

BEURTEILUNG VON BILDSCHIRMMASKEN ANHAND GEOMETRISCHER PARAMETER

Peter Elsinger, Werner Graf, Loretta Merz,  
Helmut Krueger, Zürich

**Zusammenfassung:** Die Resultate von Tullis (1984) konnten bestätigt werden. Es besteht eine hohe Korrelation zwischen geometrischen Parametern einer Bildschirmmaske und den Suchzeiten eines Objekts auf der Maske sowie der subjektiven Beurteilung durch die Probanden. Darüber hinaus konnte gezeigt werden, dass die durch Aufzeichnen der Blickbewegungen der Probanden erhaltenen Saccadendistanzen und Fixationszeiten ebenfalls hoch mit den geometrischen Parametern korrelieren. Die Blickbewegungen können aber für die ergonomische Forschung auch Aufschluss über den genauen Ort der Probleme mit einer Maske geben.

Einleitung

Durch den Einsatz von Computern wurde der Mensch in vielen Fällen von belastenden Umwelteinflüssen befreit; aber es entstanden neue Belastungen, weil ergonomische und organisatorische Aspekte nicht berücksichtigt wurden. Für die ergonomischen Aspekte der Arbeitsplatzumgebung und die Apparaturen existiert heute eine breite Literatur (z.B. Grandjean und Vigliani, 1980), die Ergonomie der Software steht jedoch noch in den Anfängen. Der grösste Teil der Arbeit am Bildschirm wickelt sich in Dialogform zwischen Mensch und Maschine ab. Ein Teilbereich des Dialogs sind die Bildschirmmasken, die Gegenstand dieser Arbeit sind (s. auch Elsinger, 1986). Für die Praxis ist es wichtig, dass sich ein Instrument finden lässt, das die Qualität von Bildschirmmasken objektiv beurteilt, ohne dass zeitaufwendige Experimente mit Probanden durchgeführt werden müssen.

Die geometrischen Masse

Tullis (1983a und 1983b) hat in seiner Literaturarbeit durch Anwendung verschiedener Kriterien vier grundlegende Merkmale für alphanumerische Bildschirmdarstellungen bestimmt, wobei

diese nicht notwendigerweise voneinander unabhängig sind und deren Herleitung nicht immer ein wohldefinierter Prozess waren:

Gesamtdichte: Sie ist der Prozentteil der aktivierten Bildschirmfläche, d.h. die Zahl der Buchstaben auf dem Bildschirm dividiert durch die maximal mögliche Zahl der Buchstaben.

Lokale Dichte: Das lokale Gewicht eines Buchstaben wird definiert durch die Summe der Gewichte der Buchstaben in der Nähe, d.h. innerhalb eines Kreises, dessen Durchmesser durch einen Blickwinkel von 5 Grad bestimmt ist, wobei das Gewicht der umliegenden Buchstaben mit dem Radius proportional abnimmt, dividiert durch das Gewicht des Kreises, d.h. das maximal mögliche Gewicht eines Buchstaben. Die lokale Dichte ist das Mittel der lokalen Gewichte aller Buchstaben.

Gruppenbildung: Um die Gruppen zu bestimmen, verwendet Tullis einen vereinfachten Algorithmus von Zahn (1971), der von den Gesetzen der Gestaltpsychologen ausging. Dadurch erhält Tullis die Anzahl der Gruppen sowie ihre durchschnittliche Grösse, wobei letztere mit der Buchstabenanzahl einer Gruppe gewichtet wurde.

Layout-Komplexität: Als Items betrachtet Tullis die festen und die variablen Felder einer Maske. Aus den Abständen der Items vom linken bzw. oberen Rand des Bildschirms erhält er durch Anwendung der Formel aus der Informationstheorie die Itemunsicherheit, die multipliziert mit der Anzahl Items die Layout-Komplexität ergibt.

