



Forschungsberichte

Band 17

Worldwide Readership
Research Symposium 10

Venice 2001

Deutsche Übersetzung von
ausgewählten Session Papers

(Übersetzung und Redaktion:
Ute Löffler Marktforschungskontor, Wiesbaden)

Vakat

Vorbemerkung 7

Session Nr./

Original-

Seite

Autor

Titel

Seite

2. THE PLANNING PROCESS

2.1 17	Erwin Ephron Ephron, Papazian & Ephron, Inc.	COUNTING CALORIES - ON THE NEED TO ADJUST ISSUE READERSHIP DATA Kalorien zählen - Die Notwendigkeit Reichweiten für Ausgaben zu justieren	9
2.2 21	Tony Jarvis, MediaCom/Grey Global Group	YOU GAVE US THE 'ANSWERS', NOW ASK US THE QUESTIONS! – HOW TO MAXIMIZE MAGAZINE ADVERTISING EFFECTIVENESS IN THE PLANNING/BUYING PROCESS Ihr gabt uns die 'Antworten', nun fragt uns die Fragen! - Wie die Werbewirkung von Anzeigen in Zeitschriften imPlanungs-/Einkaufsprozess gesteigert werden kann	13
2.3 29	Stuart Gray, BBDO Julian Baim & Martin Frankel, MRI-NOP World	AN EXAMINATION OF DIFFERENTIAL WEIGHTING OF SUCCESSIVE WAVES OF PRODUCT USAGE ESTIMATES Eine Prüfung von verschiedenen Gewichtungen aufeinander- folgender Wellen von Verbrauchsdaten	19
2.4 35	Tony Jarvis, MediaCom/Grey Global Group Stephen A. Douglas, The Douglas/ Jones Group	TO PROTOTYPE OR NOT TO PROTOTYPE, THAT IS THE QUESTION – CAN USA GUIDELINES BE EMBRACED WORLDWIDE? Vorgabe von Standards oder keine Vorgabe von Standards, das ist die Frage - Können Richtlinien der USA die ganze Welt einschließen?	21

3. QUALITY AND IMPACT OF FUSION

3.6 123	Tony Jarvis, MediaCom/Grey Global Group Leslie Wood, Leslie Wood Research	FUSION – CAN MAGAZINES BE EQUAL PLAYERS IN A FUSED DATABASE? Fusion - Können Zeitschriften gleichwertige Mitspieler in einem fusionierten Datenbestand sein?	27
------------	--	---	----

4. RECENT READING ISSUES

4.4 167	Julian Baim, Martin R. Frankel, Joseph Agresti und Risa Becker, MRI-NOP World	MAGAZINE AUDIENCE ACCUMULATION: DISCUSSION OF RESEARCH ISSUES, MODELING AND APPLICATION Zeitschriften-Leserschafts-Kumulationen: Diskussion über die Forschung, das Modelling und ihre Anwendung	33
------------	---	---	----

Session Nr./ Original- Seite	Autor	Titel	Seite
4.5 185	Gary Morgan, Michele Levine & Sergey Dorofeev, Roy Morgan Research	A NEW METHOD TO MEASURE MEDIA CASUALNESS FOR MAGAZINES AND NEWSPAPERS Eine neue Methode um gelegentliches Lesen für Zeitschriften und Tageszeitungen zu erheben	39
4.6 205	Roberto Lobl & Flavio Ferrari IBOPE Pesquisa de Midia	READERSHIP PANEL: WE DID IT! HISTORY AND RESULTS OF A HUGE CHALLENGE Leserschafts-Panel: Wir führten es durch! Entwicklung und Ergebnisse einer großen Aufgabe	43
5. DATA COLLECTION - WHAT'S NEW			
5.1 215	Johannes Schneller, Institut für Demosko- pie Allensbach	THE EFFECTS OF DIFFERENT MASTHEAD CARD DESIGNS: COLOUR VS. BLACK AND WHITE Effekte unterschiedlicher Gestaltung von Titelformen: Farbig oder Schwarzweiß	49
5.2 221	Hilary Birt, Ipsos-RSL	SELF-COMPLETION QUESTIONNAIRES: THE IMPORTANCE OF TESTING (BASED ON EXPERIENCE FROM THE BRITISH NRS) Fragebogen zum Selbstaussfüllen: Die Wichtigkeit sie zu testen (Erfahrungen von der britischen NRS - Nationale Leserschafts-Analyse)	55
5.3 233	Virginia Cable, The Wall Street Journal Valentine Appel, Consultant	TITLE CONFUSION AND PHYSICAL PROXIMITY IN A SELF ADMINISTERED QUESTIONNAIRE Titel-Verwechslungen und physische Nähe in einem Selbstaussfüller- Fragebogen	63
5.4 241	Marco Vriens, Barbara Graham, Eric Melton, Tony Incalcaterra, Millward Brown IntelliQuest Valentine Appel, Consultant	MEASURING AND MINIMIZING POSITIONAL AND TITLE CONFUSION BIASES IN PRINT MEDIA AUDIENCE ESTIMATES Erhebung und Minimierung von Positionierungs- und Titelverwechslungs- Effekten in der Leserschaftsforschung	69
5.5 255	Alke Bassler & Costa Tchaousoglou, SUMMO	SEARCHING FOR BEST CASI PRACTICES: EXPERIMENTS FOR A NEW DATA COLLECTION METHOD Auf der Suche nach der besten CASI-Praxis: Experimente für eine neue Datenerhebungs-Methode	75
5.6 265	Katherine Page, Ipsos-RSL	PERSONALISED MEDIA LISTS Personalisierte Media-Listen	85

Session Nr./ Original- Seite	Autor	Titel	Seite
6. IMPROVING FIELDWORK			
6.1 277	Hastings Withers, PMB Print Measure- ment Bureau Ivor Thompson, Thompson Lightstone (Maritz Research Inc.)	NON-RESPONDERS – AN OPPORTUNITY TO LOOK INTO THEIR HEARTS AND SOULS Verweigerer - Eine Gelegenheit, um in ihre Herzen und Seelen zu schauen	93
7. NICHE AUDIENCES			
7.4 337	Ton Jansen, R&M, Research and Marketing B.V. Vincent van den Berg, Cebuco	BREAKTHROUGH IN NEWSPAPER SECTION RESEARCH IN THE NETHERLANDS Durchbruch für Tageszeitungs-Bücher(Sektionen)-Forschung in den Niederlanden	101
7.6 359	Fred Bronner, Veldkamp Dick Suèr, WNCP	THE COMMUNICATION POWER OF CUSTOMER MAGAZINES Die Kommunikations-Kraft von Kundenzeitschriften	109
8. THE DIGITAL LANDSCAPE: THE INTERNET AS A MEDIUM AND AS A MEASUREMENT TOOL			
8.1 371	Scott C. McDonald, Condé Nast Publications	ONLINE MEDIA MEASUREMENT: AFTER THE DELUGE Online-Untersuchungen: Nach der Sintflut	
8.5 425	Richard Asquith, BMRB International	WHEN TWO WORLDS COLLIDE – OFFLINE MEASUREMENT OF ONLINE READERSHIP Wenn zwei Welten aufeinandertreffen: Offline-Messungen von Online-Leserschaften	113
8.7 445	Jay Mattlin, Knowledge Networks; Bruce Goerlich, Starcom MediaVest Group	MEASURING MAGAZINE READING VIA THE INTERNET: TESTING THE EFFECT OF NUMBER OF TITLES AND OTHER QUESTIONNAIRE DESIGN ISSUES Erhebung des Zeitschriftenlesens via Internet: Testen des Effekts der Titelanzahl und anderer Fragebogen-Design-Merkmale	125
9. THE ADVERTISING EFFECTIVENESS OF PRINT AND THE QUALITY OF THE READING EVENT			
9.1 469	Ellen Oppenheim & David Arpin, Magazine Publishers of America Robert Wyman & Gerald Pollak, Media Marketing Assessment Thomas Robinson, Precision Marketing Services	MODELING MAGAZINE EFFECTIVENESS Modelling Zeitschriften-Effektivität	133

Session Nr./ Original- Seite	Autor	Titel	Seite
9.2 475	Shinichiro Okauchi, Dentsu Inc.	SYNERGY BROUGHT ABOUT THROUGH DUPLICATION OF CONTACT WITH NEWSPAPER AND INTERNET Synergien durch Überschneidungs-Kontakte von Tageszeitungen und Internet	137
9.3 491	Cynthia Evans, The Media Edge/ Y&R / WPP	MEASURING THE QUALITATIVE VALUE OF MAGAZINES Erhebung des qualitativen Wertes von Zeitschriften	143
9.5 509	Morten Kromann-Larsen & Rolf Randrup, Taylor Nelson Sofres/ Gallup	QUALITY RATING POINT – THE NEW CURRENCY Quality Rating Point - Die neue Währung	149
10. INTEGRATING PRINT IN THE PLANNING PROCESS: INTERMEDIA			
10.1 519	Scott C. McDonald, Condé Nast Publications	USING THE 24-HOUR DAY AS THE COMMON DENOMINATOR FOR CROSS-MEDIA COMPARISONS: THE CASE OF MULTI-MEDIA MENTOR Nutzung des 24-Stunden-Tages als gemeinsamen Nenner für Cross-Media-Vergleiche: Der Multi-Media Mentor	155
10.6 567	Peter Masson, Bucknull and Masson Peter Callius, Research International	TIME RELATED MEASURES OF RELATIVE 'EFFECT' – AN ESSENTIAL INGREDIENT FOR MULTI-MEDIA CAMPAIGN EVALUATION Zeitbezogene Erhebungen des 'relativen Effekts' - Eine wesentliche Voraussetzung zur Berechnung von Media-Mix-Kampagnen	159
10.7 583	Andrew Green, OMD USA	MULTI-MEDIA OPTIMIZERS: THE NEW COUTURE? Media-Mix Optimierungsprogramme: Die neue Mode?	165
10.8 589	Bruce Goerlich, Starcom MediaVest Group	EVALUATING TV AND PRINT SALES EFFECTS Bewertung von Fernseh- und Print-Verkaufs-Effekten	169

Vorbemerkung

Mit der vorliegenden Veröffentlichung wird eine Auswahl aus insgesamt zweiundfünfzig Vorträgen, die anlässlich des 10. Readership Research Symposiums in Venedig gehalten wurden, in einer zusammenfassenden Übersetzung vorgelegt. Wie auch bei der Veröffentlichung zum vorangegangenen Readership Research Symposium 1999 (AG.MA Forschungsberichte, Band 16) versucht die Auswahl, die Themenvielfalt widerzuspiegeln.

Mit den Zusammenfassungen soll dem geneigten Leser eine schnelle Entscheidungshilfe an die Hand gegeben werden, die es ihm erlaubt, aufgrund seiner persönlichen Interessenlage zu entscheiden, welchen Vortrag er in aller Ausführlichkeit im Original lesen möchte. Keinesfalls ersetzt diese Veröffentlichung das umfangreiche Buch mit den Originalbeiträgen. Das wird auch dadurch dokumentiert, dass beispielsweise Tabellen, Graphiken oder Anhänge mit Fragebogendetails nicht in der Übersetzung enthalten sind, es sei denn, sie sind zum Verständnis notwendig gewesen.

Wir danken den Initiatoren des Readership Research Symposiums BMRB International, London und Ipsos-RSL Ltd, Harrow für die freundliche Genehmigung dieser Veröffentlichung.

Frankfurt, im Juni 2002

Jürgen Wiegand
MEDIA MICRO CENSUS GMBH

Counting calories - On the need to adjust issue readership data

Kalorien zählen - Die Notwendigkeit Reichweiten für Ausgaben zu justieren

Erwin Ephron, Ephron, Papazian & Ephron, Inc.

Wir beginnen gerade die Naturgesetze der Werbung zu verstehen. Sie könnten von den Weight-Watchers geschrieben sein. Das erste Gesetz scheint zu sein: Lass nicht die Sau raus. Bei den Verbrauchern ist eine sinkende Werbewirkung bei hohen Werbeausgaben zu beobachten. Jeder zusätzliche Kontakt bringt ein wenig weniger. Die Weight-Watchers sagen den Planern, sie mögen einige Donut für morgen aufheben. Das zweite Gesetz heißt: Nicht das Mittagessen überspringen. Kontinuierliche Werbung erhöht die Wirkung. Jede zusätzliche Werbewoche für eine Kampagne erbringt mehr als die Woche zuvor, d.h. die Spreizung von Werbemaßnahmen über die Zeit ist kostengünstiger. Diese zwei Gesetze sagen dem Planer, dass mehr Werbewochen besser sind als ein stärkeres Werbegewicht in einer Woche. Das dritte Gesetz: Esse eine ausgewogene Diät. Die Verringerung geringer Kontakte beschreibt sowohl Medien als auch Medien-Gewichte, die die Wichtigkeit von Media-Mix erklären. Marketing-Mix-Modelle weisen oft umgekehrte Erfolge zu dem Anteil der Spendings auf. Je mehr Dollar in TV oder Zeitschriften investiert werden, desto mehr geht die Wirkung pro Dollar jeweils nach unten. Es kommt immer ein Punkt, wo der nächste Dollar irgendwo anders eingesetzt werden sollte. Diesen Punkt zu finden, ist der Schlüssel zum Media-Mix. Es setzt voraus, die Zusatzkosten für den nächsten wirksamen Kontakt zu kennen und zu vergleichen. Wenn Werbungtreibende sagen, dass sie den Einsatz anderer Medien überlegen, da TV weniger kosteneffektiv ist, meinen sie, dass der Crossover-Punkt früher auftritt. Wenn wir bedenken, was wir alles schon wissen, sollten Zeitschriften einen Feiertag haben.

Wenn Print zu einem TV-Plan hinzukommt, wird das die Reichweite mehr erhöhen, da Print und TV sich weniger überschneiden als TV mit TV. Zeitschriften-Zielgruppen eignen sich besser zum Aufbau von selektiver Reichweite als TV. Diese zwei Faktoren, Verringerung der Wirksamkeit (von TV) und Zielgruppen-Marketing, werden die Media-Ausgaben unvermeidlich umstrukturieren. Die totale Dominanz von TV wird einem ausgeglichenen Media-Mix Platz machen. Nach 35 Jahren Durchs-Fenster-Zuschauen werden die Zeitschriften den Bankett-Tisch erobern.

Was ist falsch an diesem Bild? Einfach gesagt, es wird nicht so eintreffen. Es hat keine erkennbare Verschiebung von TV nach Print gegeben. Das ist das vierte Weight Watcher Gesetz, dass ich vorher nicht erwähnt habe. Zähle die Kalorien ehrlich. Und darum geht es in diesem Vortrag.

Zeitschriften gewinnen immer

An prominenter Stelle, auf der Website der Magazine Publishers of America, lautet die Überschrift: "Zeitschriften sind TV überlegen". Dieser jährliche FCB-Bericht ist jahrelang ein Dorn im Auge der TV-Networks gewesen. Zeitschriften gewinnen immer. Dieses Jahr überprüft CBS diesen Vergleich. Manche Zeitschriften haben umfangreichere Leserschaft-

ten als TV-Shows Zuschauer, aber diese Zeitschriften-Daten übertreiben damit, was der Werbungtreibende erhält. Network-TV weist die Seherschaft für eine durchschnittliche Minute aus. Zeitschriften nutzen die Leserschaftsdaten für eine durchschnittliche Ausgabe. CBS hält dagegen, dass das äquivalent für TV die Gesamt-Seherschaft wäre und die ist 10% höher für die halbe Stunde und 20% bis 30% höher für längere (zeitschriftenähnliche) Programme. Eine kürzliche Erhebung von Survivor 2 erbrachte dreifach höhere Gesamt-Reichweiten als die durchschnittliche Seherschaft.

Werbeträger- versus Werbemittel-Kontakt

Planer arbeiten mit der Zahl der Personen, die ein TV-Programm sehen oder eine Zeitschriften-Ausgabe lesen (Werbeträger-Kontakt). Die relevantere Zahl wäre aber, wie viele Personen eine durchschnittliche Anzeige sehen (Werbemittel-Kontakt). Unterschiede zwischen der Wahrscheinlichkeit eines Lesers und eines Sehers Werbung gesehen zu haben, machen den Vergleich zwischen Reichweite und Kosten pro Tausend missverständlich. Die durchschnittliche Minute beim TV ist ein guter Ersatz für den Werbemittel-Kontakt, da die Größenordnung zwischen der Reichweite für eine durchschnittliche Minute und für die durchschnittliche Seherschaft von Werbung gleich ist. Die Leserschaft einer durchschnittlichen Ausgabe ist kein Ersatz für den durchschnittlichen Werbemittel-Kontakt mit einer Anzeige, da nicht alle Leser auch Anzeigen aufschlagen.

In der Vergangenheit war dieses kein Thema. Print wurde nicht so geplant. Die Einheit sind Schaltungen und nicht Kontakte. Die Reichweiten wurden mehr zur Auswahl der Zeitschriften genutzt, wo absolute Größenordnungen nicht so wichtig waren, wie die konkurrierenden Titel. Aber beim Einsatz von Media-Mix-Plänen werden Kontakte wichtig und Reichweitendaten werden die Wahl des Mediums beeinflussen und wie wir damit umgehen - sicher nicht so, wie die Verleger das gerne hätten. Die hohen Reichweiten und niedrigen Tausender-Kosten werden die Zeitschriften schlagen. Das Ziel der Werbungtreibenden ist meistens ein Verkaufserfolg. Wenn Zeitschriften-Leser einen niedrigeren Anzeigen-Kontakt haben als TV-Seher, dann sind die niedrigeren Tausenderpreise, die durch höhere Leserschaften zustande kommen, kontraproduktiv. Werber werden die Illusion von mehr Werbedruck für weniger Geld akzeptieren und nicht adäquate Zeitschriften-Pläne erstellen.

Dann, wenn die Wahl heißt Zeitschriften oder TV, werden die Zeitschriften verlieren. Irreführende geringe Tausenderpreise werden ineffektive Zeitschriften-Pläne nach sich ziehen, die nur wenig Werbewirkung aufweisen. Und das ist das, was Werber vorfinden werden. Ich denke, CBS sollte ruhig sein. Die schlechte Art und Weise, wie Agenturen Print planen, lassen TV so gut aussehen.

Gesamt-Reichweite versus durchschnittliche Reichweite

Wir finden klare Beweise für eine Reichweiten-Inflation bei den Leserschaftsdaten, wenn wir die durchschnittlich aufgeschlagenen Seiten gegen die Lesefrequenz gezählt, anschauen. Mein Beispiel beinhaltet eine wöchentliche Zeitschrift mit hoher Auflage, hohem Leser pro Exemplar-Wert und einer Reichweite, die größer ist als die irgendeiner TV-Show.

Reichweite einer durchschnittlichen Ausgabe
Anzeigen-Leser im Vergleich zu Ausgaben-Leser
Hohe Auflage, hohe Leser pro Exemplar-Werte, MRI Frühjahr 2000

	0 gele- sen	1 gele- sen	2 gele- sen	3 gele- sen	4 gele- sen	Gesamt	%
Gelesen von Ø 4 Ausgaben							
Leser einer Ø Ausgabe (000)	1.452	7.589	8.936	4.216	12.274	34.467	100
% Verteilung	4%	22%	26%	12%	36%	100%	
% aufgeschlagene Seiten	53%	73%	82%	106%	139%	104%	
Spezifische Anzeige gesehen (000)	770	5.540	7.328	4.216	12.274	30.128	87

52% der Leser (4+22+26) sehen nicht alle Seiten einer Ausgabe (53%, 73%, 82%). Dies sind die weniger regelmäßigen Leser. 48% der Leser (12+36) schlagen alle Seiten auf, manche auch mehr als einmal. Das sind die regelmäßigen Leser. Wenn wir die Zahl der Leser in jeder Frequenzgruppe mit der Zahl der gelesenen Seiten multiplizieren, sehen wir, dass der durchschnittliche Leser 13% der Anzeigen nicht sieht, d.h. die Anzeigen-Leserschaft ist 13% geringer als die Ausgaben-Reichweite. Das Problem wird größer, wenn wir uns Reichweiten-Berechnungen anschauen. Der Einfluss zwischen gelegentlichen Lesern und dem geringeren Anteil an Anzeigen, die gesehen wurden, ist hier kritisch, da Zeit-schriften nur Reichweite aufbauen können, wenn sie unregelmäßige Leser hinzuzählen.

Werbemittel-Reichweite versus Werbeträger-Reichweite

Die nächste Analyse berechnet die Differenz zwischen Werbemittel- und Werbeträger-Reichweite, wie sie über vier Ausgaben aufgebaut wird. Es wird genauso durchgeführt, wie die Berechnung der Reichweite für eine durchschnittliche Ausgabe aus der Frequenzskala. 1 von 4 erhält das Gewicht 0.25, 2 von 4 = 0.5, 3 von 4 = 0.75 und 4 von 4 = 100.

Reichweite von vier Ausgaben nach Lesefrequenz
Hohe Auflage, hohe Leser pro Exemplar-Werte, Berechnung des Autors*

Gelesen von Ø 4 Ausgaben	0 von 4	1 von 4	2 von 4	3 von 4	4 von 4	Gesamt
Leser einer Ø Ausgabe (000)	1.452	7.589	8.936	4.216	12.274	34.467
4-Ausgaben Reichweite	1.452	30.356	17.872	5.621	12.274	67.575
Erhöhung	0	22.767	8.936	1.405	0	33.108
in %	0	69	27	4	0	100

* basierend auf MRI Frühjahrs-Daten 2000

Die kumulative Reichweite ist mehr als doppelt so groß wie die Reichweite für eine durchschnittliche Ausgabe, aber 96% (69+27) sind unregelmäßige Leser, die nicht jede Anzeige sehen. Die neu hinzugekommenen Leser aus der o.a. Berechnung sehen 24% der Anzeigen nicht.

(Tab. mit nicht gesehenen Anzeigen)

Das gleiche passiert, wenn Zeitschriften kombiniert werden, um Reichweite zu erhalten. Die unregelmäßigen Leser werden für den Reichweiten-Aufbau benötigt und die lesen nicht die ganze Ausgabe.

Die meisten Titel sind betroffen

Das ist kein gelegentlich auftretendes Problem für schnell-lebige Zeitschriften. MRI produzierte Seiten-Reichweiten für 188 Titel, davon erreichten 165 (88%) einen durchschnittlichen Seiten-Kontakt von unter 1.0 in der Frequenzgruppe 0 von 4 Ausgaben gelesen, 132 (70%) in der Frequenzgruppe 1 von 4 Ausgaben und 75 (40%) in der Frequenzgruppe 2 von 4 Ausgaben.

(Tab. mit Frequenzgruppen und Zahl der Titel mit Seiten-Kontakt unter 1.0)

Die Verteilung für eine typische Zeitschrift entspricht der U-Form. Eine große Gruppe liest 4 von 4 Ausgaben, eine kleine Gruppe 3 von 4 und eine große Gruppe weniger als 3 von 4. Personen, die weniger als 3 von 4 Ausgaben lesen, sind im allgemeinen mehr als die Hälfte aller Leser und sie sehen weniger Anzeigen.

Alle diese Berechnungen sind kein Argument gegen die Recent-Reading-Methode (Abfrage zum letzten Lesen). In der heutigen Zeit wird eine Mediagattung die Technik verwenden, die die Gesamt-Nutzerschaft einbezieht und zählt. Aber es ist dann Sache eines intelligenten Media-Einkäufers diese Reichweiten auf ein realistisches Niveau herunterzurechnen. Das ist die Aufgabe von Media-Forschung. Der Vortrag zeigt, dass wir die Leserschaftsdaten justieren können, um eine bessere Schätzung der Reichweiten vorzunehmen. Die "Seiten-Kontakte-an-der-Frequenz-Korrektur" ist ein einfacher Weg dieses zu tun, bei Gebrauch von Daten der Leserschafts-Untersuchung selber.

Die Größe der Reichweiten-Justierung wird für jeden Titel unterschiedlich sein. Einige Zeitschriften bieten das wiederholte Lesen der Seiten mit Werbung an, aber die Haupt-Planungs-Methode ist mit Reichweiten und nicht mit Brutto-Kontakten. Es zeigt sich, dass durch Leserschaftsdaten die Werbemittel-Reichweiten ziemlich übertrieben werden.

You gave us the 'answers', now ask us the questions! - How to maximize magazine advertising effectiveness in the planning/buying process

Ihr gabt uns die 'Antworten', nun fragt uns die Fragen! - Wie die Werbewirkung von Anzeigen in Zeitschriften im Planungs- / Einkaufsprozess gesteigert werden kann

Tony Jarvis, MediaCom / Grey Global Group

Zusammenfassung

Der Autor des Vortrags fasst hier etliche Vorträge von anderen Symposien oder Konferenzen zusammen, um damit auf pragmatische Weise Werbewirkung von Anzeigen in Zeitschriften bzw. deren Steigerung erklären zu können. Außerdem wurden eine Reihe von Fokus-Gruppen mit Planern und Einkäufern plus Kunden durchgeführt, um verstehen zu können, was effektiv erreicht werden kann. Dabei wurde herausgefunden, dass nur so, "neue Dimensionen für Zeitschriften-Planung und -Einkauf" empfohlen werden können. Die gefundenen "Antworten" verlangen nicht nur neue Dimensionen in Zeitschriften-Planung, -Einkauf und -Verkauf sondern werfen auch Fragen über die Messung von Zeitschriften-Leserschaften auf. Es könnte sein, dass der neue Print-Planungs-Ansatz Einfluß auf Fusionen und Media-Mix-Optimierungsprogramme hat.

Hintergrund

Die Vielzahl an Veröffentlichungen zu den World Wide Readership Symposien und anderen Veranstaltungen werden kaum im täglichen Geschäft von den Planern und Einkäufern in Agenturen genutzt. Die Ergebnisse sind oft schwer in die tägliche Anwendung zu integrieren. Dieser Vortrag soll Hilfestellungen geben, die sofort übernommen werden können.

Schlüssel-Fragen

1. Soll Print mehr von dem TV-/Print-Budget erhalten?
2. Sollten Werbungtreibende eine Print-Kampagne ohne Pretest fahren? Und, sollte Print mit TV getestet werden?
3. Sollten Werbungtreibende die Zahl der (kreativen) Kampagnenmotive für Print erhöhen?
4. Sollte Print, ähnlich wie TV, wöchentlich geplant werden, obwohl es gegen die strategische Zielgruppenausrichtung ist?
5. Sollte Print für Einzelmotive oder für Kampagnen geplant werden?
6. Sollte Print mit höheren TRPs (Target Rating Points) geplant werden, um TV-Niveau zu erreichen? Und ist es überhaupt erreichbar?
7. Sollte der "neue" Print-Planungs-Ansatz besser mit TV-Planung koordiniert werden?
8. Sollten Zeitschriften vor Network TV geplant werden?
9. Wird die Industrie weitere Forschungen unterstützen?
10. Ist die Industrie bereit, sich zu einer längst überfälligen Verbesserung der Leserschaftsforschung, auch für Lebenszyklen für Zeitschriften, zu verpflichten?

Jetziges Prinzip der TV und Zeitschriften-Werbung

Im Zeitalter des Controlling, wo nur zählt, das nächste Monatsergebnis zu erreichen, ist

der kurzfristige Verkaufserfolg der Grund für TV-Werbung. Für Print-Werbung sprechen langfristig erreichbare Marken-Aufbau-Effekte. Beides wird mit dem Marken-Wert und der Anzeigen-Erinnerung gemessen, nicht jedoch mit Abverkaufszahlen, d.h. Zeitschriften werden hierbei benachteiligt.

Wie Zeitschriften und TV arbeiten

Print-Werbung erhält sofortige Aufmerksamkeit - the power of print - und scheint schnell zu ermüden. Wie Millward Brown (UK) in der Studie AdTrack 1994 zeigt, fällt die Werbeerinnerung nach drei Kontakten signifikant zurück.

(Abb. Wear-out Effekt nach 5 Kontakten)

Andere Faktoren, die Zeitschriften gegen TV unterstützen, beinhalten:

- Der Index für die erste Kontaktchance ist bei Print 1,8 mal höher als bei TV (FIPP).
- Die Qualität von Anzeigen-Werbung wird von den Zeitschriften-Inhalten und dem Produktinteresse beeinflusst.
- Zeitschriften-Botschaften werden, wenn sie relevant erscheinen, schnell aufgenommen.
- Print-Werbung scheint häufiger dazu beizutragen, eine Marke für einen Kauf in Betracht zu ziehen, als TV-Werbung.
- TV-Werbeerinnerung zeigt keinen Wear-Out-Effekt (Millward Brown), deshalb scheint sie geeignet zu sein, langfristige Markenerinnerung, Marken-Werte und Verkaufserfolge aufzubauen.
- Millward Brown ermittelte für MPA in der Ad-Effectiveness-Untersuchung, dass für 113 Marken in zwei Jahren (1997,1998) die durchschnittlichen wöchentlichen GRPs für Zeitschriften 11 betragen.

Warum ist der Einfluss von Zeitschriften-Werbung auf den Verkauf oftmals nicht ersichtlich? Die mögliche Antwort wäre in vielen Fällen ... *weil die Print-Pläne mit so wenig Gewicht versehen werden.*

Die Voraussetzung:

Um die Effektivität von Zeitschriften-Werbung für jedes Schalthniveau steigern zu können, müssen wir das gesammelte Wissen über die Wirkungsweise von Print-Werbung anwenden. Es erfordert neue Dimensionen für Zeitschriften-Planung / -Einkauf und grundlegende Änderungen in der Art, wie Zeitschriften erhoben, analysiert, geplant und eingekauft (und verkauft) werden.

Einige Hauptmerkmale von Zeitschriften (USA):

- Niedrigere Tausenderpreise als Network TV (\$ 8 vs. \$ 18, Frauen 18-49 J.)
- Höhere Quoten als die Top-TV-Shows (Parade Reader's Digest, TV Guide, People, USA Weekend, Better Homes and Gardens, National Geographic versus ER (CBS)).
- Produziert gleich hohe Reichweiten wie TV Prime, wenn es kombiniert genutzt wird.
- Braucht Zeit, um die Gesamt-Leserschaft aufzubauen. (Geht bei TV sofort.)
- Die einzelne Anzeige wird selten getestet. (Was macht eine gute Anzeige aus?)
- Wird meist im Studio bzw. auf lokaler Basis ausgeführt und nicht in exotischer Umgebung.

- Wird typischerweise mit TV verglichen. Auf Dollar-Basis oder die Zahl der Einschaltungen nicht wöchentlicher GRPs oder Reichweiten und Frequenzen - ein "Apfel- und Birnen-Vergleich".
- Wird für die Dauer der Kampagne analysiert, geplant und gekauft und nicht auf einer wöchentlichen oder vierwöchentlichen Basis, wie TV.
- Wird den Kunden auf Flow Charts präsentiert, die keine Aussagen über die Größe der Leserschaften machen, die über die Zeit gewonnen werden. Solche Charts vermitteln eine gefährliche Illusion.

(Abb. eines konventionellen Flow Charts)

Laut der 1999er MPA-Untersuchung erreichen Zeitschriften in einem typischen Mixverhältnis (TV 80%, Zeitschriften 20% oder TV 90%, Zeitschriften 10%) die Werbeerinnerung kosteneffektiver als TV

- über alle Kategorien
- mit unterschiedlicher Budgetgröße
- mit unterschiedlichen Mix-Spendings
- in verschiedenen Konkurrenzsituationen
- für neue und etablierte Marken

(Tab. Index-Vergleich Erinnerung gegen Kosten TV und Zeitschriften - MPA Untersuchung 1999)

Einige vorgeschlagene Antworten zu den Schlüsselfragen:

Fr. 1 Sollte Print mehr Anteile von einem TV-/Zeitschriften-Budget erhalten? Ja?

Wenn Print weniger als 50% des Budgets umfasst, sollte das um wenigstens 10 Prozentpunkte erhöht werden. (John Philip Jones, Telmar Workshop 1998)

Beispiel:	Originalplan:	TV-/Zeitschriften-Budget	\$ 40 Mio.	
		10% Zeitschriften	\$ 4 Mio.	
		90% TV	\$ 36 Mio.	(5.800 GRPs)

	Neuer Plan:	TV-/Zeitschriften-Budget	\$ 40 Mio.	
		20% Zeitschriften	\$ 8 Mio.	
		80% TV	\$ 32 Mio.	(< 5.160 GRPs)

(Aufzählung verschiedener Untersuchungen, die dieses beweisen)

Fr. 2A) Sollten Werbungtreibende eine Print-Kampagne schalten, ohne sie zu testen?

Nein!

Warum sollte eine Kampagne eingesetzt werden, ohne zu wissen, wie die kreative Umsetzung arbeitet und wie die Nachricht kommuniziert wird? Zu riskant. Die Werbewirkung (langfristiger und kurzfristiger ROI) wird durch Kreativität in Kombination mit dem Werbedruck, erreicht.

Fr. 2B) ... Und, sollte Print mit TV getestet werden? Ja?

TV und Zeitschriften sollen kombiniert getestet werden. Media-Mix erhöht die Reichweite in der Zielgruppe. Gleichfalls wichtig ist jedoch die Steigerung des Kommunikationseffektes (Multiplier Effekt).

Fr. 3 Sollten Werbungtreibende die Zahl der kreativen Motive in einer Print-Kampagne erhöhen? Ja?

Da es Indikatoren auf einen schnellen Wear-Out-Effekt gibt, sollte die Zahl der Motive erhöht werden, um die Anzeigen-Wirkung zu erhöhen.

"Eine Erhöhung der Motive von vier auf sechs bringt eine Erhöhung der Werbewirkung um 35% (63% - 85%)." Alan Smith, FIPP, Canadian Advertising Research Foundation Workshop, Toronto 1997.

Fr. 4 Sollte Print, ähnlich wie TV, wöchentlich geplant werden, obwohl es gegen die strategische Zielgruppenausrichtung ist? Ja?

Die neue Zeitdimension-Pilot-Studie von MRI erschien Ende 2000 und auch Telmar und IMS bieten entsprechende Planungssoftware für Zeitschriften an. Um die Leserschaftszuwächse im Zeitverlauf darstellen zu können, müssen die Flow Charts geändert werden.

(Abb. eines "neuen" Flow Charts)

Fr. 5 Sollte Print sowohl für Einzelmotive als auch für Kampagnen geplant werden? Ja?

Millward Brown fand heraus, dass der Wear-Out-Effekt nach vier Kontakten einsetzt. Daher sollte man die Frequenz für jedes Motiv überprüfen und die Zielgruppen, die vier oder mehr Kontaktchancen haben, minimieren. Das ist ein trickreiches Konzept und sehr zeitaufwendig - aber noch einmal: es geht nicht um die hübschen Darstellungen auf den Planungscharts.

Fr. 6 Sollte Print auf höheren TRPs und Reichweiten-Niveaus geplant werden? Wahrscheinlich?

Nutze typische wöchentliche Planungsziele als Führer:

Periode	TRPs	Reichweite
1 Woche	60-80	35-50%
4 Wochen	240-320	70-80%
13 Wochen	780-1.040	85% +

Bei Einbeziehung von wöchentlichen Reichweiten muss die Zahl der einzubeziehenden Zeitschriften erhöht werden und das reduziert die Rabatte für Mehrfachbelegungen. Reichweiten-Ziele sollten niedriger als bei TV angesetzt werden oder in Kombination mit TV auf wöchentlicher oder vierwöchentlicher Basis.

Fr. 7 Sollte der "neue" Zeitschriften-Planungs-Ansatz mit TV koordiniert werden? Ja?

Obwohl der Kommunikationseffekt unterschiedlich für lang- und kurzfristige Werbewirkung sein kann, unterscheiden Verbraucher nicht, von welchem Medium sie die Botschaft erhalten haben. Bei "neuer" Planung auf wöchentlicher Basis, sollte analysiert werden, ob man damit nicht immer wieder die gleichen Personen erreicht.

"Print plus TV lassen das Werbebudget härter arbeiten, da die Werbebotschaft verschiedenartig kommuniziert wird." Alan Smith, FIPP, CARF Workshop, Toronto 1997.

Fr. 8 Sollten Zeitschriften vor TV geplant werden? Ja!

Die jetzige Praxis kann beschrieben werden mit: Upside Down Media. Ist die Planung von Zeitschriften nach TV, wie die Planung von TV-Spots vor Network?

Fr. 9 Wird die Industrie weitere Forschungen unterstützen?

Wird in weitere Forschung investiert werden?

Fr. 10 Ist die Industrie bereit, sich zu einer längst überfälligen Verbesserung der Leserschaftsforschung auch für Lebenszyklen für Zeitschriften zu verpflichten?

