

Der lange Weg von Face-to-face zu CATI

→ Die Optimierung der Hörfunkabfrage in der Media Analyse

Von Dieter K. Müller*

Die Medienforschung ist im Umbruch. Die Prüfung alternativer Meßmethoden und Erhebungstechniken, ebenso wie die Optimierung der Standardinstrumente zählen zu den besonderen Herausforderungen der Medienforschung. Seit Bestehen der Arbeitsgemeinschaft Media Analyse (AG.MA) werden die Interviews für die Erhebung der Radionutzung als Face-to-face-Befragung im Haushalt der interviewten Personen durchgeführt. Mit der MA 99 Radio geht diese Tradition zu Ende. Pünktlich zum Jahrtausendwechsel wird die Radionutzung mit der CATI-Technik (CATI = Computer-Assisted Telephone Interviewing) erhoben. (1) Parallel bereiten die gedruckten Medien den Wechsel vom Paper-pencil-interview zur Pentop-Befragung (2) vor.

Erhebungsinstrumente der Radionutzung

Seit über zehn Jahren wird die Radionutzung in der MA in der bisherigen Form durchgeführt, von geringfügigen Anpassungen abgesehen. In der Zwischenzeit hat sich aber der Hörfunkmarkt entscheidend verändert. Es war deshalb erforderlich zu prüfen, ob das Erhebungsmodell noch in der Lage ist, diesen Markt zuverlässig abzubilden.

Die im Juni vorgestellte MA 99 Radio hat insgesamt 246 verschiedene Radiosender/-programme erhoben. Von diesen 246 Angeboten sind 30 Programme werbefrei, sie müssen aus methodischen Gründen ebenfalls erhoben werden. Vier national verbreitete Sender und 62 landesweite Sender/Programme stehen als Werbeträgerangebote zur Verfügung. Eine weitere Besonderheit sind die derzeit 138 lokalen Sender/Programme, die es entsprechend der Mediengesetzgebung in vier Bundesländern gibt. Diese Besonderheit der Angebotsstruktur muß ein Erhebungsmodell natürlich berücksichtigen, wenn daraus im Ergebnis Planungsdaten für nationale, regionale und gegebenenfalls auch lokale Belegungseinheiten entstehen sollen.

In diesen Zahlen drückt sich die Dynamik des Hörfunkmarktes aus, denn beispielsweise die MA 87 hatte nur die Anforderung, 44 Sender bzw. Programme valide abzubilden. Damals reichte ein einziger Fragebogen, um dieses Angebot vollständig zu erheben. In der MA 99 Radio waren insgesamt 51 regionale Fragebogensplits im Einsatz, die sich durch die jeweils vorgegebenen Hörfunksender bzw. -programme unterscheiden.

Die Hörfunknutzung wird anhand einer persönlich-mündlichen Tagesablaufbefragung (Stichtagsbefragung) erfaßt. Im Tagesablauf werden Leittätigkeiten und Mediennutzung in Viertelstunden für den Zeitraum 5.00 bis 24.00 Uhr erhoben. Neben der Erfassung des Radiohörens wird im Tagesablauf auch die Nutzung von Komplementärmedien erfaßt. Als Teilergebnis der noch zu erläuternden Methodentests wurde in der MA 97 erstmalig auf die Erfassung der Fernsehnutzung nach Einzelsendern verzichtet, die vorher ausschnittsweise erfragt wurden, und statt dessen nur die allgemeine Fernsehnutzung ermittelt. Dafür wurde die PC-Nutzung in den Fragebogen aufgenommen, beibehalten wurde die Nutzungsabfrage von Video, Schallplatten, Tonbändern, Kassetten und CD.

Das bisher eingesetzte Tagesablaufmodell ist demnach kein statisches Instrument, sondern kontinuierlich den Markterfordernissen angepaßt worden. Eine wichtige Besonderheit der bisherigen MA-Erhebung stellt zudem die im gleichen Interview stattfindende Erhebung der Tageszeitungsnutzung dar.

Neben der MA wurden und werden zusätzlich verschiedene Untersuchungen durchgeführt, die ebenfalls Indikatoren der Radionutzung erheben. Auch dies war ein Anlaß zur vergleichenden Überprüfung der einzelnen Instrumente. Es handelt sich dabei um:

- sogenannte regionale E.M.A.s (Elektronische Medienanalysen), die wie die MA als persönlich-mündliche Interviews durchgeführt wurden. Die sogenannte E.M.A. NRW wurde vor dem Hintergrund der Entwicklungen in der AG.MA mit Beginn des Jahres 1999 auf das CATI-Verfahren umgestellt. Allerdings sind die methodischen Vorgaben deutlich von denen entfernt, die inzwischen von der AG.MA für einen CATI-Standard formuliert wurden, insofern ist dieses Instrument derzeit auch nicht MA-kompatibel. Der Unterschied der regionalen E.M.A.s zur MA war und ist im wesentlichen der reduzierte Fragenumfang, der sich auf die Abbildung der Hörfunknutzung konzentriert.
- die Erfassung der Hörfunknutzung mit Diaries (Tagebuch). Dieses Verfahren wurde vor allem von Infratest entwickelt und war lange Jahre der Erhebungsstandard der Funkanalyse Bayern (FAB). Ebenfalls mit Beginn des Jahres 1999 wurde auch diese Erhebungstechnik für die FAB aufgegeben und durch Telefoninterviews ersetzt. Auch dies erfolgte im Vorgriff auf die noch nicht endgültig festgelegten methodischen Rahmenbedingungen für eine CATI-MA, womit auch für dieses Instrument derzeit die MA-Vergleichbarkeit nur eingeschränkt gegeben ist.
- sogenannte Trendstudien der einzelnen Landesrundfunkanstalten und vieler Privatsender. Diese Instrumentarien wurden seit Ende der 80er Jahre von den Landesrundfunkanstalten entwickelt und mit zunehmender Bedeutung für die strategische Programmplanung einge-

Tagesablaufbefragung erfaßt neben Radio auch Komplementärmedien

Alternative Erhebungsinstrumente der Radionutzung basieren inzwischen auf Telefoninterviews

Dynamischer Hörfunkmarkt erfordert Anpassung der Media Analyse

* ARD-Werbung Sales & Services.

setzt. Als Ergänzungsstudien zur MA beinhalten sie andere Befragungsschwerpunkte und haben auch andere methodische Anforderungen zu erfüllen. Allerdings wurden diese Studien von Beginn an per CATI durchgeführt. (3)

- die Erfahrungen im europäischen Ausland. (4) In immer mehr Nachbarländern etablierte sich CATI zum Standard der Ermittlung der Radio-nutzung.

**Ergebnisvergleiche
verschiedener
Instrumente zeigen
Unterschiede**

Aus Ergebnisvergleichen zwischen E.M.A.s, Diaries und Telefontrends mit der MA sind abweichende Ergebnisniveaus bekannt. Die Ursachen konnten methodisch bisher nicht zuverlässig zugeordnet werden, da es zwischen den Instrumenten selten parallele Erhebungszeiträume gab und die methodischen Anforderungen sehr unterschiedlich waren. Mit dem 1995 durchgeführten Methodentest sollten erstmals die unterschiedlichen Erhebungstechniken vergleichend getestet werden.

**Prüfung verschiedener Erhebungsmethoden:
Methodentest 1995**

**Aufwendiger
Methodentest als
erste Stufe der
MA-Optimierung
durchgeführt**

Den Impuls für das umfangreichste und mit bisher insgesamt rund 2 Mio DM teuerste Forschungsexperiment in Deutschland gab die Werbewirtschaft. Werbungtreibende und Agenturen haben Ende 1994 einen Forderungskatalog an die gedruckten und elektronischen Medien in der AGMA gerichtet. Kernpunkt war die Forderung nach flexiblen, aktuellen und medienspezifischen Analysen, die den komplexer gewordenen Bedürfnissen der Nutzer – also der Werbewirtschaft – gerecht werden sollen. Die Hörfunkanbieter haben diesen Forderungskatalog ausführlich mit den Nutzern diskutiert. Dabei wurde gemeinsam die valide Sender- und Nutzungserfassung als vordringlich zu realisierende Aufgabe festgelegt.

Medienspezifische Bedürfnisse einerseits sowie kundenspezifische Anforderungen andererseits gaben also Anlaß, nach fast zehnjährigem Bestehen des Partnerschaftsmodells (5) eine Bestandsaufnahme der damaligen Gegebenheiten zu machen, um zukünftigen Erfordernissen gerecht zu werden.

**Acht verschiedene
Modellvarianten
getestet**

Die Optimierung der Hörfunkabfrage in der Media Analyse (MA) war die zentrale Aufgabe des Methodentests 1995, der auf internationaler Ebene als GRAMMS '95 (German Radio Multi-Method Survey 95) große Beachtung fand. (6) Dazu wurde ein aufwendiges Multimethoden-Design konzipiert, das insgesamt acht verschiedene Modellvarianten umfaßt. Getestet wurden fünf Face-to-face-Modelle, zwei Diary-Varianten und eine telefonische Befragung (siehe dazu Steckbrief der Modellvarianten in Tabelle 1). Die CATI-Variante wurde zunächst von der ARD alleine beauftragt. Die privaten Radiosender/Vermarkter schlossen sich erst nachträglich dieser Testalternative an.

Als Kontrollbefragung zur Validierung der in den alternativen Methodentests ermittelten Hörfunkreichweiten wurde ein telefonischer Coincidental Check vorgenommen. Zu zwölf Meßzeitpunkten mit jeweils 500 Interviews wurden die

Hörfunkreichweiten ermittelt. Damit wird ermöglicht, die Zuverlässigkeit der Hörfunkreichweiten der acht Testmodelle einzuordnen.