Seine theoretischen Überlegungen versuchte Tullis (1983b und 1984) in einem Experiment zu verifizieren. Er verwendete je 26 verschiedene Formate (gleiche Form einer Maske) von Fluglinien- und Hotellisten mit je zehn verschiedenen Datensätzen. Jeder Versuchsdurchgang bestand darin, nach einem Objekt in einem Datensatz zu suchen. Ausserdem musste der Proband für jedes Format angeben, wie leicht es zu benutzen war. Tullis mittelte die Suchzeiten der korrekt beantworteten Versuchsdurchgänge für jedes Format über alle Probanden sowie die subjektive Beurteilung über alle Probanden. Dann berechnete er die multiple Regression für die geometrischen Masse und die Suchzeit, wobei er ein  $R$  von 0.71 erhielt, sowie diejenige für die Masse und die subjektive Beurteilung mit einem  $R$  von 0.90.

### Ziel der Arbeit

In dieser Arbeit sollten die Resultate von Tullis in Form eines Pilotversuchs verifiziert werden. Zusätzlich könnten durch die Analyse des Blickverhaltens Aussagen über die mentale Beanspruchung, ein für die Ergonomie wichtiger Parameter, gemacht werden. Deshalb wurden die Blickbewegungen ebenfalls registriert, wobei die Fixationsdauer und die Saccadenlänge ausgewertet wurden. Als Bildschirmmasken wurden diejenigen von Haubner (1985) verwendet.

Suchzeit: Die Suchzeit wurde wie bei Tullis als die Zeit der Präsentation der Bildschirmmaske gemessen. Die Auswahl der Bildschirmmasken bringt es mit sich, dass sowohl ein- als auch mehrdimensionale Suchobjekte vorkommen, so dass eine getrennte Analyse notwendig war (Teichner und Mocharnuk, 1979).

Subjektive Beurteilung: Da eine Frage für eine subjektive Beurteilung einer Bildschirmmaske sehr wenig ist, wurden in dieser Arbeit jeweils nach einem Format fünf Fragen gestellt.

Augenbewegungen: Die Augenbewegungen wurden bisher im Zusammenhang mit Bildschirmmasken kaum untersucht. Es wird jedoch immer evident, dass diese beim Lesen die Charakteristika des Textes widerspiegeln (Dee-Lucas, Just, Carpenter und Dane-man, 1982). Lévy-Schoen (1981 und 1982) erhärtete in einer Reihe von Experimenten ihre Annahme, dass den Augenbewegungen ein aufgabenorientiertes Programm zugrunde liegt, das während der Ausführung der Aufgabe durch lokale Anpassungen variiert wird, was sich in den Fixationszeiten und den Saccadendistanzen niederschlägt. Dabei werden die lokalen Anpassungen immer wichtiger, je freier sich das Auge bewegen kann. So stellt sie z.B. beim Absuchen einer Zeile nach Suchobjekten einen Zusammenhang zwischen Fixationszeit und Saccadendistanz fest, der beim Lesen von Texten nicht gefunden werden kann. Möglicherweise tritt dieser Zusammenhang desto eher auf, je schlechter eine Bildschirmmaske gestaltet ist, da dann viel intensiver und ohne gute Anhaltspunkte aus dem Layout gesucht werden muss.

Da Antes (1974) bei der Analyse der Blickdaten beim Betrachten von Bildern einen zeitlichen Verlauf für die Fixationszeiten (Zunahme) und Saccadendistanzen (Abnahme) feststellte, wurde dies bei der Auswertung der Blickdaten berücksichtigt.

### Methode

Zunächst wurden die geometrischen Masse implementiert, die Tullis (1983) zusammengestellt hat. Insgesamt wurden zehn verschiedene Formate verwendet, vier von Haubner (1985), sechs selbsterstellte. Für jedes Format wurden zehn Datensätze aus dem Kfz-Handel gezeigt.

Die Arbeit wurde auf einem IBM Personal Computer AT durchgeführt, ausgerüstet mit einem hochauflösenden, monochromen Monitor mit weissem Phosphor. Der Versuch wurde in einem schallisolierten, klimatisierten Raum durchgeführt, beleuchtet mit Fluoreszenzlampen an 3 Phasen. Die Blickbewegungen wurden mit einem G+W Eye-Trac, Modell 200, der Applied Science Laboratories registriert und vom Computer gespeichert. Am Versuch nahmen 10 Personen teil, die alle die erforderliche Sehschärfe von 0.8 erreichten.

