

# **Anpassung der Funktionalität an die Wissensstruktur geübter Benutzer/innen**

Edmund Eberleh, Aachen

## **Zusammenfassung**

Geübte Benutzer eines grafikbasierten Computersystems sahen Bildschirmoberflächen aus verschiedenen Dialogzuständen. In einer folgenden Testphase bekamen sie neben anderen Proben stets einen Dialogzustand dargeboten, den sie in der Lernphase *nicht* gesehen hatten, der aber zu der Handlungssequenz und dem Handlungsfeld der gesehenen Zustände gehörte. Die Vpn beurteilten diese Dialogzustände umso häufiger fälschlicherweise als gesehen, je mehr Dialogzustände sie zuvor aus der jeweiligen Handlungssequenz gesehen hatten. Die Reaktionszeiten für richtige Zurückweisungen der fraglichen Dialogzustände stiegen in gleicher Weise an. Die Ergebnisse werden auf aktivations-theoretischer Grundlage dahingehend interpretiert, daß geübte Benutzer ihr Wissen in auf Ziele ausgerichtete Handlungssequenzen organisiert haben.

## **1 Fragestellung**

In der Regel haben Benutzer/innen eines Computers bestimmte Arbeitsaufgaben, die sie mit Unterstützung des Computers erledigen wollen. Die zur Zielerreichung notwendigen Handlungen am Computer sind durch dessen Funktionalität und Benutzungsschnittstelle bestimmt. Die Passung der Systemstruktur mit der Aufgabenstruktur oder deren Kompatibilität wird allgemein als ein wichtiger Faktor der benutzungsfreundlichen und nützlichen Softwaregestaltung angesehen (z.B. DIN 66234 T8; Kieras & Polson, 1985; Rohr, 1988; Ulich, 1986).

Entscheidend ist jedoch die Frage, in Bezug auf *welche* Merkmale die Passung vorliegen bzw. optimal sein sollte. Ich nehme an, daß hierfür neben der *konzeptuellen* Aufgaben- und Systemstruktur ihrer *mental*en Repräsentation als vermittelnder, integrierender und handlungsleitender Instanz große Bedeutung zukommt. Im folgenden berichte ich ein Experiment, welches die Wissensrepräsentation geübter Benutzer/innen von einem Computersystem und damit möglichen Arbeitshandlungen untersucht, um auf dieser Grundlage zu Aussagen

über eine benutzer- und aufgabenangemessene Gestaltung der Systemfunktionalität zu gelangen.

Die meisten vorliegenden Ansätze zur Gestaltung der Funktionalität betrachten die konzeptuellen Aufgaben- und Systemstrukturen. So benutzt Moran (1984) die Zahl der Zuordnungsregeln von externen Arbeitsaufgaben auf systembedingt mögliche Operationen als ein Maß der Systemkomplexität und Bedienleichtigkeit. Nach Card, Moran und Newell (1983) ist ein Großteil der Komplexität des Benutzerverhaltens auf die Aufgabenstruktur und deren Umsetzungsbedingungen aufgrund des Computersystems zurückzuführen, und nicht auf die Komplexität des menschlichen Denkens an sich. Rohr (1988) sieht die Effektivität eines Software-Werkzeuges zwar durch die mehr oder weniger gute Korrespondenz zwischen *mentalen* Aufgabenzuständen und Systemoperationen bestimmt, zieht hieraus jedoch den Schluß, daß versucht werden sollte, Funktionsprimitive für die reine Aufgabenlösung und deren Transformationsschritten zu finden, und daß daher detaillierte Aufgabenanalysen notwendig seien.

Den genannten Ansätzen liegt somit implizit die Annahme zugrunde, daß die mentale Repräsentation des Benutzers eine vollständige, strukturerhaltende Abbildung der Aufgabenstruktur und der Dialogstruktur darstellt, und damit aus der Betrachtung ausgeklammert werden kann. Eine Eigenschaft mentaler Modelle besteht jedoch darin, daß sie das fragliche System meistens nur unvollständig, reduziert und fehlerhaft abbilden (Kempton, 1986; Norman, 1983), und sich im Laufe des Wissenserwerbs sowohl ihre Struktur als auch ihr Inhalt ändern kann (Forbus & Gentner, 1986).

