
Eye-Tracking: Grundlagen und Anwendungsfelder

Christopher Blake

Abstract Eye-Tracking ist ein apparatives, rezeptionsbegleitendes Verfahren zur Ermittlung der Blickrichtung von Personen. Es ermöglicht genaue Aussagen darüber, welche Medieninhalte Rezipientinnen und Rezipienten wie lange und in welcher Reihenfolge betrachten. Dieser Beitrag veranschaulicht die Funktionsweise des Verfahrens anhand relevanter physiologischer und technologischer Grundlagen. Unterschiedliche aktuelle Gerätetypen (stationäre und mobile Systeme) werden vorgestellt und in Bezug auf ihre jeweiligen Vor- und Nachteile für die Verwendung im Rahmen kommunikationswissenschaftlicher Studien betrachtet. Der Beitrag stellt grundlegende qualitative und quantitative Auswertungsmöglichkeiten von Blickaufzeichnungsdaten vor und setzt sich mit der Frage auseinander, unter welchen Bedingungen Blickaufzeichnungsdaten auch als Indikatoren kognitiver Prozesse sowie ausgewählter Stimulus- und Rezipienteneigenschaften genutzt werden können. Eine Betrachtung kommunikationswissenschaftlicher Anwendungsfelder der Methode sowie anwendungsbezogener Herausforderungen beschließen den Beitrag.

Schlagwörter Blickaufzeichnung, apparative Verfahren, rezeptionsbegleitende Beobachtung, Selektionsforschung

1 Einführung

Eine gute Platte, eine Sendung im Radio und seit neuerem auch eine größer werdende Auswahl von Hörbüchern und gesprochenen Zeitungsartikeln – danach wird es eng: Die Vielfalt medialer Angebote, die sich ohne Beteiligung des Auges nutzen lassen, ist überschaubar. Die meisten traditionellen wie auch modernen Medienangebote sind visueller oder audiovisueller Natur. Das Auge rezipiert mit. Wann immer Informationen durch Medien nicht nur auditiv, sondern auch visuell vermittelt werden, wird das Auge zum maßgeblich an der Mediennutzung beteiligten Wahrnehmungsorgan. Es setzt Se-

lektionsvorgänge um und versorgt die Rezipientinnen und Rezipienten mit dem sensorischen Input, der beim Lesen der Bibel gleichermaßen entsteht wie beim Spielen eines Ego-Shooters am heimischen PC. Wenn wir Medien nutzen, dann lesen wir Texte, betrachten Bilder und Filme oder das visuell aufbereitete Ergebnis unserer Handlungen in interaktiven Spielen. Die große Rolle, die der optischen Wahrnehmung beim Zugang zu Medieninhalten zukommt, begründet die Relevanz des in diesem Kapitel vorgestellten methodischen Ansatzes (vgl. auch Bente 2004).

Verfahren zur Blickaufzeichnung bieten uns als Medienforscherinnen und Medienforschern die Möglichkeit eines Perspektivenwechsels: Wir können uns bildlich ausgedrückt in Rezipientinnen und Rezipienten hineinversetzen und durch ihre Augen sehen, welche Inhalte sie wie lange und in welcher Reihenfolge betrachten. Mehr noch: Blickaufzeichnungsdaten werden auch nutzbar gemacht, um Einblicke in kognitive und emotionale Prozesse zu gewinnen, die sich während des Betrachtens spezifischer Inhalte abspielen. Durch Blickaufzeichnungsdaten können Forscherinnen und Forscher also nicht nur durch die Augen, sondern – wenngleich in sehr beschränktem Maße – auch in den Kopf von Rezipientinnen und Rezipienten schauen. So zumindest kann das Ziel bei der Entwicklung weiterführender Indikatoren aus den im Rahmen von Blickaufzeichnungsstudien anfallenden Daten beschrieben werden. Die Validität entsprechender Maße wird im Rahmen dieses Beitrages zu diskutieren sein. Zunächst wird jedoch ein Einblick in die physiologischen und technologischen Grundlagen der Datenerhebung durch das Eye-Tracking (ET) gegeben um anschließend einen Eindruck von den Einsatzmöglichkeiten und Erkenntnispotenzialen zu vermitteln, die die Methode speziell im Kontext kommunikationswissenschaftlicher Forschung mit sich bringt. Die Besonderheiten bei der Analyse der prozessbezogenen Blickaufzeichnungsdaten werden ebenfalls erörtert. Eine Betrachtung zukünftiger Herausforderungen im Rahmen der Anwendung des Verfahrens beschließt den Beitrag.

2 Grundlagen der visuellen Wahrnehmung

Das menschliche Auge wird oft als die „schlechteste Kamera der Welt“ bezeichnet. Es ist mit vielen konstruktionsbedingten Mängeln behaftet, die dazu führen, dass wir bei der Betrachtung unserer Umwelt Verzerrungen und Unschärfen unterschiedlicher Art in Kauf nehmen müssen (Duchowski 2007: 18). Ein besonderer Mangel, der gewissermaßen die Basis der Funktionsweise des ET darstellt, ist die ungleichmäßig verteilte Auflösung unseres Sehvermögens: Während mit einer Kamera aufgenommene Fotos sowohl in der Mitte als auch am Rand gleichermaßen „scharf“, d. h. gleichermaßen hochauflösend sind, ist dies bei der Abbildung visueller Reize auf der Netzhaut des menschlichen Auges nicht der Fall (Nielsen & Pernice 2010: 6). Die Dichte der für das Sehen bei Tageslicht und das Sehen in Farbe relevanten Photorezeptoren (sogenannte Zapfen) ist nur in einem bestimmten Areal der Netzhaut so hoch, dass darauf einfallendes Licht in ein de-

tailreiches Bild umgesetzt werden kann. Dieses Areal wird als *Fovea Centralis* oder auch als *Sehgrube* bezeichnet. Die grob davon zu unterscheidende *Netzhautperipherie* hat nur ein vergleichsweise geringes Auflösungsvermögen (Joos, Rötting & Velichkovsky 2003: 143). Es werden hier nur ca. 15 bis 20 Prozent der Sehschärfe erreicht, so dass detailreiche, unverschommene Abbildungen hier nicht möglich sind (Jacob 1995, zit. nach Bente 2004: 303). Indem wir ein Objekt direkt anblicken, erreichen wir, dass dessen Abbild in den Bereich des höchsten Auflösungsvermögens auf der Netzhaut fällt (vgl. Hofer & Mayerhofer 2010: 146). Dieser hat einen Durchmesser von 0,5 bis 1 mm und umfasst einen Radius von 1° des horizontal insgesamt ca. 160° breiten menschlichen Sichtfeldes (vgl. Leven 1991: 74). Anhand der Ausrichtung der Augen kann also relativ genau bestimmt werden, welchen Bereich ihres Sichtfeldes Rezipientinnen und Rezipienten gerade hochauflösend sehen können (*foveales Sehen*) und welche Bereiche nicht (*peripheres Sehen*; Bente 2004: 303; Leven 1991: 75). Diesen Umstand macht sich das ET zu Nutze, um die von Rezipientinnen und Rezipienten visuell wahrgenommenen Inhalte bestimmen zu können.

Was bedeutet das am Beispiel konkreter Mediennutzungssituationen? Wenn wir in einer wohnzimmertypischen Entfernung von 3,5 Metern vor einem 16:9 Fernsehgerät mit 32" Bildschirmdiagonale sitzen, dann hat der foveal sichtbare Bildausschnitt einen Durchmesser von 12,2 cm. Das entspricht nur etwa 4 Prozent des Fernsehbildes. Hält man ein DIN-A4 großes Blatt zum Lesen in einem Abstand von 40 cm in den Händen, so kann man einen Bereich mit einem Durchmesser von etwa 1,4 cm foveal sehen (ca. 0,2 Prozent der Blattfläche). Der Rest des Blattes ist nur unscharf im peripheren Sichtfeld wahrnehmbar.

Die Kapazität fovealer und peripherer Wahrnehmung lässt sich anhand eines einfachen Beispiels anschaulich demonstrieren (vgl. Nielsen & Pernice 2010: 8). Wenn Sie das letzte Wort auf dieser Seite in den Blick nehmen, dann können Sie dieses Wort sowie vielleicht auch die Wörter in dessen direkter Umgebung identifizieren. Im peripheren Sichtfeld ist darüber hinaus erkennbar, dass sich auch im oberen Bereich der Seite, weit über dem von Ihnen fixierten Wort, noch Text befindet. Konkrete Wörter lassen sich hier aber nicht mehr erkennen. Würden sich Abbildungen oder andere in Form oder Farbe abgesetzte Gestaltungselemente (Tabellen, Infoboxen etc.) auf der Seite befinden, so wären auch diese im peripheren Blickfeld schemenhaft erkennbar.