Die wesentlichen Befunde sind in Tabelle 2 wiedergegeben. Neben Ergebnissen, die sich auf den Vergleich von Nutzungswerten beziehen, stehen vor allem methodische Auswertungen im Vordergrund, die mögliche Unterschiede erklären helfen. Die Ausschöpfungsquote gibt Auskunft darüber, wie gut oder wie schlecht im jeweiligen Modell die Befragten erreicht wurden. Die Ausschöpfung liegt für alle getesteten acht Modelle deutlich über 70 Prozent.

Die leicht unterschiedlichen Werte für die Interviewdauer lassen sich für die Face-to-face-Modelle durch den unterschiedlichen Umfang der Befragungsinhalte erklären. Die inhaltlich reduzierten Modelle 2 und 3 dauern entsprechend kürzer als die umfangreicheren Modelle 1, 4 und 5. Die telefonisch durchgeführten Interviews liegen mit durchschnittlich 16 Minuten deutlich unter dem Niveau der persönlichen Befragungsvarianten. Dies läßt sich mit dem stark reduzierten Befragungsumfang erklären. Darüber hinaus haben die Befragungserfahrungen gezeigt, daß sich die sehr detaillierte Abfrage der Hörfunknutzung nach Viertelstunden im Tagesablaufschema durchaus praktikabel am Telefon bewerkstelligen läßt.

Zudem war zu prüfen, inwiefern sich die beiden Verfahren der Stichprobenziehung auf die Zusammensetzung der Stichproben auswirken. Die Sample Points (Stimmbezirke) der Face-to-face-Modelle wurden aus dem ADM-Stimmbezirksdatenband gezogen. Die Auswahl der Haushalte erfolgte nach dem sogenannten Random-Route-Verfahren. Bei den Diary- und CATI-Modellen erfolgte die Ziehung auf Basis des aktuellen CD-ROM-Telefonverzeichnis der Deutschen Telekom.

Es zeigte sich, daß die per Telefonstichprobe rekrutierten Befragten im Vergleich zur Face-to-face-Stichprobe jünger, höher gebildet und im größeren Umfang berufstätig sind. Dieses Ergebnis spiegelt den allgemein bekannten Zusammenhang zwischen Alter, Schulbildung und Berufstätigkeit wider (vgl. Tabelle 3).

Wieviel Prozent der Grundgesamtheit innerhalb der letzten 14 Tage Radio gehört haben, läßt sich anhand des Weitesten Hörerkreises (WHK) feststellen. Dieser eher „weiche“ Indikator für die Hörfunknutzung überschreitet für alle Modelle die 90-Prozent-Marke.

Wie differenziert einzelne Radiosender gehört werden, darüber gibt die durchschnittliche Anzahl gehörter Programme bezogen auf den Weitesten Hörerkreis Auskunft. Hier zeigen sich deutliche Unterschiede für die einzelnen Modelle. Während beispielsweise die per Telefon befragten Personen in den letzten 14 Tagen durchschnittlich vier Pro-

Methodische Ergebnisse stehen neben Nutzungswerten im Vordergrund

Unterschiede bei Zusammensetzung der Stichprobe

Zahl durchschnittlich gehörter Programme im WHK variiert

① **Methodentest 1995: Steckbrief der Modellvarianten**

| Face-to-face-Modelle¹⁾

- Modell 1 Identisch mit der MA 95.
- Modell 2 Zu prüfen war, ob einzelne Befragungsteile des MA-Interviews, die nicht ursächlich im Zusammenhang mit der Erhebung der Hörfunknutzung stehen, die Erfassung der Hörfunknutzung beeinträchtigen. Dabei handelt es sich zum einen um die Abfrage der Tageszeitungsnutzung, die in der Elektronik-Tranche wie in der Print-Tranche der MA identisch erfolgt.
- Modell 3 Zum anderen war zu überprüfen, ob die Abfrage der Fernsehnutzung nach Einzelsendern im Tagesablauf einen Einfluß auf die Erfassung der Hörfunknutzung hat.
- Modell 4 Die zusätzliche Erfassung von Produktdaten und Einstellungen war eine der zentralen Kundenforderungen. Deshalb wurden in zwei weiteren Testvarianten Marketingzusatzfragen erhoben. Testmodell 4 entspricht Modell 3 und ist mit Marketinginformationen (Produktinformationsinteresse) ergänzt.
- Modell 5 Entspricht in den Befragungsinhalten den damaligen E.M.A.s. Im Gegensatz zu den Modellen 1 bis 4 verzichtet Modell 5 auf alle für das Fernsehen relevanten Fragestellungen, auch auf die für die Fusion von Meßdaten aus dem TV-Panel der AGF/GfK in den Datensatz der MA erforderlichen.

| Diary-Modelle²⁾

- Modell 6 Die Erfassung der Hörfunknutzung per Tagebuch bedeutet, daß die Befragungsperson ihre Hörfunknutzung für ca. eine Woche in einem Tagebuch registriert. Die dortige Erfassung der Hörfunknutzung erfolgt in einem dem MA-Interview vergleichbaren Tagesablauf.
- Modell 7 Neben der reinen Erfassung der Hörfunknutzung per Tagebuch bot es sich an, den Test um eine sogenannte Single-Source-Variante zu ergänzen. Damit war die Absicht verbunden, in Ergänzung der Erfassung der Hörfunknutzung eine Nachbefragung zu den Konsum- und Verbrauchsgewohnheiten durchzuführen.

| Telefonbefragung

- Modell 8 In der Telefonbefragung (CATI-Interview) wurde die Hörfunknutzung anhand eines vergleichbaren Tagesablaufes für den gestrigen Tag erfragt. Im Testzeitraum wurden 1 000 Telefoninterviews durchgeführt. Das telefonische Auswahlverfahren entspricht dem der Diaries.

1) Für alle Face-to-face Testmodelle wurden jeweils 1 000 Interviews durchgeführt. Erhebungszeitraum war 15. Mai bis 9. Juli 1995. Grundgesamtheit und Auswahlverfahren entsprachen der MA. Als Testgebiet wurde das Bundesland Hessen ausgewählt, das in seiner Struktur etwa dem Bundesdurchschnitt entspricht.
 2) Für jedes Testmodell wurden rund 1 000 Diaries ausgefüllt. Grundgesamtheit und Befragungszeitraum waren identisch mit den Face-to-face-Testmodellen. Im Auswahlverfahren besteht ein Unterschied: Die Auswahl der Haushalte erfolgte telefonisch. Jede Person des ausgewählten Haushalts, die älter als 14 Jahre ist, soll ein Tagebuch ausfüllen. Ergebnisunterschiede zu dem klassischen Face-to-face-Interview können also auch durch das Auswahlverfahren bedingt sein.

② **Methodentest 1995: Zentrale Ergebnisse**

Modell	Face-to-face					Diary		CATI
	1	2	3	4	5	6	7	
Methodische Aspekte								
Ausschöpfung in %	72,4	74,9	74,4	75,4	75,3	72,7	72,9	73,8
Interviewdauer in Min.	37	34	35	38	37	-	-	16
Nutzungsindikatoren¹⁾								
Weitester Hörerkreis (WHK) in %	91,5	90,5	91,0	93,7	96,0	94,6	94,8	94,4
Ø Anzahl gehörter Sender im WHK	2,5	2,4	2,7	2,4	2,6	3,1	3,3	4,0
Hörer gestern in %	75,4	71,1	76,3	76,9	80,3	74,5	74,8	81,3
Ø Anzahl „gestern gehörte“ Sender	1,2	1,2	1,3	1,2	1,3	1,2	1,2	1,5
Ø Anzahl Hörvorgänge pro Tag	1,7	1,8	1,7	1,8	1,8	1,9	2,0	3,2
Hördauer in Min.	141	149	162	164	174	144	146	223

1) Radio gesamt, Mo-So, Pers. ab 14 Jahre.

③ Methodentest 1995: Stichprobenstrukturen ungewichtet

in %	Solldaten (Hessen)	Face-to-face Modelle 1-5	Diary Modelle 6-7	CATI Modell 8
Geschlecht				
Männer	48	47	49	47
Frauen	52	53	51	53
Alter in Jahren				
14-19	6	6	8	6
20-29	17	14	15	21
30-39	17	18	22	24
40-49	15	15	20	17
50-59	17	16	17	14
60-69	13	16	11	11
70 und älter	14	16	7	8
Bildung				
Hauptschule	54	51	41	39
weiterf. Schule	27	28	33	31
Abitur	19	10	12	15
Studium		10	13	15
Berufstätig	52	43	56	59

gramme gehört haben, ergeben sich für die persönlich Befragten Werte zwischen 2,4 und 2,6. Die Diary-Modelle ermitteln durchschnittlich 3,1 resp. 3,3 Programme.

Bei CATI höherer Erinnerungswert an gestern gehörte Programme

Der sogenannte Hörer-gestern-Wert gibt die tägliche Nettoreichweite des Hörfunks an. Er liegt für das Modell 5 mit 80,3 Prozent auf ähnlich hohem Niveau wie in der Telefonbefragung (81,3 Prozent).

Die beiden Tagebuchmodelle liegen in etwa auf MA-Niveau. Auch für die gestern gehörten Programme läßt sich – analog zum Weitesten Hörerkreis – die durchschnittliche Anzahl der Programme ermitteln. Wiederum zeigt sich, daß bei Modell 8 CATI eine höhere Anzahl gestern gehörter Programme erinnert wird. Die beiden Diary-Modelle unterscheiden sich dagegen nicht vom MA-Modell. Dies sind schon erste Hinweise, daß in einer telefonischen Befragung die Hörfunknutzung weitaus differenzierter abgebildet wird als in den anderen Verfahren.