Für eine benutzer- und aufgabenbezogene Gestaltung der Funktionalität sind im Prinzip zwei Entwurfsstrategien denkbar (vgl. Shneiderman, 1987)

1. In der Entwurfsphase wird eine genaue Analyse der Arbeitsaufgaben und Ziele der künftigen Benutzer vorgenommen. Die Funktionalität des Systems wird so gewählt, daß diese Ziele vom Benutzer direkt und mit möglichst wenig Aufwand erreicht werden können. Diese Vorgehensweise hat den Nachteil der mangelnden Flexibilität des Systems für abweichende, neue Aufgaben.

2. Den Benutzer/innen werden viele elementare Funktionen bzw. Operatoren zur Verfügung gestellt, und sie kombinieren sie selber flexibel zu den für ihre Aufgaben und Arbeitsbedingungen notwendigen Handlungssequenzen. Diese Vorgehensweise birgt den Nachteil, daß das System für den Benutzer undurchschaubar und durch kaum benutzte Funktionen unnötigerweise komplex werden kann.

Sind die Dialogzustände des Computers und ihre Übergänge von Benutzern mental in zielgerichteten, voneinander unabhängigen Handlungssequenzen organisiert

(wie etwa im GOMS-Modell von Card, Moran und Newell, 1983 angenommen wird), kommt diesem Repräsentationsprinzip eine an Zielen ausgerichtete Funktionalität entgegen. Ein alternatives Repräsentationsprinzip wäre, daß mit den einzelnen Dialogzuständen alle jeweils möglichen Folgezustände fächerartig assoziiert sind und so die Handlungsmöglichkeiten in einem Dialogzustand oder das Handlungs *feld* (Oesterreich, 1981) des Benutzers abbilden. In diesem Falle erscheint eine flexible, modular kombinierbare Funktionalität passender zu sein. Das im folgenden beschriebene Experiment soll Hinweise darüber geben, welches dieser zwei Repräsentationsprinzipien die Wissensstruktur geübter Benutzer angemessener beschreibt.

Routinierte Benutzer eines graphikbasierten Computersystems bekamen Bildschirmoberflächen aus verschiedenen Dialogzuständen dargeboten. Diese Dialogzustände stammten aus jeweils einer Handlungssequenz und/oder einem Handlungsfeld. In einer unmittelbar darauffolgenden Testphase bekamen die Versuchspersonen (Vpn) mehrere Bildschirmoberflächen (Dialogzustände) dargeboten und sollten jeweils entscheiden, ob sie das Bild eben gesehen hatten oder nicht. Unter den Testproben befand sich stets ein Dialogzustand (*target*), den die Vpn in der Lernphase nicht gesehen hatten, der aber zu der Handlungssequenz und dem Handlungsfeld der gesehenen Zustände gehörte. Dieser Dialogzustand war somit der Kreuzungspunkt der aus den Zuständen der Lernphase gebildeten horizontalen Zielsequenz und des vertikalen Handlungsfelds (s. Abbildung 1).

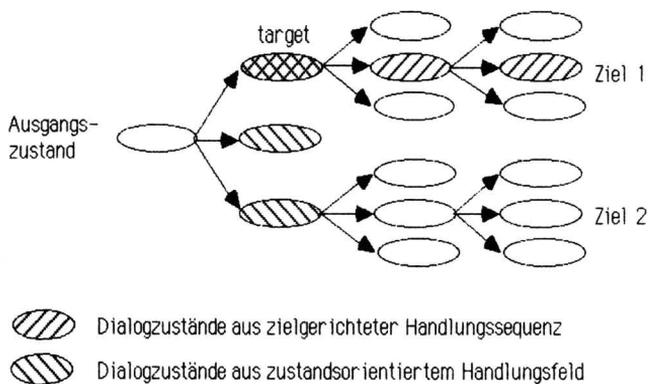


Abb. 1. Vereinfachte, schematische Darstellung der verwendeten Dialogstruktur mit Handlungssequenzen, Handlungsfeldern und dem *target*.