Entsprechend der unterschiedlichen Leistungsfähigkeit fovealer und peripherer Wahrnehmung kommen beiden Arten des Sehens unterschiedliche Funktionen zu. Das foveale Sehen wird zur Detailinspektion ausgewählter Bereiche des Blickfeldes benötigt. Um einen Gesamteindruck von z. B. einem Zeitschriftencover zu bekommen, müssen dessen unterschiedliche Elemente sukzessive fovealisiert (Bente 2004: 303) und dann zu einem mentalen Abbild zusammengesetzt werden (Hofer & Mayerhofer 2010: 145). Im peripheren Sichtfeld können markante Areale eines Stimulus ausgemacht werden (z. B. Bilder oder große Überschriften auf einem Cover), die in einem nächsten Schritt detailliert foveal in den Blick genommen werden sollen. Das periphere Sehen

dient also vor allem der Umgebungsbeobachtung und der darauf basierenden Vorbereitung von Augenbewegungen auf neue Ziele (Findlay & Gilchrist 2003, zit. nach Geise 2011: 165). Detailliertere Informationen zu den hier komprimiert und vereinfacht dargestellten Grundlagen visueller Wahrnehmung können bei Duchowski (2007) nachgelesen werden.

Man kann zwei im Rahmen der Medienforschung relevante Arten des Blickverhaltens differenzieren, die zum Zweck der Detailbetrachtung bildlicher oder textlicher Informationen sowie zur Neuausrichtung auf weitere Zielpunkte durchgeführt werden: Fixationen und Sakkaden (Joos et al. 2003). *Fixationen* sind Phasen des relativen Stillstandes der Augen. In den Blick genommene (fixierte) Stimulusbereiche werden hier foveal betrachtet. Auch wenn der Begriff ‚Fixation‘ einen Stillstand der Augen nahelegt, weichen die Augen immer wieder geringfügig vom Fixationsobjekt ab und müssen durch minimale Positionskorrekturen in die ursprüngliche Position gebracht werden (vgl. Rötting 1999b: 5 zu Drifts, Tremor und Microsakkaden). Auch um ein sich bewegendes Objekt (z. B. ein fahrendes Auto) im Blick zu behalten, können langsame, gleitende *Folgebewegungen* (auch: *smooth pursuit*) während der Fixationen erfolgen (Joos et al. 2003: 155; Yarbus 1967: 159). Aus beiden Gründen ist bei der Definition von Fixationen von einem *relativen* Stillstand der Augen die Rede. *Sakkaden* sind sehr schnelle Augenbewegungen, die dazu dienen, die Augen foveal auf ein neues Ziel auszurichten (Rötting 1999b: 4). Sakkaden und Fixationen lösen sich in einem fortwährenden Prozess wechselseitig ab. Die Initiierung einer Sakkade kann zum einen bottom-up, also unwillkürlich und reflexartig durch plötzliche auffällige Änderungen im peripheren Blickfeld hervorgerufen werden (z. B. durch Bewegungen). Zum anderen können Sakkaden auch top-down, also willentlich von Rezipientinnen und Rezipienten ausgelöst werden, um peripher wahrgenommene Objekte, die das Interesse der Betrachterinnen und Betrachter geweckt haben, genauer zu inspizieren (Godijn & Theeuwes 2003: 3). Es herrscht weitgehender Konsens darüber, dass nur im Rahmen von Fixationen visuelle Informationen wahrgenommen werden können. Während der schnellen Sakkaden ist das Sehvermögen extrem eingeschränkt (Hammoud & Mulligan 2008: 7; Leven 1991: 82). Die im Rahmen von Sakkaden „überflogenen“ Areale von Medienangeboten werden also nicht im Detail wahrgenommen.

Uneinigkeit herrscht darüber, wie lange das Auge sich mindestens in relativem Stillstand zu einem Betrachtungsobjekt befinden muss, um eine Informationsaufnahme während einer Fixation zu ermöglichen (vgl. Hofer & Mayerhofer 2010: 153; Inhoff & Radach 1998: 34). Auf Basis bisheriger theoretischer und empirischer Erkenntnisse lässt sich diese Frage nicht zufriedenstellend beantworten. Es gibt eine große Varianz der von unterschiedlichen Forscherinnen und Forschern genannten Werte; sie schwanken zwischen 50 und 240 ms (vgl. Überblick bei Hofer & Mayerhofer 2010: 153). Vielschichtiger wird das Problem der Definition angemessener minimaler Fixationsdauern zudem dadurch, dass mediale Reizvorlagen unterschiedlich komplex zu verarbeitende Inhalte (Text vs. Bild, statische vs. dynamische Inhalte) aufweisen können und dadurch auch

jeweils angepasste Fixationsdefinitionen angebracht sind (ebd.). So werden zum Lesen tendenziell längere minimale Fixationsdauern benötigt als für die Wahrnehmung einfacher bildlicher Stimuli (Leven 1991: 95). Bei dynamischen Stimuli könnte aufgrund der erhöhten Komplexität ebenfalls eine längere Fixationsdauer vonnöten sein als bei statischen.

Die angesprochene Wissenslücke ist insofern sehr problematisch, als dass sie den Kern der operationalen Definition des für ET-Untersuchungen zentralen Basismaßes betrifft. Die Definition der minimalen Fixationsdauer kann einen starken Einfluss auf die ermittelten Ergebnisse ausüben (Hofer & Mayerhofer 2010: 153; Inhoff & Radach 1998: 34). Die Folge: Die Vergleichbarkeit von Ergebnissen unterschiedlicher Forscherinnen und Forscher, die auf Basis unterschiedlicher Fixationsdefinitionen entstanden sind, ist eingeschränkt. Ein kumulativer Erkenntnisfortschritt, bei dem aktuelle Projekte auf bisher publizierte Erkenntnisse zurückgreifen, wird dadurch erschwert. Aus diesem Grund wären weitere methodische Grundlagenuntersuchungen wünschenswert, auf Basis derer Richtlinien für angemessene minimale Fixationszeiten in Abhängigkeit vom verwendeten Stimulustyp definiert werden können.

3 Grundlagen der apparativen Blickaufzeichnung

Blickaufzeichnungsverfahren stellen eine Form der rezeptionsbegleitenden apparativen Beobachtung dar, durch die die Augenstellung von Personen bestimmt wird. Aus dieser kann auf die Blickrichtung bzw. die foveal sichtbaren Areale eines Stimulus geschlossen werden. Die Bestimmung der Augenstellung kann durch unterschiedliche technische Ansätze realisiert werden (vgl. Duchowski 2007: 51–58 oder Joos et al. 2003: 146–152 für einen Überblick). Heute finden vorwiegend videobasierte Systeme Anwendung. Sie setzen die per Kamera aufgezeichneten und mit einem Detektionsverfahren identifizierten Positionen von *Pupille* und *Cornea-Reflex* in Relation zueinander (vgl. Holmqvist et al. 2011: 26–29 zur Funktionsweise der Detektionsverfahren). Der Begriff ‚Cornea-Reflex‘ bezeichnet die Spiegelung einer Lichtquelle auf der konvex geformten Hornhaut¹ des Auges (Duchowski 2007: 57). Stimmen Pupillenmitte und Mittelpunkt des Cornea-Reflexes überein, schauen Personen exakt in Richtung der Lichtquelle, die die Reflexion auf der Cornea hervorruft. Bente verdeutlicht dies anhand eines anschaulichen Beispiels: „Wie man sich aus der Betrachtung des eigenen Auges im Spiegel klar machen kann, zeigt der Reflex einer äußeren Lichtquelle auf dem Auge, zum Beispiel das Bild des Fensters, immer in die Richtung dieser Quelle. Fällt dieser Reflex [...] in die Pupillenmitte, schaut das Auge direkt auf dieses Objekt“ (Bente 2004: 313). Im Rahmen des ET werden als äußere Lichtquellen punktförmige, in das jeweilige System integrierte Infrarotlichter genutzt (Duchowski 2007: 54). Infrarotlicht ist für das menschliche Auge

1 Die Hornhaut (Cornea) ist der Bereich, auf dem Kontaktlinsen üblicherweise aufliegen.

unsichtbar und damit kein Störfaktor im Rahmen der beobachteten Mediennutzungssituationen.² Für die in das System integrierten Kameras hingegen ist infrarotes Licht sichtbar, weshalb der Cornea-Reflex von der Kamera gleichermaßen erfasst werden kann (als heller Punkt) wie die dunkle Pupille des Auges. Damit ein Eye-Tracker nicht nur bei direkter Übereinstimmung von Pupillenmitte und Mittelpunkt des Cornea-Reflexes eine genaue Bestimmung der Blickrichtung vornehmen kann, muss das System vor seiner Anwendung kalibriert werden: „we must give the eye-tracker some examples of how points in our tracked area correspond to specific pupil and corneal reflection relations“ (Holmqvist et al. 2011: 29). Zur Kalibrierung müssen nacheinander eine Reihe von Punkten mit bekannten Koordinaten fixiert werden. Darauf aufbauend ist es dem System dann möglich, aus beliebigen Relationen der Pupillenmitte zum Mittelpunkt des Cornea-Reflexes den fixierten Bereich des Sichtfeldes zu schätzen.