Der Versuch wurde in zwei Teilen an zwei aufeinanderfolgenden Tagen durchgeführt. Jeder Versuchsdurchgang gestaltete sich wie folgt: Zunächst wurde eine Suchfrage präsentiert, nach einem Tastendruck des Probanden die Maske. Hatte der Proband die Antwort gefunden, drückte er wieder eine Taste, worauf zwei Antwortmöglichkeiten erschienen. Die Suchzeit wurde als die Zeit zwischen den beiden Tastendruck gemessen. Nach den zehn Beispielen eines Formats folgte die subjektive Beurteilung. Im zweiten Teil wurde die Zahl der Datensätze pro Format auf zwei reduziert und dafür die Augenbewegungen registriert. Wegen des zu erwartenden starken Lerneffekts wurde die Reihenfolge der Formate zwar für jeden Probanden gleich gewählt, jedoch zyklisch um eines vertauscht. Jeder Proband erhielt zu jeder Maske die gleichen Suchfragen. Die Reihenfolge der Masken zum selben Format wurde zufallsmässig gewählt.

### Resultate und Diskussion

#### Suchzeit

Zunächst wurde das Mittel der Suchzeit über die Probanden für jedes Format berechnet, getrennt nach Experiment und Fragenkomplexität (ein- bzw. mehrdimensionale Fragen), wobei nur die richtig beantworteten Versuchsdurchgänge berücksichtigt wurden. Die multiplen Regressionen mit den geometrischen Massen ergaben, wie Tabelle 1 zeigt, durchwegs höhere Werte für  $R$  als bei

Tullis (1984), waren jedoch nicht signifikant, was mindestens teilweise auf die geringe Datenmenge zurückzuführen sein dürfte. Die Werte sind im zweiten Experiment ähnlich wie im ersten, trotz des Lerneffekts, d.h. das Experiment ist reproduzierbar.

Tab. 1. Multiple Korrelationen zwischen den sechs geometrischen Massen und der Suchzeit im ersten und zweiten Experiment, getrennt nach der Fragenkomplexität

	Experiment 1		Experiment 2	
	eindim. Fragen	mehrdim. Fragen	eindim. Fragen	mehrdim. Fragen
R	0.90	0.81	0.76	0.93
Signifikanz	0.28	0.56	0.68	0.19

Die im Experiment verwendeten Masken wurden ebenfalls mit dem käuflich erworbenen Programm von Tullis ausgewertet. Die erhaltenen Korrelationen waren den oben dargestellten Werten sehr ähnlich, obwohl sich vor allem die Werte für die Anzahl Items und die Itemunsicherheit deutlich unterschieden. So erstaunt es nicht, dass die von Tullis vorhergesagten Suchzeiten mit den gemessenen kaum korrelierten. Das Modell von Tullis konnte deshalb nicht bestätigt werden.

#### Subjektive Beurteilung

Die Antworten auf die fünf subjektiven Fragen wurden ebenfalls wie bei der Suchzeit gemittelt. Die multiplen Regressionen ergaben wieder durchwegs höhere Werte für  $R$  als bei Tullis (Tabelle 2), waren jedoch für die eher objektiven Fragen 2 (Information leicht / schwierig zu finden) und 3 (zu viele Daten auf dem Bildschirm) signifikant, nicht aber für die Befindlichkeitsfragen 4 (Ermüdung) und 5 (ärgerlich / erfreulich). Dass

Tab. 2. Multiple Korrelationen zwischen den sechs geometrischen Massen und der subjektiven Beurteilung

	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5
R	0.95	0.97	0.99	0.93	0.95
Signifikanz	0.11	0.05	0.02	0.19	0.12

Frage 1 (Struktur übersichtlich / unübersichtlich) nicht signifikant korreliert, kann wahrscheinlich damit erklärt werden, dass die Erfahrung der Probanden mit Bildschirmarbeiten von "sehr grosse Erfahrung" bis "keine Erfahrung" reichte.

Zwischen der vorhergesagten subjektiven Beurteilung durch das Programm von Tullis und der Antwort auf die Frage 2, welche etwa der Frage von Tullis (1984) entsprach, ergab sich keine Korrelation (s. auch Suchzeit).