Ich nehme an, daß eine Vp ihre Entscheidung aufgrund des Aktivationsgrades der jeweiligen Probe (Dialogzustand) in ihrer Gedächtnisrepräsentation trifft. Gesehene Dialogzustände sollten im Mittel einen höheren Aktivationswert haben als nicht gesehene. Liegt die mentale Aktivierung einer zu beurteilenden Probe über einer gewissen kritischen Schwelle, antwortet die Vp mit "gesehen", ansonsten mit "nicht gesehen" (vgl. Anderson, 1983). Sind die Werte der jeweils aktuellen Probe sehr weit von dieser Schwelle entfernt, kann die Entscheidung schnell getroffen werden. Liegen sie dagegen sehr dicht an der kritischen Grenze, muß auf zeitaufwendigere genauere Gedächtnissuche vorgenommen werden (Atkinson & Juola, 1974).

Je mehr Zustände die Vp nun aus der Zielsequenz bzw. dem Handlungsfeld sieht, desto stärker sollte die mentale Repräsentation der Zielsequenz bzw. des Handlungsfelds insgesamt aktiviert werden, falls die kognitive Repräsentation nach dem jeweiligen Prinzip strukturiert ist, und damit auch das nicht gesehene, aber zu der jeweiligen Struktur zugehörige target (vgl. Pohl, Colonius & Thüring, 1985). Je mehr sich die Aktivierung der targets jedoch den Aktivationswerten für positive Proben annähert, desto häufiger werden targets über der kritischen Schwelle liegen und deshalb fälschlicherweise als "gesehen" beurteilt werden. Gleichzeitig sollten die Entscheidungszeiten für richtige Zurückweisungen der targets entsprechend des genannten 2-stufigen Entscheidungsprozesses umso länger werden, je mehr die Aktivierung eines targets durch Proben aus der gleichen Handlungssequenz bzw. dem gleichen Handlungsfeld an den Aktivierungsgrad tatsächlich gesehener Dialogzustände angeglichen und damit schwerer davon zu unterscheiden wird.

## 2 Methode

### 2.1 Versuchsablauf

Das Experiment wurde für jede Versuchsperson in 9 Einzelsitzungen an aufeinanderfolgenden Tagen durchgeführt. Die Versuchsperson saß dabei vor dem Computersystem Apple Macintosh Plus. Jede Vp erhielt zu Beginn eine schriftliche Instruktion, in welcher der Ablauf des Versuchs und ihre Aufgabe erklärt wurde. Danach folgten zur Demonstration und praktischen Übung zwei beispielhafte Versuchsdurchgänge. Diese zwei Versuchsdurchgänge wurden auch zu Beginn jeder weiteren Tagessitzung als Aufwärmphase durchlaufen. Pro Sitzung

folgten dann insgesamt 20 Versuchsdurchgänge, in der Hälfte unterbrochen von einer 15-minütigen Pause, in denen die Vp Kreuzworträtsel löste. Jede Einzelsitzung dauerte insgesamt etwa 45 Minuten.

Jeder Versuchsdurchgang bestand aus einer Lern- und Testphase. In der Lernphase bekam die Vp für jeweils 5 Sekunden insgesamt 8 verschiedene Bildschirmoberflächen aus verschiedenen Dialogzuständen des Apple Macintosh dargestellt. Die Vp sollte sich die Bildschirmoberflächen konzentriert anschauen und deren *Bedeutung* erfassen.