Da sowohl die Pupille als auch der Cornea-Reflex als Referenzpunkte bei der Bestimmung der Blickrichtung berücksichtigt werden, führen leichte Kopfbewegungen, z. B. während der Betrachtung einer Website auf dem Bildschirm eines stationären ET-Systems, nicht zu gravierenden Fehleinschätzungen der Blickrichtung. Aus diesem Grund können moderne Apparaturen auch ohne Fixierung des Kopfes durch „Beißleisten“ oder Kinnauflagen genutzt werden, was der Natürlichkeit der Nutzungssituation zugutekommt, die Genauigkeit der gewonnenen Daten aber beeinträchtigt (Holmqvist et al. 2011: 51–53). Würde der Eye-Tracker hingegen nur die Pupille als Referenzpunkt verwenden, wäre es bei einer Bewegung dieses Referenzpunktes (z. B. um einige Millimeter nach rechts) nicht möglich, festzustellen, ob diese Positionsveränderung aus einer Augenbewegung oder aber einer Bewegung des Kopfes bzw. Körpers resultiert. „Two points of reference are needed to separate eye movements from head movements. The positional difference between the pupil center and corneal reflection remains relatively constant with minor head movements“ (Duchowski 2007: 56).

Aktuelle videobasierte Eye-Tracker nehmen die Bestimmung der betrachteten Koordinaten in unterschiedlich dichten Zeitintervallen von 60 bis über 1000 Hz (= Bildern pro Sekunde) auf. Sehr hohe zeitliche Auflösungen sind im Rahmen der Medienforschung, die sich vor allem für Fixationen interessiert, in vielen Fällen nicht unbedingt notwendig. Hier kann in der Regel auch auf „langsamere“ Systeme mit einer Auflösung von 60 oder 120 Hz zurückgegriffen werden (vgl. Holmqvist et al. 2011: Kap. 2.6.1 für eine Diskussion des Zusammenhangs von Geschwindigkeit und Datenqualität). Sind die Rohdaten auf beschriebenen Wege erhoben, kommt einer Analysesoftware die Aufgabe zu, die Basismaße der Fixationen und Sakkaden daraus zu ermitteln.³ Die Daten

2 Das emittierte Infrarotlicht gängiger ET Apparaturen stellt keine Gefahr für das menschliche Auge dar. Detaillierte Informationen zur Sicherheit des ET können bei Mulvey et al. (2008) nachgelesen werden.

3 Näheres zu Methoden der Ermittlung von Fixationen und Sakkaden aus Rohdaten findet sich bei Larsson (2010).

können vom System in Form von Datensätzen oder Videos ausgegeben werden, die mit den Blickverläufen unterlegt sind.

3.1 Gerätetypen: Stationäre und mobile Eye-Tracking-Systeme

Unter den videobasierten ET-Apparaturen lassen sich grob zwei verschiedene Systemtypen voneinander abgrenzen: stationäre und mobile Systeme. Am weitesten verbreitet sind *stationäre Systeme*, die äußerlich Computermonitoren ähneln und genau wie diese auf einem Tisch vor den Versuchspersonen platziert werden (siehe Abbildung 1). Sie bestehen aus einem Bildschirm, in den Infrarotlichtquellen sowie Kamera(s) zur Aufzeichnung des Augenabbildes integriert sind. Da die Augen aus der Ferne betrachtet werden und keinerlei körperlicher Kontakt zwischen Apparatur und Versuchsperson besteht, werden die Geräte auch als *Remote Eye-Tracker* bezeichnet. Mediale Stimuli, wie z. B. Internetseiten oder Filme, werden bei diesen Systemen direkt auf dem Monitor präsentiert (Geise 2011: 178; Holmqvist et al. 2011: 51). Remote Eye-Tracker sind mit einem Steuerungscomputer verbunden, über den ET-Projekte angelegt und gesteuert, Kalibrierungen durchgeführt, Echtzeitdarstellungen des Blickverlaufs nachverfolgt sowie Daten archiviert, ausgewertet und für die datenanalytische Weiterverwertung exportiert werden können.

Mobile Eye-Tracker können flexibler eingesetzt werden als ihre stationären Pendanten. Die Systeme sind auf Kopfbedeckungen (z. B. Helme & Caps) montiert oder in Brillen integriert (siehe Abbildung 1). Sie können von Versuchspersonen mit sich geführt,

Abbildung 1 Beispiele unterschiedlicher Eye-Tracking-Systeme (Links: Tobii T60XL Remote System, rechts: Brillensystem SMI Eye Tracking Glasses; Herstellerabbildungen)



d. h. auch außerhalb von Laborsituationen verwendet werden. Ein Steuerungscomputer zur Aufzeichnung der Daten wird von den Versuchspersonen in einem Rucksack transportiert. Neben Augenkamera und Infrarotlichtquelle verfügen mobile Systeme zusätzlich über eine Umgebungskamera, um das bei Körper- und Kopfbewegungen variierende Sichtfeld der Versuchspersonen aufzeichnen zu können. Eingeschränkt werden die Nutzungsszenarien mobiler Geräte dadurch, dass sie nur für jene Entfernungen präzise Schätzungen der Fixationsorte liefern, für die sie zuvor kalibriert wurden. Es ist also z. B. nicht ohne weiteres möglich, sowohl für das aus einer Entfernung von 3,5 m betrachtete Fernsehbild als auch für den aus einer Entfernung von 50 cm betrachteten Bildschirm des nebenbei genutzten Smartphones präzise Blickschätzungen zu erhalten (vgl. parallax error bei Holmqvist et al. 2011: 60; Rötting 1999a: 30). Einige Analysesoftwarepakete bieten jedoch Möglichkeiten der manuellen Korrektur des Fehlers im Rahmen der Auswertung.

3.2 Messgenauigkeit

Die Messgenauigkeit (*Accuracy*) von Remote Eye-Trackern liegt laut Herstellerangaben der beiden führenden Anbieter (SensoMotoric Instruments – SMI und Tobii) bei durchschnittlichen Abweichungen von $0,4^\circ$ – $0,5^\circ$ zwischen tatsächlichem und gemessenem Blickpunkt. Für das mobile System macht nur SMI eine Angabe zu den Abweichungen; sie liegt bei bis zu $1,0^\circ$. Methodenexperimente von Komínková, Pedersen, Hardeberg und Kaplanová (2008) oder Nevalainen und Sajaniemi (2004) ergaben ebenfalls Vorteile für stationäre Systeme, wenngleich die durchschnittlichen Abweichungen beider Gerätetypen in den Praxistests höher ausfielen als genannte Herstellerangaben ($1,3$ – $1,6^\circ$ bei stationären und $1,8^\circ$ bei mobilen Systemen).⁴ Was bedeuten diese Werte für Blickaufzeichnungen typischer Rezeptionssituationen? Betrachten Rezipientinnen und Rezipienten bspw. bei der Internetnutzung einen Monitor aus 60 cm Entfernung, würde eine durchschnittliche Abweichung von $0,5^\circ$ einer Differenz von 0,5 cm zwischen tatsächlichem und gemessenem Blickpunkt entsprechen (bei $1,8^\circ$ ca. 1,9 cm). Bei der Betrachtung eines 3,5 m entfernten Fernsehbildes lägen die Differenzen bei 3,1 cm ($0,5^\circ$ Accuracy) bzw. bei ca. 11 cm ($1,8^\circ$ Accuracy). Die erforderliche Accuracy hängt zum einen davon ab, ob die Forschungsfrage nur die Unterscheidung grober Bereiche erfordert oder auch kleinere Änderungen der Blickrichtung relevant sind. Zum anderen spielt auch der Versuchsaufbau eine Rolle, da die Höhe der Abweichungen (in cm) mit zunehmendem Stimulusabstand wächst. Es sollte also vor der Untersuchung berechnet