Um Werbekampagnen in Zeitschriften zu unterstützen, ist es notwendig,

- Kommunikationseffektivität mit den richtigen Zielen zu planen
- auf Basis von Ausgaben und Leserschaftszuwächsen zu kaufen
- wöchentlich oder monatlich zu analysieren.

Um das zu erreichen, müssen einige kritische Forschungs-Auswirkungen beachtet werden.

1. Es sind bessere Daten zur Lebensdauer von Zeitschriften notwendig.
2. Es müssen zeitbasierte Optimierungsprogramme für Zeitschriften entwickelt werden.
3. Die MRI/NTI-Fusion muss gut arbeiten.
4. Es sind schnelle zeitbasierte Media-Mix-Optimierungsprogramme notwendig, um fusionierte Daten nutzen zu können.

Zusammenfassung

Es wird eine "neue" Dimension für Zeitschriften-Planung und -Einkauf benötigt. Wenn die Richtung dieses Vortrags akzeptiert wird, wird es helfen, bei einigen Marken Media-Mix mit einem größeren Budgetanteil für Zeitschriften auszustatten. Es werden jedoch substantielle Veränderungen in der Forschung notwendig. Außerdem ist es notwendig, die Zeitschriften-Werbung von der kreativen Seite und hinsichtlich der Marken-Perspektive zu testen.

An examination of differential weighting of successive waves of product usage estimates

Eine Prüfung von verschiedenen Gewichtungen aufeinanderfolgender Wellen von Verbrauchsdaten

Stuart Gray, BBDO/Julian Baim & Martin Frankel, MRI-NOP World

1999 erklärten die Autoren, dass bei einem Zusammenspielen von mehreren Wellen von Leserschaftsdaten, bei einer stärkeren Gewichtung der neuesten Daten, bessere Vorhersagen möglich sind. Insbesondere ein Verhältnis von 20 zu 80 produziert bessere Vorhersagedaten als das traditionelle Verhältnis von 50 zu 50.

Dieser Vortrag soll über ähnliche Vorgehensweisen bei der Nutzung von Produkt-Daten berichten. Wenn die Ergebnisse hierbei ähnlich sind, wäre es eine sinnvolle Überlegung, eine Gewichtung von 20 zu 80 bei Leserschafts- und Produkt-Daten vorzunehmen.

Einige der zwischen zwei Wellen beobachteten Unterschiede in der Produktnutzung können auf Stichproben-Unterschiede zurückzuführen sein, einige können aber auch auf aktuelle Trends hinweisen. Um das zu testen, wurden verschiedene Gewichtungen der Wellen vorgenommen: einmal 40 zu 60 und einmal 20 zu 80 im Vergleich zur 50 zu 50 Gewichtung. Damit sollte das ideale Gewicht für die aktuellsten Daten und damit für Vorhersagen herausgefunden werden.

Für die Analyse wurden MRI-Daten der Wellen 31 bis 40, d.h. von März 1994 bis März 1999, genutzt.

Produktkategorie-Daten

Insgesamt war eine Gewichtung von 20 zu 80 am besten. Sie ergab bei der Vorhersage für die nächste Periode 10%ige Verbesserungen gegenüber der Gewichtung 50 zu 50. Die Gewichtung von 40 zu 60 ergab eine 4%ige Verbesserung. Da dieses nur sehr geringe Verbesserungen sind, nahmen wir an, dass Produktkategorien, die von einer Periode zur nächsten größere Änderungen aufweisen, auch mehr davon profitieren, wenn der jeweils jüngsten Erhebungswelle mehr Gewicht verliehen wird. Deshalb errechneten wir für die Daten den z-score. Bei einem z-score von 2 und mehr ergab eine Gewichtung von 20 zu 80 eine Verbesserung von 20% gegenüber einer Gewichtung von 50 zu 50. Und bei einem z-score von 3 und mehr ergab es eine Verbesserung von 27%. Auch eine Gewichtung von 40 zu 60 war besser als eine von 50 zu 50, aber nicht so gut wie 20 zu 80.

(Hinweise auf zwei Tab. im Anhang)

Marken-Daten

Insgesamt ergaben sich hier keine Unterschiede bei den verschiedenen Gewichtungen.
(Hinweis auf eine Tab. im Anhang)

Wir nahmen jedoch wieder an, dass Marken, die von einer Periode zur nächsten größere Änderungen aufweisen, auch mehr von einer höheren Gewichtung der jüngsten Welle pro-

fitieren. Auch hier wurden für die Marken wieder die z-scores errechnet und sie in diejenigen, die einen z-score von 2 und mehr bzw. 3 und mehr aufweisen, eingeteilt. Bei einem z-score von 2 und mehr ergab sich eine Verbesserung von 3% bei einer Gewichtung von 20 zu 80 gegenüber einer Gewichtung von 50 zu 50. Bei Marken mit einem z-score von 3 und mehr betrug die Verbesserung 10%. Eine Gewichtung von 40 zu 60 war bei Marken mit einem z-score von 2 und mehr genauso gut wie bei 20 zu 80; bei einem z-score von 3 und mehr aber weniger gut.

(Hinweis auf zwei Tab. im Anhang)

Außerdem wurde die Konsistenz der Vorhersagen überprüft, indem die verschiedenen Wellen einzeln betrachtet wurden, d.h. es wurden die Wellen 31/32 und 33/34 gewichtet, um die Wellen 35/36 vorherzusagen, die Wellen 33/34 und 35/36 für die Wellen 37/38 und die Wellen 35/36 und 37/38 für 39/40. In allen drei Kategorien produzierte eine Gewichtung von 20 zu 80 bei den Produktkategorien geringere Fehler als bei 50 zu 50 oder 40 zu 60. Bei den Marken-Daten war dieses Ergebnis nicht so klar. Nur eine Datengruppe zeigte geringere Fehler bei einer Gewichtung von 20 zu 80 auf. Das gleiche Ergebnis ergab sich für Marken mit einem z-score von 2 und mehr. Für Marken mit einem z-score von 3 und mehr war in allen drei Markengruppen eine Gewichtung von 20 zu 80 besser als 50 zu 50 und in zwei von drei Fällen besser als 40 zu 60.

Nachfolgend wurde noch die Annahme überprüft, ob es nicht gut wäre, den jüngsten Daten ein Gewicht von 100 zu geben, also eine Gewichtung von 0 zu 100 vorzunehmen. Dabei zeigte sich für Produktkategorien, dass diese Gewichtung genauso gut war wie 50 zu 50. Für Marken-Daten war diese Gewichtung nie so gut wie 20 zu 80.

Wir nehmen an, dass für Produkt-Daten das Sample groß genug ist, für Marken-Daten jedoch nicht. D.h. bei 0 zu 100 werden nur die Daten der jüngsten Stichprobe genommen, bei 20 zu 80 werden die Daten der jüngsten Stichprobe virtuell vervierfacht und so erlangen die jüngsten Daten mehr Einfluss.

Diese Ergebnisse unterstützen die Schlussfolgerung, dass eine Gewichtung von 20 zu 80 sowohl für Leserschafts-Daten als auch für Produktverbrauchs-Daten geeignet ist.

(6 Tabellen mit der Darstellung der durchschnittlichen relativen Fehler)

To prototype or not to prototype, that is the question - can USA Guidelines be embraced worldwide?

Vorgabe von Standards oder keine Vorgabe von Standards, das ist die Frage - Können Richtlinien der USA die ganze Welt einschließen?

Tony Jarvis, MediaCom/Grey Global Group/Stephen A. Douglas, The Douglas/Jones Group

Hintergrund

Es gibt eine Vielzahl von Zeitschriften, die eine große Zahl von Spezial-Zielgruppen bedienen und damit auch Zielgruppen-Marken unterstützen. Für viele hundert davon können die Leserschaften nicht in den syndicated Untersuchungen erhoben werden und dafür bietet die Vorgabe von Standards eine Lösung. Vorgaben, um die Leserschaften aus einer Abonnentenuntersuchung ersatzweise als "gemessene" Daten einzusetzen, ist mehr eine Kunst als eine Wissenschaft. Da jedoch der Druck besteht, Planungsdaten in die computer-gestützten Planungssysteme aufnehmen zu können, wird dieses in den USA durchgeführt. Vorgaben bestimmen die Daten, um für nichterhobene Zeitschriften, geschätzte Leserschaften in der entsprechenden syndicated Untersuchung für computer-gestützte Media-Analysen aufnehmen zu können. Das erhöht die Effizienz für Media-Planer, -Käufer und -Verkäufer. Der Vortrag beinhaltet einen historischen Überblick über die Vorgabe-Praxis in den USA, die jetzige Vorgehensweise von Agenturen und Verlegern und die neuen vorläufigen USA-Richtlinien mit Vorschlägen der Mitglieder des ARF Print Councils.

Wichtigkeit

Weltweit möchten die Werbungtreibenden zunehmend die Media-Planung von den Agenturen über alle verfügbaren Titel erstellt bekommen, unabhängig davon, ob die Titel erhoben wurden oder nicht. Diese Daten werden immer wichtiger, da die Werbungtreibenden sie für Geschäftsergebnismodelle benötigen (z.B. Print-Anteil an den Verkäufen). Und immer mehr gerät es zu einer beängstigenden Aufgabe, Leserschaften für nicht erhobene Zeitschriften zu quantifizieren, insbesondere weil so viele Titel nicht erhoben werden.

Die Zahl der erhobenen Titel erhöhte sich in den letzten 20 Jahren in den USA um 48%, insgesamt stieg die Zahl der Zeitschriften aber um 76%.

(Tab. mit Titel-Anstieg in 5-Jahres-Schritten,
1980: Zahl der Zeitschriften 10.236, Zahl der erhobenen Titel 147,
2000: Zahl der Zeitschriften 17.970, Zahl der erhobenen Titel 217)

Standard-Vorgaben sind ein Weg (in den USA ein erfolgreicher Weg) um neue oder nicht erhobene Zeitschriften und neue Zeitschriften-Kategorien in syndicated Leserschafts-Analysen einzubeziehen und wichtig für Media-Planung, -Einkauf und -Verkauf.

Warum begannen Agenturen mit den Vorgaben?

Es war für die Planer in den Agenturen unbefriedigend für so viele Zeitschriften keine Daten zu haben, da sie nicht erhoben wurden. Die Zahl der abgefragten Titel konnte jedoch auch nicht erhöht werden. Man behalf sich mit Schätzungen, die aber nicht institutionalisi-

siert wurden. Zum ersten mal erwähnt wurde es 1991 in Hong Kong auf dem Readership Research Symposium. Vorgestellt wurde ein Computer-Programm von Roger Baron.

Da die Media-Abteilungen heruntergefahren wurden, wurde es notwendig, den Vorgabe-Prozess für Standards zu zentralisieren, um den Planern und Käufern Zeit zu sparen. Viele Agenturen entwickelten für sich standardisierte Vorgehensweisen für alle Planungsgruppen. Die Agentur-Media-Forscher entwickelten Schätzungen und stimmten sie mit den Planern ab. Dieses Vorgehen variierte jedoch von Agentur zu Agentur. Die Zentralisierung erforderte geringere Personalstunden für diese Arbeit. Zentralisierung ist wie eine Klärungsstelle und standardisiert den Prozess, somit kann er erklärt und verteidigt werden. Einige Agenturen haben die Zentralisierung kritisiert, da sie meinen, die Sensibilität und die Überprüfung bleiben bei Vorgehensweisen nach Schema F auf der Strecke. Außerdem glauben sie, das Standard-Vorgaben die Verantwortung der Planer, die gesamten Media-Optionen zu überprüfen, herabsetzen.

Die jetzige Situation

Die Zahl der Zeitschriften, die nicht erhoben wurden, stieg in den neunziger Jahren und 2000 sehr stark. Im Jahr 2001 gab es einen allgemeinen wirtschaftlichen Abschwung und einige US-Zeitschriften haben ihr Erscheinen eingestellt. Media-Agenturen wie MediaCom, STAR-COM, MindShare, Foot Cone & Belding und Saatchi USA produzieren während des Jahres zwischen 100 und 500 geschützte Standard-Vorgaben für neue oder nicht erhobene Publikationen. Die Douglas/Jones Group untersuchte im August 2001 die Anzahl der allgemein zugänglichen Standards und derjenigen mit demographischen Angaben, die bei IMS zugänglich waren. Insgesamt wurden 166 Standards und demographische Angaben ermittelt. 129 davon waren Prototypen.

Alle diese Schätzungen wurden entweder von Verlagen erstellt oder von Firmen wie DJG (Douglas/Jones Group). Die großen Agenturen überprüfen die von den Verlagen herausgegebenen Schätzungen. Nach der Überprüfung bereiten sie ihre eigenen Schätzungen mit ihren eigenen Standards auf. Auf der Basis einer vertraulichen Standard-Schätzung von einigen Agenturen wurde herausgefunden, dass die Standard-Profile von Verlagen und Agenturen im allgemeinen das "aktuelle" Profil von Zeitschriften widerspiegeln, wenn sie hinterher erhoben werden. Die Gesamt-Leserschaften der Schätzung auf der Basis der Standard-Vorgaben variierten sehr weit und abweichend von den "aktuellen" syndicated Untersuchungen. Obwohl es mit den ARF-Daten kein rechtes Vorankommen gab, traten viele Agenturen dem Komitee für Standards, das allgemeine Prinzipien und Richtlinien für den besten Gebrauch formulierte, bei. Es wird erwartet, dass von dem Komitee im Namen des ARF-Print-Councils Standards entwickelt werden, die vom ARF-Vorstand geprüft werden.

Die Autoren des Vortrags sind der Meinung, dass weltweit die Verlage, die ihre eigenen Schätzungen vornehmen, Agenturen haben werden, die dieses mit mehr Vertrauen überprüfen, wenn die hier vorgestellten Regeln angewandt werden. Das wird für einige Länder eine "Revolution" darstellen.

Definition eines Standards (Prototype)

Definition: Ein Zeitschriften-Standard ist die Leserschafts-Schätzung für einen nicht erhobenen Titel und wird zum Gebrauch innerhalb der syndicated Untersuchungen, für Media-Planung, -Kauf und -Verkauf, entwickelt. Typischerweise basiert er auf existierenden erhobenen Zeitschriften oder "Hosts".

Diese Definition hat drei Varianten:

- Eine Verlags-Schätzung ist eine Schätzung der Leserschaft für eine nicht erhobene Zeitschrift ohne eine Abonnenten-Untersuchung. Im allgemeinen handelt es sich hierbei um Publikationen, die noch nicht veröffentlicht wurden. Sie bieten aber reale Anzeigen-Seiten, ein nachgemachtes Editorial und sind bereit, Auflage zu garantieren. Beispiele für Publikationen, die im vorhinein Verlags-Schätzungen veröffentlichten, sind: Oprah, National Geographic Explorer und My Generation. Einige Media-Forschungsabteilungen in großen Agenturen werden jedoch solche Verlags-Schätzungen nicht berücksichtigen. Sie fordern Planungs-Zielgruppen, die an dem neuen Titel interessiert sein werden, um diese zu überprüfen und dann empfehlen zu können. Dabei spielt der Wert der Marke und das Umfeld für die Anzeige eine Rolle.

- Eine traditionelle Standard-Schätzung präsentiert Daten, die auf ausreichend veröffentlichten Ausgaben basieren, um ein ABC oder BPA Statement (Anmerkung: ähnlich dem deutschen IVW) zu erlangen sowie aus einer Abonnenten-Untersuchung, die von einem renommierten Institut durchgeführt wurde, nachdem mindestens drei bis vier Ausgaben erschienen sind.

- Die Leserschaftsschätzungen für demographische Standards wurden seit den siebziger Jahren entwickelt. Der Prozess wurde von dem Launch von Newsweek Executive, Time Business und Time Top Management angestoßen. Beide Verlage produzierten unterschiedliche Schätzungen für jeden Titel; damit mussten die Agenturen jede Schätzung überprüfen und Entscheidungen treffen.

Agenturen verlangen folgendes, bevor sie mit Standards arbeiten:

- Eine Abonnenten-Untersuchung hoher Qualität. ABC (Audit Bureau of Circulation) hat dazu Standards vorgeschrieben. In diesen Standards sind Stichprobenauswahl und die Überprüfung der Studie enthalten. Es gibt keine Mindestausschöpfung, aber traditionellerweise sollte sie 50% betragen. Die Abonnenten-Studie ist notwendig, um "Host"-Titel mit vergleichbaren Angaben zu Demographie und Lifestyle aussuchen zu können. Eine Kopie der Durchführungs-Regeln der ABC Zeitschriften-Abonnenten-Studie ist im Anhang enthalten.

- Forschungs-Ergebnisse von Dritten können ebenfalls als Grundstock zum Aufbau von Standards oder Verlagsschätzungen benutzt werden. Mendelsohn z.B. erhob Investors Business Daily. DJG schlug einen Standard für MRI vor, der das demographische Profil übernahm und für die MMR-Daten ähnliche Verhältnisse produzierte. Die Analyse wurde in der MRI-Untersuchung in Haushalten mit \$ 75.000 und mehr Haushaltseinkommen, durchgeführt, die vergleichbar zur MMR ist. Ein weiteres Beispiel ist die Verwendung von MRI-

Produkt-Daten als Anker für die "Host"-Auswahl, z.B. Gartendaten für eine Garten-Zeitschrift.

- Jüngste Ausgaben der Zeitschriften. Agenturen wollen keine Media-Kits sondern ein komplettes Briefing, um sich ein Bild über die inhaltliche Position des Titels machen zu können.
- Auflagen-Daten inkl. ABC- oder BPA-Statement.
 - Mindestauflage 100.000 bis 250.000
 - Detaillierte Auflagen-Pläne inkl. Erscheinungsfrequenz-Änderungen und Änderungen in der Verbreitung. Erklärung der Abonnenten-Entwicklung, Einzelverbrauchs-Verbreitung oder zukünftige Pläne für den Einzelverkauf usw.
- Minimum-Frequenz von 4 Ausgaben pro Jahr. Manche Agenturen nehmen nur Standards von mindestens 6 Ausgaben auf.
- Informationen, die helfen, den Leser pro Ausgabe zu schätzen.
- Vor dem Prozess der Standard-Entwicklung wird zunächst die Media-Forschungs-Abteilung gebeten einen solchen zu entwickeln.

Die Verfahrensweisen der syndicated Untersuchungen bei Standards:

- MRI erlaubt Verlagen und Agenturen Standards zu erstellen, solange sie Abonnenten der Untersuchungen sind, die sie nutzen wollen. Prototypen müssen klar gekennzeichnet sein.
- Mendelsohn erlaubt Standards nur in den ersten fünf Jahren oder bis sie erhoben werden können. Außerdem werden sie für Titel erlaubt, die niemals erhoben werden können, wie Stagebill. Agenturen, die Abonnenten des MMR sind, können jederzeit Standards erstellen.
- IntelliQuest CIMS erlaubt Verlagen Standards für ein Jahr bevor sie Ergebnisse veröffentlichen können. Magazine, für die keine Ergebnisse nach einem Jahr vorgelegt werden können, dürfen auch nicht als Standards dargestellt werden. Agenturen können einen Titel jederzeit als Standard darstellen, so lange sie CIMS abonniert haben.
- J. D. Powers erlauben Publikationen kein Abonnement für die Standard-Erstellung. Agenturen dürfen, so lange sie Abonnenten sind, Standards erstellen.

Wann & warum Verlage Standards erstellen

- Wenn Agenturen nach Zusammensetzungs-Daten, Leser pro Exemplar-Daten oder Reichweiten fragen.
- Wenn die Auflage zeigt, dass der Titel erhoben werden könnte.

- Wenn die Verkaufsabteilung meint, dass dadurch das Verkaufspotential erhöht werden könnte - besonders gegenüber der Konkurrenz.
- Weil auch eine "gute" Abonnenten-Untersuchung nicht die Gesamt-Leserschaft abdeckt.
- Weil viele Agenturen Forschung als "Screener" nutzen.
- Weil man nicht "existiert", wenn man nicht im Computer von Planern und Einkäufern ist.

Überblick über Anforderungen an Standards
(Beschreibung der einzubeziehenden Variablen)

Überprüfung der Methoden für die Entwicklung von Standards
(In diesem Abschnitt werden Standards für ein fiktives Beispiel (Travel Holiday) entwickelt. Sie basieren weitgehend auf Daten, die das Produktinteresse abdecken.)

Computer Modelling

Für die Entwicklung von Standards werden Computer-Modelle eingesetzt. Ein gutes Beispiel dafür ist ein Programm mit dem Namen UMPIRE (Unmeasured Magazine Profile Method). (Nachfolgend wird sehr kurz die Arbeitsweise von UMPIRE und die einbezogenen Variablen beschrieben.)

Wie die Leser pro Exemplar-Werte geschätzt werden können

Der letzte Schritt ist die Schätzung des Lesers pro Exemplar auf der Basis durchschnittlicher (ABC/BPA) Auflagen. Dabei müssen die Konkurrenztitel berücksichtigt werden. Das ist ein extrem schwieriger Teil des Prozesses. Normalerweise werden zwei oder drei LpE analysiert. Einige Forscher beginnen mit dem "typischen" LpE-Wert. Roger Baron 3,0 / Cynthia Evans 2,5 / Tony Jarvis 2,8. Manche Agenturen beginnen mit 1,0.

Leser pro Exemplar-Überlegungen

Das Leserschafts-Niveau wird durch Multiplikation des LpE mit der durchschnittlichen Gesamtauflage gebildet. Bei der Entwicklung des LpE sind die Verbreitungsart, das Genre, die Verbreitung an öffentlichen Plätzen extrem wichtige Elemente. In Abonnenten-Untersuchungen werden die Mitleser erhoben. Diese Frage ist jedoch fundamental falsch, da die Befragten niemals sagen können, wie viele andere Personen ihre Ausgabe lesen.

Eine Bemerkung zu Überschneidungen

Es gibt keine Möglichkeiten Überschneidungen schätzen zu können. Die Titel, für die Standards entwickelt werden, haben i.d.R. kleine Leserschaften.

Wie werden Standards in der Media-Planung eingesetzt?

Nur mit strengster Vorsicht und nur in Computer-Analysen:

- für Rangreihen-zählungen und Kreuzzählungen - ja

- Reichweiten- und Frequenzanalysen - Nein
- Optimierungen - Nein

Nachfolgend werden in dem Vortrag Richtlinien und Regeln für Standards vorgeschlagen.

Im Anhang befinden sich ABC-Forschungs-Standards für schriftliche Abonnenten-Untersuchungen.

Fusion - Can magazines be equal players in a fused database?

Fusion - Können Zeitschriften gleichwertige Mitspieler in einem fusionierten Datenbestand sein?

Tony Jarvis, MediaCom /Grey Global Group/Leslie Wood, Leslie Wood Research

Zusammenfassung

In den USA beginnt man gerade Fusionen als Methode im Media-Mix-Vergleich in Erwägung zu ziehen. Wie eine Fusion durchgeführt wird hat einen entscheidenden Einfluß auf Media-Planungen und -Einkauf. Zeitschriften sollten in einer fusionierten Datenbasis gleichgewichtig mit TV sein. Eine getestete Methode fusionierte die Gesamt-Datenbestände von MRI und NTI. Dieses Verfahren, die Gesamt-Datenbestände zu nehmen, stellt sicher, dass die Zeitschriften ihre Leserschaften und Überschneidungen beibehalten und die robusten Produktverwendungs- und Lifestyle-Daten intakt bleiben.

Einleitung

Dieser Vortrag beschreibt, wo Fusionen auf dem amerikanischen Markt sinnvoll sind. Er zeigt, welche Daten fusioniert werden sollten und warum die Fusionen auf diese Art und Weise durchgeführt werden sollten. Die ARF definiert eine Fusion als eine "pseudo Single-Source-Datenbasis", die aus zwei oder mehreren anderen Single-Source-Datenbasen kreiert wurde. Das steht im Gegensatz zu manchen breiter angelegten Daten-Integrationen, die weniger rigorose Techniken zur Kombination von aggregierten oder zusammengefassten Ergebnissen von mehr als einer Untersuchung, anwenden. Es wird gezeigt, wie Zeitschriften fusioniert werden können, um zum einen dem US-Markt Rechnung zu tragen und zum anderen ihr "richtiges Gewicht" gegenüber TV zu behalten.

Anmerkung: Der Vortrag beinhaltet keine aktuellen Ergebnisse der Tests, da Fusion Lab die Ergebnisse zunächst zur Überprüfung vor ihren Teilnehmern präsentieren muss.

Hintergrund

The Fusion Lab testete eine Reihe von Fusions-Methoden. Zwei dieser Verfahren beinhalteten die Gesamt-Stichprobengröße von MRI (Zeitschriften- und Produktverbrauchs-Studie) und Nielsen (die Network TV-Untersuchung). Das steht im Gegensatz zu herkömmlichen Fusionen, wo in dem fusionierten Datenbestand nicht die Gesamt Stichprobengröße beider Untersuchungen enthalten ist.

Die Hauptfragen sind:

- Wofür können fusionierte Daten in den USA benutzt werden?
- Worauf achten wir bei einem fusionierten Media-Datenbestand? (Wie genau sind die fusionierten Daten?)
- Welches ist der beste Weg, um zu fusionieren und die Notwendigkeiten des US-Marktes zu berücksichtigen?

Wofür können fusionierte Daten in den USA benutzt werden?

Zur Zeit werden Fusionen für folgende Zwecke in Erwägung gezogen:

1. Produkt-Verwendung und/oder Lifestyle Ratings für TV-Shows und Tages-einteilungen
2. Media-Mix-Reichweiten und Frequenz
 - a) gegen allgemeine Produkt-Verwendung oder Lifestyle
 - b) gegen allgemeine Demographie
3. Media-Mix-Optimierung (auf Werbeträgerebene)
4. Als Entscheidungselement für Cross Media Budget-Entscheidungen

Die Autoren schlagen vor, fusionierte Media-Daten nicht als Währung beim Media-Einkauf zu nutzen.

Sind alle diese o.a. Zwecke gleich wichtig?

1. Produkt-Verwendungs-Zielgruppen Ratings für TV Shows und/oder Tageszeitungsverteilungen sind wahrscheinlich der wichtigste Faktor für Fusionen. Planer können Tagesverteilungen und Programm-Typen nach Sendern vergleichen. Media-Einkäufer können fusionierte Daten für den Vergleich von Programm-Paketen nutzen - und alles gegen Produkt- und Lifestyle-Daten.

2. Media-Mix-Reichweiten und Frequenzen sind der zweitwichtigste Grund. Planer benötigen eine bessere Möglichkeit alternative Media-Optionen vergleichen zu können, sowohl intermedial als auch intramedial. Damit können Fragen der Belegungen und/oder der Synergien zwischen den Medien beantwortet werden.

3. Media-Mix-Optimierungen sind der dritte Grund. Wenn die Optimierung hilft, den Mix zu regeln, werden die Nutzer eventuell einen erhöhten Zeitschriften-Budgetanteil beschließen. Das wird zuerst ein Schock sein, aber Planer dann davon überzeugen, ihre Zeitschriften-Budgets zu erhöhen.

4. Crossmediale Budget-Allokationen werden mit fusionierten Media-Datensätzen möglich und geben den Entscheidungen eine größere Basis.

Worauf achten wir bei einem fusionierten Media-Datenbestand?

Es gibt zwei Gebiete, bei denen die Genauigkeit kritisch ist:

1. "Gute" Produktverwendungs-Daten für TV-Shows, nicht Episode für Episode oder notwendigerweise Show für Show sondern für Programm-Typen und Tagesverteilungen für Einzel-Sender.

2. "Gute" intermediale und intramediale Überschneidungen zwischen TV-Programmen/ Tagesverteilungen und einem breiten Spektrum von Zeitschriften, für das volle Spektrum an Produkt-Verwendung und demographische Zielgruppen.

Was ist der beste Weg, um zu fusionieren und die Notwendigkeiten des US-Marktes zu berücksichtigen?

In den meisten Teilen der Welt werden Fusionen durchgeführt, indem der größere Datenbestand (MRI) in den kleineren (Nielsen) fusioniert wird. Das gibt der kleineren Datenbasis beständige Ergebnisse und reduziert die Stichprobe der größeren Datenbasis.

Vielleicht ist der ideale fusionierte Datenbestand ein vollzähliges MRI-Sample mit Televisions-Ratings, die mit Nielsen Ratings übereinstimmen. Dieses Ziel ist ein erreichbareres, als die Daten lediglich zu fusionieren. Das hat einige Gründe:

1. MRI beinhaltet bereits TV-Daten, die helfen können, die Befragten zu fusionieren.
2. Die Robustheit der Produktverwendungs- und Zeitschriften-Daten können unverändert bleiben.

Wenn Nielsen-Daten zu den MRI-Daten hinzukommen muss einiges beachtet werden:

1. Zeit: Nielsen-TV-Daten ändern sich häufig; MRI-Zeitschriften-Daten sind statischer.
2. Die Analyse-Werkzeuge, die von den zwei Datenbeständen genutzt werden, sind unterschiedlich.

Nielsen-Daten

- i. Reichweiten und Frequenzen für TV werden bei Nielsen mit Programm-Optimierern produziert, die meistens neue Daten oder kumulierte Reichweiten reflektieren und auf modellierten Reichweiten-Kurven basieren.
- ii. TV-Programm Rating-Analyse-Systeme sind sehr kompliziert und nutzen sehr zeitnahe Daten.

MRI-Daten

- iii. Reichweiten und Frequenzen für Zeitschriften bei den MRI-Daten basieren auf der Anzahl der Ausgaben, die jeder Befragte liest. Diese Daten werden in C1 und C2 übersetzt, die in Modellrechnungen mit Überschneidungen einfließen.
- iv. Produkt-Verwendungs-Daten von MRI werden oft für Kreuztabellierungs-Systeme oder andere statistische Analyse-Software genutzt. Damit werden Planer unterstützt, Zielgruppen-Märkte zu verstehen und Media-Alternativen zu nutzen.

Ein fusionierter Datenbestand soll das MRI-Daten-Format und das Nielsen-Daten-Format widerspiegeln. Da die Nielsen-Daten so oft wechseln, müssten die Nielsen-Daten von Datenbestand zu Datenbestand verändert werden oder es müsste jede Woche ein neuer Bestand erstellt werden. Die Hauptfrage ist: Wie oft muss eine komplette neue Fusion durchgeführt werden?

Mit der Fusion des kleineren Datenbestandes (Nielsen) in den größeren (MRI) werden eher die Notwendigkeiten des US-Marktes berücksichtigt. Jedoch werden auch nicht alle Befragten des kleineren Datenbestandes benutzt. Constrained Fusionen nutzen alle Befragten von beiden Stichproben und fusionieren auch das Gesamt-Gewicht jeder Person.

Fusion der Gesamt-Stichprobe

Die Technik, die Fusion Lab entwickelte, nutzt alle Befragten. Dabei ist es jedoch nicht notwendig, dass jeder Befragte ein Gewicht hat. Diese Methode nennt sich Full Sample Fusion.

Die Grundgesamtheiten beider Stichproben werden gemischt. Es werden die Nielsen Bevölkerungszahlen für Alter und Geschlecht gegen Ortsgrößen ausgewählt. Die Befragten-Gewichte für Personen aus größeren Alter/Geschlecht/Ortsgrößen-Definitionen werden von den MRI-Daten übernommen und in der Summe an den Nielsen-Daten justiert.

Für jede Person aus jeder Stichprobe wird der "ideale Partner" ausgewählt. Dabei wird die Anzahl der "Partner" begrenzt. Ist ein idealer Partner gefunden, werden die Personen verbunden und erhalten das MRI-Gewicht. Bei mehreren Partnern teilt jeder MRI-Befragte das Gewicht mit allen seinen Partnern. Der fusionierte Datenbestand ist dadurch etwas größer als der größere der beiden Datenbestände.

Weiterhin werden

1. Alle fusionierten TV-Show-Ratings eng an Nielsen angelehnt (für Geschlecht / Alter / Ortsgrößen)
2. MRI-Befragte hinsichtlich Alter / Geschlecht / Ortsgröße weitgehend belassen, wie sie sind
3. Alle Befragten von beiden Stichproben genommen.

Durchgeführt wurden drei der Fusionen MRI und Nielsen, drei Nielsen und Nielsen und drei MRI und MRI (insgesamt neun).

Ergebnisse

Bleiben der Präsentation des Vortrags vorbehalten.

Warum Fusionen der Gesamt-Stichprobe Zeitschriften betreffen

Bei üblichen Fusionen werden die Daten eines größeren Datenbestandes in einen kleineren fusioniert. Der größere Datenbestand beinhaltet im allgemeinen die Zeitschriften-Daten, d.h. die Stichprobenbasis für die Zeitschrift wird dramatisch reduziert. Das hat auf die Zeitschriften direkte Einflüsse:

1. Nur Zeitschriften mit großer Auflage haben ausreichend Fälle und können ausgewiesen werden.
2. Alle Zeitschriften-Daten ändern sich - einschließlich der Leserschaften und der Überschneidungen.
3. Wenn kleinere Titel verloren gehen und die Leserschaften sich ändern, werden Zeitschriften-Media-Pläne nicht ihr Reichweitenpotential erreichen und werden im Cross-Media-Vergleich nicht so wichtig genommen.
4. Die Reichhaltigkeit der Produktverwendungsdaten wird vermindert. Für die meisten Produktkategorien wird es weniger Befragte geben und die Lifestyle-Daten werden nicht mehr so robust sein.

Zeitschriften sollten deshalb ein Interesse daran haben, auf welche Art Fusionen durchgeführt werden und wie die Ergebnisse der fusionierten Datenbestände aussehen.

Andere kritische Dinge

Mit dem Fortschreiten der integrierten Media-Mix-Planung sollten Zeitschriften einiges beachten:

- Leserschaftszuwächse über den Zeitablauf

Wenn Zeitschriften beim Media-Mix berücksichtigt werden wollen, sollten sie Daten anbieten, die eine wöchentliche Planung ermöglichen. Verstehen, nutzen und überprüfen von Zeitschriften-Lebenszyklus-Daten zum Einbringen in einen fusionierten Datenbestand, sollte ein Hauptanliegen von Verlagen und Agenturen sein.

- Kosten der Nielsen-Daten

Zeitschriften müssen Zugang zu den Nielsen Leserschaftsdaten zu erträglichen Kosten erhalten. Ohne diese Voraussetzung haben sie keinen Zugang zu fusionierten Daten.

Zusammenfassung

Integrierte Media-Mix-Planung ist nicht so weit weg, wie man denken mag. Fusionen werden das beschleunigen. Es ist eine neue Welt, die wir betreten und jeder sollte in der Lage sein "in dem tiefen Wasser mitzuschwimmen".

Magazine audience accumulation: Discussion of research issues, modeling and application

Zeitschriften-Leserschafts-Kumulationen: Diskussion über die Forschung, das Modelling und ihre Anwendung

Julian Baim, Martin R. Frankel, Joseph Agresti und Risa Becker, MRI-NOP World

Die Messung von Leserschafts-Zuwächsen für mehr als eine Handvoll von Zeitschriften mit erträglichen Kosten ist eine entmutigende Aufgabe. Wahrscheinlich weil in diesem Bereich seit 20 Jahren keine empirische Forschung stattfand. MRI begann mit der Aufgabe Leserschafts-Zuwächse zu messen, mit der Durchführung einiger Vorstufen 1998 und 1999. Historische Versuche anderer, die theoretische Untermauerung der derzeitigen Methode, Beschreibungen und Ergebnisse sowie eine Diskussion der Ergebnisse wurden in zwei Vorträgen im November 1999 in Florenz auf dem Worldwide RRS präsentiert.

Mehrere Methoden sind eingesetzt worden, um Reichweiten-Zuwächse messen zu können. Alle haben ihre eigenen Vor- und Nachteile und stellen Vermutungen über die Lebenszyklen von Zeitschriften an sowie über die Möglichkeit, dass die Befragten das Konzept des "Zum-ersten-Mal-gelesen" verstehen. Wenn wir alles das berücksichtigen, glauben wir, die Tagebuch-Methode ist die beste Methode, um die Reichweiten zu messen. Sie erfordert die wenigsten Vorkenntnisse und bietet die besten Voraussetzungen für die Validität der Antworten. Die endgültige Untersuchung besteht aus 10.615 Befragten, die das Tagebuch über 16 Wochen ausgefüllt haben. Die Stichprobe bestand aus Panel-Teilnehmern des NPD-Panels und wurde disproportional geschichtet, um Männer, 18 bis 44 Jahre, berufstätig, überzurepräsentieren.