### Blickbewegungen

Die Saccadendistanzen und die Fixationszeiten wurden automatisch ausgewertet. Zwischen diesen konnten keine Korrelationen festgestellt werden. Offensichtlich ist das Programm, welches der Suchaufgabe zugrunde liegt, sehr flexibel und nicht nur von der Verarbeitungszeit abhängig (Lévi-Schoen, 1982).

Um einen möglichen zeitlichen Verlauf zu erkennen, wurden die Versuchsdurchgänge in drei gleich grosse Teile aufgeteilt (s. Antes, 1974). Das Mittel der Saccadendistanzen und der Fixationszeiten über die Probanden wurde für jedes Format berechnet, getrennt nach Fragenkomplexität (ein- bzw. mehrdimensionale Fragen) und nach Drittel des Versuchsdurchgangs, wobei nur die richtig beantworteten Versuchsdurchgänge berücksichtigt wurden (siehe auch Suchzeit). In Abbildung 1 und 2 sind die Saccadendistanzen und in Abbildung 3 und 4 die Fixationszeiten für die einzelnen Drittel und die verwendeten Formate aufgetragen. Wie der Wilcoxon-Test zeigt, nehmen die Saccadendistanzen bei den mehrdimensionalen Fragen vom ersten zum letzten Drittel signifikant ab (Tabelle 3) und die Fixationszeiten ebenfalls bei den mehrdimensionalen Fragen vom ersten zum letzten und auch vom zweiten zum letzten Drittel signifikant zu (Tabelle 4). Für die eindimensionalen Fragen scheint die mentale Belastung noch nicht genügend hoch zu sein, so dass dieser Effekt durch eine hohe Streuung verwischt wurde. Dass bei den eindimensionalen Fragen die Fixationszeiten vom ersten zum letzten Drittel zunahm, dürfte eher mit der Aufgabe, einer Suchaufgabe, zusammenhängen.

Die multiplen Korrelationen zwischen den geometrischen Massen und den Saccadendistanzen (Tabelle 5) nehmen mit jedem Drittel zu und sind im letzten Drittel höher als für die Suchzeit. Sie sind nicht signifikant, aber im letzten Drittel für