4 Tobii (und vermutlich auch SMI) verwendet im Gegensatz zu den genannten Untersuchungen bei der Ermittlung der Accuracy eine Kinnablage, durch die Kopfbewegungen deutlich eingeschränkt werden, was sich vorteilhaft auf die Accuracy auswirkt. Wenn sehr genaue Messungen notwendig sind, kann auch auf hier nicht eingehender thematisierte, stationäre Tower-Systeme zurückgegriffen werden (z. B. SMI iView X™ Hi-Speed), die Kopfbewegungen ebenfalls stark beschränken.

werden, welche Accuracy im Rahmen des spezifischen Settings notwendig ist. Die von der Literatur und den Herstellern als Accuracy bezeichnete Messgröße kann als *Validität* der Blickrichtungsmessung interpretiert werden. Davon abzugrenzen sind Maße zur Berechnung der *Precision*. Sie geben Auskunft darüber, wie stark die einzelnen Messpunkte bei wiederholter/fortwährender Betrachtung eines stationären Punktes variieren und entsprechen damit der *Reliabilität* des Verfahrens (weiterführende Informationen zu Precision und Accuracy bei Holmqvist et al. 2011: 34–43).

3.3 Synchronisation von Stimulusinhalten und Blickaufzeichnungsdaten

Ein zentraler Vorteil stationärer Systeme ist die Möglichkeit einer automatischen *Synchronisation* der aufgezeichneten Blickkoordinaten mit dem präsentierten Stimulusmaterial. Remote Eye-Tracker geben Koordinaten aus, die sich auf Punkte des Bildschirms beziehen, auf dem das Stimulusmaterial präsentiert wurde. Diese Rohdaten lassen sich entsprechend ohne größeren Aufwand in grafische Auswertungen oder Kennzahlen (z. B. die Fixationsdauer oder -anzahl eines Werbebanners) übersetzen. Mobile ET-Systeme bieten diese komfortable Möglichkeit in der Regel nicht. Da jegliche „Kopfbewegungen des Probanden dazu führen, dass der aufgezeichnete Bildausschnitt permanent variiert, lassen sich die Blickkoordinaten nicht unmittelbar mit den Bildinhalten synchronisieren.“ (Geise 2011: 181) Positionsänderungen mobiler Medien (z. B. Zeitungen, die beim Lesen, Blättern etc. nicht immer in exakt derselben Position gehalten werden) führen selbst bei ruhiger Kopf- und Körperhaltung zum selben Problem. In den meisten Fällen ist der einzige ohne hohen Aufwand nutzbare Output mobiler Systeme deshalb ein durch die Umgebungskamera aufgezeichnetes Video, das mit dem individuellen Blickverlauf der jeweiligen Versuchsperson überlagert ist (vgl. Holmqvist et al. 2011: 54). Nur durch aufwändige manuelle Nachcodierungen sind im Nachhinein weitere stimulusbezogene Daten zu ermitteln.

In Ausnahmefällen, wie z. B. der Blickaufzeichnung bei Bildschirmspielen, kann dies auch bei Remote Eye-Trackern der Fall sein. Die auf dem Eye-Tracker-Bildschirm dargestellten Inhalte können sich interaktivitätsbedingt von Person zu Person stark unterscheiden. Möchte man z. B. herausfinden, wie lange in ein Spiel eingebettete Werbemittel fixiert wurden, so ist auch hier eine manuelle Codierung der mit Blickaufzeichnungen überlagerten Videodaten notwendig (Schlütz et al. 2013). Interaktivität stellt jedoch nicht generell ein Problem für die Synchronisierung von Stimulusinhalten und Blickkoordinaten dar. Durch Scrolling sowie die Nutzung von Links produzierte Unterschiede in den Bildschirminhalten bei der Internetnutzung können etwa durch viele Systeme registriert und bei der Synchronisation berücksichtigt werden.

3.4 Natürlichkeit der Nutzungssituation

Um extern valide Daten durch Blickbeobachtungen generieren zu können, müssen möglichst natürliche Rezeptionssituationen hergestellt werden. Remote Eye-Tracker machen es Versuchspersonen aufgrund ihrer großen Ähnlichkeit zu PC-Monitoren leicht, die Beobachtungssituation auszublenden und sich natürlich zu verhalten. Das Tragen eines Caps, Helms oder einer Brille bei mobilen Systemen kann hingegen eine gedankliche Präsenz der Untersuchungssituation befördern. Auch ein ungewohntes Körpergefühl kann durch das Tragen der Apparatur ausgelöst werden (Komínková et al. 2008: 10). Beides sind Störfaktoren, die möglicherweise das Rezeptionsverhalten beeinflussen. Von den allgemeinen, konstruktionsbedingten Unterschieden abgesehen, muss die mit beiden Systemen erzielbare Natürlichkeit der Untersuchungssituation spezifisch für das jeweils untersuchte Medium beurteilt werden. Für Inhalte, die ohnehin am Monitor genutzt werden (z. B. Internetanwendungen, Computerspiele etc.), eignen sich stationäre Geräte besonders gut. Sollen hingegen Printmedien untersucht werden, „bei denen der natürliche Umgang mit dem Stimulus auch das Umblättern oder Falten“ des Mediums enthalten kann, so lässt sich dies nicht vollständig auf dem Bildschirm der statischen Geräte nachbilden (Geise 2011: 180). Hier haben mobile Systeme einen Vorteil.⁵ Zusammenfassend lässt sich sagen, dass bei der Auswahl eines geeigneten Blickaufzeichnungssystems immer eine von der Forschungsfrage und dem untersuchten Medium abhängige Abwägung zwischen der Natürlichkeit der Nutzungssituation, der notwendigen Genauigkeit der Messungen sowie der Praktikabilität der angestrebten Auswertungsstrategie getroffen werden muss. Eine „One-Fits-All“-Lösung existiert bisher nicht, könnte sich in Zukunft aber möglicherweise im Rahmen der zunehmenden Miniaturisierung mobiler Systeme verbunden mit einer einfacheren Stimulussynchronisierung der Daten durch die zusätzliche Erfassung von Kopfbewegungen ergeben (vgl. Head-Tracking-Systeme bei Geise 2011: 182; Holmqvist et al. 2011: 54).

4 Auswertungsansätze für Blickaufzeichnungsdaten

Wie im letzten Kapitel deutlich wurde, werden durch rezeptionsbegleitende Blickaufzeichnungen zeitlich und räumlich hochauflösende Daten generiert. Die Koordinaten fixierter Stimulusbereiche werden mindestens 60-mal pro Sekunde registriert und aufgezeichnet (s. o.). So entstehen schon im Rahmen einer dreiminütigen Beobachtung einer einzelnen Person 10 800 Datenpunkte bzw. Fälle. Um die Datenfülle im Rahmen der jeweils verfolgten Fragestellungen nutzbar zu machen, werden unterschiedliche Auswertungsansätze verfolgt, die sich als *qualitativ-deskriptiv*, *quantitativ-deskriptiv* und

5 Praktische Empfehlungen zum Einsatz mobiler Systeme mit unterschiedlichen Medien sind bei Gröblbauer und Kapf (2011: 7–13) nachzulesen.

quantitativ-schließend charakterisieren lassen. Sie erfordern ein unterschiedliches Maß an Aggregation der Rohdaten auf zeitlicher, räumlicher und personaler Ebene (Bente 2004: 318). Die Auswertungen können sich entweder auf den kompletten Stimulus oder aber auf spezifische, räumlich und zeitlich definierbare Stimulusbereiche beziehen.