Endgültige Stichprobe

	Gesamt	Männer	Frauen
Versand	24.600	14.356	10.244
Zurück erhalten	12.646	6.646	6.000
Nicht verwendbar	2.031	1.436	595
Vollständig ausgewertet	10.615	5.210	5.405
Ausschöpfung in %	43,2	36,3	52,8

Methodische Übersicht

Als die Tagebuch-Methode übernommen wurde, wurden einige kritische Punkte, die die Ergebniskurve beeinflussen, hinterfragt. Darunter waren:

- Etablierung eines Datums, an dem die Zeitschrift zum ersten Mal verkauft wird
- Etablierung einer Zeitschriften-Lebensdauer
- Bearbeitung der Tagebücher für logische Eintragungen
- Bewertung der Annehmbarkeit der Ergebnisse
- Identifizierung der Vorhersagbarkeit von Kumulationen
- Darstellung der Zeitschriften für Modelling-Zwecke
- Aufklärung der Unterschiede zwischen diesen und vorhergehenden Daten

Erst-Verkaufs-Datum

Die meisten Zeitschriften-Informationen-Systeme nennen sehr konkrete Erst-Verkaufs-Daten für die einzelnen Zeitschriften. Diese Daten werden so interpretiert, dass an diesen Tagen die spezifische Ausgabe einer Zeitschrift zum ersten Mal für Abonnenten und Käufer erhältlich ist. Während der Untersuchung gaben Befragte an, vor diesem Datum bereits diese Ausgabe der Zeitschrift gelesen zu haben. Zunächst wurde dadurch die Validität der Tagebuch-Methode angezweifelt. Es stellte sich jedoch heraus, dass die angegebenen Erst-Verkaufsdaten nicht bedeuten, dass eine Ausgabe erst an diesem Tag erhältlich ist. Besonders 14-tägliche und mit längerem Erscheinungsintervall erscheinende Zeitschriften sind vor dem angegebenen Erst-Verkaufstag erhältlich. Kommentar: Methoden, die zur Messung von Leserschaftsdaten die angegebenen Erst-Verkaufstage nehmen, statt des Datums an dem die erste Möglichkeit besteht, die Zeitschrift zu lesen, werden die Kumulationskurve beschneiden.

Das bedeutete, dass MRI zunächst die Daten ermitteln mußte, an denen die Zeitschriften zum ersten Mal gelesen werden konnten. Das erfolgte sehr pragmatisch: Die ermittelten Daten entschieden. Es war zulässig, die Zeitschrift vor dem angegebenen Erst-Verkaufstag zu lesen. Die Daten für drei wöchentlich erscheinende Publikationen unterstützten diese Entscheidung noch einmal. Bei Zeitschriften mit sehr geringen Tagebuchaufzeichnungen blieb das Problem jedoch erhalten. MRI etablierte deshalb ein allgemeines "Erst-Erhältlichkeits-Datum" innerhalb des Erscheinungsintervalls für alle Zeitschriften.

Zeitschriften-Lebensdauer

Es bestand aus der früheren Kumulations-Forschung die Vermutung, dass Zeitschriften einen maximalen Zyklus haben, nach dem es nicht mehr möglich wäre, Angaben zum ersten Lesen zu erhalten. Diese Annahme lag in der Methode begründet. Die Untersuchungsdauer belief sich auf sechs bis 13 Wochen ab dem Erstverkaufsdatum. Als Ergebnis erreichten die Kumulationskurven nach Ablauf 100%. Die Tagebuch-Methode ging nicht von diesen Vermutungen aus. Die Befragten wurden angewiesen alle Lesevorgänge einzutragen, sowohl für Zeitschriften, die zum ersten Mal gelesen wurden als auch für andere. Befragte konnten zwei, drei oder zehn Jahre alte Ausgaben angeben. Das hatte einen bedeutenden Einfluss auf die Kumulationskurven. Das erste Lesen erfolgte über einen längeren Zeitraum als zuvor vermutet wurde, und es dauerte viel länger, um die Gesamt-Leserschaft zu erreichen. Bevor diese Befunde akzeptiert wurden, wurden die Tagebuch-Einträge mit älteren Ausgaben noch einmal kontrolliert und für richtig befunden. Die Zeitschriften-Lebensdauer hat einen offensichtlichen und dramatischen Einfluss auf die Kumulationsgeschwindigkeit und unterscheidet die früheren MRI Kumulationskurven von den jetzigen.

Tagebuch-Bereinigung: Logisch und Philosophisch

Die Tagebuch-Methode erfordert die Möglichkeit, die Eintragungen der Befragten hinsichtlich ihrer Logik und Konsistenz zu überprüfen. MRI entwickelte dazu Bereinigungs-Richtlinien:

- Keine Ausgabe konnte mehr als einmal zum ersten Mal gelesen werden.
- Alle Erst-Lesevorgänge haben Vorrang vor weiteren Lesevorgängen

- Eintragungen mit unkorrekten Ausgabe-Daten wurden verbessert, wenn sie das Lesedatum betrafen
- Wenn das Datum, an dem die Zeitschrift zum ersten Mal erhältlich war, in der Tagebuch-Woche auftrat, mußte der Lesevorgang zum ersten Mal sein.
- Wenn für das Datum der Ausgabe "keine Angabe" gemacht wurde, wurde für wöchentliche, zweiwöchentliche und dreiwöchentliche Zeitschriften standardmäßig der Ausgaben-Monat eingesetzt und der Tag weggelassen.

MRI etablierte diese Regeln, nachdem viele Tagebücher sorgfältig durchgesehen wurden. Diese Regeln zeigten, ob Befragte ein gültiges Ausgaben-Datum angaben, und ob sie das Konzept des ersten Lesens verstanden hatten. Wenn die Kumulationskurven durch diese Regeln beeinflusst wurden, wurden die Erscheinungsintervalle intensiver kontrolliert.

(Tab. mit Bereinigungs-Einfluss auf die Kumulation)

"Annehmbarkeit" der Ergebnisse

Nachdem die Daten bereinigt waren, initiierte MRI eine zusätzliche Bewertung der Ergebnisse. Es wurde der Verlauf der Kumulationskurve für das erste Lesen verglichen mit den Kurven für die ersten Lesemöglichkeiten in der Woche, um ein Mißverständnis bei der Abfrage des ersten Lesens ausschließen zu können. Das Ergebnis zeigte, dass die Kurven zur ersten Lesemöglichkeit in der Woche einige Tage unter denen zum ersten Lesen lagen, aber ähnliche Verläufe hatten. Das zeigte MRI, dass die Ergebnisse robust waren und größeres Vertrauen in die Validität gesetzt werden konnte.

(Tab. mit Kurven-Darstellungen)

Entwicklung von Vorhersagen

Auch eine Stichprobe von 10.000 ist limitiert und die niedrigen Daten für einzelne Zeitschriften verlangten ein Modelling für viele der Kumulationskurven. MRI erreichte dieses durch eine (logistic) Regressionsanalyse, um die Haupt-Vorhersagen für die Kumulation herauszufinden und eine Typologie zu entwickeln, die Zeitschriften mit ähnlichen Charakteristika zusammenfasste. Reichweiten, die per Tagebuch erhoben wurden, unterscheiden sich von denen, die nach der Abfrage zum letzten Lesen erhoben wurden. Auch die Daten zum Zuhause- und Außer-Haus-Lesen sowie die demographische Zusammensetzung unterscheiden sich. Es ist jedoch nicht auf Anhieb klar, ob es notwendig ist, die Tagebuch-Daten anzupassen. Um dieses herauszufinden wurden zunächst die kritischen erklärenden Variablen und ihre Anteile am Verlauf der Kumulationskurve identifiziert. Dazu wurden demographische Merkmale, wie Geschlecht, Alter, Bildung, Berufstätigkeit, Censusregion und Ortsgröße herangezogen. Zusätzlich wurden Attribute der Zeitschrift herangezogen: Anteil der Abonnenten, Leser pro Exemplar-Werte (aus der nationalen Erhebung), Erscheinungsintervall und Gattung.

MRI nutzte ein Regressionsverfahren (logistic regression), um herauszufinden, welche Variablen eine signifikante und entscheidende Rolle in der Erklärung für den Verlauf der Kumulationskurve und ihrer Geschwindigkeit spielen. Es wurde getrennt für wöchentliche und monatliche Zeitschriften durchgeführt. Die Regressionsanalyse wurde für Erst-Leser und für den ersten möglichen Lesetag, für unterschiedliche Tage, durchgeführt: für

wöchentliche Titel Tag 0, Tag 3, Tag 7 und Tag 14; für monatliche Titel: Tag 0, Tag 7, Tag 14 und Tag 30.

Für wöchentliche Zeitschriften waren der Leseort und die Gattung substantielle und signifikante Variable für die Regressionsanalyse. Demographie spielte nur für einige Zeit-Perioden eine Rolle. Hohe Abonnementanteile hatten einen Einfluss am Tag 0, waren danach aber nicht kontinuierlich signifikant. Das läßt vermuten, dass viele Abonnenten von wöchentlichen Zeitschriften diese sofort nach Erhalt lesen und sich die Kumulation an den ersten Tagen des Lebenszyklus konzentriert. Am Ende der ersten Woche verringern die Käufer den Einfluss der Abonnenten an der Kumulationskurve. Für die monatlichen Zeitschriften wurden die gleichen Variablen herausgefunden.

MRI mußte jetzt die Ergebnisse der Regressionsanalyse anwenden, um Kumulationskurven für jede Zeitschrift aus der nationalen Untersuchung zu erstellen. Dafür wurden die Ergebnisse aus dem Tagebuch, bei zu kleinen Fallzahlen aus den jeweiligen Gattungen, übernommen.

(Tab. mit Darstellung der Regression und der Übertragung für die Gattungen)

Gleichzeitig mit diesem Modelling mußte MRI die Proportionen der Leseort-Angaben von den Tagebuch-Aufzeichnungen an die nationale Untersuchung anpassen.

(Tab. mit Darstellung der Unterschiede)

Die Unterschiede, die durch die unterschiedlichen Erhebungsmethoden zustande kommen, sind besonders bei den Außer-Haus-Anteilen für die Leseorte zu bemerken. Die Regressionsanalyse zeigte, dass diese Unterschiede den Verlauf der Kumulationskurve beeinflussen.

Vergleich mit früheren Modellen

Die MRI-Daten zeigten nicht nur die unterschiedlichen Kumulationsverläufe zwischen den Titeln mit dem selben Erscheinungsintervall, sondern auch dramatische Unterschiede im Vergleich zu früheren Modellen mit Kumulationen. Einige davon sind unterschiedlichen methodischen Ansätzen zuzuschreiben, aber einige sind auch auf unterschiedliche Ausgangsannahmen zurückzuführen. IMS und Telmar erstellten Kumulationskurven aufgrund der Ergebnisse der Politz Kumulationsuntersuchung, die vor 40 Jahren durchgeführt wurde und übertrugen es auf das Zeitschriften-Umfeld, das im Jahr 1990 vorherrschte. Da sind große Differenzen zwischen den Kurven der beiden Institute festzustellen. Die MRI-Kurven unterscheiden sich hauptsächlich auf drei Ebenen: tatsächliches Erstverkaufsdatum, Lebenszyklus und Erscheinungsintervall.

Die traditionelle "Erstes-Lesen"-Methode vermutete, dass das Lesen frühestens am angegebenen Erstverkaufstag stattfinden konnte. Seit der Tagebuch-Methode wissen wir, dass das Einsammeln der Leserschaft, insbesondere bei Monats-Titeln, bereits früher beginnt. Das ergibt einen dramatischen Effekt bei der Darstellung von wöchentlichen Reichweiten für Monatstitel.

Frühere Forschungen ließen vermuten, dass Titel mit dem gleichen Erscheinungsintervall die gleichen Lebenszyklen aufweisen. Z.B. bei IMS erreichen einige wöchentliche Titel 100% ihrer Reichweite nach fünf Wochen. Die MRI-Daten zeigen, dass diese Annahme nicht bestätigt werden kann, sondern dass die Beispiel-Titel unterschiedliche Anteile nach fünf Wochen erreichen und auch später noch Reichweite kumulieren.

(Verbale Darstellung weiterer Beispiele mit den Unterschieden zwischen MRI und IMS und Telmar - Tab. dazu im Anhang)

Media-Planung: Verwendung der Kurven

Die Kumulationsuntersuchung erbringt nicht nur neue Erkenntnisse über die Reichweitzuwächse sondern hat auch Auswirkungen auf den Planungs- und Einkaufsprozess.

(Beispiel-Darstellung, wie die neuen Kurven die Auswahl von spezifischen Ausgaben innerhalb eines Plans, der von einem Optimierungsprogramm erstellt wurde, verändern)

Auswirkungen haben dabei die valideren Ergebnisse für das Erste-Lesen auch vor dem Erstverkaufstag und die genauere Messung des Lebenszyklus einer Ausgabe. Damit kann der Planer vermeiden, Ausgaben zu belegen, die außerhalb seiner Kampagnenzeit liegen. Außerdem kann durch wöchentliche Reichweiten und GRPs besser mit anderen Medien konkurriert werden und Planer können beobachten, wie Zeitschriften in einer Mix-Kampagne arbeiten.

Zusammenfassung

Die Zusammenhänge, die bei der Messung von Reichweiten-Zuwächsen auftreten, sind sehr vielfältig. Die Schwierigkeiten und Kosten, die für die Durchführung einer solchen Untersuchung auftreten, können überwältigend sein. Dieser Vortrag zeigt die unzähligen Entscheidungen, die getroffen werden müssen, um Kumulationskurven erstellen zu können. Das beginnt mit der Auswahl der adäquaten Methode, um die Daten zu erheben. Neben theoretischen Überlegungen, wie die Leserschaft eingesammelt werden kann, muss gewährleistet werden, dass die Befragten die Abfrage des "Ersten-Lesens" verstehen. Dieses alleine ist eine entmutigende Aufgabe.

Neben den methodischen Entscheidungen müssen die Bereinigungsphase und das Modelling berücksichtigt werden. Zu jedem Zeitpunkt hat MRI die Notwendigkeit und den Einfluss der Prozeduren auf die Ergebnisse der Kurven, abgeschätzt. Wie immer wäre eine größere Stichprobe wünschenswert gewesen, damit mehr Zeitschriften mit größerer Reliabilität gemessen hätten können. Wir glauben jedoch, dass die neuen Kurven den Lebenszyklus von Zeitschriften besser abbilden und besser zwischen den Titeln differenzieren.

Zur Zeit arbeiten wir mit NOP-UK an einer Kumulationsuntersuchung für Großbritannien.

A new method to measure media casualness for magazines and newspapers

Eine neue Methode um gelegentliches Lesen für Zeitschriften und Tageszeitungen zu erheben

Gary Morgan, Michele Levine & Sergey Dorofeev, Roy Morgan Research

Zusammenfassung

Dieser Vortrag kann mit folgenden Schlußfolgerungen kurz zusammengefaßt werden:

- Eine Schätzung des gelegentlichen Lesens (oder einer Art Reichweiten Zu- und Abnahmen) ist für das Modelling von Media- (Print, TV, Radio, Internet usw.) Reichweiten- und Frequenz-Verteilungen, wichtig. Wenn die gelegentlichen Leser unterschätzt werden, wird die Reichweite für mehr als eine Belegung auch unterschätzt. Das ist sowohl für Mono-Pläne als auch für Media-Mix-Pläne wichtig.
- Wahrscheinlichkeiten für das gelegentliche Lesen für Print-Medien sind mindestens so wichtig, wie Leser pro Ausgabe-Wahrscheinlichkeiten, wenn Reichweiten- und Frequenz-Wahrscheinlichkeiten für eine größere Zahl von Belegungen gezählt werden.
- Es wird eine neue Frequenzfrage zur Messung des gelegentlichen Lesens vorgeschlagen. Die neue Frequenz-Frage bringt genauere Ergebnisse für die Schätzung des gelegentlichen Lesens als die traditionelle Frequenzfrage.
- Die neue "Frequenz-(Abfrage-)Methode" ist eine Einzel-Interview-Version der Abfrage-Methode mit Wiederholungsinterview, die bis heute als Gold-Standard zur Messung von gelegentlichem Lesen oder "Zu- und Abnahmen" gilt. Mit anderen Worten, die neue Methode nutzt ein Einzel-Interview und erreicht Ergebnisse, die bis jetzt nur mit Wiederholungsbefragungen möglich waren.
- Für Tageszeitungen ist das gelegentliche Lesen zwischen verschiedenen Wochen signifikant größer als innerhalb einer Woche. Deshalb kann gelegentliches Lesen zwischen verschiedenen Wochen nicht ersetzt werden durch das innerhalb einer Woche, weil dieses eine signifikante Unterschätzung der Reichweite für mehr als eine Belegung erbringt.
- Wenn eine gleichbleibende Methode zur Messung von Leserschaften für dieselben Zeitschriften in verschiedenen Ländern angewandt wird, werden im allgemeinen gleiche Ergebnisse für Leser pro Exemplar und gelegentliche Leser ermittelt.

Der Vortrag beinhaltet die mathematischen Formeln, um "Media-Unregelmäßigkeiten" für Reichweiten- und Frequenz-Modelling für Media-Pläne mit Print, TV, Radio, Internet und anderen Medien durchführen zu können. Eingeführt wurden "Media-Unregelmäßigkeiten" von Christopher Fry Anfang der 70er Jahre und dann weiterentwickelt von Georg Rennie. Die meisten Beispiele beziehen sich auf Print-Medien, sind aber für andere Medien adaptierbar.

Als Untersuchungsmethode wird single-source angewandt, wobei Informationen über alle Medien eingesammelt werden. Man benötigt für die Messung mindestens zwei Gelegenheiten zum Lesen, Sehen oder Hören für den Befragten. Die traditionelle Meßmethode

besteht aus einem face-to-face Wiederholungsinterview mit demselben Befragten. Mit anderen Worten, dasselbe Interview wird mit demselben Befragten noch einmal durchgeführt. Obwohl das die ideale Befragungsform ist, ist sie insbesondere für große Stichproben zu teuer. Deshalb nutzt Roy Morgan im Anschluß an das face-to-face Interview ein Tagebuch zum Selbstaussfüllen. Für Mehr-Länder-Studien (USA, GB, Neuseeland) werden zur Ermittlung der Leserschaften nur die Tagebuchfragen verwendet.

Die traditionellen Tagebuchfragen um die gelegentlichen Leser zu ermitteln sind die Fragen zum letzten Lesen (z.B. "in der letzten Woche" für wöchentliche Titel) und zur Frequenz (z.B. Zahl der gelesenen Ausgaben in den letzten 4 Wochen für wöchentliche Titel). Das Problem ist, dass mit diesem Ansatz die Frequenz-Antworten von den Befragten im allgemeinen über- oder unterschätzt werden. Als Ergebnis sind die Werte für das gelegentliche Lesen, das aus dem Tagebuch gezählt wird, meistens geringer als die tatsächlichen gelegentlichen Lesevorgänge, die auf einer Establishment-Untersuchung mit Wiederholungsinterviews basieren.

Die vorgeschlagene neue Frequenzfrage erhebt zwei statt vier Ausgaben und soll patentiert werden. Auf dem Fragebogen liegt ein Copyright. Zusammen mit der Frage zum letzten Lesen werden die Befragten gefragt, wie viele Ausgaben sie in den letzten zwei Erscheinungsintervallen gelesen haben: 0, 1 oder 2 und mehr. Die neue Methode, dieses in einem Interview zu erheben zeigt Übereinstimmung mit den Daten, die zweistufig erhoben wurden, mit einer Establishment-Untersuchung und anschließendem Tagebuch. Die Erhebung des gelegentlichen Lesens zusammen mit Reichweiten-Erhebungen ist wichtig für das Modelling von Medien-Reichweiten und Frequenz-Verteilungen. In dem Vortrag werden Beispiele gezeigt, was passiert, wenn die Daten für gelegentliches Lesen falsch sind.

Das Konzept des gelegentlichen Lesens

Die zusätzliche Reichweite von zwei Ausgaben eines Titels im Vergleich zu einer Ausgabe eines Titels ist die zusätzliche Reichweite der ersten Ausgabe über die zweite und umgekehrt. Wenn z.B. 10% die erste Ausgabe lesen aber nicht die zweite, während 15% die zweite Ausgabe lesen aber nicht die erste, beträgt die zusätzliche Reichweite $(10\% + 15\%) / 2 = 12,5\%$.

Die Definition des gelegentlichen Lesens

Wenn zwei Ausgaben eines Titels z.B. eine Reichweite von 10% und 12% haben und davon 6% beide Ausgaben lesen, dann ist die Reichweite für Exklusivleser der ersten Ausgabe: $10\% - 6\% = 4\%$ und für Exklusivleser der zweiten Ausgabe $12\% - 6\% = 6\%$. Die zusätzliche Reichweite beträgt also $(4\% + 6\%) / 2 = 5\%$. Im nächsten Schritt muss die zusätzliche Reichweite ermittelt werden, wenn

- 1) alle Ausgaben die gleiche Reichweite haben und
- 2) die Leserschaft der zweiten Ausgabe unabhängig ist von der Leserschaft der ersten Ausgabe

Die Annahme, dass die Leserschaft die gleiche ist, meint dass beide Ausgaben eine Leserschaft von 11% haben (der Durchschnitt von 10% und 12%). Wenn die Leser der zweiten Ausgabe unabhängig von den Lesern der ersten Ausgabe sind, ist die Verteilung der

11% der zweiten Ausgabe genauso wie die der ersten Ausgabe. D.h. 11% der Leser der ersten Ausgabe genauso wie 11% der Nicht-Leser der ersten Ausgabe lesen die zweite Ausgabe. Die zusätzliche Reichweite dafür wäre 11% von den 89% Nicht-Lesern der ersten Ausgabe, das sind 9,79% vom Gesamt. Konsequenterweise ist das gelegentliche Lesen $5\% / 9,79\% = 0.511$ oder 51,1%. (Darstellung in einer Formel)

Die Original-Definition des gelegentlichen Lesens nutzte nur zwei Ausgaben mit der Annahme, diese haben das gleiche Leserschafts-Niveau. Die neue Definition ist genauer, da diese Annahme verworfen wird, die Leserschaft muss nicht mehr die gleiche sein. Daher ist sie für das Modelling besser geeignet.

(Nachfolgend werden im Paper verschiedene Modellrechnung vorgestellt sowie in zwei Tabellen die Ergebnisse der Establishment-Untersuchung mit der "alten" Frequenzfrage 0 bis 4 gezeigt und mit der "neuen" Frequenzfrage 0 bis 2+. Mit der "neuen" Abfrage sind die Unterschiede signifikant geringer. Im weiteren wird die Verwendung der Betabinomial Verteilung in der Formel für das Modelling der Reichweiten-Kumulation beschrieben.

Außerdem wird noch einmal auf die Unterschiede bei Tageszeitungen zwischen den Tagen verschiedener Wochen und den Tagen innerhalb einer Woche eingegangen. Um dieses Phänomen zu untersuchen wurden in einer Establishment-Untersuchung zunächst die Tages-Reichweiten erhoben und dann mit den anschließenden Tagebucheintragungen verglichen. Dabei wurde im Tagebuch auch noch einmal nach den letzten sieben Einzeltagen und dem Lesen der Tageszeitung gefragt. - Tabelle mit den Ergebnissen)

Im Anschluss daran wird noch einmal auf die Wichtigkeit der gleichen Abfrage hingewiesen um die Ergebnisse für die gleichen Print-Medien in den verschiedenen Ländern vergleichen zu können. Dabei werden Ergebnisse aus USA, Australien und Neuseeland verglichen.

(In den Anhängen werden mathematische Prüfungen vorgestellt.)

Readership Panel: We did it! – History and results of a huge challenge

Leserschafts-Panel: Wir führten es durch! – Entwicklung und Ergebnisse einer großen Aufgabe

Roberto Lobl & Flavio Ferrari, IBOPE Pesquisa de Midia

Einleitung

Es wurden schon zu etlichen Gelegenheiten die Vor- und Nachteile der verfügbaren verschiedenen alternativen Forschungsmethoden zur Messung von Zeitschriften-Leserschaften diskutiert. Als internationaler Standard setzte sich aus pragmatischen Gründen und weil es das beste Preis-Leistungs-Verhältnis bot, die Abfrage-Methode nach dem Letzten-Lesen (Recent Reading), durch. Trotzdem erkennen die Forscher die kontinuierliche Panel-Forschung mit individuellen Leserschafts-Daten als solide und theoretisch adäquate Alternative an, um das zunehmend dynamische und fragmentierte Leseverhalten messen zu können. Die Haupthindernisse zur Einführung dieser Technik sind:

- Kosten für die Panel-Implementierung und -Unterhaltung
- Grenzen bei den Panel-Teilnehmern über das Leseverhalten hinausgehende Fragen über andere interessierende Themen (z.B. Produkte) zu beantworten
- Befürchtungen der Verleger hinsichtlich des Einflusses substantiell unterschiedlicher Indikatoren im Hinblick auf die Werbeaktivitäten der Zeitschriften.

Das vorliegende Paper präsentiert die Ergebnisse eines einzigartigen Experiments eines Zeitschriften-Leserschafts-Panel in Brasilien. Außerdem werden die Schwierigkeiten bei der Panel-Implementation und bei der Durchführung diskutiert und die Unterschiede zu traditionellen Untersuchungen (Target Group Index), die nach dem Recent Reading Modell arbeiten, aufgezeigt.

(Als nächstes wird in dem Vortrag ein Überblick über den brasilianischen Markt gegeben.)

Historie des Zeitschriften-Panels

Begonnen hat es mit dem Fernseh-Panel, das in 3.000 Haushalten mit einem Meter-System die Reichweiten misst. Die Genauigkeit und die Breite der gesammelten Informationen haben den Werbemarkt in die Lage versetzt anspruchsvolle Planungs- und Optimierungsprogramme für die TV-Ausgaben zu entwickeln. Auf der anderen Seite waren die Programme, die für die Print-Ausgaben-Planung zur Verfügung standen, unzureichend. Sie erlaubten keine tiefgehenden Analysen, überschätzten die Leserschaften und machten Vergleiche mit anderen Medien unmöglich. Aus diesem Grund wurde IBOPE beauftragt, ein Leserschafts-Panel zu entwickeln, das genauso anspruchsvoll sein sollte, wie das beim TV genutzte.

Auswahl der Panel-Abteilung

Hier standen entweder die Abteilung, die die TV-Meter-Messungen durchführt zur Verfügung, die Abteilung, die ein Verbraucher-Panel führt und somit Erfahrungen mit Tagebuch-Erhebungen hat oder eine neu zu gründende Abteilung. Ausgewählt wurde die Verbraucher-Panel-Abteilung, weil sie Erfahrungen hatte, ein Panel aufzubauen bei dem die

Befragten Fragebogen zum Selbstausfüllen bekommen und wegen des Einflusses auf die mögliche Nutzung der Erhebung des individuellen Verbraucherverhaltens.

Design der Untersuchung

Das erste Konzept beinhaltete eine Stichprobe von 5.000 Befragten in zehn großstädtischen Gebieten, disproportional aufgestockt in gehobenen Zielgruppen. Der Fragebogen wurde wöchentlich verteilt und eingesammelt und beinhaltete zusätzlich zu den Zeitschriften-Fragen, Fragen zu Verbrauchsgütern und Finanzdienstleistungen. Das Panel wurde Anfang 1997 installiert. In der zweiten Hälfte des Jahres wurde klar, dass das Projekt weit entfernt davon war, erfolgreich zu sein. Es würde zu hohe Investitionen erfordern und zu lange dauern, bis es stabil und mit den richtigen Strukturen laufen würde. Außerdem wurde festgestellt, dass die Vielzahl der abgefragten Produkte den Strukturmangel verstärkten und Ergebnisse mit geringer Qualität erbrachten.

Das führte dazu, dass das Design radikal geändert wurde und 1999 ein Neubeginn erfolgte. Die Stichprobe wurde auf 3.900 Fälle in den gehobenen Zielgruppen, in drei geographischen Regionen, reduziert. Außerdem wurde der Fragebogen überarbeitet und als erstes alle Produktfragen herausgenommen. Das Tagebuch wurde als kleiner Vordruck, den die Befragten täglich mitführen können um ihn beim aktuellen Zeitpunkt des Medienkontakts ausfüllen zu können, konzipiert. Die Tagebücher wurden monatlich verschickt, jeweils für jede Woche eins, die durch unterschiedliche Farben gekennzeichnet waren. Eingesammelt wurden sie wöchentlich. Die Teilnehmer des Panels erhielten anfangs eine Schulung und wurden dann während ihrer Mitarbeit kontinuierlich betreut.

Die Konzentration auf gehobene Zielgruppen erfolgte, weil die anderen Schichten aufgrund zu hoher Kosten oder wegen Analphabetentum kaum Zugang zu wöchentlichen Zeitschriften haben. Die ausgewählte Gruppe ist vergleichbar zu europäischen Märkten, repräsentiert jedoch nicht die brasilianische Gesamtbevölkerung.

Die Hauptschwierigkeiten bei der Panel-Rekrutierung und der -Unterhaltung

Die Stichprobenkonzentration auf gehobene Zielgruppen erforderte eine hohe Kontaktzahl pro Region. Die Panel-Teilnehmer wurden für ihre Mitarbeit belohnt. Dabei wurde ein Meilen-Programm, ähnlich wie bei Fluggesellschaften, eingesetzt.

Um die korrekte Messung der Leserschaftsdaten zu gewährleisten, wurden folgende Informationen überprüft:

- Titel: Sehr oft schrieben die Befragten den für sie gebräuchlichen Namen des Titels auf, was Probleme bei der Titelidentifikation brachte.
- Ausgabe und Jahr: Diese Angaben sind bei den Titeln nicht einheitlich. Einige haben nur eine fortlaufende Nummer, andere beginnen jedes Jahr wieder bei eins, allerdings beim Jahrestag der Zeitschriften-Gründung und nicht am Anfang des Kalenderjahres. Es ist sehr schwierig, die Ausgaben korrekt zu identifizieren. Das ist jedoch notwendig für die Berechnung der Einzel-Ausgaben-Reichweiten. Außerdem sind die Angaben für Befragte schwierig zu finden, da sie überall auf dem Titel stehen können und sehr klein gedruckt sind.

- Herausgeber: Das war zusätzliche Arbeit für die Befragten.
- Lesetag: Wichtige Information für die Ermittlung der Lesefrequenzen. Kein Problem für die Befragten.
- Weitere Informationen, die überprüft wurden: Leseort, persönliches Abonnement-Exemplar.

Um Widersprüche zu vermeiden und um die Qualität der Eintragungen zu erhöhen, wurden bei der Neugestaltung des Tagebuchs Titel-Logos eingedruckt. Dieses hatte jedoch Nachteile für Titel mit niedrigen Auflagen, deren Logos nicht im Tagebuch enthalten waren.

Das größte Problem mit den Befragten bestand darin, die Tagebücher zurückzuerhalten und die fehlende Sorgfalt bei der Eintragung der Ausgaben-Nummern. Um das Zurückschicken der Tagebücher zu verbessern, wurden verschiedene Postsysteme ausprobiert sowie auch ein Tagebuch, das im Internet ausgefüllt werden konnte. Die fehlenden Ausgaben-Nummern konnten oftmals nicht rekonstruiert werden. Es wurde entschieden, in solchen Fällen jeweils die Ausgabe zu nehmen, die vor dem Lesetag zuletzt erschienen war. Einige Titel, die mehr für "Erwachsene" sind, wie z.B. Erotik-Titel, wurden zunächst nicht angegeben, so dass für die Befragten eine Art Siegel mit der Aufschrift "Privat" entwickelt wurde.

Analyse der Daten

Die Informationen werden aus zwei verschiedenen Sichtweisen durchgesehen:

1. aus der Sicht der Mediaplanung, mit einer Reichweiten- und Frequenz-Software für Agenturen und Medien,
2. aus der Sicht des Produkt-Managements (der Zeitschriften-Titel), mit einer Software, die üblicherweise von Werbungtreibenden genutzt wird, wenn sie Verbraucher-Panel analysieren.

Dieses brachte erhebliche Vorteile für die Analyse der Daten, allerdings gab es signifikante konzeptionelle Unterschiede. Die Analyse, bei der die Zeitschriften wie ein Produkt betrachtet wurden, beurteilt die Vergangenheit und diejenige, bei der die Zeitschriften wie ein Medium betrachtet wurden, sollte die zukünftigen Aktivitäten beurteilen. Bei der Behandlung der Daten sind fehlende oder inkorrekte Informationen bei dem Zeitschriften-Verbraucher-Panel nicht so dramatisch, wie bei der Nutzung für Mediapläne. Da aber für die Media-Forschung für die korrekten Zählungen von Reichweiten und Frequenzen alle Daten korrekt sein müssen, wurden sie entsprechend bereinigt. Wenn z.B. eine Berichtswoche fehlte, wurde die entsprechende Woche des Vormonats, für die Daten des Befragten vorhanden waren, zugrunde gelegt.

Die "kulturelle" Frage

Die Schwierigkeiten ein Panel zu installieren liegen nicht nur auf technischer oder methodischer Ebene sondern auch darin, dass ein Panel andere Ergebnismiveaus bringt. Die Reichweiten, die mit der Recent-Reading-Methode (Abfrage nach dem letzten Lesen) erreicht werden und die aufgrund ihrer fantastischen und unglaublichen Höhe der Leser

pro Exemplar-Werte im Markt diskutiert werden, werden mit einem Panel nicht erreicht. Das bringt Unsicherheiten und Befürchtungen für die Zukunft der Zeitschriften-Verlage. Es brauchte ein Jahr ausführlicher Analysen, bevor die Verantwortlichen der Verlage überzeugt waren, dass die neuen Reichweiten die Realität besser widerspiegeln als die alten. Es brauchte noch länger, davon zu überzeugen, dass, wenn die Reichweiten geringer sind, die Gelegenheit da war, zu expandieren, da die Ziele der Werbungtreibenden hinsichtlich der Ausgabenquantität erreicht werden sollten. Basierend auf dem alten System machte es wenig Sinn einen Titel mehrfach zu belegen, da die Reichweiten-Zuwächse nur sehr gering waren. Der brasilianische Markt war (damals) geprägt durch einen Rückgang von 30% bei den TV-Reichweiten bei der Umstellung von Tagebüchern zu Meter-Mess-Geräten.

Ein Vergleich der Panel-Daten mit einer anderen früheren Studie, die die Reichweiten mit anderen Medien verglichen hatte, führte zu dem Schluss, das Zeitschriften ihren Konkurrenzstatus auch mit den neuen Daten beibehalten können. Es war klar, dass die Anzeigen-Verkaufsabteilungen der Verlage und der Markt insgesamt umdenken mußten.

Eine traurige Entscheidung

Ein paar Monate vor der offiziellen Einführung des neuen Meßsystems wurde das Verlags-Management des Sponsors der Untersuchung neu strukturiert und nach einer strategischen Überprüfung der Investment-Prioritäten, das Panel nicht weiter finanziert. Damit wurde das Panel in der ersten Hälfte 2001 nicht weitergeführt. Was blieb, ist die Zufriedenheit es probiert und die Schwierigkeiten ausgeräumt zu haben. Die Erfahrungen und die Vielzahl der Daten, die zwar vertraglich geschützt sind, erlauben es, eine Reihe von Untersuchungen durchzuführen, die die Dynamik des Marktes besser verständlich machen.

Arbeiten mit dem, was verfügbar ist

Zu Vergleichszwecken standen das nicht weiter geführte Panel und Daten aus dem TGI - Target Group Index, der zur gleichen Zeit durchgeführt worden war, zur Verfügung. Beim TGI wurde nach der Recent-Reading-Methode gearbeitet.

Einige wichtige Rückschlüsse können aus den beiden Studien gezogen werden:

- in welcher Größenordnung werden bei der Recent-Reading-Methode die Reichweiten überschätzt
- sind die Abweichungen proportional für alle Titel (wöchentliche und monatliche) gleich
- gibt es eine Möglichkeit die Abweichungen zu minimieren, dadurch, dass bei der Recent-Reading-Methode besser definiert wird, was ein "Leser" ist.

Diese Untersuchungen sind nicht leicht durchzuführen und benötigen Zeit. Trotzdem wollten wir die Gelegenheit nutzen, ein paar grundlegende Daten zu präsentieren.

Überschätzung beim Recent-Reading (Abfrage zum letzten Lesen)

Bei der Analyse von 44 Zeitschriften (31 monatlich erscheinende Zeitschriften, elf wöchentlich erscheinende und zwei 14-täglich erscheinende), die in beiden Studien erhoben wurden, zeigt sich, dass die durchschnittlichen Leserschaft Indizes bei der Recent-

Reading-Methode gegenüber dem Panel überschätzt werden.

