACT-R ist, dass "cognitive skills are realized as production rules" (Anderson 1993, S. 1). Diese Produktionsregeln werden in dem prozeduralen Gedächtnis abgelegt und stellen das Handlungswissen eines kognitiven Modells dar. Produktionsregeln sind in

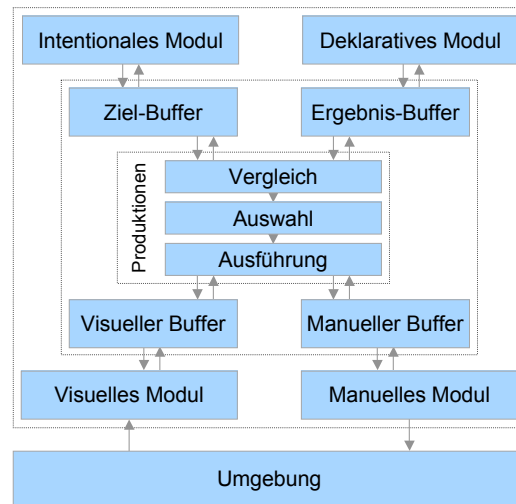


Abbildung 2: Die Informationsorganisation in ACT-R. Die Informationen in den Buffern, die mit den Modulen assoziiert sind, werden durch Produktionsregeln angefragt und geändert (nach Anderson et al., 2004).

ACT-R durch den Gebrauch von Variablen generalisierbar, durch die Restriktion eines Anwendungsbereiches einer Produktion spezifizierbar und bilden die Handlungskontrolle von einer Bedingung zur Aktion ab (d. h. wenn A dann B).

Die als atomar angenommen Produktionsregeln erlauben durch die integrierten Mechanismen zum Wissenserwerb, das heißt durch Lernprozesse, die Reduzierung von komplexen Fertigkeiten auf einfache Regeln (Anderson & Lebiere 1998). Für die Konfliktlösemechanismen ist jeder Produktionsregel eine Anzahl von Werten zugeordnet, wie zum Beispiel Wahrscheinlichkeit der Produktionen zur Zielerfüllung und Kosten der Zielerreichung bei Selektion der Produktion. Sind in einem Zyklus des Produktionensystems (festgelegte Dauer 50 Millisekunden) mehrere Produktionsregeln ausführbar, wird für jede momentan ausführbare Produktion anhand der assoziierten Parameter der erwartete Nutzen berechnet. Die Produktion mit dem höchsten Nutzen wird ausgewählt und kommt zur Ausführung. Um zu verhindern, dass es zu einer Dominanz einer Produktionsregel kommt, sind stochastische Verfahren in den Auswahlprozess integriert.

In ACT-R sind Mechanismen zum Wissenserwerb auf verschiedenen Ebenen realisiert (Anderson & Lebiere 1998). Zum einen werden neue Informationen (z. B. Objekte aus der Aufgabenumgebung oder Resultate von Produktionsregeln) in Form von Chunks im deklarativen Gedächtnis gespeichert und als gelernt der Wissensbasis hinzugefügt (symbolisches Lernen). Zum anderen stellt die Berechnung der Aktivierungslevels von Chunks und die Brauchbarkeitsbestimmung von Produktionsregeln Lernen auf der Basis von Erfahrung dar (subsymbolisches Lernen). Ein weiterer Mechanismus zum Wissenserwerb ist die Automatisierung von Handlungsabfolgen durch die Kompilation deklarativer Gedächtniselemente in Produktionsregeln (Anderson &

Lebiere 1998). Dieser *Production Compilation* genannter Mechanismus referiert auf empirische Ergebnisse, die zeigen, dass Menschen eine Prozedur ausführen können, das heißt gelernt haben, ohne die jeweiligen deklarativen Wissens Elemente benennen zu können.

Durch die Bereitstellung von peripheren Schnittstellen kann ACT-R mit den gleichen Softwaresystemen wie bei experimentellen Untersuchungen interagieren. Die Evaluation von realen Systemen wird prinzipiell ermöglicht, erfordert aber einige Einschränkungen in der Gestaltung der Mensch-Maschine-Schnittstelle, da nur einige wenige restriktive Bildelemente bereitstehen, wie beispielsweise Linien, Buttons und Text (vgl. Byrne 2003).

Der Anwendungsbereich von ACT-R reicht von der Untersuchung kognitionspsychologischer Fragestellungen, wie Entscheiden und Lernen (Anderson & Lebiere 1998), bis hin zur Modellierung des Benutzerverhaltens in Bereichen der Mensch-Maschine-Interaktion, wie beispielsweise die Kontrolle des Flugverkehrs durch Fluglotsen (Schoelles & Gray 2000) oder die Abbildung der Fahrzeugführung im Fahrsimulator (Salvucci 2001b).

2.2.2 Kognitive Modellierung

Für die kognitive Modellierung gibt es keine allgemeine Methodologie (Ritter & Larkin 1994). Der Prozess der kognitiven Modellierung kann schematisch als eine Verbindung von vier nicht linearen Phasen beschrieben werden (Wallach 1998; Stroher 1995). Nach Wallach (1998) sind diese: Die aufgabenanalytische Untersuchung des Aufgabenbereichs, die Durchführung empirischer Experimente, die Modellimplementierung und die Geltungsprüfung der Modellierung (siehe Abbildung 3).

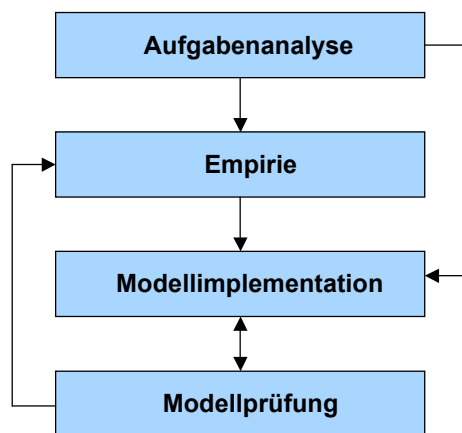


Abbildung 3: Phasen der Modellkonstruktion (nach Wallach 1998, S. 40).

Die Aufgabenanalyse dient der Untersuchung des Gegenstandsbeereichs, um den Verwendungszweck des angestrebten Modells festzulegen. Durch empirische Untersuchungen werden zuvor formulierte theoretisch begründete Aussagen bewertet. Die Bewertung basiert auf

der Sichtung und Bewertung von in diesem Schritt durchzuführenden empirischen Untersuchungen oder auf vorhandenen empirischen Befunden. In dem Schritt der Modellimplementierung werden theoretische Aussagen in Ausdrücke formaler Sprachen überführt und mit Hilfe einer geeigneten Programmiersprache umgesetzt. Die Geltungsprüfung der Modellierung ermöglicht das Aufdecken und Beheben von Lücken und Fehlern des kognitiven Systems im Bezug zu genutzten theoretischen Ansätzen und empirischen Befunden.

In der Praxis stellt sich dieser Prozess oft als dialektischer Spiralprozess dar (Dörner 1989), das heißt ein Wechsel zwischen den vier Phasen ist Bestandteil des Prozesses zur Modellanpassung. Beispielsweise kann die Modellprüfung ergeben, dass weitere empirische Untersuchungen notwendig sind, oder es zeigt sich, dass die Aufgabenanalyse nicht detailliert genug war, um das Modell zu implementieren (siehe Abbildung 3).

Modellprüfung – Die kognitive Empirie

Die kognitive Empirie stellt den Prozess der Bewertung eines kognitiven Modells dar. Diese stellt in dem Prozess der kognitiven Modellierung den Schritt der Modellprüfung dar (siehe Abbildung 3). Dabei kommen natürliche und künstliche Systeme zum Einsatz (Stroher 1995). Die Bewertung eines kognitiven Modells resultiert aus der Analyse der Beziehung zwischen der Empirie mit natürlichen kognitiven Systemen und der Empirie mit künstlichen kognitiven Systemen. Die Beziehung zwischen Computermodell und menschlicher Psyche lässt sich als eine Relation zwischen zwei Tripeln beschreiben (Zinterhof 1987), (I ; O ; S) und (I' ; O' ; S'). Dabei bezeichnet I den Input des realen Systems menschlicher Psyche und I' den Input des Computermodells. O , O' bezeichnen den zu beobachtbaren Output des jeweiligen Systems und S , S' die Übersetzung von Input zu Output (siehe Abbildung 4).

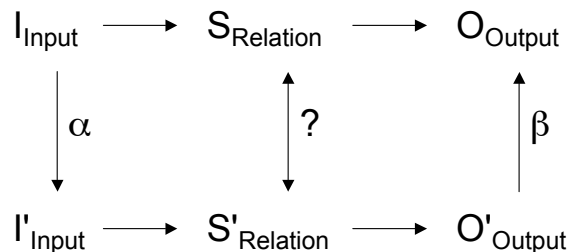


Abbildung 4: Relation zwischen natürlichem System S und künstlichem System S' . α und β stellen eine beobachtbare Abbildung der Input- und Outputrelation dar. Die Relation von S zu S' ist nur theoriegeleitet auf Grundlage von α und β zu interpretieren (nach Zinterhof 1987, S. 105ff.).

Da die interne Struktur der menschlichen Psyche, das heißt die kognitiven Prozesse zur Übersetzung des Inputs in den Output, nicht direkt zu beobachten sind, kann die Verknüpfung der beiden Tripel nur über die jeweiligen Input- beziehungsweise Outputrelationen hergestellt werden (α und β in Abbildung 4). Die Gültigkeit eines Computermodells

kann anhand der Ausprägung der Parameter α und β abgeleitet und durch diese auf die Beziehung zwischen S und S' geschlossen werden.

Durch die Verknüpfung der verschiedenen Untersuchungsverfahren, die für empirische und simulationsbasierte Untersuchungen eingesetzt werden, wird eine adäquate Methode im Rahmen der Kognitionswissenschaft ermöglicht (Stroher 1995).

Vorteile und Nachteile der kognitiven Modellierung

Für die Simulation kognitiver Prozesse werden in der Literatur Vorteile und Nachteile genannt (vgl. Boden 1987; Byrne 2003; Deppe 1977; Opwis 1992; Schäfer 1985; Wallach 1998), die in Tabelle 2 zusammenfassend dargestellt sind.

Tabelle 2: Vorteile und Nachteile der kognitiven Modellbildung.

Vorteile	Nachteile
• Möglichkeit zur formalen Überprüfung von kognitionspsychologischen Theorien • Formale Vorhersage von Verhaltensdaten • Einfache Manipulation von Simulationsexperimenten • Einfache Reproduktion von Simulationsexperimenten • Präzise Formulierung von theoretischen Ansätzen • Möglichkeit zur Kommunikation von sonst verbal formulierten Annahmen	• Kognitive Architekturen sind unvollständige Konstrukte der menschlichen Kognition • Wissensstrukturen werden durch die Entwickler in das System induziert • Detaillierte Annahmen über die Kognition und die internen Prozesse sind notwendig • <i>Theory implementation gap</i> – Unterschiede zwischen Theorie und Implementation • <i>Irrelevant specification</i> – Es werden Zusatzannahmen getroffen, die theoretisch nicht begründet sind • Keine Unterstützungswerkzeuge für die Modellentwicklung und die Analyse von Modelldaten und deren Vergleich

Die Anwendung formaler Simulationsmodelle unterstützt die Überprüfung von kognitionspsychologischen Theorien und Annahmen, die auf beobachteten Eingaben und Ausgaben natürlicher kognitiver Systeme beruhen. Es besteht die Möglichkeit inkonsistente, unvollständige und überflüssige Annahmen zu identifizieren. Anhand der durch kognitive Modelle simulierten Leistungsdaten können empirische Verhaltensdaten vorhergesagt und präzise überprüft werden. Eine einfache Wiederholung und Manipulation von kognitiven Modellierungsexperimenten ist in diesem Zusammenhang ebenfalls möglich und erlaubt den beschleunigten oder verlangsamten Ablauf des Simulationsexperiments für die genauere Analyse und Beurteilung der Veränderungen des künstlichen Systems. Die präzise Formulierung der Theorien gewährleistet eine bessere Kommunikation als verbale Formulierungen. Beispielsweise kann die Methode der kognitiven Modellierung in der

kognitiven Neuropsychologie angewandt werden. Durch einen Vergleich der Leistungsdaten von kognitiven Modellen und systematisch verletzten kognitiven Modellen (z. B. durch die Reduzierung der Wissensbasis des kognitiven Modells oder der Änderung von Funktionen für die Berechnung der Aktivierungslevels von Wissensselementen) können mögliche Erklärungen für empirisch belegte Verhaltensmuster abgeleitet werden (vgl. Stroher 1995).

Heutige kognitive Architekturen beruhen auf unvollständigen Theorien der menschlichen Kognition und beziehen nicht alle Aspekte mit ein, die für das menschliche Verhalten verantwortlich sind (z. B. Langeweile, Ästhetik, Freude und Vorlieben). Zum heutigen Zeitpunkt ist die Berücksichtigung dieser Aspekte noch nicht notwendig, da der gegenwärtige Fokus kognitiver Architekturen in der Erklärung von Kognition und Performanz liegt (Byrne 2003). Mit steigender Erkenntnis der Kognitionspsychologie kann der Fokus verändert und die Bandbreite der simulierbaren Phänomene erweitert werden.

Bei der Entwicklung kognitiver Modelle muss berücksichtigt werden, woher das Wissen der kognitiven Modelle stammt, das für die Aufgabenbearbeitung (z. B. Aufgabenstellung, Schnittstellenbenutzung) benötigt wird. Diese Wissensstrukturen müssen heute durch den Entwickler eines kognitiven Modells implementiert werden. Die Umsetzung eines Simulationsmodells basiert dementsprechend auf den Wahrnehmungen und Interpretationen der Entwickler und hat zur Folge, dass Ergebnisse im Kontext der implementierten Annahmen interpretiert werden müssen. Des Weiteren werden für die Formulierung eines kognitiven Modells detaillierte Annahmen bezüglich der Mechanismen und Strukturen benötigt, die heute schwer bis gar nicht empirisch zu gewinnen sind und eine Vielzahl von identischen empirischen Untersuchungen erfordern, um eine gesicherte statistische Datenbasis zu erhalten.

Die kognitionspsychologischen Annahmen kognitiver Architekturen und kognitiver Modelle beruhen auf verbal formulierten Theorien. In der Modellierung kann es daher zu Unterschieden zwischen Theorie und tatsächlicher Implementierung kommen, die beispielsweise auf die Abstraktion des Gegenstandsbereiches zurückzuführen sind (*Theory implemetation gap*; Cooper & Shallice 1995). Ein weiteres Problem ist das Treffen von Zusatzannahmen bei der Implementierung, die das kognitive Modell ablauffähig machen und keine eindeutige Trennung von essentiellen Teilen der Theorie und Details der Implementierung zulassen (*Irrelevant specification*; Newell 1990).

Der Aufwand der Modellerstellung und der Analyse der simulierten Daten ist groß und es werden kaum Systeme zur Unterstützung angeboten. Hinderlich ist auch, dass kognitive Architekturen ein großes und komplexes Regelwerk darstellen und nur vereinzelt Hilfen in Form von Dokumentationen und Tutorials angeboten werden (z. B. bei ACT-R). Diese Aspekte führen dazu, dass die Modellbildung auf Grundlage kognitiver Architekturen schwer anzuwenden und zu erlernen ist.

2.3 ANALYSE SIMULierter KOGNITIVER PROZESSE

„Der Entwurf und die Entwicklung von Mensch-Maschine-Systemen erfordert eine ganzheitliche und simultane Berücksichtigung der inter-

agierenden technischen und menschlichen Systemkomponenten“ (Kraiss 1995, S. 33). Die Möglichkeiten und Grenzen des menschlichen kognitiven Apparates müssen dementsprechend bei der Entwicklung von Mensch-Maschine-Systemen berücksichtigt werden. Diese Forderung ist hinsichtlich der zunehmenden Komplexität von Mensch-Maschine-Systemen, die sich in steigender Anforderung an Benutzer bei der Bedienung technischer Systeme manifestiert, notwendig.

Die empirische Bewertung von Mensch-Maschine-Systemen ist kosten- und zeitaufwendig, und aufgrund der benötigten Prototypen liegen Ergebnisse häufig erst in späteren Phasen der Systementwicklung vor. Die Ergebnisse können nachträglich nur unter großem Aufwand berücksichtigt werden. Richtlinienbasierte Methoden können schon in frühen Phasen der Systemgestaltung eingesetzt werden, bilden aber statisches Expertenwissen ab und integrieren die menschliche Komponente nur in Form von Heuristiken.

Die rechnergestützte Bewertung von Gestaltungsalternativen stellt eine Methode dar, die eine ganzheitliche Betrachtung von komplexen und dynamischen Mensch-Maschine-Systemen und darauf basierend die Generierung von Gestaltungshinweisen erlaubt (Jürgensohn 2002). Aussagen sind beispielsweise bezüglich der Aufgabenteilung von Mensch und Maschine und der Anpassung der Interaktion an die Bedingungen und Anforderungen der menschlichen Kognition möglich. Zu diesem Zweck interagiert ein kognitives Benutzermodell mit einem virtuellen Prototypen oder einem Mensch-Maschine-System. Die Simulationsdaten helfen, die Schnittstelle aufgrund der Explizitheit und Operationalität der Simulationsmodelle quantitativ und normativ zu bewerten. Somit können durch den Einsatz von kognitiven Benutzermodellen, soweit diese effizient implementiert sind (siehe hierzu Urbas, Schulze-Kissing & Leuchter 2005), die Kosten und die benötigten Ressourcen in der Systementwicklung verringert und die Anzahl der pro Zeiteinheit zu evaluierenden Gestaltungsalternativen erhöht werden. Nachteile dieser Methode liegen in der komplexen und aufwendigen Modellerstellung und Datenanalyse und in der Unvollständigkeit kognitiver Theorien, Architekturen und Modelle (Sun, Coward & Zenzen 2004). In dieser Arbeit wird die Analyse von kognitiven Simulationsexperimenten behandelt.

2.3.1 *Methoden zur Analyse von kognitiven Simulationsexperimenten*

In der Simulation von Benutzerverhalten besteht die Möglichkeit, ein Phänomen oder eine Gruppe von Phänomenen auf verschiedene Arten zu modellieren. Es gibt keine gesicherten Erkenntnisse, die einen theoretischen Ansatz bestätigen. Es ist nur die Plausibilisierung eines Ansatzes gegenüber konkurrierender Ansätze möglich (Schaub 1993). Die kognitive Empirie (siehe Kapitel 2.2.2) ermöglicht die Bewertung von kognitiven Simulationsexperimenten. Dabei werden Experimente mit natürlichen und künstlichen kognitiven Systemen betrachtet. Beispielsweise können zwei kognitive Modelle, die auf verschiedenen Annahmen beruhen, verglichen werden oder empirische und simulierte Ergebnisse werden miteinander in Bezug gesetzt.

Die Methode *Goodness-of-fit* ermöglicht den Vergleich von Daten unterschiedlicher Herkunft für einen Untersuchungsgegenstand im Bereich der Modellbildung (Schunn & Wallach 2005). Dabei werden zwei Ansätze unterschieden: Die visuelle Präsentation und die numerische Betrachtung der erfassten Daten. Die visuelle Präsentation erlaubt den visuellen Vergleich von Modellvorhersagen und beobachteten Daten und ermöglicht das Auffinden von Unterschieden und Ähnlichkeiten. Dieses Verfahren kann für die schnelle Sichtung der Daten angewandt werden und hilft problematische Datenabschnitte und systematische Auffälligkeiten zu identifizieren. Die numerische Betrachtung ermöglicht eine Aussage über die Genauigkeit der Vorhersagen. Es gibt eine Vielzahl an Methoden, die in diesem Zusammenhang angewandt werden können. Für einen Überblick siehe Schunn und Wallach (2005).

Im Bereich der kognitiven Modellierung mit ACT-R werden zunehmend Methoden entwickelt, die eine numerische Betrachtung der Simulationsdaten erlauben. Statistische Methoden können angewandt werden, um beispielsweise die kognitive Performanz von Modellen in ACT-R zu vergleichen (Baker, Corbett & Koedinger 2003). Dabei handelt es sich um Ansätze, die Modelldaten miteinander vergleichen, um die Qualität und Genauigkeit der Vorhersagen festzustellen. Für die Bewertung von Mensch-Maschine-Systemen stellen diese Verfahren noch keine ausreichende Möglichkeit dar, um eine Schnittstelle zu bewerten, da mit diesen Methoden die internen Strukturen und Mechanismen bewertet werden. Das heißt die Modellstruktur wird mit der formalisierten Theorie abgeglichen. Das Modellverhalten in Bezug zu empirischen Daten wird dabei nicht betrachtet.

2.3.2 Auswahl einer Evaluationsmethode

In heutigen Simulationsexperimenten werden aufgrund des hohen Aufwands zumeist heuristische Leistungsdaten analysiert (z. B. Ausführungszeiten und Fehler), um die Übereinstimmung zwischen Modell und Empirie auf globaler Ebene zu zeigen. Allerdings kann das Verhalten durch eine Vielzahl verschiedener kognitiver Modelle simuliert werden, die auf globaler Ebene Ähnlichkeiten, auf feinerer Ebene aber Unterschiede aufweisen (Gray, Sims & Schoelles 2006). Beispielsweise können kognitive Modelle mit unterschiedlichen Strategien für die Informationsaufnahme identische Ausführungszeiten und Fehlerraten vorhersagen. Daher ist für die Bewertung von Schnittstellen für Mensch-Maschine-Systeme die Analyse feinerer empirisch erhobener und simulierter Leistungs- und Verhaltensdaten von großem Interesse.

Für die Evaluation von Mensch-Maschine-Systemen können eine Vielzahl von heuristischen Verhaltensdaten (z. B. Ausführungszeiten, Fehler) und psychophysiologischen Daten (z. B. bildgebende Verfahren, pupillometrische Daten) erhoben werden (vgl. Funke & Spering 2006; Ward & Marsden 2003).

Die kognitive Architektur ACT-R verarbeitet während der Simulation des Benutzerverhaltens eine große Menge von Daten, die den empirisch erhobenen Daten zum Teil gleichen. Abgeleitet werden können unter anderem Blickbewegungen, Handlungsprotokolle, Zielverfolgung, die Aktivierungslevels der deklarativen Gedächtniselemente

und neurophysiologische Daten wie die Aktivierung von Gehirnarealen (Anderson & Lebiere 1998; Anderson et al. 2004). Für den Vergleich von Modell und Empirie müssen die simulierten Daten ein Korrelat in empirisch erhobenen Daten besitzen. Diese müssen einfach erhoben werden können und genaue Aussagen bezüglich der Gestaltung ermöglichen.

Viele der Leistungsdaten, die ACT-R bereitstellt, sind heute aus psychologischer Sicht noch zu ungenau, um Aussagen bezüglich der Gestaltung zu ermöglichen (z. B. neuropsychologische Daten), oder sind nur unter großem Aufwand zu erheben (z. B. Handlungsprotokolle). Ein weiterer Faktor, der die Auswahl einer Methode bestimmt, ist der Einsatz des empirischen Korrelats in der Praxis zur Bewertung von Mensch-Maschine-Systemen.

Werden alle Aspekte zusammengefasst, stellt die Verwendung der simulierten Blickbewegungen der kognitiven Architektur ACT-R eine zweckdienliche Wahl dar. Blickbewegungsdaten sind in der Praxis einfach zu erheben und werden experimentell in der Forschung und der Industrie zur Bewertung von Mensch-Maschine-Systemen eingesetzt (Rötting 2001). Untersucht werden beispielsweise Mensch-Maschine-Systeme und deren Unterstützung bei fertigungs- und regelbasierten Tätigkeiten. Die Blickbewegungsdaten des visuellen Systems von ACT-R können ausgelesen, algorithmisch aufbereitet und für die Analyse bereitgestellt werden. Somit bietet die Analyse der simulierten Blickbewegungen für die Bewertung von Mensch-Maschine-Systemen in prospektiven Phasen der Systementwicklung eine gute Grundlage für die Bewertung von Gestaltungsalternativen.

Aus diesem Grund wird in dieser Arbeit die Anwendung simulierter Blickbewegungsdaten kognitiver Modelle für die Bewertung von Mensch-Maschine-Systemen weiter betrachtet und die Ergebnisse mit empirisch erhobenen Blickbewegungsdaten in Beziehung gesetzt.

2.4 BLICKBEWEGUNGEN

Fast alle Handlungen und deren zugrunde liegenden kognitiven Prozesse sind von Augenbewegungen begleitet (Land & Hayhoe 2001). Augenbewegungen werden als die Bewegungen des Auges bezeichnet, „die allein durch Beobachtung des Auges erfasst und interpretiert werden können“ (Joos, Rötting & Velichkovsky 2003, S. 142). Als Blickbewegungen werden „solche Bewegungen des Auges bezeichnet, die in Verbindung mit den vom Auge aufgenommen Informationen interpretiert werden“ (Joos et al. 2003, S. 143). Die Erfassung und Analyse von Blickbewegungen stellt im psychologischen und arbeitswissenschaftlichen Kontext eine wichtige Methode dar, um kognitive Prozesse der visuellen Informationsaufnahme und -verarbeitung im Rahmen arbeitswissenschaftlicher Arbeitsanalysen oder der Bewertung von Mensch-Maschine-Systemen zu untersuchen (Rötting 2001).

2.4.1 Blickbewegungen als Indikatoren für kognitive Prozesse

Die Verwendung von Methoden der Blickbewegungserfassung zur Bewertung von Prozessen der Informationsaufnahme und -verarbeitung

stützt sich auf die Annahme, dass Blickbewegungen Aufschluss über höhere kognitive Prozesse geben. Dazu werden folgende Annahmen postuliert: (1) Es wird das Objekt fixiert, das Gegenstand der zentralen Verarbeitung ist (*eye-mind assumption*; Just & Carpenter 1980) und (2) die Dauer einer Fixation entspricht der Dauer der zentralen Verarbeitung (*immediacy assumption*; Just & Carpenter 1980). (3) Blickbewegungen erlauben so das Treffen von Aussagen über nicht direkt beobachtbare kognitive Vorgänge der zentralen Informationsaufnahme und -verarbeitung (Just & Carpenter 1980; Schroiff 1986).

Im Rahmen von Blickbewegungsuntersuchungen muss überprüft werden, ob die Annahmen Gültigkeit besitzen und ob eine Theorie zu Blickbewegungen als Indikator für kognitive Prozesse genutzt werden kann. In der Literatur wurden daher Anforderungen formuliert (vgl. Ballard, Hayhoe, Pook & Rao 1997; Kowler 1990; Viviani 1990 nach Funke & Spering 2006, S. 51–52), die im Folgenden kurz dargestellt werden: (1) Um den Interpretationsraum von Blickbewegungen zu beschränken, müssen bei Blickbewegungsuntersuchungen die perzeptuellen und okulomotorischen Prozesse sowie deren Berechnungsmechanismen bekannt sein. (2) Da Blickbewegungen serielle Vorgänge sind, können diese nur serielle kognitive Prozesse abbilden. Bei parallelen Prozessen können aus einem seriellen Blickmuster daher nur Aussagen über einen seriellen Prozess oder über eine Überlagerung aller parallelen Prozesse abgeleitet werden. (3) Für die Ableitung mentaler Prozesse aus Blickbewegungsmustern müssen a priori Annahmen des Probanden über Zustände und Ereignisse bekannt sein sowie idealtypische Blickmuster zur Hypothesengenerierung vorliegen. (4) Die reliable Messung des visuellen Informationsgehalts einer Szene ist nicht möglich, daher müssen Vorwissen und Vorahnungen des Probanden bekannt sein. (5) Der Informationsgehalt einer Oberfläche darf nur durch inhaltliche und nicht durch physikalische Eigenschaften (z. B. Kontrast, Farbe) bestimmt werden. (6) Bei der Analyse muss beachtet werden, dass es eine empirisch belegte partielle Unabhängigkeit von kognitiven Prozessen und Blickbewegungen gibt (Anderson et al. 2004) und dass Fixationsort und Aufmerksamkeit nicht immer koinzident sind.

Dementsprechend kann eine sinnvolle Anwendung von Methoden der Blickbewegungserfassung im Rahmen kognitionswissenschaftlicher Untersuchungen nur dann stattfinden, wenn die für die Bearbeitung einer Aufgabe zugrunde liegenden kognitiven Prozesse in Art und Ablauf im Wesentlichen bekannt sind und erhobenen Reaktionslatenzen, Fixationszeiten und Blickmustern zugeordnet werden können (Funke & Spering 2006). Die Gültigkeit der Annahmen kann in diesen Untersuchungen dann vorausgesetzt werden.

In arbeitswissenschaftlichen Untersuchungen steht die Beobachtung des Handlungsvollzugs im Mittelpunkt. Die Aufklärung kognitiver Vorgänge wird in diesem Zusammenhang nicht besonders betrachtet. Aus diesem Grund muss die Gültigkeit der Annahmen nicht geklärt werden (Rötting 2001).

2.4.2 *Parameter der Blickbewegung*

Die wichtigsten Parameter der Blickbewegungsmessung stellen die Sakkaden, die Fixationen und der sich aus diesen Parametern zusammensetzende Blickpfad dar. Neben diesen Parametern gibt es eine Vielzahl von Parametern (z. B. räumliche Dichte der Fixationen, Länge des Blickpfades, lokale und globale Fixationspfade), die einen Einblick in die verschiedenen Betrachtungsebenen von Arbeitsprozessen sowie kognitiver Arbeitsstrukturen erlauben.

Sakkaden sind schnelle, ruckartige Bewegungen des Auges, die unwillkürlich oder gezielt ausgelöst werden können (Carpenter 1988) und dazu dienen, Informationen auf die Fovea, den Punkt des schärfsten Sehens des Auges, zu bringen. Sie können beispielsweise bezüglich ihrer räumlichen und zeitlichen Eigenschaften, ihrer Schnelligkeit und ihres zeitlichen Verlaufs analysiert werden. Die durchschnittliche Sakkadenweite liegt bei 15 Grad und die Sakkadendauer zwischen 50 und 60 Millisekunden. Eine Übersicht zu den wichtigsten Eigenschaften der Parameter, ihrer Messung und Analyse ist in Rötting (2001) zu finden.

Fixationen stellen keine Bewegung des Auges dar, sondern sind als Zustand des Auges definiert, „bei dem das Auge sich bezüglich eines Sehobjektes in ‚relativem‘ Stillstand befindet“ (Rötting 2001, S. 68). Dabei wird zwischen dem relativen Stillstand des Auges und dem relativen Verweilen auf einem Sehobjekt unterschieden. Diese Unterscheidung muss bei der Operationalisierung der Fixationen auf Basis der Sakkaden durchgeführt werden (z. B. örtliche und zeitliche Abweichungen der Sakkaden). Die Bestimmung der Fixationsdauer hat einen entscheidenden Einfluss auf die ermittelten Daten. Als minimale Fixationsdauern werden in der Literatur Werte zwischen 60 Millisekunden und 120 Millisekunden angegeben, die je nach Untersuchung angepasst werden können (Rötting 2001). Wichtige Aspekte der Analyse sind die Fixationsdauer und die Fixationshäufigkeit eines Sehobjektes oder einer definierten Region des Sehfeldes.

Der Wechsel von Sakkaden und Fixationen wird als Blickpfad definiert und erlaubt den Vergleich des Blickverhaltens zwischen verschiedenen Beobachtern und unterschiedlichen Stimulusmaterialien sowie die Quantifizierung von beobachteten Blickverhalten bezüglich eines hypothetischen optimalen Blickverhaltens.

Parametersystematik

Für die Nutzung von Augen- und Blickbewegungen in arbeitswissenschaftlichen Untersuchungen schlägt Rötting (2001) eine Parametersystematik vor. Diese klassifiziert die Parameter der Augen- und Blickbewegung zum einen nach Eigenschaften der Parameter (inhärente Parametersystematik) und zum anderen nach einer anwendungsorientierten Systematik (applikative Parametersystematik).

In der anwendungsorientierten Systematik werden die Parameter drei aufgabenanalytischen Ebenen zugeordnet: (1) der Bestimmung der physiologischen Kosten bei der Aufgabebearbeitung, (2) der Erklärung, Bemessung und Prognose von Zeitverbräuchen und (3) die Analyse der Bedingungen der Mensch-Maschine-Interaktion. Jede Ebene ist in Untergruppen eingeteilt, die die Auswahl geeigneter Parameter für einen speziellen Fall erleichtern (siehe Tabelle 3).

In dieser Arbeit wird die Ebene der Analyse der Bedingungen der Mensch-Maschine-Interaktion weiter betrachtet. Diese erlaubt die „analytische Beschreibung und damit die Möglichkeit zum Vergleich unterschiedlicher Bedingungen der Mensch-Maschine-Interaktion in mehreren Dimensionen (z. B. Anforderungen, Stabilität, Fehler)“ und die Beschreibung und Analyse des Handlungsvollzugs bei der Interaktion mit Mensch-Maschine-Systemen (Rötting 2001, S. 173). Das Ziel ist die Bewertung von Mensch-Maschine-Systemen auf der Grundlage von simulierten Blickbewegungen kognitiver Modelle anhand der vorgeschlagenen Parameter.

Die dritte Ebene ist unterteilt in Parameter für (1) die Analyse der örtlich-räumlichen Gestaltung, (2) der Analyse von Suchprozessen, (3) der Analyse von Informationsaufnahme- und -verarbeitungsprozessen und (4) der Analyse der Ebene der Informationsverarbeitung (siehe Tabelle 3). Eine Kombination von Parametern aller vier Gruppen erlaubt somit eine umfassende Untersuchung von Schnittstellen von Mensch-Maschine-Systemen.

2.4.3 *Methoden der Blickbewegungserfassung*

Für die Registrierung von Augen- und Blickbewegungen gibt es verschiedene Verfahren, die je nach Untersuchungsgegenstand eingesetzt werden können. Umfassende Darstellungen sind zu finden bei Carpenter (1988) und Rötting (2001). Für die Erfassung der Augen- und Blickbewegungen werden heute zumeist zwei Klassen von Verfahren genutzt. Die erste Klasse beschreibt die getrennte Erfassung der horizontalen und vertikalen Bewegungen des Auges. Diese Methoden beruhen auf analogen Messwerten, wie zum Beispiel elektronische Spannung, die mittels computergestützter Methoden weiterverarbeitet werden können. Die zweite Klasse fasst videobasierte Methoden zusammen. Mittels lichtempfindlicher Sensoren wird ein Bild des Auges aufgezeichnet und einer computergestützten Bildverarbeitung zugeführt, um charakteristische Merkmale zu extrahieren und die Augenbewegung zu bestimmen (vgl. Funke & Spring 2006; Rötting 2001).

Ein Beispiel der ersten Klasse stellt die Methode des Elektrookulogramms dar. Die Methode basiert auf der Potentialdifferenz zwischen dem negativen Pol der Retina (Netzhaut) und dem positiven Pol der Kornea (Hornhaut). Das Auge kann daher als Dipol angenommen werden. Unter Verwendung von Oberflächen Elektroden kann eine Änderung des elektrischen Feldes bei der Positionsänderung des Auges gemessen und die Augenbewegung aufgezeichnet werden. Aufgrund von Einschränkungen des Messbereichs und der geringen örtlichen Genauigkeit bei Gesichtsfeldern über 30 Grad ist die Elektrookulogramm-Messung im Bereich der Kognitionspsychologie nur bedingt einsetzbar und wird eher im klinischen Kontext eingesetzt.

Beispielhaft für die zweite Klasse steht die Korneareflex-Methode. Bei dieser Methode wird die Augenbewegung durch die Reflexion von Licht auf der Kornea erfasst. Dabei wird ausgenutzt, dass sich das auf die Kornea auftreffende Licht in einer punktförmigen Quelle, dem Korneareflex oder erstem Purkinje-Bild, spiegelt und ca. 3,5 Millimeter hinter der Augenoberfläche zu liegen scheint. Durch eine

Tabelle 3: Parametersystematik der Verwendbarkeit von Parametern der Augen- und Blickbewegungen im Rahmen von arbeitswissenschaftlichen Aufgabenanalysen (Rötting 2001, S. 187ff.).