Qualitativ-deskriptive Auswertungen basieren maßgeblich auf der visuellen Inspektion von Rohdaten, die zuvor auf unterschiedliche Art und Weise grafisch aufbereitet und ggf. aggregiert wurden. Die grundlegendste Art der visuellen Inspektion, bei der keine Datenaggregation erfolgt, ist die Betrachtung von Videos, die mit den individuellen Blickverläufen einzelner Versuchspersonen (Blickpfade bzw. Scanpaths) unterlegt sind. Fixationen werden hier meist mit Punkten gekennzeichnet, die sich mit zunehmender Fixationsdauer vergrößern. Sakkaden werden durch Linien visualisiert, die einzelne Fixationen verbinden. Durch Gaze-Plots lassen sich die Blickpfade einer oder mehrerer Personen in einem einzelnen Bild visualisieren, d.h. es findet eine zeitliche und (optional) auch eine personenbezogene Aggregation statt. Durch die chronologische Nummerierung der Fixationen gemäß ihrer zeitlichen Abfolge bleibt der prozesuale Charakter der Daten trotz statischer Abbildung für die Analyse erhalten (siehe Abbildung 2). Attention-Maps stellen auf verschiedene Art und Weise dar, wie intensiv unterschiedliche Bereiche eines Stimulus betrachtet wurden. Heat-Maps und Lumination-Maps sind besonders verbreitete Typen dieser Art (vgl. Holmqvist et al. 2011: Kap. 7). Je rötlicher Stimulusareale in Heat-Maps eingefärbt sind, desto intensiver wurden sie betrachtet.⁶ Lumination-Maps zeigen nur jene Bereiche des Stimulus hell an, die betrachtet wurden; der Rest des Stimulus bleibt schwarz. Attention-Maps können auf Basis einzelner oder mehrerer Versuchspersonen generiert werden. Sie kombinieren also eine zeitliche mit einer (optionalen) personenbezogenen Aggregation. Bee-Swarm-Darstellungen sind Videos, in denen durch Punkte für eine Vielzahl von Versuchspersonen gleichzeitig visualisiert wird, welche Stimulusbereiche sie gerade betrachten (jeder Punkt im Bild entspricht der durch eine Person betrachteten Koordinate). Hier findet also nur eine personenbezogene Aggregation statt; die zeitliche und räumliche Auflösung der Daten bleibt erhalten. Bei qualitativ-deskriptiven Auswertungen besteht generell eine räumliche Kopplung der Daten zum Stimulus. Aggregationen finden nur auf zeitlicher und personaler Ebene statt.

Die quantitativ-deskriptive Aufbereitung von Rohdaten durch Kennzahlen beinhaltet auch Aggregationen auf räumlicher Ebene. Für den kompletten Stimulus oder einzelne Ausschnitte können quantifizierende Parameter wie z. B. die Fixationsdauer oder -anzahl berechnet werden. Um die Parameter nur für spezifische Ausschnitte (z. B. ein Werbebanner) zu generieren, müssen so genannte Areas of Interest (AOI) in der Analysesoftware definiert werden. Dadurch findet technisch gesehen eine Aggregation der Blickdaten auf räumlicher Ebene (z. B. im Areal des Werbebanners), auf zeitlicher Ebene

6 Die Visualisierung kann auf Basis unterschiedlicher Basisdaten, wie z. B. der Fixationsdauer, der relativen Fixationsdauer oder der Fixationsanzahl generiert werden.

Abbildung 2 Beispiele für Visualisierung von Blickaufzeichnungsdaten durch Gaze-Plot, Heat-Map und Lumination-Map (v.l.n.r.; eigene Darstellung)



(Gaze-Plot der ersten 5 Sek. d. Betrachtung, N=1, Fixationsdefinition: Minimaldauer = 100 ms, Maximaldistanz zwischen zwei aufeinanderfolgenden Punkten = 50 pixel/sample bei 60 Hz; Attention-Maps auf Basis der Fixationsdauer in den ersten 10 Sek. der Betrachtung, N=57. System: Tobii T60XL)

(z. B. Summierung der Fixationsanzahl über die gesamte Präsentationsdauer) sowie optional auch auf personenbezogener Ebene statt (Fixationsanzahl einzelner Personen vs. Fixationsanzahl eines (Teil-)Samples). Die Definition beweglicher, zeitlich begrenzter AOI ist ebenfalls möglich, z. B. um Fixationen zu erfassen, die auf ein quer über den Bildschirm fahrendes Auto entfallen. Die ermittelten quantifizierenden Parameter lassen sich exportieren und im Rahmen quantitativ-schließender, d. h. inferenzstatistischer Auswertungen weiterverarbeiten. So könnte etwa geprüft werden, ob sich Eltern und Kinder hinsichtlich der Fixationsdauer von Alterskennzeichnungen auf Mediencovern systematisch unterscheiden (Jöckel, Blake & Schlütz 2013).

Die Ausführungen machen deutlich, dass die ursprünglich hochauflösenden Informationen von Eye-Tracking-Daten bei quantitativen Auswertungen stark aggregiert und dabei ihres ehemals prozesshaften Charakters „beraubt“ werden. Es gibt jedoch vielversprechende Ansätze, mit denen die Struktur von Blickverläufen paarweise (Dewhurst et al. 2012) oder sogar für große Fallzahlen (Feusner & Lukoff 2008, zit. nach Holmqvist et al. 2011: 278) quantitativ beschreibbar und vergleichbar gemacht werden könnte, sofern eine spezifische Fragestellung dies verlangt. Die hier vorgestellten Auswertungsansätze sind nicht umfassend – sie sollen eher einen grundsätzlichen Eindruck verschiedener analytischer Zugänge vermitteln. Für erschöpfendere Darstellungen sei auf Holmqvist et al. (2011: Kap. 6–8) verwiesen.

5 Indikatoren

Durch Blickaufzeichnungen soll nicht ausschließlich herausgefunden werden, welche Stimulusbereiche wie lange und in welcher Reihenfolge betrachtet werden (vgl. Abschnitt 4). Es existieren unzählige, auf Eye-Tracking-Daten basierende Indikatoren, die zur Abbildung verschiedenster Phänomene genutzt werden. Einige davon werden auch herangezogen, um Aussagen über Stimuli sowie deren Nutzerinnen und Nutzer zu treffen (stimulus- und rezipientenbezogene Indikatoren).

Rezipientenbezogene Indikatoren lassen sich grob den Kategorien *Stimulusinteresse*, *Workload* bei der Wahrnehmung und Verarbeitung, *Verarbeitungstiefe*, *aufgabenspezifische Expertise* sowie *emotionale Reaktionen*⁷ auf Stimuli zuordnen. Wenig aussagekräftig sind Blickaufzeichnungsdaten hingegen, wenn es um die Messung von Motiven, subjektiven Bewertungen, Einstellungen, Wissen oder eine differenzierte Erfassung des Rezeptionserlebens geht. Dieses Defizit ist klassischerweise mit der Anwendung von Beobachtungsverfahren verbunden (vgl. auch den Beitrag von Gehrau und Schulze im vorliegenden Band). Um es auszugleichen, können Blickaufzeichnungen mit weiteren Erhebungsmethoden kombiniert werden. Geise (2011: 216–240) liefert einen Überblick möglicher Mehrmethoden-Designs sowie spezifischer Erkenntnispotenziale, die damit verbunden sind. Indikatoren, die den Stimulus bzw. definierte Stimulusbereiche beschreiben, lassen sich grob den Kategorien der *sprachlichen und gestalterischen Komplexität von Inhalten*, ihrer *Aufmerksamkeitsstärke* sowie, im Falle interaktiver Angebote, auch deren *Usability* zuordnen.

Eine umfassende Darstellung der in der Forschung verwendeten Indikatoren wäre an dieser Stelle weder möglich noch zielführend. Für einen Überblick sei auf die sehr kenntnis- und umfangreiche Auseinandersetzung mit ET-Maßen bei Holmqvist et al. (2011) verwiesen. Ziel dieses Abschnittes ist es stattdessen, beispielhaft die Möglichkeiten und Grenzen der Nutzung von Blickaufzeichnungsdaten als Indikatoren für kommunikationswissenschaftlich relevante Phänomene zu verdeutlichen. Wie auch bei vielen physiologischen und reaktionszeitbasierten Beobachtungen (vgl. die Beiträge von Fahr & Hofer sowie Hefner im vorliegenden Band) sind auch beim ET teils mehrere Interpretationsschritte notwendig, um von Rohdaten auf kognitive Prozesse, Rezipientenmerkmale (z. B. das Interesse an bestimmten Inhalten) oder Stimulusmerkmale (z. B. die Verständlichkeit eines Textbausteins) schließen zu können. „By itself, eyetracking can do one thing and one thing only: discover where users look at.“ (Nielsen & Pernice 2010: 10). Jeder von den Rohdaten abstrahierende Interpretationsschritt ist mit spezifischen Annahmen verbunden, deren Plausibilität theoretisch begründet und optimalerweise auch empirisch validiert werden muss. Zwei der grundlegenden Annahmen,

7 Emotionale Reaktionen können nur sehr basal in Form des „emotional arousal“ erfasst werden, das durch die Pupillenweite angezeigt wird (Holmqvist et al. 2001: 394). Eine differenzierte Messung unterschiedlicher emotionaler Zustände ist hingegen nicht möglich.

auf deren Basis weitergehende Interpretationen von Blickaufzeichnungsdaten üblicherweise basieren, werden im Folgenden kurz vorgestellt und diskutiert.