(Tabelle)

Die Abweichungen bei der internen (kumulierten) Reichweite und bei externen Überschneidungen sind geringer. Bei den wöchentlichen Titeln sind die Abweichungen größer als bei den monatlichen Titeln. Zwei Gründe scheinen bei Recent-Reading-Untersuchungen für diese Effekte hauptsächlich die Ursache zu sein: das wiederholte Lesen (einer alten Ausgabe) und die "Marken-Stärke". Es soll hier nicht weiter untersucht werden. Wir möchten aber - wie bereits andere Autoren auch - darauf hinweisen, dass die Standard-Leser-Definition in Recent-Reading-Untersuchungen keine absoluten und unbeeinflussten Leserschafts-Daten messen kann.

Neugestaltung der Leser-Definition

Es wurde versucht einen Filter zu finden, der die Überschätzungen von Leserschafts-Daten in Recent-Reading-Untersuchungen minimieren konnte. Dabei konzentrierten wir uns auf das Problem des wiederholten Lesens. Zur Verfügung standen Informationen über: Heftherkunft, Erst-Leser, Lesedauer, Anzahl der Lesevorgänge pro Ausgabe.

Wir überlegten, dass

- die wöchentlichen Titel, die intensiver gelesen werden (wie z.B. aktuelle Illustrierte), ihre Leserschaft eher unter Erst-Lesern im Erscheinungsintervall haben und mindestens eine Stunde gelesen werden.
- wöchentliche Titel, die weniger intensiv gelesen werden, nur in geringen Anteilen weitergegeben werden und mindestens eine halbe Stunde gelesen werden
- monatliche Titel mindestens zweimal im Erscheinungsintervall gelesen werden sollten
- einige Titel aufgrund ihres Charakters üblicherweise nicht ausgeliehen werden.

Bei Benutzung dieser Filter-Variablen für Recent-Reading-Untersuchungen, haben wir neue Leserdefinitionen entwickelt.

(Tabelle)

Die Ergebnisse sind für wöchentliche und für monatliche Titel näher an den Werten des Panels.

Schlussfolgerungen

Die Arbeit ist noch nicht beendet. Das Ziel ist es, zu demonstrieren, dass bei einer Neudefinition der Leser, wenn man einige Filter definiert, der Effekt des wiederholten Lesens minimiert werden kann.

Wenn man alle Schwierigkeiten berücksichtigt und die Kosten eines Panels bedenkt, sollte man über alternative Lösungen nachdenken, mit denen die Qualität der Leserschaften, die von Recent-Reading-Untersuchungen ermittelt werden, verbessert werden können.

The effects of different masthead card designs: colour vs. black and white

Effekte unterschiedlicher Gestaltung von Titelformen: Farblich oder Schwarzweiß

Johannes Schneller, Institut für Demoskopie Allensbach

Hintergrund

Die internationale Printforschung stützt sich ganz überwiegend auf eine bestimmte Methodenkonstellation: auf persönliche Interviews und auf die Erhebungsmethode, die die Leserschaftsdaten mit der Frage nach dem letzten Lesen ermittelt. 79% von 62 auf dem Readership Research Symposium in Florenz dargestellten Untersuchungen machen von der Möglichkeit Gebrauch, im Interview zu den einzelnen Werbeträgern Identifikations- und Erinnerungshilfen vorzulegen.

Diesen visuellen Hilfen sind zwei Aufgaben zugeordnet:

- Sie sollen die einzelnen Publikationen 'benennen' und die in Frage stehende Publikation gegen andere, mit denen sie möglicherweise verwechselt werden kann, abgrenzen.
- Sie sollen als Erinnerungshilfe funktionieren und Assoziationen, Stichworte liefern, wenn einzelne Lesevorgänge aus dem Gedächtnis abgerufen werden müssen.

Bei der Gestaltung der Erinnerungshilfen ergeben sich eine Reihe unterschiedlicher Möglichkeiten, z. B.

- Die Erinnerungshilfen können das Logo der Zeitschrift oder Zeitung abbilden oder nur den Namen in einheitlicher Schrift enthalten.
- Die Namen oder Logos der Titel können zu mehreren auf einer Vorlage abgebildet werden, oder es können Einzelkärtchen mit dem Namen oder Logo nur einer Publikation verwendet werden.
- Die Logos können farblich oder schwarzweiß reproduziert werden.

Besonders die Frage nach schwarzweißer oder farbiger Gestaltung der Logos wird in Deutschland in den letzten Jahren wieder intensiver diskutiert. Durch die nach wie vor ansteigende Zahl der Publikationen wächst die Gefahr von Titelverwechslungen. Viele Verlage versuchen, die Markenannuität und damit den Wiedererkennungswert ihrer Zeitschriften durch die konsequente Verwendung bestimmter Farbcodes zu steigern. Diese charakteristische Farbwahl könnte dann bei der Gestaltung von Logokarten genutzt werden, um die Identifikationsleistung dieser visuellen Hilfen zu erhöhen.

In Deutschland verwendet die Allensbacher Markt- und Werbeträgeranalyse Schwarzweiß-Reproduktionen des Titellokos, die den ganzen nutzbaren Raum auf der Titelform (105 x 74 mm) einnehmen. Dagegen verwendet die Media-Analyse, herausgegeben von der Arbeitsgemeinschaft Media-Analyse, farbige Logoreproduktionen, die auf der Titelform ähnlich platziert sind wie das Originallogo auf der Titelseite der Zeitschrift.

Anlage der empirischen Überprüfungen

Es wurde eine Logo- und Titelseitenanalyse der 100 reichweitenstärksten Publikationen und ein breit angelegtes Experiment mit farbigen und schwarzweiß gestalteten Titelformen durchgeführt.

(1) Logo- und Titelseitenanalyse

Die Logo- und Titelseitenanalyse untersucht, in welchem Umfang und nach welchen Regeln die Verlage zur Gestaltung der Titelformen und der Titelseiten Farbe einsetzen. Untersucht wurden die 100 reichweitenstärksten Publikationen der Allensbacher Markt- und Werbeträgeranalyse (AWA 2000): 94 Zeitschriften, 3 Wochen- und 3 Tageszeitungen. Die Analyse stützt sich bei jedem Titel auf mindestens 6 Ausgaben, die den Zeitraum von Juli/August 2000 bis August/September 2001 abdecken.

(2) Split-ballot-Experiment mit farbigen und schwarzweiß gestalteten Titelformen

Das Experiment berücksichtigt 49 Zeitschriftentitel aus verschiedenen Zeitschriftengattungen, darunter Magazine zum Zeitgeschehen, Programmzeitschriften, Frauenzeitschriften und Special-Interest-Titel zu den Themen Garten, Auto, Reisen, Computer.

(Titelliste im Anhang)

Für jeden Titel wurde eine Karte mit der Schwarzweiß-Reproduktion des Titelformen und eine zweite Karte mit farbigem Titelformen erstellt. Mit Ausnahme der Farbgebung waren beide Versionen identisch. Ermittelt wurden die weitesten Leserkreise, die Lesefrequenzen und die Leser im Erscheinungsintervall.

Das Experiment wurde in zwei aufeinanderfolgende Mehr-Themen-Umfragen mit insgesamt 4.226 Interviews eingeschlossen. In der repräsentativen Halbgruppe A (2.116 Interviews) wurden die farbigen Titelformen vorgelegt und in Halbgruppe B (2.110 Interviews) die schwarzweiß gestalteten Pendanten. Die Interviews wurden durchgeführt im März/April und im April/Mai 2001.

Ergebnisse der Titelseiten- und Logoanalyse

(1) Farbgestaltung der Titelseiten

Farbe auf der Titelseite verwenden 99 der 100 untersuchten reichweitenstärksten Printpublikationen. Auch bei den Zeitungen werden mehrheitlich zumindest Zusatzfarben verwendet. Die Hälfte der Titel (51) setzt auf eine einheitliche Farbgestaltung der ganzen Titelseite und benutzt Farbelemente, die sich bei allen Ausgaben wiederholen. Dazu zählen einzeln oder in Kombination

- die wiederkehrende Verwendung einer dominierenden Farbe, einer bestimmten Farbkombination oder die wiederkehrende Wahl bestimmter Farbtönen wie Erdfarben, Pastellfarben oder besonders kräftige, leuchtende Farben (30 Titel),
- die Wahl immer gleicher Hintergrundfarben (24 Titel) oder
- die wiederkehrende Verwendung eines farbigen Rahmens (9 Titel).

(2) Farbgestaltung der Titellokos

Die Logos von 77 Titeln sind mehrfarbig gestaltet; darunter sind 53 Titel, die das Logo in einem Rahmen mit eigener Hintergrundfarbe präsentieren. 23 Titel verwenden einfarbige Logos. 71 Titel verwenden das Logo in immer gleicher Farbgestaltung. 29 Titel variieren die Farbgestaltung des Logos von Ausgabe zu Ausgabe.

Ergebnisse des Split-Ballot-Experiments zu farbigen und schwarzweißen Titeln

- Für den weitesten Leserkreis ergibt sich im Durchschnitt der 49 Titel kein Unterschied zwischen der Verwendung von farbigen und schwarzweiß gestalteten Titellokos. Geringe Unterschiede ergeben sich bei der Reichweite. Bei der Vorlage von farbigen Titellokos werden um durchschnittlich 6% geringere Reichweiten ermittelt (Index 94) gegenüber der Vorlage von schwarzweiß gestalteten Titellokos.
- Vielfach wird vermutet, daß eine komplexere Gestaltung der Erinnerungshilfen vor allem den unregelmäßigen Lesern einer Publikation eine treffsicherere Titelidentifikation ermöglicht. Ein Einfluß ist nicht zu erkennen.
- Unterschiede ergeben sich zwischen den ausgewählten wöchentlich, 14-täglich und monatlich oder seltener erscheinenden Titeln. Während bei den wöchentlichen Titeln mit Farblogos sogar höhere Leser pro Nummer-Reichweiten gemessen werden (Index 104), führen Farblogos bei den 14-täglichen Titeln und den Titeln, die monatlich oder seltener erscheinen, zu deutlich niedrigeren Reichweiten als die Verwendung von Schwarzweiß-Logos (Index 91 und Index 88).
- Dieser Zusammenhang ist aber weitgehend auf die unterschiedliche Verwendung von Farbe bei der Logo-Gestaltung der einzelnen Ausgaben zurückzuführen. Wir haben in der Analyse unterschieden zwischen Titeln, die die Farbgebung des Titelschriftzuges und Logos bei allen Ausgaben in stets derselben Weise wiederholen, und Titeln, die bei der Farbgebung in der beschriebenen Weise variieren. Die Ergebnisse zeigen, daß der Reichweitenunterschied zwischen der Verwendung von farbigen und schwarzweißen Logokarten besonders groß ist bei Titeln, die die Farbe des Titellokos von Ausgabe zu Ausgabe wechseln (Index 91).

(Tab. mit Ergebnissen)

Farbige Logokarten, obwohl die Logofarbe vieler Titel von Ausgabe zu Ausgabe wechselt?

Die Logokarte dient der Identifizierung einer Zeitschrift oder Zeitung, nicht der Identifikation einer bestimmten Ausgabe dieser Publikation. Sie sollte daher nur jene Elemente der Titelseite enthalten, die von Ausgabe zu Ausgabe wiederkehren. Dies gilt auch für die Farbgestaltung. Die Logoanalyse der 100 reichweitenstärksten Printtitel ergab, daß 29 Titel mit Logos in wechselnden Farben erscheinen. In diesen Fällen kann die Farbe auf der Titelnkarte nur willkürlich festgelegt werden. Für das Titelnkartexperiment wurde z.B. "rot" als besonders aufmerksamkeitsstarke Farbe gewählt.

Das Titelnkartexperiment zeigt, daß bei Titeln mit Logos in wechselnden Farben niedrigere Reichweiten ermittelt werden, wenn farbig statt schwarzweiß gestaltete Logokarten vorgelegt werden. Dies könnte auf eine Irritation von Befragten zurückzuführen sein, die mit dem Titel eine andere Farbe assoziieren als diejenige, die für die Logokarte ausgewählt

wurde. Bei Titeln, die die Logofarben nicht variieren, stimmt die Farbe auf der Logokarte überein mit der Farbe, die der Befragte dem Titel zuordnet.

ADAC-Titel und GEO-Familie

Bei zwei Titeln, bei denen die Gefahr von Verwechslungen besonders hoch erscheint und die Farbinformation eventuell eine besonders wichtige Identifikationshilfe ist, wurde eine zusätzliche Variante der Titelfarbe getestet. Es sind ADAC reisemagazin und GEO Special. ADAC reisemagazin (Auflage rd. 230 Tsd.) ist ein separater Titel der ADAC motorwelt, einer Zeitschrift des größten deutschen Automobilclubs. GEO Special (Auflage rd. 120 Tsd.) ist eine von drei Zeitschriften der engeren GEO-Familie. Bei beiden Reisezeitschriften wurden Logokarten getestet mit einem Textzusatz, der explizit auf die Verwechslungsgefahr mit dem Muttertitel hinweist.

Beide Testversionen

- Farbige Logokarten anstelle von schwarzweißen Logokarten und
- Titelfarben mit einem Textzusatz, der explizit auf die Verwechslungsgefahr mit dem Muttertitel hinweist, anstelle von Titelfarben ohne diesen Zusatz führen zu gleichgerichteten Effekten.

ADAC reisemagazin und GEO Special - Farbige Titelfarben und Titelfarben mit Textzusatz mit gleichgerichtetem Effekt

	Titelfarbe mit farbigem Logo (s/w = 100) Index	Titelfarbe mit Textzusatz (ohne Textzusatz = 100) Index
Weitester Leserkreis	95	90
Leser pro Nummer	86	90

Basis: Bevölkerung ab 16 Jahre

Quelle: Allensbacher Archiv, IfD-Umfragen 7005, 7006

Zusammenfassung

- Ob farbig oder schwarzweiß gestaltete Logokarten vorgelegt werden, beeinflusst die Reichweitenergebnisse nur geringfügig.
- Bei der überwiegenden Mehrzahl der getesteten Titel ergeben sich keine Anhaltspunkte dafür, daß durch farbige Titelfarben Verwechslungen vermieden werden können.
- Bei der scheinbar so leichten Aufgabe, farbige Logokarten zu gestalten, ergibt sich eine grundsätzliche praktische Schwierigkeit. Ein erheblicher Teil der Titel variiert die Farbgestaltung der Logos von Ausgabe zu Ausgabe. Das Titelfarbenexperiment stützt die Vermutung, daß farbige Titellogos in einer willkürlich festgelegten Farbe in diesen Fällen zu Irritationen führen können.
- Es wurden auch unter anderem zwei Titel getestet, die nach den Richtlinien des "corporate design" einer Zeitschriften-Familie gestaltet sind. Es wird vermutet,

dass gerade bei diesen Titeln die Logofarbe eine zur korrekten Titelidentifizierung besonders wertvolle Information darstellt. In beiden Fällen wurde ein Textzusatz getestet, der auf die Verwechslungsgefahr mit dem jeweiligen Muttertitel hinweist. Der Test zeigt, daß Farblogo und Textzusatz den gleichen Effekt haben.

- Die Testergebnisse legen nahe, in der AWA auch künftig schwarzweiß gestaltete Titelkarten zu verwenden. Für die Gestaltung der Titellogos sollten nur Merkmale verwendet werden, die bei allen Ausgaben wiederkehren. Wo Farbe einen ganz herausgehobenen Wiedererkennungswert hat und eine besondere Verwechslungsgefahr besteht, können unter Umständen spezielle Textzusätze verwendet werden. (Anhang: Liste der getesteten Titel)

Self-completion questionnaires: The importance of testing (Based on experience from the british NRS)

Fragebogen zum Selbstaussfüllen: Die Wichtigkeit sie zu testen (Erfahrungen von der britischen NRS - Nationale Leserschafts-Analyse)

Hilary Birt, Ipsos-RSL

Einleitung

Es gibt viele Designfaktoren, die die Ergebnisse bei Nutzung eines Selbstaussfüllers für Leserschaftsfragen beeinflussen können: die Länge, die Anweisungen und die Beispiele, die Fragen-Reihenfolge, das Layout für das Antwortschema, maschinengeschriebene Titelpräsentation oder mit Abbildungen usw. Augenscheinlich können auch subtile Unterschiede einen Einfluss haben. Das bedeutet, dass sorgfältige Tests vor dem Beginn der Untersuchung notwendig sind. Der Vortrag stellt Methoden und Ergebnisse der Tests alternativer Selbstaussfüller für die britische NRS vor. Es bestand die Vorstellung, dass ein Selbstaussfüller genutzt werden kann, wenn es schwierig sein sollte ein normales face-to-face Interview (CAPI) zu erhalten.

Es wurden zwei Versionen für den Test vorgeschlagen. Eine Version war in Reihenfolge und Filterung so, wie der face-to-face NRS-Fragebogen. Die andere Version war kürzer mit einer anderen Fragenreihenfolge. Die Kriterien für die Beurteilung der Selbstaussfüller-Designs waren die Einfachheit des Ausfüllens, die Erzeugung "korrekter" Ergebnisse (im Hinblick auf die Abbildung der Realität des Leseverhaltens des Befragten) und die Vergleichbarkeit der Ergebnisse mit der Standard-CAPI-NRS. Die beschriebenen Tests wurden qualitativ angelegt und sollten die ersten beiden Kriterien beleuchten. Im Anschluß daran wird ein quantitativer Feldtest durchgeführt, um die Vergleichbarkeit mit den Standard-NRS-Daten zu überprüfen, bevor der Selbstaussfüller in die NRS übernommen wird.

Hintergrund

Bei der NRS-Erhebung wird die CAPI-Methode angewandt. Seit einigen Jahren gibt es nun Bedenken wegen sinkender Ausschöpfungsquoten. Dieses gilt besonders im Gebiet von London, wo Berufstätigkeit und Zugangsprobleme (z.B. durch Haussprechanlagen) die Arbeit für Interviewer schwierig machen. Im November 1999 wurde Ipsos von der NRS beauftragt eine gemischte Methode auszuprobieren. Dabei sollte herausgefunden werden, ob zwei verschiedene Interview-Methoden und verschiedene Feld-Management-Techniken die Ausschöpfungsquote in einigen Fällen erhöhen könnte. Die Ergebnisse dieses Tests ließen darauf schließen, daß der Einsatz von Selbstaussfüllern, nach einem vergeblichen Versuch ein persönliches Interview zu bekommen, die Ausschöpfungsquote erhöhen könnte.

Als wir den Selbstaussfüller entwickelten, gab es drei Schlüsselfragen die wir erörterten:

- Der Selbstauffüller musste klar sein, gut auszufüllen und nicht mißverständlich, weil kein Interviewer da sein würde, um Fragen beantworten zu können. Ein zu komplizierter oder einschüchternder Fragebogen würde den Befragten nicht ermutigen, ihn auszufüllen.

- Über die Methode, nach der die Befragten die Fragen beantworten sollten, mußte nachgedacht werden, da sie Einfluß auf die Aufzeichnung der Leserschafts-Daten haben konnte.
- Der Selbstausfüller sollte, so weit wie möglich, kompatibel zum CAPI-Fragebogen sein, da die Daten zusammengespielt werden sollten. Es war jedoch klar, dass die Selbstausfüll-Methode auf jeden Fall andere Ergebnisse bringen würde.

Als Ergebnis der Diskussion dieser Fragen, wurde entschieden, zwei alternative Fragebogen-Versionen mit qualitativen Tests zu überprüfen. Die erste Version, die wir die "erweiterte" Version nannten, enthielt eine ausführlichere Form der Leserschaftsfragen, die dem NRS-Fragebogen ähnlicher waren (inkl. eines Filters zum Lesen im letzten Jahr und mit einer für alle Erscheinungsintervalle gleiche, vollständige Skala zur Abfrage des letzten Lesens).

(Abb. Fragebogenseite)

Die zweite Version, die wir die "kurze" Version nannten, nutzte eine vereinfachte Abfrage der Leserschaftsfragen, ähnlich einer Abfrage, die bereits für Entscheidungsträger genutzt wurde. In der Version gab es keinen Filter zum Lesen im letzten Jahr und die Frage zum letzten Lesen beinhaltete eine simple ja/nein Antwort getrennt für die Erscheinungsintervalle.

(Abb. Fragebogenseite)

Ein weiterer Unterschied zwischen den beiden Versionen war die Anordnung der Titel. In der "erweiterten" Version wurden die Titel nach Gattungen sortiert und in der "kurzen" Version nach Erscheinungsintervallen.

Methode

Bei der qualitativen Untersuchung wurden 30 Studio-Interviews an zwei Tagen im September 2000 durchgeführt. Die Befragten wurden vorrekrutiert. Quotenmerkmale waren Geschlecht, Alter, soziale Schicht und Berufstätigkeit. Zu Beginn wurden die Befragten durch einen Interviewer unterwiesen und gebeten, eine der Fragebogen-Versionen auszufüllen. Danach wurde ihnen die jeweils andere Fragebogen-Version zum Ausfüllen gegeben. (Die Hälfte der Stichprobe begann mit der "erweiterten" Version und die andere Hälfte mit der "kurzen" Version.) Dabei wurde jede Befragungsperson beobachtet und jeweils die Zeit zum Ausfüllen der Leserschaftsfragen und der anderen Fragen gestoppt (wobei die anderen Fragen inkl. der Demographie nur einmal in der jeweils ersten Version ausgefüllt wurden).

Danach wurde ein Gespräch mit dem Befragten durchgeführt, bei dem es im wesentlichen um zwei Fragen ging

- die (empfundene) Mühe(losigkeit), mit der die zwei verschiedenen Versionen ausgefüllt werden konnten (aus Sicht der Befragten)
- eine Überprüfung, wie die verschiedenen Versionen hinsichtlich der Fragentechnik arbeiteten und inwieweit sie das aktuelle Leseverhalten widerspiegeln.

Mühe(losigkeit) des Ausfüllens

Beide Versionen wurden im allgemeinen akzeptiert. Es gab nur wenige Fragen während des Ausfüllens. Das bedeutet, dass die Anforderungen an die Befragten sehr klar waren und dass beide Versionen aufrichtig beantwortet wurden. Es gab jedoch einige negative Kommentare. Bei beiden Versionen waren die Anleitungen ziemlich lang und wortreich. Die meisten Befragten sagten, sie hätten sie gelesen, aber es handelte sich hier um eine spezifische Interviewsituation. Die Frage ist, ob sie in einer normalen Interviewsituation zu Hause auch gelesen worden wäre.

Bei beiden Versionen gab es die Tendenz, die Anleitung nicht noch einmal anzuschauen, nachdem man sie einmal gelesen hatte. Dieses führte dazu, in den Kopfspalten des Antwortschemas alle wichtigen Informationen aufzuführen, damit es korrekt ausgefüllt werden konnte. Einige Befragte wußten nicht, ob es um gekaufte Titel ging oder ob gelesen ausreichte. Die Antwortvorgabe "nein" bei der Frage 1 ("erweiterte" Version - Filterfrage Lesen im letzten Jahr, "kurze" Version - Frage zum letzten Lesen) irritierte einige Befragte. Sie meinten, wenn sie bei einigen Titeln "ja" ankreuzen, ergäbe sich das "nein" für die restlichen Titel automatisch. Die Befragten unterscheiden klar, welche Version sie präferieren. Manchmal wird die Version, die als zweites ausgefüllt wurde, als einfacher empfunden, das kann jedoch ein Positions-Effekt sein.

Gründe für die Präferenzierung der "erweiterten" Version

- Bei der für alle Erscheinungsintervalle gleichgehaltenen Skala zur Abfrage des letzten Lesens mussten die Befragten nicht umdenken.
- Die detailliertere mehrstufige Leseskala wurde als besser empfunden, um sich an den letzten Lesevorgang zu erinnern und erschien den Befragten eher das Leseverhalten zu reflektieren.
- Einige meinten, das Ausfüllen sei einfacher, weil die Kästchen enger beieinander seien.
- Die im Fragebogen schräg gestellten Spaltenbezeichnungen wurden von einigen bevorzugt.

Negative Aussagen zur "erweiterten" Version

- Die Frequenzfrage wurde lediglich als Wiederholung der Frage zum letzten Lesen verstanden und führte zu Konfusionen.
- Die Antwort "nicht in den letzten 12 Monaten" in der Frage zum letzten Lesen wurde als redundant zur Frage 1 verstanden, wo bereits gefragt wird, ob der Titel in den letzten 12 Monaten gelesen wurde.
- Während einige Befragte meinten, die eng beieinander stehenden Kästchen seien von Vorteil, meinten andere, dadurch sieht es kompliziert aus.

- Und während einige die schräg gestellten Spaltenbezeichnungen gut fanden, fanden andere sie schwierig zu lesen.

Plus-Punkte für die "kurze" Version

- Die Gruppierungen nach Erscheinungsintervallen halfen den Befragten, sich auf eine Gruppe zu konzentrieren.
- Dadurch, dass es im Vergleich zur "erweiterten" Version weniger Kästchen gab, meinten einige, es würde besser aussehen und leichter auszufüllen sein.
- Es wurde als Vorteil empfunden, dass die numerischen Aussagen in die Frequenzskala integriert waren, da man dadurch die Anleitung nicht noch einmal lesen musste, um es verstehen zu können.

Auf der Negativ-Seite der "kurzen" Version

- Die (je nach Erscheinungsintervall) wechselnden Zeitperioden für das letzte Lesen verwirrte einige.
- Es gab einige Unsicherheiten, ob die Antwortvorgabe "nein" bei der ersten Frage ausgefüllt werden muss.

Wie wurden die Versionen ausgefüllt?

Die "erweiterte" Version wurde meistens zunächst vertikal (Fr. 1 - Lesen in den letzten 12 Monaten) ausgefüllt und dann horizontal für Frage 2 (Letztes Lesen) und Frage 3 (Frequenz) für die Titel, für die bei Frage 1 mit "ja" geantwortet wurde.

(Abb. Fragebogenseite)

Einige füllten für die Titel, für die in Frage 1 "ja" geantwortet wurde, sofort die Fragen 2 und 3 aus.

(Abb. Fragebogenseite)

In einigen Fällen wurde "nein" in Frage 1 nicht ausgefüllt, sondern nur das "ja". Einige füllten den Fragebogen Seite für Seite aus, einige nach Titelgruppen. Eine kleine Minderheit füllte die Fragen 1 bis 3 horizontal aus.

Die "kurze" Version wurde vertikal ausgefüllt. Einige gingen bei Frage 1 zunächst die Ja-Spalte durch und dann die Nein-Spalte - manchmal fehlten jedoch auch die Nein-Antworten - und dann Frage 2.

(Abb. Fragebogenseite)

Einige Befragten gingen die ja/nein-Antworten von Frage 1 gemeinsam durch und beantworteten dann die Frage 2.

(Abb. Fragebogenseite)

Es gab eine Tendenz die Frage 1 zum letzten Lesen als Filter für die Frage 2 für die Frequenz zu betrachten. Eine kleine Minorität füllte den Fragebogen horizontal aus. Im allgemeinen wurden die Fragen nacheinander für die Erscheinungsintervalle beantwortet.

Beurteilung der Fragebögen

Aus Forschungssicht gab es mehr Probleme mit der "kurzen" Version. Die "erweiterte" Version bietet gegenüber der "kurzen" Version eine Reihe von Vorteilen (diese sind ähnlich den Vorteilen, die die Befragten anführten, aber nicht dieselben):

- Der Hauptvorteil der "erweiterten" Version war, dass die Frage 1 (Lesen in den letzten 12 Monaten) ganz klar als Filter für die Fragen 2 und 3 fungierte. Dadurch, dass mit einem sehr weiten Zeithorizont begonnen wird, hilft es den Befragten sich auch an weit zurückliegende Lesevorgänge zu erinnern.
- Die detaillierte mehrstufige Zeitskala zum letzten Lesen diente als effektive Erinnerungshilfe. Verglichen mit der "kurzen" Version, bei der man nur "ja" oder "nein" für den jeweiligen Titel und die Lesezeit angeben kann.
- Die für alle Titel gleichermaßen geltende Zeitskala zum letzten Lesen wurde als konsistenter Bestandteil des Fragebogens empfunden. Das machte die "erweiterte" Version weniger anfällig für Fehler, da die Befragten die veränderten Erscheinungsintervalle nicht bemerkten, wie es bei der "kurzen" Version der Fall war.
- Und letztlich, die zusammengedrückten Kästchen brachten die Fragen besser zueinander und führten dazu, dass die Fragen 2 und 3 horizontal ausgefüllt wurden.

Der einzige beobachtete Nachteil der "erweiterten" Version bestand darin, dass die Frequenzskala nicht durchgängig in den Kopfspalten numerisch dargestellt wurde.

Es wurden jedoch mehr Probleme mit der "kurzen" Version beobachtet:

- Es bestand die Tendenz Frage 1 als Filter für Frage 2 zu betrachten. In einigen Fällen wurde die Reichweite dadurch überschätzt, dass Befragte sagen wollten, dass sie den Titel schon gelesen haben und um die Frage 2 auszufüllen. Auf der anderen Seite wurde die Frage 1 korrekt ausgefüllt, aber bei der Antwort "nein", wenn der Titel nicht im Erscheinungsintervall gelesen wurde, Frage 2 nicht ausgefüllt. Das führte dann zu einer Unterschätzung der kumulierten Reichweite. Der Trend ging jedoch zum ersten Beispiel, d.h. die Reichweite wurde eher überschätzt.
- Die Nein-Antwort bei Frage 1 wurde oft nicht ausgefüllt, da die Befragten meinten, wenn sie nicht "ja" angeben, heißt das automatisch "nein". (Dieses kann auch mit den weiter auseinanderliegenden Kästchen zu tun haben.)
- Die alleinige Vorgabe des Erscheinungsintervall-Zeitraums war keine effektive Erinnerungshilfe. Ganz im Gegensatz zu der Skala in der "erweiterten" Version, war es hier eine Frage von "im Erscheinungsintervall sein" oder "nicht" und führte dazu eher "ja" anzugeben.

- Die wechselnden Zeitintervalle der Titel verwirrten einige Befragte. Das war eine wichtige Beobachtung, da es zeigte, dass die Befragten auch bei regelmäßig gelesenen Titeln nicht wußten, in welchen Intervallen sie erscheinen.
- Und als letzter Punkt: Das großzügige Layout mit mehr Abstand zwischen den Kästchen führte zu einer vertikalen Beantwortung, die mehr Fehler produzierte.

Punkte für die zukünftige Forschung

Folgende Hauptpunkte mußten vor einem quantitativen Test berücksichtigt werden:

- Die Länge und das Layout der Anleitung, um maximale Klarheit zu erreichen.
- Das Layout der Kopfspaltenbeschriftung für mehr Klarheit und der Einbezug aller wichtigen Informationen.
- Einführung klarerer Angaben zur Reihenfolge der Fragen.
- Die Titelgruppierungen: Sollten sie nach Erscheinungsintervallen oder nach Gattungen vorgenommen werden?

Ein weiterer Punkt war, ob Mini-Titelkarten eingeführt werden sollten, um die Erinnerung zu stützen. Beide Versionen wurden von den Befragten akzeptiert und benötigten ähnliche Ausfüllzeiten. Die "erweiterte" Version war aber anscheinend weniger fehleranfällig. Es wurde aber trotzdem beschlossen auch die "kurze" Version mit wesentlichen Änderungen weiter zu testen. Es wurden zwei weitere Versionen getestet: Eine war die "erweiterte" Version mit minimalen Änderungen und die andere eine veränderte "kurze" Version. In beiden Versionen wurden die Anleitungen einfacher und kürzer und beinhalteten ein Beispiel.

Die Hauptänderung bei der "erweiterten" Version bestand darin, dass für die Frage 1 (Lesen in den letzten 12 Monaten) lediglich "ja" vorgegeben wurde und ein Pfeil zur Frage 2 führte und dann ein Pfeil von Frage 2 zu Frage 3, um ein horizontales Ausfüllen zu unterstützen. Bei der Frequenzskala wurden auch die numerischen Werte aufgenommen. Und am rechten Rand wurden die Titel-Logos abgebildet, während am linken Rand der maschinengeschriebene Titel stand.

(Abb. Fragebogenseite)

Die neue "kurze" Version unterschied sich substantiell von der getesteten Version. Es wurde ein Filter hinzugenommen, in dem nach dem Lesen in den letzten 12 Monaten gefragt wird (nur "ja" Antwort); die Reihenfolge der Fragen zum letzten Lesen und der Frequenz wurde umgekehrt, so dass die eine nicht mehr als Filter für die andere dienen konnte (erst Frequenzfrage, dann Frage zum letzten Lesen mit Vorgabe des Erscheinungsintervalls und ja/nein-Antwort). Auch hier wurden am rechten Rand die Titel-Logos aufgenommen.

(Abb. Fragebogenseite)

Weitere Tests

Die beiden veränderten Versionen wurden wieder qualitativ bei 22 Befragten getestet. Die Änderungen betrafen die Probleme, die im ersten Test beobachtet worden waren. Interessanterweise wurden die kürzeren, einfacheren Anleitungen positiv empfunden, das Beispiel jedoch als überflüssig betrachtet. Der lediglich positive Filter zum Lesen in den letzten 12 Monaten führte dazu, dass leere Seiten nicht ungewöhnlich waren. Das führte dazu, dass einige Befragte meinten, sie hätten den Fragebogen nicht richtig beantwortet. Ein Kästchen auf jeder Seite für "keinen Titel auf dieser Seite in den letzten 12 Monaten gelesen" würde da eventuell helfen.

Titelkarten (Logos)

In den früheren Tests waren keine Logos enthalten. Jetzt wurden sie als Check benutzt, wenn man sich bei einem Titel nicht sicher war. Nur eine Minderheit orientierte sich von vornherein an den Logos, was jedoch dazu führen konnte, dass sie den Fragebogen von rechts nach links ausfüllten. In der "kurzen" Version konnte das Verwirrung zwischen den Fragen zum letzten Lesen und zum Lesen in den letzten 12 Monaten stiften; mit der Tendenz, dass Frage 2 und Frage 1 gar nicht ausgefüllt wurden. Eine Minderheit fand die Logos überflüssig. Es wurde beschlossen, die Logos beizubehalten und die maschinengeschriebenen Namen auf der linken Seite auch.

Vergleichende Beurteilung der geänderten Versionen

Obwohl beide Versionen gegenüber dem ersten Test Verbesserungen brachten, wurde doch die "erweiterte" Version favorisiert:

- Die erweiterte Skala zum letzten Lesen schien mehr Optionen an die Erinnerung zu geben und die Befragten meinten, sie würden vollständigere Angaben über ihr Lesen machen.
- Die Fragenreihenfolge vom letzten Lesen zur Frequenz wurde als besser empfunden.
- Die Titelreihenfolge nach Gattungen war besser als nach Erscheinungsintervallen, da es weniger Titelverwechslungen gab.

Ergebnis des Tests

Es wurde beschlossen, die "erweiterte" Version einem quantitativen Test zu unterziehen. Es wurden dazu noch einige geringfügige Änderungen vorgenommen: das Beispiel wurde entfernt und das Kästchen "keinen der Titel auf dieser Seite in den letzten 12 Monaten gelesen" eingeführt.

(Abb. Fragebogenseite)

Was wurde gelernt?

Die wichtigste Lektion, die gelernt wurde war, dass selbst subtile kleine Aspekte des Fragebogendesigns Einfluß auf das Ausfüllen haben können: Design des Antwortschemas hinsichtlich des horizontalen oder vertikalen Ausfüllens und des Einflusses auf die Ergebnisse; die Fragenreihenfolge und die Klarheit der Anleitung.

Für die Forscher war es von großem Interesse, dass nicht notwendigerweise die "kurze" Version befragtenfreundlicher war. Qualitative Tests gaben wertvolle Hinweise auf die Gestaltung des Selbstausfüllers. Es bleibt jedoch einem quantitativen Test vorbehalten, die Ergebnisse der Leserschafts-Daten und ihrer Vergleichbarkeit mit der Standard-CAPI-NRS zu überprüfen.