Abb. 3. Fixationszeiten für eindimensionale Fragen

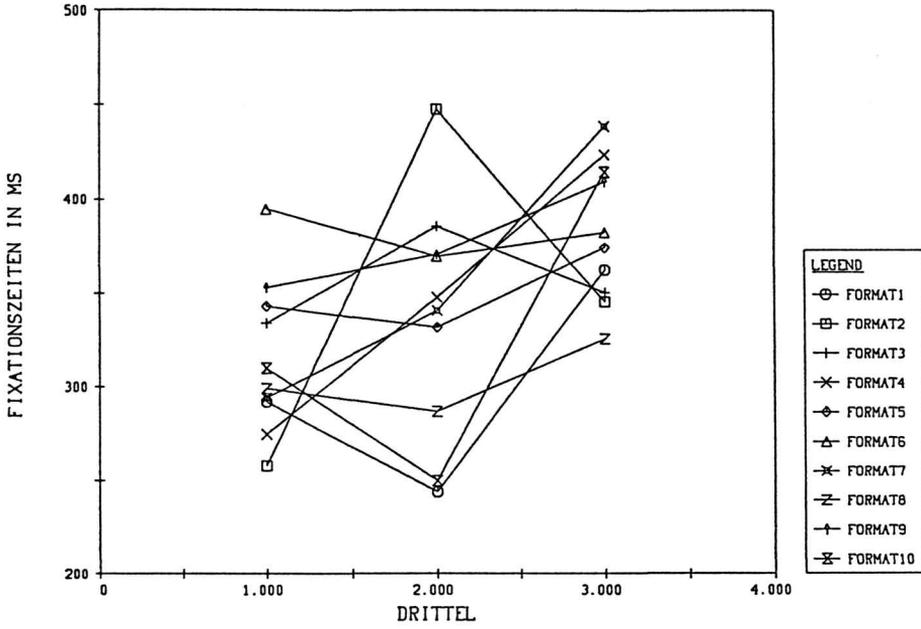


Abb. 4. Fixationszeiten für mehrdimensionale Fragen

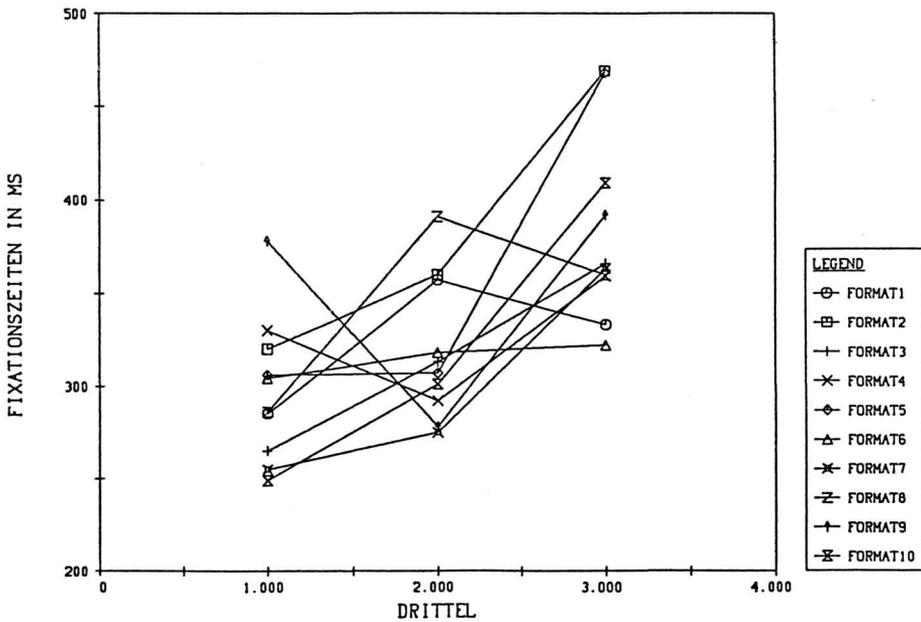


Abb. 1. Saccadendistanzen für eindimensionale Fragen

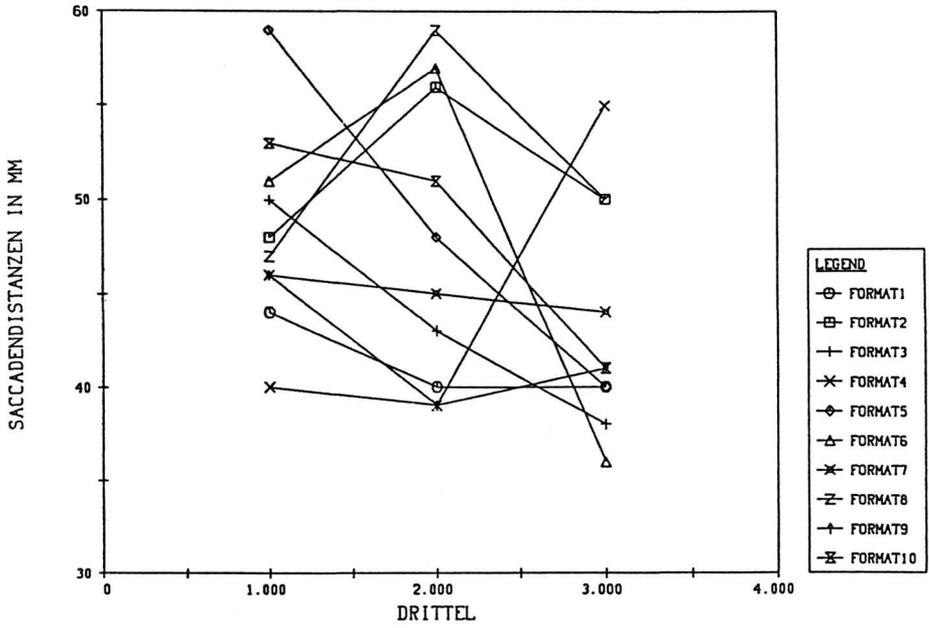
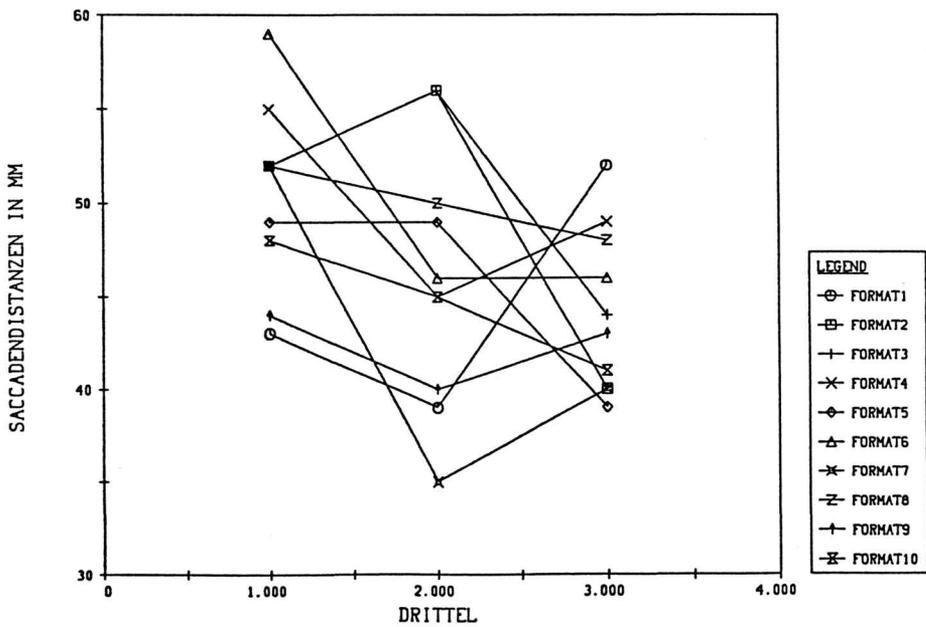