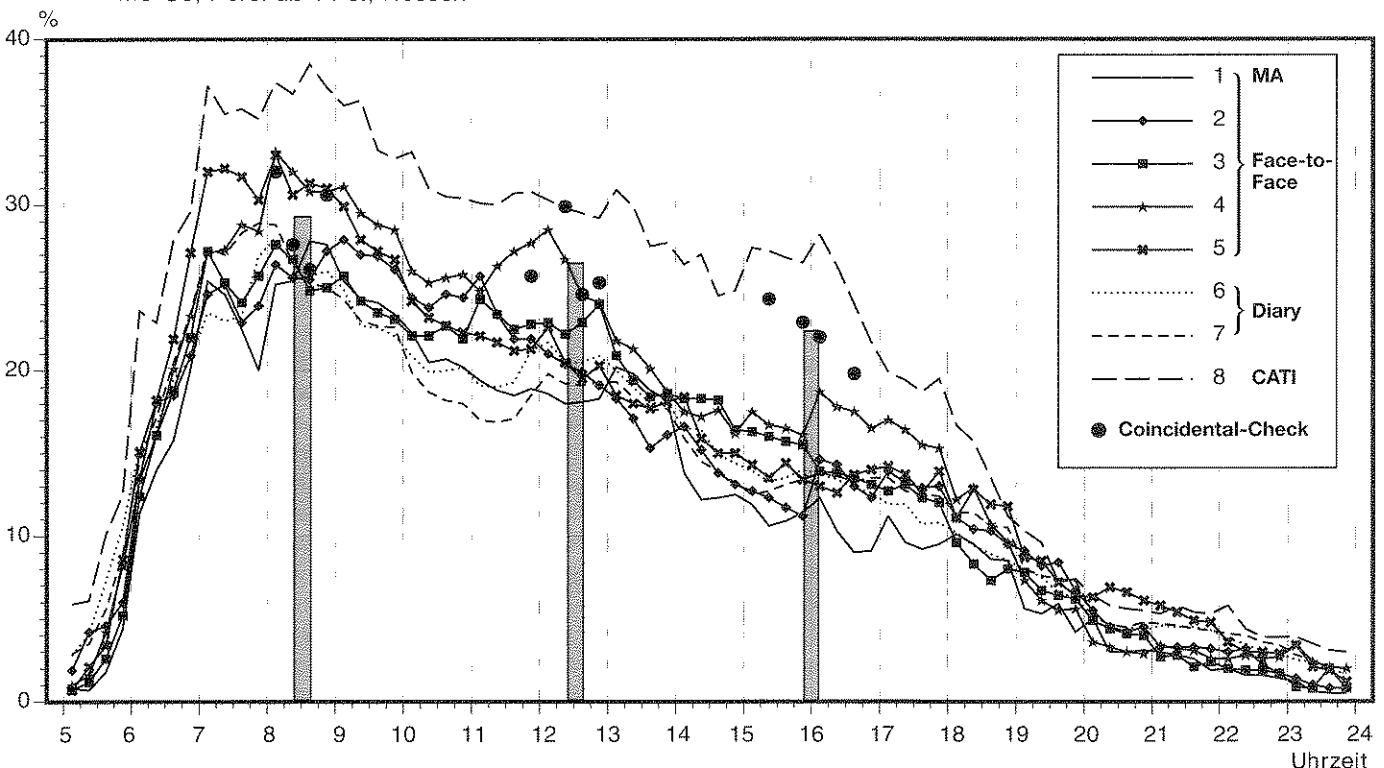
Vergleicht man die tägliche Hördauer, werden die sich bereits abzeichnenden Unterschiede noch deutlicher. Die telefonische Befragung produzierte mit einer durchschnittlichen Hördauer von 223 Minuten den höchsten Wert und lag damit rund 60 Prozent über dem Basiswert von Modell 1 (MA). Die Face-to-face-Modelle 3, 4 und 5, in denen die Fernsehnutzung im Tagesablauf nur pauschal als Sammelkategorie erhoben wurde, liefern im Vergleich zu den Modellen 1 und 2 sowie den Tagebuchversionen mit jeweils senderspezifischer Erhebung des Fernsehens tendenziell höhere Hördauerwerte.

Wie sich die Hörfunknutzung über den Tag hinweg verteilt, zeigen die Viertelstundenreichweiten (vgl. Abbildung 1). Als Vergleichsmaßstab sind die Reichweitenwerte des Coincidental Checks und die Durchschnittswerte aus den jeweils vier Meßzeitpunkten morgens, mittags und nachmittags als Balkendiagramm abgetragen. Die telefonisch er-

Wert für Radiohördauer steigt, wenn Fernsehnutzung pauschal erhoben wird

Telefonisch erfragte Hörfunkreichweiten sind im Tagesverlauf höher

Abb.1 Methodentest 1995
Viertelstundenreichweiten / Radio gesamt
Mo-So, Pers. ab 14 J., Hessen



fragten Hörfunkreichweiten lagen über den gesamten Tagesverlauf hinweg über denen sämtlicher Vergleichsmodelle und auch höher als die im Coincidental Check ermittelten durchschnittlichen Nutzungswerte. Die Tagebuchmodelle lagen morgens auf MA-Niveau, mittags leicht sowie nachmittags deutlicher darüber.

Anhand der Abbildung zeigt sich aber auch, daß diejenigen Erhebungsvarianten, die die Fernsehnutzung im Tagesablauf lediglich als gesamthafte Tätigkeit abfragen, dem Maßstab des Coincidental Checks am nächsten kommen. Offensichtlich wird hier einmal mehr die Erfahrung bestätigt, daß sich die detaillierte Abfrage weiterer Mediengattungen im Tagesablauf reichweitenmindernd auf den Hörfunk auswirkt. Bei der Abfrage der Fernseheinzelnder wird davon ausgegangen, daß sich die Dominanz dieses Mediums auf die Erinnerungsleistung der Befragten niederschlägt. Das gelegentlich unbewußt und häufig nebenbei gehörte Radio wird besser und vollständiger erinnert, wenn der Befragte nicht durch eine differenzierte Abfrage der Fernsehnutzung abgelenkt wird. Diese Vermutung wird bestätigt, wenn man die erfragte Fernsehnutzung mit der von der GfK gemessenen vergleicht. Die telefonisch erfragten Reichweiten kommen den Meßdaten tendenziell am nächsten.

Face-to-face-MA scheint Radio-nutzung zu unterschätzen

Radio wird überwiegend als Hintergrundmedium genutzt. Ursache ist die nahezu unbeschränkte Verfügbarkeit, die es erlaubt, Radio an jedem Ort und zu jeder Zeit hören zu können. Aus diesem Grunde ist es auch wichtig zu überprüfen, inwieweit sich die Radionutzung in Verbindung mit einzelnen Tätigkeiten bei den verschiedenen Testvarianten unterscheidet. Es zeigt sich, daß die Diary-Varianten im Vergleich zu den Face-to-face-Erhebungen eine signifikant höhere Außer-Haus-Radionutzung (Indikator Hördauer) dokumentieren, die inhäusige Radionutzung jedoch ebenso deutlich niedriger liegt. Diese Tendenz gilt auch für CATI. Die bereits festgestellte höhere Hördauer bei der CATI-Erhebung resultiert somit in erster Linie aus der besseren Erfassung der Außer-Haus-Radionutzung (vgl. Tabelle 4). Die Diary-Erhebungstechnik und insbesondere die CATI-Technik sind offensichtlich besser als die MA geeignet, auch die Außer-Haus-Radionutzung zuverlässig abzubilden. Die Face-to-face-MA scheint demnach offenkundige Schwächen bei der Abbildung mobiler Mediennutzungsvorgänge zu haben und somit insgesamt die Radionutzung eher zu unterschätzen.

Die Ergebnisse des Methodentests 1995 lassen sich folgendermaßen zusammenfassen:

1. Die gleichzeitige Erhebung der Tageszeitungsnutzung hat keinen wesentlichen Einfluß auf die Erinnerungsleistung hinsichtlich der Hörfunknutzung.
2. Bei den Face-to-face-Modellen weisen die Varianten, bei denen die Fernsehnutzung nicht nach Einzelsendern erhoben wurde, eine geringfügig höhere Hörfunknutzung auf. Dies läßt sich dadurch erklären, daß die starke Fokussierung der Befragten auf Fernsehen wegfällt und diese sich besser auf ihre gestrige Hörfunknutzung konzentrieren können. Das gelegentlich unbewußt und häufig nebenbei gehörte Medium Radio wird besser und vollständiger erinnert, wenn die Befragten nicht durch eine differenzierte Abfrage der Fernsehnutzung abgelenkt werden. Auf die gleichzeitige Abfrage von Fernseheinzelndern in einer Tagesabläuferhebung, die der Ermittlung der Radionutzung dient, sollte deshalb verzichtet werden.
3. Zusatzfragen, die eine weitere Zielgruppenbeschreibung für das Marketing ermöglichen, hatten bei dem hier festgelegten Umfang keinen Einfluß auf das Hörfunkreichweiteniveau. Die Erhebung von Radio- oder Tageszeitungsreichweiten könnte deshalb – sofern das Interview freie Kapazitäten hat – mit Marketingzusatzfragen ergänzt werden.
4. Die Abfrage via Tagebuch weist weitgehend identische Nutzungsvorgänge für Radio auf wie die Face-to-face-Befragung. Die Diary-Testvarianten haben somit auch die in der Face-to-face-MA ermittelten Daten zur Radionutzung validiert.
5. Die Telefonbefragung eignet sich grundsätzlich zur Erhebung der Hörfunknutzung. Es hat sich gezeigt, daß bei der telefonischen Befragung die Nutzung differenzierter erinnert und angegeben wird. Ein Erklärungsansatz hierfür könnten auch die unterschiedlichen demografischen Zusammensetzungen der Stichproben sein. In der Telefonstichprobe sind die Jüngeren, die formal besser Gebildeten und die Berufstätigen leicht überrepräsentiert. Im Gegensatz dazu sind diese Gruppen bei den anderen Erhebungstechniken unterrepräsentiert. Die Telefonbefragung ist demnach ein mediengerechtes Modell für Radio und eine mögliche Alternative zur Face-to-face-MA.
6. Insgesamt bestätigt der Test, daß die Hörfunknutzung mit klassischen Befragungstechniken durch die MA zuverlässig abgebildet werden kann.

**Fazit
Methodentest 1995:
Mediengerechte
Optimierung**

④ **Methodentest 1995: Hördauer in Minuten**
Radio gesamt, Mo-So, Pers. ab 14 J.

