INTERAKTIONSGRAMMATIK UND KOGNITIVER AUFWAND

Eine Pilotstudie am Beispiel des XS-2 Systems

David Ackermann und Thomas Greutmann, Zürich

Die Analyse der Ueberlegungszeiten der Benutzer eines Computersystems lässt Schlüsse über den kognitiven Regulationsaufwand und die mentale Strukturierung der Aufgabenstellung zu. Es bestehen grosse individuelle Unterschiede in der mentalen Strukturierung und im Lernprozess.

Ausgangslage

Card u. a. (1983) gehen in ihrem GOMS-Modell von der Annahme aus, dass alle Operationen ungefähr die gleiche "Ueberlegungszeit" erfordern. Ihre Annahme belegen sie anhand der Untersuchung einer einfachen Texteditieraufgabe. Auch Kieras & Polson (1984), Polson & Kieras (1984,1985) nehmen in ihrem vieldiskutierten Ansatz an, dass die Bedienung eines Texteditors mit einem invariablen Satz von Produktionsregeln beschrieben werden kann und dass die damit verbundenen Denkprozesse für alle Versuchspartner gleich lang dauern. Würden diese Annahmen zutreffen, so wäre der Grundstein für eine neue Art von Taylorismus im Sinne eines "one best way" gelegt und der Entwurf von Dialogschnittstellen sehr einfach. Jensen (1982) weist aber beispielsweise individuelle Unterschiede in Ueberlegungszeiten nach.

Ausgehend von der Hypothese der semantischen Informationsübertragung von Krause (1982) haben wir an anderer Stelle (Ackermann 1983) die Hypothese formuliert, dass bei Nichtübereinstimmung der realen Aufgabenstruktur mit der mentalen Repräsentation entweder ein zusätzlicher Retransformationsprozess notwendig ist oder Handlungsfehler auftreten. Beides führt zu einem erhöhten kognitiven Regulationsaufwand, der sich in einer Verlängerung der Ueberlegungszeiten nachweisen lassen sollte.

Fragestellung

(1) Es fragt sich nun, wie die zeitliche Struktur der Operationsabfolge als Indikator für die kognitive Strukturierung der Aufgabenstellung und den Verlauf des Lernprozesses erfasst und dargestellt werden kann.

(2) Anhand dieser Darstellung soll die vorstehend erwähnte Hypothese geprüft werden, ob Retransformationsprozesse und Handlungsfehler tatsächlich einen erhöhten kognitiven Regulationsaufwand bewirken.

(3) Ergänzend soll das subjektive Beanspruchungserleben der Versuchspartner im Hinblick auf die Abschätzung der mentalen Beanspruchung erfasst werden.

Versuchsplan und Aufgabenstellung

An den Experimenten zur Untersuchung des XS - 2 - Systems, von denen wir einen Ausschnitt darstellen, nahmen 29 Studenten der Universität (Psychologie / Rechtswissenschaften) und der ETH (Informatik / Elektrotechnik) teil. Die Studenten wurden entsprechend ihrer Studienrichtung und Computererfahrung, wie in Tab. 1 aufgestellt, in Gruppen eingeteilt.

Tab. 1: Gruppeneinteilung Studienrichtung und Computererfahrung der Versuchsparter im XS-2 - Experiment

(1) Computerexperten, meist Informatikstudenten höherer Semester
(2) Informatikstudenten im 1. Semester mit APPLE-MacIntosh-Erfahrung
(3) "Mischgruppe", d.h. Versuchspartner, die nicht in die vorerwähnten Gruppen eingeordnet werden können, die aber über Computererfahrung verfügen
(4) Versuchspartner ohne Computererfahrung, meist Studenten anderer Abteilungen bzw. Fakultäten.

Diese Gruppeneinteilung wurde im Hinblick auf die spätere Auswertung der Ergebnisse vorgenommen, um unterschiedliche Vorerfahrungen bei der Interpretation berücksichtigen zu können. Der zeitliche Ablauf der Versuche ist in Tab. 2 dargestellt.

Der Versuchsleiter überwachte den Arbeitsablauf, beantwortete Fragen und führte Protokoll über spezielle Probleme der Versuchspartner, Fragen und Erklärungen wie auch über eigene Beobachtungen und Eindrücke. Alle Aktionen der Benutzer am XS-2 - System wurden auf Logfiles aufgezeichnet.