Ebene 1 – Bestimmung der „physiologischen Kosten“	
Veränderungen des Arousal • Sakkadenlatenz • Anzahl der Sakkaden • Sakkaden pro Zeiteinheit • Fixationsdauer • Fixationen pro Zeiteinheit Ermüdung und Vigilanz • Sakkadengeschwindigkeit • Sakkadendauer • Genauigkeit der Sakkade • Anzahl glissadischer Augenbewegungen	Veränderungen des visuellen Feldes • Sakkadenweite • Sakkadenausdehnung Veränderungen des Blickverhaltens bei einer zusätzlichen Aufgabe • Übergangshäufigkeiten • Entropie • Markov-Matrizen • „Aufwand“ der Bewegung des Auges • Fixationsvektor • Gewichtetes Suchgebiet
Ebene 2 – Erklärung, Bemessung und Prognose von Zeitverbräuchen	
Dauer von Informationsaufnahme und -verarbeitungsprozessen • Fixationsdauer • Fixationen pro Zeiteinheit • Sakkaden pro Zeiteinheit • Kumulierte Fixationsdauer • Verweildauer	Dauer von Übergangszeiten • Sakkadendauer • Sakkadenweite • Kumulierte Übergangszeit • Fixationen-Sakkaden-Verhältnis • Dauer von Suchzeiten • Blickpfaddauer • Suchzeit
Ebene 3 – Analyse der Bedingungen der Mensch-Maschine-Interaktion	
Analyse der örtlich-räumlichen Gestaltung • Verteilung der Augenbewegungsdaten • Fixationsort • Fixationsdichte • Räumliche Dichte • Kumulierte Fixationsdauer • Verweildauer • Relative Häufigkeit der Fixation von Objekten oder Regionen • Übergangshäufigkeiten Analyse von Suchprozessen • Anzahl der Fixationen • Anzahl der Sakkaden • Länge des Blickpfades • Übergangsdichte • Konvexe Hüllfläche • Sakkadenweite • Kumulierte Übergangszeit • Fixationen-Sakkaden-Verhältnis • Augenbewegungsgeschwindigkeit	Analyse von Informationsaufnahme und -verarbeitungsprozessen • Statistische Analyse der Übergangshäufigkeiten • Sakkadenweite-Fixationsdauer-Diagramm • Autokorrelation • Räumliche Clusterung • Lokaler Fixationspfad • Richtungsänderungen • Markov-Matrizen • Entropie • Veränderungen der relativen Häufigkeit der Fixation von Objekten oder Regionen Analyse der „Ebenen“ der Informationsverarbeitung • Fixationsdauer • Fixationen pro Zeiteinheit • Sakkaden pro Zeiteinheit • Verteilung der Fixationsdauer • Laterale Augenbewegungen

Augenbewegung kommt es zu einer Verschiebung des Korneareflexes, die mit einem videobasierten System aufgezeichnet werden kann. Es werden kopfgestützte, das heißt die Videokameras sind direkt am Kopf befestigt, und berührungslose Systeme für die Erfassung des Korneareflexes angeboten. Da auf diese Weise nur die Bewegung des Auges relativ zum Kopf gemessen wird, müssen zusätzlich Methoden für die Registrierung der Kopfposition genutzt werden (z. B. mittels optischer, elektromagnetisch- oder ultraschallbasierter Verfahren). Kopfgestützte Systeme sind genauer in der Blickregistrierung als berührungslose Systeme und werden in der Kognitionspsychologie aus diesem Grund bevorzugt.

Auswertung von Blickbewegungen

Die in Blickbewegungsuntersuchungen erhobenen Blickbewegungsdaten liegen in einer hohen zeitlichen und räumlichen Auflösung vor. Während die Erhebung der Daten mit heutigen Erhebungstechniken im Bezug zu Aufwand und Zeitverbrauch relativ einfach durchzuführen ist, stellt die Auswertung eine Schwierigkeit dar. Gründe dafür sind die hohe Auflösung der Daten und somit die Datenmenge, das mögliche Verrauschen der Daten und eine daraus resultierende Anpassung sowie die erhebliche individuelle Variabilität von unterschiedlichen Probanden (Funke & Spering 2006). Trotz der weiten Einsatzgebiete von Blickbewegungsmessungen existieren für die automatische Analyse von Blickbewegungen nur wenige formale Methoden. Die meisten Analysemethoden werden informell und nur mit wenigen Details beschrieben (Salvucci & Anderson 2001).

Blickbewegungsdaten können automatisiert mit gängigen Auswertungsprogrammen (wie z. B. MATLAB; The MathWorks 2003) ausgewertet werden. Zu diesem Zweck müssen eigenständige Algorithmen entwickelt werden, die für einen Anwendungsfall gelten und nur selten direkt auf weitere Untersuchungen übertragbar sind. Einige Hersteller von Blickbewegungsmessgeräten bieten eigene Auswertungsprogramme an (z. B. BeGaze; SensoMotoric Instruments 2006), die eine erste Auswertung der Daten ermöglichen. Für kognitionswissenschaftliche oder arbeitswissenschaftliche Fragestellungen können diese Verfahren allerdings nur bedingt eingesetzt werden, da allein grundlegende Parameter ausgewertet werden.

Neben diesen Anwendungen gibt es Verfahren, die die Berechnung von Sequenzen von Sakkaden und Fixationen aus den Rohdaten erlauben (Cabiati, Pastormerlo, Schmid & Zambambieri 1983; Stark & Ellis 1981; Tole & Young 1981). Weitere Ansätze versuchen anhand von *Tracing*-Verfahren Blickbewegungen mit den zugrunde liegenden kognitiven Verfahren zu verknüpfen (z. B. die auf kognitiven Modelldaten beruhende Protokollanalyse; Salvucci & Anderson 2001).

Hinsichtlich der Parameter von Augen- und Blickbewegungen müssen die gewählten Verfahren jeweils neu umgesetzt, implementiert und an die jeweilige Untersuchung angepasst werden.

2.4.4 *Simulation von Blickbewegungen*

Einige kognitive Architekturen erlauben die Simulation von Blickbewegungen auf der Basis von kognitionspsychologischen Theorien (Anderson & Lebiere 1998; Meyer & Kieras 1997a). Dabei werden nicht die Augenbewegungen simuliert, sondern es wird determiniert, ob, wann, wie und in welchem Detaillierungsgrad ein Objekt wahrgenommen werden kann. Dabei werden neben menschlichen Beschränkungen (z. B. nur ein visuelles System) heuristische Maße genutzt, um beispielsweise die Geschwindigkeit der Aufmerksamkeitsverschiebung oder die Fixationsdauer von Objekten zu berechnen. Künstliche Blickbewegungen in Simulationsexperimenten erfüllen die Anforderungen an eine Theorie der Blickbewegungen als Indikator für kognitive Prozesse (siehe Kapitel 2.4.1) und können daher für kognitionspsychologische und arbeitswissenschaftliche Untersuchungen genutzt werden. Es bleibt zu prüfen, inwieweit kognitive Architekturen die Phänomene der natürlichen Blickbewegungen simulieren und inwieweit diese für die Bewertung von Mensch-Maschine-Systemen genutzt werden können.

In ACT-R ist eine Theorie der visuellen Aufmerksamkeit implementiert, die es kognitiven Modellen in der kognitiven Architektur ACT-R ermöglicht, Umgebungsinformationen wahrzunehmen und in höheren kognitiven Prozessen zu verarbeiten.

Das visuelle System von ACT-R

Das visuelle System von ACT-R besteht zum einen aus einer Theorie der visuellen Wahrnehmung und der visuellen Aufmerksamkeit und zum anderen aus deren Bezug zu Prozessen der höheren Kognition (Anderson & Lebiere 1998). Das Ziel der perzeptuellen Komponente ist das Bereitstellen einer psychologisch plausiblen Theorie, die keine neue Theorie darstellt, sondern auf Vorhandenem aufbaut. Aus diesem Grund wurde in ACT-R eine Theorie integriert, die als Synthese der *Attentional-spotlight*-Theorie (Posner 1980), der *Feature-synthesis*-Theorie (Treisman & Sato 1990) und der *Attentional*-Theorie (Wolfe 1994) angesehen werden kann.

Konzeptionell basiert das visuelle System auf der perzeptuellen Komponente von EPIC. In EPIC wurde eine erfolgreiche Strategie entwickelt, die es erlaubt, ohne reale Sensoren und ohne die Integration aller Details der menschlichen Perzeption Blickbewegungen abzubilden. Dabei werden die grundlegenden Ausführungszeiten und die Ausgaben des perzeptuellen Systems in angenäherter Form berechnet (Meyer & Kieras 1997b).

In ACT-R wurde dieser Ansatz in Teilen übernommen und um neue Aspekte erweitert. Das visuelle System von ACT-R ist im Gegensatz zu EPIC in zwei Module mit jeweiligen Buffern unterteilt. Das *Visual-location*-Modul repräsentiert das dorsale *Where*-System und ist für die visuelle Suche und die Lokalisierung von Objekten in der Szene verantwortlich. Das *Visual-object*-Modul repräsentiert das ventrale *What*-System und ermöglicht den Wechsel von Aufmerksamkeit zwischen Objekten oder einer visuellen Ortsangabe (Anderson et al. 2004).

Macht eine Produktion eine Anfrage an das *Where*-System, dann wird basierend auf den angegebenen Beschränkungen der Produktion eine

Ortsangabe (in Form eines Chunks) zurückgegeben. Beschränkungen sind visuelle Eigenschaften des angefragten Objekts (z. B. Farbe, Text) oder die örtliche Ausrichtung (z. B. oben). Sind mehrere passende Objekte vorhanden, wird ein Objekt zufällig ausgewählt.

Durch das *Where*-System hat ein kognitives Modell in ACT-R das Wissen über den Ort und einige visuelle Eigenschaften (z. B. Farbe, Text) aller Objekte, die für das System momentan zugänglich sind. Um ein Objekt zu identifizieren, muss das kognitive Modell eine Anfrage an das *What*-System stellen. Die Anfrage muss einen Chunk mit der jeweiligen Ortsangabe enthalten. Das *What*-System richtet die visuelle Aufmerksamkeit auf den in dem Chunk repräsentierten Ort, verarbeitet das Objekt und speichert es in Form eines Chunks im deklarativen Gedächtnis ab.

Die Wahrnehmung von Umgebungsinformationen ist in ACT-R durch die aktive Aufmerksamkeitsverschiebung des visuellen Systems realisiert. Dazu müssen kognitive Prozesse aktiv ausgeführt werden, um den Aufmerksamkeitsfokus zu ändern und die Umgebungsinformation wahrzunehmen. Weitere Wahrnehmungsprozesse, wie Aufmerksamkeitsablenkung und peripheres Sehen, sind in der aktuellen Version von ACT-R noch nicht formalisiert. Infolgedessen ist in ACT-R zusammenfassend eher eine Theorie der visuellen Aufmerksamkeit als eine Theorie der Wahrnehmung implementiert.

ERWEITERUNGEN DES PERZEPTUELLEN SYSTEMS Für das perzeptuelle System von ACT-R wurden in letzter Zeit verschiedene Erweiterungen entwickelt, die Eigenschaften der menschlichen Informationsaufnahme und -verarbeitung abbilden, die bisher in ACT-R nicht integriert waren.

Das berechenbare Modell EMMA (*Eye Movements and Movement of Attention*; Salvucci 2001a), das der kognitionspsychologisch begründeten Verknüpfung von visueller Aufmerksamkeit und Augenbewegung Rechnung trägt, kann beispielsweise in ACT-R integriert werden. Dadurch wird die Theorie der visuellen Aufmerksamkeit in ACT-R um Phänomene der Augenbewegung erweitert (z. B. ausgelassene oder mehrere Fixationen auf einem Objekt).

Neben den Ansätzen zur Integration der Augenbewegung werden ebenfalls Ansätze vorgeschlagen, die die visuellen Suchprozesse von Objekten in einer Szene betreffen, um beispielsweise die Suche in komplexen Szenen zu unterstützen. Dabei werden die vorhandenen Methoden der visuellen Suche um weitere Aspekte ergänzt oder neue Methoden entwickelt (Nicholson, Byrne & Fotta 2006). Beispielsweise werden Verfahren vorgeschlagen, die es ermöglichen, den Suchraum mehrfach zu klassifizieren (lokale und globale Eigenschaften von Objekten) oder angereicherte Attribute für die Suche zu verwenden (z. B. Suche im Umkreis von 3 cm). Problematisch bei diesen Ansätzen ist, dass der Modellierer die Klassifizierungen des Suchraums und die Strategien der Blickbewegungen schon bei der Entwicklung eines Modells durchführen muss. Die Blickbewegungen basieren demnach auf der Aufgabenanalyse und den subjektiven Einflüssen des Entwicklers. Somit sind die Blickbewegungen keine generische Eigenschaft der kognitiven Architektur, sondern werden aus den Erfahrungen und Beobachtungen der Entwickler abgeleitet.

ACT-R UND REALE EXPERIMENTALUMGEBUNGEN ACT-R bietet die Möglichkeit, dass empirische Untersuchungen und die Simulation von Nutzerverhalten an einer Experimentalumgebung durchgeführt werden können. Das bedeutet, dass künstliche und natürliche Systeme mit einer Umgebung interagieren. Zu diesem Zweck werden einfache Elemente (z. B. Text, Druckknöpfe, Linien) angeboten (Anderson & Lebiere 1998).

In der Forschung werden momentan Werkzeuge und Programme entwickelt, die eine Anbindung von kognitiven Modellen in ACT-R an reale Prototypen ermöglichen (vgl. Urbas & Leuchter 2005; Riedl & St. Amant 2002; Halbrügge, Deml, Färber & Bardins 2007).

Das Programm AGImap (*ACT-R Graphical Interface Mapper*; Urbas & Leuchter 2005) ermöglicht die Kommunikation von kognitiven Modellen mit einer realen Experimentalumgebung. Dazu werden graphische Elemente der realen Benutzungsschnittstelle auf die der Umgebung von ACT-R übertragen und über eine Kommunikationsebene aktualisiert. Zusätzlich werden generalisierte Methoden in Form von Produktionen angeboten, die die Interaktion mit den Elementen der Benutzungsschnittstelle erlauben. Dabei werden beispielsweise Methoden für die visuelle Informationsaufnahme oder das Betätigen eines Druckknopfes bereitgestellt.

Ein weiterer Ansatz basiert auf einer automatischen Segmentierung der Benutzungsschnittstelle und einer anschließenden Bereitstellung der Informationen in Form von durch ACT-R interpretierbaren Schnittstellenbeschreibungen (SegMan; Riedl & St. Amant 2002).

Eine Mischform aus den zwei zuvor vorgestellten Ansätzen stellt ACT-CV (Halbrügge et al. 2007) dar. Das System erlaubt die Interaktion von ACT-R mit beliebigen graphischen Benutzungsschnittstellen. Die Inhalte der Schnittstelle werden mittels maschineller Bildverarbeitung ausgelesen und anhand von vorher erstellten Bilddaten abgeglichen. Die erkannten Inhalte werden an ACT-R übergeben.

Die einzelnen Systeme befinden sich noch in frühen Phasen der Entwicklung. Im Gebiet der komplexen und dynamischen Mensch-Maschine-Schnittstellen ist das Werkzeug AGImap geeignet, um dynamische Informationen zwischen ACT-R und der realen Repräsentation auszutauschen (z. B. Veränderungen von Anzeigen). ACT-CV hat bezüglich der Bearbeitung von dynamischen Inhalten noch nicht genügend schnelle Algorithmen, und es ist fraglich, wie dynamische Inhalte erkannt und übertragen werden. SegMan stellt eine mögliche Variante dar, ist heute aber noch nicht für die aktuelle Version von ACT-R verfügbar.

2.5 ABLEITEN DER FRAGESTELLUNG

Kognitive Modellierungsmethoden stellen in Form von formalen, quantitativen und ablauffähigen Computermodellen eine Evaluierungsmethode von Mensch-Maschine-Systemen dar, die kognitionspsychologische Aspekte integriert und eine einfache und kostengünstige Wiederhol- und Überprüfbarkeit von verschiedenen Gestaltungsalternativen ermöglicht. Die Anwendung von kognitiven Modellierungsmethoden für die Bewertung von Mensch-Maschine-Systemen kann helfen, bereits in frühen Phasen die Komponente Mensch in den Entwicklungsprozess von Mensch-Maschine-Systemen zu integrieren. Der

Entwicklungsaufwand in nachfolgenden Phasen kann reduziert und die Entwicklung von besser an den Menschen angepassten Systemen gefördert werden.

Ausgehend von den in diesem Kapitel vorgestellten Rahmenbedingungen ist zu überprüfen, inwieweit simulierte Leistungsdaten kognitiver Modelle auf einer feinen Auflösungsebene für die formale Bewertung von Mensch-Maschine-Systemen in frühen Systementwicklungsphasen genutzt werden können.

Die Analyse von Blickbewegungsdaten wird in der Forschung und der Industrie genutzt, um Mensch-Maschine-Systeme zu bewerten und um wissenschaftliche Fragestellungen zu untersuchen (siehe Kapitel 2.4). Blickbewegungsdaten können von natürlichen und künstlichen Systemen erfasst werden und stellen gegenüber der Betrachtung von Performanzdaten eine feinere Betrachtungsebene dar. Aus diesem Grund werden Performanzdaten und Blickbewegungsdaten im Rahmen dieser Arbeit genutzt, um die Fragestellungen zu untersuchen. Für die Simulation der Blickbewegungen wird die kognitive Architektur ACT-R eingesetzt. Die Analyse der Blickbewegungsdaten orientiert sich an der dritten Ebenen der vorgestellten applikativen Parametersystematik, die eine Analyse und Bewertung der Bedingungen der Mensch-Maschine-Interaktion ermöglicht. Dabei werden Parameter, die Sakkaden betreffen, nicht betrachtet, da Sakkaden von ACT-R nicht simuliert werden. Ebenfalls werden Augenbewegungen nicht behandelt, da die Analyse hinsichtlich der visuellen Aufmerksamkeit bei der Interaktion mit Mensch-Maschine-Systemen für eine erste Bewertung von Gestaltungsalternativen in frühen Systementwicklungsphasen ausreichend ist und Augenbewegungen einen zu feinen Auflösungsgrad für die Bewertung darstellen würden. Erweiterungen von ACT-R, wie EMMA, werden daher für die Simulation des Blickverhalten nicht benötigt.

In dieser Arbeit wird überprüft, ob die Analyse und der Vergleich von simulierten Blickbewegungen gegenüber der auf einfachen heuristischen Maßen beruhenden Bewertung von Mensch-Maschine-Systemen (wie z. B. Fehler und Ausführungszeiten) eine Erweiterung des heutigen Methodenrepertoires darstellt. Eine wichtige Frage in diesem Zusammenhang ist, ob die kognitive Architektur ACT-R das menschliche Blickverhalten adäquat genug abbilden kann, um die Anwendung von feinen Leistungsdaten für die Bewertung von Mensch-Maschine-Systemen zu ermöglichen. Dabei ist eine Abweichung von bis zu 20 % von empirisch erfassten Daten für die Bewertung in frühen Phasen der Systementwicklung möglich, da Prozessparameter und Simulationsparameter in diesen Phasen mit ähnlicher Abweichung zu dem zukünftigen Endprodukt vorliegen (Card et al. 1983; Urbas, Schulze-Kissing & Leuchter 2005).

Des Weiteren wird überprüft, ob unterschiedliche theoretische Annahmen zu Informationsaufnahme- und -verarbeitungsprozessen, die in kognitiven Modellen formalisiert sind und auf der Betrachtungsebene der Performanzdaten ähnliche Leistungsdaten simulieren, zu unterscheidbaren Blickdaten führen. Diese Überprüfung ist notwendig, um zu zeigen, dass die Betrachtung von feinen Leistungsdaten einen Vorteil für die Bewertung von kognitiven Modellen darstellt.

Abschließend wird geprüft, ob die simulationsgestützte Bewertung von unterschiedlichen Schnittstellen für ein Mensch-Maschine-System anhand von kognitiven Modellen möglich ist und zu empirisch vergleichbaren Ergebnissen führt.

Für die prozessbegleitende Unterstützung der Analyse und Vergleiche von simulierten und empirischen Blickbewegungsdaten wird ein Werkzeug entwickelt, das in prospektiven Phasen der Systemgestaltung eingesetzt werden kann und eine erste Bewertung von Gestaltungsalternativen auf Basis von simulierten Leistungsdaten ermöglicht. Dabei ist zu beachten, dass die Ergebnisse des Werkzeugeinsatzes so genau wie möglich, aber nur so genau wie nötig sind, um für den Einsatz in prospektiven Systementwicklungsphasen anwendbar zu sein.

Zusammenfassend wird in dieser Arbeit überprüft, ob die Bewertung von Mensch-Maschine-Schnittstellen anhand von kognitiven Benutzermodellen für fertigungs- und regelbasierte Tätigkeiten möglich ist. Im Speziellen wird dies an dem Beispiel von Blickbewegungsdaten überprüft, die mit einem Werkzeug aufbereitet und verdichtet werden.

In diesem Kapitel wird der *Simulation Trace Analyzer* (SimTrA) vorgestellt. SimTrA ermöglicht die Analyse und den Vergleich empirischer und simulierter Blickbewegungsdaten und unterstützt den dialektischen Spiralprozess der kognitiven Modellbildung in zwei Aspekten: Zum einen bei dem Vergleich von simulierten und empirischen Daten für die Geltungsprüfung der Modellierung und zum anderen bei der Datenanalyse empirischer Untersuchungen (siehe Kapitel 2.2.2). Durch die automatisierte Analyse simulierter Leistungsdaten und deren Vergleich mit empirischen und simulierten Daten wird die Bewertung von Mensch-Maschine-Systemen und die Bewertung kognitionspsychologischer Theorien durch den Einsatz kognitiver Nutzermodelle ermöglicht.

Der Aufbau dieses Kapitels folgt einem Phasenmodell strukturierter Softwareentwicklung (Balzert 1998). In einer Voruntersuchung wurde gezeigt, dass simulierte Blickbewegungen kognitiver Modelle mit bekannten Analysemethoden verarbeitbar sind. Darauf aufbauend werden Anforderungen an ein Werkzeug zur Datenanalyse definiert und SimTrA konzeptionell beschrieben. Anschließend wird die Realisierung von SimTrA und die Implementierung dargestellt. Das Kapitel schließt mit der formalen Verifikation der Implementierung des Werkzeuges SimTrA.

3.1 VORUNTERSUCHUNG

Vor der Konzepterstellung und der prototypischen Realisierung von SimTrA wurde überprüft, ob simulierte Blickbewegungsdaten kognitiver Modelle der kognitiven Architektur ACT-R verwertbare Daten zu Verfügung stellen, um Methoden für die Analyse von Blickbewegungen anzuwenden. Das Ziel der Voruntersuchung war, zu zeigen, dass ein kognitives Modell, das auf globaler Ebene im Vergleich zu empirischen Leistungsdaten ein ähnliches Verhalten zeigt, ebenfalls auf der Ebene der Blickbewegungen vergleichbare Leistungsdaten simuliert. Des Weiteren sollte vor der prototypischen Entwicklung von SimTrA die Analyse und der Vergleich kognitiver Modelldaten und empirischer Leistungsdaten betrachtet und angewendet werden. Zu diesem Zweck wurden simulierte und empirisch erfasste Blickbewegungsdaten bei einer Kontroll- und Steuerungsaufgabe mit einem Ausschnitt aus einem Prozesskontrollsystem erhoben, analysiert und die Ergebnisse verglichen.

3.1.1 Untersuchungsaufgabe

Die Kontroll- und Steuerungsaufgabe besteht in der Interaktion eines Operators mit einer komplexen und dynamischen Schnittstelle zur Steuerung einer Destillationskolonne (siehe Abbildung 5). Die Aufgabe

stellt einen Ausschnitt aus einem Aufgabenszenario für die Untersuchung der Zeitschätzung von Operateuren in komplexen und dynamischen Mensch-Maschine-Schnittstellen dar (Schulze-Kissing 2007). Die zu bearbeitende Aufgabe besteht darin, den chemischen Prozess der Destillation zu starten und den Flüssigkeitspegel zwischen zwei Pegelgrenzen zu stabilisieren, wobei die Heizung für den Destillationsprozess angeschaltet sein muss. Durch das Regulieren von Heizung und Regelventil kann der Flüssigkeitspegel beeinflusst und gesteuert werden. Ist die Heizung angeschaltet, wird mit einer kurzen Verzögerung ein Verdampfungsprozess der Flüssigkeit gestartet und der Pegel sinkt. Befindet sich der Pegel unterhalb der unteren Grenze, wird die Heizung automatisch ausgeschaltet, um eine Beschädigung der Anlage zu verhindern. Durch das Ausschalten der Heizung wird der Verdampfungsprozess gestoppt. Durch das Öffnen und Schließen des Regelventils wird der Abfluss aus dem Tank gesteuert. Mit dem Öffnen des Regelventils fließt eine festgelegte Menge der Flüssigkeit aus dem Tank ab. Durch das Schließen des Regelventils wird der Abfluss gestoppt.

Der Benutzer interagiert mit der Schnittstelle durch die der Maus, indem auf die dargestellten Symbole für die Heizung (H1202) und das Regelventil (S1202) geklickt wird (siehe Abbildung 5). Das Klicken ändert beispielsweise den aktuellen Zustand der Heizung von an nach aus beziehungsweise des Regelventils von offen nach zu.

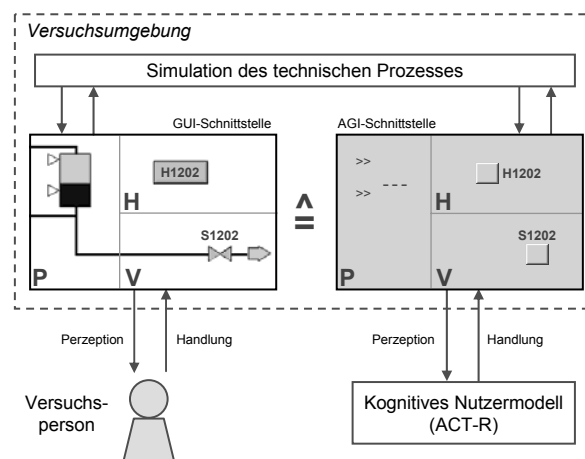


Abbildung 5: Schematische Darstellung der Empirie (links) und des Simulationsexperiments (rechts). P: Pegel, V: Regelventil, H: Heizung, GUI: Graphical User Interface, AGI: ACT-R Graphical Interface.

Die technischen Prozesse der Destillationskolonne und die Bedienschnittstelle des Menschen und des kognitiven Modells wurden getrennt voneinander implementiert. Für das Simulationsexperiment wurde die menschliche Bedienschnittstelle (GUI: Graphical User Interface) in eine für das Benutzermodell lesbare Notation überführt (AGI: ACT-R Graphical Interface) und an die Prozesse der Destillationskolonne angeschlossen (siehe Abbildung 15). Dafür wurde das Werkzeug AGImap (siehe Kapitel 2.4.4) verwendet.

3.1.2 Untersuchungsgruppen

Für die Untersuchung wurden die Leistungsdaten eines kognitiven Modells und von Versuchspersonen erfasst. Diese Daten bildeten die Grundlage für die Analyse und den Vergleich der simulierten und empirischen Leistungsdaten bei der Interaktion mit der bereitgestellten Mensch-Maschine-Schnittstelle.

Kognitives Modell

Für das Simulationsexperiment wurde in ACT-R 6.0 ein kognitives Modell zur Steuerung der simulierten Destillationskolonne erstellt. Das Modell wurde mit den Wissensstrukturen eines erfahrenen Benutzers des simulierten Prozesses ausgestattet, das in einer vorangegangenen Untersuchung mit Experten erhoben wurde. Um das Modell an das durch externe Reize geleitete Verhalten des Menschen (bottom-up) anzupassen, wurde es nach dem Prinzip der minimalen Kontrolle der internen Zustände implementiert (Taatgen 2007). Das bedeutet, dass Abläufe und die Auswahl von Produktionen nicht in einer festgelegten Sequenz innerhalb des kognitiven Modells kodiert sind. Die Auswahl der Produktionen und das Verhalten wird durch die aktuelle Belegung der Buffer (siehe Kapitel 2.2.1) determiniert. Somit stellen die wahrgenommen Systemzustände der Schnittstelle (z. B. Heizung an oder aus) und deren interne Repräsentation die Grundlage für das simulierte Verhalten dar.

Stichprobe

Die empirische Studie wurde mit 28 Studenten der Technischen Universität Berlin durchgeführt (12 weibliche und 16 männliche Versuchspersonen; Altersdurchschnitt: 24,3 Jahre). Sie erhielten 10 Euro für die Teilnahme an dem Versuch.

3.1.3 Untersuchungsdesign

Als unabhängige Variable wurden zwei verschiedene Szenarien entwickelt, die sich in dem Verhältnis von Zu- und Abfluss der Flüssigkeit in den Behälter unterscheiden. Dies führt zu einer unterschiedlichen Anzahl von notwendigen Interaktionen in der vorgegebenen Zeit von 90 Sekunden bei der korrekten Steuerung der Destillationskolonne (Szenario 1: n Interaktionen; Szenario 2: zweimal n Interaktionen).

Als abhängige Variablen wurden die Blickbewegungen (X- und Y-Koordinaten und Zeitstempel) während der Interaktion und die Performanz der Benutzer erhoben.

Die Untersuchung wurde im Zeitraum von drei Wochen im März und April 2006 an der Technischen Universität Berlin durchgeführt. Jede Versuchsperson bearbeitete die Aufgabe in beiden Szenarien. Die Versuchsdauer lag bei durchschnittlich 30 Minuten.

3.1.4 Untersuchungsablauf

Die Untersuchung begann mit einer Einführung der Versuchsperson in den Untersuchungsgegenstand. Anschließend wurde die Versuchsperson zwei Meter von der Projektion der Aufgabenumgebung platziert und der Helm für die Erfassung des Augenbildes aufgesetzt und festgeschnallt. Nach der Kalibrierung des Systems für die Blickbewegungsdatenerfassung wurde ein Training für den Umgang mit der Aufgabe durchgeführt, um die wichtigsten Funktionen zu erlernen. Die Aufgabe wurde als gelernt betrachtet, wenn die Versuchsperson den Flüssigkeitspegel dreimal für 30 Sekunden mit eingeschalteter Heizung zwischen den festgelegten Pegelgrenzen stabilisieren konnte. Daran schlossen die zwei verschiedenen Aufgabenszenarien mit Blickbewegungsdatenerfassung an. Jedes Szenario wurde 90 Sekunden lang bearbeitet. Der Versuch schloss mit einer Information über die weitere Verwendung der erfassten Daten.

Für die Simulationsdaten wurde das kognitive Modell dreimal 90 Sekunden ausgeführt, um überprüfen zu können, ob das Modell unterschiedliches Verhalten in der Steuerung der Prozesskolonne zeigt.

3.1.5 Geräte

Die Blickbewegungserfassung wurde mit der iView X HED (*head mounted eye tracking device*) der Firma SMI (SensoMotoric Instruments 2007) durchgeführt. Dabei handelt es sich um ein kopfgestütztes System zur Erfassung des Korneareflexes (siehe Kapitel 2.4.3). Die Aufnahmezeit der Blickbewegungen liegt bei 50 Hz (d. h. alle 20 Millisekunden wird der Blickpunkt erfasst). Das Blickbewegungsmesssystem wurde mit einer 13-Punkt-Kalibrierung eingemessen.

3.1.6 Untersuchungsauswertung

Die Analyse der Performanz wurde anhand der benötigten Interaktionen und dem Stabilisierungsfaktor der Füllstandsanzeige durchgeführt, die aus den Logdateien der Prozesssimulation automatisiert ausgelesen wurden. Der Stabilisierungsfaktor stellt die relative Zeit dar, in der sich der Flüssigkeitspegel zwischen der oberen und der unteren Grenze der Füllstandsanzeige befindet.

Für die Analyse der empirischen Blickbewegungen wurden aus den Rohdaten der erfassten Blickbewegungen mittels der Analysesoftware BeGaze der Firma SMI (SensoMotoric Instruments 2006) Fixationen berechnet. Zur Identifizierung einer Fixation wurden die minimale Fixationsdauer bei 100 Millisekunden und die Abweichungen des Blickwinkels bei bis zu 2 Grad festgelegt. Diese Werte haben sich in vorangegangenen Blickbewegungsstudien bewährt und können ebenfalls in der Literatur gefunden werden (siehe Kapitel 2.4.2).

Für die Erfassung der Blickbewegungsdaten des kognitiven Modells wurde ein Skript in das kognitive Modell integriert, das zu jedem Produktionszyklus (d. h. alle 50 ms) und nach jeder ausgeführten Aktion des visuellen Moduls (siehe Kapitel 2.2.1) den Aufmerksamkeitsfokus des kognitiven Modells und den Zeitstempel erfasst. Nach

der Erfassung der Daten wurden Fixationen bestimmt. Die minimale Fixationsdauer wurde bei 100 ms festgelegt. Eine Abweichung des Blickwinkels musste nicht betrachtet werden, da der Aufmerksamkeitsfokus der kognitiven Modelle in einem fest definierten Bereich lag.

Die Analyse der Blickbewegungen wurde exemplarisch anhand der lokalen Fixationspfade, die das aktuelle perzeptuelle Geschehen widerspiegeln, untersucht (Groner, Walder & Groner 1984). Für dieses Verfahren wird das Stimulusmaterial (d. h. die Schnittstelle) in Blickgebiete (*Areas of Interest*, AOIs) eingeteilt. Die theoretisch möglichen Dreierfolgen von AOIs werden gebildet (entspricht den lokalen Fixationspfaden), und es wird geprüft, wie häufig sich jede dieser Kombinationen in dem Datenmaterial wieder findet. Sequenzen von Fixationen, die in dieselbe AOI fallen, werden dabei als eine Fixation behandelt (z. B. AOIs: 1,2 – Fixationen: 2122112 – Bereinigte Fixationen: 21212 – Dreierfolgen: 121: 1, 212: 2).

Die Schnittstelle der Steuerungs- und Kontrollaufgabe wurde in drei AOIs eingeteilt: Pegel (P), Heizung (H) und Regelventil (V; siehe Abbildung 5). Anschließend wurden die lokalen Fixationspfade für die 2 Gruppen (Modell und Mensch) berechnet. Mit den drei gewählten AOIs sind 12 verschiedene lokale Fixationspfade möglich.

3.1.7 Ergebnisse und Diskussion

Die Blickbewegungsdaten konnten von 17 Versuchspersonen korrekt aufgenommen werden. Diese Datensätze wurden für die weitere Analyse verwendet.

Bei der Performanz zeigte sich, dass auf dieser Betrachtungsebene das simulierte Benutzerverhalten das empirisch erhobene Verhalten gut simuliert (siehe Tabelle 4).

Tabelle 4: Performanz der Steuerungs- und Kontrollaufgabe der zwei Gruppen (kognitives Modell und Empirie) für die zwei Szenarien. Dargestellt sind die Mittelwerte für die Anzahl der Schaltungen der Heizung und des Regelventils und des Stabilisierungsfaktors.

	Szenario 1			Szenario 2		
	Heizung	Ventil	SF	Heizung	Ventil	SF
Kognitives Modell	1	10,33	100 %	2	21	91 %
Mensch	0,23	11,33	97 %	0,5	25,33	93 %

SF: Stabilisierungsfaktor

Die lokalen Fixationspfade der zwei Szenarien wurden verglichen, um festzustellen, ob sich diese Kennwerte aufgrund der unterschiedlichen Komplexität der Aufgabe unterscheiden. Die Unterschiede wurden nicht signifikant. Für die weitere Analyse wurden die Daten der Szenarien zusammengefasst. Da die Modelldaten kaum Varianz zeigen, können die folgenden Ausführungen nur auf einer deskriptiven Ebene gemacht werden.

Für die Analyse der Blickbewegungen wurden die relativen Häufigkeiten der lokalen Fixationspfade berechnet. Lokale Fixationspfade mit einer Häufigkeit kleiner als 3 % wurden aus der Analyse ausge-

schlossen. Die folgenden 6 wichtigsten Fixationspfade decken 85 % der Modelldaten (d. h. die Interaktionen) ab und wurden für den Vergleich Mensch und Modell benutzt: PVP, VPV, PHP, HPV, VPH und HPH.

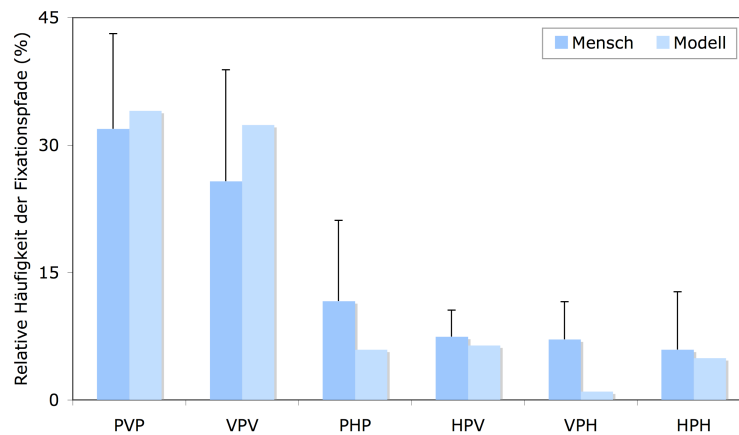


Abbildung 6: Relative Häufigkeiten der 6 wichtigsten lokalen Fixationspfade für Modell und Versuchspersonen mit Standardabweichung. H: Heizung, V: Regelventil, P: Pegel.

Es konnte festgestellt werden, dass diese sechs lokalen Fixationspfade für den Menschen ebenfalls am häufigsten auftreten (sie decken 80 % der empirischen Daten ab) und dass die ersten vier lokalen Fixationspfade von dem Modell akzeptabel simuliert wurden. Die Modelldaten liegen bei fünf lokalen Fixationspfaden innerhalb einer Standardabweichung der empirischen Daten (siehe Abbildung 6). Die durchschnittliche Abweichung der vorher aufgezeigten lokalen Fixationspfade durch das Modell von den tatsächlich beobachtbaren des Menschen liegt bei 2,8 %.