In der folgenden Testphase bekam die Vp insgesamt fünf weitere Bildschirmoberflächen dargestellt. Sie mußte bei jedem Bild durch Druck auf eine von vier Tasten entscheiden, ob sie dieses in der unmittelbar vorausgehenden Lernphase gesehen hatte oder nicht. Die vier Tasten bildeten eine Ratingskala von "ganz sicher gesehen" über "sicher gesehen" auf "sicher nicht gesehen" und "ganz sicher nicht gesehen". Nach dem Tastendruck verschwand das Bild und das nächste erschien nach 500 ms. Für jedes Bild wurde die Antwortkategorie und Reaktionszeit registriert. Die Vp wurde instruiert, möglichst alle Antwortkategorien auszuschöpfen und so schnell und genau wie möglich zu antworten.

## 2.2 Material

Auf der Grundlage der Zustandsübergangnetzwerke des Textverarbeitungsprogramms MacWrite und des Zeichenprogramms MacDraw wurden Dialogzustände (targets) ausgewählt, die *sowohl* einer zielgerichteten, häufig auftretenden Handlungssequenz mit mindestens 5 Zuständen *als auch* gleichzeitig einem entsprechend großem Handlungsfeld angehörten (vgl. Abb. 1). Für jedes Programm wurden 10 derartige targets aus möglichst allen Handlungsfeldern und Positionen der Handlungssequenzen ausgewählt.

Für jede Lernphase wurde eine Testsequenz entwickelt, in der das jeweilige target als negative Probe, Bilder aus der Lernphase als positive Proben und Zustände aus anderen aber ähnlichen Netzwerkteilen als Distraktoren bzw. negative Proben vorkamen. Die Zahl positiver und negativer Proben wurde pro Person über alle Versuchsdurchgänge ausgeglichen. Die Reihenfolge der Bilder der Lern- und Testphasen wurde für jede Vp randomisiert. Die Reihenfolge der insgesamt 180 Versuchsdurchgänge wurde ebenfalls für jede Vp unterschiedlich durch Randomisierung festgelegt. Versuchsdurchgängen mit Proben aus MacWrite folgten jedoch stets Versuchsdurchgänge mit Dialogzuständen aus MacDraw, um die Interferenz zwischen den Lernsequenzen zu verringern.

### 2.3 Versuchsplan

Dem Experiment lag ein 3x3 - faktorieller, vollständig gekreuzter Versuchsplan mit Meßwiederholung über beide Faktoren zugrunde. Unabhängig voneinander variiert wurde die Zahl der Dialogzustände aus der zielgerichteten Handlungssequenz des targets (im folgenden kurz "Zielsequenz" genannt) und die Zahl der Dialogzustände aus dem Handlungsfeld des targets (im folgenden kurz "Handlungsfeld"), mit jeweils 0, 2 und 4 Zuständen. Für jede der sich so ergebenden 9 Versuchsbedingungen wurden pro Person 20 Versuchsdurchgänge durchgeführt, die durch jeweils 10 verschiedene targets aus den Programmen MacWrite und MacDraw definiert waren. Die insgesamt 20 verschiedenen targets wurden in jeder Versuchsbedingung verwendet, so daß das Reizmaterial für die Anwartsicherheit und Reaktionszeit in jeder Bedingung identisch war.

### 2.4 Versuchspersonen

Am Experiment nahmen 10 Studenten und eine Studentin der RWTH Aachen teil. Alle hatten zuvor schon seit längerer Zeit privat und beruflich mit dem Apple Macintosh und den Programmen MacWrite und MacDraw gearbeitet, und sind als hochgeübte Benutzer einzuschätzen.

## 3 Ergebnisse

Die Auswertung erfolgte im Paradigma der Signal-Entdeckungs-Theorie (vgl. McNicol, 1972). Die Signal-Entdeckungs-Theorie betrachtet die Sensitivität von Personen bei der Entdeckung von Signal-Reizen innerhalb einer Reihe von Stör-Reizen oder von "Rauschen". Im Falle des vorliegenden Experimentes waren die in der Lernphase gezeigten Dialogzustände die Signale, die die Vpn in der Testphase von den negativen Stör-Proben unterscheiden sollten.