Tab. 2: Versuchsablauf

1.	HAKEMP-Fragebogen (Kuhl 1983)	10 min
2.	Bilder 1 (Fragebogen)	15 min
3.	EZ-A Eigenzustandsskala (Nitsch 1974)	3 min
4.	Experimente mit XS-2 (Stelovsky 1984)	120 min
	a) Taschenrechner "TRNorm"	
	b) Autozeichen mit "BILDER ZEICHNEN"	
	c) Taschenrechner "TRPol"	
	d) Spreadsheet: Verbuchen der Autoverkäufe	
	e) Taschenrechner nach Wahl zur Ueberprüfung	
5.	EZ-B Eigenzustandsskala	3 min
6.	Bildschirmzeichnen aus dem Gedächtnis	10 min
7.	Bilder 2 (Fragebogen)	15 min
8.	Interview zur Dialogstruktur	10 min

In dieser Arbeit beschränken wir uns auf die Auswertung der beiden Taschenrechner-Aufgaben (für einen Ueberblick siehe Ackermann (1987) und Greutmann (1986)). Mit beiden Rechnern mussten drei einfache Rechnungen gelöst und die Resultate summiert werden. Für den Rechner "TRNorm" sind die Rechnungen in Tab. 3 dargestellt. Für den zweiten Rechner "TRPol" waren die Aufgaben vergleichbar.

Tab. 3: Die Rechungen für TRNorm

```
22 * 15 + 18
(13 + 5 * 8) / (256 - 300 + 30)
(34 + 17) * (34 + 17) / 113
```

Diese Aufgaben sind "auf einen Blick" leicht zu lösen, für die Arbeit mit den einfachen Rechnern "TRNorm" und "TRPol" muss man sie jedoch entsprechend den Prioritäten der Operatoren umstrukturieren, was allerdings nicht allzu schwierig ist.

Auswertungskonzepte

Wir nehmen an, dass sich in der beobachtbaren zeitlichen Abfolge der Operationen die antizipierte, im Arbeitsgedächtnis gespeicherte Handlungsstruktur abbildet, soweit sie bewusstseinspflichtig ist. Bei gewohnten / gelernten Operationen sind gemäss unseren Annahmen kürzere zeitliche Abstände zwischen den einzelnen Operationen zu erwarten als wenn ein neues Ziel gesetzt und eine neue Operationsfolge entworfen werden muss.

Die Ausführungszeiten einer Handlung verringern sich bekanntlich mit zunehmender Uebung. Das "Power law of practice" (Card u. a., 1983, S. 57) besagt, dass die zur Ausführung einer Operation benötigte Zeit T im n-ten Operationszyklus gemäss einem Potenzgesetz ermittelt bzw. vorhergesagt werden kann:

$$T_n = T_1 n^{-a}$$

Diese mathematische Beziehung soll uns ermöglichen, aus den für die einzelnen Operationen benötigten Zeiten den zu erwartenden Uebungsgewinn unter Berücksichtigung individueller Unterschiede herzuleiten.

Wir können das Vorgehen des Versuchspartners am System in verschiedene Zeitabschnitte gliedern: (1) die Zeit, während der das System ein Kommando ausführt, (2) die Zeit, während welcher der Versuchspartner einen Befehl eingibt und (3) die Zeit, die er offensichtlich zum Ueberlegen seines weiteren Vorgehens braucht. Aus den Logfiles können wir diese Zeitabschnitte genau berechnen. Die "Ausführungszeit" eines Befehls beginnt mit der ersten Interaktion, beispielsweise mit dem Start einer Mausbewegung oder einem Tastendruck. Die Zeitspanne, während welcher der Versuchspartner keine Interaktion vornimmt, bezeichnen wir als "Ueberlegungszeit". Wir nehmen an, dass diese Ueberlegungszeit Informationen über den kognitiven Regulationsaufwand beinhaltet, d. h. diese Zeit muss aufgewendet werden, um die nächsten Operationen zu planen.