Die Ergebnisse der Voruntersuchung lassen den Schluss zu, dass simulierte Blickbewegungsdaten mit Methoden zur Analyse von Blickbewegungsdaten verarbeitet werden können. Die Analyse von feingranularen Leistungsdaten erlaubt einen besseren Einblick in das Verhalten der kognitiven Modelle und einen feineren Vergleich mit empirischen Leistungsdaten. Die Studie zeigt, dass der Vergleich von simulierten Daten kognitiver Modelle mit empirischen Daten auf deskriptiver Ebene möglich ist. Die Abweichung der Häufigkeiten von 2,8 % zwischen Modell und Empirie liegt im Rahmen der Anforderungen der Genauigkeit von Daten für die Bestimmung der Gebrauchstauglichkeit in frühen Phasen der Systemgestaltung (ca. 20 % Abweichung; Card et al. 1983; Urbas, Schulze-Kissing & Leuchter 2005). Somit können simulierte Blickbewegungen kognitiver Benutzermodelle für die Vorhersage von Benutzerverhalten in frühen Phasen der Systemgestaltung eingesetzt werden, um Mensch-Maschine-Schnittstellen zu evaluieren.

Die experimentellen Bedingungen für die softwareunterstützte Analyse von simulierten Blickbewegungen wurden erfüllt. Ein Werkzeug für die Analyse und den Vergleich von simulierten und empirisch erfassten Blickbewegungsdaten kann auf Grundlage der Voruntersuchung als Unterstützung in der kognitiven Modellbildung betrachtet werden.

3.2 ANFORDERUNGSANALYSE

Der *Simulation Trace Analyzer* soll die Analyse und den Vergleich von natürlichen und künstlichen Leistungsdaten im arbeitswissenschaftlichen und kognitionspsychologischen Kontext ermöglichen. Aus diesem Grund stellen sich neben den Anforderungen der Gebrauchstauglichkeit weitere Anforderungen an ein Werkzeug zur Verarbeitung von Leistungsdaten. Die Anforderungen wurden im engen Dialog mit Mitarbeitern der Forschungsgruppe und auf Grundlage eigener Erfahrungen in der Durchführung von Datenanalysen entwickelt. Die Arbeitsgruppe setzt sich aus Psychologen, Ingenieuren und Informatikern zusammen und stellt eine gute Auswahl der zu erwartenden Endanwender dar. Die Mitglieder der Arbeitsgruppe wurden bei der Analyse von Daten psychologischer Untersuchungen beobachtet und konnten ihre Wünsche und Vorstellungen hinsichtlich eines Werkzeugs zur unterstützenden Datenanalyse äußern. Die zusammenfassenden Anforderungen werden im Folgenden kurz dargelegt. Zur besseren Referenzierung werden die Anforderungen in jedem Abschnitt abschließend aufgelistet (dabei steht A für Anforderung).

3.2.1 Datenintegration

Die Daten von natürlichen und künstlichen Systemen werden größtenteils auf unterschiedliche Art erfasst (z. B. bei der Blickbewegungsdatenerfassung) und müssen vor einer Analyse in der Skalierung und Auflösung der erfassten Daten angepasst werden. Somit muss ein Werkzeug zur Verarbeitung von künstlichen und natürlichen Leistungsdaten die Integration von Daten unterschiedlichen Ursprungs und deren Skalierung ermöglichen.

A1 Integration von Daten unterschiedlichen Ursprungs

3.2.2 Datenverarbeitung

In psychologischen und arbeitswissenschaftlichen Experimenten werden für die Datenanalyse die Daten in einem ersten Schritt gesichtet und auf fehlerhafte Datensätze überprüft. Zu diesem Zweck werden schon während der Durchführung des Experiments Fehler in der Datenaufzeichnung notiert. Die Aufzeichnungen dienen bei der Datensichtung als eine Orientierung für den Ausschluss von Datensätzen aus der weiteren Analyse. Ebenfalls können die Daten nach der ersten Analyse betrachtet und Datensätze ausgeschlossen werden, die festgesetzten Anforderungen nicht entsprechen oder die trotz erfolgreicher Aufzeichnung der Daten Fehler aufweisen. Die Analyse der Daten muss folglich zum einen ohne Ausschluss und zum anderen mit Ausschluss von Daten ermöglicht werden.

Aufgrund der großen Datenmenge und der langen Analysezyklen müssen Daten und Analyseparameter persistent abgelegt werden können. Ebenfalls muss der Analyse- und Vergleichsprozess an definierten Punkten abbrechbar und wieder aufnehmbar sein.

In Experimenten werden Versuche, Versuchspersonen und Durchläufe mit Bezeichnungen versehen, um die durchgängige Interpretation

und Verfolgung der zu bewertenden Daten zu ermöglichen. Somit ist zu beachten, dass die vergebenen Bezeichnungen für die zu integrierenden Datensätze durch den Benutzer eingebbar und nach Wunsch anpassbar sind und über den gesamten Prozess hinweg erhalten bleiben.

Sind Analysen und Vergleiche durchgeführt, muss ein Werkzeug das Speichern der ausgewählten Daten und Parameter ermöglichen. Somit können Berechnungen einfach wiederholt und eine neue Berechnung mit veränderten Daten und Parametern auf der Grundlage schon durchgeführter Berechnungen durchgeführt werden.

Der softwaregestützte Vergleich der Daten muss einen Vergleich von bis zu drei Datengruppen ermöglichen. Somit können komplexe Vergleiche durchgeführt werden, ohne dass die Darstellung und damit die individuelle Bewertung der Daten zu unübersichtlich wird. Ebenso wird gewünscht, die Datengruppen einzeln bearbeiten zu können, um beispielsweise einzelnen Versuchspersonen mit einer Gruppe von Versuchspersonen oder einfache Paarvergleiche durchzuführen.

Die Voruntersuchung (siehe Kapitel 3.1) zeigt, dass simulierte Daten eines kognitiven Modells zum Teil keine Varianz aufweisen, somit sollten für den Vergleich und die Analyse von simulierten und empirischen Blickbewegungsdaten deskriptive Verfahren zur Anwendung kommen. In frühen Phasen der Systemgestaltung sind einfache Vergleiche ausreichend für die Bewertung von Mensch-Maschine-Schnittstellen (Urbas, Schulze-Kissing & Leuchter 2005).

- A2 Analysen mit und ohne Ausschluss von Datensätzen
- A3 Persistente Ablage von Daten und Analyse- und Vergleichsparametern
- A4 Vergabe von Datenbezeichnungen
- A5 Abspeichern und Integration von Parametern einer Datenanalyse
- A6 Vergleich von bis zu drei Datengruppen
- A7 Anpassen der Datengruppen für den Vergleich
- A8 Vergleiche auf deskriptiver Ebene

3.2.3 Analyse- und Vergleichsmethoden

Bei der Bewertung von Leistungsdaten unterschiedlichen Ursprungs muss gewährleistet sein, dass die Daten mit gleichen Methoden verarbeitet werden, um die Integrität der Ergebnisse zu ermöglichen. Dementsprechend müssen die Algorithmen des Werkzeuges eine Verarbeitung der Daten implementieren, die eine Unterscheidung der Daten auf Ebene der Bezeichnungen, aber keine in der algorithmischen Verarbeitung ermöglicht.

Neben integrierten Methoden zur Analyse und für den Vergleich der Daten sollte die Möglichkeit bestehen, dass neue Methoden einfach in das System und den Arbeitsprozess zu integrieren sind. Ein modularer Aufbau des Werkzeugs ermöglicht eine Anpassung des angebotenen Repertoires der Datenverarbeitung.

A9 Gleiche Analyse- und Vergleichsmethoden für empirische und simulierte Daten

A10 Modularer Aufbau der Analyse- und Vergleichsmethoden

3.2.4 *Datensichtung*

Für die Datensichtung müssen die Ergebnisse in graphischer und in tabellarischer Form abgelegt werden. Die graphische Repräsentation der Daten ermöglicht eine schnelle Sichtung und das Auffinden von Unregelmäßigkeiten in diesen. Die Ablage der Ergebnisse in Tabellen sollte in einfacher textbasierter Form erfolgen, um eine Integration der Analysen und Vergleiche in externe Programme (z. B. Werkzeuge für die weiterführende statistische Analyse wie SPSS und MatLab) zu ermöglichen. Für eine bessere Übersicht sollten alle Daten des Systems in einer einheitlichen Struktur abgelegt werden, deren Ursprung durch den Benutzer festzulegen ist. Dadurch können die Daten in bestehende Strukturen eingefügt werden, und der Umgang mit den zumeist großen Datenmengen wird vereinfacht.

A11 Bereitstellen von graphischen Ausgaben

A12 Bereitstellen von Datentabellen

A13 Ablegen der Daten in strukturierter Form

3.3 KONZEPT

In den vorangegangenen Kapiteln wurden Anforderungen und theoretische Überlegungen für ein Werkzeug zur Analyse und den Vergleich von Leistungsdaten natürlicher und künstlicher kognitiver Systeme dargestellt. Darauf aufbauend wird nachfolgend das Konzept für ein Werkzeug zur Unterstützung der notwendigen Verarbeitungsprozesse vorgestellt. Das Werkzeug kann in den verschiedenen Phasen des Entwicklungsprozesses von Mensch-Maschine-Systemen eingesetzt werden, um eine formale Bewertung der Gestaltung von Mensch-Maschine-Schnittstellen auf Grundlage kognitiver Benutzermodelle zu unterstützen (siehe Kapitel 2.1.2). Des Weiteren ist durch den Einsatz des Werkzeuges eine Plausibilisierung kognitionspsychologischer Theorien, die in kognitiven Modellierungsmethoden formalisiert sind, im Bezug zu empirischen Untersuchungen und weiteren formalen Methoden möglich.

Das Werkzeug wird mit SimTrA bezeichnet. SimTrA ist ein Akronym für *Simulation Trace Analyzer*. In Abbildung 7 sind die drei Teilsysteme von SimTrA in schematischer Form dargestellt: (1) die Aufbereitung der Rohdaten, (2) die Analyse der aufbereiteten Daten und (3) der Vergleich oder die wiederholte Analyse von simulierten und empirischen Daten. Der modulare Aufbau von SimTrA erlaubt eine unabhängige Weiterentwicklung und Ergänzung der einzelnen Teilsysteme (Anforderung A10, siehe Kapitel 3.2.3). Konzeptionell ist SimTrA für die Analyse und den Vergleich von verschiedenen Leistungsdaten kognitiver Modelle ausgelegt. SimTrA wurde prototypisch für die Analyse und den

Vergleich von Blickbewegungsdaten implementiert (zur Auswahl der Blickbewegungsdaten siehe Kapitel 2.3.2).

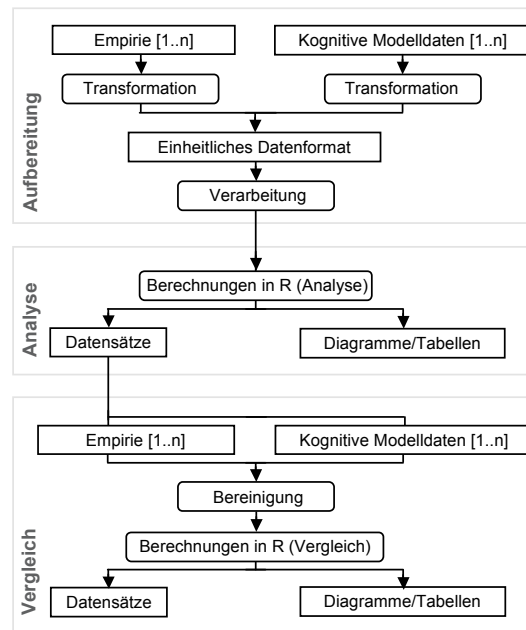


Abbildung 7: Schematische Darstellung der Komponenten von SimTrA.

Das Werkzeug SimTrA ist als abschließender Baustein einer Prozesskette für die kognitive Modellierung im Bereich der Bewertung von Mensch-Maschine-Systemen zu betrachten. Das Ziel der Prozesskette liegt in der Erhöhung der Effizienz der Methode Benutzermodellierung (vgl. Urbas, Schulze-Kissing & Leuchter 2005). Die in der Arbeitsgruppe entwickelte prototypische Prozesskette besteht aus vier Komponenten: (1) der funktionalen Übertragung der Interaktions- und Rückmeldeelemente einer Mensch-Maschine-Schnittstelle in eine von kognitiven Architekturen interpretierbaren Repräsentation mit dem Werkzeug AGI-map (*ACT-R Graphical Interface Mapper*; Urbas & Leuchter 2005), (2) der Entwicklung von kognitiven Modellen mit dem Werkzeug HTAmap (*Hierarchical Task Analysis Mapper*; Heinath & Urbas 2007) auf Grundlage einer Aufgabenanalyse und vordefinierten Interaktionsmustern, (3) der Simulation des Nutzerverhaltens mit der kognitiven Architektur ACT-R (Anderson et al. 2004) und (4) der Analyse und dem Vergleich der erfassten Leistungsdaten mit dem Werkzeug SimTrA.

Die ersten zwei Prozessschritte sind eng miteinander verbunden und stellen die Erstellung eines lauffähigen Benutzermodells in der kognitiven Architektur ACT-R dar. In dem Schritt der Simulation werden die Leistungsdaten für die Analyse und den Vergleich generiert, die an SimTrA übergeben werden.

Das Werkzeug SimTrA ist von der Wahl der kognitiven Architektur unabhängig, da die internen Verarbeitungsstrukturen auf einem von den Rohdaten unabhängigen Format beruhen (siehe Abbildung 7).

Ebenso muss die Modellerstellung nicht notwendigerweise in der vorgeschlagenen Prozesskette zur Benutzermodellierung erfolgen, da die in dem Schritt der Simulation generierten Daten für die Analysen und Vergleiche genutzt werden. Somit kann das Werkzeug SimTrA für die Analyse und den Vergleich von Leistungsdaten kognitiver Modelle beliebiger kognitiver Architekturen, die Blickbewegungsdaten generieren, angewandt werden (Anforderung A1, siehe Kapitel 3.2.1).

3.3.1 Prozess der Datenverarbeitung

Der Prozess der Datenverarbeitung ist in SimTrA in vier einzelne Schritte unterteilt: (1) Aufbereitung, (2) Analyse, (3) Bereinigen und (4) Vergleich. In Abbildung 8 sind die vier Schritte, der Informationsfluss und die jeweils abgespeicherten Daten schematisch dargestellt. Nach jedem Prozessschritt werden die aufbereiteten Daten, die verarbeiteten Daten und die Prozessparameter der Analysen und Vergleiche automatisch in Dateien abgelegt (Anforderung A3, siehe Kapitel 3.2.2). Dies ist in einer Baumstruktur, deren Ursprungsverzeichnis durch den Benutzer festgelegt werden kann, realisiert (Anforderung A13, siehe Kapitel 3.2.4). Die Sicherung der Daten kann auch explizit durch den Benutzer erfolgen (Anforderung A5, siehe Kapitel 3.2.2). Alle Parameter und Ergebnisse der Analysen und Vergleiche werden in graphischer und tabellarischer Form in der Baumstruktur gespeichert und können für weiterführende Bewertungen genutzt werden (Anforderung A4, siehe Kapitel 3.2.2; Anforderungen A11 und A12, siehe Kapitel 3.2.4).

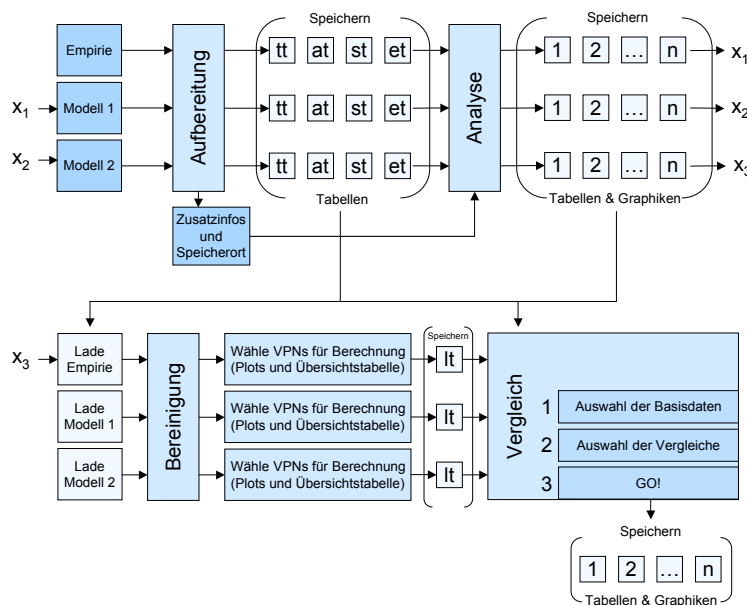


Abbildung 8: Die Verarbeitungsschritte von SimTrA: Dargestellt ist der Informationsfluss und die jeweils abgelegten Dateien bei der Verarbeitung von drei Datensätzen. at: Tabelle der AOIs, et: Tabelle der Elemente der kognitiven Architektur, lt: Tabelle der ausgewählten Versuchspersonen, st: Tabelle der ausgewählten Zeitintervalle, tt: Tabelle der Blickbewegungsdaten, VPN: Versuchspersonen.

3.3.2 Aufbereitung der Rohdaten

Die Aufbereitung der Rohdaten der Blickbewegungsdatenerfassung stellt den ersten Prozessschritt in SimTrA dar. Die simulierten Blickbewegungsdaten werden in SimTrA integriert und in einem nachfolgenden Schritt in Fixationsdaten transformiert (X- und Y-Koordinaten und Anfangs- und Endzeit der Fixation). Zu diesem Zweck werden Blickzuwendungen auf gleiche X- und Y-Koordinaten zusammengefasst. Dabei wird überprüft, ob bei gleichen Koordinaten eine gerichtete Aufmerksamkeitszuwendung zwischen den Blickzuwendungen durch das kognitive Modell initiiert wurde (in ACT-R: *move-attention*). Somit können nachfolgende Fixationen auf einen gleichen Blickort unterschieden werden, die in kognitiven Modellen nicht durch eine Ablenkung des Auges unterscheidbar wären. Die empirisch erfassten Rohdaten werden in Form von Fixationsdaten in SimTrA integriert. Diese bestehen mindestens aus den X- und Y-Koordinaten und der Anfangs- und der Endzeit der Fixationen. Für die Berechnung der Fixationsdaten werden in der Literatur eine Vielzahl von Algorithmen genannt (siehe Kapitel 2.4.3). Aus diesem Grund wird die Berechnung der Fixationsdaten nicht in SimTrA implementiert, sondern es wird auf vorhandene Algorithmen zurückgegriffen.

Die zu integrierenden Daten können sich in Datenformat und Auflösungen unterscheiden (z. B. aufgrund von unterschiedlichen Systemen zur Blickbewegungsdatenerfassung und der Blickbewegungsdaten kognitiver Architekturen). In SimTrA werden die jeweiligen Daten eines Blickprotokolls durch spezielle Importalgorithmen in ein standardisiertes Format transformiert (Anforderung A1, siehe Kapitel 3.2.1). Durch den Import werden die Daten automatisch um Parameter erweitert, die aus den Eingangsdaten ausgelesen oder anhand von Benutzereingaben eingefügt werden. Dazu zählen der Datenherkunftsbezeichner (dID, data-ID), die Versuchsdurchführungsnummer (tID, trail-ID) und die Versuchspersonennummer (sID, subject-ID; Anforderung A4, siehe Kapitel 3.2.2). Die Bezeichnungen bleiben über den gesamten Verarbeitungsprozess der Eingangsdaten konstant und dienen der internen Referenzierung und der Zuordnung durch den Benutzer von SimTrA.

Für die Verarbeitung von Daten mit unterschiedlicher Auflösung können in SimTrA die importierten Daten in X- und Y-Richtung skaliert und um einen Offset auf der X- und der Y-Achse verschoben werden. Somit können Blickbewegungsdaten aus kognitiven Simulationsexperimenten und empirischen Experimenten in ihrer Auflösung angeglichen und in nachfolgenden Prozessschritten mit gleichen Rahmenparametern analysiert und verglichen werden (Anforderung A9, siehe Kapitel 3.2.3).

In diesem Prozessschritt werden ebenfalls Möglichkeiten angeboten, um die für die Analysen benötigten Parameter einzugeben. Der Benutzer kann das Blickgebiet festlegen und dieses in regelmäßige oder in unregelmäßige Gebiete (AOI) aufteilen. Zusätzlich können Zeitintervalle definiert werden, die für die Interpretation der Blickbewegungen genutzt werden, um spezielle Aspekte oder Verläufe analysieren zu können.

3.3.3 Analyse der Daten

In dieser Komponente werden die Blickbewegungsdaten analysiert. Die Auswahl der Algorithmen leitet sich aus der dritten Ebene der applikativen Parametersystematik von Rötting (2001) ab. Diese ermöglicht die Analyse der Bedingungen der Mensch-Maschine-Interaktion (siehe Kapitel 2.4.2). Analysemethoden, die sich auf Sakkaden beziehen, wurden ausgeschlossen, da heutige kognitive Architekturen keine Sakkaden simulieren. Die Analysemethoden der Ebene der Informationsverarbeitung wurden ebenfalls ausgeschlossen, da diese nur „tendenzielle“ Rückschlüsse auf die Mensch-Maschine-Interaktion zulassen (Rötting 2001, S. 186). Aus den verbleibenden drei Kategorien wurden Analysemethoden ausgewählt, die eine Betrachtung der Blickbewegungen im Rahmen der Mensch-Maschine-Interaktion erlauben.

Die ausgewählten Methoden stellen einen Ausschnitt aus allen möglichen Analysemethoden dar. Der modulare Aufbau von SimTrA erlaubt die Integration von beliebigen Algorithmen in den bereitgestellten Prozess.

Analyse der örtlich-räumlichen Gestaltung

Die Analyse von Blickbewegungen stellt bei der Evaluation der örtlich-räumlichen Gestaltung einen wichtigen Aspekt für die Mensch-Maschine-Interaktion dar. In diesem Kontext wurden die Analysemethoden für die Fixationsdichte, die räumliche Dichte und die relative Häufigkeit der Fixation von Objekten oder Regionen, die kumulierte Fixationsdauer und die Übergangshäufigkeiten realisiert.

FIXATIONSDICHTE Für die Bestimmung der Fixationsdichte wird das Blickgebiet in regelmäßige Felder unterteilt und die Anzahl oder die Dauer der Fixationen in diesem Feld bestimmt. Die Angabe der Werte kann absolut oder relativ erfolgen. Hohe Werte für die Fixationsdichte weisen auf besonders informationstragende Bereiche (Felder) hin, die bei der Gestaltung von Mensch-Maschine-Schnittstellen betrachtet werden sollten (Mackworth & Morandi 1967).

RÄUMLICHE DICHTE Für die Analyse der räumlichen Dichte wird das Blickgebiet in ein regelmäßiges Raster unterteilt. Anschließend wird die Anzahl der Felder bestimmt, in die mindestens eine Fixation gefallen ist (Isfort 1986). Werden die Werte prozentual ausgegeben, weisen kleine Werte der räumlichen Dichte auf eine gerichtete globale visuelle Suche hin (Goldberg & Kotval 1999). Dabei ist zu beachten, dass Vergleiche von Schnittstellen nur mit gleicher Aufteilung des Blickgebietes durchgeführt werden können und die Werte der räumlichen Dichte bei dem Vergleich in relativer Weise zu betrachten sind.

RELATIVE HÄUFIGKEIT DER FIXATION VON OBJEKTEN Für die Bestimmung der relativen Häufigkeit der Fixation von Objekten oder Regionen wird das Blickfeld in AOIs aufgeteilt und die Anzahl der in die AOIs fallenden Fixationen bestimmt. Ein hoher Wert deutet auf ein Gebiet hin, das eine hohe visuelle Aufmerksamkeit erfordert. Bei der Evaluation von Mensch-Maschine-Systemen können im Vergleich

wichtigere bzw. informationshaltigere Bereiche einer Schnittstelle identifiziert werden (Göbel 1999). Darauf basierend, kann die Gestaltung an die Ergebnisse einer Analyse angepasst werden.

KUMULIERTE FIXATIONSDAUER Für die Analyse der kumulierten Fixationsdauern wird das Blickfeld in Felder aufgeteilt, und alle Fixationen innerhalb eines Feldes werden summiert. Ein hoher Wert weist auf einen Ort mit hoher visueller Aufmerksamkeit hin (Rötting 2001). Dieser Wert kann für den Vergleich von unterschiedlichen Aufgabenbedingungen genutzt werden und erlaubt eine schnelle Identifikation der Unterschiede und Gemeinsamkeiten der Blickbewegungen.

ÜBERGANGSHÄUFIGKEITEN Zur Bestimmung der Übergangshäufigkeiten wird das Blickgebiet in regelmäßige oder selbstgewählte Gebiete aufgeteilt und die Anzahl der Übergänge zwischen den definierten Bereichen aufsummiert (vgl. Ellis & Smith 1985). Der prozentuale Wert der Übergänge zwischen zwei Gebieten erlaubt eine Beurteilung der Güte der räumlichen Gestaltung von Schnittstellen. Viele verschiedene Übergänge zwischen Gebieten einer Schnittstelle deuten auf einen ineffizienten Blickpfad und somit auf eine nicht adäquate Gestaltung der Schnittstelle hin (Goldberg & Kotval 1998). Aus den Übergangshäufigkeiten kann Top-down-Verhalten, das heißt gelerntes Verhalten, abgeleitet werden (Rötting 2001).

Analyse von Suchprozessen

Das Auffinden von Informationen innerhalb einer Schnittstelle stellt einen wichtigen Aspekt bei der Interaktion von Mensch und Maschine dar. Um die visuelle Suche analysieren zu können, wurden Analysealgorithmen für die Bestimmung der Anzahl der Fixationen und der Übergangsdichte integriert.

ANZAHL DER FIXATIONEN Die Anzahl der Fixationen wird innerhalb eines definierten Zeitintervalls oder bis zur Lösung einer Aufgabenstellung bestimmt (Boersema, Zwaga & Adams 1989). Der Wert kann Hinweise zur Beschreibung von Suchprozessen liefern. Ein kleiner Wert von Fixationen bis zur Zielerreichung deutet auf ein salientes Ziel, einen gut trainierten Benutzer oder eine einfache Aufgabe hin. Ein hoher Wert kann ein Hinweis auf eine Behinderung der visuellen Suche darstellen (Goldberg & Kotval 1998).

ÜBERGANGSDICHTE Für die Bestimmung der Übergangsdichte wird auf die Übergangshäufigkeiten zurückgegriffen (siehe Kapitel 3.3.3). Es wird die Anzahl der Übergänge zwischen zwei Blickgebieten größer als 1 bestimmt und in Relation zu allen möglichen Übergängen gesetzt. Ein hoher Wert kann ein Indikator für einen langen verteilten Blickpfad sein. Kleine Werte weisen auf eine direkte visuelle Suche hin (Goldberg & Kotval 1998, 1999).

Analyse von Informationsaufnahme- und -verarbeitungsprozessen

Die Analyse von Informationsaufnahme- und -verarbeitungsprozessen ermöglicht die Untersuchung der Art der Informationsverarbeitung und die Auswirkungen der aufgenommenen Informationen auf die Informationsverarbeitungsprozesse und somit auf das Blickverhalten des Benutzers. In diesem Zusammenhang wurden Analysemethoden für die statistische Analyse der Übergangshäufigkeiten und der lokalen Fixationspfade in SimTrA integriert.

STATISTISCHE ANALYSE DER ÜBERGANGSHÄUFIGKEITEN Bei der statistischen Analyse der Übergangshäufigkeiten wird untersucht, ob das erfasste Blickverhalten zufällig ist oder eine statistische Abhängigkeit besteht. Zufällig bedeutet in diesem Zusammenhang, dass Blickgebiete mit gleicher Wahrscheinlichkeit betrachtet werden und die Übergänge zwischen den Blickgebieten zufällig sind. Weichen die Übergangshäufigkeiten von den erwarteten Häufigkeiten ab, kann von einem statistisch abhängigen Blickverhalten gesprochen werden (Ellis & Smith 1985; Ellis & Stark 1986). Statistisch abhängiges Blickverhalten beruht auf aufgenommenen Informationen während einer Fixation, die den Ort der nächsten Fixation bestimmt (*Close-loop-control*; Ellis & Smith 1985) oder auf von der aufgenommenen Information unabhängigen Informationsverarbeitungsprozessen (*Open-loop-control*; Ellis & Smith 1985). Zusammenfassend wird aus den statistischen Abhängigkeiten der Blickbewegungen die bedingte Information (H_c) einer Übergangstabelle berechnet (Ellis & Stark 1986). Der H_c -Wert erlaubt eine Aussage über die Zufälligkeit der erfassten Blickbewegung. Der H_c -Wert hat ein Maximum, wenn die Blickbewegung gleichförmig verteilt ist, und ein Minimum, wenn die Blickzuwendungen nur auf einer AOI liegen. Dies bedeutet, dass ein kleiner H_c -Wert auf eine höhere statistische Abhängigkeit der Blickbewegungen hinweist.

LOKALE FIXATIONSPFADE Für die Analyse der lokalen Fixationspfade wird das Blickgebiet in selbstgewählte Gebiete eingeteilt und alle theoretischen Triple von AOIs ermittelt. Die Häufigkeit der gebildeten Triple wird in der Sequenz der Fixationen der erfassten Blickbewegungsdaten ermittelt. Dabei werden aufeinander folgende Fixationen in eine AOI als eine Fixation betrachtet (Groner et al. 1984). Lokale Fixationspfade zeigen das aktuelle perzeptuelle Geschehen (*Bottom-up-Prozesse*). Geordnet nach Häufigkeit, geben sie Hinweise auf die wichtigsten visuellen Fixationspfade. Die Analyse von lokalen Fixationspfaden helfen Informationsaufnahmeprozesse zu evaluieren.

3.3.4 *Bereinigung der Daten*

Die Bereinigung der Daten erlaubt, die Datenbasis für nachfolgende Vergleiche und erneute Analysen anzupassen. In dieser Komponente können fehlerhafte Datensätze und Datensätze, die festgelegten Kriterien nicht entsprechen, ausgeschlossen werden (Anforderung A2, siehe Kapitel 3.2.2).

Dem Benutzer wird für die jeweiligen Versuchsdurchläufe, Versuchspersonen und Zeitintervalle eine Darstellung der Fixationen und der

festgelegten AOIs angezeigt. Zusätzlich kann eine Tabelle betrachtet werden, die Informationen der Blickbewegungen zusammenfasst. In dieser Tabelle werden das obere und das untere Perzentil einer frei wählbaren Tabellenspalte angezeigt. Anhand dieser Informationen können Datensätze deselektiert werden, um diese aus der weiteren Verarbeitung auszuschließen.

3.3.5 *Vergleich der Daten*

Der Vergleich oder die erneute Analyse von Daten schließt den Prozess von SimTrA ab. Dafür können die analysierten Daten von bis zu drei Gruppen (z. B. Modell 1, Modell 2, Mensch) in diese Komponente geladen werden (Anforderung A6, siehe Kapitel 3.2.2). Jede Gruppe kann hinsichtlich der Auswahl von Versuchsdurchläufen, Versuchspersonen und Zeitintervallen angepasst werden (Anforderung A7, siehe Kapitel 3.2.2). Abschließend werden die Vergleichs- oder Analysemethoden ausgewählt. Diese beruhen auf den in der Analysekomponente integrierten Methoden (siehe Kapitel 3.3.3). Der Vergleich von Gruppen wird mit deskriptiven Methoden durchgeführt (Anforderung A8, siehe Kapitel 3.2.2).

3.4 REALISIERUNG DES WERKZEUGS SIMTRA

Basierend auf dem vorangegangenen vorgestellten Konzept wurde ein Komponentenmodell entwickelt, das als Grundlage für die Realisierung von SimTrA dient (siehe Abbildung 9). Dabei wird zwischen Komponenten für die Benutzerinteraktion und die Bearbeitung der Analyse- und Vergleichsalgorithmen unterschieden. Die Schnittstellen für die Benutzerinteraktion sind in Java (Sun Microsystems 2007) und die algorithmischen Berechnungen in R, einer Software für die statistische Analyse (R Development Core Team 2006), realisiert. Voraussetzung für die Nutzung von SimTrA sind eine lauffähige Java-Umgebung und eine Installation der Software R. Beide Softwaresysteme sind frei im Internet zu beziehen.

Die jeweils benötigten Daten werden in einer definierten Baumstruktur abgelegt (siehe Abbildung 9), auf die alle Komponenten zugreifen, um benötigte Daten auszulesen. Die einzelnen Komponenten sind modular angelegt und einfach um neue Aspekte und Algorithmen zu erweitern.

Für die Implementierung von SimTrA wurde eine benutzerzentrierte Entwicklung von Prototypen gewählt, die iterativ um neue Aspekte erweitert wurden. Nach jedem Iterationsschritt wurde die Funktionalität und die Handhabung der Prototypen überprüft, um Folgefehler ausschließen zu können. Im Nachfolgenden werden einzelne Aspekte jeder Komponente dargelegt und die jeweiligen Benutzungsschnittstellen beschrieben.

3.4.1 *Allgemeine Angaben*

Das Zusammenspiel der einzelnen Komponenten wird durch die Bereitstellung von allgemeinen Informationen und Parametern ermöglicht.

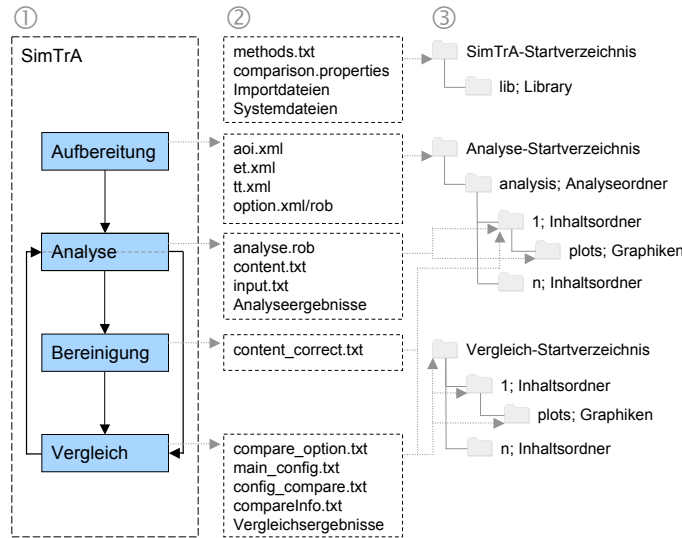


Abbildung 9: 1: Komponentenmodell von SimTrA. 2: Die wichtigsten Dateien, die angelegt und abgerufen werden. 3: Verzeichnisstruktur von SimTrA.

Dazu zählen beispielsweise die Steuerungsinformationen für die Algorithmen, die Standardparameter und die Ablagehierarchie für Dateien. In dem Startverzeichnis von SimTrA sind verschiedene Dateien abgelegt, die Parameter und Standardinformationen für die Vergleiche und die erneute Analyse von Blickbewegungsdaten vorhalten (z. B. comparison.properties; siehe Abbildung 9). Der Benutzer kann in diesen Dateien Angaben ändern und hinzufügen, um eigene Standardeinstellungen zu setzen oder um SimTrA zu erweitern.

Für neue Analysen oder Vergleiche kann das jeweilige Standardverzeichnis von SimTrA durch ein selbstgewähltes Verzeichnis des Benutzers ersetzt werden. In dem Standardverzeichnis werden alle Parameter und Dateien für die Analyse oder den Vergleich abgelegt. Die Ergebnisse der Analyse oder der Vergleiche werden zusammen mit den relevanten Parametern in einem Unterverzeichnis (Namensbereich 1..n) abgelegt. Für graphische Ausgaben wird in dem Verzeichnis ein Ordner mit der Bezeichnung *Plots* bereitgestellt (siehe Abbildung 9). Die einzelnen Analysen und Vergleiche sind voneinander unabhängig, da alle relevanten Parameter und Dateien in dem jeweiligen Analyse- oder Vergleichsverzeichnis abgelegt und zugreifbar sind. Alle Pfadangaben in SimTrA werden von dem jeweiligen Standardverzeichnis gesetzt.

3.4.2 Aufbereitung der Rohdaten

Für die Aufbereitung der Rohdaten und die Eingabe von zusätzlichen Informationen wurde die Aufbereitungsschnittstelle erstellt (siehe Abbildung 10). Die Schnittstelle besteht aus zwei Bereichen. Der erste Bereich beinhaltet die Eingabe und Manipulationsmöglichkeiten für den Benutzer, der zweite Bereich stellt die Daten und Informationen dar.



Abbildung 10: Benutzungsoberfläche von SimTrA für die Datenaufbereitung.

Datenintegration

Im ersten Schritt kann der Benutzer das Standardverzeichnis für die Datenanalyse auswählen (Schaltfläche *Set Project Path*). Das Standardverzeichnis stellt das Verzeichnis dar, von dem aus für die durchzuführende Analyse die Pfade für das Ablegen aller weiteren Dateien relativ gesetzt werden und das die Grundinformationen für die Analyse enthält. Befinden sich in dem gewählten Verzeichnis durch SimTrA lesbare Dateien, werden diese eingelesen und dargestellt.

Für die Datenintegration (siehe Abbildung 10) kann über die Schaltfläche *Import Data* eine Datei ausgewählt werden, die in SimTrA importiert werden soll. Anhand der Dateierweiterung wird der jeweilige Importer ausgewählt, die Datei in das standardisierte Datenformat von SimTrA überführt und in dem gewählten Standardverzeichnis abgespeichert. Die Importer sind in Java und jeweils für ein Eingangsformat realisiert, das durch die Dateierweiterung spezifiziert wird. Neue Importer können hinzugefügt werden, indem eine neue Java-Klasse (d. h. ein Importer) und die gewünschte Dateierweiterung für den Import integriert wird (z. B. *.soar für Blickbewegungen aus Soar). In SimTrA werden Importer für die Blickbewegungen aus ACT-R und für den Import des Datenformats der BeGaze-Software bereitgestellt.