Abb. 2. Saccadendistanzen für mehrdimensionale Fragen



Tab. 3. Signifikanzen nach Wilcoxon-Test für die Abnahme der Saccadendistanzen

	Eindim. Fragen	Mehrdim. Fragen
Erstes zu zweitem Drittel	$\underline{p} = 0.646$	0.110
Zweites zu drittem Drittel	$\underline{p} = 0.139$	0.636
Erstes zu drittem Drittel	$\underline{p} = 0.139$	0.028

Tab. 4. Signifikanzen nach Wilcoxon-Test für die Zunahme der Fixationszeiten

	Eindim. Fragen	Mehrdim. Fragen
Erstes zu zweitem Drittel	$\underline{p} = 0.508$	0.139
Zweites zu drittem Drittel	$\underline{p} = 0.074$	0.022
Erstes zu drittem Drittel	$\underline{p} = 0.007$	0.005

die mehrdimensionalen Fragen nicht mehr weit davon entfernt. Der zeitliche Verlauf ist interessant, jedoch ist eine Interpretation schwierig, da die einfachen Korrelationen sehr uneinheitlich sind (Elsinger, 1986).

Tab. 5. Multiple Korrelationen zwischen den sechs geometrischen Massen und den Saccadendistanzen

	Eindimensionale Fragen			Mehrdimensionale Fragen		
	Erstes Drittel	Zweites Drittel	Drittes Drittel	Erstes Drittel	Zweites Drittel	Drittes Drittel
$R$	0.57	0.65	0.92	0.72	0.84	0.95
Signifikanz	0.93	0.86	0.22	0.76	0.48	0.10

Bei den multiplen Korrelationen zwischen den geometrischen Massen und den Fixationszeiten (Tabelle 6) ist das Bild uneinheitlicher. Sie fallen für die eindimensionalen Fragen über die Drittel leicht, steigen aber eher für die mehrdimensionalen Fragen. Die Korrelationen sind sehr hoch, jedoch nicht signifikant.