	Face-to-face Modelle 1-5	Diary Modelle 6-7	CATI Modell 8
Hördauer gesamt	158	145	223
davon: zu Hause	113	89	128
außer Haus	43	56	90
beim Autofahren	18	23	34
während der Arbeit	20	26	46

Positive Bewertung der CATI-Technik, aber Optimierungen erforderlich

Die abschließende Bewertung im Kreis der Radiosender und auch innerhalb der AG.MA war einheitlich: Die CATI-Technik wurde als geeignete Alternative zum bisherigen Erhebungsinstrument angesehen. Allerdings war unstrittig, daß es noch einer Reihe von Optimierungen und zusätzlicher Prüfungen bedürfe, um CATI im Regelbetrieb als Ersatz für die Face-to-face-MA einsetzen zu können. Darüber hinaus hat der Methodentest 1995 grundsätzlich die Zuverlässigkeit der Radio-MA bestätigt, zudem aber auch verdeutlicht, daß unterschiedliche Erhebungstechniken, die alle auf der Erinnerungsleistung der Befragten basieren, bei den Indikatoren der Hörfunknutzung vergleichbare Nutzungsmuster aufweisen. Insofern werden mit Recall-Techniken (Face-to-face oder CATI) im Prinzip immer ähnliche Nutzungsstrukturen mit geringen Niveauverschiebungen reproduziert. Solange Radiometersysteme (7) methodisch und technisch noch nicht serienreif sind, werden deshalb die bisher bekannten Relationen der Radionutzung zumindest erhebungstechnisch keine gravierenden Veränderungen erfahren.

In Abstimmung mit den Kundenvertretern in der AG.MA und mit OWM und GWA wurde daraufhin für die folgende MA 97 weiterhin eine Face-to-face-Befragung mit aus dem Test abgeleiteten Änderungen beim Fragebogen – unter anderem keine Fernseheinzelsenderabfrage im Tagesablauf – festgelegt. Gleichzeitig wurde auf Kundenwunsch der MA-Fragebogen im Hinblick auf die Nutzung von PC und Multimedia ergänzt.

Optimierungen der CATI-Technik für die Media Analyse

Schwierige Umsetzung des A-4-Fragebogens in PC-Bildschirmformat

Der Methodentest 1995 hat vor allem hinsichtlich der CATI-Technik eine Vielzahl von Fragen offen gelassen, auch deshalb, weil eine Optimierung des CATI-Instruments nicht zu der Aufgabenstellung des Methodentests zählte. Ein kritischer Punkt ist vor allem die Abfrage der Radionutzung für den gestrigen Tag im sogenannten Tagesablaufschemata. Was in der gedruckten Version den Umfang von drei aufklappbaren Seiten im A-4-Format hat, muß so umgesetzt werden, daß es auf einem kleinen PC-Bildschirm ohne Informationsverlust während des Interviews schnell und zügig handhabbar ist.

Komplexes Tagesablaufschemata

Der Fragebogen zum Tagesablauf ist ein höchst komplexes Befragungsinstrument, in dem neben der senderbezogenen Zuordnung der Radionutzung derzeit insgesamt elf Kategorien für Aktivitäten und vier Kategorien für die Nutzung von Komplementärmedien (8) einzutragen sind. Die Struktur dieses Tagesablaufschemas basiert in der heutigen Form im wesentlichen auf Erkenntnissen von Methodentests aus dem Jahre 1988. (9) Ein wesentlicher Befund war dabei die Bedeutung der Tätigkeitsabfrage für die Hörfunknutzung. Die Tätigkeitsabfrage zwingt einerseits den Interviewer, mit dem Befragten zusammen den gesamten gestrigen Tag durchzugehen, andererseits hat sie sich als unverzichtbare Erinnerungstütze für die gestrige Hörfunknutzung und deren zeitliche Zuordnung erwiesen.

Es galt also zunächst, dieses Tagesablaufschemata ohne wesentlichen Informationsverlust in eine praktikable CATI-Version umzusetzen. Als Voraussetzung für einen weiteren Feldtest mußte zudem das CATI-Interview diejenigen Befragungsinhalte integrieren, die in einem MA-Interview aus unterschiedlichen Gründen unverzichtbar sind (z.B. Fusionsbindeglieder für die Erstellung einer Intermedia-Datei). Diese Anforderungen wurden beim Methodentest 1995 nicht berücksichtigt, weshalb die Länge des CATI-Interviews (Modell 8) auch nur durchschnittlich 16 Minuten betrug.

Feinabstimmung und Fragebogenerstellung für das Telefoninterview wurden zunächst auf Basis der Software Quancept vorgenommen, da diese Software in den meisten Marktforschungsinstituten vorhanden ist und bereits für Telefoninterviews eingesetzt wird. Damit sollte sichergestellt werden, daß bei einer möglichen CATI-MA ein Grundprinzip der AG.MA beibehalten werden kann: die Feldarbeit auf mehrere Marktforschungsinstitute aufzuteilen und mit einem für alle Institute verbindlichen einheitlichen Fragebogen zu arbeiten, um mögliche Institutshandschriften zu minimieren.

Der so entwickelte neue CATI-Fragebogen wurde im April 1997 einem ersten Handlingtest unterzogen. 22 Interviewer führten 347 Interviews in Hessen durch und berichteten über ihre Erfahrungen. Der neue Fragebogen wurde als leicht verständlich eingestuft, zwei Drittel der Befragten fanden das Interview interessant, drei Viertel hatten sogar Spaß daran. Vielversprechend war auch die subjektive Unterschätzung der Interviewdauer durch die Befragten. Während das Interview tatsächlich im Durchschnitt 26 Minuten lang war, wurde die subjektiv empfundene Interviewdauer auf durchschnittlich 21 Minuten geschätzt. Die Voraussetzungen für einen Feldtest waren also erfüllt.

Aufgabenstellung des Feldtests von Herbst 1997 war vor allem die Beantwortung methodischer Fragen. Neben der Zielsetzung, erstmals Ausschöpfungsquoten für Telefoninterviews zu erhalten, galt das Hauptinteresse der Befragtenstruktur einer Telefonstichprobe und deren möglichen Auswirkungen auf die Medienergebnisse. Daneben sollte natürlich auch das Handling des CATI-Interviews auf breiterer Basis und die Koordination mehrerer Institute erprobt werden.

Zwischen dem 27. Oktober und dem 17. Dezember 1997 wurden von drei Instituten insgesamt 3114 CATI-Interviews durchgeführt. Als Befragungsgebiet wurde Bayern gewählt, da die dortige Hörfunklandschaft mit landesweiten Radioprogrammen und einer Vielzahl von Lokalstationen für einen derartigen Test als besonders geeignet erschien. Die Auswahl der Befragungspersonen geschah in einem zweistufigen Zufallsverfahren:

Einheitlicher Fragebogen für verschiedene Institute

CATI-Pretest im April 1997 vielversprechend

CATI-Feldtest im Herbst 1997 sollte methodische Fragen beantworten

5 CATI Feldtest Herbst 1997: Ausschöpfung

	CATI Feldtest Herbst 1997			zum Vergleich MA 97 ET Bayern		
	absolut	in % ¹⁾	in % ²⁾	absolut	in % ¹⁾	in % ²⁾
Telefonnummern/Adressen gesamt	4 552	100,0	-	5 320	100,0	-
Neutrale Ausfälle	383	8,4	-	10	0,2	-
Bereinigte Stichprobe	4 196	91,6	100,0	5 310	99,8	100,0
Summe Interviews	3 114	-	74,7	4 607	-	86,8
Summe erreichte Haushalte	3 020	-	72,4	4 376	-	82,4
Ausschöpfung in Haushalten mit 4 u. mehr Personen						
Erfaßte Zielpersonen	1 682	100,0	-	738	100,0	-
Durchgeführte Interviews	935	55,6	-	600	81,3	-
2 Interviews im Haushalt	188	-	20,1	462	-	77,0
1 Interview im Haushalt	747	-	79,9	138	-	25,0

1) Gesamt.
2) Basis: bereinigte Stichprobe.

Die Auswahl des Haushaltes erfolgte als Telefonnummer auf Basis der aktuellen CD-ROM der Deutschen Telekom nach vorherigem Aussortieren der Firmenadressen. Den Instituten wurde dabei eine feste Anzahl von Telefonnummern vorgegeben. Die Auswahl der Befragungspersonen im Haushalt selbst wurde wie in der MA anhand eines sogenannten Schwedenschlüssels (10) vorgenommen. In Haushalten mit vier und mehr Personen ab 14 Jahren wurde, wie im MA-Interview, auch eine zweite Befragungsperson ausgewählt. Die Interviews wurden in den CATI-üblichen Vorabendzeiten durchgeführt, mit der Möglichkeit von Terminvereinbarungen zu beliebigen Zeiten. Bei mehreren erfolglosen Kontakten einer Telefonnummer wurde dieser Haushalt zur Erzielung eines Interviews dann auch tagsüber angerufen.

Ausschöpfungsgrad lag bei 72,4 %

Die Qualität einer Random-Stichprobe hängt von der Ausschöpfung des Bruttoansatzes ab, da sogenannte qualitätsneutrale und nichtneutrale Ausfälle (11) zu Verzerrungen des Gesamtergebnisses führen können. Der Ausschöpfungsgrad der bei diesem Feldtest eingesetzten Telefonstichprobe lag mit 72,4 Prozent über dem im ZAW-Rahmenschema (12) vorgegebenen Schwellenwert von 70 Prozent. Damit wurde der Nachweis geführt, daß auch mit einer Telefonstichprobe die Voraussetzungen für die repräsentative Ermittlung der Hörfunknutzung erfüllt werden können.