5.1 Grundlegende Annahmen der Eye-Tracking Forschung

Die erste grundlegende Annahme besagt, dass man aus Fixationen ableiten kann, womit sich Personen zeitgleich kognitiv auseinandersetzen (Inhoff & Radach 1998: 37; Leven 1991: 87). Die Daten einiger sehr sorgfältig konzipierter Untersuchungen (z. B. Just & Carpenter 1976, 1980; Russo & Rosen 1975; Yarbus 1967) liefern empirische Evidenzen für die Tragfähigkeit der Annahme: „Eye movements reflect the human thought processes; so the observer’s thought may be followed to some extent from records of eye-movements (the thought accompanying the examination of the particular object)“ (Yarbus 1967: 190). Obschon die intuitiv plausible Prämisse durch konkrete Untersuchungsergebnisse mit unterschiedlichen Stimulusmaterialien gestützt wird (Texte: Just & Carpenter 1976; Russo & Rosen 1975; Bilder: Yarbus 1967; oder gezeichnete Objekte: Just & Carpenter 1980), kann nicht grundsätzlich von einer Übereinstimmung zwischen kognitivem und visuellem Fokus ausgegangen werden. Aufmerksamkeitsverschiebungen auf nicht fixierte, periphere Bereiche des Blickfeldes sind auch unabhängig von Sakkaden möglich, die die Augen foveal auf diesen Bereich ausrichten (Findlay & Gilchrist 1998: 296). Davon abgesehen können wir uns natürlich auch mit Dingen auseinandersetzen, die gänzlich außerhalb unseres momentanen Sehbereichs liegen. Neben diesen allgemeinen Einschränkungen sind auch speziell auf Mediennutzungssituationen bezogene Generalisierbarkeitsabwägungen der Prämisse zu diskutieren. Im Rahmen der grundlagenpsychologischen Untersuchungen, die als empirische Hinweise für die Belastbarkeit der Prämisse dienen, werden Versuchspersonen meist mit konkreten Aufgaben konfrontiert, die sie zudem häufig unter Zeitdruck lösen müssen. Diese Rahmenbedingungen könnten eine Voraussetzung oder ein begünstigender Faktor für die Übereinstimmung von optischem und kognitivem Fokus sein (Bente 2004: 310). In natürlichen Mediennutzungssituationen kann nicht grundsätzlich von einer gleichermaßen konzentrierten und zielgerichteten Rezeptionsweise ausgegangen werden. Vielmehr ist diesbezüglich mit einer hohen intra- und intermedialen Varianz zu rechnen. Je nachdem, ob Personen ein wichtiges Spiel ihres favorisierten Fußballclubs im TV verfolgen oder sich durch ein beliebiges Programm beiläufig beim Verspeisen ihres Abendessens unterhalten, ist mit einer unterschiedlich konzentrierten Nutzung zu rechnen (intramediale Varianz). Bei der Nutzung textlastiger oder interaktiver Medien (z. B. Zeitungen oder Bildschirmspiele) dürfte die Konzentration aufgrund der aktiven Einbindung beim Lesen oder Steuern des Spielgeschehens tendenziell höher ausfallen als beim Fernsehen (intermediale Varianz). Eine Übertragbarkeit der empirischen Erkenntnisse zur Übereinstimmung von Blickrichtung und kognitivem Fokus sollte also vor allem dann gegeben sein, wenn Personen ein klares Ziel bei der Betrachtung medialer Stimuli verfolgen

bzw. auf deren Nutzung konzentriert sind. Ist dies nicht der Fall, ist die Übertragbarkeit der Ergebnisse und somit die Gültigkeit der Annahme fraglich.

Eine zweite Prämisse ist, dass die Dauer der Fixation von Arealen mit der Dauer der kognitiven Beschäftigung mit diesen Arealen korrespondiert (Joos et al. 2003: 156). Auch für diese als *Eye-Mind Assumption* bekannte Annahme (vgl. Just & Carpenter 1980: 330) finden sich überzeugende empirische Evidenzen (z. B. Just & Carpenter 1976, 1980). Anhand bekannter Alltagsphänomene, wie z. B. dem in Gedanken versunkenen Starren auf einen Punkt (Leven 1991: 97), kann stellvertretend verdeutlicht werden, dass auch die Eye-Mind Assumption keine generelle Gültigkeit besitzt. Um als akzeptable Grundannahmen verwendet werden zu können, müssen die genannten Annahmen jedoch auch „nicht für jede einzelne Fixation gelten, sondern nur“ auf den Großteil der Fixationen zutreffen (Geise 2011: 198). Dies kann für beide Prämissen grundsätzlich angenommen werden (Duchowski 2007: 12; Joos et al. 2003: 154).

5.2 Interpretationsoffenheit der Maße

Von der Gültigkeit oben genannter Annahmen ausgehend, sind weitere Interpretationsschritte notwendig, um auf spezifische Rezipienten- oder Stimulusmerkmale schließen zu können. Am Beispiel der Fixationsdauer, einer der am häufigsten verwendeten Messgrößen (Inhoff & Radach 1998: 40; Joos et al. 2003: 155), kann dies beispielhaft verdeutlicht werden. Wenn man Fixationen und deren Länge als Ergebnis bewusster Selektionsentscheidungen auffasst, kann die Fixationsdauer z. B. als Maß für das Interesse von Personen an definierten Stimulusbereichen (z. B. einer Infografik) gewertet werden. Eine andere Interpretation derselben Messgröße wäre, sie als Maß der Komplexität betrachteter Stimulusbereiche zu interpretieren. Die zugrunde liegende Annahme wäre hier, dass die für die Verarbeitung der Inhalte notwendige Zeit mit steigender Komplexität wächst. Das Beispiel macht deutlich, dass Blickaufzeichnungsdaten für sich genommen oft „inhaltlich unspezifisch und damit interpretationsoffen“ sind (Geise 2011: 200; Joos et al. 2003: 160). Die inhaltliche Bedeutung von Maßen kann nur unter Berücksichtigung von Kontextmerkmalen erschlossen werden (Bente 2004: 303). Dazu gehören z. B. die Art präsentierter Stimuli sowie die Versuchspersonenanweisungen, mit denen gearbeitet wurde. Die Interpretationsoffenheit der Maße wird in gewähltem Beispiel auch auf einer weiteren Ebene deutlich: Eine einzelne Messgröße kann sowohl als Indikator für *Stimuluseigenschaften* (Komplexität) als auch *Rezipienteneigenschaften* (Interesse) interpretiert werden, da das Blickverhalten sowohl top-down durch Motive und Ziele von Rezipientinnen und Rezipienten als auch bottom-up durch Stimuluseigenschaften beeinflusst werden kann (Bucher & Schumacher 2006: 352). Diese doppelte Interpretationsoffenheit bezieht sich nicht allein auf die hier beispielhaft betrachtete Fixationsdauer. Auch andere Maße, die auf Fixationen, Sakkaden, Blickpfaden oder Rohdaten basieren, sind davon betroffen. Einheitliche Standards bei der Interpre-

tation von Indikatoren sind selbst in Bezug auf festgelegte Kontexte bisher eher selten (Geise 2011: 199). Eine sorgfältige Auseinandersetzung mit möglichen Lesarten von Indikatoren im Rahmen der Studienkonzeption ist deshalb unerlässlich. Hilfreiche Kriterien zur Auswahl geeigneter Maße für gegebene Fragestellungen liefern Holmqvist et al. (2011: 76–77).