Nach erfolgreicher Integration der Blickbewegungsdaten wird der Wertebereich für sID (Versuchspersonennummer) und tID (Versuchsdurchlaufnummer) sowie die Anzeige der Fixationen aktualisiert. Durch direkte Eingabe einer Nummer neben der Anzeige des Wertebereichs kann ein Datensatz direkt ausgewählt werden, der in der Datenanzeige dargestellt wird.

Durch die Schaltfläche *Save Options* können alle aktuellen Einstellungen gesichert werden. Die Schaltfläche *Load Options* erlaubt die Integration von Informationen aus vorangegangenen Analysen (z. B. AOIs, Skalierungsfaktoren). Um die Informationen in das aktuelle Pro-

jekt zu integrieren, muss das Ursprungsverzeichnis der zu ladenden Informationen gewählt werden. Die jeweiligen Daten werden in das aktuelle Standardverzeichnis übernommen. Durch den Befehl `Delete All` werden alle getätigten Angaben gelöscht und aus dem Verzeichnis entfernt.

Durch Betätigen der Schaltfläche `Start Analyses` wird der Dialog zur Auswahl der Analysemethoden geöffnet, der den Abschluss der Datenaufbereitung darstellt (siehe Kapitel 3.4.3).

Datenmanipulation

Die Datenmanipulation erlaubt die Skalierung der eingelesenen Fixationen, um Datensätze aus unterschiedlichen Experimenten (z. B. Simulationsexperiment, empirisches Experiment) auf eine vergleichbare Auflösung zu skalieren. Die Datenmanipulation wird für alle eingelesenen Fixationen der aktuellen Aufbereitung durchgeführt. Die Fixationen können in X- und Y-Richtung skaliert und verschoben werden. Die Eingabe der Skalierungsfaktoren und der Verschiebungsfaktoren wird durch den Benutzer vorgenommen.

Eingabe von zusätzlichen Angaben

Mit Hilfe dieser Eingabemöglichkeiten können weitere Angaben für die Analyse gemacht werden. In dem Dialogfeld `Define Areas Of Interest (AOIs)` kann der Benutzer Blickgebiete erstellen, die für die Analyse von Relevanz sind. Dafür kann der Benutzer die Blickgebiete direkt definieren, indem die obere linke Ecke mit X- und Y-Position, die Höhe und die Breite eingegeben werden. Durch zweimaliges Klicken innerhalb des Anzeigebereichs (für obere linke und untere rechte Ecke des Blickgebiets) kann eine AOIs ebenfalls direkt in der Anzeige der Fixationen gesetzt werden. Die gesetzten AOIs sind später veränderbar und können gelöscht werden (Schaltfläche `Del`). Um eine AOI anzupassen, wird in dem Feld `Actual AOI` die gewünschte Nummer eingegeben. Anschließend können die vorher beschriebenen Werte der AOIs verändert werden.

Ein Raster kann automatisch über das Blickfeld gelegt werden. Dazu kann der Benutzer die Anzahl der Spalten und der Reihen angeben. Wird ein Raster über das Blickfeld gelegt, werden alle AOIs gelöscht und das Raster angezeigt. Die einzelnen Felder des Rasters können nicht verschoben oder verändert werden. Eine Änderung des Rasters ist durch erneute Definition der Spalten- oder Reihenanzahl möglich. Wird eine AOI gesetzt, wird das vorhandene Raster gelöscht.

Da für die Analysen der Blickbewegungen mindestens eine AOI oder ein Raster vorhanden sein muss, werden die Analysealgorithmen nur ausgeführt, wenn eine der zwei Bedingungen erfüllt ist.

In dem Dialogbereich `Define Screen (Rectangle)` kann der Benutzer das Blickfeld numerisch eingeben und verändern. Das Blickfeld wird in der Anzeige mit einem schwarzen Rahmen dargestellt. Zusätzlich kann ein Hintergrundbild ausgewählt beziehungsweise gelöscht werden, das hinter den Fixationen in der Datenanzeige eingefügt wird. Dadurch wird die Definition von AOIs vereinfacht und es kann ein Überblick der

erfassten Fixationen im Bezug zu der zu untersuchenden Schnittstelle gewonnen werden.

Der Dialogbereich *Define Times of Interest* ermöglicht die Definition von Zeitintervallen, die bei der Analyse der Blickbewegungsdaten gesondert berechnet werden sollen. Die Zeitintervalle werden mit einem Start- und einem Endzeitpunkt in Millisekunden angegeben. Erstellte Zeitintervalle können gelöscht werden.

Datenanzeige

In der Datenanzeige werden die Fixationen eines Datensatzes, die erstellten AOIs beziehungsweise das Raster und das Blickfeld angezeigt. Unterhalb des Blickfeldes werden die Zeitintervalle und die Datenbezeichner dargestellt. Die Farben der angezeigten Informationen korrespondieren mit der Darstellung innerhalb der Anzeige. Alle Änderungen der Daten und Parameter durch den Anwender werden übernommen und innerhalb des Anzeigebereichs sofort aktualisiert.

Bei Blickbewegungsdaten aus der kognitiven Architektur ACT-R werden zusätzlich die Elemente der Versuchsumgebung dargestellt. Die Anzeige der Elemente basiert auf dem Blickprotokoll des kognitiven Modells, in dem die Umgebungsinformationen abgespeichert sind.

Automatische Ablage von Parameterdateien

Nach der Aufbereitung liegen verschiedene Daten vor, die in dem gewählten Standardverzeichnis abgelegt und für die Analyse genutzt werden. Wenn von dem Benutzer gewünscht, können die Dateien in einem Text-Editor geändert und angepasst werden. Bei erneutem Start von SimTrA werden die veränderten Daten übernommen, auf Konsistenz geprüft und verarbeitet. Sind die Daten nicht konsistent mit der Spezifikation von SimTrA, werden die jeweiligen Daten nicht angezeigt und müssen erneut eingegeben oder in der Datei angepasst werden.

Im Nachfolgenden ist eine Liste aller abgelegten Dateien für die Parameter von SimTrA in dem Schritt der Datenaufbereitung dargestellt.

TABELLE DER FIXATIONEN UND ZEITSTEMPEL – TT.XML Die Datei enthält die Fixationen einer Versuchsgruppe, die in SimTrA angezeigt werden und aus den Rohdaten gewonnen wurden. Jede Fixation besteht aus den X- und Y-Koordinaten, dem Zeitpunkt, der Versuchspersonennummer und der Versuchsdurchlaufnummer.

TABELLE DER ZEITINTERVALLE – ST.XML Die Datei enthält die festgelegten Zeitintervalle, die bei der Analyse gesondert analysiert werden. Jedes Zeitintervall besteht aus einer Nummer, der Anfangs- und der Endzeit.

TABELLE DER BLICKGEBIETE – AOI.XML Die Datei enthält die gesetzten AOIs oder die Informationen eines Rasters für die Analyse. Jede AOI hat eine Nummer und wird durch die X- und Y-Koordinaten der oberen linken und der unteren rechten Ecke definiert.

TABELLE DER ANZEIGEELEMENTE – ET.XML Stammen die Blickbewegungsdaten aus der kognitiven Architektur ACT-R, enthält diese Datei die angezeigten Elemente der virtuellen Versuchsumgebung von ACT-R. Abgelegt sind die Elemente mit X- und Y- Koordinaten, Zeitstempel, den Attributen des Elements (z. B. Farbe, Zustand) und deren Belegung.

TABELLE DER OPTIONEN – OPTION.XML In der Datei option.xml sind die Benutzereingaben und -manipulationen abgelegt. Dazu zählen die Skalierungsfaktoren und die Verschiebung auf der X- und Y-Achse, die Spalten- und Zeilenanzahl eines gewählten Rasters und die Definition des Blickgebiets.

3.4.3 Analyse der Daten

In dem Dialog zur Auswahl der Analysemethoden kann der Benutzer die Analysemethoden bestimmen und einzelne Standardwerte überschreiben (siehe Abbildung 11). Ist keine R-Version ausgewählt, muss vor dem Starten der Analyse eine R-Version gewählt werden, indem in dem Dialogfenster der bin-Ordner von der zu verwendenden R-Version ausgewählt wird.

Sollen im Schritt des Vergleichs Datengruppen verglichen werden, deren Blickbewegungsdaten mit der gleichen Methode erhoben wurden (d. h. die Datenherkunftsbezeichner in den Datentabellen sind gleich), kann der Datenherkunftsbezeichner (dID: data-ID) in diesem Fenster geändert werden.

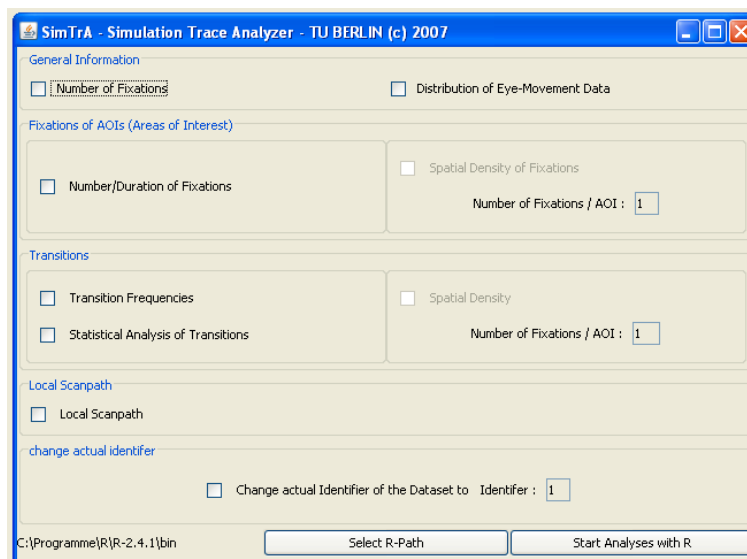


Abbildung 11: Benutzungsoberfläche von SimTra für die Auswahl von Analysemethoden.

Durch den Start der Analyse wird der aktuelle Pfad des Analyseverzeichnisses und die aktuellen Parameter der Analyse in der Datei

analyse.rob abgelegt (siehe Abbildung 9). Anschließend wird die aktuelle Pfadangabe an R übergeben und die Analyse gestartet. Dabei wird von R auf die Dateien in dem Analyseverzeichnis zugegriffen. Eine Analyse besteht aus 3 Schritten: (1) Erstellung einer Grundtabelle, (2) Anwendung der Analysemethoden und (3) Speichern der Ergebnisse.

Erstellung der Grundtabelle

In dem ersten Schritt wird aus den aufbereiteten Daten eine Grundtabelle erstellt (input.txt), die als Basis für die Analysemethoden dient und alle Informationen integriert. Die Tabelle enthält für jede Fixation folgende Werte:

FIXATION X- und Y-Koordinaten und Dauer der Fixation.

ZEITANGABEN Start- und Endzeit der Fixation. Die Angabe erfolgt als absolute Zeitangabe aus den Log-Dateien und als relative Zeit, gemessen vom Startpunkt des Versuchsdurchlaufs bis zum aktuellen Zeitpunkt.

DATENBEZEICHNUNGEN Alle Datenbezeichner (Datenherkunft, Versuchsdurchlaufnummer, Versuchspersonennummer).

BLICKGEBIET Angabe des Blickgebiets und dessen Koordinaten, in das die aktuelle Fixation fällt.

ZEITINTERVALL Angaben zu dem Zeitintervall, in das die Fixation fällt.

ELEMENTE Angabe aller Elemente in dem aktuellen Blickgebiet und die Angabe des am wahrscheinlichsten fixierten Elements (nur für kognitive Modelle in ACT-R).

KOGNITIVE ANGABEN Angabe der Wechselkosten zwischen zwei Fixationen auf unterschiedliche Orte.

Anwendung der Analysemethoden

Für die Analyse der Blickbewegungen wird die Grundtabelle iterativ in einzelne Tabellen zerlegt (je eine Tabelle für eine Gruppe, bestehend aus Datenherkunftsbezeichner, Versuchsdurchlauf, Versuchspersonennummer und Zeitintervall) und anhand der aus der Literatur extrahierten Analysemethoden (siehe Kapitel 3.3.3) algorithmisch verarbeitet. Dabei wird für jede Teilgruppe eine graphische Ausgabe der Fixationen und Blickgebiete angelegt. Die Ergebnisse der Analyse jeder Teilgruppe werden an Ergebnistabellen für jede Analysemethode angehängt.

Für die einzelnen Analysen werden soweit möglich alle Angaben in absoluter und relativer Form abgespeichert. Jede Tabelle enthält die Datenbezeichner der Daten, um die durchgängige Identifikation und Zuordnung der Daten zu ermöglichen und zusätzlich die wichtigsten Parameter der Analysen. Alle Werte, die für die jeweilige Analyse nicht vorhanden sind oder nicht eingegeben wurden (z. B. nicht gesetzte Zeitintervalle, keine AOIs), werden in den Ergebnistabellen mit NA (*not available*) angegeben.

Nach dem Zusammenfassen der einzelnen Tabellen wird für jede Ausgabetabelle eine Tabelle erstellt, die Mittelwerte und Standardfehler der zusammengefassten Werte bereitstellt.

Speichern der Ergebnisse

Nach Beendigung der Analysen werden alle Informationen in dem Analyseverzeichnis gespeichert (siehe Abbildung 9). Dabei werden für jede Analysemethode die Ergebnistabellen, die Durchschnittstabellen, die Grundtabelle und die Hilfstabellen abgelegt. Zusätzlich werden graphische Ausgaben der Ergebnistabellen erstellt. Die Graphiken werden zum einen als .png für den schnellen Überblick und zum anderen als hochauflösendes .ps abgelegt. In der nachfolgenden Auflistung sind die abgelegten Dateien dargestellt. Für die Übersicht der Analysemethoden siehe Kapitel 3.3.3.

INPUT.TXT In dieser Datei sind die aufbereiteten Blickbewegungsdaten für die Analyse der Daten abgespeichert (Grundtabelle).

SHRINK.TXT Diese Datei fasst die Fixationen der Grundtabelle, die hintereinander in eine AOI gefallen sind, zu einer Fixation zusammen. Diese Tabelle wird als Grundlage für die Analyse der lokalen Fixationspfade, der Übergangshäufigkeiten und den damit zusammenhängenden Algorithmen genutzt.

GENERAL.TXT UND **GENERALMEAN.TXT** In dieser Tabelle sind allgemeine Informationen zusammengefasst, wie die Anzahl der Fixationen in einem Versuch, die Versuchsdauer und Durchschnittswerte für die Fixationsdauer.

AOI.TXT UND **AOIMEAN.TXT** Die Dateien stellen Informationen über die Fixation von Blickgebieten bereit. Aus den Informationen können Aussagen bezüglich der relativen Häufigkeit der Fixation von Objekten und der kumulierten Fixationsdauer abgeleitet werden.

AOIS.TXT UND **AOISMEAN.TXT** In diesen Tabellen sind die Informationen der räumlichen Dichte und der Fixationsdichte zusammengefasst.

TRANS.TXT UND **TRANSMEAN.TXT** Diese Tabellen ermöglichen Aussagen zu den Übergangshäufigkeiten zwischen zwei Blickgebieten.

TRANSS.TXT UND **TRANSSMEAN.TXT** Diese Dateien stellen Informationen zu der Übergangsdichte bereit.

STATISTICAL.TXT UND **STATISTICALMEAN.TXT** Diese Tabellen enthalten Informationen zu der statistischen Analyse der Übergangshäufigkeiten.

SCANP.TXT UND **SCANPMEAN.TXT** In diesen Dateien können Informationen zu den lokalen Fixationspfaden abgerufen werden.

CONTENT.TXT Diese Datei beinhaltet alle Datenbezeichner für die analysierten Datengruppen und wird für die Bereinigung und den Vergleich der Daten benötigt.

3.4.4 Bereinigung der Daten

Die Bereinigung der Daten wird in der Benutzungsschnittstelle zur Datenbereinigung durchgeführt (siehe Abbildung 12). Aus der Prozesssicht der Datenverarbeitung findet vor einer erneuten Analyse oder einem Vergleich von Daten die Bereinigung der Daten statt. Sollen bereits bereinigte Daten erneut analysiert oder verglichen werden, kann mit der Schaltfläche *Just Start Comparison* die Bereinigung übersprungen und der nächste Prozessschritt gestartet werden.

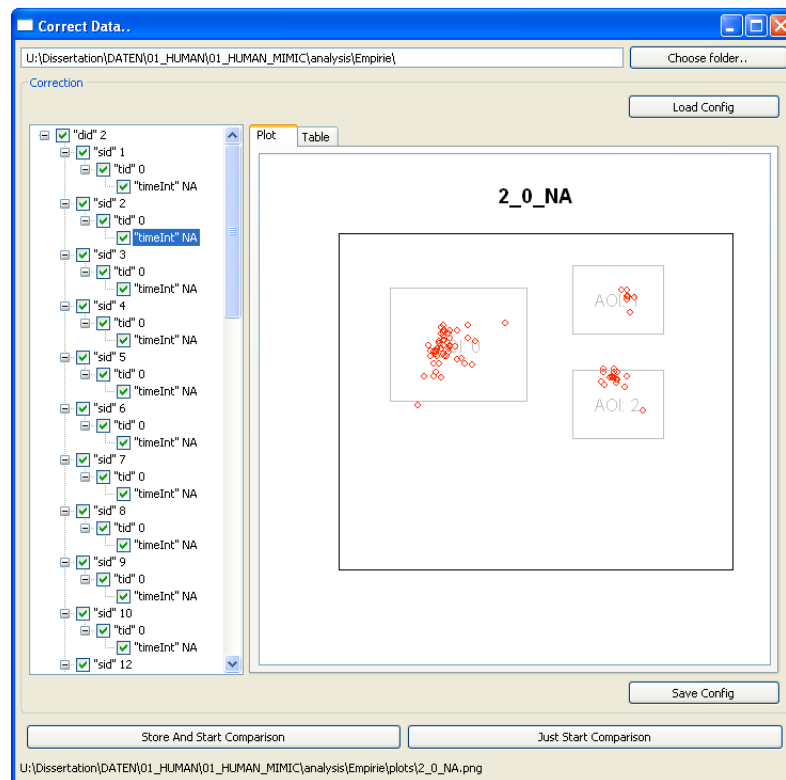


Abbildung 12: Benutzungsoberfläche von SimTrA für die Bereinigung von Datensätzen.

In dieser Komponente können jeweils einzelne Datengruppen sukzessiv betrachtet und bereinigt werden. Mit der Schaltfläche *Choose Folder...* wird eine Datengruppe für die Bereinigung ausgewählt. In dem bereitgestellten Dialogfenster zum Import der Datengruppe muss dafür der gewünschte Analyseordner aus dem Analyseverzeichnis gewählt werden. Basierend auf der Inhaltsdatei in dem Ordner, werden die Datenbezeichner in einer Baumstruktur dargestellt (content.txt; angelegt von der Analysekomponente). Für jeden Datensatz wird eine graphische Übersicht der Fixationen und AOIs angeboten (Auswahl über den Reiter *Plot*), die auf der vorangegangenen Analyse beruht. Zusätzlich werden die allgemeinen Informationen der Blickbewegungen angezeigt (gelesen aus der Datei general.txt) und das obere und untere Perzentil einer festgelegten Spalte berechnet und farblich

gekennzeichnet (Auswahl über den Reiter *Table*). Die Spalte und die prozentuale Größe für die Berechnung des Perzentils können durch den Benutzer in der Konfigurationsdatei festgesetzt werden (comparison.properties.txt; siehe Abbildung 9). Durch Selektion oder Deselektion der Datenbezeichner in der Baumstruktur können Datensätze für die weitere Verarbeitung ein- oder ausgeschlossen werden.

Nach dem Abschluss der Bereinigung der Daten kann der Benutzer den Vergleich oder die erneute Analyse mit den bereinigten Daten starten (Schaltfläche *Store And Start Comparison*). Die Konfiguration der Auswahl wird automatisch in den Analyseordner der Datengruppe geladen. Der Benutzer kann nun auswählen, ob der nächste Prozessschritt gestartet oder ob eine weitere Datengruppe bereinigt werden soll.

Durch die Schaltfläche *Save Config* kann die Konfiguration der Auswahl in einem beliebigen Verzeichnis abgespeichert werden, um beispielsweise für eine Datengruppe verschiedene Konfigurationen vorzuhalten. Für den nachfolgenden Prozessschritt muss eine Konfigurationsdatei in dem Analyseordner der Datengruppe liegen und als content_correct.txt bezeichnet sein (diese wird als Standardauswahl für den Speicherort und den Dateinamen angeboten). Durch die Schaltfläche *Load Config* kann eine Konfiguration von einem beliebigen Verzeichnis aus geladen werden.

3.4.5 Vergleich der Daten

Die letzte Komponente von SimTrA ermöglicht den Vergleich oder die erneute Analyse von bereinigten Daten. Dafür steht die Benutzungsoberfläche für die Vergleichsvorbereitung bereit (siehe Abbildung 13).

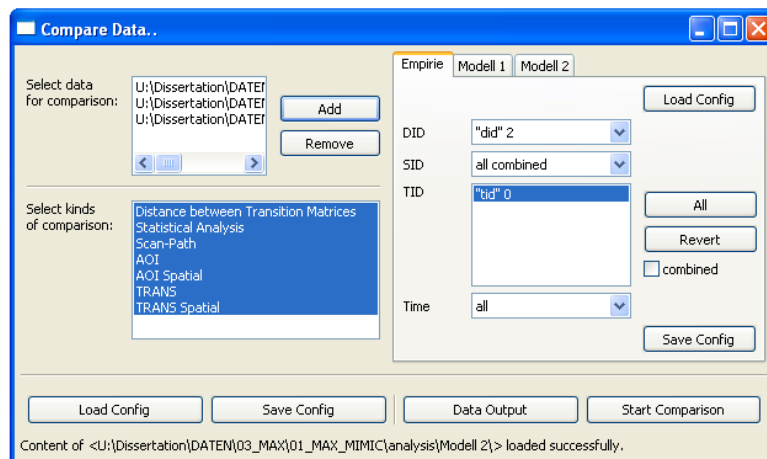


Abbildung 13: Benutzungsoberfläche von SimTrA für den Vergleich von Daten-
gruppen.

Für die Verarbeitung der Daten können bis zu drei Datengruppen in die Komponente integriert werden (Schaltfläche *Add*). Auszuwählen ist der jeweilige Analyseordner der Datengruppe. Von dort werden die Eingabedaten gelesen und die gewünschten Konfigurationen geladen.

Für jede Datengruppe wird auf der rechten Seite der Benutzeroberfläche eine Ansicht dargestellt, die für die Datenbezeichner und die Zeitintervalle der jeweiligen Datengruppe alle Informationen bereitstellt. In der Ansicht einer Datengruppe kann der Benutzer die Datengruppe auf bestimmte Aspekte einschränken (d.h. Teilmengen bilden), die für den anschließenden Vergleich genutzt werden. Versuchspersonen können einzeln ausgewählt werden oder gemittelt in den Vergleich einfließen. Die Versuchsdurchführungen können in beliebiger Kombination gewählt werden. Zusätzlich kann über ein Auswahlkästchen die Option zugeschaltet werden, alle Daten der Versuchsdurchführungen zusammenzufassen. Die Auswahl der Zeitintervalle verhält sich analog zu der Auswahl der Versuchspersonen. Die Konfiguration jeder Datengruppe kann einzeln abgespeichert (Schaltfläche `Save Config` in der Ansicht der Datengruppe) und erneut geladen werden (Schaltfläche `Load Config` in der Ansicht der Datengruppe).

Die Auswahl der Methoden für die Analyse der Daten, die verglichen werden sollen, erfolgt in der Auswahlliste `Select kinds of comparison`. In dieser Liste können die Methoden selektiert und deselektiert werden. Die Liste wird aus der Datei `methods.txt` in dem Startverzeichnis von SimTrA geladen und kann erweitert werden, wenn neue Methoden in das Werkzeug SimTrA integriert werden.

Eine Gesamtkonfiguration einer Auswahl (d.h. Datengruppen, Analysemethoden und Teilmengen der Datengruppen) kann gesondert abgespeichert (Schaltfläche `Save Config`) und erneut geladen werden (Schaltfläche `Load Config`). Somit kann der Arbeitsprozess unterbrochen oder erneute Datenanalysen mit abgeänderten Teilmengen der Datengruppen und Parameter durchgeführt werden.

Die Schaltfläche `Data Output` erlaubt die Benennung eines Ergebnisordners. In diesem Zusammenhang werden Parameter für die gewählten Analysemethoden abgefragt. Darunter fallen die Länge der lokalen Fixationspfade (Standardwert 3), die Anzahl der Fixationen in einem Feld für die räumliche Dichte (Standardwert 1) und die Anzahl der Übergänge für die Übergangshäufigkeiten (Standardwert 1). Somit können verschiedene Parameter einfach abgeändert und verschiedene Analysen sukzessiv durchgeführt werden.

Der abschließende Vergleich der Datengruppen oder die erneute Analyse einer Datengruppe wird durch `Start Comparison` gestartet. Die Konfigurationsdateien (`main_config.txt` und `compare_option.txt`) werden in den entsprechenden Ordnern abgelegt und die Analyse mit R gestartet. Die Bereitstellung der Ergebnisse erfolgt in zwei Schritten: (1) Datenanalyse und (2) Datenintegration.

DATENANALYSE Die Daten einer Datengruppe werden einzeln analysiert und die Ergebnisse in dem Analyseordner der Datengruppe (Wertebereich 1..n) abgelegt (siehe Kapitel 3.4.3). Der Pfad zu den Ergebnisdaten jeder Datengruppe wird an die Datenintegration übergeben.

DATENINTEGRATION In diesem Schritt werden für jede ausgewählte Analysemethode die Ergebnisdaten der Datengruppen zusammengefasst. Je nach Auswahl der Teilmengen der Datengruppen werden die Daten gruppiert (nach Versuchspersonen, nach Versuchsdurchläufen

oder nach beiden Bezeichnern zusammen). Wurde ein Datenbezeichner zusammengefasst, wird dieser in der Ergebnistabelle mit NA (*not available*) dargestellt. Sind beispielsweise alle Versuchspersonen einer Datengruppe für den Vergleich zusammengefasst worden, wird dies in der Spalte für die Versuchspersonennummern (sID) in den betreffenden Zeilen der Datengruppen NA angezeigt.

Für den Vergleich von Datengruppen werden die Datenbezeichner für die Herkunft der Daten (dID) neu gesetzt, um zu verhindern, dass Daten mit gleicher Herkunft durch die Verarbeitung zusammengelegt werden. In der Datei info.txt im Ergebnisordner werden die originalen und die neuen Datenbezeichner abgelegt.

In den Ergebnistabellen eines Vergleichs (beispielsweise aoi_compare.txt) sind die Informationen für alle beteiligten Datengruppen und die wichtigsten Parameter aufgeführt (wie z. B. Anzahl der Versuchspersonen und Auswahl der Gruppierung). Zusätzlich zu den Ergebnistabellen werden für jede Analyseverfahren die Eingangsdaten, die für den Vergleich genutzt wurden, abgelegt (beispielsweise aoi_data.txt), um einen schnellen Zugriff auf die Vergleichsdaten zu ermöglichen. Alle Ergebnistabellen werden in dem gewählten Ergebnisverzeichnis in einem neuen Ordner (Wertebereich 1..n) abgelegt.

Anschließend werden graphische Ausgaben erstellt und in dem Unterverzeichnis *Plots* abgelegt. Als Datengrundlage der graphischen Ausgaben sind für jede Analyseverfahren Dateien mit der Endung *_transdata.txt* abgelegt. Die Graphiken werden entsprechend zu der Analysekomponente zum einen als .png für den schnellen Überblick und zum anderen als hochauflösendes .ps erzeugt. Sollen die Graphiken individuell angepasst werden, kann der Benutzer die Datei compareInfo.txt im Ergebnisverzeichnis bearbeiten und den Vergleich erneut ausführen. In dieser Datei können *Cut-off*-Kriterien für die Datendarstellung und eine Spalte festgelegt werden, nach der in ansteigender Reihenfolge die Daten sortiert dargestellt werden sollen.

3.5 VERIFIKATION DES WERKZEUGS SIMTRA

Die Verifikation der Implementierung stellt einen wichtigen Bestandteil für die Realisierung des Werkzeugs SimTrA dar. Anhand von Black-Box-Tests (Sommerville 2001) kann das System aufgrund der Ein- und Ausgaben überprüft werden.

Begleitend zu der Implementierung von SimTrA wurden Modultests und algorithmische Tests durchgeführt. Diese stellten die spezifikationsgemäße Realisierung der einzelnen Komponenten sicher. Abschließend wurden verschiedene Integrationstests durchgeführt.

In diesem Abschnitt wird ein Integrationstest mit einem reduzierten Testbett dargestellt. Das Ziel dieses Integrationstests ist die Überprüfung, ob für die Testdaten die Analyse- und Vergleichsalgorithmen die korrekten Ausgabewerte berechnen und ob das Zusammenwirken von Java und R spezifikationsgemäß funktioniert. In diesem Zusammenhang wurde ebenfalls geprüft, ob die notwendigen Dateien angelegt und für die einzelnen Systemkomponenten zugreifbar sind (Black-Box-Test).

Für die Verifikation in SimTrA wurde eine reduzierte Testeingabe spezifiziert und der Prozess der Datenanalyse mit der standardisierten Konfiguration in allen Prozessschritten durchlaufen. Vor jedem Prozessschritt wurden die erwarteten Ergebnisse manuell erhoben und berechnet und diese mit den Systemausgaben verglichen. Zusätzlich wurden die von dem System abgelegten Dateien mit den konzeptionell erwarteten verglichen.

Während der Verifikationsphase wurden die Anforderungen an ein Werkzeug zur Datenanalyse (siehe Kapitel 3.2) exemplarisch überprüft.

Testdaten

Die Testdaten für den Black-Box-Test stellen ein fiktives Blickprotokoll einer Versuchsperson dar. Die Fixationszeiten wurden so gewählt, dass eine einfache Überprüfung der Systemausgaben möglich ist.

Das Blickprotokoll besteht aus 20 Fixationen, die abwechselnd auf zwei unterschiedliche Punkte, (B1(50,150) und B2(250,150)), gerichtet sind (siehe Abbildung 14). Eine Fixation dauert 950 Millisekunden, der Wechsel von einem Fixationspunkt zu dem anderen Fixationspunkt 50 Millisekunden. Somit ergibt sich für das Blickprotokoll eine Dauer von 20 Sekunden, wobei die erste Fixation bei 50 Millisekunden stattfindet. Für jeden Blickpunkt gibt es 10 Fixationen (je 9,5 Sekunden).

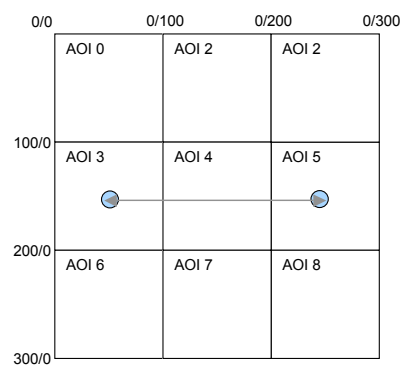


Abbildung 14: Übersicht der Testdaten.

Aufbereitung

Für die Aufbereitung wurde das Blickprotokoll in SimTrA importiert und ein Blickgebiet von 300 mal 300 Pixeln mit einem 3 mal 3 Raster angelegt (siehe Abbildung 14). Zusätzlich wurde eine schematische Abbildung der Versuchsanordnung geladen und ein Zeitintervall von 10000 bis 20000 Millisekunden erzeugt.

Alle erwarteten Systemausgaben der Komponente Aufbereitung wurden erzeugt und in der Verzeichnisstruktur in den spezifizierten Ordnern abgelegt.

Analyse

Für die Analyse der Daten wurden alle Analysemethoden ausgewählt und die Analyse gestartet. In Tabelle 5 sind die erwarteten und die von SimTrA berechneten Ergebnisse zusammengefasst. Es zeigt sich, dass die Systemausgaben mit den erwarteten Werten übereinstimmen. Alle Ergebnisdateien (d.h. Tabellen und graphische Ausgaben) und Steuerungsdateien wurden angelegt und in den jeweiligen Ordnern der Verzeichnisstruktur abgelegt.

Tabelle 5: Von SimTrA berechnete Werte der Analyse für die Testdaten im Bezug zu den manuell errechneten Werten.

		SimTrA	Manuell
Änzahl von Fixationen	Gesamt	20	✓
	Intervall	10	✓
Fixationsdichte und Räumliche Dichte	Gesamt	22,2 %	✓
	Intervall	22,2 %	✓
Relative Häufigkeiten der Fixationen	Gesamt	AOI3: 50 %, AOI5: 50 %	✓
	Intervall	AOI3: 50 %, AOI5: 50 %	✓
Kumulierte Fixationsdauer	Gesamt	AOI3: 9,5 Sek., AOI5: 9,5 Sek.	✓
	Intervall	AOI3: 4,75 Sek. AOI5: 4,75 Sek.	✓
Übergangshäufigkeiten	Gesamt	AOI3>5: 52,63 %, AOI5>3: 47,37 %	✓
	Intervall	AOI3>5: 55,56 %, AOI5>3: 44,44 %	✓
Übergangsdichte	Gesamt	2,47 %	✓
	Intervall	2,47 %	✓
Statistische Analyse der Häufigkeiten	Gesamt	$H_c: -0,006$	✓
	Intervall	$H_c: -0,027$	✓
Lokale Fixationspfade	Gesamt	AOI3>5>3: 50 %, AOI5>3>5: 50 %	✓
	Intervall	AOI3>5>3: 50 %, AOI5>3>5: 50 %	✓

Sek. : Sekunden

Bereinigung und Vergleich

Für die Bereinigung wurden die zuvor analysierten Daten ein zweites Mal analysiert und abgespeichert. Beide Datensätze wurden in die Komponente zur Bereinigung der Daten integriert und betrachtet. Die Daten mussten für den Vergleich nicht bereinigt werden. Anschließend wurden die Daten für den Vergleich zweimal geladen und alle möglichen Vergleiche ausgeführt.

Bei der Bereinigung und dem Vergleich wurden alle Konfigurations- und Ergebnisdateien angelegt und in der Verzeichnisstruktur abgelegt. Hinsichtlich der deskriptiven Vergleiche wurden die Ergebnistabellen mit den Werten der Analyse verglichen, um die korrekte Übertragung der Analysedaten zu gewährleisten. Abschließend wurden die Datenbezeichner überprüft.

Die Datenbezeichner der Ergebnistabellen wurden spezifikationsgemäß erstellt. Alle Daten für die Vergleiche wurden aus den Analysetabellen richtig übernommen. Abgeleitete und zusammengefasste Werte wurden adäquat berechnet. Alle graphischen Ausgaben wurden erstellt.

Ergebnis

Bei der Verifikation von SimTrA konnte bei der Ausführung der Algorithmen für die Analyse und den Vergleich der Blickbewegungsdaten kein Fehlverhalten festgestellt werden. Die Ergebnisse von SimTrA entsprechen den zuvor manuell erhobenen und berechneten Werten.

Die einzelnen Schritte der Verifikationsphase zeigten ebenfalls, dass die implementierte Benutzungsoberfläche in Java die effektive und effiziente Verarbeitung von Blickbewegungsdaten ermöglicht und das Zusammenwirken von Java und R problemlos funktioniert. Alle Steuerungs- und Konfigurationsdateien wurden angelegt und konnten zwischen den verschiedenen Komponenten ausgetauscht werden. Es zeigte sich, dass die Anforderungen bezüglich eines Werkzeugs zur Datenanalyse, die konzeptionell in SimTrA integriert sind, in der Implementierung umgesetzt wurden.

In diesem Kapitel wird die Anwendung von SimTrA beispielhaft an einer experimentellen Interfacestudie unter Verwendung von simulierten und empirisch erhobenen Blickbewegungsdaten gezeigt. Das Untersuchungsdesign der Interfacestudie wird eingeführt und anschließend werden die Ergebnisse diskutiert. Dieses Kapitel schließt mit einer Beschreibung von weiteren experimentellen Studien, in denen SimTrA für die Analyse von Blickbewegungsdaten eingesetzt wurde.

4.1 EXPERIMENTELLE HAUPTSTUDIE

Mit der Hauptstudie soll zum einen gezeigt werden, dass SimTrA die Analyse und den Vergleich von empirischen und simulierten Blickbewegungsdaten im Bereich der Mensch-Maschine-Interaktion unterstützt. Zum anderen soll der Vorteil der Analyse von feinen Simulationsdaten auf der Ebene der Blickbewegungsdaten für die Bewertung des Simulationsverhaltens gegenüber den Standardanalysen (z. B. Fehler, Ausführungszeiten) dargestellt werden.