Die Ausschöpfungsquote bisheriger MA-Face-to-face-Interviews von meist über 80 Prozent wurde damit allerdings nicht erreicht. Dies ist bei der sehr kurzen Feldzeit von sieben Wochen für die Testdurchführung auch nicht verwunderlich. Die Erhebungszeit einer MA-Welle ist mit zwölf Wochen fast doppelt so lang. Der erzielte Ausschöpfungsgrad von über 70 Prozent bei relativ kurzer Feldzeit wurde deshalb als ein gutes, optimierbares Niveau bewertet.

Demgegenüber blieb die Ausschöpfung des Zweitinterviews in Vier- und Mehrpersonenhaushalten im Vergleich zur MA unterdurchschnittlich (vgl. Tabelle 5). Dieses Ergebnis dürfte sicherlich zum einen aus der eingeschränkten Erreichbarkeit der konkreten Zielpersonen und deren Kooperationsbereitschaft, zum anderen aus der im Gegensatz zu Face-to-face-Befragungen totalen Kontrolle des Interviewers im CATI-Studio resultieren. Als Konsequenz wurde festgehalten, bei einer CATI-MA auf ein Zweitinterview in Großhaushalten zu verzichten. Insgesamt zeichneten die drei beteiligten Institute ein positives Bild für die Realisierbarkeit der Radioerhebung via CATI.

Zweitinterviews unterdurchschnittlich ausgeschöpft

Um eine Befragungsperson zu erreichen, waren im Durchschnitt 4,8 Telefonanrufe erforderlich. 20 Prozent der Interviews konnten bereits mit dem ersten Telefonkontakt erzielt werden. Nach dem dritten Telefonkontakt war bereits die Hälfte der Interviews durchgeführt. Bei 10 Prozent der Interviews waren mehr als zehn Telefonkontakte zur Durchführung eines erfolgreichen Interviews erforderlich. Die durchschnittliche Interviewdauer lag bei 22 Minuten. Das aus dem Pretest bekannte Phänomen der subjektiven Unterschätzung wiederholte sich, die Befragungspersonen schätzten die Interviewdauer mit knapp 17 Minuten erfreulich kurz ein.

Bei der nachfolgenden Ergebnisdarstellung ist auch zu berücksichtigen, daß bei der Stichprobenanlage nur im Telefonverzeichnis eingetragene Haushalte ausgewählt werden konnten, womit ca. 31 Prozent der bayerischen Haushalte in der Stichprobe nicht berücksichtigt werden konnten. Zur Überprüfung der möglichen Ausschöpfungsrate mußte dieser Aspekt allerdings in Kauf genommen werden. Zudem lag der Interviewschwerpunkt in der Zeit zwischen 15.00 und 20.30 Uhr, wodurch die Erreichbarkeit von Personen, die beruflich außer Haus tätig sind, verbessert wird.

Die Ergebnisse des CATI-Tests wurden, soweit möglich, mit den Daten der ersten Welle der Radiotranche zur MA 98 in Bayern verglichen. Ein

CATI erreicht jüngere und mobile Personen besser

⑥ CATI Feldtest Herbst 1997: Strukturvergleiche
ungewichtet (Index gewichtet = 100)

	Soll ¹⁾		Methodentest 1995 (CATI)		CATI Feldtest Herbst 1997		MA 98 Radio 1. Welle Bayern	
	%		%	Index	%	Index	%	Index
Geschlecht								
Männer	48		47	98	46	97	44	93
Frauen	52		53	102	54	102	56	107
Alter in Jahren								
14-19	7		6	100	8	108	6	92
20-29	16		21	124	17	109	15	92
30-39	19		24	141	26	137	19	102
40-49	15		17	113	16	108	16	104
50-59	17		14	82	14	83	16	97
60-69	13		11	85	11	87	15	119
70 und älter	13		8	57	8	58	12	92
Bildung								
Hauptschule	56		39	98	44	91	55	99
Weiterf. Schule	29		31	97	28	109	28	97
Abitur	7		15	115	14	109	7	106
Studium	9		15	107	13	111	10	111
Berufstätig	55		59	109	60	109	54	98

1) Sollvorgaben für MA 98 Radio, Bayern.

Vergleich der ungewichteten Stichproben zeigt teilweise signifikante Unterschiede (vgl. Tabelle 6). Tendenziell liegt die Face-to-face-Stichprobe näher am gewichteten Bevölkerungsdurchschnitt. Bei der CATI-Stichprobe ist wiederum ein klarer Trend zu jüngeren, mobilen Personen erkennbar. Dies bestätigt die Vermutung, daß CATI Bevölkerungsteile, die bei Face-to-face schwerer zu erreichen sind, besser ausschöpft. Die CATI-Stichprobe erreicht nicht nur überproportional Personen unter 50 Jahren, insbesondere 30- bis 39jährige, auch der Anteil von Personen mit höherem Bildungsniveau ist dadurch überdurchschnittlich. Diese Bildungsstruktur findet in der Berufsstellung mit einem geringeren Anteil von Blue-Collar-Berufen ihre Entsprechung. Auch nach erfolgter Gewichtung heben sich diese Strukturunterschiede nicht gänzlich auf. Der höhere soziale Status der CATI-Befragten bleibt bestehen.

Diese Ergebnisse lassen den Schluß zu, daß die Wahl der Befragungsmethode die Zusammensetzung der Stichprobe beeinflusst. Aus heutigem Kenntnisstand bleibt dabei offen, welche Strukturen „richtiger“ sind, das heißt die „wahre“ Grundgesamtheit repräsentieren. Von Vertretern der Werbewirtschaft in der AG.MA wurde in diesem Zusammenhang die Auffassung vertreten, daß auch bei Face-to-face-Stichproben ein Teil der Bevölkerung nicht erreicht werde, so daß diesbezügliche Vergleiche zwischen persönlich-mündlicher Befragung und CATI nicht hilfreich seien. Betrachte man die Anforderungen für die Erhebung der Radionutzung, könne CATI als akzeptable Methode angesehen werden.

Face-to-face-Interview erfolgt eine regional gesplittete Vorlage von Radio-Titelkarten, während im CATI-Interview die für den Befragungsort relevanten Sender automatisch eingespielt und vom Interviewer vorgelesen werden. Einstrahlende Sender werden im Face-to-face-Interview mit Listenvorlage und mit einer zusätzlichen offenen Frage erhoben, im CATI-Interview wird nur offen nach der Nutzung sonstiger Sender gefragt.

Die weicheren Nutzungsindikatoren für Bekanntheit und Weitesten Hörerkreis (WHK) liegen bei beiden Instrumenten auf gleichem Niveau, allerdings ist die Zahl der in den letzten zwei Wochen (WHK) durchschnittlich gehörten Programme mit 3,7 im CATI-Test deutlich höher als in der MA. Dagegen werden bei CATI an einem durchschnittlichen Tag mit 79,3 Prozent weniger Personen als Radiohörer ermittelt als bei der Face-to-face-MA. Diese Personen hören jedoch mit durchschnittlich 258 Minuten länger Radio und nutzen auch mehr Sender. Die Angaben erfolgten also differenzierter, und auch die Varianz der Angaben ist höher (vgl. Tabelle 7). Dies gilt sowohl bei den Senderangaben als auch bei den Tätigkeiten.

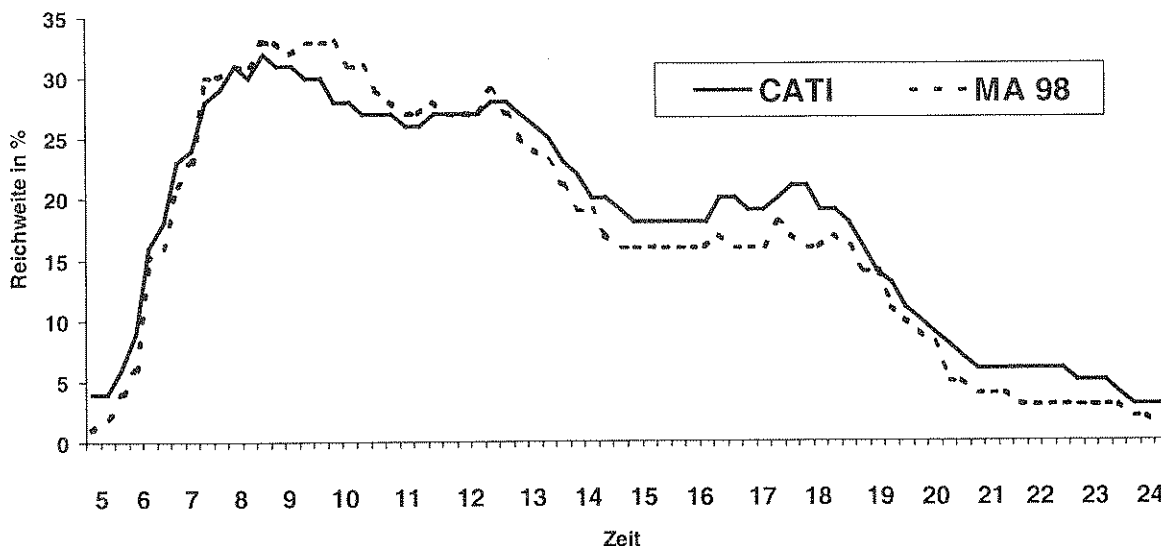
Eine wesentliche Erkenntnis des Methodentests von 1995 wurde somit bestätigt: Die Angaben der im CATI-Interview erfaßten Hörer sind insgesamt reichhaltiger. Sie nutzen nicht nur mehr Sender, sondern sie hören auch länger Radio, sowohl insgesamt über den ganzen Tag verteilt als auch bei jedem einzelnen Nutzungsvorgang. Im CATI-Interview wird das außerhäusliche Radiohören besser abgebildet. Insgesamt entspricht die mit CATI ermittelte Radionutzung in ihrem Tagesverlauf weit-

Außerhäusliche Radionutzung mit CATI besser abgebildet

Radionutzung fällt bei CATI differenzierter aus

Im CATI-Interview wie im persönlich-mündlichen MA-Interview liegen für die Ermittlung der Radionutzung vergleichbare Fragetexte zugrunde. Im

Abb. 2 CATI-Feldtest Herbst 1997
Viertelstundenreichweiten / Radio gesamt
Mo-So, Pers. ab 14 J.