Ergebnisse

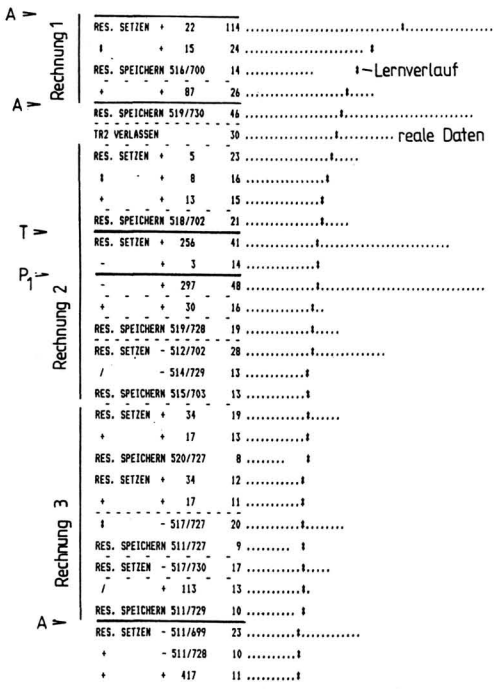
Wir haben die Ueberlegungszeiten der Versuchspartner für die Operanden + - * / (Tastatureingaben TRNorm) und "Operand eingeben" (TRPol) nach dem "Power law of practice" berechnet, um den Verlauf des Lernprozesses darzustellen. In Abb. 1 sind die Lernkurven und die realen Daten für zwei Versuchspartner graphisch dargestellt. Zusätzlich werden links die Nummer der Rechenaufgabe sowie die gewählten Befehle, die verwendeten Speicherzellen und eingegebenen Zahlen angegeben.

Die beiden Lernverläufe weichen deutlich voneinander ab. Kurve (A) ist typisch für die Interaktion eines ungeübten, Kurve (B) für die eines erfahrenen Benutzer. Kurve (A) fällt am Anfang steil ab, bleibt aber auch noch am Ende der Aufgabe auf einem vergleichsweise hohen Wert. Kurve (B) dagegen verläuft flach: Die

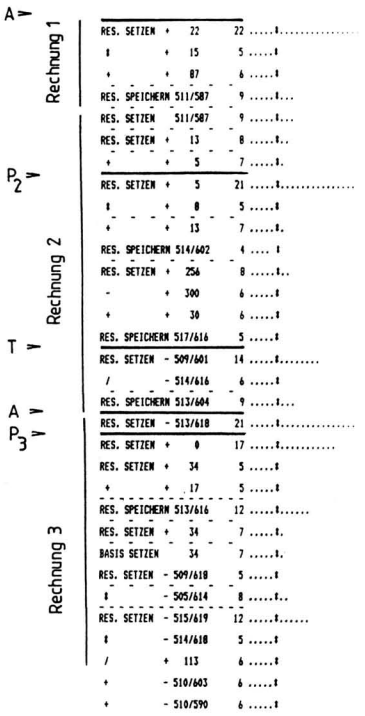
Interaktion findet mit konstanter Geschwindigkeit statt, die Ueberlegungszeiten sind am Anfang und am Ende vergleichsweise kurz.

Abb. 1: Reale Ueberlegungszeiten und gerechnete Lernkurve für zwei Versuchspartner

Versuchspartner (A)



Versuchspartner (B)



Es fällt auf, dass bei beiden Versuchspartnern einige Ueberlegungszeiten deutlich länger sind als die gemäss der errechneten Lernverläufe zu erwartenden Werte. Entsprechend unserer Hypothese sollten solche längeren Ueberlegungszeiten dann auftreten, wenn entweder infolge eines Handlungsfehlers oder einer Fehlinterpretation ein zusätzlicher Retransformationsprozess erforderlich ist und deshalb ein zusätzlicher kognitiver Regulationsaufwand geleistet werden muss oder die Aufgabenstellung für die nächste Handlungssequenz mental zu strukturieren ist. Ueberlegungszeiten, welche grösser sind als die gemäss der Lernverläufe berechneten Werte, sind in Abb. 1 im Handlungsablauf

entsprechend markiert. Dabei wurde nach dem Betrag der Abweichung unterschieden, wobei wie folgt klassiert wurde (immer nur Abweichungen nach oben): (1) Abweichungen kleiner als 50 %, (2) Abweichungen zwischen 50 und 100% und (3) Abweichungen grösser als 100%. Diese Abweichungen sind durch entsprechend dicht ausgelegte Strichlinien in der Handlungsabfolge markiert. Durchgezogene Linien bedeuten Abweichungen grösser als 100%. Diese Strichlinien werden wir im folgenden als Ueberlegungszeitschwellen (Ackermann 1987) bezeichnen.