Je mehr Dialogzustände die Vpn aus den zielgerichteten Handlungssequenzen des targets in der Lernphase gesehen hatten, desto häufiger beurteilten sie in der Testphase fälschlicherweise auch das target selber als "gesehen" (sog. falsche Alarme) (s. Abb. 2). Dialogzustände aus den Handlungsfeldern des targets führten nicht zu diesem Effekt. Eine Ausnahme bildet allerdings die Versuchsbedingung mit zwei Dialogzuständen aus der Zielsequenz bei keinem Dialogzustand aus dem Handlungsfeld. Dem Anstieg bei den targets steht eine gleichbleibende oder unsystematisch variierende Anzahl falscher Alarme bei den restlichen negativen

Proben gegenüber, deren Mittelwert etwa bei dem Wert der falschen Alarme bei keinem Dialogzustand aus der Handlungssequenz liegt.

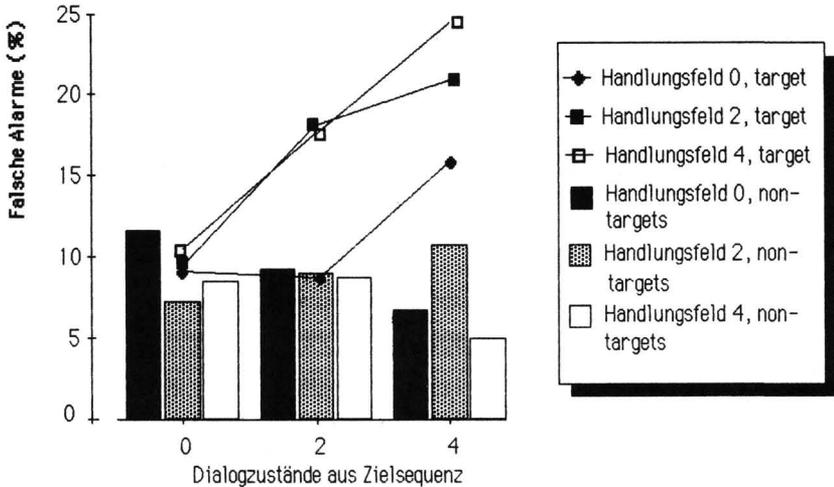


Abb. 2. Häufigkeiten falscher Alarme auf targets und auf die restlichen negativen Proben in Abhängigkeit von den Versuchsbedingungen.

Eine höhere Zahl falscher Alarme kann jedoch alleine durch eine Verringerung des Schwellenwertes der Aktivierung auftreten, ab dem eine Vp eine Probe als gesehen wertet. Bei einer derartigen Verringerung des Entscheidungskriteriums *erhöht* sich gleichzeitig die Trefferrate, so daß sich die Diskriminationsleistung *insgesamt* nicht verändert. Daher wurde pro Versuchsbedingung die Zahl der falschen Alarme und Treffer in einer ROC (Receiver Operating Characteristic) - Kurve gegenübergestellt.

Als Maß für eine gemäß den Hypothesen evtl. auftretende verringerte Diskriminationsleistung bei den targets in Abhängigkeit von den Versuchsbedingungen verwendete ich die Fläche  $P(A)$  unter den ROC-Kurven, die mit abnehmender Diskriminationsfähigkeit kleiner wird. Da dieses Maß jedoch rechtsschief verteilt ist, wurde die jeweilige Fläche mit  $2 \arcsin \sqrt{P(A)}$  transformiert (McNicol, 1972). Die auf diese Weise pro Versuchsbedingung erhaltenen über die Vpn gemittelten Werte sind in Abb. 3 dargestellt. Bis auf die schon als Ausnahme erwähnte Versuchsbedingung zeigt sich auch in der gesamten Diskriminationsleistung eine Verringerung über die Anzahl der Dialogzustände aus der



Abb. 3. Mittlere Flächen unterhalb der ROC-Kurven pro Versuchsbedingung.