Title confusion and physical proximity in a self administered questionnaire

Titel-Verwechslungen und physische Nähe in einem Selbstausfüller-Fragebogen

Virginia Cable, The Wall Street Journal/Valentine Appel, Consultant

Zusammenfassung

Bei der Nutzung eines Online-Fragebogens zum Selbstausfüllen zeigte sich, dass - entgegen der Theorie und entgegen anderer Forschungsergebnisse - wenn verwechslungsfähige Titel zusammengefaßt gezeigt werden, höhere Reichweiten erreicht werden, als wenn die Titel den Befragten einzeln in zufälliger Reihenfolge präsentiert werden. Mit der Untersuchung wurde außerdem herausgefunden, dass, wenn nur wenige Titel abgefragt werden, die Reichweiten höher sind, als wenn viele Titel abgefragt werden, und dass die zuerst abgefragten Titel höhere Reichweiten aufweisen als die zuletzt abgefragten.

Hintergrund und Zielsetzung

Seit einiger Zeit versucht das Wall Street Journal die Mediaforscher davon zu überzeugen, bei Untersuchungen die Tageszeitungen und Zeitschriften vergleichen, für beide Gruppen die gleichen Abfrage-Prozeduren einzusetzen. Mediamark Research Inc. (MRI), das führende Leserschaftsforschungs-Unternehmen in den USA, nutzt zwei sehr unterschiedliche Verfahren. Bei Zeitschriften wird ein 6-Monats-Filter eingesetzt und die Befragten, die den Filter passieren, werden sofort gefragt, ob sie irgendeine Ausgabe im letzten Erscheinungsintervall gelesen haben - letzter Monat, letzte Woche oder was auch immer. Tageszeitungen werden separat vor den Zeitschriften ganz anders abgefragt. Statt eines 6-Monats-Filters wird ein 7-Tage-Filter eingesetzt. Diejenigen, die den Filter passieren, werden gefragt, ob sie in der letzten Zeit irgendeine Ausgabe gelesen haben und bei Angabe von "gestern" werden sie als Leser qualifiziert. Dieses ist die Standard-Abfrage, die von der U.S. Advertising Research Foundation ARF in den Richtlinien für Tageszeitungs-Forschung empfohlen wird.

Als das Wall Street Journal die 1994er-Untersuchung veröffentlichte, in der dargestellt wurde, dass ein 7-Tage-Filter geringere Reichweiten erbringt, als ein 6-Monats-Filter, ändert MRI die eigenen Erhebungsprozeduren erheblich. Das Ergebnis wurde vom Wall Street Journal auf dem 1997er Vancouver-Symposium vorgestellt: Sowohl für das Wall Street Journal als auch für USA Today hatten sich die LpE-Werte verdreifacht - und MRI veröffentlichte sie nicht. Spätere Untersuchungen von Simmons und von Audits and Survey ließen darauf schließen, dass die Erhöhung der LpE-Werte bei MRI darauf zurückzuführen waren, dass das Wall Street Journal, USA Today und The New York Times mit einer Vielzahl von Zeitschriften-Logos gezeigt wurden.

Das Wall Street Journal unternahm den Versuch dieses Ergebnis besser zu verstehen und führte eigene Untersuchungen durch, um die Auswirkungen der folgenden drei Fragebogen-Modifizierungen zu überprüfen:

1. Separierung einer bestimmten Gruppe von Titeln und Messung dieser Gruppe außerhalb der Hauptgruppe der Titel.
2. Untersuchung des Positionierungseffekts der separierten Titel-Gruppe, Messung einmal vor und einmal nach der Abfrage der Haupt-Titel-Gruppe.
3. Gruppierung der Titel nach Namen oder nach inhaltlichen Ähnlichkeiten und nicht in einer zufälligen Reihenfolge.

Durchführung

Die Untersuchung wurde Online in einer Stichprobe des Intelliquest Technology Panels durchgeführt. Das Panel beinhaltet Entscheidungsträger im Computer-Bereich, die bereits schon einmal befragt wurden und weiteren Online-Befragungen zustimmten. Auswahlkriterien der Stichprobe waren: 25 Jahre alt und älter sowie berufstätig. Ungefähr die Hälfte der in Frage kommenden Zielpersonen hatten ein Haushaltseinkommen von \$ 75.000 und mehr. Zur Verfügung standen 10.000 potentiell zu befragende Zielpersonen, die per e-mail kontaktet und zur Teilnahme an der Untersuchung eingeladen wurden. Ein Paßwort führte zu einer von neun Fragebogenversionen. Jede Fragebogenversion beinhaltete die Logos derselben 90 Titel. Der Durchführungszeitraum war Anfang August 2001. Ausgewertet werden konnten 4.146 Interviews.

Die Untersuchung erfaßte die Leserschaftsdaten für die 90 Titel mit einem 6-Monats-Filter, an den sich die Lesefrequenz-Frage anschloss. Die 90 Titel wurden in zwei Gruppen abgefragt: einmal sechs Titel und dann die 84 anderen Titel. Alle Titel wurden mit schwarz-weiß Logos abgebildet, jeweils sechs pro Seite, d.h. insgesamt 15 Seiten. Die Entscheidung von sechs Titeln pro Seite basierte darauf, was noch gut lesbar war.

(Abb. einer Computer-Bildschirm-Seite mit sechs Titel-Logos)

Die Reihenfolge der Abfrage in den neun Fragebogenversionen:

Reihenfolge	A n=650	B n=467	C n=466	D n=455	E n=455	F n=455	G n=451	H n=456	I n=491
Erst-Vorlage	6 TZ	6 TZ	6 W-ZS	6 W-ZS	84 R keine TZ	84 G keine TZ	84 R keine W-ZS	84 G keine W-ZS	90 R
Zweit-Vorlage	84 R keine TZ	84 G keine TZ	84 R keine W-ZS	84 G keine W-ZS	6 TZ	6 TZ	6 W-ZS	6 W-ZS	

Erläuterung: 6 TZ = 6 Tageszeitungen: The New York Times, The Wall Street Journal, USA Today, The Financial Times, The Los Angeles Times, The Washington Post. 6 W-ZS = 6 Wirtschafts-Zeitschriften: Business Week, Forbes, Fortune, Inc., Entrepreneur, The Economist. R = Random: Die Titelpräsentation erfolgte in zufälliger Reihenfolge; jeweils 6 pro Seite. G = Gruppirt: Die Titel wurden nach inhaltlichen oder Namens-Ähnlichkeiten zu

jeweils sechs Titeln zusammengefasst und präsentiert. Jeder Sechser-Block wurde in zufälliger Reihenfolge gezeigt und auch die Reihenfolge innerhalb des Blocks erfolgte zufällig. Die Möglichkeit der Random-Anordnung war mit ein Grund für die Wahl einer Online-Befragung. So waren nicht einmal zwei Fragebogen exakt identisch.

Ablauf am Beispiel von Version A: Zunächst wurden die sechs Tageszeitungen gezeigt und für die Titel, die den Filter passierten wurde danach die Frequenz-Frage gestellt. Dann wurden die 84 anderen Titel-Logos in zufälliger Reihenfolge jeweils zu sechs Titeln zusammengefasst gezeigt. Nachdem alle 84 Titel die Filterfrage durchlaufen hatten, wurde für die Titel, die den Filter passierten, die Frequenz-Frage gestellt. Eine Ausnahme in den Fragebogenversionen war die Version I, in der alle 90 Titel per zufälliger Reihenfolge abgefragt wurden. Allerdings waren hier keine Ergebnisunterschiede zu den Versionen festzustellen, bei denen nur 84 Titel per Zufalls-Reihenfolge abgefragt wurden, so dass in den nachfolgenden Ergebnisdarstellungen diese neunte Fragebogenversion nicht berücksichtigt wurde.

Filter-Ergebnisse

1. Separierung von sechs Titeln

Zunächst wurde die Auswirkung der Trennung der Tageszeitungen und der Wirtschafts-Zeitschriften begutachtet. Die relevanten Daten sind in der nachfolgenden Tabelle in der Mitte zu finden.

Tageszeitungen

	Getrennte Abfrage			Nichtgetrennte Abfrage		
	Erster Block	Letzter Block	Gesamt	Gesamt nicht		
			Separiert	Separiert	Gruppirt	Random
Basis	(916)	(922)	(1.838)	(1.817)	(907)	(910)
	%	%	%	%	%	%
Irgendeine	80,0*	75,3	77,6*	72,8	77,8*	67,8
Wall St. Journal	46,5*	40,0	43,2*	40,6	44,5*	36,6
USA Today	64,3*	59,0	61,6*	58,7	62,8*	54,6
New York Times	27,8	25,3	26,6*	24,0	24,0	24,0
Washington Post	14,1	11,8	13,0	12,7	13,1	12,3
LA Times	8,4	9,4	8,9	8,6	8,7	8,6
Financial Times	5,3	3,6	4,4	3,4	4,3	2,4

* signifikant verschieden (P<0.5 zweiseitig) gegenüber der nächsten Spalte rechts

Bei der Trennung der sechs Tageszeitungen von den restlichen 84 Zeitschriften ergeben sich 77,6%, die mindestens einen der sechs Tageszeitungs-Titel durch den Filter bringen, gegenüber 72,8%, wenn die Tageszeitungen nicht getrennt abgefragt wurden. Die Netto-Differenz von 7% ist auf dem .05 Niveau statistisch signifikant. Ebenso die Er-

gebnisse für Wall Street Journal, USA Today und The New York Times.

2. Positionierung

Hier wurden die Ergebnisse zwischen den Gruppen verglichen, die zuerst die Tageszeitungen oder die Wirtschafts-Zeitschriften beantwortet haben und die zuerst die restlichen Titel beantwortet haben (Spalten mit der Überschrift Erster Block, Letzter Block). Auch die hier festgestellte Netto-Differenz von 6% zugunsten der Abfrage der Tageszeitungen an erster Stelle ist statistisch signifikant. Ebenfalls signifikant sind die Differenzen beim Wall Street Journal und bei USA Today. Ein ausgeprägteres Ergebnis wurde bei den Wirtschafts-Zeitschriften beobachtet. (siehe nachfolgende Tabelle)

3. Gruppierte Titel-Vorlage vs. Zufalls-Reihenfolge

Die Ergebnisse stehen in den Spalten ganz rechts. Für die Tageszeitungen ergibt sich eine 15%ige statistisch signifikante Differenz zugunsten der gruppierten Vorlage (77,8% vs. 67,8%) mit Vorteilen insbesondere für das Wall Street Journal und USA Today bei der gruppierten Vorlage. Wiederum sind die Ergebnisse für die Wirtschafts-Zeitschriften entsprechend.

Wirtschafts-Zeitschriften

	Getrennte Abfrage			Nichtgetrennte Abfrage		
	Erster Block	Letzter Block	Gesamt Separiert	Gesamt nicht Separiert	Gruppiert	Random
Basis	(906) %	(911) %	(1.817) %	(1.838) %	(921) %	(917) %
Irgendeine	57,2*	41,8	50,0*	46,1	48,4*	43,8
Business Week	37,5*	20,5	29,0*	24,2	26,4*	22,1
Forbes	23,6*	18,1	20,8	20,6	22,7*	18,6
Fortune	26,5*	15,7	21,1*	17,7	18,1	17,3
Inc.	18,1*	13,2	15,6	13,1	13,7	12,5
Entrepreneur	12,6*	8,9	10,8	9,5	9,4	9,6
The Economist	8,3*	6,1	7,2	7,8	8,3	7,3

* signifikant verschieden ($P < 0.5$ zweiseitig) gegenüber der nächsten Spalte rechts

Die anderen 78 Zeitschriften

Für jeden Befragten wurde die Anzahl der durch den Filter gekommenen Zeitschriften gezählt um somit den Durchschnitt zu berechnen. Dieses erfolgte für die Fragebogen-Versionen in denen die Reihenfolge zufällig war und für die Versionen, in denen die Titel gruppiert vorgelegt wurden sowie für die Versionen, in denen die Titel zuerst abgefragt wurden und in denen sie an zweiter Stelle standen.

Zwischen der Abfrage an erster oder zweiter Stelle gab es keine signifikante Abweichung. Wenn sie jedoch gruppiert vorgelegt wurden, erhöhte sich der Durchschnitt von 11,6 Titel auf 12,6 Titel. Dieser Unterschied ist statistisch signifikant.

(Tab. mit durchschnittlichen Filterpassagen)

Reichweiten-Ergebnisse

Für jeden Titel wurden die durchschnittlichen Ausgaben-Reichweiten errechnet, in dem die Frequenz in Dezimalwerte umgerechnet (z.B. $3/4 = 0.75$), addiert und durch die Zahl der Befragten geteilt wurden. Nachfolgend die Ergebnisse für die Tageszeitungen. Statistisch signifikant sind nur die Unterschiede zwischen dem Wall Street Journal und der New York Times beim Vergleich zwischen getrennter Vorlage und nicht-getrennter Vorlage.

Tageszeitungs-Reichweiten

	Getrennte Abfrage			Nichtgetrennte Abfrage		
	Erster Block	Letzter Block	Gesamt Separiert	Gesamt nicht Separiert	Gruppiert	Random
Basis	(906)	(911)	(1.817)	(1.838)	(921)	(917)
	%	%	%	%	%	%
Wall St. Journal	18,2	18,4	18,3*	16,5	17,7	15,3
USA Today	19,7	19,3	19,5	19,0	20,0	18,1
New York Times	10,9	11,4	11,2*	9,6	9,1	10,0
Washington Post	4,8	5,2	5,0	5,4	5,3	5,4
LA Times	3,5	4,2	3,8	3,4	3,2	3,7
Financial Times	1,9	1,1	1,5	1,0	1,2	0,9

* signifikant verschieden ($P < 0.5$ zweiseitig) gegenüber der nächsten Spalte rechts

Die Ergebnisse für die Wirtschafts-Zeitschriften

Business Week hat eine signifikant höhere Reichweite, wenn die Wirtschafts-Zeitschriften getrennt von den anderen Zeitschriften abgefragt werden. Business Week und Fortune haben signifikant höhere Ergebnisse, wenn sie zuerst abgefragt werden. Forbes hat eine signifikant höhere Reichweite, wenn der Titel gruppiert abgefragt wird.

Wirtschafts-Zeitschriften

	Getrennte Abfrage			Nichtgetrennte Abfrage		
	Erster Block	Letzter Block	Gesamt	Gesamt nicht		
			Separiert	Separiert	Gruppiert	Random
Basis	(916)	(922)	(1.838)	(1.817)	(907)	(910)
	%	%	%	%	%	%
Business Week	16,1*	9,7	12,9*	10,0	11,0	9,0
Forbes	9,3	7,9	8,6	8,1	9,1*	7,1
Fortune	9,9*	6,5	8,2	7,0	7,2	6,7
Inc.	7,4	5,7	6,6	5,1	5,6	4,6
Entrepreneur	5,0	3,9	4,4	3,4	3,4	3,3
The Economist	3,1	2,4	2,8	3,2	3,7	2,8

* signifikant verschieden ($P < 0.5$ zweiseitig) gegenüber der nächsten Spalte rechts

Die nächste Tabelle zeigt die durchschnittliche Reichweite für die 78 anderen Zeitschriften. Bei gruppierter Vorlage sind die Ergebnisse signifikant höher als bei zufälliger Vorlage.

(Tab. mit Durchschnitts-Reichweite 78 Titel inkl. Standard-Abweichung)

Zusammenfassung

Bei Abfrage einer kleinen Gruppe von Titeln bevor eine größere Anzahl von Titeln abgefragt wird (so wie MRI vorgeht), ergeben sich höhere Reichweiten. Wenn die Titel gruppiert vorgelegt werden und nicht in zufälliger Reihenfolge (wie MRI vorgeht), sind die Reichweiten ebenfalls höher.

Verblüfft hat uns, dass sich bei der Gruppierung der Titel auch die Reichweiten erhöht haben und sie nicht, wie vermutet wurde, zurückgehen. Ein Rückgang der Reichweiten wurde in anderen Untersuchungen festgestellt. Zuerst dachten wir daran, aufgrund dieses verblüffenden Ergebnisses den Vortrag zurückzuziehen, aber dann entschieden wir, nachdem wir die Ergebnisse nochmals überprüft haben, es nicht zu tun. Immerhin kann es manchmal wichtiger sein, ein Ergebnis nicht zu reproduzieren als es zu tun. Warum unser Ergebnis nicht die Ergebnisse anderer bestätigt, hat viele Gründe. Wir sind ganz sicher, dass es kein statistischer Zufallstreffer war. Vielleicht liegt es daran, dass die Erhebung Online durchgeführt wurde, und dass jeweils sechs Logos zusammen auf einer Seite gezeigt wurden.

Erhebung und Minimierung von Positionierungs- und Titelverwechslungs-Effekten in der Leserschaftsforschung

*Marco Vriens, Barbara Graham, Eric Melton, Tony Incalcaterra, Millward Brown IntelliQuest
Valentine Appel, Consultant*

Bei der Erhebung von Leserschaften mit Selbstausfüllern in Form von Papierfragebögen wurden zwei Möglichkeiten festgestellt, durch die die Ergebnisse verzerrt werden können: Positionierungseffekte und Titelverwechslungen. Dieser Vortrag behandelt den Einfluß von zwei Arten von Positionierungseffekten sowie den Einfluß von Titelverwechslungen auf die Filterdurchgänge und die Lesefrequenz. Außerdem wird aufgezeigt, wie die Verzerrungen minimiert werden können, wenn ein Web-Fragebogen genutzt wird.

Potentielle Verzerrungen in der Leserschaftsforschung

Selbstausfüller werden im allgemeinen zur Bestimmung der Höhe der Reichweiten genutzt. Studien haben gezeigt, dass die Position eines Titels den Filterdurchgang und die Reichweiten beeinflussen kann. Wenn die Titel in einer festen Reihenfolge vorgegeben werden, kann die Position an der der Titel abgefragt wird, Einfluß auf die Filterfrage und die Frequenzfrage haben, insbesondere wenn alle Befragten Fragebögen mit der gleichen Reihenfolge ausfüllen. Das nennen wir den reinen Positionierungs-Bias. Wenn ein solcher Bias besteht, kann er durch die zufällige Vorgabe der Titel minimiert werden. Zweitens wurde herausgefunden, dass es einen weiteren Bias gibt, wenn direkt im Anschluß an die Filterfrage weitere Fragen zum Lesen gestellt werden. Hierbei lernen die Befragten sehr schnell, dass sie durch die Angabe, dass sie den Titel nicht lesen, die Länge des Interviews beeinflussen können. Dementsprechend werden sie ihre Antworten anpassen, indem sie z.B. nur regelmäßig gelesene Titel angeben. Das nennen wir den Positionierungs-Bias bei Filterfragen. Der Effekt ist, dass Titel die später oder am Ende des Fragebogens abgefragt werden, unterschätzt werden. Wenn die Filterfrage und die Frequenzfrage getrennt werden und die Titel in zufälliger Reihenfolge abgefragt werden, können diese o.a. Positionierungseffekte vermieden werden.

Sind jedoch die Filterfrage und die Frequenzfrage getrennt und die Titel-Reihenfolge wird zufällig gewählt, kann ein weiterer Bias auftreten. Untersuchungen haben gezeigt, dass bei Titeln, die ähnlich klingen, wie z.B. PC Magazine und PC World, jeder Titel höhere Reichweiten erreicht, wenn sie voneinander getrennt an unterschiedlichen Positionen abgefragt werden, als wenn sie zusammen in einer Gruppe abgefragt werden. Dieses ist der Titelverwechslungs-Bias. Dieser Bias kann durch entsprechende Gruppierungen von ähnlichen Titeln vermieden werden, da die Befragten dann beide Titel gemeinsam sehen.

Frühere Forschungen zum Positionierungs-Bias

(Es werden einige Untersuchungen angeführt, die sich mit dem Thema beschäftigt haben und den Positionierungs-Bias nachgewiesen haben. Dabei wurde z.B. beobachtet, dass i.d.R. der zuerst abgefragte Titel Vorteile gegenüber dem an zweiter Stelle abgefragten Titel hat. - Außerdem wurden auch Positionierungseffekte bei der Stellung der verschiede-

nen Erscheinungsintervalle nachgewiesen. Wenn monatlich erscheinende Titel als erstes abgefragt werden, erzielen sie eine höhere Reichweite als wenn sie am Schluß abgefragt werden. usw. Weitere frühere Forschungen haben nachgewiesen, dass Folgefragen zu einem Titel die Reichweiten minimieren.)

Frühere Forschungen zum Titelverwechslungs-Bias

(Auch in diesem Kapitel werden Untersuchungen zitiert, die sich bereits mit den Titelverwechslungen beschäftigt haben. U.a. wird auch auf ein Experiment von 1988 verwiesen, in dem ein Titel abgefragt wurde, den es aber gar nicht gab, der lediglich einen fast gleichen Namen eines existierenden Titel hatte. Drei Viertel der Leser des echten Titels gaben an, auch den anderen Titel zu lesen.)

In dem Vortrag sollen die diskutierten Biases in einer web-basierten Untersuchung mit experimentellen Design dargestellt werden. Es wurde herausgefunden, dass die Anzahl der Titel, die den Filter passieren, bei web-basierten Untersuchungen höher ist, als bei Paper-pencil-Untersuchungen.

Außerdem soll die Größenordnung des jeweiligen Bias eingeschätzt werden. Insbesondere dann, wenn mehr als ein Bias auftritt, können wir die relative Wichtigkeit einschätzen. Und schließlich wollen wir die Möglichkeiten einer web-basierten Abfrage testen und aufzeigen.

Basierend auf den früheren Forschungsergebnissen erwarten wir:

1. dass in dem Moment, wo Zusatzfragen direkt nach der Filterfrage gestellt werden, die Anzahl der Titel, die den Filter passieren geringer ist, als wenn keine Zusatzfragen direkt danach gestellt werden.
2. dass in Untersuchungen, bei denen die Titel nach einer festen Reihenfolge abgefragt werden, die Titel, die mehr am Ende stehen, geringere Filterpassagen aufweisen als wenn die Reihenfolge der Titel zufällig erfolgt.
3. dass ein Fragebogen, in dem ähnliche Titel nicht zusammen abgefragt werden, wegen der Titelverwechslungen entweder zu höheren Filterpassagen oder zu geringeren Filterpassagen führt, weil dadurch geringere Erinnerungsimpulse gegeben werden, als wenn die Titel zusammen in einer Gruppe abgefragt werden.
4. Der Effekt, der für alle Verzerrungen insgesamt auftritt, ist schwer vorherzusagen, da beide Biases (Positionierungs- und Titelverwechslungs-/Erinnerungs-Bias) sich gegenseitig aufheben können.

Design des Experiments

Selbstaussfüller und Paper-pencil-Untersuchungen erlauben uns nicht, alle Arten der Verzerrungen simultan zu überprüfen und zu reduzieren. Online-Untersuchungen können jedoch entsprechend entwickelt werden. Es wurden fünf Versionen für fünf gleiche Stichproben mit ca. 600 Befragten entwickelt. Es wurden 50 Zeitschriften einbezogen.

(Tab. mit Übersicht über 5 Fragebogen-Versionen)

Die 50 Titel wurden je nach Titelähnlichkeit oder nach inhaltlicher Ähnlichkeit in zehn Gattungen eingeteilt.

In der 1. Experiment-Version (Maximaler Bias) wurden ausdrücklich alle Biases einbezogen, damit ergibt das den maximalen Bias.

- Ähnliche Titel wurden nicht zusammengefaßt.
- Nach der Filterfrage werden für jeden Titel sofort die Zusatzfragen gestellt. (Lesefrequenz, wie ausführlich Werbung für Computer-Hard- und -Software, Internet usw. gelesen werden und wie viele Seiten insgesamt gelesen werden)
- Die Titel werden in einer festen Reihenfolge vorgegeben.

In der 2. Experiment-Version (Zwei Positionierungseffekte) werden die beiden Positionierungseffekte zugelassen, aber nicht der Titelverwechslungseffekt.

- Ähnliche Titel werden zusammen abgefragt.
- Nach der Filterfrage werden für jeden Titel sofort die Zusatzfragen gestellt.
- Die Titel werden in einer festen Reihenfolge nach den o.e. Gattungen vorgegeben.

in der 3. Experiment-Version (Reiner Positionierungseffekt) wurde nur der reine Positionierungseffekt zugelassen.

- Ähnliche Titel werden zusammen abgefragt.
- Die Filterfrage, die Frequenzfrage und die anderen Zusatzfragen werden getrennt abgefragt.
- Die Titel werden in einer festen Reihenfolge nach den o.e. Gattungen vorgegeben.

In der 4. Experiment-Version (Titelverwechslungseffekt) wurden nur die Titelverwechslungen zugelassen.

- Die ähnlichen Titel werden nicht zusammengefaßt.
- Die Filterfrage ist getrennt von den Zusatzfragen.
- Die Titel-Reihenfolge wird für jeden Befragten in zufälliger Reihenfolge vorgegeben.

Die 5. Experiment-Version (Minimum Bias) eliminiert alle Effekte.

- Ähnliche Titel werden zusammengefasst abgefragt.
- Die Filterfrage ist getrennt von den Zusatzfragen.
- Die Gattungen werden in zufälliger Reihenfolge vorgegeben.

Der Web-Fragebogen

Die 50 Titel wurden aus 91 Titeln einer anderen Untersuchung (CIMS Vers. 8.0) ausgewählt. Titel mit geringer Auflage wurden dabei nicht berücksichtigt. Im Anhang werden die gebildeten Gattungen vorgestellt, sowie die Reihenfolge bei ungruppiertem Abfrage. Außerdem wird im Anhang dargestellt, wie es aussieht, wenn die Filterfrage separat gestellt wird und wenn gleich im Anschluss daran die Zusatzfragen gestellt werden. Den Befragten wurden jeweils sieben Titel pro Bildschirmseite gezeigt. Um zur nächsten Seite zu gelangen mussten alle Fragen einer Seite beantwortet werden. Im Anschluss an die Leserschaftsfragen wurden einige Fragen zu der Entscheidungskompetenz beim Kauf von technischen Produkten gestellt.

Stichprobe

Vom IntelliQuest-Technik-Panel mit ca. 27.000 Teilnehmern wurde eine Stichprobe von 7.500 Adressen gezogen. Dabei wurden elf Segmente gebildet, die eine ähnliche Untersuchung abbilden (Dun & Bradstreet Office Intensive business sample). Jeder Befragte wurde zufällig einer der fünf Experiment-Versionen zugeordnet. Für die Teilnahme wurde ein Incentive in der Größenordnung von ca. \$ 7,50 angeboten. Nach dem ersten Versand des Fragebogens wurde 48 Stunden danach eine Erinnerung verschickt. Insgesamt wurden in den einzelnen Versionen 602, 606, 601, 600 und 604 Interviews durchgeführt.

Die Bias-Maße

Der Unterschied zwischen Version 5 und 2 zeigt, wie groß der gesamte Positionierungseffekt ist. Der Unterschied zwischen Version 5 und 3 zeigt, wie groß der reine Positionierungseffekt ist, d.h. wenn die Titel-Reihenfolge fest vorgegeben wird. Der Unterschied zwischen Version 5 und 4 zeigt, wie groß der Titelverwechslungs-Effekt ist. Der Unterschied zwischen Version 2 und 3 zeigt, wie groß der Effekt ist, wenn nach der Filterfrage sofort die Zusatzfragen gestellt werden. Der Unterschied zwischen Version 5 und 1 zeigt, wie groß der Effekt ist, wenn alle Biases gleichzeitig auftreten.

Ergebnisse

Durchschnittliche Anzahl der Titel, die den Filter passierten nach Experiment-Versionen

	Vers. 1 Maximaler Bias	Vers. 2 Zwei Posi- tionierungs- effekte	Vers. 3 Reiner Posi- tionierungs- effekt	Vers. 4 Titel verwechs- lungseffekt	Vers. 5 Minimaler Bias
Basis	602	606	601	600	604
Ø Anzahl Titel nach Filterpassage	6,69	7,01	10,04	9,01	9,26
Standardabweichung	5.69	5.59	7.03	6.96	6.70
Ø Anzahl der gelesenen Titel*	3,64	3,78	4,80	4,44	4,69
Standardabweichung	6.69	5.59	7.03	6.96	6.70

* gebildet aus Wahrscheinlichkeiten zur Frequenz 4 von 4 = 1, 3 von 4 = 0,75, 2 von 4 = 0,5, 1 von 4 = 0,25, weniger als 1 von 4 = 0,10

Es werden signifikante Unterschiede zwischen den Versionen 2 und 5 erwartet, weil hier die Differenz in den Filterpassagen gemessen wurde. Außerdem beinhaltet die Version 2 den reinen Positionierungseffekt. Wir bemerken jedoch, dass der reine Positionierungseffekt nicht durch die Ergebnisse von Version 3 und 5 aufgedeckt werden kann, weil es keinen Grund gibt, warum die Anzahl der durchschnittlichen Filterpassagen durch die zufällige Titel-Vorlage beeinflusst werden soll. Um den reinen Positionierungseffekt herauszubekommen müssen Einzel-Titel oder Titelgruppen betrachtet werden. Wie erwartet,

ist die Differenz zwischen Version 2 und 5 statistisch signifikant und sehr bedeutsam. Wenn der Einfluß des Abfragens von Zusatzfragen direkt nach der Filterfrage wegfällt, geben die Befragten mehr gelesene Titel an.

Die Ergebnisse zur Titelverwechslung sind weniger eindeutig. Die Ergebnisse von Version 4 und 5 sind nicht statistisch signifikant. Es wurde jedoch schon vorher erwähnt, dass der Titelverwechslungseffekt die Filterpassagen sowohl größer als auch kleiner werden lassen kann.

Bei Betrachtung von Titelgruppen kann dieser Effekt eventuell eher beleuchtet werden, genauso wie der reine Positionierungseffekt.

(Tab. mit Ergebnissen der Titelgruppen)

Wenn man hier wiederum die Versionen 2 und 5 betrachtet, wird ein deutlicher Effekt sichtbar, wenn die Zusatzfragen im Anschluß an die Filter gestellt werden:

Differenz für Zeitschriften Gruppe A: $1,17 - 1,14 = 0,03$ Verringerung um 2%

Differenz für Zeitschriften Gruppe F: $0,72 - 0,54 = 0,18$ Verringerung um 25%

Differenz für Zeitschriften Gruppe J: $0,67 - 0,39 = 0,28$ Verringerung um 42%

Dieser Positionierungseffekt wird ganz besonders für die Zeitschriften deutlich, die am Ende des Fragebogens abgefragt werden. Die Ergebnisse von Version 3 und 5 zeigen, dass auch ein reiner Positionierungseffekt existiert. Die meisten Ergebnisse für die Filterpassagen sind in Version 3 höher als in Version 5: Das wurde nicht erwartet.

Allerdings haben die Ergebnisse nach Titelgruppen keinen deutlichen Hinweis auf Titelverwechslungseffekte gebracht. In der nachfolgenden Tabelle wurden die Ergebnisse nochmals nach Einzeltiteln (mit Dummy-Namen) analysiert. Dabei treten einige Unterschiede zwischen Version 4 und 5 auf. Es bleibt jedoch, dass die Filterpassagen bei beiden Versionen keine deutlichen Unterschiede aufweisen. Vielleicht liegt es daran, dass die Befragten der Stichprobe eher beide Titel lesen, so dass sie sie nicht verwechseln können, da sie ja Leser beider Titel sind und wenn man sie zusammen vorgibt, wird eventuell eher die Erinnerung verbessert.

(Tab. mit einigen (Dummy-)Einzeltiteln)

Zusammenfassung

Verzerrungen in Leserschafts-Untersuchungen bleiben ein Thema. Man kann jedoch mit web-basierten Fragebögen die Effekte der drei Biases, die in diesem Vortrag besprochen wurden, minimieren. Die Experimente zeigten nicht nur die Vorteile eines Online-Interviews, sondern auch das relative Maß der Effekte, wenn ein Online-Interview nicht möglich ist.

Sehr stark ist der Effekt, wenn im Anschluß an die Filterfrage sofort für jeden Titel Zusatzfragen gestellt werden. Titel, die erst gegen Ende der Befragung abgefragt werden, können in der Reichweite bis zu 40% unterschätzt werden.

Unsere Experiment-Ergebnisse konnten die Erwartungen zum reinen Positionierungseffekt nicht unterstützen. Das liegt zum einen vielleicht daran, dass bei der Filterfrage nur "ja/nein" zu antworten ist und zum anderen, dass nur 50 größere Titel abgefragt wurden.

Die Ergebnisse zur Titelverwechslung sind ebenfalls konträr zu dem, was wir erwartet haben. In der Literatur ist darüber auch nur wenig Experimentelles zu finden. Wir sind jedoch der Meinung, dass bei einer Abfrage die Gruppierung von ähnlichen Titeln versus einer Nicht-Gruppierung Effekte hat. Gruppierungen können nicht nur Titelverwechslungen eliminieren, sondern auch bei der Erinnerung helfen.

Was die vorliegende Untersuchung überhaupt nicht behandelt, sind andere Dinge im Zusammenhang mit Web-Studien, wie Formatierung, Anzahl der Pixel, Qualität der Seiten usw. Außerdem wurde auch die Präsentation der Logos (farbig oder schwarz-weiss) oder von Titelseiten nicht behandelt. Zusätzlich zu den Verzerrungen durch verschiedene Effekte sollten auch solche Design Elemente überprüft werden. Dazu gehören auch: die Länge des Media-Teils und die Position des Titels darin, die Anzahl der abzufragenden Titel und die Gesamtlänge der Untersuchung. Zukünftige Forschungen sollten diese Dinge berücksichtigen.

Es wird oft gesagt: Je mehr wir wissen, desto mehr müssen wir wissen. Basierend auf unseren Ergebnissen sollten folgende Dinge bei web-basierten Leserschafts-Untersuchungen berücksichtigt werden:

1. Bei web-basierten Untersuchungen ist es für die Befragten schwieriger, die Länge des Media-Teils der Untersuchung abschätzen zu können und deshalb ist die Anzahl der Titel, die den Filter passieren, höher.
2. Ein Web-Fragebogen erlaubt es, die Filterfrage von den anderen (Zusatz-)Fragen zu trennen und die Reihenfolge der Titel zufällig vorzunehmen, wobei ähnliche Titel zusammen gezeigt werden können. Das reduziert potentielle negative Effekte von drei Biases: von dem Bias, wenn im Anschluß an die Filterfrage sofort für jeden Titel Zusatzfragen gestellt werden, den reinen Positionierungsbias und den Titelverwechslungsbias.

(Anhang mit Titelauswahl und -gruppierung, Fragebogenseiten)

Searching for best CASI practices: Experiments for a new data collection method

Auf der Suche nach der besten CASI-Praxis: Experimente für eine neue Datenerhebungs-Methode

Alke Bassler & Costa Tchaoussoglou, SUMMO

Synopse

Dieser Vortrag beinhaltet die Fortsetzung eines Vortrags zu einem Experiment mit Computer Assisted Self Interviewing (CASI) (= Selbstausfüller, bei dem die Testperson den Fragebogen am Computer selbst beantworten) von Bart Soels und Costa Tchaoussoglou, gehalten in Florenz 1999. In den letzten drei Jahren wurde die CASI-Methode mit positiven Ergebnissen für die holländische nationale Leser-Analyse getestet.

Die Mehrheit der Testpersonen ist sehr gut in der Lage, den Fragebogen am Computer selbst auszufüllen, auch wenn sie keine Erfahrung im Umgang mit Computern haben. In den wenigen Fällen, in denen der Interviewer das Ausfüllen übernehmen musste, hatte das keine Auswirkungen auf die Reichweiten. Unabhängig davon, dass es keinen Interviewer-Bias gibt - und das ist der Hauptvorteil gegenüber CATI-Interviews - konnten aber Ergebnis-Unterschiede zwischen den Instituten festgestellt werden. Das ist überraschend, da Instituts-Unterschiede normalerweise immer mit den unterschiedlichen Arbeitsweisen der Interviewer begründet werden.

CASI kann in face-to-face Interviews genauso genutzt werden, wie in Computer-Panels. Die Ergebnis-Unterschiede zwischen den beiden Untersuchungsformen resultieren aus den unterschiedlichen Stichproben. Wenn die Befragten vergleichbar sind, sind die Ergebnis-Unterschiede sehr klein oder verschwinden ganz. Eine repräsentative Stichprobe ist also so wichtig wie immer. Bei der Rekrutierung von Befragten für ein face-to-face Interview ist ein Telefon-Screening einer face-to-face Rekrutierung überlegen. Neben dem fehlenden Interviewer-Bias wurde durch CASI und die größere Gewissenhaftigkeit und Sorgfalt beim Selbstausfüllen und dem fehlenden Zeitdruck, die Qualität der erhobenen Daten verbessert. Dieser Effekt wird vermindert, wenn die (voraussichtliche) Antwort, um Zeit zu gewinnen, durch Weglassen gegeben wird, so dass die Befragten die Antwort nur bei Differenzen ändern müssen. Es ist ratsam, dass die Befragten immer eine Antwort geben müssen.