⑦ **CATI Feldtest Herbst 1997**

Radio gesamt, Mo-So, Pers. ab 14 J., Bayern

	CATI Feldtest	MA 98 Radio 1. Welle	Index (MA=100)
Weitester Hörerkreis (WHK) in %	94,7	94,3	100
bayerische Sender	93,5	92,5	101
einstrahlende Sender	27,9	18,5	151
Ø Anzahl gehörter Sender im WHK	3,7	2,6	142
Hörer gestern in %	79,3	89,2	89
bayerische Sender	76,1	85,5	89
einstrahlende Sender	9,4	6,2	152
Ø Anzahl „gestern gehörte“ Sender	1,3	1,2	108
Hördauer in Min.	205	193	106
Verweildauer in Min.	258	216	119

gehend der vergleichbaren MA-Erhebung (vgl. Abbildung 2). Die dabei erkennbare höhere Radionutzung am Nachmittag entspricht den aus den Tests von 1995 gewonnenen Erkenntnissen.

Tageszeitungen sollen nicht mit CATI-MA erfaßt werden

Die Erfassung der Tageszeitungsnutzung erfolgte in der MA bisher sowohl in der Pressemedientranche (PT) als auch in der Elektronikbranche (ET). Es war deshalb zu prüfen, ob mit veränderter Erhebungstechnik auch die Nutzung der Tageszeitungen valide erhoben werden kann. Als Ergebnis kann festgehalten werden, daß CATI bei allen Zeitungsgattungen eine höhere Nutzung im weitesten Leserkreis ermittelt hat. Dieser Effekt ist ver-

gleichsweise gering bei überregionalen Kaufzeitungen, aber um so stärker bei den regionalen Abonnementzeitungen. Insgesamt sind stärkere Überschneidungen zwischen den einzelnen Titeln erkennbar, bei einem insgesamt höheren Anteil seltener Leser. Bezogen auf den Indikator „Leser pro Nummer“ (LpN) sind niedrigere Reichweiten der regionalen Abonnementzeitungen und der überregionalen Kaufzeitungen erkennbar. Dem stehen deutlich höhere LpN-Reichweiten der regionalen Kaufzeitungen und der überregionalen Abonnementzeitungen gegenüber. Im CATI-Interview ist der Anteil der Personen, die gestern keine Tageszeitung gelesen haben, signifikant höher als in der MA-Erhebung, gleichzeitig gibt es auch im LpN eine stärkere Überschneidung der Nutzung der einzelnen Titel.

Die Verlagsvertreter haben zwischenzeitlich entschieden, die Tageszeitungsnutzung nicht in einer künftigen CATI-MA erheben zu lassen, sondern andere Alternativen als Ersatz für die in der Elektronikbranche erhobenen Daten zu prüfen. Ein überzeugendes Argument ist die methodische Anforderung, die Nutzung von gedruckten Medien durch optische Erinnerungshilfen (Titelkarten) zu stützen, was im Rahmen eines Telefoninterviews technisch nicht möglich ist.

Wie fast jedes Experiment, so hat auch dieser Feldtest in vielen offenen Punkten Klarheit geschaffen, und damit die Entwicklung eines Standardinstruments für eine MA Radio per CATI entscheidend vorangebracht, gleichwohl blieben einige Fragen offen. Im Mittelpunkt des Interesses stand der Versuch, für die vorliegenden Ergebnisunterschiede Erklärungen zu finden. Systembedingte Einflußfaktoren liegen in der Stichprobentechnik, der speziellen CATI-Technik und der Umsetzung des Fragebogens in ein CATI-gerechtes Fragenmodell. Dies soll nachfolgend im einzelnen erläutert und Lösungsansätze aufgezeigt werden.

Feldtest bringt Entwicklung einer CATI-MA voran

Lösungsansätze für Probleme der CATI-Technik

Fehlen optischer Erinnerungshilfen erfordert Präzision bei Frageitems

Im CATI-Interview ist die Vorlage von optischen Erinnerungshilfen grundsätzlich nicht möglich. Erinnerungshilfen haben eine besondere Relevanz für die Nutzungsabfrage von gedruckten Medien, sind aber auch hilfreich bei anderen Fragestellungen, zum Beispiel der Ermittlung von Besitzgegenständen und Kaufabsichten. Die übliche Listenvorlage eines Face-to-face-Interviews wird im CATI-Interview ersetzt durch das Vorlesen von Einzelitems, mit der Folge, daß der Befragte sich zu jedem Item sofort äußern muß, ohne die Folgeitems zu kennen. Die Schlußfolgerung war, daß größte Aufmerksamkeit und Präzision auf die Bezeichnung der Einzelitems (z.B. was ist eine Kühl-/Gefrierkombination?) zu legen ist, wie auch in deren Abfolge im Gesamtzusammenhang. Dieses Problem zeigte sich auch bei der Abfrage der Statistik, insbesondere bei der Zuordnung zu Bildung und Berufstätigkeit. Insgesamt eine Problematik, die relativ einfach gelöst werden konnte. Die abweichende CATI-Befragtenstruktur ist also teilweise auch auf eine nicht eindeutige Abfrage im Test zurückzuführen. Als Konsequenz wurde u.a. in der MA aus dem bisherigen „Haushaltsvorstand“ der „Haupteinkommensbezieher“.

Keine signifikanten Unterschiede bei Prestigeantworten

Darüber hinaus wurde die Hypothese aufgestellt, die CATI-Befragten würden am Telefon die Unwahrheit sagen, durch die Anonymität des Telefoninterviews also dazu neigen, sich sozial höher einzustufen. Zu diesem Zweck wurde ein Validierungsexperiment durchgeführt. Im Raum München wurden 193 verkürzte CATI-Interviews mit zusätzlichen Fragen zu Besitz und Einstellungen durchgeführt. Die Angaben zu Besitz und Wohnverhältnissen sollten anschließend durch ein Interview im Haushalt verifiziert werden. Bei 100 von 193 CATI-Interviewten wurden Face-to-face-Nachbefragungen im Haushalt realisiert. Im Ergebnis zeigten sich keine signifikanten Unterschiede, die Hypothese konnte also nicht bestätigt werden.

Tagesablaufschemata mit CATI offenbar gut zu handhaben

Ein zentrales Ergebnis des Feldtests 1997 lautet: An einem durchschnittlichen Tag hören weniger Personen Radio als in der MA, sie hören aber mehr Sender, haben eine längere Verweildauer, und die Varianz der Angaben im Tagesablaufschemata ist höher. Als mögliche Ursache für dieses Ergebnisseniveau wurde die PC-Datenmaske und die Handhabung des Tagesablaufschemas genannt. Ein Validierungstest sollte diese Hypothese prüfen. Dazu wurde ein verkürztes CATI-Interview in zwei Testvarianten mit je 35 Befragten durchgeführt. Neben dem im Feldtest 1997 eingesetzten Tagesablauf mit Bildschirmmaske wurde in einer zweiten Variante der gedruckte Tagesablauf aus dem MA-Interview (Paper-pencil) als Vorlage zum Eintragen der abgefragten Antworten eingesetzt. Der Interviewer hatte also die Möglichkeit, im Gespräch mit dem Befragten die Tagesablaufangaben in eine große und übersichtliche Vorlage einzutragen, ohne die technischen Restriktionen eines PC-Bildschirms bewältigen zu müssen. Ein Einfluß des Befragungsinstrumentes auf die Tagesablaufangaben war

nicht feststellbar. Die Interviewer empfanden die Handhabung der Bildschirmmaske sogar als angenehmer und zuverlässiger.

Ein Anliegen des Feldtests 1997 war die Überprüfung der Ausschöpfungsmöglichkeiten einer CATI-Stichprobe. Aus dieser Sicht entschied man sich für ein Auswahlverfahren, dem das Telefonverzeichnis zugrunde liegt. Das angewandte Verfahren sah für jedes Institut eine repräsentative Stichprobe vor. Die Stichproben wurden einheitlich gezogen, um Firmenadressen bereinigt. Den Instituten wurden die Telefonnummern vorgegeben. Zu jeder Nummer wurde ein Protokoll erstellt. Damit konnte eine exakte Ausschöpfungsquote bei voller Transparenz der Durchführung ermittelt werden.

Gegenüber der Face-to-face-Stichprobe bietet das CATI-Verfahren die Möglichkeit der exakten Aussteuerung der Interviews nach Wochentagen und Tageszeit, auch in Abhängigkeit von der Antreffbarkeit des Befragten. Im Unterschied zur Face-to-face-Stichprobe sind die Interviews nicht örtlich geklumpt. Demgegenüber bedingt ein Telefoninterview die räumliche Distanz zum Befragten, es setzt voraus, daß die Befragungsperson mit dem Telefon umgehen und einem abstrakten Interview intellektuell folgen kann. Andererseits sind non-verbale Einflüsse des Interviewers ausgeschlossen. Telefoninterviewer sind im Durchschnitt jünger als Face-to-face-Interviewer. Es gibt also eine Reihe technisch bedingter Unterschiede zwischen den einzelnen Erhebungstechniken, die Einfluß auf die Befragtenstruktur haben können. (13)

Ein grundsätzlicher Nachteil der CATI-Stichprobe des Feldtests 97 war die unvollständige Abdeckung der Grundgesamtheit durch die Beschränkung auf die im Telefonverzeichnis eingetragenen Telefonnummern. Dieses Stichprobenverfahren war für die Aufgabenstellung des Tests angemessen, aber nicht auf eine CATI-MA übertragbar. Das gilt übrigens auch für das von den Marktforschungsinstituten üblicherweise eingesetzte Random-Last-Digit-Verfahren bei Telefonbefragungen. Hier werden von vorhandenen Telefonnummern die letzte bzw. die beiden letzten Ziffern abgeschnitten und im Zufallsverfahren neu generiert. Der große methodische Nachteil dieses Verfahrens ist, daß man kaum zwischen real existierenden und künstlich generierten Nummern unterscheiden kann und somit die Ziehung einer Bruttostichprobe unmöglich ist. Eine repräsentative Telefonumfrage auf dieser Basis ist somit nur eingeschränkt möglich. Auch eine Telefonstichprobe muß eine methodische Grundvoraussetzung erfüllen, nämlich jeder Person die gleiche Auswahlchance gewähren.