Zu diesem Zweck wurde eine experimentelle Studie durchgeführt. Während der Interaktion mit drei alternativen Gestaltungen einer Mensch-Maschine-Schnittstelle für eine komplexe Steuerungs- und Kontrollaufgabe wurden Leistungsdaten erfasst und anschließend analysiert. Für die Simulation wurden zwei kognitive Modelle entwickelt, die sich hinsichtlich der Implementierung der kognitiven Prozesse unterscheiden. Das erste Modell simuliert starres regelbasiertes Verhalten, das ohne Situationseinflüsse einem Regelkreis folgt und die Aufgabe bearbeitet (Top-down-Verhalten). Das zweite Modell simuliert flexibles Verhalten, das für die Bearbeitung der Aufgabe Umgebungsinformationen berücksichtigt und für die Auswahl der Aufgabenschritte integriert (Bottom-up-Verhalten). Die Schnittstellen unterschieden sich in der Anordnung der Interaktions- und Rückmeldeelemente.

Betrachtet wurden die Leistungsdaten (d. h. die Performanzdaten und Blickbewegungen) der Versuchsgruppen bei der Interaktion mit drei alternativen Schnittstellen. In der experimentellen Hauptstudie wurden zum einen die drei Schnittstellen aufgrund der Leistungsdaten für jede Versuchsgruppe untersucht und Unterschiede betrachtet und bewertet. Zum anderen wurden die kognitiven Modelle miteinander verglichen und im Bezug zu den empirischen Daten betrachtet.

4.1.1 Aufgabe

Die Aufgabe in der Untersuchung stellt die Kontrolle und Steuerung einer Teilaufgabe einer komplexen und dynamischen Schnittstelle zur Steuerung einer Destillationskolonne dar. Sie entspricht der in der Voruntersuchung zu SimTrA genutzten Aufgabe (siehe Kapitel 3.1.1). Das Ziel besteht in der Destillation einer Flüssigkeit in einem Tank (siehe

Abbildung 15). Zusammengefasst wird der Flüssigkeitspegel durch das Regeln eines Regelventils (S1202) und einer Heizung (H1202) gesteuert, wobei die Heizung für den Destillationsprozess angeschaltet sein muss. Fällt der Flüssigkeitspegel unter die untere Grenze, wird die Heizung automatisch ausgeschaltet, um eine Schädigung der Prozessanlage zu verhindern. Die einzelnen Bedienelemente werden über die bereitgestellte Schnittstelle mit der Maus bedient. Einmaliges Klicken auf eine Bedienfläche ändert den aktuellen Zustand des Bedienelements (z. B. von *auf* nach *zu* für das Regelventil).

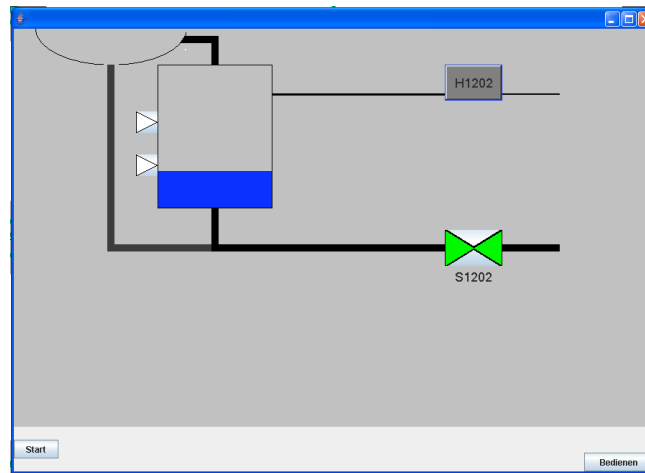


Abbildung 15: Ansicht der Versuchsumgebung für die Kontroll- und Steuerungsaufgabe (H1202: Heizung, S1202: Regelventil).

Für die Untersuchung wurde eine Prozesssimulation genutzt. Diese wurde für alle Versuchsgruppen mit gleichen initialen Werten gestartet (d. h. der Flüssigkeitspegel lag bei 26 %) und ausgeführt. Die obere Grenze für die Kontroll- und Steuerungsaufgabe wurde bei 80 % und die untere Grenze bei 30 % des Füllung der Pegelstandsanzeige festgelegt. Die Prozessdaten wurden alle 333 Millisekunden in der Mensch-Maschine-Schnittstelle aktualisiert.

4.1.2 Versuchsgruppen

Für die Untersuchung wurden drei Versuchsgruppen unterschieden, die analysiert und verglichen wurden. Zum einen wurde das Interaktionsverhalten von realen Versuchspersonen erfasst und zum anderen das Interaktionsverhalten von zwei unterschiedlichen kognitiven Modellen, die sich in der internen Struktur der Aufgabenabarbeitung und der visuellen Suche unterscheiden.

Versuchspersonen

Die Versuchspersonen waren 34 Studenten der Technischen Universität Berlin, davon 19 männliche und 15 weibliche (Altersdurchschnitt: 28,4 Jahre), und trugen keine Brille. Für die Teilnahme an der Untersuchung erhielten sie 10 Euro.

Kognitive Modelle

Für die kognitiven Modelle wurde in einer qualitativen Expertenstudie ($n = 9$) das Interaktionswissen für gelernte Anwender bei der Kontroll- und Steuerungsaufgabe erhoben. Für die Kontroll- und Steuerungsaufgabe wurde die Position des Flüssigkeitspegels durch die Experten in fünf Bereiche eingeteilt: unter der unteren Grenze, niedrig, optimal, hoch und über der oberen Grenze. Zusätzlich musste der Trend des Flüssigkeitspegels betrachtet werden, wenn dieser fallend war. Es zeigte sich, dass für die erfolgreiche Abarbeitung der Aufgabe acht Regeln zu lernen sind (siehe Tabelle 6), die für die Steuerung ausreichend sind. Jede Regel basiert auf einem Gesamtzustand der Schnittstellenelemente der Experimentalumgebung, aus dem eine Handlungsanweisung abgeleitet wird. Die acht Regeln wurden als Grundlage für die Entwicklung der kognitiven Modelle genutzt und in Produktionsregeln und deklarative Wissens Elemente umgesetzt. Die Zustände der Schnittstellenelemente werden in dem Zielzustand der kognitiven Modelle intern repräsentiert. Die Bereiche niedrig und hoch werden mit 20 % über bzw. 20 % unter den jeweiligen Grenzen initialisiert.

Alle Bedien- und Aufmerksamkeitsprozesse der kognitiven Modelle werden über die bereitgestellten Funktionen und Produktionsregeln von AGImap (siehe Kapitel 2.4.4) realisiert.

Tabelle 6: Regeln für die Steuerung und Kontrolle der untersuchten Aufgabe. Der Pegel der Flüssigkeit im Tank ist in fünf Abschnitte eingeteilt: über der oberen Grenze, hoch, optimal, niedrig und unter der unteren Grenze. Die Heizung kann an- oder ausgeschaltet sein. Das Regelventil kann geöffnet oder geschlossen sein.

Regel	Bedingungsteil			Ereignisteil	
	Heizung	Regelventil	Pegel	Trend	Handlung
1		geschlossen	über		Öffnen des Regelventils
2	aus	geöffnet	über		Anschalten der Heizung
3	aus		hoch		Anschalten der Heizung
4	an	geschlossen	hoch		Öffnen des Regelventils
5	aus		optimal		Anschalten der Heizung
6	aus		niedrig		Anschalten der Heizung
7	an	geöffnet	niedrig	fallend	Schließen des Regelventils
8		geöffnet	unter		Schließen des Regelventils

TOP-DOWN-ORIENTIERTES KOGNITIVES MODELL Das Top-down-Modell implementiert eine feste interne Struktur der Abläufe durch die Anwendung von Zustandsvariablen in den Produktionsregeln (Top-down-Verhalten). Jede Produktionsregel besitzt einen einzigartigen Zustand im Bedingungsteil (d. h. eine Zustandsvariable), der die Auswahl der Produktionsregel begründet, und einen nachfolgenden Zustand im Ereignisteil, der die Auswahl der nächsten Produktionsregel bestimmt. Durch die starre interne Struktur wird das beobachtbare Modellverhalten bestimmt.

Das kognitive Modell erfasst fortlaufend die relevanten Bedienelemente in einer von dem Modell vordefinierten Reihenfolge und legt die

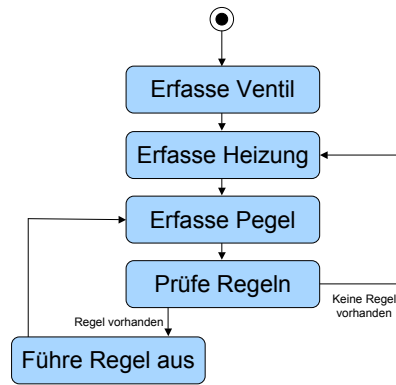


Abbildung 16: Schematische Darstellung des internen Ablaufzykluses des Top-down-orientierten kognitiven Modells.

Zustände in einer internen Repräsentation ab. Zu diesem Zweck wurde ein Regelkreis für die Aufgabenbearbeitung implementiert (siehe Abbildung 16). In den ersten drei Schritten werden die Bildelemente in einer festgelegten Reihenfolge betrachtet und der Zustand jedes Bedienelements (Heizung und Regelventil) sowie des Flüssigkeitspegels mit dem aktuellen Trend in der internen Repräsentation abgelegt. Anschließend wird der aktuelle Zustand der internen Repräsentation genutzt, um eine Handlung auszuwählen, das heißt es wird überprüft, ob eine der acht Regeln zu der aktuellen Datenrepräsentation passt (siehe Tabelle 6). Kann eine Regel ausgewählt werden, wird diese ausgeführt und der neue Zustand des geänderten Bedienelements (d. h. Heizung oder Regelventil) wird abgelegt. Kann keine Regel ausgewählt werden, wird die Heizung auf ihren Zustand überprüft und der Zustand in der internen Repräsentation abgelegt. Somit wird die Heizung kontinuierlich auf ihren Zustand überprüft. Der Sonderfall, das heißt der Flüssigkeitspegel sinkt unter die untere Grenze und die Heizung wird automatisch ausgeschaltet, wird ebenfalls durch das kognitive Modell berücksichtigt. Anschließend startet der Regelzyklus bei dem Erfassen des Flüssigkeitspegels und des Trends erneut.

Die Erfassung eines Schnittstellenelements beinhaltet die Ausrichtung der visuellen Aufmerksamkeit auf das jeweilige Element, die Erfassung des aktuellen Zustands (Farbe oder Position), die Interpretation des Zustandes und die Speicherung des Zustandes in der internen Repräsentation des kognitiven Modells. Die Ausführung einer Regel beinhaltet die Ausrichtung der Aufmerksamkeit auf das jeweilige Interaktionselement, das Schalten des Bedienelements mit der Maus und das Ablegen des aktuellen Zustandes in der internen Repräsentation des kognitiven Modells.

Die Top-down-Implementierung erlaubt die Abbildung von stark regelbasiertem Verhalten. Die Reaktion auf unvorhergesehene Zustände des Systems sind nur bedingt möglich und der Abbruch von angestoßenen Abläufen ist nicht möglich. Eine gestartete Aktion wird auch dann durchgeführt, wenn die Ausführung aufgrund einer Änderung der Experimentalumgebung nicht mehr notwendig ist (z. B. Schalten der Heizung).

BOTTOM-UP-ORIENTIERTES KOGNITIVES MODELL Das Bottom-up-Modell basiert auf der Annahme der minimalen Kontrolle von Zuständen der menschlichen Kognition (Taatgen 2007). Die Annahme besagt, dass das menschliche Verhalten durch minimal notwendige interne Kontrollstrukturen gesteuert wird, die zur Ausführung einer Aufgabe notwendig sind. Somit werden vorhandene Umgebungsinformationen genutzt, um das Verhalten dynamisch zu steuern (Bottom-up-Verhalten).

Für kognitive Modelle in der kognitiven Architektur ACT-R folgt, dass Informationen (z. B. in den Buffern der Module; siehe Kapitel 2.2.1), die zu einem Zeitpunkt vorhanden sind, integriert und für die Auswahl des folgenden Produktionsschrittes berücksichtigt werden müssen. Für darüber hinaus notwendige Kontrollstrukturen können zusätzliche Zustandsvariablen eingeführt werden, die zum einen Produktionen verketten und zum anderen Gruppen von parallel aktivierten Produktionsregeln definieren (entspricht einer geschichteten Top-down-Steuerung). Für ein kognitives Modell bedeutet dies, dass die Produktionsregeln nicht in einer festen Reihenfolge implementiert sind, sondern je nach aktuellem Zustand des künstlichen kognitiven Systems ausgewählt werden. Das kognitive Modell kann dementsprechend angefangene Ausführungsabläufe abbrechen und Änderungen in der Experimentalumgebung schnell wahrnehmen und adäquat reagieren.

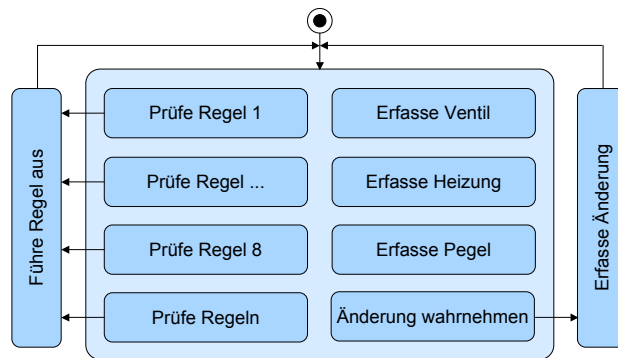


Abbildung 17: Schematische Darstellung des Bottom-up-orientierten kognitiven Modells.

Das Bottom-up-Modell nutzt die interne Repräsentation der Schnittstellenelemente und die Zustände der Buffer des kognitiven Modells, um eine Produktionsregel zur Ausführung auszuwählen. In Abbildung 17 ist die interne Organisation des Bottom-up-Modells schematisch verkürzt dargestellt. Die Gruppierung von Aufgaben bedeutet eine theoretisch parallele Aktivierung dieser Aufgaben.

Für die Aufgabebearbeitung wurden Produktionen implementiert, die für jedes Schnittstellenelement, dessen Zustand unbekannt ist, die Aufmerksamkeit auf das Bedienelement richten, den Zustand erkennen und diesen in die interne Repräsentation ablegen. Weiterhin wurde für jede Regel zur Steuerung und Kontrolle der Prozessaufgabe (siehe Tabelle 6) eine Produktionsregel erstellt, die die jeweiligen Zustände der Schnittstellenelemente als Eingangsbedingung besitzt und die entsprechende Handlung in ihrem Aktionsteil anstößt. Sobald eine

Bedingung durch den aktuellen Zustand der internen Repräsentation erfüllt ist, wird die Produktionsregel aktiviert und kann zur Ausführung ausgewählt werden. Die Ausführung der Handlung wird über die Verwendung von Zustandsvariablen realisiert.

Änderungen in der Experimentalumgebung werden durch das visuelle Modul wahrgenommen und in dem visuellen Buffer bereitgestellt. Eine Produktionsregel wurde eingeführt, die aktiviert ist, wenn eine Änderung der Experimentalumgebung durch das visuelle Modul festgestellt wurde und keine Zustandsvariable im aktuellen Zielzustand vorhanden ist. Diese Regel stößt einen Prozess an, der die visuelle Aufmerksamkeit auf das geänderte Element richtet, den Zustand erfasst und diesen in die interne Repräsentation der Schnittstellenelemente integriert (entspricht den Aufgaben *Änderung wahrnehmen* und *Erfasse Änderung* in Abbildung 17). Die Erfassung einer Zustandsänderung wird ebenfalls durch Zustandsvariablen gruppiert.

Das kognitive Modell reagiert dementsprechend auf Änderungen in der Experimentalumgebung und integriert neue Informationen in die interne Repräsentation der Schnittstellenelemente. Eine Handlung (d. h. das Schalten der Heizung oder des Regelventils) wird abhängig von der internen Repräsentation ausgeführt.

4.1.3 Versuchsdesign

Die zweite unabhängige Variable, neben den Versuchsgruppen (siehe Kapitel 4.1.2), ist die Art der Schnittstelle, für die drei unterschiedliche Alternativen gestaltet wurden. Diese unterscheiden sich in der Anordnung der Interaktions- und Rückmeldeelemente (siehe Abbildung 18) und werden im Folgenden als Szenarien bezeichnet. In der ersten Schnittstelle sind die Interaktion und die Rückmeldung für ein Bedienelement in einem Element innerhalb der schematischen Abbildung der Destillationskolonne zusammengefasst. In der zweiten Schnittstelle werden für jedes Bedienelement ein Interaktionselement und ein Rückmeldeelement eingeführt. Die Interaktionselemente werden in einer zusätzlichen Leiste unterhalb der schematischen Darstellung der Destillationskolonne bereitgestellt. Die Rückmeldeelemente befinden sich innerhalb der schematischen Abbildung der Destillationskolonnen. Die dritte Schnittstelle stellt für die Interaktionselemente eine Leiste und für die Rückmeldeelemente eine Leiste unterhalb bzw. oberhalb der schematischen Abbildung bereit. In der dritten Schnittstelle werden die Rückmeldeelemente in Form von farbigen Leuchten dargestellt.

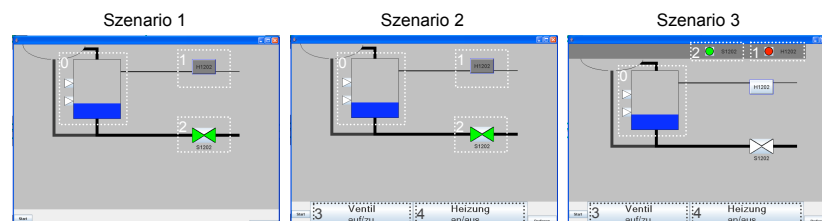


Abbildung 18: Drei Schnittstellen für die Untersuchung mit den für die Analyse festgelegten AOIs (Szenario 1, Szenario 2 und Szenario 3).

Als abhängige Variablen wurden die Blickbewegungen (X- und Y-Koordinaten und Zeitstempel) erfasst und die Performanz der Anwender in der Kontroll- und Steuerungsaufgabe erhoben.

Die empirische Studie wurde innerhalb von zwei Wochen im Februar 2007 an der Technischen Universität Berlin durchgeführt. Jede Versuchsperson bediente in einem Versuchsablauf alle drei Schnittstellenszenarien. Die Reihenfolge wurde permutiert. Die Versuchsdauer lag bei durchschnittlich 30 Minuten.

Die Simulationen wurden für jedes Szenario und jedes Modell hintereinander ausgeführt. Dafür wurden für jeden Durchlauf das deklarative Gedächtnis und alle von dem kognitiven Benutzermodell abgelegten Informationen aus der Simulationsumgebung gelöscht.

4.1.4 *Versuchsablauf*

Der Versuch wurde mit einer kurzen Einführung des Versuchsaufbaus und der nachfolgenden Schritte begonnen. Anschließend wurden die demographischen Daten erhoben. Nach den einführenden Schritten wurde die Versuchsperson gebeten, sich vor den Arbeitsplatz zu setzen, so dass die Maus bequem zu erreichen und der Bildschirm gut zu sehen ist. Im Folgenden wurde die zu bearbeitende Aufgabe anhand einer kurzen Präsentation eingeführt. Daran schloß eine Lernphase aller drei Schnittstellen an, um die Bedienoberflächen und die Aufgabe kennen zu lernen. Das System wurde als verstanden und gelernt angenommen, wenn der Anwender den Flüssigkeitspegel für jede Schnittstelle dreimal zwischen den zwei Grenzen mit eingeschalteter Heizung für 10 Sekunden stabilisieren konnte. Nachfolgend wurde der Helm zur Erfassung des Augenbildes aufgesetzt, festgeschnallt und das System für die Blickbewegungserfassung kalibriert. Anschließend wurden die Schnittstellen der drei Szenarien durch die Versuchsperson bedient und die Blickbewegung aufgezeichnet. Jedes Szenario wurde 45 Sekunden lang bearbeitet. Der Versuch endete mit der Informierung der Versuchspersonen über die Weiterverarbeitung und Nutzung der Ergebnisse des Versuchs.

Für die Simulationsstudie wurden die zwei kognitiven Modelle für jedes Szenario 30-mal 45 Sekunden lang ausgeführt und die Daten erfasst. Dabei wird jeder Modelldurchlauf als ein Element der Versuchsgruppe betrachtet.

4.1.5 *Geräte*

Die Blickbewegungserfassung erfolgte wie im Vorversuch (siehe Kapitel 3.1.5). Die Schnittstellen (600 x 400 Pixel) für die Interaktion wurden auf einem 30-Zoll-Monitor dargestellt und durch eine Maus bedient (siehe Abbildung 19). Die Auflösung, Größe und Anordnung der Schnittstellen waren für die kognitiven Modelle und die Versuchspersonen gleich. Der Abstand der Versuchspersonen von dem Bildschirm lag bei ungefähr 70 Zentimetern, der simulierte Abstand der kognitiven Modelle lag bei 70 Zentimetern.

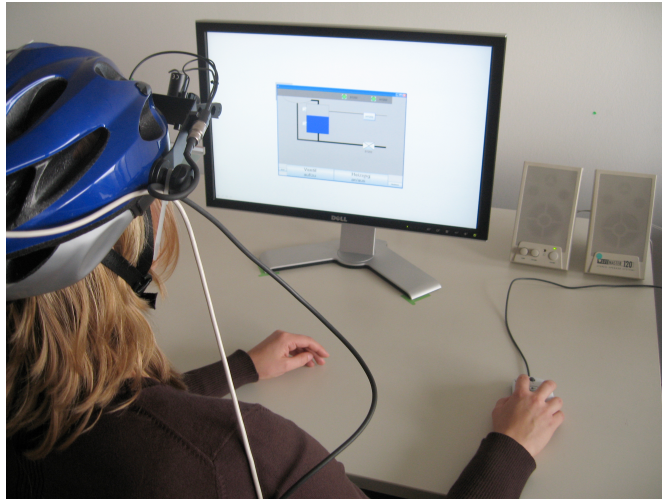


Abbildung 19: Ansicht des Versuchsaufbaus für die Versuchspersonen.

4.1.6 Versuchsauswertung

Für die Auswertung der Performanz der Anwender bei der Kontroll- und Steuerungsaufgabe wurden die Log-Dateien der Prozesssimulation mit der Software R ausgewertet. Dabei wurden die Anzahl der Interaktionen und der Stabilisierungsfaktor ermittelt. Der Stabilisierungsfaktor stellt die Dauer des Pegels innerhalb der Füllstandsanzeige zwischen den zwei Grenzen in Relation zu der Ausführungsdauer von 45 Sekunden dar. Die Interaktionen wurden nach Schaltungen der Heizung und des Regelventils unterteilt.

Die empirisch erhobenen Blickbewegungsdaten wurden mit der Software BeGaze in Fixationen transformiert (minimale Fixationsdauer: 100 Millisekunden, mögliche Abweichungen des Blickwinkels: 2 Grad). Für die simulierten Blickbewegungsdaten wurde die minimale Fixationsdauer software-seitig bei 100 Millisekunden festgelegt. Die Analyse und der Vergleich der Blickbewegungsdaten wurde mit SimTrA durchgeführt. Das Blickgebiet (Größe der Schnittstelle) wurde einheitlich festgelegt ($Ecke_{linksOben}(20/20)$ und $Ecke_{rechtsUnten}(820/620)$). Die empirischen Daten wurden innerhalb von SimTrA anhand eines eingefügten Hintergrundbildes einheitlich angepasst (X-Skalierung: 4, Y-Skalierung: 4, X-Offset: -120 und Y-Offset: -120). Für jedes Szenario und jede Versuchsgruppe wurden die Fixationsdaten analysiert und jeweils in einem Analyseordner abgelegt. Für die Analyse der relativen Häufigkeit der Fixationen, der kumulierten Fixationsdauer, der Übergangshäufigkeiten, der Anzahl der Fixationen, der statistischen Analyse der Übergangshäufigkeiten und der lokalen Fixationspfade wurden 3 bzw. 5 AOIs festgelegt (siehe Abbildung 18). Für die Analyse der Fixationsdichte, der räumlichen Dichte und der Übergangsdichte wurde ein 10×10 Gitter angelegt. Dieses Gitter wurde gewählt, um eine feine Betrachtung der Blickbewegungsdaten zu ermöglichen. Anschließend wurden die Daten bereinigt und für jedes Szenario die drei Versuchsgruppen miteinander verglichen.

Für den Vergleich der Blickbewegungsdaten der einzelnen Szenarien und für einen übersichtlichen Vergleich der Versuchsgruppen wurden die gewählten AOIs in Szenario 2 und Szenario 3 funktional zusammengefasst und die Analyse und der Vergleich der Fixationspfade und der statistischen Analyse der Übergangshäufigkeiten aller Szenarien und Versuchsgruppen erneut durchgeführt. Hierzu wurden die AOIs 2 und 3 für das Regelventil und die AOIs 1 und 4 für die Heizung zusammengefasst.

Die Auswertung wird deskriptiv durchgeführt. Inferenzstatistische Maße kommen nicht zur Anwendung, da in der Voruntersuchung gezeigt wurde, dass simulierte Daten kognitiver Modelle zum Teil keine Varianz aufweisen (siehe Kapitel 3.1). In der Auswertung wird von Ähnlichkeiten gesprochen. Diese beziehen sich auf die Standardabweichungen und Mittelwerte der jeweils betrachteten Faktoren. Dabei wird überprüft, ob die Mittelwerte innerhalb einer Standardabweichung liegen.

4.1.7 Experimentelle Hypothesen

In der Studie wurden drei experimentelle Hypothesen exemplarisch untersucht, um zeigen zu können, dass die Analyse und der Vergleich simulierter und empirischer Leistungsmaße mit SimTrA die Betrachtung von experimentalpsychologischen Fragestellungen ermöglicht.

In der ersten Hypothese wird angenommen, dass die drei unterschiedlich gestalteten Schnittstellen unterscheidbare Leistungsdaten induzieren. Für die Performanz wird angenommen, dass die unterschiedliche Anordnung der Schnittstellenelemente und die Aufteilung in Interaktions- und Rückgabeelemente für die drei Schnittstellen Unterschiede erzeugen (Hypothese 1a). Dies kann durch den erhöhten Koordinationsaufwand begründet werden. Die Interfaces werden ebenfalls als unterscheidbar angenommen hinsichtlich der induzierten Fixationspfade und der bedingten Abhängigkeit der Blickbewegungen (Hypothese 1b). Diese Annahme stützt sich darauf, dass unterschiedliche Anordnungen und Entfernungen der Interaktions- und Rückmeldeelemente unterschiedliche Kosten für die visuelle Suche erzeugen. Je größer die Entfernung der Elemente voneinander, desto größer die Kosten für die visuelle Suche. Somit ergeben sich für die visuelle Suche und die drei Schnittstellen unterschiedliche Kosten, da diese aufgrund der größer werdenden Entfernung der Schnittstellenelemente von Szenario 1 nach Szenario 3 zunehmen.

Im Weiteren wird angenommen, dass die zwei in den kognitiven Modellen abgebildeten unterschiedlichen theoretischen Annahmen zu Informationsaufnahme- und -verarbeitungsprozessen zu unterscheidbaren Leistungsdaten führen (Hypothese 2). Das Top-down-Modell steuert die Informationsaufnahme und die Interaktionsprozesse nach einem festen Schema. Das Bottom-up-Modell reagiert flexibel auf Umgebungsreize. Dieser Unterschied wird in einer unterscheidbaren Häufigkeit von wichtigen Fixationspfaden der zwei kognitiven Modelle zu sehen sein.

Die dritte Annahme besagt, dass das flexible und situationsabhängige Verhalten des Bottom-up-Modells das menschliche Verhalten bei

der Bedienung der komplexen und dynamischen Mensch-Maschine-Schnittstelle besser präzisieren kann als das einem festen Ablaufschema folgende Top-down-Modell (Hypothese 3).

4.1.8 Ergebnisse

Die Ergebnisse der Untersuchung werden getrennt nach der Analyse der Performanz und der Blickbewegungen dargestellt. In jedem Abschnitt werden die Daten im Bezug zu den drei gestellten Hypothesen betrachtet.

Performanz

Für die Analyse der Performanz wurden die Log-Dateien nach der ersten Analyse bereinigt. Datensätze mit einem Stabilisierungsfaktor kleiner als 60 % und mit mehr als 20 Interaktionen für die Heizung oder das Regelventil wurden aus der weiteren Analyse ausgeschlossen. Nach der Bereinigung konnten für das erste Szenario 31, für das zweite Szenario 33 und für das dritte Szenario 32 Datensätze von je 34 der empirischen Untersuchung genutzt werden. Für die Analyse der Simulationsdaten konnten jeweils alle 30 Datensätze genutzt werden. In Abbildung 20 sind die Ergebnisse schematisch zusammengefasst. Dargestellt sind der Mittelwert und die Standardabweichung der betrachteten Faktoren für jedes Szenario und für jede Versuchsgruppe. Zusätzlich wurden die gemittelten Daten jeder Versuchsgruppe für die drei Szenarien der Abbildung hinzugefügt.

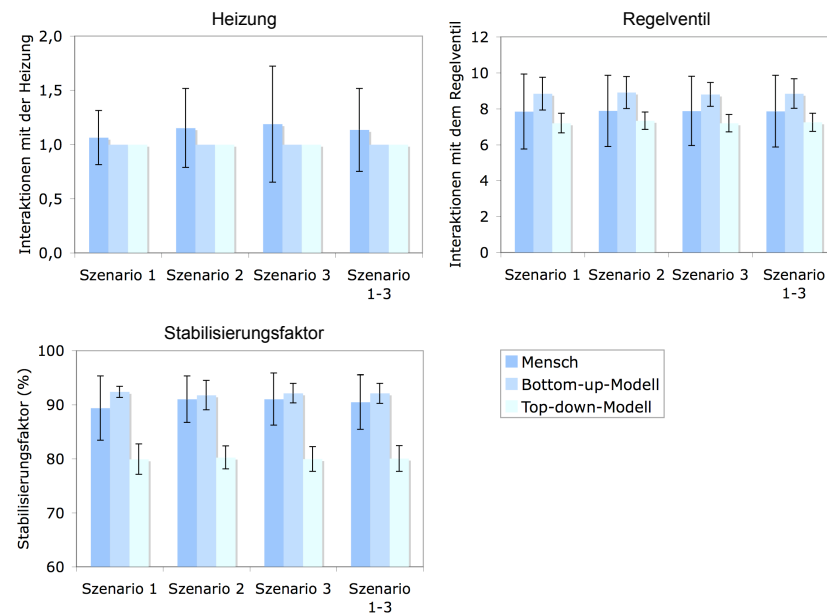


Abbildung 20: Übersicht der betrachteten Faktoren für die Bewertung der Performanz der drei Versuchsgruppen. Dargestellt sind der Mittelwert und die Standardabweichung für jeden betrachteten Faktor für die drei Szenarien und zusammengefasst über die drei Szenarien.

ERGEBNISSE ZUR HYPOTHESE 1 Für die Betrachtung, ob die drei unterschiedlichen Schnittstellen unterschiedliche Leistungsdaten induzieren, wurden die erfassten Performanzdaten in der Kontroll- und Steuerungsaufgabe der verschiedenen Szenarien für jede Versuchsgruppe miteinander verglichen. In Abbildung 20 und Tabellen 7 und 8 werden die Performanzdaten für die drei Versuchsgruppen und die drei untersuchten Szenarien dargestellt. Angegeben werden die mittlere Anzahl der Schaltungen der Heizung und des Regelventils und der mittlere Stabilisierungsfaktor in Prozentwerten mit der jeweiligen Standardabweichung.

Tabelle 7: Performanz der Menschen für die drei Szenarien. Dargestellt sind die Mittelwerte und die Standardabweichung der betrachteten Faktoren.

	Heizung		Regelventil		Stabilisierungsfaktor	
	MW	SA	MW	SA	MW	SA
Szenario 1	1,06	0,25	7,84	2,08	89,39	5,94
Szenario 2	1,15	0,36	7,88	1,98	91,03	4,30
Szenario 3	1,19	0,54	7,88	1,93	91,03	4,83

MW: Mittelwert und SA: Standardabweichung

Für die Versuchspersonen zeigt sich, dass die Mittelwerte für die Stabilisierungsfaktoren, die Anzahl der Interaktionen mit der Heizung und mit dem Regelventil jeweils sehr ähnlich sind. Für die empirischen Daten kann kein Unterschied zwischen den drei Szenarien festgestellt werden. Die erhobenen Befunde entsprechen somit nicht den Erwartungen von Hypothese 1.

Tabelle 8: Performanz des Bottom-up-Modells und des Top-down-Modells für die drei Szenarien. Dargestellt sind die Mittelwerte und die Standardabweichung der betrachteten Faktoren.

	Bottom-up-Modell					
	Heizung		Regelventil		Stabilisierungsfaktor	
	MW	SA	MW	SA	MW	SA
Szenario 1	1,00	0,00	8,83	0,91	92,40	1,03
Szenario 2	1,00	0,00	8,90	0,88	91,78	2,74
Szenario 3	1,00	0,00	8,80	0,66	92,14	1,82

	Top-down-Modell					
	Heizung		Regelventil		Stabilisierungsfaktor	
	MW	SA	MW	SA	MW	SA
Szenario 1	1,00	0,00	7,20	0,55	79,93	2,80
Szenario 2	1,00	0,00	7,33	0,48	80,25	2,12
Szenario 3	1,00	0,00	7,20	0,48	79,96	2,28

MW: Mittelwert und SA: Standardabweichung

Die Mittelwerte der simulierten Daten der kognitiven Modelle für den Stabilisierungsfaktor, die Anzahl der Interaktionen mit der Heizung und dem Regelventil sind ebenfalls für die jeweilige Versuchsgruppe

und die drei Szenarien sehr ähnlich. Im Bezug zu den Performanzdaten unterstützen die Befunde für das Bottom-up-Modell und das Top-down-Modell Hypothese 1 nicht.

Zusammenfassend kann zwischen den drei Szenarien innerhalb einer Versuchsgruppe kein unterschiedliches Verhalten auf der Betrachtungsebene der groben Leistungsdaten festgestellt werden. Das heißt, dass die drei unterschiedlichen Schnittstellen keine unterschiedliche Performanz induzieren. Zusammenfassend unterstützt die deskriptive Befundlage der Stichprobe Hypothese 1 nicht.

ERGEBNISSE ZUR HYPOTHESE 2 Für die Hypothese, dass die zwei verschiedenen theoretischen Annahmen der kognitiven Modelle unterschiedliche Leistungsdaten induzieren, wurden die Performanzdaten der kognitiven Modelle in Relation zueinander betrachtet. Da keine Unterschiede innerhalb der Versuchsgruppen für die drei Szenarien festgestellt wurden, wurden die Daten der Versuchsgruppen für die Überprüfung von Hypothese 2 zusammengefasst betrachtet. Die Performanzdaten sind in Abbildung 20 (Szenario 1–3) und Tabelle 9 abgebildet.

Tabelle 9: Durchschnittliche Performanz der drei Versuchsgruppen für die zusammengefassten Szenarien. Dargestellt sind die Mittelwerte und die Standardabweichung der betrachteten Faktoren.

	Heizung		Regelventil		Stabilisierungsfaktor	
	MW	SA	MW	SA	MW	SA
Mensch	1,13	0,38	7,86	2,00	90,48	5,02
Bottom-up-Modell	1,00	0,00	8,84	0,82	92,11	1,86
Top-down-Modell	1,00	0,00	7,24	0,50	80,04	2,40

MW: Mittelwert und SA: Standardabweichung

Der Vergleich der Performanzdaten zeigt, dass der Mittelwert des Stabilisierungsfaktors des Bottom-up-Modells um durchschnittlich 10,44 % über dem Stabilisierungsfaktor des Top-down-Modells liegt. Für den Stabilisierungsfaktor kann daher zwischen den zwei kognitiven Modellen ein Unterschied angenommen werden.

Für die Anzahl der Interaktionen mit dem Regelventil zeigt sich ebenfalls ein Unterschied. Das Bottom-up-Modell simuliert durchschnittlich 1,6 Interaktionen mehr als das Top-down-Modell. Ein Grund dafür kann in der freien Ablaufstruktur des Bottom-up-Modells liegen, das somit flexibel auf Systemänderungen reagieren kann (siehe Kapitel 4.1.2). Im Zusammenhang mit dem höheren Mittelwert für den Stabilisierungsfaktor des Bottom-up-Modells kann geschlossen werden, dass in kritischen Situationen (z. B. Pegelstand im Bereich hoch oder niedrig) das Bottom-up-Modell schneller reagieren und somit den Prozess besser steuern konnte. Das heißt, das Regelventil wurde geschlossen, bevor der Pegelstand den Grenzbereich verlassen konnte, wodurch mehr Interaktionen notwendig waren.

Beide kognitiven Modelle simulieren durchschnittlich einen Mittelwert von einer Interaktion mit der Heizung. Es kann kein Unterschied zwischen den zwei Modellen festgestellt werden. Es ist aber zu be-

achten, dass die zu bearbeitende Aufgabe das (mindestens) einmalige Einschalten der Heizung erfordert, um den Destillationsprozess zu starten. Somit ist die einmalige Interaktion mit der Heizung aus Gründen der Aufgabenumgebung notwendig. Die identische Anzahl an Interaktionen mit der Heizung kann daher nicht als Unterscheidskriterium genutzt werden, sondern deutet auf eine korrekte Implementierung der kognitiven Modelle hin.