Ueberlegungszeiten, die weniger als 50% über den vom Lernverlauf "vorausgesagten" Zeiten liegen, können durch die "zufällige" Verteilung der Ueberlegungszeiten erklärt werden. Auch im Bereich 50 - 100 % können "Zufallseinflüssen" nicht ausgeschlossen werden. Ueberschreitet aber die reale Ueberlegungszeit den mit Hilfe des "Power law of practice" berechneten Wert um mehr als 100%, so kann dies nicht mehr durch Zufall erklärt werden. Die in Abb. 1 dargestellten langen Ueberlegungszeiten sind also nicht zufällig zustande gekommen. Die mit A bzw. mit T markierten Ueberlegungszeitschwellen sind auf den Anfang einer Aufgabe bzw. einer Teilaufgabe zurückzuführen. Die Aufgabe / Teilaufgabe muss mental strukturiert werden. Die mit P_1, P_2 und P_3 bezeichneten Ueberlegungszeitschwellen lassen sich durch Interaktionsfehler der Versuchspartner erklären. Die Planung der Korrekturen bewirkt gemäss unserer Hypothese einen erhöhten kognitiven Aufwand. Bei P_1 wurde eine falsche Zahl eingetippt, bei P_2 wurde die Aufgabe falsch strukturiert (Multiplikationen müssen zuerst gerechnet werden), so dass neu begonnen werden musste. Vor der Ueberlegungszeitschwelle P_3 wurde irrtümlicherweise ein falscher Befehl mit der Maus angewählt.

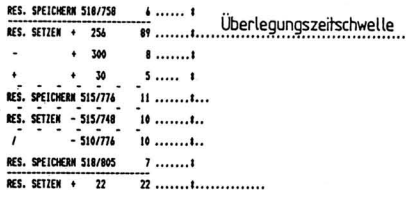
Die Analyse der Ueberlegungszeiten lässt nicht nur Schlüsse über erhöhten kognitiven Regulationsaufwand, sondern auch über die kognitive Strukturierung der Aufgabenstellung zu. Dies wird in Abb. 2 verdeutlicht.

Abb. 2 zeigt denselben Ausschnitt aus der Taschenrechneraufgabe für zwei verschiedene Versuchspartner. Beide Personen verwenden genau die gleichen Befehle, um die Aufgabe zu lösen. Die kognitive Strukturierung ist jedoch verschieden, was man aus den unterschiedlich langen Ueberlegungszeitschwellen sehen kann. Der wesentliche Unterschied besteht darin, dass Ver-

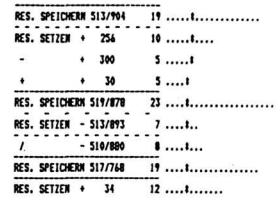
suchspartner X das Abarbeiten eines Teils der Berechnung und das anschliessende Zwischenspeichern des Resultats als eine zusammengehörige Einheit plant. Versuchspartner Y hingegen plant zuerst nur das Abarbeiten der Berechnung. Vor dem Abspeichern des Zwischenresultats benötigt er dann eine lange Ueberlegungszeit, offenbar um das Abspeichern zu planen. Die langen Ueberlegungszeiten sind vor jedem Speicherbefehl ("RES. SPEICHERN") festzustellen.

Abb. 2: Beispiel für unterschiedliche kognitive Strukturierung einer identischen Abfolge elementarer Interaktionen

Struktur X



Struktur Y



Um allfällige wechselseitige Beziehungen zwischen kognitiven Regulationsaufwand und subjektivem Beanspruchungserleben nachzuweisen, wurde versucht, das subjektive Beanspruchungserleben mit Hilfe der Eigenzustandsskala EZ von Nitsch (1974) festzuhalten. Bei fast allen Versuchspartnern sank die Grösse "Beanspruchung" sensu Nitsch, was eher auf die Dauer des Experiments (2 Stunden) zurückgeführt werden muss.

Fazit

(1) Die Ergebnisse stützen die Ausgangshypothese, dass bei Nicht-Uebereinstimmung der realen Aufgabenstruktur mit der mentalen Repräsentation entweder ein zusätzlicher Retransformationsaufwand notwendig ist oder Handlungsfehler auftreten, welche einen erhöhten kognitiven Regulationsaufwand bewirken.