Unabhängig von den Überlegungen der AG.MA wurde vom Zentrum für Umfragen, Methoden und Analysen (ZUMA) der Universität Mannheim ein neues Design für Telefonstichproben entwickelt. (14) Dieses Modell wurde unter Mitwirkung der AG.MA vom Arbeitskreis Deutscher Marktfor-

Stichprobentechnik für CATI optimieren

Nachteil CATI: Bei Rückgriff auf Telefonverzeichnis wird Grundgesamtheit unvollständig abgedeckt

Neues Stichproben-design für Telefonbefragungen

schungsinstitute (ADM) als Grundlage für die Erstellung einer nationalen Telefonstichprobendatei ausgewählt. Die Fertigstellung dieser Stichprobendatei erfolgte Mitte 1999, so daß alle Vorbereitungen für eine Radio MA per CATI ab Herbst 1999 getroffen werden konnten.

Da das Stichprobendesign zum ersten Mal angewendet wird, folgt hier eine kurze Beschreibung der Vorgehensweise. Eine ausführliche Dokumentation der ADM-Telefonstichprobendatei wird in Kürze vom ADM veröffentlicht. Die Stichprobenziehung erfolgt mit den Vorgaben und unter der Aufsicht der AG.MA.

Auswahlverfahren zur Erfassung aller vorhandenen Telefonnummern

Das Grundproblem ist, daß es keinen uneingeschränkten Zugriff auf eine vollständige Telefonnummerndatei der Bundesrepublik Deutschland gibt. Es gibt übrigens auch keinen Zugriff auf Melderegister, was für die Erstellung einer Flächenstichprobe für Face-to-face-Studien ebenso nützlich wäre. Nur ein Teil (ca. 60 bis 70%) der Telefonteilnehmer lassen ihren Anschluß veröffentlichen. Ziel des ADM war es nun, eine Auswahlgrundlage zu schaffen, die es ermöglicht, eine Zufallsauswahl unter allen vorhandenen Telefonnummern zu treffen.

Als Basis für die Auswahlgrundlage der ADM-Telefonstichprobendatei dient eine aktuelle offizielle Nummerndatei, die alle veröffentlichten Telefonnummern enthält. Für die Erstellung der Auswahlgrundlage wird nach dem vorgenannten Gabler/Häder-Verfahren der Rufnummernbestand in sogenannte „Hunderterblöcke“ aufgeteilt, das heißt, bei allen eingetragenen Rufnummern werden die letzten zwei Ziffern abgeschnitten und mit den Ziffern 00 bis 99 aufgefüllt. Allen Telefonnummern wird eine Gemeindekennziffer nach Verteilung der eingetragenen Nummern zugeordnet. Blöcke mit mehreren Gemeindezuordnungen werden geteilt, so daß für jede Gemeinde eine Anzahl von Blöcken entsteht, die eindeutig dieser Gemeinde zugeordnet sind. Außerdem wird für jeden Nummernblock die Eintragsdichte errechnet (Anteil der eingetragenen Telefonnummern).

Nach der Entfernung der gewerblichen Anschlüsse ergibt sich eine Grundgesamtheit von ca. 550 000 Nummernblöcken mit ca. 53 Millionen Telefonnummern (30 Millionen eingetragene und 23 Millionen generierte Telefonnummern). Diese Datei enthält nahezu 100 Prozent aller in Deutschland vorhandenen privaten Telefonanschlüsse. Für die MA 2000 wurde nun auf Basis dieser Auswahlgrundlage ein Design für die Ziehung und Abarbeitung der Stichprobe von der AG.MA entwickelt.

Stichprobenziehung mit systematischer Zufallsauswahl

Ziel der Stichprobenziehung ist es, jeder vorhandenen Telefonnummer die gleiche Auswahlchance entsprechend der regionalen Verteilung der Grundgesamtheit zu geben. Zur Ziehung der Stichprobe werden die Nummernblöcke hierarchisch nach Regionen geschichtet und innerhalb der Schichtungen nach Gemeinden, Stadtteilen und Bedeu-

tungsgewicht der Blöcke (Eintragsdichte) angeordnet. Die Ziehung erfolgt per computergesteuerter mehrstufiger systematischer Zufallsauswahl.

1. Stufe: Ziehung des Blocks
2. Stufe: Zufallsauswahl im Block. Pro Block wird eine Nummer bearbeitet.
3. Stufe: Zufallsauswahl im Haushalt nach Auflistung der Personen im Haushalt.

Grundsätzlich wurde darauf geachtet, die methodischen Vorgaben der Face-to-face-Stichprobenziehung soweit wie möglich auf CATI zu übertragen.

Für die MA 2000 sollen 54 000 Interviews mit sechs Instituten durchgeführt werden. Die Feldlaufzeit verteilt sich auf zwei Wellen à acht Ersteinsatz- und zwei Nachwochen. Unter Berücksichtigung eines großen Anteils von nicht-existenten Nummern unter den generierten Nummern und einer angestrebten Ausschöpfung von über 70 Prozent wurde ein doppelter Ansatz der Bruttostichprobe vereinbart. Das heißt, es werden 108 000 Blöcke gezogen.

Die gezogenen Blöcke werden so angeordnet, daß sie einer regionalen und zeitlichen Gleichverteilung entsprechen. Bei sukzessivem Einladen und Abarbeiten der Stichprobe ist somit gewährleistet, daß zu jedem Zeitpunkt eine optimale Verteilung vorliegt. Ein täglicher Soll/Ist-Abgleich ermöglicht eine gleichmäßige Verteilung hinsichtlich Regionalität, Wochentage und Feldlaufzeit.

Die Interviews werden in der Regel in der Zeit von 17.00 Uhr bis 21.00 Uhr durchgeführt. Zu diesem Zeitpunkt ist die Mehrheit der Grundgesamtheit zu Hause anzutreffen. Ein zusätzlicher Test der ARD ergab, daß eine Ausweitung der Befragungszeit keine wesentlichen Vorteile für die Stichprobenstruktur erbringt und wegen zu geringer Ausschöpfung am Vor- und Nachmittag nicht wirtschaftlich ist. Nach erfolglosen Versuchen am Abend kann der Einsatz tagsüber jedoch die Ausschöpfung erhöhen.

Ein computergesteuertes Sample Management System steuert die Anzahl der Kontaktversuche unter Berücksichtigung von Wiedervorlageregeln. Bei Nichterreichen werden mindestens 15 Kontaktversuche pro Telefonnummer an zehn verschiedenen Tagen unternommen. Nach zehn Kontaktversuchen wird die Telefonnummer auch tagsüber ausgewählt. Pro Sample Point darf nur eine Telefonnummer bearbeitet werden. Ist die erste Telefonnummer ein neutraler Ausfall, dann und nur dann darf die zweite Telefonnummer ausgewählt werden, u.s.w. Auch damit wird die methodische Kompatibilität zur bisherigen Stichprobenanlage sichergestellt, wobei der Vorteil der direkten Kontrolle im Institut noch qualitätsverbessernd hinzukommt.

Fazit

Mit dem Start der Erhebung der MA 2000 Radio per CATI ab Anfang Oktober 1999 gehen fünf Jahre experimenteller Forschung erfolgreich zu Ende. Ausgelöst durch die Forderung der Werbe-

MA 2000: Für 54 000 Interviews werden 108 000 Blöcke gezogen

Abarbeitung der Stichprobe

Mindestens 15 Kontaktversuche an zehn Tagen

MA 2000 Radio ab Oktober 1999 per CATI

wirtschaft nach bestmöglichen medienindividuellen und mediengerechten Erhebungsinstrumenten wurde von der AG.MA ein Industriestandard für repräsentative Telefonbefragungen entwickelt, der eine neue Qualität in der Marktforschung bedeuten wird.

Die Zielsetzung, die 1994 mit der Konzeption der ersten Methodentests verbunden war, wurde erreicht. Mit der gelungenen Übertragung der MA-Methode für die repräsentative Ermittlung von Mediennutzung auf die CATI-Technik, wurde für Radio ein eigenes Erhebungsinstrumentarium geschaffen, das den Besonderheiten der Nutzung dieses Mediums gerecht wird. Die CATI-Technik scheint nach allen Erfahrungen für die Abbildung der Radionutzung medienadäquater zu sein als andere Erhebungstechniken, vor allem da sie mobile Zielgruppen besser erreicht und damit auch die Mobilität der Nutzung zuverlässiger abbilden kann. Die reichhaltigeren, stärker differenzierenden Nutzungsangaben, die insgesamt höhere Varianz bei der ermittelten Radionutzung sind ein überzeugendes Argument.

Kompatibilität zu senderinternen Trenduntersuchungen

Hinzu kommt die Kompatibilität zu den senderinternen Trenduntersuchungen und die dadurch ermöglichten Synergieeffekte. Erstmals können mit CATI die Reichweitenerhebung und die ergänzende Programmbegleitforschung wieder methodisch vergleichbar durchgeführt werden, was zudem die inhaltliche Optimierung der Senderuntersuchungen erleichtert. Vor allem aber kann damit eine einheitliche Währung für die Programmplanung einerseits und die Werbeplanung andererseits hergestellt werden.