Die erfassten Performanzdaten ermöglichen den Schluss, dass die zwei kognitiven Modelle unterschiedliches Verhalten erzeugen. Dieses ist auf die zugrunde liegenden unterschiedlichen theoretischen Annahmen der Interaktions- und Wahrnehmungsprozesse zurückzuführen. Die deskriptiv-statistischen Befunde der Stichprobe stimmen mit den Erwartungen aus Hypothese 2 gut überein.

ERGEBNISSE ZUR HYPOTHESE 3 Um zu überprüfen, ob das Bottom-up-Modell das menschliche Verhalten besser vorhersagen kann als das Top-down-Modell, wurden die Performanzdaten der kognitiven Modelle mit den empirischen Performanzdaten verglichen. Da keine Unterschiede innerhalb der Versuchsgruppen für die drei Szenarien festgestellt wurden, konnten die Daten der drei Versuchsgruppen für die Bewertung der dritten Hypothese ebenfalls zusammengefasst betrachtet werden. In Abbildung 20 (Szenario 1–3) und Tabelle 9 sind die Performanzdaten dargestellt.

Es zeigt sich, dass der Mittelwert für den Stabilisierungsfaktor des Bottom-up-Modells (92,11 %) ähnlich zu dem der empirischen Daten (90,48 %) ist. Die simulierten Daten des Top-down-Modells liegen 10,44 % unterhalb der empirischen Daten. Dieser Befund ist konform mit den Vorhersagen von Hypothese 3.

Beide Modelle simulieren Werte für die Anzahl der Interaktionen mit der Heizung und dem Regelventil, die jeweils mit den empirisch erfassten Daten für die Faktoren vergleichbar sind. Es kann nicht differenziert werden, welches der zwei kognitiven Modelle hinsichtlich der Bedienung der Heizung und des Regelventils eher das menschliche Verhalten prädiziert.

Da auf deskriptiver Ebene für die Anzahl der Interaktionen keine Unterschiede festgestellt werden konnten und der Stabilisierungsfaktor einen Unterschied aufweist, entsprechen die Befunde tendenziell den Erwartungen von Hypothese 3.

Blickbewegungen

Für die Analyse der Blickbewegungen wurden die Daten in der Komponente zur Bereinigung der Daten von SimTrA betrachtet und bereinigt. Bei den empirischen Daten wurden alle Datensätze aus der Analyse ausgeschlossen, bei denen die Anzahl von Fixationen im oberen oder unteren Perzentil lag und die aufgrund der angezeigten Graphiken der Blickbewegungsdatenverteilung als fehlerhaft eingestuft wurden. Für die Analyse der Simulationsdaten wurden alle Datensätze übernommen. In Tabelle 10 sind die allgemeinen Informationen für die Versuchsgruppen und die einzelnen Szenarien dargestellt (d. h. die Anzahl der Fixationen, die Fixationsdauer und die Wechselkosten zwischen zwei Fixationen).

Tabelle 10: Anzahl der Versuchspersonen und Mittelwerte und Standardabweichung der allgemeinen Informationen der Blickbewegungsdaten für jede Versuchsgruppe und die drei Szenarien.

Szenario 1							
	DS	Fixationen		Dauer		Wechselkosten	
		MW	SA	MW	SA	MW	SA
Mensch	23	58,30	21,73	0,84	0,38	0,08	0,05
Bottom-up-Modell	30	199,00	2,19	0,17		0,06	
Top-down-Modell	30	173,77	3,78	0,21		0,05	

Szenario 2							
	DS	Fixationen		Dauer		Wechselkosten	
		MW	SA	MW	SA	MW	SA
Mensch	24	55,50	24,45	0,92	0,39	0,08	0,05
Bottom-up-Modell	30	200,60	1,97	0,17		0,06	
Top-down-Modell	30	174,33	2,30	0,21		0,05	

Szenario 3							
	DS	Fixationen		Dauer		Wechselkosten	
		MW	SA	MW	SA	MW	SA
Mensch	24	58,25	4,71	0,88	0,44	0,07	0,10
Bottom-up-Modell	30	190,57	1,81	0,18		0,06	
Top-down-Modell	30	73,57	2,57	0,21		0,05	

DS: Datensätze, MW: Mittelwert und SA: Standardabweichung

Die Daten zeigen, dass die durchschnittliche Anzahl der Fixationen bei den empirischen Daten geringer und die durchschnittliche Fixationsdauer größer ist als die simulierte Anzahl der Fixationen und die Fixationsdauer. Ein Grund liegt darin, dass die Daten der kognitiven Modelle sehr genau erfasst werden konnten und eine Abweichung im Sehwinkel für die Berechnung der Fixationen nicht betrachtet wurde. Somit werden für die folgenden Analysen der Blickbewegungsdaten relative Werte oder kumulierte Zeitmaße genutzt und angegeben.

Für die Betrachtung der Hypothesen werden in dieser Arbeit nicht alle analysierten Daten diskutiert. Für jede Hypothese werden geeignete Daten dargestellt. Die lokalen Fixationspfade (siehe Kapitel 3.3.3) werden für die Überprüfung aller drei Hypothesen genutzt, da in vorangegangenen Studien festgestellt wurde, dass dieser Blickparameter sehr robust ist und viele Informationen weiterer Blickparameter integriert (z. B. Transitionen, Übergangsdichte). Als Entscheidungskriterium für die Auswahl der häufigsten Fixationspfade wurde die kumulierte Häufigkeit der als häufig ausgewählten Fixationspfade und der Kurvenverlauf betrachtet. Für die Betrachtung des Kurvenverlaufs wurden die Fixationspfade in eine Rangreihenfolge gebracht und hinsichtlich der Verflachung der Kurve betrachtet.

ERGEBNISSE ZUR HYPOTHESE 1 Für die Überprüfung, ob bei der Interaktion mit den drei verschiedenen Szenarien unterscheidbare Blick-

bewegungsdaten erzeugt werden, wurden die Fixationspfade der funktional zusammengefassten AOIs betrachtet. In Abbildung 21 sind für jede Versuchsgruppe die 12 möglichen Fixationspfade (bei 3 AOIs) in den drei Szenarien dargestellt. Zusätzlich wurden die statistische Analyse der Übergangshäufigkeiten für die drei Szenarien und die drei Versuchsgruppen mit zusammengefassten AOIs betrachtet (siehe Abbildung 22).

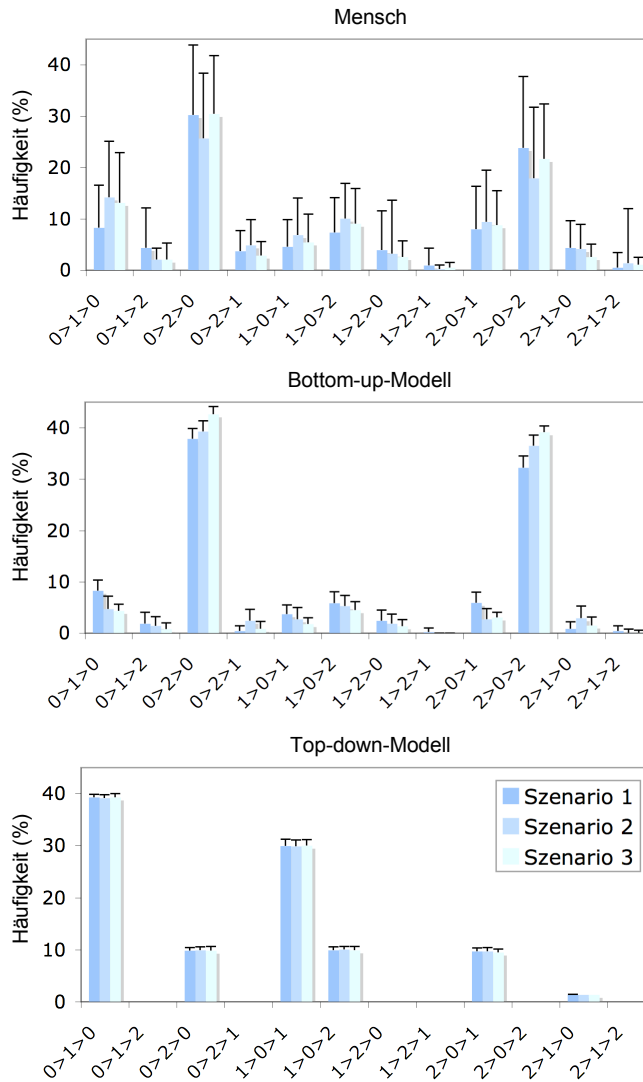


Abbildung 21: Fixationspfade der funktional zusammengefassten AOIs für die drei Versuchsgruppen und die drei Szenarien mit Standardabweichungen.

Für die Versuchsgruppe der Menschen zeigt sich, dass in allen drei Szenarien fünf Fixationspfade mit einer Häufigkeit größer als 7 % vorkommen ($0>1>0$, $0>2>0$, $1>0>2$, $2>0>1$ und $2>0>2$). Diese lokalen Fixationspfade können als wichtige Fixationspfade des Menschen gekennzeichnet werden. In den drei Szenarien ergeben sich für die kumulierte

Anzahl der wichtigsten fünf Fixationspfade ähnliche Werte (Szenario 1: 77,74 %; Szenario 2: 77,27 %; Szenario 3: 83 %), die somit ca. 80 % der erfassten Daten erklären. Die Rangreihenfolge der fünf Fixationspfade ist in Szenario 2 und Szenario 3 für den Menschen gleich. Szenario 1 unterscheidet sich in der Position des Blickpfades $1>0>2$. Alle relativen Häufigkeiten der Fixationspfade liegen innerhalb einer Standardabweichung der zwei anderen Szenarien. Die relativen Häufigkeiten der lokalen Fixationspfade sind über die drei Szenarien nicht gleich verteilt, liegen jeweils aber dicht zusammen. Aus den lokalen Fixationspfaden kann für die empirischen Daten kein eindeutiger Unterschied zwischen den drei Szenarien abgeleitet werden. Hinweise auf einen Unterschied zwischen den drei Szenarien bietet die Betrachtung der statistischen Analyse der Übergangshäufigkeiten (siehe Abbildung 22, links). Hier zeigt sich, dass das Blickverhalten in Szenario 1 eine größere statistische Abhängigkeit besitzt (geringerer H_c -Wert) als das Blickverhalten in Szenario 2 und Szenario 3 (mit zusammengefassten AOIs). Im Zusammenhang mit den unterschiedlichen Werten der lokalen Fixationspfade in den drei Szenarien und dem Unterschied in der statistischen Analyse der Übergangshäufigkeiten kann Hypothese 1 für die empirischen Daten teilweise belegt werden. Es zeigt sich, dass das Blickverhalten des Menschen in Szenario 2 und Szenario 3 ähnlich ist und sich von dem Blickverhalten in Szenario 1 unterscheidet.

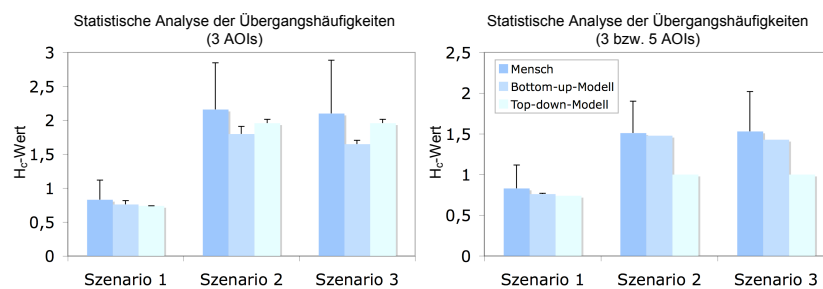


Abbildung 22: Statistische Analyse der Übergangshäufigkeiten der drei Szenarien für die drei Versuchsgruppen. Dargestellt sind die H_c -Werte mit einer Standardabweichung für die funktional zusammengefassten AOIs (links) und die 3 bzw. 5 AOIs (rechts).

Für das simulierte Blickverhalten des Bottom-up-Modells zeigt sich, dass in allen drei Szenarien Fixationspfade größer als 4 % vorkommen ($0>1>0$, $0>2>0$, $1>0>2$, $2>0>1$ und $2>0>2$). Diese werden als die wichtigsten Fixationspfade des Bottom-up-Modells festgelegt. Die fünf Fixationspfade erklären für die drei Szenarien ca. 90 % der erfassten Daten (Szenario 1: 90,15 %; Szenario 2: 88,47 %; Szenario 3: 93,67 %). Die Rangreihenfolge in Szenario 2 und Szenario 3 ist gleich. Szenario 1 unterscheidet sich in der Position des Fixationspfades $1>0>2$. Anhand der lokalen Fixationspfade kann angenommen werden, dass in Szenario 2 und Szenario 3 die Blickbewegungsdaten des Bottom-up-Modells ähnlich sind und in Szenario 1 tendenziell ein Unterschied festzustellen ist. Dieser Unterschied wird durch die statistische Analyse der Übergangshäufigkeiten mit zusammengefassten AOIs noch bestärkt. In Abbildung 22 (links) ist für die simulierten Daten des Bottom-up-

Modells zu sehen, dass die H_c -Werte für Szenario 2 und Szenario 3 ähnlich sind und mehr als ein halbes Mal über dem Wert für Szenario 1 liegen. Für das Bottom-up-Modell kann Hypothese 1 teilweise belegt werden. Es zeigt sich, dass das Blickverhalten des Bottom-up-Modells in Szenario 2 und Szenario 3 ähnlich ist und sich von dem Blickverhalten in Szenario 1 unterscheidet.

Für das Verhalten des Top-down-Modells zeigt sich, dass in allen drei Szenarien ähnliche Fixationspfade vorliegen. Die wichtigsten Fixationspfade (größer als 9 %) sind: $0>1>0$, $1>0>1$, $1>0>2$, $0>2>0$ und $2>0>1$. Die 5 Fixationspfade erklären ca. 98 % der erfassten Daten (Szenario 1: 98,63 %; Szenario 2: 98,62 %; Szenario 3: 98,63 %). Die Rangreihenfolge der Fixationspfade ist in allen drei Szenarien gleich. Eine Differenzierung der Szenarien ist auf Grundlage der Fixationspfade nicht möglich. Die statistische Analyse der Übergangshäufigkeiten für die 3 (zusammengefassten) AOIs zeigt, dass die H_c -Werte für Szenario 2 und Szenario 3 gleich und mehr als doppelt so groß sind wie der H_c -Wert von Szenario 1 (siehe Abbildung 22, links). Für das Blickverhalten des Top-down-Modells kann daher Hypothese 1 teilweise belegt werden. Aus Abbildung 22 geht hervor, dass Szenario 2 und Szenario 3 ähnliches Blickverhalten erzeugen, das unterschiedlich zu dem Blickverhalten von Szenario 1 ist.

Zusammenfassend kann Hypothese 1 für die Blickbewegungsdaten mit deskriptiv-statistischen Methoden teilweise gestützt werden. Für alle Versuchsgruppen zeigte sich, dass die Interaktion mit der Mensch-Maschine-Schnittstelle in Szenario 2 und Szenario 3 ein ähnliches Blickverhalten erzeugt. Das Blickverhalten in Szenario 1 scheint von den zwei anderen Szenarien unterschiedlich zu sein. Im Vergleich zu den Performanzdaten zeigt sich, dass die Blickbewegungsdaten eine detailliertere Bewertung der Schnittstellen ermöglichen.

ERGEBNISSE ZUR HYPOTHESE 2 Für die Überprüfung, ob die Interaktion der zwei kognitiven Modelle unterschiedliches Verhalten erkennen lässt, wurden die Fixationspfade der zwei kognitiven Modelle (siehe Abbildung 23) und die statistische Analyse der Übergangshäufigkeiten der nicht funktional zusammengefassten AOIs (siehe Abbildung 22, rechts) in den jeweiligen Szenarien betrachtet und für die zwei Versuchsgruppen miteinander verglichen. Dabei wurde auf die funktionale Zusammenfassung der AOIs verzichtet, um die Daten detaillierter betrachten zu können.

Den Fixationspfaden größer als 3 % des Bottom-up-Modells wurden die entsprechenden Werte der zwei kognitiven Modelle zugeordnet. Die ausgewählten Fixationspfade erklären für das Bottom-up-Modell im ersten Szenario 93,83 %, im zweiten Szenario 74,57 % und im dritten Szenario 83,77 % der Daten. In Szenario 1 zeigt sich, dass die zwei häufigsten Fixationspfade des Bottom-up-Modells nicht von gleicher Wichtigkeit für das Top-down-Modell sind. Dies gilt ebenfalls umgekehrt für die zwei wichtigsten Fixationspfade des Top-down-Modells ($0>1>0$ und $1>0>1$).

In Szenario 2 und Szenario 3 werden von den fünf häufigsten Fixationspfaden vier beziehungsweise drei nicht von dem Top-down-Modell simuliert. Die relative Häufigkeit der lokalen Fixationspfade ist im Vergleich der kognitiven Modelle sehr unterschiedlich. Die zwei häufigsten

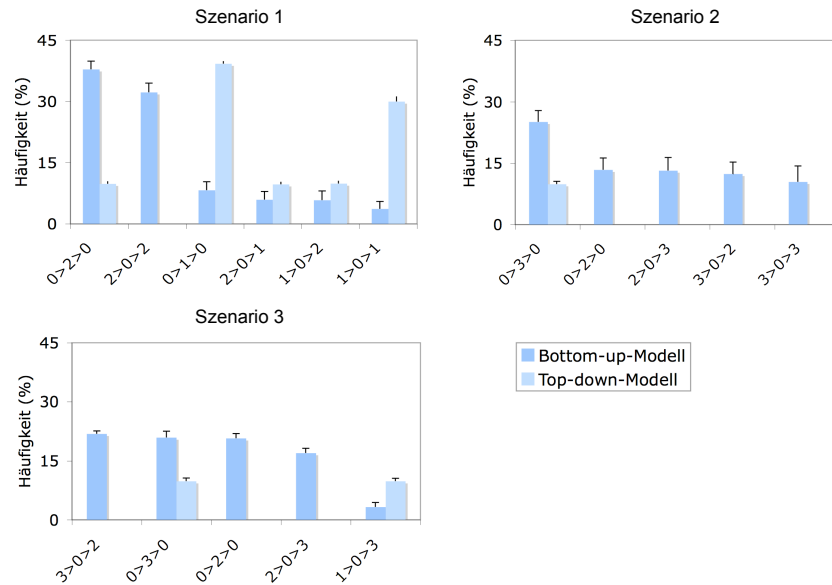


Abbildung 23: Die lokalen Fixationspfade des Bottom-up-Modells größer als 3 % und die zugeordneten lokalen Fixationspfade des Top-down-Modells für die drei untersuchten Szenarien.

Fixationspfade des Top-down-Modells (1>0>1 und 0>1>0) werden von dem Bottom-up-Modell in diesen zwei Szenarien nicht als wichtige Fixationspfade simuliert.

Die durchschnittliche Abweichung der lokalen Fixationspfade zwischen Bottom-up-Modell und Top-down-Modell liegt für die betrachteten häufigsten Fixationspfade bei 20,89 % (Szenario 1), 12,93 % (Szenario 2) und 15,44 % (Szenario 3).

Auf Grundlage der Fixationspfade der zwei kognitiven Modelle kann das simulierte Blickverhalten als unterscheidbar angenommen werden. Diese Annahme wird durch die statistische Analyse der Übergangshäufigkeiten verstärkt (siehe Abbildung 22, rechts). Es zeigt sich, dass das Blickverhalten des Bottom-up-Modells in Szenario 2 und Szenario 3 weniger statistisch abhängig ist, als das Blickverhalten des Top-down-Modells. In Szenario 1 ist die statistische Abhängigkeit für beide kognitiven Modelle ähnlich, auch wenn die lokalen Fixationspfade zeigen, dass es unterschiedliche Prioritäten in dem Blickverhalten für die Fixationspfade gibt. Werden die H_c -Werte für die zusammengefassten AOIs betrachtet (siehe Abbildung 22, links), ist dieser Unterschied für Szenario 1 und Szenario 2 nicht so deutlich zu erkennen. Durch das Zusammenlegen der AOIs scheint das Blickverhalten besser auf die drei AOIs verteilt zu sein, wodurch die statistische Abhängigkeit abnimmt.

Zusammenfassend kann anhand der lokalen Fixationspfade und der statistischen Analyse der Übergangshäufigkeiten die Hypothese 2 auf deskriptiver Ebene gestützt werden. Die unterschiedlichen theoretischen Annahmen über die Informationsaufnahme- und -verarbeitungsprozesse der zwei kognitiven Modelle führen zu unterscheidbarem Verhalten.

ERGEBNISSE ZUR HYPOTHESE 3 Für die Überprüfung der Annahme, dass das Bottom-up-Modell das empirische Blickverhalten besser abbilden kann als das Top-down-Modell, wurden die lokalen Fixationspfade mit zusammengefassten AOIs (siehe Abbildung 21), die statistische Analyse der Übergangshäufigkeiten (siehe Abbildung 22, rechts), die kumulierte Fixationsdauer der AOIs mit 3 beziehungsweise 5 AOIs (siehe Abbildung 24) und die Übergangsdichte (siehe Abbildung 25 und Kapitel 3.3.3) betrachtet.

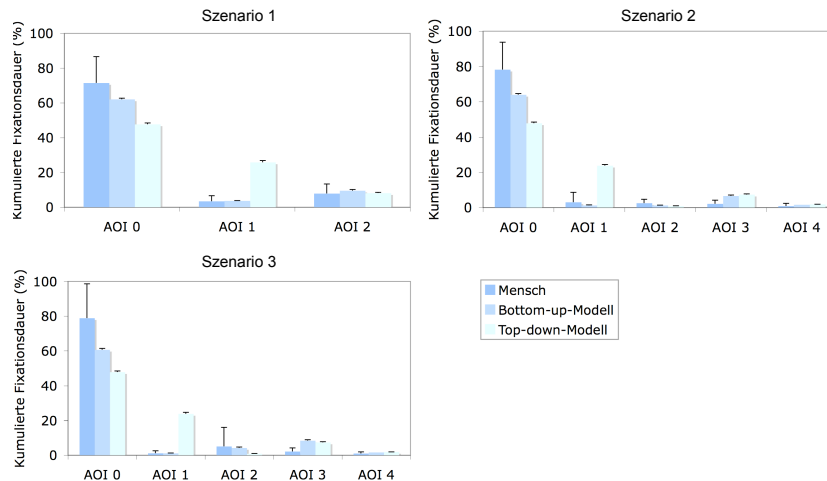


Abbildung 24: Kumulierte Fixationsdauern für die drei Versuchsgruppen in den drei untersuchten Szenarien.

In diesem Abschnitt erfolgt zuerst der Vergleich der Blickdaten des Bottom-up-Modells mit den empirisch erhobenen Blickdaten. Anschließend werden die Blickdaten des Top-down-Modells mit den Blickdaten der Versuchspersonen verglichen.

In Abbildung 21 ist zu sehen, dass alle Fixationspfade des Menschen in ähnlicher Tendenz durch das Bottom-up-Modell simuliert werden. Für den weiterführenden Vergleich der Fixationspfade wurden die 5 wichtigsten Fixationspfade des Menschen genutzt (siehe Kapitel 4.1.8). Das Bottom-up-Modell simuliert in Szenario 1 die gleichen Fixationspfade am häufigsten und in gleicher Rangreihenfolge wie in den empirischen Daten gefunden. In Szenario 2 und Szenario 3 werden die gleichen Fixationspfade am häufigsten simuliert. Die Rangreihenfolge unterscheidet sich in der Position des Fixationspfades $1 > 0 > 2$. Die durchschnittliche Abweichung der Fixationspfade zwischen den empirischen Daten und den Daten des Bottom-up-Modells liegt bei 8,11 % (alle Fixationspfade: 4,34 %). Die H_c -Werte der statistischen Analyse für die empirischen und die durch das Bottom-up-Modell simulierten Daten sind sehr ähnlich (siehe Abbildung 22, rechts). Die Abweichung der H_c -Werte beträgt durchschnittlich 5,65 %. Die kumulierten Fixationszeiten zeigen für alle drei Szenarien und die zwei Versuchsgruppen ähnliche Tendenzen. Die Reihenfolge der kumulierten Fixationsdauer ist wie folgend: Pegel (Mensch 76,19 %; Bottom-up-Modell: 62,30 %), Regelventil (Mensch 6,66 %; Bottom-up-Modell: 10,00 %) und Heizung (Mensch 3,21 %; Bottom-up-Modell: 3,09 %). Die durchschnittliche Ab-

weichung von Mensch und Bottom-up-Modell liegt bei 3,71 % (Szenario 1), 4,51 % (Szenario 2) und 5,17 % (Szenario 3). Die Übergangsdichte zeigt für das Bottom-up-Modell, dass die visuelle Suche von Szenario 3 über Szenario 2 bis zu Szenario 1 effizienter wird. Dieser Trend entspricht den empirischen Daten.

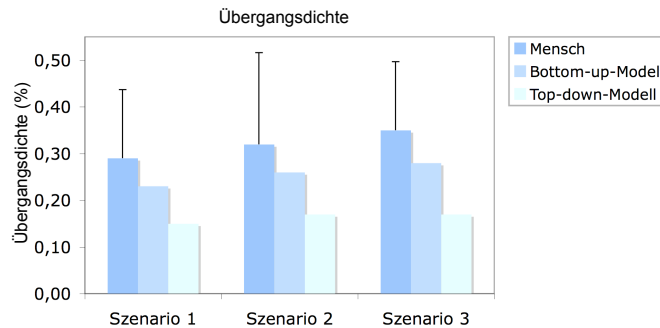


Abbildung 25: Mittlere Übergangsdichte für die drei Versuchsgruppen und die drei Szenarien. Grundlage für den Vergleich ist ein 10 x 10 Gitter, das über das Blickfeld gelegt wurde.

Das Top-down-Modell simuliert die Hälfte der menschlichen Fixationspfade. Von den fünf wichtigsten Fixationspfaden des Menschen werden vier von dem Top-down-Modell simuliert. Bei zwei Fixationspfaden ist die simulierte Häufigkeit ähnlich. Die Rangreihenfolge der lokalen Fixationspfade stimmt in keinem Szenario mit den empirischen Daten überein. Die durchschnittliche Abweichung zwischen Top-down-Modell und Empirie liegt bei 13,9 % (alle Fixationspfade: 8,96 %) im Vergleich zu dem Bottom-up-Modell deutlich schlechter. Die statistische Analyse der Übergangshäufigkeiten zeigt, dass die statistische Abhängigkeit im Vergleich zu den empirisch erhobenen Daten in Szenario 2 und Szenario 3 unterschiedlich und in Szenario 1 ähnlich ist (siehe Abbildung 22, rechts). Die Abweichung der H_c -Werte beträgt durchschnittlich 26,41 %. Die Reihenfolge der kumulierten Fixationsdauern ist unterschiedlich zu denen der empirischen Daten (Pegel: 47,77 %; Heizung: 25,72 %; Regelventil: 8,22 %). Die Abweichung zwischen Top-down-Modell und empirischen Daten liegt für die kumulierten Fixationsdauern durchschnittlich bei 12,68 % (Szenario 1), 8,03 % (Szenario 2) und 8,02 % (Szenario 3). Auffällig ist, dass die kumulierte Fixationsdauer für alle drei Szenarien für die AOIs der Heizung viel höher als die der empirischen Daten ist. Ein Grund ist in der Regelstruktur des Top-down-Modells zu finden, wodurch die Heizung kontinuierlich überwacht wird. Die Übergangsdichte zeigt, dass die visuelle Suche in Szenario 2 und Szenario 3 ähnlich und zu Szenario 1 unterscheidbar ist. Dieses Verhalten ist bei dem Menschen nicht zu finden.

Das Bottom-up-Modell simuliert tendenziell das menschliche Blickverhalten mit einer durchschnittlichen Abweichung aller lokalen Fixationspfade und der kumulierten Fixationsdauern kleiner als 5 %. Das Top-down-Modell simuliert ein stark regelbasiertes Blickverhalten, das bei dem Menschen in ähnlicher Weise nicht zu finden ist. Die Abweichungen der kumulierten Fixationshäufigkeiten und der lokalen Fixationspfade des Top-down-Modells sind für alle Szenarien doppelt

so hoch. Zusammenfassend zeigt sich, dass die Befunde den Vorhersagen von Hypothese 3 entsprechen. Das Bottom-up-Modell simuliert das menschliche Verhalten besser als das Top-down-Modell.

4.1.9 Diskussion

Die Ergebnisse der Untersuchung zeigen, dass die Analysen und Vergleiche der Blickbewegungsdaten eine genauere und weiter führende Betrachtung von Leistungsdaten ermöglichen als die alleinige Betrachtung der Performanzdaten. Durch die Analyse der Performanzdaten konnte Hypothese 1 (unterschiedliche Schnittstellen induzieren unterschiedliche Leistungsdaten) nicht gestützt werden. Die Analyse der Blickbewegungsdaten ermöglichte eine Differenzierung dieses Ergebnisses. Die unterschiedlichen Schnittstellen in Szenario 2 und Szenario 3 induzieren ein ähnliches Blickverhalten, das von dem durch Schnittstelle 1 induzierten Blickverhalten unterscheidbar ist. Die von Hypothese 2 (die unterschiedlichen theoretischen Annahmen der kognitiven Modelle induzieren unterscheidbares Verhalten) vorhergesagten Effekte traten bei beiden Arten von Leistungsdaten auf. Die dritte Hypothese (das Verhalten des Bottom-up-Modells ist näher an dem menschlichen Verhalten als das Verhalten des Top-down-Modells) wird durch die Performanzdaten teilweise unterstützt. Die Analyse der Blickbewegungsdaten konnte die dritte Hypothese stützen.

Die Prozessaufgabe und die unterschiedlichen Schnittstellen waren in der Untersuchung leicht gewählt, um die Aspekte der Informationsaufnahme- und -verarbeitungsprozesse isoliert betrachten zu können. Das Steuerungswissen konnte daher durch die Versuchspersonen gut gelernt und auf wenige Regeln verkürzt in die kognitiven Modelle integriert werden. Die unterschiedlichen Schnittstellen sind einfache Variationen der räumlichen Gestaltung. Vorstellbar ist, eine weitere Studie mit komplexeren Schnittstellen oder einer komplexeren Steuerungs- und Kontrollaufgabe durchzuführen, die eine größere Wissensbasis erfordert. Beispielsweise könnten geänderte Bedienelemente bereitgestellt werden, die den Abfluss der Flüssigkeit aus dem Tank oder die Heizleistung der Heizung in prozentualen Schritten regeln oder die Prozessvariablen (z. B. den Pegelstand) digital darstellen.

In dieser Studie wurde eine korrekt funktionierende Prozesssimulation genutzt, um das zu steuernde und zu kontrollierende System zu simulieren. Eine interessante Variation stellt die fehlerhafte Simulation mit nicht aufgabenspezifischem Verhalten der Schnittstellenelemente dar (z. B. fehlerhafte Anzeige der aktuellen Zustände oder Ausfall der Zustandsanzeigen).

In der Studie wurden Abweichungen der Modelle von dem beobachteten Verhalten des Menschen festgestellt (Bottom-up-Modell: durchschnittlich 5 %; Top-down-Modell: durchschnittlich 10 %). Die erfassten Daten der kognitiven Modelle liegen innerhalb des möglichen Abweichungspotentials von 20 % für die Simulation in frühen Phasen der Systemgestaltung, die darauf zurückzuführen ist, dass Prozesswerte für die Simulation in diesen Phasen mit einer Ungenauigkeit von bis zu 20 % bereitstehen (Card et al. 1983; Urbas, Schulze-Kissing & Leuchter 2005).

Für die Abweichung sind verschiedene Gründe vorstellbar. In der gewählten Steuerungs- und Kontrollaufgabe wurde die Flüssigkeitsanzeige aufgrund einer Vorstudie für die kognitive Modellierung in fünf feste Bereiche eingeteilt (siehe Kapitel 4.1.2). Eine mögliche Variation der Modelle stellt die Veränderung der Größe dieser Bereiche dar und die Betrachtung und der Vergleich der Leistungsdaten. Eine Verbesserung der Blickbewegungsdaten könnte durch eine bessere Abbildung der Grenzbereiche erreicht werden.

In beiden kognitiven Modellen sind für die Informationsaufnahme und die Interaktion einfache Strategien integriert. Die Modelle basieren auf zwei theoretischen Annahmen über die Informationsaufnahme- und -verarbeitungsprozesse. Neben den ausgewählten Ansätzen und integrierten Strategien gibt es weitere Theorien der menschlichen Informationsaufnahme- und -verarbeitungsprozesse (wie z. B. Salienzenkarten; Itti & Koch 2001) sowie Verhaltensweisen (z. B. visuelle Update-Circle der abgelegten Informationen im Arbeitsgedächtnis; Yarbus 1967), die in den betrachteten kognitiven Modellen und der genutzten kognitiven Architektur noch nicht integriert sind. Eine Integration dieser Ansätze in kognitive Architekturen und Modelle kann eine Verbesserung der erfassten Leistungsdaten im Bezug zur Passung an empirische Daten ermöglichen. Die Anpassung des visuellen Moduls von ACT-R an empirische Ergebnisse stellt somit eine notwendige Erweiterung dar. Zum Beispiel kann das visuelle Modul um visuelle Suchalgorithmen erweitert werden, die bekannte Heuristiken und Strategien formalisieren und integrieren (wie z. B. Bayes'sche Heuristiken).

Vorstellbar ist die Integration neuer Strategien in die kognitiven Modelle, die, ausgehend von der Situation oder den deklarativen Gedächtnisinformationen des kognitiven Modells, die kognitiven Prozesse verändern und die Abweichung der simulierten Leistungsdaten zu den empirischen Daten verringern können. Anstatt der implementierten visuellen Suche der Bedienelemente könnte ein Prozess implementiert werden, der ein deklaratives Gedächtniselement der Position abrufen und die Ausrichtung der visuellen Aufmerksamkeit auf die abgerufene Position richtet.

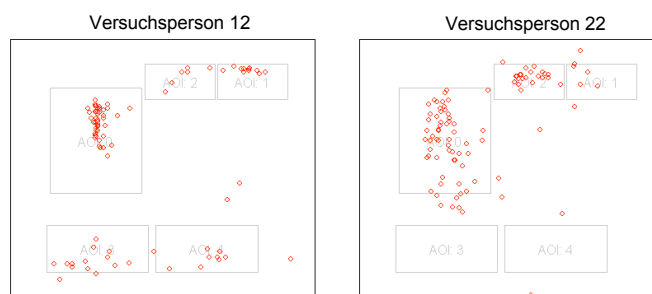


Abbildung 26: Die Blickverteilung von zwei unterschiedlichen Versuchspersonen während der Interaktion mit der Mensch-Maschine-Schnittstelle in Szenario 3.

Ein weiterer Grund für die Abweichungen der simulierten Blickbewegungsdaten von den empirischen Blickbewegungsdaten könnte darin zu finden sein, dass es innerhalb der Versuchspersonen unter-

scheidbare homogene Gruppen gibt, die eine Klasse von Verhalten, das heißt eine bestimmte Strategie, nutzen, um die Prozessaufgabe zu bearbeiten. In der vorliegenden Untersuchung wurde davon ausgegangen, dass die Gruppe der Versuchspersonen homogen ist. Eine Betrachtung der Blickbewegungsdaten zeigt, dass dies nicht für alle Versuchspersonen vorausgesetzt werden darf (siehe Abbildung 26). Es ist zu erkennen, dass die Interaktion in einem Szenario für die zwei ausgewählten Versuchspersonen unterschiedliche Blickverteilungen induziert. Eine Clusteranalyse könnte helfen, einzelne homogene Gruppen zu identifizieren. Anschließend wäre es möglich, die Simulationsdaten der kognitiven Modelle durch den Vergleich einer dieser Gruppen zuzuordnen, um daraus Aussagen über die empirische Abbildbarkeit abzuleiten. Ebenfalls können aus den homogenen Gruppen Bearbeitungsstrategien abgeleitet und anhand dieser weitere kognitive Modelle erstellt werden.

Ebenfalls ist bei dem Vergleich von simulierten und empirisch erfassten Blickdaten zu berücksichtigen, dass es beim Menschen eine partielle Unabhängigkeit von kognitiven Prozessen und Blickbewegungen gibt (Anderson et al. 2004). Daraus ist zu schließen, dass bei der Betrachtung der empirischen Daten ein Fixationsort nicht mit der Aufmerksamkeit des kognitiven Systems übereinstimmen muss. Bei den simulierten Daten kann aber davon ausgegangen werden, dass die Aufmerksamkeit des kognitiven Systems auf den Fixationsort gerichtet ist. Somit können Unterschiede zwischen kognitivem Modell und Mensch erklärt werden.

Die Hauptstudie zeigt einen Unterschied zwischen den zwei betrachteten kognitiven Modellen. Es ist festzustellen, dass die zwei wichtigsten Fixationspfade des Top-down-Modells ($0 > 1 > 0$ und $1 > 0 > 1$) in den drei Szenarien nicht oder nicht mit ähnlicher Häufigkeit von dem Bottom-up-Modell simuliert werden. Aus der Implementierung der Modelle ist dieser Unterschied zu erklären. Das Bottom-up-Modell wechselt nur zwischen AOIs, wenn der Wechsel für die Bedienung des Interaktionselements wichtig ist. Somit sind die Wechsel zwischen Regelventil (AOI 2) und Pegel (AOI 0) häufiger zu finden als die Wechsel zwischen Heizung (AOI 1) und Pegel (AOI 0). Das Top-down-Modell überprüft fortlaufend den Zustand der Heizung und des Pegels. Der Wechsel zwischen Heizung (AOI 1) und Pegel (AOI 0) findet dementsprechend sehr oft statt. Dieses Verhalten ist in den Daten zu finden.