6 Anwendungsbereiche

Die durch Blickaufzeichnungen gewonnenen Basisinformationen, welche Inhalte wie lange und in welcher Reihenfolge genutzt werden, können sich in verschiedenen kommunikationswissenschaftlichen Anwendungsbereichen als nützlich erweisen. In erster Linie natürlich im Rahmen der Selektionsforschung, deren zentrales Anliegen die Beschreibung und Erklärung des Auswahlverhaltens von Rezipientinnen und Rezipienten ist. Medienselektion ist ein mehrstufiger Prozess, der sowohl a) eine grundsätzliche Entscheidung für die Mediennutzung, b) die Auswahl eines spezifischen Mediums sowie auch c) die innerhalb eines Mediums gewählten Beiträge und d) Beitragselemente betrifft. Die hohe zeitliche und räumliche Auflösung von Blickaufzeichnungsdaten macht Selektionsentscheidungen auch auf den Ebenen c) und d) messbar: Es kann präzise beschrieben werden, welche Beiträge und welche Beitragselemente (z. B. Bilder, Überschriften oder Textbestandteile) eines Mediums genutzt werden. In Kombination mit rückblickenden Befragungen anhand von Blickaufzeichnungsvideos ist für bewusst getroffene Selektionsentscheidungen zudem eine Verbalisierung von Gründen möglich. So kann z. B. ermittelt werden, warum das Lesen eines Artikels an einer bestimmten, per Blickaufzeichnung identifizierten Stelle abgebrochen wurde. Ohne Blickaufzeichnungen, wie sie etwa in der vielbeachteten „Eyes on the News“ Studie (Garcia & Stark 1991) angewendet wurden, sind innermediale Selektionsentscheidungen der Ebenen c) und d) nicht ähnlich detailliert beschreibbar und erklärbar. Speziell die Werbeforschung nutzte deshalb in der Vergangenheit das ET intensiv, um zu erforschen, welche inhaltlichen und gestalterischen Charakteristika dazu führen, dass Werbemittel im Rahmen von Selektionsprozessen berücksichtigt und möglichst lange betrachtet werden (z. B. Blake, Schlütz & Herbst 2012; Pieters & Wedel 2004). Die gewonnenen empirischen Erkenntnisse können dazu beitragen, bestehende Theorien und Modelle der Medienselektion zu prüfen bzw. sie auf Basis identifizierter Muster und Regelmäßigkeiten induktiv weiterzuentwickeln. Selektionsentscheidungen können natürlich nicht nur in Abhängigkeit von Stimulusmerkmalen, sondern auch abhängig von Rezipientenmerkmalen untersucht werden: Eine Beispielfragestellung dazu wäre etwa, inwieweit sich Frauen, die mit ihrem Körper zufrieden sind, von jenen, auf die dies nicht zutrifft, im Hinblick auf die Selektion bzw. Vermeidung hochattraktiver Frauendarstellungen in Medien unterscheiden (vgl. Knobloch-Westerwick & Romero 2011).

Durch Befragungen oder alternative Beobachtungsverfahren (z. B. Analysen von Browserdaten genutzter Webinhalte oder Mouse-Tracking-Verfahren) kann das Selektionsverhalten aufgrund mangelnder Gedächtniskapazitäten, einer mangelnden räumlichen Auflösung der Beobachtungsmethode oder einer Beeinträchtigung des natürlichen Selektionsverhaltens durch ihre Anwendung (speziell beim Mouse-Tracking) nicht gleichermaßen detailliert und natürlich erfasst werden. Zuverlässige Messungen sind hier meist nur für die Ebenen a), b) und teilweise auch c) möglich. Spätestens auf Ebene d) bleiben Blickaufzeichnungen jedoch das Mittel der Wahl zur Erfassung des Selektionsverhaltens. Bei der Interpretation des erfassten Nutzungsverhaltens ist zu beachten, dass auch das ET anfällig für sozial erwünschtes Verhalten sein kann: Das Wissen um die Nachvollziehbarkeit des eigenen Blickverlaufs kann das Selektionsverhalten beeinflussen (Risko & Kingstone 2011).

Eine exakte Erfassung genutzter Inhalte auf allen Selektionsebenen kann auch jenseits der Beschreibung und Erklärung des „Selektionsverhaltens an sich“ sinnvoll sein. Im Rahmen experimenteller Medienwirkungsstudien können Blickaufzeichnungsdaten beispielsweise zur Generierung von Moderatorvariablen verwendet werden, die Aufschluss über die Betrachtungsdauer der kritischen, manipulierten Stimulusbereiche geben. Möchte man etwa die Körperzufriedenheit von Leserinnen eines Zeitschriftenartikels, in dem hochattraktive Frauendarstellungen enthalten sind, mit jener der Leserinnen eines Artikels vergleichen, der weniger attraktive Frauendarstellungen enthielt, so wäre die tatsächliche Betrachtungsdauer der manipulierten Fotos eine relevante Moderatorvariable. Sie könnte dazu beitragen, die Innergruppenstreuung besser zu verstehen und differenziertere Aussagen zum Wirkungsprozess zu machen. Auch für korrelative Studien ist es mitunter relevant, Zusammenhänge zwischen Betrachtungsdauer oder -häufigkeit einzelner Stimulusbereiche und einer abhängigen Größe herzustellen. Ein Beispiel dafür ist die Untersuchung des Zusammenhangs zwischen der Betrachtungsdauer von Werbemitteln und der Erinnerung an die beworbene Marke (vgl. z. B. Blake et al. 2012).

Auch im Rahmen von Studien zum Rezeptionserleben ergeben sich neue Perspektiven aus der Möglichkeit zur zeitlich und räumlich exakten Beschreibung betrachteter Medieninhalte, wenn diese z. B. mit ebenfalls rezeptionsbegleitenden Verfahren zur Ermittlung von Bewertungen (etwa RTR-Messungen, vgl. den Beitrag von Maurer im vorliegenden Band) kombiniert werden. Da moderne Remote Eye-Tracking-Systeme es erlauben, Videos der Gesichter der Versuchspersonen während der Rezeption aufzuzeichnen, bestünde auch die Möglichkeit zur rezeptionsbegleitenden Messung von Emotionen über automatisierte Facial Action Coding Systeme, deren Entwicklung immer weiter vorangetrieben wird (vgl. Hamm, Kohler, Gur & Verma 2011 oder www.affectiva.com für eine kommerzielle Anwendung). Dies würde eine unmittelbare Verknüpfung von betrachteten Inhalten mit darauf folgenden Emotionen ermöglichen, was gerade im Bereich der Unterhaltungsforschung eine sinnvolle Perspektive wäre. Es

könnte so etwa ermittelt werden, wie Rezipientinnen und Rezipienten spezifische Inhalte in Filmen, Werbeanzeigen oder ausgewählte Textpassagen bewerten oder wie sie auf diese Inhalte emotional reagieren.

7 Herausforderungen

Blickaufzeichnungen können, vor allem in Kombination mit anderen Erhebungsverfahren wie beispielsweise der Befragung, eine sinnvolle Erweiterung des kommunikationswissenschaftlichen Methodeninventars darstellen. Sie tragen zur Erforschung relevanter Fragestellungen bei, die sonst nicht oder ungenauer untersucht werden könnten. Der technische Fortschritt hat dafür gesorgt, dass die Methode heute wesentlich kostengünstiger, einfacher und unter natürlicheren Rahmenbedingungen einsetzbar ist als früher (vgl. Duchowski 2007). Die Kehrseite des vereinfachten Zugangs: Die „Plug’n Play“ Bedienbarkeit moderner Systeme kann zu einem unreflektierten Umgang mit der scheinbar ohne intensive Einarbeitung verwendbaren Methode führen. Ein weiteres Problem der ET-Forschung ist, dass bisher keine Standards für das Reporting grundlegender Kennzahlen zur Güte der Messungen in Publikationen entwickelt wurden. Um die Qualität gewonnener Daten und damit auch die Vertrauenswürdigkeit der Ergebnisse nachvollziehbar zu machen, wäre dies ein wichtiger Schritt (vgl. hierzu auch McConkie 1981 sowie Inhoff & Radach 1998: 30). Ein weiterer für die Zukunft wünschenswerter Punkt betrifft die Aushandlung stimulusabhängiger Standardkriterien bei der Definition wichtiger Basisgrößen von Blickaufzeichnungsstudien, wie etwa Fixationen und Sakkaden. Da unterschiedliche Definitionen hier für teils stark veränderte Ergebnisse sorgen können, würde dies für eine bessere Vergleichbarkeit der in unterschiedlichen Studien gewonnenen Ergebnisse sorgen. Schließlich existiert auch ein Bedarf an Evaluationsstudien zur Validierung von weiterführenden ET-basierten Indikatoren innerhalb verschiedener Kontexte. Zusammengefasst besteht also eine der zentralen Herausforderungen für die zukünftige ET-Forschung darin, klare Standards im Umgang mit der Methode auszuhandeln und zu kommunizieren.

Literaturtipps

Duchowski, Andrew (2007). *Eye Tracking Methodology. Theory and Practice* (2. Aufl.). London: Springer.