(2) Die Analyse der Ueberlegungszeiten lässt Schlüsse auf die kognitive Strukturierung der Aufgabe zu. Hier bestehen deutliche interindividuelle Unterschiede, selbst bei gleichen Folgen von elementaren Interaktionen. Die Lernprozesse verlaufen eben-

falls individuell unterschiedlich. Gemäss dem Prinzip der differentiellen Arbeitsgestaltung von Ulich (1978, 1983) müssen solche interindividuellen Unterschiede bei der Gestaltung von Dialogsystemen berücksichtigt werden, um eine optimale Interaktion und Entwicklung der Persönlichkeit zu gewährleisten. Ergebnisse experimenteller Untersuchungen weisen darauf hin, dass die entsprechende Wahlmöglichkeiten enthaltende Dialogsysteme effizienter sind als vorgegebene (Ackermann & Nievergelt, 1985; Ackermann 1986). Individuellen Unterschieden kann z.B. dadurch Rechung getragen werden, dass ein Sortiment von verschiedenen Dialogstrukturen angeboten wird, aus dem sich dann jeder Benützer die von ihm bevorzugte Struktur zusammenstellen kann.

Herrn Prof. Dr. E. Ulich danken wir für die ausführliche Diskussion dieser Arbeit und die Durchsicht des Manuskripts.

Literatur

Ackermann, D. (1983): Robi Otter oder die Suche nach dem operativen Abbildsystem. Interner Bericht der Studienarbeiten im SS 1983.

Ackermann, D. & Nievergelt, J. (1985): Die Fünffingermaus - eine Fallstudie zur Synthese von Hardware, Software und Psychologie. In: Software-Ergonomie '85. Mensch-Computer-Interaktion. Stuttgart, B.G. Teubner.

Ackermann, D. (1986): A pilot study on the effects of individualization in man-computer-interaction. In: Mancini, G., Johannsen, G. & Martennson, L. (Hrsg.): Analysis, Design and Evaluation of Man-Machine Systems Proceedings of the 2nd IFAC/IFIP/IFORS/IEA Conference, Varese, September 1985. London: Pergamon Press, S. 293-297

Ackermann, D. (1987): Handlungsspielraum, Mentale Repräsentation und Handlungsregulation am Beispiel der Mensch-Computer-Interaktion. Inauguraldissertation der Philosophisch-historischen Fakultät der Universität Bern.

Card, S.K.; Moran, T.P. & Newell, A. (1983): The Psychology of Human-Computer Interaction. London & Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Greutmann, T. (1986): Task structure, dialog grammar and action grammar of the user, illustrated with the interactive system XS-2. Unveröffentlichte Diplomarbeit in Informatik, ETH Zürich.

Jensen, A.R. (1982): Reaction time and psychometric g. In: Eysenck, H. (Hrsg.): A model for intelligence. Berlin, Springer.

Kieras, D.E. & Polson, P.G. (1985): An approach to the formal analysis of user complexity. Int. Journal of Man-Machine Studies, 22 (1985).

Krause, B. (1982): Semantic information processing in cognitive processes. Zeitschrift für Psychologie, 190(1) (1982).

Kuhl, J. (1983): Motivation, Konflikt und Handlungskontrolle. Berlin, Springer.

Nitsch, J.R. (1974): Die hierarchische Struktur des Eigenzustandes - ein Approximationsversuch mit Hilfe der Binärstrukturanalyse (BISTRAN). DIAGNOSTICA, 20 (1974).

Polson, P.G. & Kieras, D.E. (1984): A formal description of user's knowledge of how to operate a device and user complexity. In: Behaviour Research Methodes, Instruments & Computers, 16(2) (1984).

Polson, P.G. & Kieras, D.E. (1985): A quantitative model of the learning and performance of text editing knowledge. CHI 85 Proceedings, S. 205-212.

Stelovsky, J. (1984): XS-2: The User Interface of an Interactive System. Diss. ETH No. 725, Zürich.

Ulich, E. (1978): Ueber das Prinzip der differentiellen Arbeitsgestaltung. Industrielle Organisation, 47 (1978).

Ulich, E. (1983): Differentielle Arbeitsgestaltung - ein Diskussionsbeitrag. Zeitschrift für Arbeitswissenschaft, 1 (1983).

David Ackermann
Thomas Greutmann
Lehrstuhl für Arbeits- und Organisationspsychologie
ETH Zürich
Nelkenstr. 11
CH - 8092 Zürich