Die Test-Ergebnisse haben gezeigt, dass CASI eine nützliche Datenerhebungsmethode darstellt, die in den verschiedensten Situationen, bei denen Befragte hinter einem (Computer-) Bildschirm sitzen, angewendet werden kann. Daher ist CASI nicht nur gegenwärtig eine geeignete Methode sondern wird auch in der vor uns liegenden Zeit nützlich sein.

1. Hintergrund und Zielsetzung

1999 berichteten Soels/Tchaoussoglou über ein Experiment von SUMMO über die Einsatzfähigkeit von Computer-Assisted Face-to-Face Interviews für die holländische nationale Leser-Analyse (SummoScanner). Bis Juli 2001 nutzte der SummoScanner CATI-Interviews.

Aus zahlreichen Gründen wurde eine alternative Erhebungsmethode zum Telefoninterview für wünschenswert bzw. notwendig erachtet.

a) Änderungen in der Medien-Landschaft

In den 16 Jahren des Bestehens musste der SummoScanner mit einer schnellen Änderung der Medienumgebung fertig werden. Die Entwicklung kann mit einem Wort zusammengefasst werden: Mehr! Mehr Zeitschriften, mehr Tageszeitungs-Sektionen, mehr Fernsehsender, mehr Medien (z.B. Internet). Besonders der Anstieg der Zahl der abzufragenden Titel haben die Studie überbeansprucht, und dies besonders, weil die Zeit, die Befragte für ein Telefon-Interview akzeptieren, begrenzt ist. Die Titelzahl zu begrenzen, war durch die Gesetzgebung unmöglich. Des weiteren befriedigten mehr Zeitschriften die gleichen Leserschaften. Diese Zeitschriften hatten oftmals nicht nur gleiche Inhalte sondern auch sehr ähnliche Titel. Dieses erhöhte das Problem der Titel-Verwechslungen - ein Problem, das für Telefon-Interviews größer ist als für Erhebungsmethoden, bei denen Visualisierungen genutzt werden können.

b) Änderungen in Gesellschaft und Forschung

Obwohl die Niederländer im Durchschnitt höchstens zweimal im Jahr für Meinungsumfragen angerufen werden, empfinden sie es als sehr lästig - weil die Bevölkerung auch nicht zwischen Marktforschung und Telemarketing unterscheidet. Die Teilnahmebereitschaft an Marktforschungs-Untersuchungen geht generell zurück. (SummoScanner: von 1989 bei 60% zu unter 40% in 2000.) Auch der Kostenvorteil für große repräsentative Telefon-Stichproben gegenüber anderen Erhebungsmethoden verliert an Boden. Ein weiteres Problem ist, dass man Interviewer benötigt. Dieses ist sowohl ein methodisches Problem (Interviewer-Bias z.B. durch das "Hindurchheilen" durch ein Interview) als auch ein logistisches Problem (gute Interviewer sind kaum zu bekommen und es wird in Zukunft noch schlechter werden). Der Vorteil der Computerunterstützung bei Telefon-Interviews kann inzwischen auch für face-to-face Interviews (CAPI) per Laptop oder CASI sowie für Internet-Umfragen in Anspruch genommen werden.

Aus diesen Gründen entschied sich SUMMO 1997 die Befragungstechnik der nahen Zukunft für die holländische Leser-Analyse zu entwickeln. Tests dazu waren: 1999 eine qualitative Pilot-Studie, Frühjahr 2000 eine quantitative Pilot-Studie, Herbst 2000 eine Stichproben-Pilot-Studie. Die qualitative Pilot-Studie zeigte, dass CASI eine vielversprechende Erhebungsmethode ist, die immer genutzt werden kann, wenn ein Computer-Bildschirm zur Verfügung steht. Der Fragebogen kann auf vielen Wegen zum Befragten gelangen, das wiederum kann die Repräsentativität der Stichprobe erhöhen. Die Befragten können den Fragebogen, wie, wann und wo sie wollen, beantworten.

Herausgefunden werden sollte im anschließenden quantitativen Pilot-Test die beste Methode zur Durchführung der CASI-NRS (holländische LA).

- a) Ermittlung der Möglichkeit, Erhebungsmethoden zu kombinieren
- b) Tests, die Filterfragen schneller zu machen
- c) Untersuchung der Änderung im Hinblick auf die Reichweiten
- d) Untersuchung des Phänomens der Instituts-Unterschiede
- e) Ermittlung des besten Stichproben-Verfahrens

2. Design

a) Qualitative Pilot-Studie

Feldarbeit: 2 Institute (Interview-NSS und NIPO), jedes Institut führte ca. 350 face-to-face Interviews durch. Zusätzlich: ca. 300 Interviews von NIPO in einem Access-Pool per e-mail. Beide Gruppen erhielten im Anschluß an das Interview ein Papier-Selbstaussfüller.

Splits: Im face-to-face Interview wurden zwei Splits erhoben: CASI vs. CAPI (Interview-NSS). Um den Interviewer-Bias zu eliminieren, füllen die Befragten den Fragebogen idealerweise selbst aus (CASI). Die qualitative Vorstudie hat gezeigt, dass dieses für die Mehrheit kein Problem darstellt, auch wenn sie noch nie einen Computer bedient haben. Bei der NRS kann es jedoch einen kleinen Anteil von Personen geben, wo der Interviewer den Fragebogen für den Befragten ausfüllen muss. (CAPI). Um herauszufinden, ob das Auswirkungen auf die Reichweiten haben kann, wurden die Interviews, die Interview-NSS durchzuführen hatte, in zwei Gruppen aufgeteilt: eine Gruppe (50%) füllte den Fragebogen selbst aus (CASI), bei der anderen Gruppe (50%) füllten die Interviewer den Fragebogen aus (CAPI).

Filterbeschleunigung durch Fehler-Coding (NIPO)

Die qualitative Vorstufe zeigte, dass - verglichen mit der Abfrage des Generalfilters pro Titel - EML-Filter (Extended Media-Lists) zu erheblichen Zeitgewinnen führten. Besonders, wenn Titel mit kleinen Reichweiten hinzukommen, ist er sehr effizient. Zunächst werden von den Befragten Titelgruppen identifiziert, von denen mindestens einer in den letzten 12 Monaten gelesen wurde. Die Frequenz-Frage wird dann für alle Titel der ausgewählten Titelgruppe gestellt.

Normalerweise lese ich ... von den Ausgaben von ...

keine - so gut wie keine - 1 von 4 - 2 von 4 - 3 von 4 - meistens jede - jede

Da einige Befragte meinten, sie hätten bereits Titel ausgeschlossen, die sie nicht lesen (vielleicht nur in Gedanken oder weil sie es gesagt haben) und weil bei den meisten Titeln "keine" angegeben wurde, wurde getestet, ob man nicht bei allen Titeln im voraus "keine" angibt, so dass die Befragten dann die Antwort nur für die Titel, die sie lesen, ändern müssen.

Übersicht Quantitative Vorstufe

Face-to-face Interview-NSS Split 1: CASI (ohne "keine"-Vorgabe) n = 199	Face-to-face NIPO Split 3: CASI (mit "keine"-Vorgabe) n = 166	Access pool NIPO Split 5: Computer Access Pool (ohne "keine"-Vorgabe) n = 311
Split 2: CAPI (ohne "keine"-Vorgabe) n = 174	Split 4: CASI (ohne "keine"-Vorgabe) n = 170	

Stichprobe

Die Stichprobe war eine eingeschränkte Random-Walk-Stichprobe. Die Interviewer erhielten Startadressen und sollten rund um diese Startadressen (links, rechts und gegenüber) drei Interviews realisieren. Für jedes Interview stand ein Cluster mit fünf aufeinanderfolgende Privatadressen zur Verfügung, die kontaktet werden konnten. Bei Nichtantreffbarkeit konnten zwei Wiederholungsbesuche gemacht werden. Das Cluster galt als abgearbeitet und wurde nicht weiter benutzt, wenn entweder ein Interview realisiert wurde oder nach drei Besuchen (auch, wenn kein Interview zustande kam). Die Adressen sollten möglichst abends kontaktet werden und die Interviews über alle Wochentage verteilt werden. Außerdem erhielten die Befragten Incentives (10 Gulden für das Interview und 15 Gulden für den Selbstausfüller im Anschluß).

b) Stichproben-Test

Nachdem in der quantitativen Vorstufe gezeigt wurde, dass ein face-to-face Ansatz keine akzeptable Stichprobe erbringt, wurde ein Telefon-Screening getestet. Idealerweise sollten die Befragten

1. einen Brief erhalten
2. danach per CATI interviewt werden und
3. am Ende des Interviews sollte ein Termin vereinbart werden, bei dem die Leserschafts-Untersuchung zu Ende durchgeführt werden sollte.

Dazu mussten die Adressen und die Telefonnummern zusammengebracht werden. Zwei Vorgehensweisen wurden getestet:

1. Ziehung eines Samples mit Adressen und Verknüpfung mit Telefon-Nummern (Version A)
2. Ziehung eines Samples mit Telefonnummern und Verknüpfung mit Adressen (Version T)

Selbstverständlich gab es hierbei Ausfälle, wenn z.B. die Telefonnummern nicht gefunden werden konnten. Für jede Version gab es zwei Möglichkeiten: 1. erfolgreiche Verknüpfung, 2. keine Verknüpfung.

Übersicht Stichprobe

A1

- Adresse bekannt
- erfolgreiche Verknüpfung mit Telefonnummer / Name
- personalisierter Einführungsbrief
- Telefon-Screening

A2

- Adresse bekannt
- keine Verknüpfung mit Telefonnummer / Name
- nicht-personalisierter Einführungsbrief
- face-to-face Kontakt

T1

- bekannte Telefonnummer
- erfolgreiche Verknüpfung mit Adresse / Name
- personalisierter Einführungsbrief
- Telefon-Screening

T1

- bekannte Telefonnummer
- keine Verknüpfung mit Adresse / Name
- kein Brief
- Telefon-Screening (kalter Anruf)

3. Kombination der Datenerhebungs-Methoden

In der Vorstufe (und in der nachfolgenden holländischen NRS) wurden zwei Datenerhebungs-Methoden kombiniert:

- a) CAPI und CASI
- b) Face-to-face und Access pool

In der quantitativen Vorstufe wurde analysiert, welche Auswirkungen die verschiedensten Methoden auf die Leserschafts-Daten haben. Aufgrund der geringen Fallzahlen werden die Ergebnisse nicht für Einzel-Titel sondern nur für alle Titel gesamt, alle Zeitschriften und alle Tageszeitungen ausgewiesen.

a) CASI vs. CAPI

Es gibt Unterschiede, die aber nicht signifikant sind. Die größten (nicht signifikanten) Unterschiede treten bei den Wochenzeitungen auf, wenn der Interviewer die Fragen stellt. Offensichtlich wird dieses durch das sozial gewünschte Antwortverhalten verursacht. Dieser Effekt ist eine weitere Bestätigung für unsere Präferenz auf Interviewer zu verzichten.

Vergleich der Leserschaftsdaten CASI / CAPI

	CASI	CAPI
Alle Titel		
In den letzten 12 Monaten gelesen	1.231%	1.252%
Leser pro Nummer	513%	505%
Alle Zeitschriften		
In den letzten 12 Monaten gelesen	898%	902%
Leser pro Nummer	350%	339%
Alle Tageszeitungen		
In den letzten 12 Monaten gelesen	218%	234%
Leser pro Nummer	79%	90%

b) Face-to-face vs. Access pool

Die Leserschaftsdaten der face-to-face Interviews (Split 4) und des Access pools (Split 5) unterscheiden sich signifikant für die meisten Titel-Gruppen (nur die Interviews vom gleichen Institut und ohne "keine"-Vorgabe wurden verglichen). Dies ist hauptsächlich auf die Unterschiede in den beiden Stichproben zurückzuführen: Im Access pool sind mehr Männer, mehr verheiratete und im Durchschnitt vier Jahre jüngere Befragte enthalten. Mehr Personen sind berufstätig haben (logischerweise) einen Computer und bessere Computer-Kenntnisse.

Vergleich face-to-face / Access pool

	Face-to-face	Access pool
Basis:	n = 170	n = 311
Anzahl ausgewählter Filterpassagen	10,85	13,05
In den letzten 12 Monaten gelesen		
Alle Titel	1.700%	2.166%
Alle Zeitschriften	1.289%	1.663%
Alle Tageszeitungen	267%	350%

Wenn in beiden Stichproben die gleiche Art von Personen verglichen werden (Pers., die einen Computer haben und berufstätig sind) gibt es keinen signifikanten Unterschied bei der Filterpassage und bei den Leserschaftsdaten für die letzten 12 Monate. Einen signifikanten Unterschied gibt es für das Lesen von Tageszeitungen in den letzten 12 Monaten. Dieser Unterschied wird von der höheren Anzahl von Männern im Access pool verursacht. Wenn das Geschlecht gleich gehalten wird, wird der Unterschied geringer und verschwindet bei den Frauen fast ganz. Das bedeutet, dass die Befragten unabhängig davon, wie der Fragebogen verteilt wird, den Fragebogen auf die gleiche Art und Weise ausfüllen. Die Daten von den verschiedenen Erhebungsmethoden können also ohne Bias kombiniert werden (Hinweis auf Paper von Candon und Fagot, RRS Berlin 1995).

Vergleich face-to-face / Access pool zwischen berufstätigen Computer-Besitzern

	Face-to-face	Access pool
Männer		
Basis:	n = 27	n = 106
Anzahl ausgewählter Filterpassagen	12,00	13,41
In den letzten 12 Monaten gelesen		
Alle Titel	1.885%	2.253%
Alle Zeitschriften	1.437%	1.660%
Alle Tageszeitungen	348%	432%
Frauen		
Basis:	n = 40	n = 70
Anzahl ausgewählter Filterpassagen	13,35	13,69
In den letzten 12 Monaten gelesen		
Alle Titel	2.245%	2.387%
Alle Zeitschriften	1.835%	1.891%
Alle Tageszeitungen	290%	341%

Offensichtlich muss die Gesamt-Stichprobe für die face-to-face Interviews und die Access pool-Interviews sorgfältig gewichtet werden, um eine Nichtrepräsentativität zu vermeiden. Außerdem sollte die Access pool Stichprobe geschichtet werden, um zu hohe Gewichtungsfaktoren zu vermeiden. Es wird erwartet, dass die Access pool Stichprobe in Zukunft repräsentativer wird. CASI ist für eine mehrkanalige Verteilung des Fragebogens gut geeignet.

4. Maßnahmen, die dazu führen, dass der Fragebogen schneller ausgefüllt werden kann

Um herauszufinden, ob die Zeit zum Ausfüllen des Fragebogens reduziert werden kann, wurde die Frequenzfrage überarbeitet. Die Frequenzfrage wird gestellt, wenn ein Filter mit einer Gruppe von Titeln passiert wurde, d.h. wenn mindestens ein Titel der Gruppe in den letzten 12 Monaten gelesen wurde.

(Bsp. Titel-Gruppe mit Filterfrage)

Es werden drei bis fünf Titel auf einmal auf dem Bildschirm gezeigt und die gleichen drei bis fünf Titel erscheinen in der Frequenzfrage. Im Durchschnitt werden davon nur ein bis zwei Titel gelesen, d.h. die Antwort auf die Frequenzfrage, wieviele Ausgaben gelesen werden, lautet meistens "keine". Wenn dieses bereits im Voraus ausgefüllt ist und nur bei Nichtzutreffen - wenn ein Titel gelesen wird - geändert werden muss, sollte das zu einem Zeitgewinn führen. Selbstverständlich birgt das die Gefahr, dass die Befragten weniger akkurat antworten und Titel übersehen. Um diese Frage zu klären, wurden Splitversionen getestet: die face-to-face Interviews von NIPO wurden zur Hälfte mit der ausgefüllten Antwort "keine" auf die Frequenzfrage durchgeführt (Split 3 und 4).

Ein Vergleich der beiden Splits erbrachte drei Minuten Zeitgewinn (Gesamt-Interviewdauer: 34 Minuten). Außerdem führte die vorgegebene Antwort zu einer Verringerung der Reichweiten, obwohl in beiden Gruppen die Filterpassagen mit zehn gleich hoch waren. Die Gruppe mit der vorgegebenen Antwort hatte im Vergleich eine signifikant geringere Reichweite für die meisten Zeitschriften als die Gruppe ohne die vorgegebene Antwort. Mit anderen Worten, weniger Titel passierten den Frequenz-Filter. Tageszeitungen wurden davon nicht berührt. Eine Erklärung dafür ist, dass, wenn eine Antwort gegeben werden muss, die Befragten sorgfältiger nachdenken. In der qualitativen Vorstufe wurde beobachtet, dass sich die Befragten Zeit für die Erinnerung ließen. Es ist plausibel, dass diese Extra-Anstrengung zu valideren Frequenz-Schätzungen führt. Es wird deshalb nicht empfohlen, Fragen mit vorgegebenen Antworten (für "keine Ausgabe gelesen") einzusetzen, wenn sie entscheidend für die Höhe der Leserschaftsdaten sind.

5. Leserschaftsdaten und Institutshandschrift

Was geschieht, wenn Leserschaftsdaten nicht länger per CATI, sondern per CASI erhoben werden? Anhang A gibt einen ersten Eindruck über den Generalfilter "in den letzten 12 Monaten gelesen" von CASI (nur face-to-face Interviews), die Werte werden indiziert auf die CATI-Zahlen dargestellt. Mit Ausnahme für die wöchentlichen Titel, die vom Methodenwechsel profitieren, gibt es für die meisten Titel eine Verringerung der Reichweiten. Dieses resultiert aus einer Änderung der Filterfrage: Während im SummoScanner nach "jemals gelesen" gefragt wurde, beinhaltet der neue Filter den Zeitrahmen "in den letzten 12

Monaten". Später wird gefragt: "Wie viele Ausgaben lesen Sie im allgemeinen?" Während der alte Filter sehr weit gefaßt wurde, basiert der neue auf gewohnheitsmäßigem Verhalten in einem festgelegtem Zeitraum (auch wenn die Antwort "meist keine" sein kann). Die Reichweiten-Verringerung gilt nicht für alle Titel-Gattungen: Besonders die Titel, die wahrscheinlich von der Titel-Verwechslung am Telefon profitieren, wie Sportzeitschriften, Auto-Zeitschriften oder Heim-und-Garten-Zeitschriften, verlieren Reichweite. Das stützt die Vermutung, das Titel-Verwechslungen mit visuellen Vorlagen reduziert werden.

Am meisten beeindrucken die Ergebnisunterschiede zwischen den beiden Instituten. Obwohl das Studien-Design identisch ist, misst Interview-NSS viel geringere Reichweiten als NIPO. Die Reichweiten sind im NIPO-Sample 30% höher als im Interview-NSS-Sample. Um herauszufinden, ob diese Unterschiede auf unterschiedliche Stichproben zurückzuführen sind, wurde eine Regressionsanalyse durchgeführt. Als erklärende Variable wurden Institut, Geschlecht, Alter, Berufstätigkeit und Computer-Besitz benutzt. Der Unterschied - obwohl gering - wird hauptsächlich den Instituten zugeschrieben: von 2,6 Filterpassagen, die von NIPO-Befragten mehr angegeben wurden, werden 1,8 auch den demographischen Merkmalen zugeordnet. Es gibt keinen signifikanten Unterschied bei Tageszeitungen.

Unterschiede Institute

	"Institut" als Erklärung	"Institut" und "Demographie" als Erklärung
Anzahl Filterpassagen	2,59	1,88
Anzahl gelesener Zeitschriften	3,89	2,87
Anzahl gelesener Tageszeitungen	0,388*	0,098*

n = 916, * nicht signifikant

Normalerweise werden Instituts-Handschriften den unterschiedlichen Interviewern der Institute zugeschrieben. Da diese Untersuchung die CASI-Methode nutzt, kann das nicht die Hauptklärung sein. Die einzig mögliche Erklärung kann die Interview-Software sein, die geringe Unterschiede zwischen den Instituten aufweist. (Z.B. unterschiedliche Plazierung des "Weiter"-Button.) Es wird daher empfohlen, bei zukünftigen Untersuchungen, bei Beauftragung von mehr als einem Institut, auch geringe Details identisch zu halten.

6. Ausschöpfung

In der quantitativen Vorstufe wurde ein face-to-face Ansatz gewählt (fünf Adressen in drei Clustern). Die Ergebnisse blieben unbefriedigend. Es wurden zu viele Frauen befragt und zu wenig junge Leute. Außerdem gab es keine gute Repräsentativität hinsichtlich Region, Berufstätigkeit und Computer-Besitz. Kurz gesagt, wir interviewten die Alten und die Hausfrauen - diejenigen, die zu Hause waren.

Ausschöpfung face-to-face Interviews

	pro Cluster		pro Adresse	
	n =	%	n =	%
Komplett	709	67	709	22
Verweigert	} 342	} 33	1684	51
Nicht angetroffen			875	27
Gesamt	1.051	100	3.268	100

Die Verweigerungsrate mag hoch erscheinen, war aber keine Überraschung, da sie mit Erfahrungen aus anderen Forschungsprojekten übereinstimmt. Das Problem wird dadurch verursacht, dass Personen, die nicht zu Hause sind, durch Personen ersetzt werden, die zu Hause sind. Es wurde deshalb entschieden ein Telefon-Screening einzusetzen. Dieses wurde in einem Stichproben-Test ausprobiert.

(Anlage der Stichprobe siehe unter 2.)

A: A1 - Adresse, Name und Telefonnummer bekannt

A2 - nur Adresse bekannt (Kontakt an Haustür)

T: T1 - Adresse, Name und Telefonnummer bekannt

T2 - nur Telefonnummer bekannt ("kalter" Anruf)

Das Telefon-Screening konnte in A1, T1 und T2 durchgeführt werden. Die Ergebnisse von A2 sind fragwürdig, da wahrscheinlich nicht alle Interviewer alle Kontakt-Versuche aufgelistet haben.

Ausschöpfung

	A1	A2	A Ges.	T1	T2	T Ges.
Screening	49%	-	49%	53%	36%	48%
face-to-face Interview	21%	34%	?	17%	11%	

Es gibt nur einen kleinen Unterschied zwischen A und T. Bei beiden beträgt die Ausschöpfungsquote für das Screening ca. 50% und für das folgende face-to-face Interview ca. 20%. "Kalte" Anrufe erbringen eine geringere Ausschöpfungsquote, fallen jedoch aufgrund ihrer geringen Anzahl kaum für die Gesamt-Ausschöpfungsquote ins Gewicht.

Die Zusammensetzung der Stichprobe ist repräsentativer als bei der quantitativen Vorstufe. Von den beiden (jetzt genutzten) Vorgehensweisen (A und T) ist keine besser als die andere. Das Adress-Sample weist jedoch theoretische Vorteile gegenüber dem Telefon-Sample auf. (Hinweis auf Mobil-Telefonnummern, ISDN-Anschlüsse mit mehreren Nummern und die damit unterschiedlichen Auswahlchancen usw. - Kuusela & Vikki 1999)

7. Zusammenfassung

Die Vorstufen haben aufgezeigt, was zu tun und was zu lassen ist, wenn man eine Leserchaftsuntersuchung per CASI durchführen will:

- Die Nutzung visueller Vorlagen reduziert Titel-Verwechslungen
- Das Ausfüllen der Fragebogen durch den Befragten verhilft zu mehr Sorgfalt. Jedoch nur dann, wenn keine Antwort bereits vorgegeben wird.

- Falls es notwendig ist, dass der Interviewer den Fragebogen ausfüllt, hat es keine Auswirkungen auf die Reichweiten.
- Die Verteilung des Fragebogens über verschiedene Medien-Kanäle ist sehr gut möglich.
- Die beobachtete Instituts-Handschrift erinnert daran, selbst kleinste Details vorzugeben, wenn die Untersuchungs-Durchführung bei mehr als einem Institut liegt, aber auch, wenn nur ein Institut beauftragt wird.
- Obwohl das Telefon nicht mehr als Erhebungsinstrument akzeptiert wird, kann es hervorragend zur Rekrutierung der Zielpersonen eingesetzt werden.

Beim Verfassen dieses Vortrags war den Autoren nicht bekannt, wann die neue Leserschafts-Untersuchung in den Niederlanden startet und ob tatsächlich die CASI-Methode dafür eingesetzt wird. Der Grund liegt in der Gründung eines neuen JIC (Verband), der sich bislang nicht einigen konnte, ob Zeitschriften und Tageszeitungen gemeinsam oder getrennt, in einer eigenen Untersuchung, abgefragt werden sollen. Es kann auch sein, dass die jetzigen Ergebnisse den Prozess eher verlangsamen: Je mehr man weiß, desto weniger wird getan. (Hinweis auf belgische Kollegen, die aus diesem Grund zahlreiche Änderungen in der nationalen Leserschafts-Untersuchung nie getestet haben - Haute 1997).

Wir sind davon überzeugt, dass CASI die Datenerhebungsmethode der Zukunft ist.

Personalised Media Lists

Personalisierte Media-Listen

Katherine Page, Ipsos-RSL

Synopse

Dieser Vortrag zeigt einen neuen Weg, wie die Anzahl der Titel, zu denen eine Zielperson befragt wird, reduziert werden kann - unter Beibehaltung oder sogar bei Ausweitung der Gesamt-Titel-Liste.

Bei PML® handelt es sich um eine Technik, die herausfiltert, zu welchen Titeln die Zielpersonen befragt werden sollen. Sie basiert auf einer geringen Anzahl von demographischen Fragen und Fragen zum Interesse, die zu Beginn des Interviews gestellt werden. Aufgrund der Komplexität dieser Filterung kann PML nur bei Computer-Interviews angewandt werden. Genauso, wie alle Titel gezeigt werden, bei denen eine große Wahrscheinlichkeit besteht, dass sie gelesen werden, werden stichprobenartig Titel gezeigt, die nur eine geringe Wahrscheinlichkeit haben, gelesen zu werden. Folglich wird der Hauptteil der Leserschaftsdaten direkt von den Befragten erhoben und es besteht eine Basis um die relativ wenigen "fehlenden" Daten zu ergänzen.

Für den Einsatz der PML gibt es verschiedene Varianten. Der Vortrag zeigt einen Prototyp, der mit den Daten, die bereits verfügbar sind, weiterentwickelt werden kann.

Hintergrund

Das Konzept der PML resultiert aus einer lang anhaltenden Debatte über die Zahl der Titel, die jeder Befragte bei der NRS zu beantworten hat. Zur Zeit werden rund 300 Tageszeitungen und Zeitschriften abgefragt plus zusätzlich eine Auswahl von Tageszeitungs-Sektionen. Es ist nur möglich eine solch große Anzahl von Titeln abzufragen, weil seit 1984 die EML-Technik eingesetzt wird. (EML = Extended Media-List. Dabei handelt es sich um Gruppen-Titelkarten, die zunächst vom Befragten sortiert werden, ob sich darauf mindestens ein Titel befindet, der in den letzten 12 Monaten gelesen wurde. Auf diese Art werden die meisten EML-Karten aussortiert, bevor die Abfrage nach Einzel-Titeln beginnt.)

Die Einführung der EML ließ die Titel-Liste merklich ansteigen, die aber seit den letzten 10 Jahren hinlänglich gleichbleibend ist.

(Graph. Länge der Media-Liste im Zeitablauf)

Unabhängig davon wird die Länge der Media-Liste seit langem diskutiert. Die Tageszeitungen befürchten, dass es ein Ungleichgewicht bei der Abfrage von Tageszeitungen (54 plus eine Auswahl regionaler Tageszeitungen) und Zeitschriften (237) gibt. Auf der anderen Seite möchten die Nutzer (Werbungtreibende und Agenturen) gerne mehr Titel abfragen, die bisher nicht berücksichtigt wurden.

Eine Lösung wäre, die Media-Liste in zwei oder mehr Sub-sample aufzusplitten und die Daten dann zu fusionieren. Hierzu wurden in UK Untersuchungen durchgeführt. 1995 wurde untersucht, wie gut eine solche Fusion funktioniert. Diese Technik wurde nicht für die NRS aber in anderen Ländern benutzt.

Unabhängig von der Effizienz und der Qualität von Fusionen haben gesplittete Stichproben eine Reihe von Nachteilen, so wird z.B. der Standard-Fehler erhöht. Das wiederum hat Auswirkungen auf die "Reichweiten-Währung". Insbesondere für kleinere Titel stellt die Stichprobenreduktion ein Problem dar. Außerdem werden dem Befragten vielleicht gerade die Titel nicht gezeigt, an denen er ganz besonders interessiert ist.

Neben der Diskussion um die Länge der Media-Liste wurden Ende der neunziger Jahre auch die Möglichkeiten der neuen Interviewtechniken betrachtet. Besonders interessant war die Technik, bei der alle Erinnerungshilfen per Computer-Bildschirm gezeigt wurden und nicht mehr als Papiervorlage. (siehe Double-screen-Technik in Frankreich) Beim Einsatz von Computern können auch noch weitere Vorteile genutzt werden, z.B. komplexere Filter, als wenn ein Interviewer mit Papier-Material hantiert.

Dieses hatte Ipsos-RSL vor Augen, als man begann über das Konzept der PML, basierend auf Fragen zur Demographie und zu den Haupt-Interessen, nachzudenken und der NRS 1999 zu präsentieren. Als die Ausschreibung für die NRS im Juli 2000 erfolgte, war eine der Anforderungen, dass die Titel, die von jedem Befragten zu beantworten waren, reduziert werden sollten, und dass dazu Vorschläge von den Anbietern erwartet würden.

Zusammenfassung der Methode

Für die PML-Technik gibt es fünf Haupt-Elemente:

1. Bevor das Interview mit der Abfrage der Leserschaftsfragen beginnt, werden eine kleine Anzahl von Demographie-Fragen und Fragen zu den Haupt-Interessen gestellt. Auf dieser Basis wählt die Computer-Software die abzufragenden Titel aus.
2. Die Media-Liste wurde bereits in Abhängigkeit der Demographie- und Interessen-Variablen in Cluster eingeteilt. Wenn der Befragte für ein Cluster ausgewählt wird, werden ihm alle Titel dieses Clusters gezeigt. Da konkurrierende Titel in der gleichen Gruppe sind, werden sie auch gleichermaßen mit vorgelegt.
3. Für jedes Cluster wurden frühere Untersuchungsergebnisse analysiert, um die Variablen herauszufinden, die die Vorhersage, ob ein Titel in dem Cluster gelesen werden würde, am besten einschätzen konnten. Drei Ebenen wurden für die Cluster definiert: "wird voraussichtlich gelesen", "wird voraussichtlich nicht gelesen" und "ist sehr unwahrscheinlich, dass der Titel gelesen wird".
4. Allen Befragten, die auf der Ebene "wird voraussichtlich gelesen" eingestuft werden, werden alle Titel des entsprechenden Clusters gezeigt. Nur einer kleinen Stichprobe derjenigen, die sich auf den anderen beiden Ebenen befinden, werden diese Titel gezeigt.
5. Die Daten, die auf den geringeren Ebenen von den Befragten eingesammelt werden, werden dazu verwendet, die nicht direkt ermittelten Informationen zu schätzen bzw. "aufzufüllen". Dieses wird durch eine Reihe von Mini-Fusionen erreicht.

Diese Methode wird auf Kundenwunsch nur für Zeitschriften angewandt. Die Tageszeitungen und die Sonntagszeitungen werden komplett bei allen Befragten abgefragt.

Demographie-Fragen und Fragen zum Haupt-Interesse

Die Vorfragen, die die zu zeigenden Titel auswählen, dürfen nicht zu lang sein. Nicht geeignet sind Fragen zur sozialen Schicht.

Nach der Analyse früherer NRS-Daten werden folgende Fragen empfohlen:

- Geschlecht
- Alter
- Haushalts-Typ (Anteil Erwachsener und Kinder)
- Familienstand
- Zukunftspläne
- Haupt-Interessen

Die Haupt-Interessen sind im allgemeinen die wichtigste Variable, diskriminieren aber noch besser in Verbindung mit den demographischen Angaben. Außerdem wird dadurch sichergestellt, dass allen demographischen Zielgruppen auch die Zielgruppen-Titel gezeigt werden, unabhängig von ihren Haupt-Interessen. Motorsport-Magazine können z.B., unabhängig davon, ob die Befragten angeben an Motorsport interessiert zu sein, allen Männern im Alter von 15 bis 34 Jahren vorgelegt werden, wenn dieses die Haupt-Zielgruppe ist.

Die vor den Leserschaftsfragen gestellten Fragen könnten jedoch einen Einfluß auf die Leserschaftsdaten haben. Dieses wurde in einem Test seit April 2001 überprüft, bei dem je zur Hälfte die Fragen vor den Leserschaftsfragen gestellt wurden und je zur Hälfte direkt nach den Leserschaftsfragen. Nach den Ergebnissen der ersten drei Monate kann ein Einfluß ausgeschlossen werden

Einteilung der Zeitschriften

Mittels Clusteranalysen wurden anhand früherer NRS-Daten die Titelgruppierungen vorgenommen. Wenn ein Befragter zu einem Titel-Cluster befragt wird, werden ihm jeweils alle Titel des Clusters gezeigt. Die Clusteranalyse erbrachte eine Vielzahl von Lösungsmöglichkeiten. Folgende 16 Cluster-Lösung wurde ausgewählt:

	Anzahl Titel		Anzahl Titel
1 Frauen-Zeitschriften	29	9 Männer, Musik und Film	11
2 Zeitschriften für junge Frauen	13	10 Männer-Zeitschriften	9
3 Eltern-Zeitschriften	5	11 TV-Zeitschriften	10
4 Lifestyle	19	12 Männer-Interessen-Zeitschriften	32
5 Zeitschriften für ältere Frauen	7	13 High Flyers	8
6 Ältere Zeitschriften	6	14 Hobby-Zeitschriften	5
7 Teenager-Zeitschriften	13	15 Garten, Landleben und Klassik	18
8 Fußball und schnelle Autos	9	16 Handels-Zeitschriften	3

Vorhersage der Lesewahrscheinlichkeit

Für jedes Cluster wurden die Variablen ausgewählt, die am besten vorhersagen würden, ob ein Befragter wahrscheinlich irgendeinen Titel des Clusters im letzten Jahr gelesen haben würde. Lesen im letzten Jahr wurde aus zwei Gründen ausgewählt: Damit wird die größtmögliche Lesergruppe ausgewählt und ein positiver "Leser-im-letzten-Jahr-Filter" gilt für alle nachfolgenden Leserschaftsfragen. Eine CHAID-Analyse identifizierte die härtesten Variablen um die Leserschaft zu beschreiben.

Der nächste Schritt war die Schichtung der Stichprobe. Dieses wurde wieder für jedes Cluster einzeln vorgenommen. Die drei Ebenen sind:

1. hoch = höchste Wahrscheinlichkeit ein Leser zu sein
2. gering = geringe Wahrscheinlichkeit ein Leser zu sein
3. sehr gering = sehr unwahrscheinlich ein Leser zu sein.

Die Entscheidung, wie die Befragten zu dem jeweiligen Cluster eingestuft werden, muss zwei Dinge berücksichtigen: zum einen sollte die Unterscheidung der Leserschaften zwischen den drei Ebenen so effizient wie möglich vorgenommen werden, damit es differenzierte; auf der anderen Seite sollten so viele Leser wie möglich zu den Einzel-Titeln befragt werden. Die Cluster, die in die Gruppe "hoch" fielen, wurden alle gezeigt, aber nur eine Auswahl aus denen der "geringen" und "sehr geringen" Gruppe.

Beispiel:

Männer, Musik & Film (11 Titel)

Beschreibung der Ebenen		% der ungewichteten NRS-Stichprobe	% von Befr., die mind. 1 Titel im letzten Jahr gelesen haben	% von Befr., die LpN von mind. 1 Titel sind	LpN-Anteile (mind. 1 Titel)
Hoch	Alter 15 - 24 J. Männer 25 - 44 J.	26 %	72 %	78 %	30 %
Gering	Frauen 25 - 34 J. Interesse an Kunst u. Musik usw. Männer 45 - 54 J. Interesse an Kunst u. Musik usw.	12 %	15 %	13 %	11 %
Sehr gering	sonst. Männer sonst. Frauen	62 %	13 %	9 %	2 %

Basis: NRS Okt. 98 - Sept. 99

Wie zu sehen ist, bilden 26% des NRS-Samples 72% der Leser im letzten Jahr für dieses Cluster und 78% des LpN. In diesem Cluster waren elf Titel enthalten. Wenn man den LpN für individuelle Titel über die drei Ebenen betrachtet, werden die Einteilungen noch aussagekräftiger. Nachfolgend ein Beispiel für vier Einzel-Titel: FHM, einer der führenden Männer-Titel und drei Musik-Zeitschriften.