Zielsequenz des targets. Die Entscheidungszeiten für die als ganz sicher eingeschätzten richtigen Zurückweisungen der targets als "nicht gesehen" sind pro Versuchsbedingung in Abb. 4 zu sehen. Je mehr Dialogzustände die Vpn aus der Zielsequenz des targets gesehen haben, desto länger brauchten sie für ihre richtige Antwort.

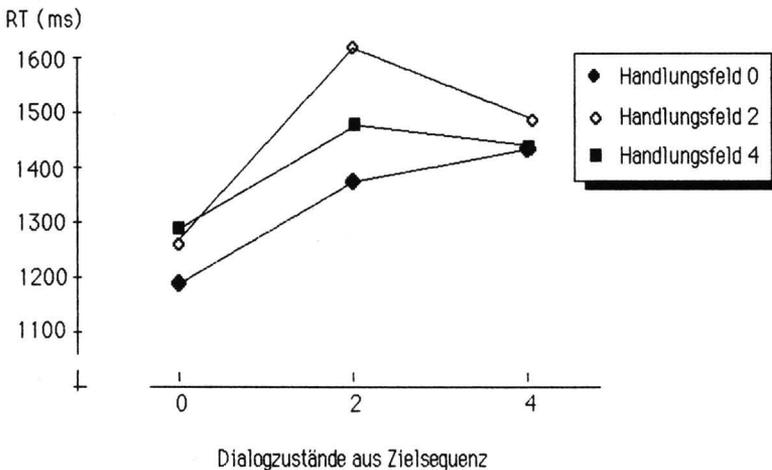


Abb. 4. Mittlere Entscheidungszeiten für als ganz sicher eingeschätzte richtige Zurückweisungen der targets.

Die berichteten Werte für die Flächen unter den ROC-Kurven und die Entscheidungszeiten wurden mit einer zweifaktoriellen, multivariaten Varianzanalyse auf statistisch signifikante Unterschiede geprüft. In der multivariaten Analyse ergibt sich nach dem Hotelling  $T^2$ -Test ein hochsignifikanter Einfluß der Zielsequenz auf Diskriminationsleistung und Entscheidungszeit ( $F_{(4,6)} = 10.83$ ;  $p < .01$ ). Der Einfluß des Handlungsfeldes ist in der gleichen Analyse auf dem 5%-Niveau signifikant ( $F_{(4,6)} = 9.04$ ;  $p < .05$ ). Werden die abhängigen Variablen getrennt voneinander (univariat) analysiert, liegt beim Faktor Handlungsfeld nur ein signifikanter Einfluß auf die Diskriminationsleistung vor, während beim Faktor Zielsequenz beide Einzelanalysen einen signifikanten Einfluß aufzeigen.

#### 4 Diskussion

Die Daten zeigen einen deutlichen Einfluß der Anzahl der Dialogzustände aus der zielgerichteten Handlungssequenz des targets auf Entscheidungsgenauigkeit und -zeit. Ein Einfluß auf die Genauigkeit liegt zwar auch für die Anzahl der Dialogzustände aus dem Handlungsfeld des targets vor, ist aber im wesentlichen auf eine einzelne abweichende Versuchsbedingung zurückzuführen. Da für diesen Datenpunkt keine plausible inhaltliche Erklärung zu finden ist, nehme ich ihn als aufgrund der beschränkten Versuchspersonenstichprobe zufällig entstanden an. Die Abnahme der Diskriminationsfähigkeit durch vermehrte Darbietung von Dialogzuständen aus einer Handlungssequenz kann nicht durch eine Interferenz aufgrund der oberflächlichen graphischen Ähnlichkeit der Bilder erklärt werden. Eine derartige Ähnlichkeit liegt - wenn überhaupt - nur bei Zuständen aus dem gleichen Handlungsfeld vor. Unter diesen Annahmen sprechen die Daten in Übereinstimmung mit bereits vorliegenden (Eberleh, im Druck) für die Hypothese, daß das Wissen geübter Benutzer/innen über die Aufgabenbearbeitung am Computer in zielgerichteten Handlungssequenzen repräsentiert ist.