In der Hauptstudie wurde gezeigt, dass das Bottom-up-Modell das menschliche Verhalten besser simuliert als das Top-down-Modell. Auf Grundlage der Blickbewegungsdaten des Bottom-up-Modells können die drei Schnittstellen bewertet werden. Anhand der statistischen Analyse der Übergangshäufigkeiten und der Fixationspfade zeigt sich, dass die Schnittstelle in Szenario 1 zu den Schnittstellen von Szenario 2 und Szenario 3 unterscheidbar ist. Wird zusätzlich die Übergangsdichte betrachtet, zeigt sich, dass die Schnittstelle in Szenario 1 die effektivste visuelle Suche ermöglicht und die Schnittstelle in Szenario 3 die visuelle Suche mit dem längsten Suchpfad induziert. Als wichtige Blickgebiete zeigten sich die Flüssigkeitsanzeige und das Regelventil. Diese Ergebnisse werden von den empirischen Daten gestützt. Die Schnittstelle in Szenario 1 hat zusammengesetzte Interaktions- und Rückmeldeelemente, wodurch die Kosten der visuellen Suche minimal sind. Die Integration der Interaktion und der Rückmeldung in einem Bildelement scheint im Bezug zu der visuellen Suche

besser zu sein und induziert den kürzesten Suchpfad. Hinsichtlich der Performanz ist dieser Unterschied nicht zu finden. In komplexen und dynamischen Mensch-Maschine-Schnittstellen stellt die effiziente visuelle Suche einen wichtigen Faktor für die Gestaltung einer Schnittstelle dar. Die empirischen Daten und die Daten des Bottom-up-Modells zeigen, dass die Variante der Mensch-Maschine-Schnittstelle in Szenario 1 am geeignetsten ist.

Es zeigt sich, dass die Analyse von feinen Leistungsdaten einen detaillierten Einblick in das Modellverhalten ermöglicht und eine gute Grundlage für die Diskussion von Modellunterschieden bietet. SimTrA kann somit für die vergleichende Analyse von kognitiven Modellen eingesetzt werden.

Die Hauptuntersuchung wurde durchgeführt, um die Anwendung von SimTrA für die Gestaltung von Schnittstellen in frühen Phasen der Systemgestaltung zu überprüfen. Die bereitgestellten Daten ermöglichen die Analyse und den Vergleich der untersuchten Schnittstellen und die Bewertung der drei experimentalpsychologischen Fragestellungen.

In SimTrA konnten der Analyse- und Vergleichsprozess und die Verarbeitung der Blickbewegungsdaten effizient und effektiv durchgeführt werden. Die integrierte Analyse, Bereinigung und der Vergleich der Blickbewegungsdaten ermöglichte nach wenigen Schritten die Betrachtung von Ergebnissen der Analysen und Vergleiche. Eine durchschnittliche Analyse dauerte fünf Minuten. Zeitlich und algorithmisch stellt die Anwendung von SimTrA gegenüber der aufwendigen Analyse von Hand einen großen Vorteil dar. Zusätzlich werden die Daten in verwertbarer Form zur Verfügung gestellt und sind in externe Programme überführbar und weiterverwendbar. Die deskriptive Analyse ist für eine Betrachtung der simulierten und empirischen Blickbewegungsdaten auf diesem Betrachtungslevel gut geeignet, um die simulierten Blickbewegungsdaten und die zugrunde liegenden Mensch-Maschine-Schnittstellen bewerten zu können. Die Anwendung von inferenzstatistischen Analysen ist vorstellbar und sollte in weiteren Entwicklungen von SimTrA berücksichtigt werden. Die Integration und Verarbeitung der Daten in SimTrA fand den Arbeitsprozess begleitend statt und konnte an Unterbrechungen und erneute Analysen der Blickbewegungsdaten individuell und situationsabhängig angepasst werden. Weiterführende Bewertungen waren nicht Teil der Untersuchung (z. B. motorische Aktivitäten oder Gedächtnisabrufe). Die Erweiterung um Analysen und Vergleiche ist aufgrund des Komponentenmodells von SimTrA aber möglich.

Zusammenfassend zeigt die Hauptstudie, dass SimTrA ein Werkzeug ist, das den Prozess der Analyse und des Vergleichs von empirischen und künstlichen Leistungsdaten unterstützt und den Analysten ein einfach zu benutzendes Methodenrepertoire an die Hand gibt.

4.2 WEITERE EXPERIMENTELLE STUDIEN

Das Werkzeug SimTrA wurde in verschiedenen Arbeitsgruppen und Projekten für die Analyse und den Vergleich von empirischen und simulierten Blickbewegungsdaten eingesetzt. Im Folgenden werden drei experimentelle Studien exemplarisch aufgeführt und kurz beschrieben.

4.2.1 Blocks World

In dieser von den CogWorks Laboratories der Rensselaer University in Troy, USA, durchgeführten Studie wurde die Entwicklung von Strategien zur Aufgabenlösung experimentell überprüft. Blocks World ist eine einfache Aufgabe, deren Ziel es ist, aus einem *Target Window* ein Muster von farbigen Blöcken in das *Workspace Window* zu übertragen, dazu müssen die farbigen Blöcke des *Resource Window* genutzt werden (siehe Abbildung 27; Gray, Schoelles & Sims 2004). Das *Target Window* wird verdeckt dargestellt und kann durch das Bewegen der Maus in das *Target Window* aufgedeckt werden. Verlässt die Maus das Fenster, wird es wieder verdeckt. Ausgehend von der Hypothese, dass kleine Variationen der Aufgabengestaltung zu einer Entscheidung zwischen Interaktions-intensiven und Gedächtnis-intensiven Strategien des Menschen führen, wurde in dem Aufgabenszenario die Dauer von der Aktivierung bis zur Aufdeckung des *Target Window* von 0 bis 3200 Millisekunden (*lockout time* – Zeit der Verzögerung) variiert.

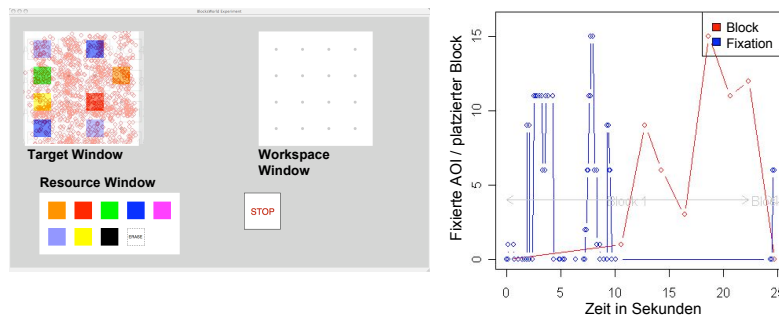


Abbildung 27: Die Aufgabenumgebung von Blocks World (nach Gray et al. 2004) mit Blickbewegungsdaten einer Versuchsperson auf das *Target Window* (rechts). Ansicht aus SimTrA der fixierten AOIs und der nachfolgend platzierten farbigen Blöcke für eine Versuchsperson (links).

In vorangegangenen Analysen wurden die Performanzdaten analysiert und mit Daten von Simulationsmodellen verglichen (siehe dazu Gray et al. 2004). Mit SimTrA wurden die aufgezeichneten Blickbewegungsdaten bei der Aufgabenanalyse analysiert, um die Fixationsanzahl, Fixationsdauer und die Fixationsreihenfolge nach dem Aufdecken des *Target Window* zu analysieren (siehe Abbildung 27). Das Ziel der Blickbewegungsdatenanalyse war, festzustellen, ob die unterschiedlichen *lockout times* zu unterscheidbaren Strategien in den Informationsaufnahme Prozessen führen. Für die Analyse wurde ein Importalgorithmus für die empirischen Blickbewegungsdaten entwickelt. Zusätzlich wurde SimTrA erweitert, so dass bei dem Import der Blickbewegungsdaten automatisch die jeweiligen Zeitintervalle für die zu untersuchenden Abschnitte aus den Rohdaten gelesen und für jede Versuchsperson individuell abgelegt werden konnten. Zusätzlich wurde ein Algorithmus entwickelt, der die fixierten Blöcke und deren Reihenfolge mit den tatsächlich platzierten Blöcken vergleicht und graphisch aufbereitet (siehe Abbildung 27, rechts). In der Abbildung

sind die Fixationen der AOIs (Fixation) und die tatsächlich platzierten Blöcke (Block) für einen Versuchsdurchlauf mit einer Versuchsperson dargestellt. Eine Fixation auf die AOI o bedeutet, dass sie erfasst wurde, aber keiner relevanten AOI zugeordnet werden konnte. Es ist zu sehen, dass bei dieser Versuchsperson nicht alle Blöcke (AOIs) fixiert wurden, die anschließend platziert wurden, und dass Blöcke fixiert wurden, die für die Aufgabenlösung nicht relevant waren. Dieser Effekt konnte bei allen Versuchspersonen festgestellt werden. Eine mögliche Ursache liegt in der räumlichen Nähe der farbigen Blöcke. Durch die periphere Wahrnehmung können beispielsweise mit einem Blick mehr Informationen aufgenommen werden (wie z. B. Farbe und Platzierung), ohne dass jedes Objekt einzeln betrachtet werden muss. Aus diesem Untersuchungsergebnis wurde abgeleitet, dass in nachfolgenden Untersuchungen schwerer zu unterscheidende Objekte genutzt werden sollten, um das Blickverhalten besser analysieren zu können (z. B. komplexe geometrische Figuren, chinesische Schriftzeichen).

4.2.2 Vernetztes Fahren

In dem Verbundprojekt *Vernetztes Fahren* der Human-Factors-Consult GmbH, der Humboldt Universität zu Berlin und der Technischen Universität Berlin wurden Fragen zur Mensch-Maschine-Interaktion beim vernetzten Fahren im Bezug zu Bedienbarkeit, Akzeptanz und Sicherheit betrachtet (Vernetztes Fahren 2007). Unter vernetztem Fahren wird jegliche Interaktion eines Kraftfahrzeuges mit anderen Kraftfahrzeugen, Personen oder Einrichtungen verstanden. Die Informationen können über Systeme für vernetzte Dienste durch den Kraftfahrzeugführer eingegeben, empfangen und automatisch verteilt werden.

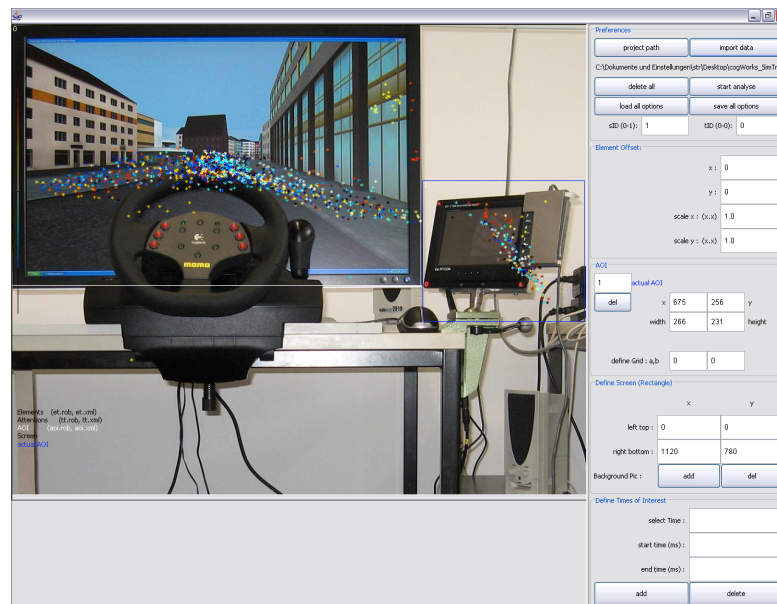


Abbildung 28: Ansicht der Blickbewegungsdaten einer Versuchsperson in SimTra mit zwei AOIs (für Fahraufgabe und Nebenaufgabe).

Die Bewertung von Bedienkonzepten von prototypischen Mensch-Maschine-Schnittstellen für vernetzte Dienste stand im Mittelpunkt eines Teilprojektes, das am Zentrum Mensch-Maschine-Systeme der Technischen Universität Berlin angesiedelt war. In dem Teilprojekt wurden kognitive Modelle für die Benutzung der unterschiedlichen Mensch-Maschine-Schnittstellen für die vernetzten Dienste erstellt. Diese ermöglichten in Kombination mit visuellen, kognitiven und motorischen Ressourcenprofilen die Bewertung der Schnittstellenvarianten und deren Bedienkonzepte. Die Frage war, wie viele und welche Ablenkungen durch die Nebenaufgabe neben der Fahraufgabe an den Kraftfahrzeugführer gestellt werden können, ohne die Fahrsicherheit zu beeinträchtigen. Die kognitiven Modelle wurden mit der Entwicklungs- und Simulationsumgebung mtGOMS (*MultiTasking-GOMS*; Urbas, Heinath & Leuchter 2007) entwickelt und stellten die Grundlage für die simulationsgestützte Bewertung der Schnittstellen dar. Im Kern der Entwicklungs- und Simulationsumgebung mtGOMS steht ein Algorithmus zur Vorhersage der gegenseitigen Beeinflussung von der zu gestaltenden Nebenaufgabe (hier die Bedienung des vernetzten Dienstes) und einer gegebenen Hauptaufgabe (hier das sichere Fahren). Der Algorithmus benötigt eine geeignete Darstellung der Teilaufgaben und Ressourcenprofile zur generischen Darstellung der Hauptaufgabe und Aufgabenmodelle der Nebenaufgabe. Für die erstellten kognitiven Modelle wurden visuelle, kognitive und motorische Ressourcenprofile erhoben.

SimTrA wurde eingesetzt, um aus empirischen Blickbewegungsdaten bei der Fahrzeugführung in einer Fahrsimulation visuelle Ressourcenprofile für verschiedene Fahrsituationen bei gleichzeitiger Bedienung der Nebenaufgabe zu erstellen. Die Blickbewegungsdaten wurden in SimTrA importiert und zwei AOIs festgelegt, um in der Analyse der Daten bei der Unterscheidung zwischen der Umgebungssimulation und der Nebenaufgabe unterscheiden zu können (siehe Abbildung 28). Anschließend wurden die Blickbewegungsdaten mit SimTrA verarbeitet. Die Reihenfolge der Fixationen und deren Dauer wurden genutzt, um visuelle Ressourcenprofile abzuleiten. Die mit SimTrA aufbereiteten Blickbewegungsdaten konnten erfolgreich genutzt werden, um die visuellen Ressourcenprofile für die Modellierung mit dem Werkzeug mtGOMS zu erstellen.

4.2.3 Hierarchical Task Analysis Mapper

In diesem Projekt wird das Ziel verfolgt, den Entwicklungsprozess kognitiver Modelle zu unterstützen und zu beschleunigen, um den Aufwand der Modellbildung zu verringern (Heinath & Urbas 2007). Darauf aufbauend, wurde der *Hierarchical Task Analysis Mapper*-Ansatz (HTAmap-Ansatz) zur automatisierten Erstellung von ACT-R Modellen auf der Grundlage von hierarchischen Aufgabenanalysen und Aktivitätsmustern entwickelt. Die automatisiert erstellten HTAmap-Modelle beinhalten Produktionsregeln und Steuerungsinformationen, die für die Generierung und den Ablauf der Modelle genutzt werden, aber aus kognitionspsychologischer Sicht nicht immer notwendig sind. Ein wichtiger Aspekt stellt daher die Überprüfung der Güte der Model-

le dar, um sicherzustellen, dass die zusätzlichen Informationen die Simulationsgüte nicht beeinträchtigen.

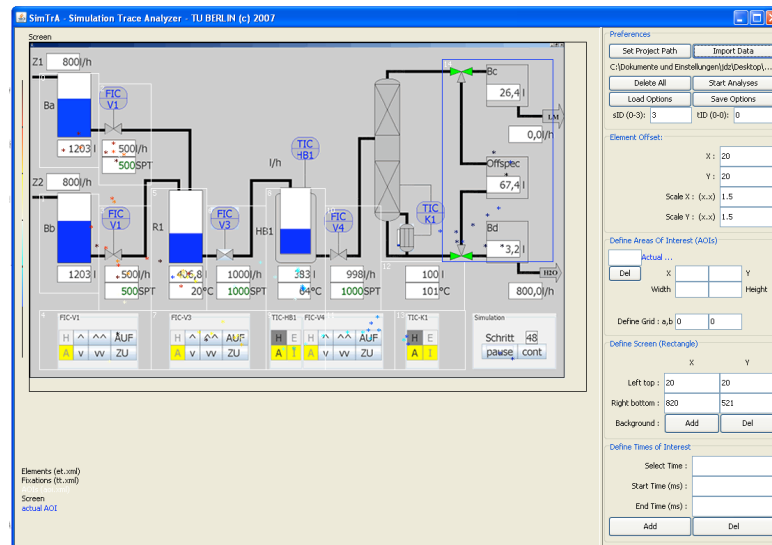


Abbildung 29: Ansicht der Versuchsumgebung mit Blickbewegungen einer Versuchsperson bei der Interaktion mit der Abwasseraufbereitungsanlage in SimTrA. Eingezeichnet sind die AOIs für die 4 verschiedenen Teilprozesse der Aufgabe: Anlieferung, Homogenisierung, Trennung und Produktlager.

In einer experimentellen Studie mit Versuchspersonen ($n = 10$) und einem HTAmap-Modell wurden die Leistungsdaten bei der Interaktion mit einer komplexen und dynamischen Mensch-Maschine-Schnittstelle erfasst und verglichen. Die Aufgabe stellt den Anfahrprozess einer Prozessanlage zur Abwasseraufbereitung dar (d. h. den Startprozess bis zur erfolgreichen Trennung von sauberem Wasser und Verunreinigung). Das Ziel der Abwasseraufbereitungsanlage ist es, Abwässer, die in industriellen Reinigungsprozessen anfallen, aufzubereiten. Dabei werden Gemische aus ca. 60 % Wasser und 40 % Lösungsmittel getrennt. Die Abwasseraufbereitungsanlage besteht aus vier getrennt zu startenden Teilprozessen: Anlieferung, Homogenisierung, Trennung und Produktlager (siehe Abbildung 29). In den einzelnen Teilprozessen ist die Steuerung von Regelventilen und Heizungen notwendig.

Für die Implementierung des HTAmap-Modells wurden in einer Vorstudie Aktivitätsmuster und Interaktionsregeln bestimmt. Auf der Grundlage eines starren Regelkreises wurden die Aktivitätsmuster in einem Modell zusammengefasst (Top-down-Verhalten).

Das entwickelte HTAmap-Modell wurde anhand von Performanzdaten (d. h. die Art, Reihenfolge und Anzahl der Interaktionen) und Blickbewegungsdaten bewertet. Für die Analyse der Blickbewegungsdaten wurden die simulierten und die empirischen Daten in SimTrA importiert und anhand der Transitionshäufigkeit analysiert und verglichen.

Für die Performanz zeigt sich, dass die Leistungsdaten des automatisiert erstellten HTAmap-Modells und der empirischen Daten ähnlich

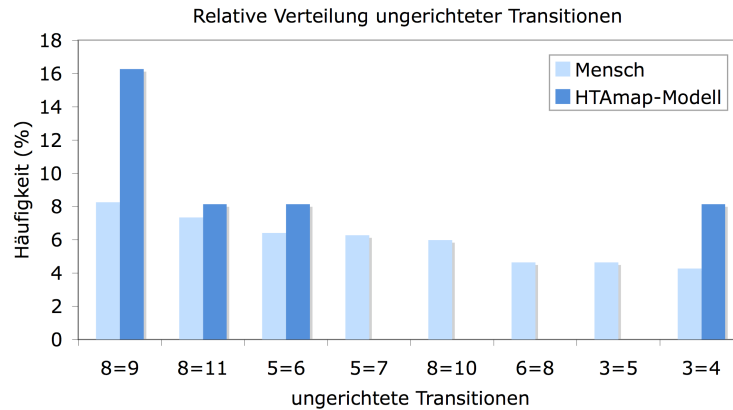


Abbildung 30: Auszug aus den Analysen der Blickbewegungsdaten der Untersuchung mit einem automatisiert erstellten kognitivem Modell in HTAmap und empirischen Daten. Dargestellt sind die ungerichteten Transitionen (größer als 4 % für die empirischen Daten).

sind. Die Analyse und der Vergleich der ungerichteten Transitionen (Häufigkeit größer als 4 % für die empirischen Daten) der Blickbewegungsdaten zeigt, dass Teilaspekte der empirischen Blickbewegungsdaten abgebildet werden können (siehe Abbildung 30). Die drei häufigsten Transitionen der Menschen und eine erhöhte Häufigkeit für die wichtigste Transition wurden von dem HTAmap-Modell simuliert. Das gesamte Spektrum der empirischen Daten wird durch das HTAmap-Modell nicht simuliert. Eine ausführliche Bewertung der Daten ist zu finden in Heinath (in Vorbereitung).

Die Ergebnisse dieser Untersuchung finden sich in ähnlicher Form für den Vergleich des Top-down-Modells mit den menschlichen Leistungsdaten bei der experimentellen Hauptstudie (siehe Kapitel 4.1). Die Analyse und der Vergleich der simulierten und empirischen Blickbewegungsdaten in SimTrA konnte helfen, die Güte der HTAmap-Modelle abzuschätzen und zu bewerten. Auf der Ebene der Performanzdaten bilden die HTAmap-Modelle das menschliche Verhalten identisch ab. Es zeigt sich jedoch, dass auf der Ebene der Blickbewegungen Aspekte fehlen, die in weiteren Studien im kognitiven Aufgabenmodell ergänzt werden müssen.

DISKUSSION

In diesem Kapitel wird die in dieser Arbeit untersuchte Fragestellung kritisch diskutiert. Die Bewertung von Mensch-Maschine-Systemen auf Basis von simulierten Leistungsdaten wird betrachtet und deren heutige Möglichkeiten aufgezeigt. Anschließend wird das entwickelte Werkzeug SimTrA bewertet und im Kontext der Arbeit diskutiert. Den Abschluss bildet die Beschreibung von möglichen Anwendungsgebieten des Werkzeuges SimTrA.

5.1 DIE BEWERTUNG VON MENSCH-MASCHINE-SYSTEMEN DURCH KOGNITIVE BENUTZERMODELLE

In dieser Arbeit wurde die Methode und das Werkzeug SimTrA entwickelt, um den Analyse- und Vergleichsprozess von simulierten und empirischen Leistungsdaten zu unterstützen. Die Analysen und die Vergleiche wurden am Beispiel von Blickbewegungsdaten durchgeführt. Die experimentelle Hauptstudie (siehe Kapitel 4.1) hat deutlich gezeigt, dass das entwickelte Werkzeug SimTrA geeignet ist, bei der Entwicklung und Bewertung von Mensch-Maschine-Systemen den Entwickler zu unterstützen. Die Analyse und der Vergleich von simulierten und empirischen Blickbewegungsdaten ist ohne besondere psychologische Kenntnisse möglich. SimTrA ermöglicht durch das bereitgestellte Methodenrepertoire die explorative und hypothesengeleitete Untersuchung der erfassten Daten.

Die Ergebnisse dieser Arbeit zeigen, dass kognitive Modellierungsmethoden für die Bewertung von Mensch-Maschine-Systemen eingesetzt werden können. Im Bereich der Gestaltung von komplexen und dynamischen Mensch-Maschine-Systemen stellt die feinere Betrachtungsebene einen wichtigen Aspekt dar. Heutige Systeme integrieren mehr und mehr Aufgaben und Schnittstellen, die durch den Benutzer kontrolliert und gesteuert werden müssen. Die Schnittstellen für komplexe und dynamische Mensch-Maschine-Systeme müssen dementsprechend bestmöglich an die menschlichen kognitiven Prozesse angepasst werden. Die reine Betrachtung der Performanz für die Bewertung der Gestaltung einer Schnittstelle kann daher nicht ausreichend sein. Die vorgeschlagene Methode dieser Arbeit erweitert die Bewertung um visuelle Aspekte, die für komplexe und dynamische Systeme von besonderer Bedeutung sind. Es wurde gezeigt, dass die Betrachtungsebene der feinen Leistungsdaten gegenüber der Betrachtungsebene der gröberen Leistungsdaten einen Vorteil darstellt und eine genauere Betrachtung des Simulationsverhaltens erlaubt. Eine detaillierte Bewertung von Mensch-Maschine-Systemen auf der Grundlage von simulierten Leistungsdaten wird ermöglicht.

Die experimentelle Hauptstudie zeigt, dass die kognitiven Architekturen und die kognitiven Modelle im visuellen Bereich nicht das gesamte Spektrum und die Vielfalt des menschlichen Verhaltens abbil-

den. SimTrA konnte genutzt werden, um Schwachstellen der kognitiven Modelle aufzuzeigen. Die Analyseergebnisse konnten Tendenzen aufzeigen, die im Vergleich zu empirischen Daten ebenfalls zu finden waren. SimTrA ermöglichte durch die Analyse- und Vergleichsmöglichkeiten die Ableitung von potentiellen Erweiterungen der visuellen Komponente von ACT-R und der untersuchten kognitiven Modelle (z. B. die Integration von Update-Circle oder von Salienzkarten). Erweiterungen der kognitiven Architekturen und der kognitiven Modelle werden zukünftig die Simulationsgüte der kognitiven Modelle erhöhen und eine genauere Vorhersage der Interaktionsprozesse ermöglichen. Die Simulationsgüte der Modelle wurde in der Hauptstudie als hinreichend eingestuft. Es bleibt zu prüfen, ob durch verbesserte Modelle und die Integration von neuen kognitionspsychologischen Theorien zur Informationsaufnahme und -verarbeitung die Modellpassung zu empirischen Daten verbessert werden kann.

Als Baustein für die kognitive Modellbildung in Entwicklungsprozessen von Mensch-Maschine-Systemen trägt SimTrA das Potential in sich, das Wissen der Psychologie in den Arbeitsprozess zu integrieren. Die formale Abbildung der kognitiven Prozesse stellt eine Schnittstelle der beteiligten Fachrichtungen dar. SimTrA leistet einen Beitrag für die zukünftige Berücksichtigung kognitiver Modellierungsmethoden für die Bewertung von Mensch-Maschine-Systemen.

Mit heutigen kognitiven Modellierungsmethoden können fertigungs- und regelbasierte Tätigkeiten gut modelliert und simuliert werden. Weitere Forschung ist im Bereich der wissensbasierten Tätigkeiten und Handlungen anzusiedeln, um diese beispielsweise in Form von mentalen Modellen und kausalen Wissensbasen in kognitive Benutzermodelle zu integrieren. Dieser Aspekt wird im Hinblick auf die zunehmende Automatisierung von Mensch-Maschine-Systemen in der Entwicklung von wachsender Bedeutung sein. Es bleibt daher zu prüfen, wie Konstrukte für das wissensbasierte Handeln in der Kognitionspsychologie formalisiert und in kognitive Architekturen integriert werden können. Eine Anforderung an die Entwickler stellt dabei die durchgängige Möglichkeit der Beobachtung und der Analyse der beteiligten Prozesse dar. Das Ziel muss es sein, transparente kognitive Systeme zu entwickeln, die menschliche kognitive Prozesse offenlegen und bei komplexen Aspekten und Betrachtungsweisen der menschlichen Kognition keine weitere Blackbox darstellen, sondern die detaillierte Untersuchung der kognitiven Prozesse erlauben. Durch die Integration von Strukturen für das wissensbasierte Handeln in kognitive Architekturen wird eine große Bandbreite der menschlichen Kognition abgebildet. Bei dem Einsatz kognitiver Benutzermodelle für die Vorhersage von Interaktionsverhalten stellt dies einen wichtigen Entwicklung dar, die das Anwendungsgebiet erweitert und die Untersuchung von vielen Aspekten der Schnittstellengestaltung ermöglicht, die heute nur sehr schwer zu analysieren sind.

5.2 DAS WERKZEUG SIMTRA

SimTrA wurde für die Analyse und den Vergleich von simulierten und empirischen Blickbewegungsdaten implementiert. Zu diesem Zweck

können die Blickbewegungsdaten um Informationen angereichert werden (z. B. AOIs und Zeitintervalle). SimTrA ermöglicht die Bewertung von visuellen Suchprozessen, der Informationsaufnahme und -verarbeitung bei der Interaktion und der örtlich-räumlichen Gestaltung von Mensch-Maschine-Systemen. Es konnte gezeigt werden, dass SimTrA die Datenverarbeitung unterstützt und ein geeignetes und erweitertes Methodenrepertoire bereitstellt, das die Bewertung von Mensch-Maschine-Schnittstellen auf Grundlage simulierter Blickbewegungen ermöglicht. Der Prozess der Analysen und der Vergleiche wird prozessbezogen und prozessbegleitend unterstützt.

SimTrA kann als eine Methode betrachtet werden, die den Einsatz des Goodness-of-fit-Methodenrepertoires aufgreift. Zum einen wird die schnelle Auswertung der Daten und die Betrachtung der Daten auf der graphischen Ebene ermöglicht. Zum anderen werden Methoden und Daten für die weiterführende Analyse bereitgestellt. Modelldaten und empirische Daten können daher bei einer ersten Betrachtung bewertet und anschließend bereinigt weiterverarbeitet werden.

Das konzeptionelle Datenformat von SimTrA ermöglicht die Integration von beliebigen Blickbewegungsdaten durch den Import. Dabei können die Blickbewegungsdaten in das Datenformat von SimTrA transformiert werden. Somit sind die Analyse und der Vergleich von Blickbewegungsdaten verschiedener kognitiver Architekturen (wie z. B. EPIC und ACT-R) und von verschiedenen Systemen zur Erfassung von Blickbewegungen möglich. Die Integration von zusätzlichen Importalgorithmen (bisher für Blickbewegungsdaten von ACT-R und von zwei Anlagen für die Erfassung von Blickbewegungen) stellt eine sinnvolle Erweiterung von SimTrA dar.

Das Methodenrepertoire von SimTrA ermöglicht deskriptive Analysen und deskriptive Vergleiche beliebiger Blickbewegungsdaten künstlicher und natürlicher Systeme. Die Hauptstudie konnte zeigen, dass die gewählten Analyseparameter für die Bewertung von simulierten und empirischen Blickbewegungsdaten gut gewählt sind und helfen, ein breites Spektrum von arbeitswissenschaftlichen und experimentalphysologischen Fragestellungen zu bearbeiten und zu überprüfen. Es bleibt festzustellen, ob inferenzstatistische Methoden für die Bewertung von simulierten Blickbewegungsdaten geeignet sind. Es ist davon auszugehen, dass inferenzstatistische Methoden nicht in allen Vergleichsbedingungen anwendbar sind. In diesem Zusammenhang müssten Vorbedingungen für die Anwendung dieser Methoden ebenfalls überprüft werden. Vor einer Integration ist zu überprüfen, ob bereits bestehende Systeme für die inferenzstatistische Analyse von Daten genutzt werden können.

Neben der Integration zusätzlicher Methoden für die Analyse und den Vergleich von Blickbewegungen ist es aufgrund des Komponentenmodells von SimTrA möglich, weitere Komponenten für die Analyse und den Vergleich von zusätzlichen Aspekten der menschlichen Kognition zu integrieren. Vorstellbar sind beispielsweise die Integration von Methoden für die Analyse und den Vergleich von motorischen Informationen oder zu deklarativen Gedächtnisinhalten. Prinzipiell sind alle Leistungsdaten, die von kognitiven Modellen simuliert werden, geeignet, um mit SimTrA verarbeitet zu werden. Hinsichtlich der empirischen Validität und Überprüfbarkeit muss sich zeigen, welche

Maße angewandt werden können. Durch die Integration weiterer Maße könnten zusammengesetzte Analysen durchgeführt werden, wie beispielsweise die Analyse von motorischen und visuellen Daten zur Untersuchung der Hand-Auge-Koordination.

In SimTrA ist die Anzahl der Vergleichspartner aufgrund experimentalphysiologischer Überlegungen auf drei Versuchsgruppen beschränkt. Die Anwendung von SimTrA in unterschiedlichen Kontexten zeigte, dass diese Größe gut gewählt war. Diese Beschränkung kann im Bedarfsfall in der Vergleichskomponente angepasst werden, da die Komponente generisch implementiert ist und eine beliebige Anzahl von Vergleichspartnern zulässt.

Der bereitgestellte Arbeitsprozess der Datenverarbeitung in SimTrA kann einfach und flexibel an Unterbrechungen und neue Bedingungen der Arbeitsschritte angepasst werden. Die persistente Datenspeicherung ermöglicht die mehrmalige Durchführung von Auswertungen gleicher Daten (d. h. deren Analyse und Vergleich) mit unterschiedlichen Konfigurationen oder angepassten Datensätzen. Die umfassende und einfache Analyse vieler einzelner Aspekte wird somit ermöglicht.

Das Ablagekonzept von Ergebnis- und Datendateien in SimTrA kann in beliebige Datenstrukturen der Anwender eingefügt werden, da für jede Analyse und jeden Vergleich eigene Steuerungsdateien angelegt werden, die mit den jeweiligen Daten verknüpft sind. SimTrA erfüllt die aufgestellten Anforderungen an ein Werkzeug zur Datenverarbeitung (siehe Kapitel 3.2).

5.3 MÖGLICHE ANWENDUNGSGEBIETE VON SIMTRA

Das Werkzeug SimTrA ermöglicht die Bewertung von Mensch-Maschine-Systemen auf der Grundlage von simulierten und empirisch erfassten Blickbewegungsdaten. Die Ergebnisse ermöglichen neben der Bewertung von Schnittstellen ebenfalls eine Anpassung der Gestaltung. SimTrA stellt die benötigten Informationen bereit, um diese zu ermöglichen. Andere müssen nun zeigen, dass der Einsatz von kognitiven Benutzermodellen die Anpassung von Schnittstellen ermöglicht und zu Schnittstellen führt, die die menschliche Kognition besser unterstützen und einbeziehen.

Neben der Nutzung von SimTrA für die Bewertung von Mensch-Maschine-Systemen in frühen Systementwicklungsphasen kann das Werkzeug in allen weiteren Entwicklungsphasen eingesetzt werden. Für die Bewertung von bestehenden Mensch-Maschine-Systemen können beispielsweise kognitive Modellierungsmethoden genutzt werden, um optimale Interaktionsmuster zu identifizieren. Weichen empirische Daten von den identifizierten Mustern ab, können beispielsweise Problemfälle in der Gestaltung aufgedeckt werden, die überprüft und gegebenenfalls angepasst werden müssen.

Anhand der experimentellen Hauptstudie konnten neben dem vorrangigen Einsatzgebiet von SimTrA weitere Anwendungsgebiete identifiziert werden. Für die kognitive Modellierung (siehe Kapitel 2.2.2) stellt SimTrA ein Werkzeug dar, das die Modellbildung in verschiedenen Prozessschritten unterstützt und neue Betrachtungsebenen ermöglicht. Für die Modellprüfung ermöglicht SimTrA neben der Nutzung

von Performanzdaten kognitiver Modelle die Nutzung von simulierten Blickbewegungsdaten. Des Weiteren können auf Grundlage empirischer Daten neue Modellvarianten mit SimTrA ohne großen Mehraufwand getestet und bewertet werden. Dieser Schritt wird vereinfacht, da zusätzlich simulierte oder empirische Daten (z. B. nach Änderung der Versuchsumgebung) in jedem Prozessschritt den Analysen und Vergleichen hinzugefügt werden können. Für die Erfassung von empirischen Daten als Datenbasis für die kognitive Modellbildung kann SimTrA ebenfalls eingesetzt werden, um Blickbewegungsdaten aufzubereiten und in geeigneter Form darzustellen. Anhand dieser Daten können für die Modellimplementierung beispielsweise verschiedene Strategien identifiziert werden, die in kognitiven Modellen integriert werden können.

Ein weiterer Anwendungsfall stellt das Testen von kognitionspsychologischen Theorien dar, die in kognitiven Modellen formalisiert werden. Mit SimTrA können Unterschiede zwischen kognitiven Modellen und den zugrunde liegenden kognitionspsychologischen Theorien erkannt werden. Zum anderen kann durch den Vergleich der Simulationsdaten mit empirischen Daten abgeleitet werden, welcher theoretische Ansatz eher dem menschlichen Verhalten entspricht.

SimTrA stellt für die kognitive Modellbildung einen wichtigen Baustein dar, der ein weites Feld von Möglichkeiten eröffnet. Der Benutzer wird bei der Untersuchung von arbeitswissenschaftlichen und experimentalpsychologischen Fragestellungen algorithmisch unterstützt. Das Werkzeug übernimmt die schwierige Aufgabe der Datenverarbeitung, entlastet somit den Benutzer und ermöglicht die Konzentration auf die arbeitswissenschaftlichen und experimentalpsychologischen Fragestellungen.

ZUSAMMENFASSUNG UND AUSBLICK

In diesem Kapitel wird die Arbeit zusammengefasst und ein Ausblick auf die kognitive Modellbildung für die Bewertung von Mensch-Maschine-Systemen gegeben.