% der Befragten, die LpN von ... sind

Beschreibung der Ebenen		FHM	Q Magazine	Mixmag	Select
Hoch	Alter 15 - 24 J. Männer 25 - 44 J.	19,5 %	3,4 %	3,1 %	1,7 %
Gering	Frauen 25 - 34 J. Interesse an Kunst u. Musik usw.	5,9 %	1,5 %	0,4 %	0,3 %
	Männer 45 - 54 J. Interesse an Kunst u. Musik usw.				
Sehr gering	sonst. Männer sonst. Frauen	0,7 %	0,2 %	*	0,1 %

Basis: NRS Okt. 98 - Sept. 99

Auswahl der Befragten für die drei Ebenen

Um für die nicht abgefragten Titel, die nicht zur "hohen" Cluster-Ebene mit der höchsten Wahrscheinlichkeit des Lesens gehören, eine Basis für das Modellierung der Daten zu ermitteln, wurden einige der Befragten zusätzlich mit den Titeln aus den "geringen" und den "sehr geringen" Cluster-Ebenen konfrontiert. Bei den "geringen" Clustern wurde die Hälfte der Stichprobe herangezogen und bei den "sehr geringen" ein Viertel.

Das bedeutet, dass die Anzahl der Titel, die jedem Befragten gezeigt werden, unterschiedlich ist. Selbst wenn die Befragten die gleichen demographischen Merkmale und die gleichen Interessen aufweisen, kann die Anzahl unterschiedlich sein, je nachdem, ob sie in die Stichprobe gelangen, denen die Titel aus der "geringen" und/oder "sehr geringen" Cluster-Ebene gezeigt werden. Außerdem variieren die einzelnen Cluster in ihrer Größe.

Die Verteilung der Titel pro Befragten wird nach einer Simulation der PML (mit der Annahme, dass gleichzeitig sechs Titel pro Bildschirmseite gezeigt werden) folgendermaßen sein:

Anzahl der Titel inkl. Tageszeitungen	Anzahl der Zeitschriften	% der NRS-Befr.	kumulierte %-Zahl der NRS-Befr.
weniger als 144	weniger als 90	3,5	3,5
144 - 168	90 - 114	10,1	13,6
174	120	5,8	19,4
180	126	5,5	24,9
186	132	5,9	30,8
192	138	5,4	36,2
198	144	7,3	43,5
204	150	7,8	51,3
210	156	8,1	59,4
216	162	8,2	67,6
222	168	7,4	75,0
228	174	4,3	79,3
234	180	4,7	84,0
240	186	3,7	86,7
246	192	3,0	90,7
252	198	3,3	94,0
258	204	1,7	95,7
264 - 300	210 - 246	4,3	100,0

Der Zwang, allen die gleiche Titel-Anzahl vorzulegen, ist ein Hauptpunkt der Diskussion. Dabei wird argumentiert, dass die Titel Vorteile haben, deren Reichweite hauptsächlich von Befragten stammen, die nur wenige Titel vorgelegt bekommen. Solch ein Argument resultiert aus der Annahme, dass lange Media-Listen einen größeren Befragten- / Interviewer-Effekt aufweisen und dadurch die Reichweiten heruntergezogen werden. Mögliche Interviewer-Manipulationen bei den Filterfragen, um eine "kurze" Liste zu erhalten, werden ebenfalls befürchtet, obwohl die Komplexität der Filterungen und die Stichproben-Prozeduren das kaum zulassen. Durch die zufällige Zuordnung der Cluster, die auf der "geringen" und der "sehr geringen" Ebene vorgelegt werden, wird es für einen Interviewer unmöglich, eine bestimmte Antwortkombination für die Vorfragen vorherzusagen.

Auf der anderen Seite reduziert die PML nicht nur die Zahl der Titel pro Befragten, sondern zeigt ihm die für ihn relevanten Titel. Außerdem können Befragte, die viele Titel lesen auch am besten mit einer langen Media-Liste umgehen. Umgekehrt, Befragte, die sehr wenig lesen, werden am meisten durch eine lange Media-Liste eingeschüchtert und/oder gelangweilt.

Anteile der direkt (erhobenen) Leserschaften

Die Mehrzahl der Leserschaften werden durch die Methode direkt erhoben (und nicht durch Fusion erstellt). Durch die Nutzung früherer NRS-Daten konnten wir den PML-Prozess und die Ergebnisse simulieren. Die direkt erhobenen Leserschaften variieren selbstverständlich von Titel zu Titel, im Durchschnitt sind es 89%. Wenn die Interessen-Frage überarbeitet wird, erwarten wir mehr als 90%. Das sind höhere Anteile als wenn die Media-Liste einfach auf verschiedene Subsamples aufgeteilt worden wäre. Wenn die Media-Liste auf zwei Stichproben aufgeteilt wird, werden lediglich 50% der Leserschaften direkt eingesammelt.

Alle nicht direkt erhobenen Leserschaften werden durch Mini-Fusionen modelliert. Wenn es z.B. 15 Cluster gibt, werden ca. 30 Mini-Fusionen durchzuführen sein (je eine für das Cluster mit "geringer" Ausprägung und mit "sehr geringer" Ausprägung). Es sollte an dieser Stelle erwähnt werden, dass die Datenbasis mit lediglich 11% sehr gering sein wird.

Einfluss des Standard-Fehlers

Da die Leserschaften nicht mehr in der Gesamt-Stichprobe erhoben werden, gibt es einen Einfluss auf den Standard-Fehler. Wenn man noch einmal die simulierte PML-Untersuchung anschaut, erhöht sich der Standard-Fehler um durchschnittlich 15%. Es gibt jedoch Unterschiede zwischen den Titeln:

Anstieg des Standard-Fehlers im Vergleich zur jetzigen NRS	% der veröffentlichten Zeitschriften-Titel (ohne TV-Zeitschriften)
weniger als 10%	23,3 %
10 - 19 %	52,9 %
20 - 29 %	17,5 %
30 - 39 %	6,3 %

Das ist gut vergleichbar mit einer einfachen Aufteilung der Media-Liste in zwei Hälften auf zwei Stichproben. Das hätte einen Anstieg des Standard-Fehlers von mehr als 40 % zur Folge.

Da jedoch in der Media-Planung i.d.R. nach Zielgruppen geplant wird und diese Zielgruppen auch den Kernzielgruppen der Zeitschriften entsprechen und diese wiederum alle direkt nach ihrem Leseverhalten für die jeweilige Titelgruppe gefragt werden, gibt es auch keinen Anstieg des Standard-Fehlers.

Test

Erste Tests, dass die PML-Prozeduren praktikabel sind, sind bereits erfolgt. Wenn Ergebnisse einer sechs Monats-Stichprobe der NRS vorliegen (18.000 Befr.), wird die Interessen-Frage für die PML überarbeitet werden.

Die britische NRS wird im Januar 2002 von CAPI auf Double-sreen-CAPI umgestellt. Ab Juli 2002 soll dann ein PML-Test mit einem halben NRS-Sample durchgeführt werden.

Zusammenfassung

Der kommerzielle Druck, möglichst viele Titel in einer Untersuchung zu erheben, ist groß und hat vielleicht den größten Einfluß auf die Entwicklung von Methoden. Allerdings muss dieser Druck gegen die Überfrachtung von Befragten, wenn zu viele Titel vorgelegt werden, ausbalanciert werden. Insbesondere, da bekannt ist, dass dieses auf einige Titel einen größeren Einfluss haben kann als auf andere.

Wenn die Anforderungen gestellt werden, die Titelanzahl pro Befragten zu reduzieren, scheint PML eine vielversprechende Option zu bieten. Selbstverständlich sind zu dem hier vorgestellten Modell viele Variationen möglich und es gibt keinen Zweifel, dass es weitere Entwicklungen gibt, wenn ein vollständiger quantitativer Test durchgeführt wurde.

(Anhang: Ergebnisse Interessenfrage vor und nach Leserschaftsfragen gestellt)

Non-Responders - An opportunity to look into their hearts and souls

Verweigerer - Eine Gelegenheit, um in ihre Herzen und Seelen zu schauen

Hastings Withers, PMB Print Measurement Bureau

Ivor Thompson, Thompson Lightstone (Maritz Research Inc.)

Synopse

Sinkende Ausschöpfungsquoten sind ein Hauptthema in der Marktforschung. Besonders bedeutend sind sie in der Leserschaftsforschung, wo die Reichweiten-Messungen für einen Titel unterschiedlich zwischen Befragten und Verweigerern sein können. Um diesem Vorzubeugen verlangen Leserschaftsuntersuchungen eine hohe Ausschöpfung. Nicht-Leserschaftsuntersuchungen erwähnen die Ausschöpfung meistens gar nicht.

Besonders junge Leute und Männer sind schwer zu erreichen. Aber es werden nicht alle demographischen oder Lifestyle-Merkmale von den Ausschöpfungsquoten beeinflusst. Unterscheiden sich die Leserschaftsdaten von Verweigerern von denen der Befragten?

Die PMB (Print Measurement Bureau) Return-to-sample-Untersuchung bietet eine gute Gelegenheit dieses zu überprüfen. Die Lesedaten für Titel, die in dieser Untersuchung abgefragt werden, sind sowohl für diejenigen bekannt, die an der Untersuchung teilnehmen, als auch für diejenigen, die nicht geantwortet haben. Die Ergebnisse zeigen, dass es verschiedene (Verhaltens-)Muster gibt. Insgesamt war die Medien-Nutzung bei den Befragungsteilnehmern höher als bei den Verweigerern. Verweigerer lesen nicht so viel. Bei Betrachtung von Einzel-Titeln werden jedoch einige durch geringe Ausschöpfungsquoten bevorzugt, während andere niedrigere Reichweiten ohne höhere Ausschöpfungsquoten haben.

1. Einleitung - Das Problem

In einigen Untersuchungen kann aufgrund der Stichprobenquelle Einblick in Aussehen und Verhalten von Verweigerern genommen werden, da ausgewählte demographische Merkmale zur Verfügung stehen und somit ein Vergleich angestellt werden kann. In einigen Gebieten stehen auch Daten der Industrie oder von staatlichen Statistikorganisationen zur Verfügung um Besitzdaten vergleichen zu können. In Kanada z.B. gibt es Daten über den Besitz von TV-Sets, Reisen außerhalb Kanadas usw. Alle diese Quellen geben einen wichtigen Input, um einen potentiellen Bias von Verweigerungen untersuchen zu können. Auch auf dem San Francisco-Symposium haben Hans Vorster und John Smart versucht, diese Frage durch Rückschluß-Verfahren zu beantworten. Sie kannten die demographische Verteilung jeder Stichprobeneinheit und ergänzten damit die Verweigerer. Damit versuchten sie dann, über Modellrechnungen, Leserschaftsdaten zu ergänzen. Sie kamen zu dem Ergebnis, dass von sechs Titeln, drei durch Verweigerungen Reichweiten-Verluste erlitten, einer davon profitierte und zwei nicht betroffen waren. Seither wurde nichts weiter unternommen.

Ebenfalls auf dem San Francisco-Symposium präsentierten wir ein Paper, das den Einfluß von Wiederholungsanrufen auf die demographische Verteilung und die Leserschaft bestimmter Titel aufzeigte. Das Ergebnis zeigte, das einige Merkmale von schwierig zu erreichenden Zielgruppen beeinflusst werden und einige nicht. Dazu waren jedoch sieben Wiederholungsanrufe notwendig.

2. Methode - Die Möglichkeit

Die Haupt-Leserschaftsuntersuchung von PMB beinhaltet eine Wiederholungsbefragungskomponente. Diejenigen Personen, die an der Hauptuntersuchung teilnahmen, werden später gefragt, ob sie an weiteren Untersuchungen teilnehmen. Dadurch können Kunden weitere Befragungen vornehmen und die Daten können mit der ersten Untersuchung kombiniert werden.

Die PMB 2000 Untersuchung umfasste 24.072 Befragte. 1.867 davon hatten falsche Anschriften oder gaben an, nicht wieder befragt werden zu wollen. 22.205 Befragte erhielten einen 12-seitigen Fragebogen mit einem Incentive von \$ 5. Das Mailing beinhaltete daneben einen personalisierten Brief und einen frankierten Rückumschlag. Denen, die bis zu einem bestimmten Datum nicht geantwortet hatten, wurde noch einmal ein Fragebogen zugeschickt. Mehr Versuche wurden nicht gemacht. Insgesamt wurden 6.524 Fragebögen zurückgeschickt. Weitere 1.639 kamen unzustellbar zurück. Die Ausschöpfung betrug somit 32%. Die Daten wurden mit den Daten der Original PMB Leserschaftsuntersuchung zusammengefasst. Diese kombinierte Datenbasis gibt uns eine einzigartige Möglichkeit, die Charakteristiken der Verweigerer zu betrachten. Andere Forschungen haben gezeigt, dass Personen, die an Wiederholungsbefragungen teilnehmen, sich ähnlich verhalten, wie diejenigen, die zum ersten Mal befragt werden.

Unterschieden sich also die Befragten, die bei der zweiten Untersuchung teilnahmen, von denen die Verweigerten?

3. Analyse - Die Einblicke

Wir erwarten (durch andere bekannte Ergebnisse, z.B. von Cable/Jennings und Apple), dass die Verweigerer älter sind und keine Kinder haben. Die Original PMB Datenbasis von 24.052 war unsere Grundgesamtheit. Unterschieden sich also die 6.524 Wiederholungsbefragten von den 24.072? Der Vergleich wurde auf Basis von Indizes dargestellt. Die Original-Datenbasis von 24.072 wurde gleich 100 gesetzt. In der Original-Studie betrug z.B. der Männer-Anteil 47,1%, in der Wiederholungsbefragung 43,4% = Index 92,0. D.h. Männer waren in der Wiederholungsbefragung unterrepräsentiert.

3.1 Demographie

Alter	Index
12-17 J.	100,6
18-24 J.	64,4
25-34 J.	73,9
35-49 J.	99,1
50-64 J.	126,8
65 J. u. älter	116,5

Geschlecht	Index
Männer	92,0
Frauen	107,1

Wie in vielen Untersuchungen, wo keine Anstrengung unternommen wird, Wiederholungsanrufe durchzuführen, werden junge Leute unterrepräsentiert und ältere überrepräsentiert. Ebenfalls unterrepräsentiert werden Männer. Dieses Ergebnis zu Alter und Geschlecht geht einher mit der Mobilität der Zielgruppen.

Betrachtet man Bildung und Einkommen zeigen sich ausgewogene Ergebnisse, mit Ausnahme für die Gruppe mit Einkommen von \$ 75.000 und mehr und Befragte mit hoher Bildung.

(Tab. mit Indizes für Bildung und Einkommen)

(Tab. mit Berufstätigkeit und Beruf)

Auch in den Bereichen Berufstätigkeit und Beruf bleiben die Ergebnisse konsistent zu den Ergebnissen von Cable/Jennings und Apple. Zielgruppen, die schwer zu erreichen sind, wie Berufstätige oder Manager, antworten weniger häufig als andere.

In Kanada spielt außerdem die Region und die Sprache eine Rolle. Oftmals wird das Augenmerk auf die großen städtischen Gebiete gerichtet und diese sind mehrheitlich englischsprachig. Die Ergebnisse in der nachfolgenden Tabelle sind so erwartet worden und konsistent. Die größten urbanen Regionen in Kanada sind Quebec, Ontario und British Columbia, die hier niedrigere Indizes aufweisen. Ebenfalls Alberta, wo es auch zwei große urbane Gebiete gibt. Reflektiert das die verfügbare Zeit (der Befragten) oder die Mobilität in diesen Gegenden? Die Ausschöpfung für die zwei offiziellen Sprachen (englisch und französisch) ist nahezu identisch.

(Tab. mit Region, Ortsgröße, Sprache)

Die letzte Tabelle mit demographischen Merkmalen zeigt die Ergebnisse für Familienstand und Personen im Haushalt.

(Tab. mit Familienstand, Haushaltsgröße, Kinder im Haushalt)

Hier sind die Ergebnisse zu denen von Cable/Jennings und Apple nicht konsistent. Personen mit kleinen Kindern antworten weniger auf Befragungen und erfordern somit eine größere Anstrengung um ihre Teilnahme zu erreichen. Die Ergebnisse decken sich jedoch

mit der Ausschöpfung nach der Zahl der Anrufe, die in San Francisco vorgestellt wurden. Danach hatten viele Befragte, die erst nach wiederholten Anrufen teilnahmen, Kinder im Haushalt.

Zusammenfassend kann festgestellt werden, dass wiederholte Anrufe zur Erhöhung der Ausschöpfung helfen, um folgende Zielgruppen zu erreichen:

1. junge Befragte
2. Männer
3. höhere Einkommen / gehobene Berufe

Andere Zielgruppen zeigen annehmbare Ergebnisse:

1. Haushaltseinkommen (mit Ausnahme der o.e. höheren Einkommen ab \$ 75.000)
2. Bildung
3. Sprache

Während diese Ergebnisse zeigen, dass bestimmte Zielgruppen mehr Anstrengung verlangen, um sie zu erreichen, sagt es noch nichts aus, ob sich ihre Produktverwendung oder ihr Leseverhalten unterscheiden.

3.2 Produkte und Dienstleistungen

In der nachfolgenden Tabelle werden Finanzprodukte und Dienstleistungen dargestellt.

(Tab. mit Finanzprodukten und Dienstleistungen)

Die Nutzung von Telefonbanking, Kreditlinienbesitz, Pfandbrief- / Hypotheken-Besitz ist ähnlich für Befragungsteilnehmer und Verweigerer. Lebensversicherungs- und Investmentfonds-Inhaber sind mehr bei den Teilnehmern vertreten; ebenfalls diejenigen, die Grundbesitz besitzen oder anschaffen wollen. Je mehr Kreditkarten jemand besitzt, desto eher nimmt er an der Befragung teil.

(Tab. mit Kreditkartenbesitz)

Das gleiche gilt für PKW-Besitz.

(Tab. mit PKW-Besitz)

Weitere Ergebnisse für Besitzer von Fotoapparaten, Computer und Telefon / Anrufbeantwortern.

(Tab. mit Fotoapparaten, Computer, Telefon)

3.3 Media-Profile

Zunächst wurde das allgemeine Media-Profil betrachtet. Ausgewiesen werden die höchste, die mittlere und die niedrigste Gruppe von fünf gebildeten Nutzergruppen.

Media-Profile - Fünfereinteilungen

Zahl der Ausgaben pro Monat	Index
hoch	112,1
mittel	99,6
niedrig	79,1

TV-Fünftel	Index
hoch	112,5
mittel	105,5
niedrig	84,9

Radio-Fünftel	Index
hoch	101,0
mittel	101,5
niedrig	92,6

Interessant ist, dass die niedrigen Fünftel signifikant weniger bei der Wiederholungsbefragung antworten. Das bestätigt Ergebnisse der demographischen Auswertung. Je aktiver die Personen sind, desto weniger Zeit wenden sie für Medien auf (Print und TV). Beim Radiohören ist der Abstand nicht so deutlich.

Für Tageszeitungen treffen nicht die gleichen Ergebnisse zu. Es gibt nur geringe Unterschiede zwischen Befragungsteilnehmern und Verweigerern.

Media-Profile - Tageszeitungsleserschaften

	Index
Gestern gelesen	104,5
Samstag gelesen	106,6
Sonntag gelesen	104,8
Regionale Tageszeitung gelesen	106,2

Da die vorhergegangenen Daten für Print-Medien geringere Anteile bei den Wenig-Lesern zeigen, stellt sich die Frage, ob das auch für Einzel-Titel zutrifft. Das kann verneint werden. Es gibt zwischen den Einzel-Titeln kein einheitliches Bild. Ausgewählte Titel werden nachfolgend dargestellt.

Media-Profile - Tageszeitungsleserschaften

Gattung	Index
Wirtschaft / Finanzen	
Titel A	82,4
Titel B	98,3
Titel C	95,9
Titel D	99,7
Titel E	105,9
In Flight	Index
Titel A	90,3
Titel B	96,6
Unterhaltung	Index
Titel A	81,7
Titel B	90,0
Titel C	92,3
Titel D	98,6
Titel E	109,6
Titel F	111,1
Titel G	111,7
Titel H	125,4
General Interest	Index
Titel A	96,3
Titel B	112,6
Titel C	115,2
Titel D	127,7
Titel E	130,4

Wirtschafts-Titel haben mit einer Ausnahme einen Index unter 100. Das korreliert mit Berufstätigkeit und gehobenen Berufen, die weniger an der Befragung teilgenommen haben. Bei den Unterhaltungs-Zeitschriften sind die mit einem Index über 100 TV-Programm-Zeitschriften und die anderen Kino- oder Theater-Zeitschriften, d.h. die mobilen Personen, die nicht zu Hause fernsehen, werden weniger erreicht. Bei den General Interest-Titeln haben bis auf einen (ein Außer-Haus-Aktivitäten-Magazin) alle Indizes über 100. D.h. Personen, die verweigert haben, lesen diese Titel weniger.

(Tab. mit Titeln für Wohnen, Frauen-Zeitschriften, Nachrichtenmagazine - alle Titel mit Indizes über 100)

Für diese Titel würde die Reichweite sinken, wenn die Ausschöpfungsquote erhöht wird. Es wird deutlich, dass einige Titel durch eine geringe Ausschöpfung profitieren und andere nicht.

4. Zusammenfassung

Es wird gezeigt, dass es zwischen den Befragungsteilnehmern und denen, die verweigern definitive Unterschiede gibt. Wie erwartet, sind dies deutliche Unterschiede beim Geschlecht und beim Alter. Bei Einkommen, Bildung und Sprache sind die Unterschiede geringer.

Bei der Produktverwendung gibt es gemischte Ergebnisse. Für einige Produkte ist die Nutzung höher, für andere geringer. Es scheint so, dass Produkte, die mehr von mobilen Personen genutzt werden, unterrepräsentiert werden.

Bei Tageszeitungen gibt es nur geringe Unterschiede. Es gibt jedoch deutliche Unterschiede bei den Zeitschriften und beim TV (weniger beim Radio). Verweigerer konsumieren weniger Medien als Befragungsteilnehmer. Daher könnte man allgemein sagen: Je höher die Ausschöpfungsquote, desto niedriger die Reichweite. Das gilt jedoch nicht für alle Einzel-Titel, von denen einige durch hohe Ausschöpfungsquoten profitieren. Zusammenfassend kann festgestellt werden, dass weiterhin große Anstrengungen unternommen werden müssen, um die Ausschöpfungsquote zu halten bzw. zu erhöhen, um die "wahre Welt" abbilden zu können.

Breakthrough in newspaper section research in the Netherlands

Durchbruch für Tageszeitungs-Bücher(Sektionen)-Forschung in den Niederlanden

Ton Jansen, R&M, Research and Marketing B.V./Vincent van den Berg, Cebuco

Einleitung

Für Anzeigen-Platzierungen in Tageszeitungen muss man den gleichen Preis zahlen - egal, ob der Platz auf Seite 2 im 1. Buch ist oder auf Seite 6 im 9. Buch. Ist das fair? Es kommt drauf an!

In dem Fall, wo die Anzeigenkontaktchance auf beiden Plätzen gleich ist, kann man sagen, dass es fair ist. Anzeigenkontaktchance ist das Schlüsselwort. Werbungtreibende und Media-Agenturen vermuten, dass Leser von Tageszeitungen nicht alle Bücher (Sektionen) lesen. Es wird vermutet, dass, je weiter hinten das Buch steht, desto geringer die Anzeigenkontaktchancen sind. Und deshalb sollten die Anzeigenpreise für diese Platzierungen geringer sein als für Bücher mit höheren Anzeigenkontaktchancen. Wenn die Vermutungen richtig sind, wäre das OK so.

Aber wie kann man herausbekommen, ob diese Annahmen für jede Tageszeitung an jedem Tag richtig sind? Wird die Samstags-Ausgabe genauso gelesen wie eine Werktags-Ausgabe? Wahrscheinlich nicht. Die einzige Lösung ist die Durchführung einer Untersuchung, die alle Tageszeitungen über alle Wochentage beinhaltet. Und die wurde in den Niederlanden durchgeführt. Cebuco übernahm die Initiative und begann ein Forschungsmodell zu entwickeln. Das musste reliabel, valide und für alle holländischen Tageszeitungen gültig sein und das innerhalb kurzer Zeit, zu akzeptablen Kosten. Bevor das Modell fertig gestellt werden konnte, wurden drei Untersuchungen durchgeführt:

1. Eine Vor-Untersuchung in der zweiten Hälfte 1998
2. Eine Pilot-Studie in der ersten Hälfte 1999
3. Ein ausführlicher Modell-Test in der zweiten Hälfte 1999 und der ersten Hälfte 2000

1. Vor-Untersuchung, 2. Hälfte 1998

Das Ziel dieser Voruntersuchung war es, herauszufinden, ob die Daten per Telefonbefragung oder per schriftlicher Befragung die o.a. Anforderungen erfüllen. In den frühen 90er Jahren wurden in den Niederlanden eine Reihe von Tageszeitungs-Sektionen-Untersuchungen face-to-face durchgeführt. Die face-to-face Methode wurde von Cebuco von vornherein ausgeschlossen, da weder zeitliche noch kostenmäßige Anforderungen erfüllt werden. In die Voruntersuchung wurden zwei Tageszeitungen einbezogen: eine lokale und eine nationale Morgenzeitung. Von beiden Titeln wurden die Donnerstag- und die Samstag Ausgaben getestet. Die Grundgesamtheit waren Leser der jeweiligen Ausgabe ab 13 Jahren in Abonnentenhaushalten. In den Niederlanden sind 90% der Tageszeitungs-Auflage Abo-Exemplare und nur 10% gehen in den Einzelverkauf.

Für jede der vier Gruppen (zwei Wochentage von zwei Tageszeitungen) wurden aus der Abonentendatei zwei Random-Stichproben gezogen. Eine Stichprobe wurde per Telefon befragt und die andere erhielt einen schriftlichen Fragebogen. Die Telefonbefragung fand teilweise an dem Erscheinungstag der Ausgabe statt und teilweise am Tag danach. Haupterhebungspunkt war, zu untersuchen, ob die Leser jede einzelne Seite der Ausgabe aufgeschlagen haben.

Der schriftliche Fragebogen wurde zusammen mit der jeweiligen Ausgabe zugeschickt. Jede einzelne Seite musste von dem Befragten markiert werden: mit einem "J" für ja, habe ich aufgeschlagen und mit einem "N" für nein, habe ich nicht aufgeschlagen. Die Netto-Stichprobengröße betrug 100 bis 125 Fälle für die Telefoninterviews und 135 bis 160 Fälle für die postalische Erhebung. Die Ergebnisse der Voruntersuchung, die für die jeweilige Methode sprachen, waren:

Pros der postalischen Methode

- Validität der Ergebnisse
- Möglichkeit alle Seiten der Ausgabe zu erheben (war für Telefon-Interviews zu lang)
- Kosten pro Befragten
- Ausschöpfungsquoten (Postalisch mit Incentive: 73%, Telefon: 23% im Durchschnitt)

Pros der telefonischen Methode

- geringerer logistischer Aufwand
- Erhebungszeit

Die Validität der Ergebnisse und die Kosten pro Befragten waren zwei wichtige Gründe für die Entscheidung die Pilot-Studie postalisch durchzuführen.

2. Pilot-Studie, 1. Hälfte 1999

Die Pilot-Studie sollte helfen, die postalische Erhebung zu optimieren. Insgesamt wurden acht Gruppen gebildet:

- Registrierung, ob die Seiten aufgeschlagen wurden, in der Ausgabe oder in einem separaten Fragebogen
- Benutzte Skala, um das Aufschlagen der Seiten festzuhalten: 2er-, 3er-, 5er-Skala
- Anzahl der Bücher (Sektionen), die erhoben werden (alle oder - zufällig ausgewählt - die Hälfte)
- Abonnenten vs. Einzel-Käufer

In die Pilot-Studie wurde nur eine Ausgabe einbezogen. Ausgewählt wurde die Samstags-Ausgabe der national verbreiteten Tageszeitung De Telegraaf.

Die Abonnenten wurden aus dem Datenbestand des De Telegraaf ausgewählt, die Einzelkäufer wurden an zehn Verkaufsstellen am Samstag rekrutiert. Die Netto-Stichprobengröße betrug 75 bis 110 Interviews pro Gruppe. Bei der Überprüfung der Ergebnisse wurde folgendes betrachtet:

- Validität
- Befragten Antworten
- Interne Antworten (Antworten für alle Seiten)
- Tiefe der Informationen (nur Kontaktchancen oder auch Lesevorgänge)
- Kosten
- Durchführbarkeit

Die Methode, die am vielversprechendsten aussah wurde für den ausführlichen Test ausgewählt.

3. Ausführlicher Test, 2. Hälfte 1999 / 1. Hälfte 2000

Für die Registrierung, ob die Seiten aufgeschlagen wurden oder nicht, wurde folgende Methode gewählt:

- Registrierung in einem separaten Fragebogen
- Nutzung einer 5er-Skala (mit 4 Antwortmöglichkeiten, die die Kontaktchance qualifizieren)
- Registrierung aller Bücher und Seiten in einer Ausgabe

Die 5er-Skala beinhaltete folgende Kategorien:

- angeschaut oder gelesen: (fast) alles auf der Seite
- angeschaut oder gelesen: ungefähr die Hälfte dieser Seite
- angeschaut oder gelesen: etwas auf dieser Seite
- Seite wurde aufgeschlagen, es wurde jedoch nichts angeschaut oder gelesen
- Seite wurde nicht aufgeschlagen

Wie wurden die Kontaktchancen definiert? Die 5er-Skala wurde nach folgendem Schema recodet:

angeschaut oder gelesen: (fast) alles	Seite aufgeschlagen
angeschaut oder gelesen: ungefähr die Hälfte	Seite aufgeschlagen
angeschaut oder gelesen: etwas	Seite aufgeschlagen
Seite wurde aufgeschlagen, nichts angeschaut oder gelesen	Seite aufgeschlagen
Seite wurde nicht aufgeschlagen	Seite nicht aufgeschlagen
weiß nicht / keine Angabe	Seite nicht aufgeschlagen

Alle Befragten, die keine einzige Seite als aufgeschlagen angaben, wurden aus der Stichprobe entfernt. Als nächstes entschieden wir, dass, wenn nur die linke oder nur die rechte Seite als aufgeschlagen angegeben wurde, auch die gegenüberliegende Seite als aufgeschlagen galt, da Doppelseiten nicht einzeln aufgeschlagen werden können.

Die durchschnittliche Kontaktchance wurde nach folgender Formel berechnet:

$(1 \times \text{Chance für Titelseite des Buches} + 2 \times \text{Chance für Doppelseite 1} + 2 \times \text{Chance für Doppelseite 2} + 2 \times \text{Chance für Doppelseite i} + 1 \times \text{Chance für Rückseite des Buches}) / \text{geteilt durch die Anzahl der Seiten des Buches.}$

Bei Nutzung der 5er-Skala wurde es möglich, nicht nur Kontaktchancen anzuschauen, sondern auch Lesevorgänge für jede Seite. Der Test wurde für zwei lokale und zwei nationale Tageszeitungen durchgeführt. Eine der Lokalzeitungen war eine Abendzeitung, die anderen Morgenzeitungen. Es wurden zwei ähnliche Wellen im Nov./Dez. 1999 und im März/April 2000 durchgeführt. In beiden Wellen wurden die Ausgaben aller Wochentage einbezogen.

Da die Pilot-Studie nur geringe Ergebnisunterschiede zwischen Abonnenten und Einzelkäufern zeigte und der EV-Anteil in den Niederlanden bei Tageszeitungen sehr gering ist, wurde der ausführliche Test nur mit Abonnenten durchgeführt. Die Brutto-Stichprobe betrug 250 Adressen pro Tageszeitung pro Tag. Jeder Befragte nahm nur an einem Tag an dem Test teil. Mit einem Vorankündigungsbrief des Chefredakteurs wurde die Untersuchung eine Woche vorher angekündigt. In 83% wurde der Fragebogen am Tag nach dem Erscheinen der Ausgabe empfangen. Die Hälfte der Stichprobe erhielt ein Incentive in Form eines Lotterieloses (Wert ca. \$ 2). Falls mehr als ein Leser der Ausgabe im Haushalt vorhanden war, wurde als erstes eine Zufallsauswahl durchgeführt. Dazu musste ein Haushaltsmitglied alle Leser auflisten.

Die Ergebnisse des Tests

Allgemein

Die Ergebnisse erscheinen plausibel und valide. Die Struktur der Ergebnisse stimmt mit denen anderer Untersuchungen überein.

Stabilität der Ergebnisse

Die Kontaktchancen vergleichbarer Bücher, z.B. Finanzen, unterschieden sich von Tageszeitung zu Tageszeitung. Das bedeutet, dass wir jede einzelne Tageszeitung in die Sektions-Untersuchung einbeziehen müssen.

(Graph. mit Ergebn. Finanz-Bücher einzelner Tageszeitungen)

Die Kontaktchancen vergleichbarer Bücher eines Tageszeitungs-Titels unterscheiden sich von Tag zu Tag. Das heißt, wir müssen tatsächlich alle Wochentage in die Untersuchung einbeziehen.

(Graph. mit Ergebn. Sport- und Finanz-Buch über die Wochentage für eine Tageszeitung)

- Die Kontaktchancen für vergleichbare Bücher (Sektionen) einer Tageszeitung, die in beiden Wellen enthalten waren, unterschieden sich nur wenig. Die durchschnittliche Differenz betrug nur 3,3%. D.h. es besteht keine Notwendigkeit, die Untersuchung mehr als einmal pro Jahr durchzuführen. Allerdings nur unter der Voraussetzung, dass die Struktur der Zeitung nicht geändert wird.
- Die Kontaktchancen und mehr noch das Leseverhalten für Einzelseiten, waren von einer Welle zur nächsten sehr unterschiedlich. Das verwundert nicht, da das sehr abhängig ist von dem jeweiligen Inhalt der Seite.

Trennschärfe der Ergebnisse

- Im allgemeinen sind die Kontaktchancen für spezifische Zielgruppen höher als im Durchschnitt. Das sagt einiges über die Validität der Ergebnisse.

Ausfüllzeitpunkt des Fragebogens

- Die Zustellung durch die holländische Post ist zufriedenstellend. 83% erhielten den Fragebogen am nächsten Tag nach Absendung.
- Es gibt keine systematische Abhängigkeit zwischen den Kontaktchancen und dem Zeitpunkt des Fragebogenausfüllens.

Incentive

- Ein Incentive erhöht die Ausschöpfungsquote: im Durchschnitt 67% vs. 44%.
- Ein Incentive hat keinen Einfluß auf die Struktur.
- Es gibt keine Auswirkungen des Incentives auf die Höhe der Kontaktchancen.

Modell-Durchführbarkeit

- Die Durchführung der Untersuchung war problemlos. R&M ist in der Lage die Untersuchung für alle 30 holländischen Tageszeitungen gleichzeitig durchzuführen.

In der ersten Hälfte dieses Jahres wurde in zwei Wellen jede Ausgabe einer Woche innerhalb von einem Monat untersucht. Dabei wurden Abonnenten aller holländischen Tageszeitungen einbezogen. Für De Telegraaf wurde zusätzlich eine Stichprobe mit Einzelkäufern einer Samstags-Ausgabe, untersucht.

4. Gespräche mit den Markt-Partnern

Nachdem 1999 bereits die PMA (Platform for Media Agencies) eine Tageszeitungsuntersuchung für Tageszeitungs-Bücher (Sektionen) mit der face-to-face Methode durchführen ließ, gab es jetzt zwei Untersuchungen auf dem Markt. Ein Komitee von Media Agenturen, Werbungtreibenden und Tageszeitungs-Vertretern begutachtete beide Methoden hinsichtlich der Reliabilität, der Validität und der Kosten für die Durchführung. Beide Methoden sind reliabel und valide. Hinsichtlich der Kosten ist jedoch die postalisch schriftliche Methode günstiger. Damit wurde beschlossen, eine große Untersuchung für alle holländischen Tageszeitungen mit der postalischen Methode durchzuführen.

5. Einige Ergebnisse

Erste Ergebnisse werden auf aggregiertem Niveau vorgestellt. D.h. die Ergebnisse werden als gewichteter Durchschnitt über alle 30 Tageszeitungen ausgewiesen.