Für die Bereitstellung und Gestaltung der Funktionalität von Software-Werkzeugen ergibt sich daraus die Folgerung, daß auf Basis einer sorgfältigen Analyse der Benutzerziele die Funktionalität ein direktes und möglichst einfaches Erreichen dieser Ziele ermöglichen sollte. Die dadurch bewirkte Reduktion der Dialogkomplexität (vgl. Frese, 1987; Streit & Eberleh, 1988) verschafft dem Benutzer Freiräume für die Bewältigung anspruchsvollerer inhaltlicher Aufgaben und ist Voraussetzung für eine anzustrebende ganzheitliche Aufgabenbearbeitung.

## Literatur

- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 261-295.
- Atkinson, R. C., & Juola, J. F. (1974). In D.H. Krantz, R.C. Atkinson, R.D. Luce & P. Suppes (Eds.), *Contemporary developments in mathematical psychology*, Vol. 1. San Francisco: Freeman.
- Deutsches Institut für Normung (1988) (Hrsg.). DIN 66234 Teil 8, Bildschirmarbeitsplätze: Grundsätze ergonomischer Dialoggestaltung. Berlin: Beuth.
- Eberleh, E. (in press). The structure of mental models: goal directed or state centered? In F. Klix, N. Streitz, H. Wandke & Y. Waern (Eds.), *MACINTER II: Man-Computer Interaction Research*. Amsterdam: North-Holland.
- Frese, M. (1987). A theory of control and complexity: implications for software design and integration of computer systems into the work place. In M. Frese, E. Ulich & W. Dzida (Eds.), *Psychological issues of human-computer interaction in the work place*. Amsterdam: North-Holland.
- Forbus, K. D., & Gentner, D. (1986). Learning physical domains: Toward a theoretical framework. In R. S. Michalski, J. G. Carbonell & T. M. Mitchell (Eds.), *Machine learning, An artificial intelligence approach* (Vol. 2) (pp. 311-348). Los Altos: Morgan Kaufmann Publishers.
- Kempton, W. (1986). Two theories of home heat control. *Cognitive Science*, 10, 75-90.
- Kieras, D. E., & Polson, P. G. (1985). An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies*, 4, 365-394.
- McNicol, D. (1972). *A primer of signal detection theory*. London: Allen & Unwin.
- Moran, T.P. (1984). Getting into a system: External-internal task mapping analysis. In A. Janda (Ed.), *Human Factors in Computing Systems* (pp. 45-49). Amsterdam: North-Holland.
- Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 1-14). Hillsdale: Erlbaum.
- Oesterreich, R. (1981). *Handlungsregulation und Kontrolle*. München: Urban & Schwarzenberg.
- Pohl, R., Colonius, H., & Thüring, M. (1985). Recognition of script-based inferences. *Psychological Research*, 47, 59-67.
- Rohr, G. (1988). *Grundlagen menschlicher Informationsverarbeitung*. In H. Balzert, U. Hoppe, R. Opperman, H. Peschke, G. Rohr & N.A. Streitz (Hrsg.), *Einführung in die Software-Ergonomie*. Berlin: de Gruyter.
- Shneiderman, B. (1987). *Designing the user interface: Strategies for effective human-computer interaction*. Reading, Mass.: Addison-Wesley.
- Streitz, N. A., & Eberleh E. (1988). *Mentale Belastung und kognitive Prozesse bei komplexen Dialogstrukturen*. Abschlußbericht an die Bundesanstalt für Arbeitsschutz. Institut für Psychologie Nr. 1/48, RWTH Aachen.
- Ulich, E. (1986). Aspekte der Benutzerfreundlichkeit. In W. Remmele & M. Sommer (Hrsg.), *Arbeitsplätze morgen. Berichte des German Chapter of the ACM*, Band 27, S. 102-122. Stuttgart: Teubner.

Dipl.-Psych. Edmund Eberleh  
 RWTH Aachen  
 Institut für Psychologie  
 Jägerstr. 17-19  
 D-5100 Aachen