6.1 ZUSAMMENFASSUNG

Die Bewertung und Gestaltung von Mensch-Maschine-Systemen bereits in frühen Phasen der Systementwicklung wird heute anhand von modellbasierten Methoden durchgeführt. Diese können aufgrund ihrer computerbasierten Formalisierung bereits in diesen Phasen eingesetzt werden. Mit der Anwendung kognitiver Modellierungsmethoden, die Verhalten auf Basis kognitionspsychologischer Theorien simulieren, können kognitive Anforderungen eines Mensch-Maschine-Systems an den Benutzer schon während des Entwicklungsprozesses formal überprüft werden. Abgeleitete Aussagen aus den Simulationsdaten ermöglichen, auf Grundlage von kognitiven Simulationsexperimenten die Schnittstelle nach ergonomisch und arbeitswissenschaftlich relevanten Kriterien zu bewerten (z. B. an die visuelle Informationsaufnahme und -verarbeitung). Aufgrund der formalen und quantitativen Leistungsdaten kognitiver Simulationsexperimente sind die Analyse und der Vergleich dieser Daten werkzeugunterstützt und automatisiert möglich. Das Potential der kognitiven Modellierung ist in Wissenschaft und Industrie bekannt. Die Methode wird aber nur vereinzelt eingesetzt. Gründe sind in der schwierigen Entwicklung kognitiver Modelle und der komplexen Analyse von Simulationsdaten kognitiver Modelle zu finden. Bis heute gibt es noch keine geeigneten Werkzeuge, die für die umfassende Erstellung und Analyse feingranularer kognitiver Modelle im Anwendungskontext der Bewertung von Mensch-Maschine-Systemen eingesetzt werden können.

In dieser Arbeit wurde die Methode und das Werkzeug *Simulation Trace Analyzer* (SimTrA) eingeführt, das die Analyse und den Vergleich von simulierten und empirischen Leistungsdaten im Kontext der Gestaltung von Mensch-Maschine-Systemen prozessbezogen unterstützt. In diesem Zusammenhang wurde untersucht, ob simulierte Leistungsdaten kognitiver Modellierungsmethoden das Potential bieten, in prospektiven Phasen der Systementwicklung eingesetzt zu werden, und ob diese mit empirischen Daten vergleichbare Aussagen simulieren.

Zunächst wurde der Entwicklungsprozess von Mensch-Maschine-Systemen eingeführt und verschiedene Methoden für deren Bewertung vorgestellt, die den Lebensphasen eines Mensch-Maschine-Systems zugeordnet werden. Anschließend wurde die kognitive Modellbildung diskutiert und darauf aufbauend argumentiert, dass kognitive Modellierungsmethoden für die Gestaltung und Evaluation von Mensch-Maschine-Schnittstellen eingesetzt werden können, um in frühen Phasen der Systementwicklung menschliche Bedürfnisse und Anforder-

rungen in die Gestaltung zu integrieren. Es wurde dargestellt, dass die Analyse kognitiver Simulationsdaten in aktuellen Forschungsarbeiten auf einer groben Betrachtungsebene durchgeführt wird (wie z. B. Ausführungszeiten und Fehlerraten). Kognitive Modelle stellen neben diesen Leistungsdaten weitere quantitative Daten bereit, die theoretisch die Bewertung des simulierten Verhaltens auf einer feineren Betrachtungsebene ermöglichen (z. B. motorische oder visuelle Leistungsdaten).

Es konnte gezeigt werden, dass für die Gestaltung und Bewertung von Mensch-Maschine-Systemen insbesondere die Analyse von Blickbewegungen einen wichtigen Aspekt darstellt. Sie unterstützt beispielsweise die Bewertung der örtlich-räumlichen Gestaltung, der visuellen Suchprozesse und der Informationsaufnahme und -verarbeitung. Blickbewegungen können mit aktuellen kognitiven Architekturen in vereinfachter und reduzierter Form abgebildet werden.

SimTrA wurde exemplarisch für die Analyse und den Vergleich von simulierten und empirischen Blickbewegungsdaten entwickelt. Der Prozess der Datenverarbeitung ist in SimTrA in drei Komponenten unterteilt: (1) die Aufbereitung der Rohdaten, (2) die Analyse der aufbereiteten Daten und (3) der Vergleich der verarbeiteten Daten. Für die Analyse und den Vergleich der Blickbewegungsdaten werden durch SimTrA Parameter bereitgestellt, die eine deskriptive Bewertung der Blickbewegungsdaten und der Mensch-Maschine-Interaktion ermöglichen.

Die Praktikabilität des theoretisch ausgearbeiteten und prototypisch implementierten Ansatzes wurde in einer experimentellen Studie zur Gestaltung von unterschiedlichen Schnittstellen eines Mensch-Maschine-Systems einer dynamischen und komplexen Aufgabenumgebung gezeigt. In der Studie wurde die Interaktion von zwei unterschiedlichen kognitiven Modellen in der kognitiven Architektur ACT-R und von Versuchspersonen mit dem Mensch-Maschine-System untersucht.

Die Ergebnisse der experimentellen Studie lassen den Schluss zu, dass simulierte Blickbewegungsdaten kognitiver Modelle das menschliche Verhalten abbilden und für die Bewertung von Mensch-Maschine-Systemen genutzt werden können. Die Abweichung der Simulationsdaten des favorisierten kognitiven Modells liegen in einem Bereich von $\pm 10\%$. Es ist zu sehen, dass die kognitionspsychologischen Theorien für eine tendenzielle Bewertung ausreichen. Erweiterungen der kognitiven Architekturen und der kognitiven Modelle sind aber notwendig, um das menschliche Verhalten genauer und in einem breiteren Spektrum abzubilden. Des Weiteren zeigt sich, dass kognitive Modelle, die auf der Ebene der Performanz (wie z. B. Anzahl von Interaktionen, Zielerfüllung) nicht zu unterscheiden sind, anhand der fein aufgelösten Blickbewegungsdaten differenziert werden können. Die Betrachtung der feinen Leistungsdaten kognitiver Modelle stellt dementsprechend einen Vorteil gegenüber der groben Betrachtungsebene dar. Ferner verdeutlicht die experimentelle Hauptstudie und die Anwendung von SimTrA im Kontext weiterer Untersuchungen, dass SimTrA geeignet ist, den Benutzer in umfangreicher Art und Weise bei der Analyse und dem Vergleich von simulierten und empirischen Blickbewegungsdaten zu unterstützen.

Die Methode und das Werkzeug SimTrA stellt eine Erweiterung des bestehenden Methodenrepertoires der kognitiven Modellbildung und der kognitiven Psychologie für das Testen von kognitionspsychologischen Theorien dar. Für den Einsatz von kognitiven Modellierungsmethoden in frühen Phasen der Entwicklung von Mensch-Maschine-Systemen ist SimTrA ein wichtiger Baustein, der die formale und computergestützte Bewertung von Mensch-Maschine-Schnittstellen auf Grundlage von kognitiven Simulationsdaten ermöglicht und somit das Wissen der Kognitionspsychologie in Systementwicklungsprozesse integriert.

6.2 AUSBLICK

Die Komplexität von zukünftigen Mensch-Maschine-Systemen wird im Vergleich zu heutigen Systemen weiter zunehmen, mehr Funktionen und Schnittstellen werden zusammengefasst und dem Benutzer gegenübergestellt. Der Fokus der Gestaltung sollte sich weiter von der Betrachtung der reinen Funktionalität auf die menschengerechte Auslegung verschieben. Bei der Entwicklung zukünftiger Mensch-Maschine-Systeme werden kognitive Aspekte daher von zunehmender Bedeutung sein. Neben der menschengerechten Gestaltung von Mensch-Maschine-Systemen wird die formale Bewertung durch kognitive Modelle einen Wettbewerbsvorteil darstellen, der ähnlich heutiger Auditierungsverfahren (z. B. der Blaue Umweltengel, TÜV) einen Qualitätsstandard setzen kann.

Die Berücksichtigung kognitiver Prozesse bei der Entwicklung von Mensch-Maschine-Systemen stellt eine herausfordernde Aufgabe für zukünftige Anwendungen dar. Der Mensch ist ein komplexes, dynamisches und flexibles System, dessen Strukturen und Prozesse heute durch die Wissenschaft nicht vollständig geklärt sind. Die heutigen kognitiven Modellierungsmethoden bilden einen kleinen Teil der menschlichen Kognition ab. Es ist zu erwarten, dass kognitive Architekturen zunehmend mehr kognitionspsychologische Theorien vereinen und Aspekte der menschlichen Kognition abbilden werden. Ein Ziel werden die *Unified Theories of Cognition* von Newell (1990) sein. Es wird sich herausstellen, ob das menschliche Verhalten in formalen computerbasierten Strukturen abgebildet werden kann oder ob das menschliche Verhalten mehr als die algorithmische Verarbeitung von Informationen ist. Sicher ist, dass das Spektrum der zu simulierenden kognitiven Prozesse in kognitiven Architekturen durch weitere Forschung erweitert und die Vorhersagegüte des menschlichen Verhaltens erhöht wird. In vielen Bereichen werden kognitive Architekturen erweitert und an neue Erkenntnisse angepasst. Beispielsweise werden Aspekte wie die Zeitschätzung (Dzaack, Trösterer, Pape & Urbas 2007) und die visuelle Suche (Wu, Remington & Lewis 2006; Tamborello & Byrne 2006) in verschiedenen Forschungsprojekten bearbeitet und in kognitive Architekturen integriert. Zunehmend werden kognitive Modellierungsmethoden genutzt, um Erkenntnisse und Befunde aus neuropsychologischen Forschungsprojekten zu untersuchen und zu erklären (Anderson et al. 2004).

Für die Bewertung von Mensch-Maschine-Systemen wird heute auch die Methode der Anthropometrie eingesetzt. Diese Methode ermöglicht

die Simulation der menschlichen Interaktion mit technischen Systemen auf Grundlage von biologischen und physischen Beschränkungen (wie z. B. Muskelkraft und Sehwinkel; Remlinger & Bubbs 2007; Siebertz, Rausch, Rasmussen & Tørholm 2007). Die Anthropometrie beschreibt den Menschen allerdings auf funktionaler Ebene und vernachlässigt die involvierten kognitiven Prozesse. Kognitive Aspekte werden in Form von Ausführungszeiten und Erfahrungswerten in die Modelle integriert. Die Synthese von kognitiven Modellen und anthropometrischen Modellen kann somit die Schwächen beider Ansätze überwinden und eine umfassende Simulation des Menschen ermöglichen.

Das Werkzeug SimTrA gibt einen Einblick in die heute mögliche Analyse simulierter kognitiver Prozesse. Durch die Erweiterungen kognitiver Architekturen wird sich das Bild kognitiver Modellierungsmethoden ändern. Der heutige Einblick wird vergrößert, und eine Vielzahl von ergonomisch und arbeitswissenschaftlich relevanten Kriterien wird mit kognitiven Modellierungsmethoden untersucht werden können. Neue Betrachtungsebenen der kognitiven Prozesse werden durch eine weitere Beschäftigung mit kognitiven Modellen und der Analyse der simulierten Leistungsdaten ermöglicht. Durch SimTrA wird ein Werkzeug bereitgestellt, das genau diese Aspekte der kognitiven Modellierung unterstützt. Im Rahmen der Bewertung von Mensch-Maschine-Systemen werden somit neue und spannende Fragestellungen bearbeitet werden können, die mit den an der Interaktion beteiligten kognitiven Prozessen verknüpft sind. Welche Schnittstelle ist für die kognitiven Prozess am besten geeignet? Welche Schnittstelle unterstützt in Situationen mit hoher Arbeitsbelastung die menschliche Kognition? Welche Schnittstelle ist aus Sicht der kognitiven Struktur des menschlichen Geistes gut gestaltet? Solche und ähnliche Fragen werden durch eine Analyse der kognitiven Prozesse und deren Wechselwirkungen und durch den Fortschritt der Erkenntnis über die menschliche Kognition in der Kognitionspsychologie theoretisch gestellt und anhand von kognitiven Modellierungsmethoden praktisch überprüfbar.

Diese Arbeit leistet einen Beitrag, um kognitive Aspekte bei fertigungs- und regelbasierten Tätigkeiten in Entwicklungsprozesse von Mensch-Maschine-Systemen zu integrieren. Das Ziel des Einsatzes kognitiver Benutzermodelle für die Bewertung von Mensch-Maschine-Schnittstellen stellt die optimale Gestaltung der Schnittstellen aus dem Blickwinkel der menschlichen Kognition dar. Das durch SimTrA bereitgestellte Methodenrepertoire ermöglicht die einfache und kostengünstige Analyse von Simulationsdaten kognitiver Modelle, die mit realen oder virtuellen Prototypen von Mensch-Maschine-Systemen interagieren. Durch die Integration der kognitiven Benutzermodellierung in die verschiedenen Phasen der Entwicklungsprozesse technischer Systeme, wird die ganzheitliche Betrachtung der drei Aspekte, die das Verhalten des Menschen bei der Interaktion mit technischen Systemen beeinflussen (d. h. die Aufgabe, die Umgebung und die Kognition), ermöglicht. Somit können technische Systeme durchgängig an die Grenzen und Möglichkeiten der menschlichen Kognition angepasst werden. Das entwickelte Werkzeug SimTrA stellt einen Ausgangspunkt dar, um diese Entwicklung schon heute zu unterstützen und von diesem Punkt aus neue Fragestellungen zu entwickeln und zu untersuchen.

LITERATURVERZEICHNIS

- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1987). Methodologies for studying human knowledge. *Behavioral and Brain Sciences*, 10(3), 467–477.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C. & Qin, Y. (2004). An integrated theory of mind. *Psychological Review*, 111, 1036–1060.
- Anderson, J. R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396–408.
- Baggen, R. & Hemmerling, S. (2002). Evaluation und Benutzbarkeit in Mensch-Maschine-Systemen. In K.-P. Timpe, T. Jürgensohn & H. Kolrep (Hg.), *Mensch-Maschine-Systemtechnik – Konzepte, Modellierung, Gestaltung, Evaluation* (S. 233–284). Düsseldorf: Symposium Publishing.
- Baker, R. S., Corbett, A. T. & Koedinger, K. R. (2003). Statistical techniques for comparing act-r models of cognitive performance. In *Proceedings of the tenth annual act-r workshop* (S. 129–134). Pittsburgh, PA: Carnegie Mellon University.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K. & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20, 723–767.
- Balzert, H. (1998). *Lehrbuch der Software-Technik: Software-Management, Software-Qualitätssicherung, Unternehmensmodellierung*. Heidelberg, Berlin: Spektrum Akademischer Verlag.
- BITV. (2002). *Verordnung zur Schaffung barrierefreier Informationstechnik nach dem Behindertengleichstellungsgesetz*. <http://bundesrecht.juris.de/bundesrecht/bitv/gesamt.pdf> [23.10.2007].
- Boden, M. (1987). *Artificial intelligence and natural man*. London: The Massachusetts Institute of Technology Press.
- Boersema, T., Zwaga, H. J. G. & Adams, A. S. (1989). Conspicuity on realistic scenes: An eye-movement measure. *Applied Ergonomics*, 20(4), 267–273.
- Byrne, M. D. (2003). Cognitive architectures. In J. Jacko & A. Sears (Hg.), *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (S. 97–117). Mahwah, NJ: Lawrence Erlbaum Associates.
- Byrne, M. D. (2007). Local theories versus comprehensive architectures: The cognitive science jigsaw puzzle. In W. D. Gray (Hg.), *Integrated models of cognitive systems* (S. 431–443). New York, NY: Oxford University Press.
- Cabiati, C., Pastormerlo, M., Schmid, R. & Zambarbieri, D. (1983). Computer analysis of saccadic eye movements. In R. Groner, C. Menz, D. F. Fisher & R. A. Monty (Hg.), *Eye movements and*

- psychological functions: International views* (S. 19–29). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Card, S. K., Moran, T. P. & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carpenter, R. H. S. (1988). *Movements of the eyes*. London: Pion Limited.
- Chandrasekaran, B. (1990). What kind of information processing is intelligence? In D. Partridge & Y. Wilks (Hg.), *The foundations of artificial intelligence* (S. 14–46). Cambridge, MA: Cambridge University Press.
- Cooper, R. P. (2002). *Modelling high-level cognitive processes*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cooper, R. P. & Shallice, T. (1995). Soar and the case for unified theories of cognition. *Cognition*, 55, 115–149.
- Deppe, W. (1977). *Formale Modelle in der Psychologie*. Stuttgart: Kohlhammer.
- DIN EN ISO 9241-11. (1998). *Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten – Teil 11: Anforderungen an die gebrauchstauglichkeit*. Berlin: Beuth.
- Dörner, D. (1989). Modellbildung und Simulation. In E. Roth (Hg.), *Sozialwissenschaftliche Methoden* (S. 337–350). München: Oldenbourg Verlag.
- Dörner, D. (1994). Heuristik der Theoriebildung. In T. Hermann & W. H. Tack (Hg.), *Methodologische Grundlagen der Psychologie* (S. 343–388). Göttingen: Hogrefe.
- Dörner, D. (1995). *Die Logik des Mißlingens*. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag GmbH.
- Dzaack, J., Trösterer, S., Pape, N. & Urbas, L. (2007). A computational model of retrospective time estimation. *Cognitive Systems Research*, 8(3), 208–215.
- Ellis, S. R. & Smith, J. D. (1985). Patterns of statistical dependency in visual scanning. In R. Groner, G. W. McConkie & C. Menz (Hg.), *Eye movements and human information processing* (S. 221–238). Amsterdam: Elsevier.
- Ellis, S. R. & Stark, L. (1986). Statistical dependency in visual scanning. *Human Factors*, 28(4), 421–438.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381–391.
- Freed, M. A., Shafto, M. G. & Remington, R. W. (1998). Using simulation to evaluate designs: The apex approach. In S. Chatty & P. Dewan (Hg.), *Engineering for human-computer interaction* (S. 207–223). Dordrecht: Kluwer Academic.
- Freise, A. (2003). Der Nutzen gut bedienbarer Produkte. *Siemens - Pictures of the Future*, 2, 65.
- Funke, J. & Spering, M. (2006). Methoden der Denk- und Problemlöseforschung. In *Denken und Problemlösen (=Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition, Band 8: S. 675–744)*. Göttingen: Hogrefe.
- Gigerenzer, G. (1988). Woher kommen Theorien über kognitive Prozesse? *Psychologische Rundschau*, 39, 91–100.
- Gigerenzer, G. (1991). From tools to theories: A heuristic discovery in cognitive psychology. *Psychological Review*, 98, 254–267.

- Gluck, K. A., Ball, J. T. & Krusmark, M. A. (2007). Cognitive control in a computational model of the predator pilot. In W. D. Gray (Hg.), *Integrated models of cognitive systems* (S. 13–28). New York, NY: Oxford University Press.
- Göbel, M. (1999). Blickbewegungsregistrierung bei der Gestaltung von Fahrerarbeitsplätzen des ÖPNV. In M. Rötting & K. Seifert (Hg.), *Blickbewegungen in der Mensch-Maschine-Systemtechnik* (S. 118–132). Sinzheim: Pro Universitate Verlag.
- Goldberg, J. H. & Kotval, X. P. (1998). Eye movement-based evaluation of the computer interface. In S. K. Kumar (Hg.), *Advances in occupational ergonomics and safety* (S. 529–532). Amsterdam: IOS Press.
- Goldberg, J. H. & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, 24(6), 631–645.
- Gray, W. D. (2007). Composition and control of integrated cognitive systems. In W. D. Gray (Hg.), *Integrated models of cognitive systems* (S. 3–12). New York, NY: Oxford University Press.
- Gray, W. D., Schoelles, M. J. & Sims, C. R. (2004). Learning to choose the most effective strategy: Explorations in expected value. In *Proceedings of the sixth international conference on cognitive modeling* (S. 112–117). Mahwah, NJ: Lawrence Erlbaum Associates.
- Gray, W. D., Sims, C. R. & Schoelles, M. J. (2006). Musings on models of integrated cognition: What are they? what do they tell us that simpler approaches cannot? how do we evaluate them? In *Proceedings of the seventh international conference on cognitive modeling* (S. 12–13). Trieste: Edizioni Goliardichi.
- Greif, S. & Grediga, G. (1988). A critique and empirical investigation of „one-best-way-models“ in human-computer interaction. In M. Frese, E. Ulrich & W. Dzida (Hg.), *Psychological issues of human-computer interaction in the work place* (S. 357–377). Amsterdam: North-Holland Publishing Co.
- Groner, R., Walder, F. & Groner, M. (1984). Looking at faces: Local and global aspects of scanpaths. In A. G. Gale & F. Johnson (Hg.), *Theoretical and applied aspects of eye movement research* (S. 523–533). Amsterdam: Elsevier.
- Habel, C., Kanngießner, S. & Strube, G. (1990). Editorial: Was ist Kognitionswissenschaft. *Kognitionswissenschaft*, 1, 1–3.
- Halbrügge, M., Deml, B., Färber, B. A. & Bardins, S. (2007). ACT-CV – Erweiterung von ACT-R um maschinelle Bildverarbeitung. In *Tagungsband der 49. Fachausschusssitzung Antropotechnik der Deutsche Gesellschaft für Luft- und Raumfahrttechnik* (S. 313–331). Bonn: DGLR.
- Heddicke, V. (2002). Gestaltung multimodaler Schnittstellen. In K.-P. Timpe, T. Jürgensohn & H. Kolrep (Hg.), *Mensch-Maschine-Systemtechnik – Konzepte, Modellierung, Gestaltung, Evaluation* (S. 203–232). Düsseldorf: Symposion Publishing.
- Heinath, M. (in Vorbereitung). *Entwicklung einer Beschreibungsmethode zur Abbildung von Aufgabenmodellen in kognitive Simulationsmodelle*. Berlin: Technische Universität Berlin.
- Heinath, M. & Urbas, L. (2007). Simplifying the development of cognitive models using a pattern-based modeling. In *Proceedings of*

- the 10th ifac/ifip/ifors/iea symposium on analysis, design, and evaluation of human-machine-systems – CD-ROM-version.* Seoul: IFAC.
- Henninger, S. (2001). A methodology and tools for applying context-specific usability guidelines to interface design. *Interacting with Computers*, 12(3), 225–243.
- Holzinger, A. (2005). Usability engineering methods for software developers. *Communications of the ACM*, 48(1), 71–74.
- Isfort, A. (1986). Extraktion von Indikatoren zur Beschreibung von Blickverhalten. In L. J. Issing, H. D. Mickasch & J. Haacka (Hg.), *Blickbewegung und Bildverarbeitung* (S. 129–136). Frankfurt am Main, Bern und New York: Peter Lang.
- Itti, L. & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194–203.
- Johnson-Laird, P. N. (1988). *The computer and the mind*. Cambridge, MA: Harvard University Press.
- Joos, M., Rötting, M. & Velichkovsky, B. M. (2003). Die Bewegungen des menschlichen Auges: Fakten, Methoden und innovative Anwendungen. In G. Rickheit, T. Hermann & W. Deutsch (Hg.), *Handbuch der Psycholinguistik* (S. 142–168). Berlin: De Gruyter.
- Jürgensohn, T. (2002). Bedienermodellierung. In K.-P. Timpe, T. Jürgensohn & H. Kolrep (Hg.), *Mensch-Maschine-Systemtechnik – Konzepte, Modellierung, Gestaltung, Evaluation* (S. 107–148). Düsseldorf: Symposion Publishing.
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329–354.
- Kaindl, H., Lutz, B. & Trippold, P. (1998). *Methodik der Softwareentwicklung: Vorgehensmodell und State of the Art der professionellen Praxis*. Wiesbaden: Vieweg Verlag.
- Kieras, D. E. (1985). The why, when, and how of cognitive simulation: A tutorial. *Behavior Research Methods, Instruments, and Computers*, 17(2), 279–285.
- Kieras, D. E. & Polson, P. G. (1988). An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies*, 22, 365–394.
- Kindsmüller, M. C., Leuchter, S., Schulze-Kissing, D. & Urbas, L. (2004). Modellierung und Simulation menschlichen Verhaltens als Methode der Mensch-Maschine-System-Forschung. *MMI-Interaktiv*, 7, 4–16.
- Klix, F. (1966). Beziehungen zwischen Experimentalpsychologie und Volkswirtschaft. In F. Klix, J. Siebenbrodt & K.-P. Timpe (Hg.), *Ingenieurpsychologie und Volkswirtschaft* (S. 9–34). Berlin: Deutscher Verlag der Wissenschaften.
- Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In *Eye movement and their role in visual and cognitive processes* (S. 1–70). Amsterdam: Elsevier.
- Kraiss, K.-F. (1995). Modellierung von Mensch-Maschine-Systemen. In H.-P. Willumeit & H. Kolrep (Hg.), *ZMMS-Spektrum – Verlässlichkeit von Mensch-Maschine-Systemen* (Bd. 1, S. 15–35). Berlin: Technische Universität Berlin.
- Land, M. F. & Hayhoe, M. M. (2001). In what way do eye movements contribute to everyday activities? *Vision Research*, 41, 3559–3565.
- Laubsch, J. (1985). Techniken der Wissensdarstellung. In C. Habel

- (Hg.), *Künstliche Intelligenz. Frühjahrsschule Dassel – März 1984* (S. 48–93). Berlin: Springer.
- Mackworth, N. H. & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception and Psychophysics*, 2(11), 547–552.
- Marrenbach, J. R. (2001). *Werkzeug-basierte Evaluierung der Benutzungsfreundlichkeit interaktiver Endgeräte mit normativen Benutzermodellen*. Aachen: Shaker.
- Meyer, D. E. & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance: Part 1. basic mechanisms. *Psychological Review*, 104, 3–65.
- Meyer, D. E. & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. accounts of psychological refractory-period phenomena. *Psychological Review*, 104, 749–791.
- Neisser, U. (1967). *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.
- Newell, A. (1973). Productions systems: Models of control structures. In W. G. Chase (Hg.), *Visual information processing* (S. 283–308). New York, NY: Academic Press.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A. & Simon, H. A. (1972). *Human problem solver*. Englewood Cliffs, NJ: Prentice Hall.
- Nicholson, S., Byrne, M. D. & Fotta, M. E. (2006). Modifying act-r for visual search of complex displays. In *Proceedings of the thirteenth annual act-r workshop 2006* (S. 108–113). Pittsburgh, PA: Carnegie Mellon University.
- Norman, D. (1986). Cognitive engineering. In D. Norman & S. Draper (Hg.), *User centered system design* (S. 64–94). Hillsdale, NJ: Erlbaum.
- Opwis, K. (1992). *Kognitive Modellierung – Zur Verwendung wissensbasierter Systeme in der psychologischen Theoriebildung*. Bern: Verlag Hans Huber.
- Payne, S. J. & Green, T. R. G. (1986). Task-action grammars: A model of the mental representation of task languages. *Human Computer Interaction*, 2, 93–133.
- Pomberger, G. & Blaschek, G. (1996). *Software Engineering: Prototyping und objektorientierte Software-Entwicklung*. München: Hanser.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Pylyshyn, Z. (1989). Computing in cognitive science. In M. I. Posner (Hg.), *Foundations of cognitive science* (S. 51–91). Cambridge, MA: The Massachusetts Institute of Technology Press.
- Rasmussen, J. (1982). Human errors. a taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents*, 4, 311–335.
- Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs and symbols, and other distinctions in human performance models. *IEEE Transactions on System, Man and Cybernetics*, SMC-13(3), 257–266.
- R Development Core Team. (2006). *R: A language and environment for statistical computing*. Vienna.

- Reinhart, G., Lindemann, U. & Heinzl, J. (1996). *Qualitätsmanagement – Ein Kurs für Studium und Praxis*. Berlin: Springer.
- Remlinger, W. & Bubb, H. (2007). Erweiterte Möglichkeiten der Sichtanalyse mit dem Menschmodell RAMSIS. In *Tagungsband der 49. Fachausschusssitzung Antropotechnik der Deutsche Gesellschaft für Luft- und Raumfahrttechnik* (S. 189–198). Bonn: DGLR.
- Riedl, M. O. & St. Amant, R. (2002). Toward automated exploration of interactive systems. In *Proceedings of intelligent user interfaces* (S. 135–142). New York, NY: ACM.
- Ritter, F. E. & Larkin, J. H. (1994). *Developing process models as summaries of hci action sequences*. Nottingham: University of Nottingham.
- Rötting, M. (2001). *Parametersystematik der Augen- und Blickbewegungen für arbeitswissenschaftliche Untersuchungen*. Aachen: Shaker.
- Salvucci, D. D. (2001a). An integrated model of eye movements and visual encoding. *Cognitive Systems Research*, 1, 201–220.
- Salvucci, D. D. (2001b). Predicting the effects of in-car interface use on drive performance: An integrated model approach. *International Journal of Human-Computer Studies*, 55, 85–107.
- Salvucci, D. D. (2005). A multitasking general executive for compound continuous tasks. *Cognitive Science*, 29, 457–492.
- Salvucci, D. D. & Anderson, J. R. (2001). Automated eye-movement protocol analysis. *Human-Computer Interaction*, 16, 39–86.
- Salvucci, D. D. & Lee, F. J. (2003). Simple cognitive modeling in a complex cognitive architecture. In *Proceedings of the SIGCHI conference on human factors in computing systems* (S. 265–272). New York, NY: ACM.
- Schäfer, R. E. (1985). *Denken*. Berlin: Springer.
- Schaub, H. (1993). *Modellierung der Handlungsorganisation*. Bern: Verlag Hans Huber.
- Schoelles, M. J. & Gray, W. D. (2000). Argus prime: Modeling emergent microstrategies in a complex simulated task environment. In N. A. Taatgen & J. Aasman (Hg.), *Proceedings of the third international conference on cognitive modeling* (S. 260–270). Veenendaal: Universal Press.
- Schroiff, H.-W. (1986). Zum Stellenwert von Blickbewegungsdaten bei der Mikroanalyse kognitiver Prozesse. In L. J. Issing, H. D. Mickasch & J. Haack (Hg.), *Blickbewegung und Bildverarbeitung* (S. 57–82). Frankfurt am Main, Bern und New York: Peter Lang.
- Schulze-Kissing, D. (2007). *Eine Mikroanalyse des zeitlichen Fehlers in der Mensch-Maschine -Interaktion*. Berlin: Technische Universität Berlin.
- Schunn, C. D. & Wallach, D. (2005). Evaluating goodness-of-fit in comparison of models to data. In W. Tack (Hg.), *Psychologie der Kognition: Reden und Vorträge anlässlich der Emeritierung von Werner Tack* (S. 115–154). Saarbrücken: University of Saarland Press.
- SensoMotoric Instruments. (2006). *Begaze 1.2 user manual – document version 1.06.03*.
- SensoMotoric Instruments. (2007). *Webseite der SensoMotoric Instruments GmbH*. <http://www.smivision.com> [4.12.2007].
- Siebertz, K., Rausch, J., Rasmussen, J. & Tørholm, S. (2007). Simulation des menschlichen Bewegungsapparates zur Innenraumgestaltung von Fahrzeugen. In *Tagungsband der 49. Fachausschusssitzung An-*

- tropotechnik der Deutsche Gesellschaft für Luft- und Raumfahrttechnik* (S. 199–222). Bonn: DGLR.
- Smith, B. C. (1982). *Reflection and semantics in a procedural language* (Bericht). Cambridge, MA: The Massachusetts Institute of Technology.
- Sommerville, I. (2001). *Software Engineering*. München: Pearson Studium.
- Stark, L. & Ellis, S. R. (1981). Scanpath revisited: Cognitive models of direct active looking. In D. F. Fisher, R. A. Monty & J. Senders (Hg.), *Eye movements: Cognition and visual perception* (S. 193–226). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stroher, H. (1995). *Kognitive Systeme – Eine Einführung in die Kognitionswissenschaft*. Opladen: Westdeutscher Verlag.
- Strube, G. (2001). Kognitive Modellierung. In G. Strube (Hg.), *Wörterbuch der Kognitionswissenschaft – CD-ROM-Version*. Stuttgart: Klett-Cotta.
- Sun, R., Coward, L. A. & Zenzen, M. J. (2004). On levels of cognitive modeling. *Philosophical Psychology*, 18(5), 613–637.
- Sun Microsystems. (2007). *Webseite der Sun-Microsystems: Java Technologie*. <http://java.sun.com> [4.12.2007].
- Taatgen, N. A. (2001). A model of individual differences in learning air traffic control. In *Proceedings of the fourth international conference of cognitive modeling* (S. 221–216). Mahwah, NJ: Lawrence Erlbaum Associates.
- Taatgen, N. A. (2007). The minimal control principle. In W. D. Gray (Hg.), *Integrated models of cognitive systems* (S. 369–379). New York, NY: Oxford University Press.
- Tack, W. H. (1995). Wege zu einer differentiellen kognitiven Psychologie. In K. Pawlik (Hg.), *Bericht über den 39. Kongress der Deutschen Gesellschaft für Psychologie in Hamburg* (S. 172–185). Göttingen: Hogrefe.
- Tamborello, I. F. P. & Byrne, M. D. (2006). Adaptive but non-optimal visual search behavior in highlighted displays. In *Proceedings of the seventh international conference on cognitive modeling* (S. 316–321). Trieste: Edizioni Goliardichi.
- The MathWorks. (2003). *The MathWorks Deutschland - MATLAB and Simulink for Technical Computing*. <http://www.mathworks.de> [6.1.2008].
- Timpe, K.-P., Jürgensohn, T. & Kolrep, H. (2002). *Mensch-Maschine-Systemtechnik – Konzepte, Modellierung, Gestaltung, Evaluation*. Düsseldorf: Symposion Publishing.
- Timpe, K.-P. & Kolrep, H. (2002). Das Mensch-Maschine-System als interdisziplinärer Gegenstand. In K.-P. Timpe, T. Jürgensohn & H. Kolrep (Hg.), *Mensch-Maschine-Systemtechnik – Konzepte, Modellierung, Gestaltung, Evaluation* (S. 9–40). Düsseldorf: Symposion Publishing.
- Tole, J. R. & Young, L. R. (1981). Digital filters for saccade and fixation detection. In D. F. Fisher, R. A. Monty & J. W. Senders (Hg.), *Eye movements: Cognition and visual perception* (S. 248–256). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Treisman, A. M. & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology*, 16, 459–478.

- Urbas, L. (1999). *Entwicklung und Realisierung einer Trainings- und Ausbildungsumgebung zur Schulung der Prozeßdynamik und des Anlagenbetriebs im Internet*. Düsseldorf: VDI Verlag.
- Urbas, L., Dubrowsky, L., Dzaack, J. & Heinath, M. (2005). *Entwicklungsmodelle in der Praxis: Ergebnisse einer Interviewstudie* (Bericht). Berlin: ZMMS, Technische Universität Berlin.
- Urbas, L., Heinath, M. & Leuchter, S. (2007). Bedienermodellgestützte Bewertung des Ablenkungspotenzials von Komfortsystemen im KFZ in frühen Phasen der Systementwicklung. *iCom - Zeitschrift für interaktive und kooperative Medien*, 2, 21–29.
- Urbas, L. & Leuchter, S. (2005). Model based analysis and design of human-machine dialogues through displays. *KI – Künstliche Intelligenz*, 4, 45–51.
- Urbas, L., Schulze-Kissing, D. & Leuchter, S. (2005). Werkzeuge für die Erstellung kognitiver Nutzermodelle. In K. Karrer, B. Gauss & C. Steffens (Hg.), *Beiträge zur Mensch-Maschine-Systemtechnik aus Forschung und Praxis* (S. 247–264). Düsseldorf: Symposion Publishing.
- Vernetztes Fahren. (2007). *Webseite des Verbundprojektes Vernetztes Fahren*. <http://www.vernetztes-fahren.de> [4.12.2007].
- Viviani, P. (1990). Eye movements in visual search: Cognitive, perceptual and motor control aspects. In E. Kowler (Hg.), *Eye movement and their role in visual and cognitive processes* (S. 353–393). Amsterdam: Elsevier.
- Wallach, D. (1998). *Komplexe Regelungsprozesse – Eine kognitionswissenschaftlicher Analyse*. Wiesbaden: Deutscher Universitäts Verlag.
- Ward, R. D. & Marsden, P. H. (2003). Physiological responses to different web page designs. *International Journal of Human-Computer Studies*, 59, 199–212.
- Weimer, J. (1995). *Research techniques in human engineering*. Englewood Cliffs, NJ: Prentice Hall.
- Wickens, C. D. & Holland, J. G. (2000). *Engineering psychology and human performance*. Upper Saddle River, NJ: Prentice Hall.
- Willumeit, H., Gediga, G. & Hamborg, K.-C. (1996). Isometrics: Ein Verfahren zur formativen Evaluation von Software nach ISO 9241/10. *Ergonomie & Informatik*, 27, 5–12.
- Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, 1, 202–238.
- Wu, S.-C., Remington, R. W. & Lewis, R. (2006). Modeling the scheduling of eye movements and manual response in performing a sequence of discrete tasks. In *Proceedings of the 28th annual conference of the cognitive science society* (S. 900–905). Mahwah, NJ: Lawrence Erlbaum Associates.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum Press.
- Zinterhof, P. (1987). Mathematische Systemtheorie. In E. Roth (Hg.), *Sozialwissenschaftliche Methoden* (S. 105–121). München: Oldenbourg Verlag.