Kostenvorteile der CATI-Technik

Weitere Vorteile leiten sich direkt aus der CATI-Technik ab. Das sind einmal die Kosten. MA-CATI-Interviews sind – nach derzeitiger Kalkulation – um gut ein Viertel preiswerter als Face-to-face-Interviews. Die bei Hörfunk derzeit gegebene Forschungsquote von 0,6 Prozent – also der Anteil der Aufwendungen zur Reichweitenermittlung an den Nettowerbeerlösen – liegt bei anderen Mediengattungen zum Teil deutlich niedriger und wird im Zuge der Umstellung auf die CATI-Technik auch beim Hörfunk sinken. Allerdings ist der finanzielle Spielraum zur Weiterentwicklung der MA dringend erforderlich. Die von der Werbewirtschaft geforderte Basisdatenerhebung, in der Grobstrukturen von Verbrauch und Konsum erhoben werden, wird von den Medien finanziert. Eine Kostenbeteiligung von Radio war nur bei Reduzierung der bisherigen Kosten denkbar. Darüber hinaus erfordert die MA Radio ungeachtet der Erhebungstechnik hohe Fallzahlen, um möglichst alle Sender hinreichend genau erfassen und darstellen zu können. Das Beispiel Berlin/Brandenburg, wo für die MA 2000 Radio fast die dreifache Sollfallzahl erhoben wird, zeigt den Bedarf. Auch die diskutierten Optionen wie Anzahl der jährlichen Berichterstattungen, Reichweiten nach Wochentagen o.ä. erfordern zur Lösung auch einen finanziellen Spielraum.

Andere Vorteile von CATI leiten sich direkt aus der Erhebungstechnik ab. Dazu zählen die Mög-

lichkeit der größeren Streuung der Stichprobe, die maximale Kontrolle der Interviews und die geringen Interviewereinflüsse, und die zeitlichen Vorteile bei der Aktualisierung des Fragenprogramms ebenso wie bei der Datenerfassung und -aufbereitung. CATI wird also künftig auch die Aktualität der Radiodaten verbessern helfen.

Damit wurde im Ergebnis mehr als nur eine Kundenanforderung erfüllt. Eine CATI-MA scheint dem geforderten optimalen, mediengerechten Erhebungsinstrument zu entsprechen und weitere Optionen offenzuhalten. Nachdem Fernsehen bereits 1985 ein eigenes – außerhalb der AG.MA organisiertes – Forschungsinstrument (15) etabliert hat, folgte 1986 im Rahmen des sogenannten AG.MA-Partnerschaftsmodells erstmals die getrennte Erhebung von gedruckten und elektronischen Medien. (16) Mit dem Umstieg auf CATI wird nun auch Radio erstmals über ein eigenständiges Erhebungsinstrumentarium verfügen. Vielleicht war es auch diese Perspektive, die während der Entwicklungsphase zu kritischen Stimmen geführt hat, die insgesamt jedoch zur Optimierung des Instruments beigetragen haben.

Gleichwohl sind Restrisiken für die MA Radio nicht ausgeschlossen. Die neu entwickelte Telefonstichprobentechnik wurde vorher noch nie großflächig getestet und kommt mit der MA 2000 Radio erstmals zum Einsatz. Es wird eine besondere Aufgabe sein, während der laufenden MA methodische Prüfungen und gegebenenfalls Korrekturen und Optimierungen vorzunehmen, was die CATI-Technik auch jederzeit erlaubt. Deshalb ist auch kaum daran zu zweifeln, daß die praktischen Erfahrungen die theoretischen Erkenntnisse bestätigen werden.

Bei allen Neuerungen bleiben jedoch die methodischen Kernelemente der MA-Befragung unverändert. Die Stichprobenanlage erfüllt die höchsten Anforderungen an Zufallsauswahl und Repräsentativität. Die Ermittlung der Radionutzung basiert auch weiterhin auf dem Modell der Stichtagsbefragung, die schon 1949 von der „Hörerforschung Ernst“ im Auftrag des Bayerischen Rundfunks entwickelt wurde (17) und bis heute als Standardinstrument dient. Die Methode bleibt also unverändert, es ändert sich lediglich die Erhebungstechnik. Aber dies ist schon eine ganze Menge.

Anmerkungen:

- 1) Eine ausführliche Dokumentation der Forschungsexperimente wird demnächst in der AG.MA-Schriftenreihe vorgelegt.
- 2) Eine Face-to-face-Befragung wird normalerweise anhand eines vorstrukturierten Fragebogens (paper) durchgeführt, in dem der Interviewer mit einem Bleistift (pencil) die Antworten markiert bzw. einträgt. Man nennt das deshalb auch ein „Paperpencil-Interview“. Inzwischen wird immer häufiger der gedruckte Fragebogen durch einen speziellen tragbaren Computer ersetzt. Dabei hat die Befragungsperson die Möglichkeit, die Antworten selbstständig mit einem elektronischen Stift (pen) auf dem Display des Laptops (= Pentop) zu markieren. Pentop wird auch bezeichnet als CAPI-Interview (Computer Assisted Personal Interview).

CATI-MA scheint mediengerechtes Erhebungsinstrument für Radio zu sein, weitere Optimierungen durch Praxiserfahrung

Methodische Kernelemente der MA bleiben unverändert

- 5) Vgl. Buß, Michael: Leistungsfähigkeit und Grenzen der Media Analyse. In: Linder-Braun, Christa (Hrsg.): Radioforschung. Konzepte, Instrumente und Ergebnisse aus der Praxis. Opladen/Wiesbaden 1998, S. 77-82; ders.: Das System des SDR-Demometer: Der Trend zur monatlichen Repräsentativbefragung. In: ebd., S. 11-23.
- 4) Vgl. Menneer, Peter: Radio Audience Measurement Systems Across Europe. In: ESOMAR Publication: First ESOMAR Radio Research Symposium, Paris 1995. Amsterdam 1995.
- 5) Vgl. Müller, Dieter K.: Das AG.MA-Partnerschaftsmodell wird neu definiert. Eröffnen sich durch den Wiedereintritt der Fernsehsender in die Arbeitsgemeinschaft Media Analyse neue Perspektiven intermedialer Paritäten? In: Media Perspektiven 6/1997, S. 320-329.
- 6) Vgl. Gane, Roger: Radio Audience Measurement in the Future. In: ESOMAR Publication: Second Radio Research Symposium, Warschau 1997. Amsterdam 1997.
- 7) Vgl. Müller, Dieter K.: Radiometer als optionales Instrument der Hörschaftsforschung. Bringen passive elektronische Meßinstrumente Fortschritt oder Verwirrung? In: Media Perspektiven 2/1998, S. 70-75.
- 8) Aktivitäten: im Haus: Körperpflege/Anziehen, Essen, Hausarbeit, Berufsarbeit; außer Haus: unterwegs im Auto, unterwegs Bahn/Bus, Einkaufen/Besorgen, Berufsarbeit, Schule, Studium, Besuche machen, Kneipe/Gaststätte/Restaurant. Hinzu kommt als Tätigkeit „Schlafen“, wobei nicht nach „im Haus“ bzw. „außer Haus“ unterschieden wird. Komplementärmedien: Schallplatten, Tonband, Kassetten, CD hören, Fernsehen, Video sehen, mit PC beschäftigen.
- 9) Vgl. Franz, Gerhard/Jürgen Piffnerling: Hörerdaten auf dem Prüfstand - Die ARW-Methodentests 1988. In: Media Perspektiven 10/1989, S. 655-668.
- 10) Schwedenschlüssel: Eine in der Umfrageforschung verwendete Reihe von Zufallsziffern, mit deren Hilfe es interviewen möglich ist, am Ort und der Stelle des Interviews bei der Auswahl der Befragungsperson unter den Haushaltsangehörigen eine Zufallsauswahl zu treffen; vgl. Koschnick, Wolfgang J.: Standard-Lexikon für Mediaplanung und Mediaforschung in Deutschland. Zweite überarbeitete und erweiterte Auflage. München u.a. 1995, Band 2, S. 1-570.
- 11) Die Bruttostichprobe ist die noch nicht um die qualitätsneutralen und nichtneutralen Ausfälle bereinigte Ausgangsstichprobe. Im Gegensatz zum nichtneutralen Ausfall, etwa wenn der Adressat nicht anzutreffen ist oder ein Interview verweigert, hat ein sogenannter qualitätsneutraler Ausfall keinen verzerrenden Einfluß auf die Repräsentanz der Stichprobe und stellt somit keinen systematischen Fehler dar. Neutrale Ausfälle liegen dann vor, wenn der Adressat z.B. verzogen, verstorben, unbekannt oder krank ist, oder aber wenn er sich als nicht zur Grundgesamtheit der Befragten gehörig herausstellt, im Falle der MA beispielsweise eine Person ausländischer Staatsangehörigkeit; vgl. Koschnick (Anm. 10), Band 1, S. 196ff.
- 12) Vgl. Zentralverband der deutschen Werbewirtschaft (Hrsg.): ZAW-Rahmenschema für Werbeträger-Analysen. Bonn 1994.
- 13) Vgl. zu diesem Thema eine Analyse des INRA-Instituts: Mosmann, Helmut: Spurenlese - Das Antwortverhalten RAPI-CAPI-CATI. In: Broschüre von INRA Deutschland. Mölln o.J.
- 14) Vgl. Gabler, Siegfried/Sabine Häder: Überlegungen zu einem Stichprobendesign für Telefonumfragen in Deutschland. In: ZUMA-Nachrichten 41, Jg. 21, November 1997.
- 15) Vgl. Buß, Michael: Das System der GfK-Fernsehforschung: Entwicklung und Nutzen der Forschungsmethode. In: Klingler, Walter/Gunnar Roters/Oliver Zöllner (Hrsg.): Fernsehforschung in Deutschland. Themen - Akteure - Methoden. Baden-Baden 1998, S. 787-813.
- 16) Vgl. Ernst, Otmar: (Medien-)Perspektiven einer Partnerschaft. Anmerkungen zum Status der Arbeitsgemeinschaft Media-Analyse e.V. (AG.MA). In: Media Perspektiven 3/1989, S. 146-156.
- 17) Vgl. Bessler, Hansjörg: Hörer- und Zuschauerforschung. In: Bausch, Hans (Hrsg.): Rundfunk in Deutschland. Band 5. München 1980.