Eine weitere Auswertemethode für die Blickbewegungen wäre eine Häufigkeitsanalyse, d.h., die Fixationszeiten und Saccadendistanzen werden in Intervalle aufgeteilt. Es ist nämlich

Tab. 6. Multiple Korrelationen zwischen den sechs geometrischen Massen und den Fixationszeiten

	Eindimensionale Fragen			Mehrdimensionale Fragen		
	Erstes Drittel	Zweites Drittel	Drittes Drittel	Erstes Drittel	Zweites Drittel	Drittes Drittel
R	0.94	0.90	0.82	0.59	0.87	0.78
Signifikanz	0.14	0.29	0.52	0.92	0.39	0.64

denkbar, dass der Mittelwert nicht für sinnvolle Aussagen geeignet ist, falls kurze und lange Zeiten bzw. Distanzen ihre Verteilung über die Formate ändern.

Aufgrund der hohen multiplen Korrelationen der Fixationszeiten bzw. Saccadendistanzen mit den geometrischen Parametern kann gesagt werden, dass die Blickbewegungen für die ergonomische Forschung ein ebenso geeignetes Hilfsmittel wie die Suchzeiten sind, um Aussagen über die Maskengestaltung machen zu können. Darüber hinaus erhält man durch die Analyse der Blickbewegungen Informationen über den Suchprozess selbst, so dass man bestimmen kann, an welcher Stelle der Maske Probleme auftreten.

#### Literatur

- Antes, J.R. (1974): The Time Course of Picture Viewing. *Journal of Experimental Psychology*, 103, 62-70.
- Dee-Lucas, D., Just, M.A., Carpenter, P.A., and Daneman, M. (1982): What Eye Fixations Tell Us About the Time Course of Text Integration. In: R. Groner and P. Fraisse (Eds.), *Cognition and Eye Movements*. Amsterdam, New York, Oxford: North-Holland Publishing Company.
- Elsinger, P. (1986): Beurteilung von Bildschirmmasken anhand geometrischer Parameter. Diplomarbeit, Eidgenössische Technische Hochschule Zürich, Zürich.
- Grandjean, E., and Vigliani, E. (1980): *Ergonomic Aspects of Visual Display Terminals*. London: Taylor & Francis Ltd.
- Haubner, P. (1985): Strukturaspekte der Informationsgestaltung auf Bildschirmen. In: H.-W. Bodmann (Ed.), *Aspekte der Informationsverarbeitung*. Berlin, Heidelberg, New York: Springer-Verlag.
- Lévy-Schoen, A. (1981): Flexible and/or Rigid Control of Oculo-

- motor Scanning Behavior. In: D.F. Fisher, R.A. Monty, and J.W. Senders (Eds.), *Eye Movements: Cognition and Visual Perception*. Hillsdale, N.J.: Erlbaum.
- Lévy-Schoen, A. (1982): Perception, Exploration and Eye Displacements. In: R. Groner, and P. Fraisse (Eds.), *Cognition and Eye Movements*. Amsterdam, New York, Oxford: North-Holland Publishing Company.
- Teichner, W.H., and Mocharnuk, J.B. (1979): Visual Search for Complex Targets. *Human Factors*, 21, 259-275.
- Tullis, T.S. (1983a): The Formatting of Alphanumeric Displays: A Review and Analysis. *Human Factors*, 25, 657-682.
- Tullis, T.S. (1983b): Predicting the Usability of Alphanumeric Displays. Teil-Dissertation, Rice University, Houston, Texas.
- Tullis, T.S. (1984): A Computer-Based Tool for Evaluating Alphanumeric Displays. *Interact*, Conference Paper.
- Zahn, C.T. (1971): Graph-Theoretical Methods for Detecting and Describing Gestalt Clusters. *IEEE Transactions on Computers*, 20, 68-80.

Peter Elsinger  
 CAP Gemini (Schweiz) AG  
 Geschäftsstelle Finanz- und  
 Dienstleistungsunternehmen  
 Brauerstrasse 60  
 CH-8004 Zürich

Werner Graf  
 Dr. Loretta Merz  
 Prof. Dr. Dr. Helmut Krueger  
 Eidgenössische Technische  
 Hochschule Zürich  
 Institut für Hygiene  
 Arbeitsphysiologie  
 ETH-Zentrum  
 CH-8092 Zürich