Geise, Stephanie (2011). Eyetracking in der Kommunikations- und Medienwissenschaft: Theorie, Methode und kritische Reflexion. *Studies in Communication | Media*, o. Jg, 149–263.

Holmqvist, Kenneth, Nyström, Marcus, Andersson, Richard, Dewhurst, Richard, Jarodzka, Halszka & van de Weijer, Joost (2011). *Eye Tracking. A Comprehensive Guide to Methods and Measures*. Oxford: University Press.

Literatur

- Bente, Gary (2004). Erfassung und Analyse des Blickverhaltens. In Roland Mangold, Peter Vorderer & Gary Bente (Hrsg.), *Lehrbuch der Medienpsychologie* (S. 297–324). Göttingen: Hogrefe.
- Blake, Christopher, Schlütz, Daniela & Herbst, Berhard (2012). Die Wahrnehmung von Online-Werbung: Eine Eyetracking-Studie. In Hannes Haas & Katharina Lobinger (Hrsg.), *Qualitäten der Werbung – Qualitäten der Werbeforschung* (S. 265 – 282). Köln: von Halem.
- Bucher, Hans-Jürgen, Schumacher, Peter (2006). The relevance of attention for selecting news content. An eye-tracking study on attention patterns in the reception of print and online media. *Communications*, 31, 347–368.
- Dewhurst, Richard, Nyström, Marcus, Jarodzka, Halszka, Foulsham, Tom, Johansson, Roger & Holmqvist, Kenneth (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior Research Methods*, 44, 1079–1100.
- Duchowski, Andrew (2007). *Eye Tracking Methodology. Theory and Practice. Second Edition*. London: Springer.
- Findlay, John M. & Gilchrist, Iain, D. (1998). Eye Guidance and Visual Search. In Geoffrey Underwood (Hrsg.), *Eye Guidance in Reading and Scene Perception* (S. 295–312). Amsterdam: Elsevier.
- Garcia, Marion R. & Stark, Pegie (1991). *Eyes on the News*. St. Petersburg (Florida): Poynter.
- Geise, Stephanie (2011). Eyetracking in der Kommunikations- und Medienwissenschaft: Theorie, Methode und kritische Reflexion. *Studies in Communication | Media*, o. Jg, 149–263.
- Godijn, Richard & Theeuwes, Jan (2003). The Relationship Between Exogenous and Endogenous Saccades and Attention. In Jukka Hyönä, Ralph Radach & Heiner Deubel (Hrsg.), *The Mind's Eye. Cognitive and Applied Aspects of Eye Movement Research* (S. 3–26). Amsterdam: Elsevier.
- Grüblbauer, Johanna & Kapf, Isabella (2011). Eyetracking Hardware- und Softwarelösungen. In Svenja Hagenhoff (Hrsg.), *Arbeitsbericht Nr. 2/2011 des Instituts für Medienwirtschaft*. Fachhochschule St. Pölten.
- Hamm, Jihun, Kohler, Christian G., Gur, Ruben C. & Verma, Ragini (2011). Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders. *Journal of neuroscience methods*, 200, 237–256.
- Hammoud, Riad I. & Mulligan, Jeffrey B. (2008). Introduction to Eye Monitoring. In Riad I. Hammoud (Hrsg.), *Passive Eye Monitoring. Algorithms, Applications and Experiments* (S. 1–19). Berlin: Springer.
- Hofer, Natalie & Mayerhofer, Wolfgang (2010). Die Blickregistrierung in der Werbewirkungsforschung: Grundlagen und Ergebnisse. *der markt*, 49, 143–169.
- Holmqvist, Kenneth, Nyström, Marcus, Andersson, Richard, Dewhurst, Richard, Jarodzka, Halszka & van de Weijer, Joost (2011). *Eye Tracking. A Comprehensive Guide to Methods and Measures*. Oxford: University Press.

- Inhoff, Albrecht W. & Radach, Ralph (1998). Definition and Computation of Oculomotor Measures in the Study of Cognitive Processes. In Geoffrey Underwood (Hrsg.), *Eye Guidance in Reading and Scene Perception* (S. 29–53). Amsterdam: Elsevier.
- Jöckel, Sven, Blake, Christopher & Schlütz, Daniela (2013). The influence of age-rating label salience on perception and evaluation of media. An eye-tracking study. *Journal of Media Psychology* 25(2), 83–94.
- Joos, Markus, Rötting, Matthias & Velichkovsky, Boris M. (2003). Spezielle Verfahren I: Bewegungen des menschlichen Auges: Fakten, Methoden und innovative Anwendungen. In Gert Rickheit, Theo Herrmann & Werner Deutsch (Hrsg.), *Psycholinguistik. Ein internationales Handbuch* (S. 142–168). Berlin: de Gruyter.
- Just, Marcel A. & Carpenter, Patricia A. (1976). Eye Fixations and Cognitive Processes. *Cognitive Psychology*, 8, 441–480.
- Just, Marcel A. & Carpenter, Patricia A. (1980). A Theory of Reading: From Eye Fixations to Comprehension. *Psychological Review*, 87, 329–354.
- Komínková, Barbora, Pedersen, Marius, Hardeberg, Jon Y. & Kaplanová, Marie (2008). Comparison of eye tracking devices used on printed images. In Bernice E. Rogowitz & Thrasyvoulos N. Pappas (Hrsg.), *Proceedings on Human Vision and Electronic Imaging XIII (Vol. 6806)* (S. 68061I-1–68061I-12). Bellingham: International Society for Optical Engineering.
- Knobloch-Westerwick, Silvia & Romero, Joshua P. (2011). Body Ideals in the Media: Perceived Attainability and Social Comparison Choices. *Media Psychology* 14, 27–48.
- Larsson, Linnéa (2010). *Event detection in eye-tracking data*. Masterarbeit am Lund University Humanities Lab. Verfügbar unter: <http://www.humlab.lu.se/www-transfer/research/publications/studentpapers/Larsson10.pdf> [31. 08. 2012].
- Leven, Wilfried (1991). *Blickverhalten von Konsumenten. Grundlagen, Messung und Anwendung in der Werbeforschung*. Heidelberg: Physica.
- McConkie, George W. (1981). Evaluating and reporting data quality in eye movement research. *Behavior Research Methods & Instrumentation*, 13, 97–106.
- Mulvey, Fiona, Villanueva, Arantxa, Sliney, David, Lange, Robert, Cotmore, Sara, Donegan, Michael (2008). *D5.4 Exploration of safety issues in Eyetracking*. Verfügbar unter: <http://www.cogain.org/results/reports/COGAIN-D5.4.pdf>
- Nevalainen, Seppo & Sajaniemi, Jorma (2004). Comparison of Three Eye Tracking Devices in Psychology of Programming Research. In E. Dunican & T. Green (Hrsg.), *Proceedings of the 16th Workshop of the Psychology of Programming Interest Group* (S. 151–158).
- Nielsen, Jakob & Pernice, Kara (2010). *Eyetracking Web Usability*. Berkeley: New Riders.
- Pieters, Rik & Wedel, Michel (2004). Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-Size Effects. *Journal of Marketing*, 68, 36–50.
- Risko, Evan F. & Kingstone, Alan (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Attention, Perception, & Psychophysics*, 73, 291–296.
- Rötting, Matthias (1999a). Methoden zur Registrierung von Augenbewegungen. In Matthias Rötting & Katharina Seifert (Hrsg.), *Blickbewegungen in der Mensch-Maschine-Systemtechnik* (S. 19–34). Sinzheim: Pro Universitate.

Rötting, Matthias (1999b). Typen und Parameter von Augenbewegungen. In Matthias Rötting & Katharina Seifert (Hrsg.). *Blickbewegungen in der Mensch-Maschine-Systemtechnik* (S. 1–18). Sinzheim: Pro Universitate.

Russo, J. Edward & Rosen, Larry D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition*, 3, 267–276.

Schlütz, Daniela, Blake, Christopher, Charrier, Daniela, Dyszy, Agnes, Emde, Katharina, Hirt, Florian, Keldenich, Felix, Link, Elena & Sponer, Julia (2013). Werbung in Bildschirmspielen: Eine Eye-Tracking-Studie zur Rolle von Markenbekanntheit und Werbeintensität für die Wahrnehmung und Erinnerung von In-Game Advertising. In Thomas Schierl & Jörg Tropp (Hrsg.), *Wert und Werte der Marketing-Kommunikation* (S. 281–298). Köln: von Halem.

Yarbus, Alfred L. (1967). *Eye Movements and Vision*. New York: Plenum Press.