Ein wichtiges Ergebnis ist, dass es nicht stimmt, dass die Kontaktchancen für die hinteren Bücher sinken. In einer Zahl von Fällen haben hintere Bücher höhere Kontaktchancen-Scores als Bücher, die weiter vorne stehen.

(Graph. mit Kontaktchancen-Scores für Bücher über die Wochentage)

Ein anderes wichtiges Ergebnis zeigt, dass Abonnenten mit Interesse an bestimmten Gebieten für die Bücher, die diese Themen behandeln, höhere Kontaktchancen haben als der Durchschnitt. Das ist selbstverständlich nicht neu. Die Tatsache jedoch, dass es auf

breiter Basis für alle holländischen Tageszeitungen gemessen wurde, öffnet neue Möglichkeiten dieses als Zielgruppenmerkmal in der Planung zu nutzen.

(Graph. mit Kontaktchancen von Interessierten an bestimmten Gebieten)

6. Qualitäts-Index-System (QIS)

Ende 1999 wurde ein neues Anzeigen-Preis-System für Tageszeitungen entwickelt. Das neue Modell basiert auf Quality Rating Points, so wie sie in Dänemark angewandt werden. Die Philosophie dahinter heißt, dass Tageszeitungen nicht nur Quantität in Form von Millimetern anbieten, sondern auch Qualität in Form von Farbigkeit, inhaltlichem Umfeld und Reichweite. Diese Parameter werden im dänischen Modell auf objektive Art bewertet und in die Preisgestaltung einbezogen.

In den Niederlanden wird die neue Währung Quality Rating Points, die jetzige Währung, die Millimeter, ersetzen. Die Quality Rating Points werden durch die Multiplikation der Größe in Millimetern mit Auflagenindizes (pro Tag), Farbigkeit und Buch-Reichweite (Sektions-Reichweite), errechnet. Der größte Unterschied zum dänischen QRP-Modell ist, dass aufeinanderfolgende Buchungsklassen mit dem Preis einer Anzeige verbunden werden. Diese Buchungsklassen bestimmen die Art des Discounts und Privilegien, die ein Werbungtreibender erhält.

Quality Index System (QIS):
$\begin{aligned} \text{Quality points} &= (\text{Anzeigen-})\text{Größe (in mm)} + \text{Farbigkeit (mm oder Index)} \times \text{Tagesauflage (Index)} \times \text{Kontaktchance (Index)} \\ &\quad \times \\ &\quad \text{Schaltungen} \\ &\quad \times \\ &\quad \text{Preis pro Quality Point} \\ &\quad \times \\ &\quad \text{Buchungsklasse} \\ &= \\ &\quad \text{Preis pro Anzeige oder Kampagne} \end{aligned}$

(Begründung des Anzeigen-Modells)

7. Verwendung der Bücher-Forschung (Section research) innerhalb des Modells

Alle Kontaktchancen-Indizes werden für jeden Wochentag einzeln errechnet. Für die Kontaktchancen-Indizes pro Buch (Sektion) bilden die Durchschnitte der gesamten Tageszeitung die Basis. Die durchschnittliche Kontaktchance pro Tageszeitung ist der Durchschnitt über alle Bücher (Sektionen) der Tageszeitung. Für jeden Tag wird die Kontaktchance über die Gesamt-Tageszeitung = 100 gesetzt.

Im Quality Index System unterscheiden wir zwei Arten von Büchern (Sektionen):

1. Ein physisch einzelnes Buch (Sektion) innerhalb der Tageszeitung.
2. Eine Reihe von Seiten innerhalb derselben Tageszeitung mit dem gleichen Thema, z.B. Sport, Wirtschaft usw. Besser beschrieben als Themen-Sektion.

Beide Indizes werden auf die gleiche Art errechnet:

Der Kontaktchancen-Index pro Buch (Sektion) pro Tag = Kontaktchancen-Score pro Sektion pro Tag / Kontaktchancen-Score pro Tageszeitung pro Tag x 100.

Der Kontaktchancen-Index pro Themen (-Sektion) pro Tag = Kontaktchancen-Score pro Themen-Sektion pro Tag / Kontaktchancen-Score pro Tageszeitung pro Tag x 100

$$\begin{aligned} \text{z.B. Kontaktchancen-Score pro Tageszeitung am Montag} &= 0.94 \\ \text{Kontaktchancen-Score des 2. Buches (2. Sektion) /} & \\ \text{des Themas Sport am Montag} &= 0.90 \\ &= 0.90 / 0.94 \times 100 = 96 \end{aligned}$$

8. Neue differenzierte Preise

Die Ergebnisse der Formel werden mit den neuen Preisen multipliziert. Prinzipiell ist es der gleiche Preis für alle Werbungtreibenden. Man könnte jedoch unterschiedliche Preise einführen, z.B. für Stellenanzeigen oder für bestimmte Zielgruppen. Der neue Preis wird durch den Discount abgesenkt, der in den Buchungs-Klassen gewährt wird.

Die Reaktion der Media-Agenturen und der Werbungtreibenden auf dieses Modell war sehr positiv. Es wird ein Zugang zu den Daten ermöglicht, um eigene Analysen und Media-Planung zu ermöglichen. Dazu wurden einheitliche statistische Regeln vorgegeben, um die Validität und Reliabilität der Ergebnisse sicherzustellen.

Die Regeln beinhalten:

- A. Kontaktchancen-Scores pro Gesamt-Tageszeitung können genutzt werden
- B. Kontaktchancen-Scores pro Buch (Sektion) pro Titel können genutzt werden
- C. Kontaktchancen-Scores pro Themen-Sektion pro Titel können genutzt werden
- D. Kontaktchancen-Scores pro Seite pro Titel können nur genutzt werden, wenn mindestens vier ähnliche Fälle pro Titel verfügbar sind.
Das heißt: n=600 (minimum von 4 Seiten) - Die Seiten haben den gleichen Themen-Code.

Die verschiedenen Lese-Scores können nicht genutzt werden. Die präsentierten Ergebnisse müssen dokumentiert werden.

9. Einfluß

Durch die neuen Planungsgrundlagen werden Tageszeitungen attraktiver. Bei der Planung muss jetzt auch auf die richtige Platzierung geachtet werden. Verleger erhalten mehr Einblick in die Präferenzen ihrer Leserschaft.

10. Die Zukunft

Das neue Modell wird im November 2001 etabliert sein. Im Januar und April 2002 wird ein Finetuning vorgenommen. Ab Januar 2002 werden die Vorbereitungen für die Implementation, die für Januar 2003 vorgesehen ist, starten. Die Hauptaufgabe ist, Software und Computer-Systeme zu entwickeln und die Anwender zu schulen.

The communication power of customer magazines

Die Kommunikations-Kraft von Kundenzeitschriften

Fred Bronner, Veldkamp/Dick Suèr, WNCP

Zusammenfassung

In dem Vortrag werden die Ergebnisse von Untersuchungen aus den letzten 20 Jahren zum Thema Kundenzeitschriften zusammengefasst. Das Hauptergebnis kann als 5-A-Modell beschrieben werden: Audience (Leserschaft), Attention (Aufmerksamkeit), Attitude to the brand (Einstellung zur Marke), Adherence (Leser-Blatt-Bindung), Active Experience (aktive Erfahrungen). Basierend auf der Literatur, in der Forschungsergebnisse beschrieben werden, kann den Kundenzeitschriften eine rosige Zukunft vorhergesagt werden.

Einleitung

Kundenzeitschriften spielen zunehmend eine wichtigere Rolle in neuen personalisierten Marketing Strategien. Das wird auch dadurch unterstrichen, dass wir z.Zt. eine Untersuchung zur Entwicklung eines Kundenmagazins für Internet-Provider durchführen. Mit dem Kundenmagazin soll das Wir-Gefühl gestärkt werden. Wenn jemand das vor drei Jahren vorhergesagt hätte, hätten die Leute gelacht.

Der Weltverband der Kundenzeitschriftenverleger (World Network of Custom Publishers - WNCP) entschied letztes Jahr ein extensives Forschungsprogramm, das aus zwei Teilen besteht:

- a) Desk-Research. Hierbei sollen die Stärken und Schwächen von Kundenzeitschriften als Mediagattung unter Berücksichtigung so vieler aktueller Untersuchungen wie möglich, aufgezeigt werden.
- b) Nutzung des Desk-Research-Inputs um ein internationales Standard-Forschungsinstrument für Kunden zu entwickeln.

In dem Vortrag werden die Ergebnisse des Desk-Research präsentiert.

Die 5 'A's der Kundenzeitschriften (Graph. Darstellung der 5 'A's)

Audience (Leserschaft):

Ein Trend der Zukunft werden Veröffentlichungen auf Datenbank-Basis sein. Zunehmend tritt eine Organisation in den Dialog mit dem Kunden ein. Die Information wird in einer Datenbank gespeichert und öffnet den Weg für kundenorientierte Kommunikation.

Attention (Aufmerksamkeit):

Durch Informations-Überfrachtungen wird es zunehmend schwieriger, Aufmerksamkeit zu erreichen. Kundenzeitschriften sind beim Erreichen der Zielgruppe sehr erfolg-

reich. Das lässt sich durch viele Untersuchungen belegen.

Attitude to brand (Einstellung zur Marke):

Hierunter werden Ergebnisse präsentiert, die zeigen, dass Kundenzeitschriften zu a) Marken-Erinnerung, b) Marken-Kenntnis, c) Marken-Image, beitragen.

Adherence (Leser-Blatt-Bindung):

Kundenzeitschriften schaffen Bindung, Vertrauen, Wir-Gefühl, Loyalität und können in Kunden-Loyalitätsprogrammen eingesetzt werden.

Active Experience (aktive Erfahrungen):

Kundenzeitschriften haben die Funktion, Erfahrungen zu vermitteln: Tipps geben (beihilflich sein), Ausdruckskraft (z.B. um bestimmte Lifestyles zu vermitteln), unterhalten.

In unserer Studie wollen wir nicht auf Verhaltensaspekte eingehen. Nachfolgend werden die 5 'A's ausführlich beschrieben. Dabei werden auch die jeweils herangezogenen Untersuchungen zitiert.

Die Zukunft

Die Zukunft für Kundenzeitschriften sieht vielversprechend aus. Ausgehend von dem o.a. 5-A-Modell, zeigen die meisten der zitierten Untersuchungen auch Zukunftsaspekte auf.

Audience (Leserschaft) und die Zukunft

Die meisten Untersuchungen zeigen ein Ansteigen der zukünftig zur Verfügung stehenden Informationen über Kunden. Zudem wird die Nutzung dieser Informationen anwachsen. Personalisierte und auf Verbraucher zugeschnittene Medien haben eine rosige Zukunft. Mit anderen Worten, die Zahl der auf die Verbraucher maßgeschneiderten Zeitschriften wird zunehmen. Die Datenbank-Technologie wird diese Entwicklung unterstützen und die notwendigen Konditionen dafür schaffen. Verleger und Kunden müssen jedoch mit selektiveren Sektionen umgehen lernen. Zur Zeit wird noch zuviel Bedeutung auf Name, Adresse und Wohnort gelegt.

Attention (Aufmerksamkeit) und die Zukunft

Aufgrund der Informations-Überfrachtung und der zunehmenden Diskrepanz zwischen Angebot und Nachfrage, werden die Verbraucher die Angebote mehr und mehr selektiv nutzen. Forschungen haben gezeigt, dass die meisten Menschen nur das lesen, was mit ihrer eigenen Meinung übereinstimmt. Dabei haben Kundenzeitschriften eine Trumpfkarte in der Hand. Eine Kundenzeitschrift wird aufgrund einer bereits bestehenden (Kunden)Beziehung ausgesandt. Damit hat der Verbraucher bereits eine positive Einstellung gegenüber der Zeitschrift. Kundenzeitschriften spielen eine große Rolle in Loyalitäts-Programmen, die immer häufiger von den Firmen entwickelt werden. Alle Anstrengungen, die die Verbraucher-Loyalität einbeziehen, geben Kundenzeitschriften einen "Schub".

Active Experience (Aktive Erfahrungen) und die Zukunft

Die Literatur zeigt, dass Kundenzeitschriften am meisten gelesen werden, wenn persönliche Ziele befriedigt werden können, wenn Mehrfach-Funktionen erfüllt werden, oder wenn breite Erfahrungen mit dem Medium vorliegen (bietet Informationen, regt an, bietet Gesprächsstoff, bietet Tipps usw.) Die Zeitschrift muss informieren, unterhalten und Gelegenheit zum Vergleich und zur Identifikation bieten. Wenn der redaktionelle Inhalt dieses vermitteln kann, ist die Zukunft in diesem Punkt für Kundenzeitschriften ebenfalls gesichert.

Dieses kann zukünftig auch für eine Interaktion zwischen Kundenzeitschriften und dem Web bedeutend werden. Die Kundenzeitschrift wird eine zunehmende Rolle bei der Unterstützung der Website spielen. Dabei wird die Zeitschrift Informationen beinhalten, die "nice-to-know" ("gut zu wissen") sind und die Website wird Informationen, die "need-to-know" ("notwendig zu wissen") beinhalten. Auf der anderen Seite wird zunehmend beobachtet, dass Internet-Firmen (sogen. Dotcoms) Kundenzeitschriften einführen, um Informationen über den Gebrauch ihrer Website zu vermitteln.

Der nächste Schritt

Basierend auf diesem Desk Research entwickelten wir ein internationales Standard-Untersuchungs-Werkzeug für Kundenzeitschriften. Alle 'A's wurden in einigen Fragen operationalisiert. Im Sommer 2001 wurde der Fragebogen entwickelt und getestet. In der 2. Hälfte 2001 werden in einigen Ländern die ersten vergleichbaren Ergebnisse vorliegen. Der Kundenzeitschriften Monitor wird mehr und bessere Einblicke in Nischen-Leserschaften bringen.

Online Media Measurement: After the deluge

Online-Untersuchungen: Nach der Sintflut

Scott C. McDonald, PH.D., Condé Nast Publications

Hintergrund

Seit dem Symposium 1995 in Berlin wurden viele Vorträge über das Wachstum des Internet und des World Wide Web gehalten. Dabei wurden Parallelen zwischen Online-Nutzer-Untersuchungen und Untersuchungen über Nutzer anderer Medien aufgezeigt. Die Verwandtschaft zwischen Online und Print wurde dargestellt und der Gebrauch von Online-Methoden zur Durchführung von Media-Forschung. Ich für meinen Teil, habe eine detaillierte Kritik über die Methoden site-centric, ad-centric und user-centric vorgestellt. Das vorliegende Papier wird die Entwicklung hierzu in den letzten zwei Jahren aufzeigen.

Kein Konsens über die (ökonomischen) Termini für Online-Werbung

Wie 1999 berichtet, gibt es zwischen den Online-"Verkäufern" und Online-"Käufern" für Werbung keine Einigung, wie oder was als Maßstab gemessen werden kann. Ein Kompromiss sagt, es sollten Impressions (Nutzungen) und "Click-Through"-Raten (Antworten auf Werbung) herangezogen werden. Von den Werbungtreibenden wird das Internet immer noch als ein Antwortmedium angesehen, deshalb werden Bezahlungen (für Werbung) auf der Basis der Click-Through-Raten gemacht. Online-Anbieter haben es versäumt, die Werbungtreibenden davon zu überzeugen, dass das Online-Medium als reines "Markenpflege"-Instrument anzusehen ist.

Site-Centric

Hierbei werden die Seiten, die vom Original-Web-Server abgerufen werden, gezählt. Das ergibt einen kompletten Überblick über die Server-Aktivitäten. Ein weiterer Vorteil ist, dass alle Seiten-Abrufe gezählt werden, egal ob sie von Zuhause, Internet-Cafés, Arbeitsplatz usw. erfolgen, oder aus einem anderen Land oder von einer Nicht-Computer-Plattform (z.B. PDA, TV usw.).

Es gibt jedoch - wie bereits 1999 in Florenz erwähnt - auch eine Reihe von Problemen: Es wird nicht gemessen, ob die Site auch angekommen ist (der Abruf reicht). Abrufe über Proxy-Server oder Local Cache werden nicht erfaßt. Es zählt der Abruf der Site und nicht der Anzeigen. Man kann keine demographischen Daten der Nutzer liefern, da man nicht weiß, ob ein Nutzer dahintersteht. Es können auch Robots, Spider, Agents usw. sein.

Es hat sich in den letzten zwei Jahren nicht viel verändert; es gibt jedoch einige wenige Entwicklungen:

Site-Registration-Campaigns: Einige Sites haben sich entschlossen Nutzer-Registrierungen einzuführen, um damit auf den Nutzer (und seine demographischen Daten) zurückzukommen. Der Nutzer bekommt dafür einige Vergünstigungen, z.B. ein persönliches Aktien-Depot, kostenlose Drucke von digitalen Photos, Frequent Flyer Punkte usw. Große Sites,

wie Yahoo, behaupten, sie hätten Millionen registrierter Nutzer. Im Zusammenhang mit Cookies können sie teilweise detaillierte Profile ihrer Nutzer bieten. Da jedoch mehrere Personen einen Computer nutzen können (oder eine Person kann mehrere Computer nutzen) sind solche Kampagnen weit entfernt von akkuraten Nutzer-Messungen.

Harmonisierung der Methoden: Bereits 1999 wurde gefordert, allgemeine Standards für Server Log Data zusammenzustellen. Fortschritte sind in diesem Zusammenhang sehr, sehr langsam. Bis 2000 wurde kaum etwas getan, obwohl die Dotcom-Unternehmen viel Geld hatten. Erst Anfang 2001, als sich das Blatt wendete, beschloß das IAB PriceWaterhouse-Coopers zu beauftragen, eine Untersuchung bei den Mitgliedern durchzuführen, wie die Ausweisungs-Methoden für Site-Centric-Messungen aussehen. Ergebnisse liegen noch nicht vor.

Führung durch Beispiele:

CNN.com beauftragte das Media Rating Council (MRC) über ihre Website ein Audit durchzuführen. Damit erhofft man sich, Standards für Site-Centric- und Ad-Centric-Erhebungen durch die MRC vorgegeben zu erhalten. Das Audit ist noch nicht beendet.

Ad-Centric

Die Erhebung von Anzeigen war auf dem Florenz-Symposium noch ziemlich neu, nahm aber schnell zu. Seit Websites Inhalte von einem Server erhalten und die Anzeigen von einem anderen Server dazu kommen, sind die Aktivitäten der Ad-Server relevanter für die Werbungtreibenden. Für Ad-Centric-Messungen gilt das Gleiche wie für Site-Centric-Messungen. Man weiß nicht, wer die Werbung anschaut und kann ebenfalls Überschneidungen nicht ausweisen. Reichweiten und Frequenzen sind schwer zu definieren. Auch für Über- und Unterschätzungen gilt das Gleiche wie für Site-Centric.

Ad-Centric-Daten für Ausgleich nach der Kampagne

Da alle Beteiligten jetzt mit den Ergebnis-Berichten vertraut sind, sind diese Daten die meist genutzte Datenquelle für Online-Werbung. Sie werden als Brutto-Impressions genutzt, da Netto-Werte nicht messbar sind (s.o.).

Übertragung auf Ad Networks

1999 wuchsen Ad Networks als Drittanbieter. Das ermöglichte Werbungtreibenden die konsolidierte Nutzung und Abrechnung von Ad-Servern, brachte jedoch Schwierigkeiten mit den Medien-Sites, da diese keine dritte Partei akzeptierten, sondern die Anzeigen selber haben wollten. Außerdem wiesen die Ad Networks geringere Ergebnisse aus, als die Medien-Sites von ihren eigenen Logfile-Servern. Bevor das geklärt werden konnte, verschwanden viele Ad Networks wieder. Heute sind nur noch Double Click und 24/7 tätig, die auch als Service-Büros fungieren. Die großen Media-Sites haben heute mehr Einfluss auf die Anzeigen-Bestände als 1999, weisen aber auch zunehmend Ad-Centric-Ergebnisse aus.

User-Centric

Der User-Centric-Ansatz ist am vertrautesten für alle, die mit anderen Medien aufgewachsen sind, da dieses Modell den Erhebungen von Fernsehseherschaften ähnlich ist. Es werden große Verbraucher-Panel rekrutiert, und mit einer Software werden die Nutzungen gemessen. Wie in Florenz dargestellt, gibt es hier einige Vorteile gegenüber Site- und Ad-Centric. So wird z.B. die empfangene Site gemessen, unabhängig vom Caching, die demographischen Daten der Nutzer werden geliefert und es ist objektiv und unabhängig von den Media-Verkäufern. Das Verfahren hat jedoch auch Nachteile. Es stehen nur wenig Daten unterhalb der Top-Level-Domains zur Verfügung und damit ist es nicht vergleichbar mit der Ebene auf der gebucht wird, dem Sub-Domain-Level. Außerdem können Messungen außerhalb von Zuhause, die jedoch wichtig sind, nicht erfolgen. Die Arbeitsweise der Meter selbst ist nicht von unabhängigen Institutionen überprüft worden. Was hat sich seit 1999 getan?

Wettbewerb und Konsolidierung

MediaMetrix und NetRatings haben den Crash überstanden. Mediametrix ist mit Jupiter Communications zusammengegangen. NetRatings, die mehrheitlich Nielsen Media Research gehören, expandierten auf internationaler Ebene mit ACNielsen. Andere Anbieter gaben auf: PCData, NetValue.

Internationalisierung

Sowohl MediaMetrix als auch NetRatings expandierten in viele Länder und besetzten einige Plätze von NetValue. Diese Internationalisierung erfolgte meistens durch Joint-Ventures mit lokalen Forschungsinstituten, so dass die Panel-Rekrutierungsprozeduren nicht grenzüberschreitend einheitlich sind.

Diversifikation

Beide Anbieter haben ihre Angebote an die Kunden diversifiziert. MediaMetrix bietet durch Akquisition von AdRelevance Tracking-Daten für eine Vielzahl von Sites an. Ähnliches bietet auch NetRatings vom eigenen Panel an. Da NetRatings zu Nielsen gehört, entsprechen die Definitionen und Klassifizierungen dem Nielsen MonitorPlus Service.

Rechtsstreit

Am Anfang des Jahres stritt sich MediaMetrix mit NetValue um das Patent für das Meter-System. Das überraschte etwas, da ja beide Systeme angeblich so unterschiedlich sein sollten.

Herabsetzung

Das Platzen der Dot.com-Blase hat die Aufmerksamkeit, die den monatlichen Rankings zufiel, reduziert. Das ist wahrscheinlich gut, da es mindestens die Querelen um die Richtigkeit der Rankings reduziert hat.

Pop-Ups und Pop-Unders

Pop-Up-Werbung gab es schon 1999. Pop-Under-Werbung ist neu. Sie guckt unter dem geöffneten Bildschirm-Fenster hervor und kommt an die Oberfläche, wenn ein Surfer versucht, sie zu schließen. Oftmals ist diese Werbung in langen Strings verbunden, so dass der

Nutzer die Site neu aufruft, da er genervt ist. Damit wird er bei dem Meter-System noch einmal gezählt. Pop-Unders werden daher immer mehr genutzt. Die Site x10 erhöhte damit ihren Site-Traffic von 232.000 Visitors im Januar auf 34 Millionen Visitors im Juni. Dadurch entstand ein Streit über die Rechtmäßigkeit.

Keine Audits bisher

1999 begann die MRC ein Audit bei MediaMetrix, das aber noch nicht abgeschlossen wurde, da die Entwicklung nur extrem langsam vorangeht.

Messungen am Arbeitsplatz

Beide Firmen bieten Panels mit Internetnutzern am Arbeitsplatz an. MediaMetrix fragte die Panel-Teilnehmer, die am Zuhause-Panel teilnehmen, ob sie die Software auch an ihrem Arbeitsplatz installieren. Das taten jedoch nur sehr wenige. NetRatings hat ein neues Arbeitsplatz-Panel aufgebaut. Dabei "überlappen" sich die Teilnehmer mit dem Zuhause-Panel. Beide Panel sind keinesfalls perfekt. Die Probleme für Arbeitsplatz-Panel sind sehr schwer zu handhaben.

Alles in allem gab es keine großen Entwicklungen in den letzten zwei Jahren; es blieb alles beim alten. Bei einer Einteilung in drei Gruppen von Viel-Nutzern, Mittel-Nutzern und Wenig-Nutzern, die ähnlich beim Fernsehen gebildet werden, zeigte sich, dass die Viel-Nutzer ca. 85% der Online-Zeit auf sich vereinigen, die Mittel-Nutzer ca. 14% und die Wenig-Nutzer nur ca. 2%. D.h. Werbung auf einer Web-Site hat eine 85%ige Chance von einem Viel-Nutzer gesehen zu werden. Aus dieser Perspektive ist jede Schätzung von überschneidungsfreien Reichweiten irreführend - die Masse der Kontakte mit einer Site erfolgt mit relativ wenig Nutzern.

Verteilung der Zuhause-Nutzer nach Nutzungs-Minuten, Oktober 2000

	Nutzer	Minuten	% der Nutzer	% der Minuten
Wenig	25.911	1.382.959	33,3	2,0
Mittel	25.856	9.198.897	33,3	13,5
Viel	25.946	57.379.528	33,4	84,4
Gesamt	77.713	67.961.384	100,0	99,9

Quelle: unveröffentlichte MediaMetrix-Daten

Neue Markt-Teilnehmer

Es gibt einige neue Anbieter im Online-Marktforschungsmarkt, die teilweise exotische Namen haben. Keiner dieser Anbieter ist aber bis jetzt national tätig und ihre Methoden sind ziemlich undurchsichtig und mysteriös.

Eine neue Firma ist Comscore.com, die eine aggressive PR betreibt und finanziell gut gepolstert ist. Dazu tragen Partner bei, wie Flatiron Partners, Accel, Siebel Systems und Forrester research. Außerdem gibt es eine Vereinbarung mit Double Click, die Diameter heißt. Diameter wird die Rating-Produkte von Comscore, die NetScore genannt werden, verkaufen. Außerdem hat Comscore eine Verbindung zu NetValue, die die Comscore-Technologien in Großbritannien, Frankreich und Deutschland einsetzen wollen. Dafür dürfen sie

die Comscore-Daten der USA nutzen.

Comscore hat einen enormen Panel-Bestand von 1,5 Mio. Online-Nutzern. Rekrutiert wurden sie über eine Kombination verschiedener Stichproben-Techniken. So wurden RekrutierungsBanner über das Web gestreut. Den Panel-Teilnehmern werden folgende Benefits angeboten: Schnellerer Web-Service und Downloads / Chancen bei einer Lotterie zu gewinnen / eine Möglichkeit die Zukunft des Web zu beeinflussen.

Einige Banner sind bezahlte Werbung, einige sind auf Kommissionsbasis. Wird jemand durch die o.a. Benefits angesprochen, wird gesagt, sie müssen ein Applet downloaden, das es erlaubt, die Computernutzung zu beobachten, dass die Daten vertraulich behandelt werden und dass einige demographische Daten erhoben werden. Neu ist das Applet, was heruntergeladen werden soll. Es arbeitet nicht wie ein Meter (à la MediaMetrix oder NetRatings) sondern es leitet die Nutzungsvorgänge auf einen von 200 Comscores Proxy-Servern um. D.h. jede Seite, die dann abgerufen wird, geht durch einen der Proxy-Server. Wegen der Größe und der Power dieser Proxy-Server kann das Versprechen der schnelleren Webnutzung eingehalten werden. Comscore berichtet von einer 50%igen Umschlagsrate (Zu- und Abgänge) - vielleicht ein Indikator dafür, dass die versprochene Geschwindigkeit schwer einzuhalten ist. Oder die Nutzer entscheiden sich, dass sie das Applet nicht mehr installiert haben möchten.

Comscore führt jede Woche 1.000 Telefoninterviews durch, um eine Gewichtungsbasis für das riesige 1,5 Mio.-Panel bereitzustellen. Das soll wahrscheinlich den Kunden vorspiegeln, dass das Panel repräsentativ ist. Jeder, der nur einige statistische Kenntnisse hat, weiss jedoch, dass dieses keineswegs der Fall sein kann, da es sich um eine selbstselektierende Stichprobe handelt.

Außerdem behauptet Comscore, dass die Meßmethode auch erlaubt, das Dollar-Volumen von Online-Einkäufen festzuhalten (Buying Power Index - BPI) - das kann weder MediaMetrix noch NetRatings. Wenn das stimmt, ist das eine attraktive Quelle für Online-Medien, auch wenn diese keine Verkäufe auf ihren Sites anbieten. So kann man aber z.B. nachschauen, welchen BPI die Besucher von Vogue.com haben und das für Werbung nutzen.

Ein weiterer Punkt, der Comscore zugute kommt sind die großen Fallzahlen, da es ein großes Panel ist. Die Tendenz geht dahin größere statt kleinerer Zahlen (wie bei MediaMetrix und NetRatings) zu bevorzugen.

Comscore vereinigt so etwas wie User-Centric- und Site-Centric-Erhebungen. Obwohl Nutzer rekrutiert werden, werden die Page Requests durch die Logfiles auf dem Proxy-Server generiert. Das ist ziemlich faszinierend, aber es gibt auch einige Problem:

- Die Rekrutierungsmethode über Banner, da hier ein Bias von Viel-Nutzern entsteht. Außerdem handelt es sich um eine selbstselektierende Stichprobe.
- Comscore-Einheiten für die Berichterstattung sind nicht Personen, sondern Maschinen. Es können keine überschneidungsfreien Daten geliefert werden, z.B. für Zuhause und Arbeitsplatz, da es für Comscore aussieht wie zwei verschiedene

Maschinen, auch wenn nur ein Nutzer dahintersteckt. Andererseits kann auch nicht unterschieden werden zwischen zwei Nutzern eines Computers, da das wie eine Maschine gezählt wird. Eventuell können Haushalts-Daten geliefert werden, wenn man einen Computer pro Haushalt annimmt.

- Comscores Applet kann Firmen-Firewalls nicht überwinden. D.h. das Arbeitsplatz-Panel ist nicht repräsentativer als das von MediaMetrix oder NetRatings, da Angestellte großer Firmen notwendigerweise dadurch unterrepräsentiert werden.
- Comscores Versprechen eines schnelleren Zugangs wird Breitband-Nutzer unterrepräsentieren, da diese bereits den schnellsten Zugang haben, den es gibt.
- Comscore misst keine Aktivitäten innerhalb des AOL-Systems. Da AOL einen Marktanteil von 50% hat, ist dieses ein signifikanter Mangel. Zur Zeit kann nur MediaMetrix dort messen.
- Comscores Methode misst keine Nicht-HTML-Sites, die in der Online-Welt zunehmen, d.h. sie kann die Nutzung von Web-Applikationen, wie RealPlayer, Napster, MusicMatch usw. nicht messen. Zur Zeit bietet MediaMetrix das breiteste Angebot dieser Nicht-HTML-Aktivitäten an, aber auch MediaMetrix hat Schwierigkeiten den konstanten Strom von Innovationen der Web-Media-Elemente einzuholen.

Zusammenfassung

Es gab keine enormen Veränderung bei der Online-Messung seit dem letzten Symposium. Es gab einige Konsolidierungen und Kürzungen/Streichungen in der Dot.com-Welt. Es gab einige interessante Dinge bei der Ausbreitung der Pop-Up- und Pop-Under-Werbung. Das hat die Diskussion angestoßen, was gezählt werden sollte und hat die Verwundbarkeit der User-Centric-Panel aufgezeigt. Und es gibt einen neuen Markt-Teilnehmer: Comscore - mit einer einzigartigen und vielversprechenden Meßmethode - aber auch mit zahlreichen Einschränkungen, die ihren Anspruch unterlaufen, das Feld aufrollen zu wollen.

When two worlds collide - Offline Measurement of Online Readership

Wenn zwei Welten aufeinandertreffen: Offline-Messungen von Online-Leserschaften

Richard Asquith, BMRB International

Einleitung

Die meisten Verlage bieten Online-Versionen ihrer Haupt-Titel an. Das bedeutet, dass zunehmend die verschiedenen Nutzerschaften identifiziert werden müssen, um das Produkt weiterzuentwickeln und damit die Nutzerschaften zu erhöhen. Außerdem bieten die Verlage zunehmend Werbe-Packages für beides an. Es wird eine akzeptierte Messmethode benötigt, die das Online-Lesen misst und neben der traditionellen Erhebung von Print-Leserschaften bestehen kann. Die Möglichkeit von Erhebungen in den Online-Panels per Metern ist durch die bestehenden Einschränkungen nicht gegeben.

Leserschaftsforscher müssen sich zunehmend mit dem Lesen der elektronischen Versionen von traditionellen Print-Titeln beschäftigen. Die Befragten werden häufig verwirrt, da sie bei der Abfrage nach dem Lesen eines Titels auch an die elektronischen Versionen denken. Das wird jedoch bei der Abfrage nicht berücksichtigt. Da die Leserschaften der Online-Versionen jedoch anwachsen, müssen wir sorgfältig die Validität der Online-Messungen im Zusammenspiel mit der traditionellen Leserschaftsforschung sicherstellen.

Dieser Vortrag soll aufzeigen, was wir zur Erhebung von Online-Leserschaften in Offline-Untersuchungen gelernt haben.

Hauptgrenzen der Meter-Leserschaftsdaten

Die Unterrepräsentation von Arbeitsplatz-PCs in den Online-Erhebungen per Meter ist eine besondere Beunruhigung für Verlage, besonders von Business-to-Business-Publikationen. Eine große Tageszeitung in Großbritannien ist der Meinung, dass sie in Meter-Panels um 80% unterschätzt wird. Da es keine Daten zur Überprüfung gibt, bieten zumindest Erinnerungs-Daten eine Annäherung.

Gestützte Erinnerung

Website	Zuhause-Nutzer		Alle Nutzer		Erhöhung
	%	Mill.	%	Mill.	
Bevölkerung		12,98		16,64	
The Times / Sunday Times	15	1,95	13	2,16	+ 11%
Guardian Unlimited	12	1,56	12	2,00	+ 28%
ft.com	12	1,56	11	1,80	+ 15%
Electronic Telegraph	10	1,30	9	1,50	+ 15%

Quelle: BMRB Internet Monitor (Febr. 2001)

Die gezeigten Titel sind alles große Tageszeitungen. Bei denjenigen Personen, die das Internet ausschließlich Nicht-Zuhause nutzen (Arbeitsplatz, Internet-Café, Uni usw.) sind die Leserschaften im allgemeinen niedriger als bei Zuhause-Nutzern (die auch diejenigen einschließen, die zusätzlich am Arbeitsplatz nutzen). Wenn man sich die absoluten Zahlen ansieht, wird das jedoch kompensiert.

Die Größe der Online-Panels begrenzt die Möglichkeit Daten für Nischen-Titel bereitzustellen. In Großbritannien haben die drei Haupt-Meter-Panels von Jupiter Media Metrix, Nielsen NetRatings und NetValue jeweils unter 10.000 Teilnehmer. Im Vergleich dazu: die nationale Leserschafts-Untersuchung für Print-Medien hat eine Stichprobengröße von 36.000.

Es gibt aber Lösungsmöglichkeiten. In den USA hat comScore z.B. ein Panel von ca. 600.000 PCs, von denen 25.000 am Arbeitsplatz stehen. Das ist proportional immer noch unterrepräsentiert, aber durch die hohe Zahl meint man, man kann die korrekten Proportionen durch Gewichtung herstellen.

Die Meter-Panel messen ausschließlich Online-Versionen von Veröffentlichungen und bemühen sich auch nur, die Online-Nutzerschaft zu repräsentieren. Unser Ziel ist es, Online- und Offline-Leserschaften abzubilden.

Die Recall-Untersuchung

Da die Publikums-Zeitschriften in ihrer gedruckten Form immer noch mehr Leser haben als ihre Online-Versionen, macht es Sinn, eine kombinierte Untersuchung, die beides misst, offline durchzuführen. Damit können die etablierten Währungen beibehalten werden. Wir wissen, dass Erinnerung an das Online lesen alle Probleme beinhaltet, die wir sonst auch haben: Titelverwechslungen, Erinnerungsfehler, Überschätzungen von großen Titeln usw. Diese Probleme werden noch verschlimmert, wenn die Befragten sich an Websites erinnern sollen, die sie in letzter Zeit besucht haben. Lt. ICANN gibt es 41 Millionen registrierte Domain-Namen und einige Sites haben alternative Adressen bzw. verschiedene internationale Sites.

Die Hauptpunkte zur Messung von Online-Publikationen via Recall-Untersuchung sind:

- "Orts"verwechslungen (Online-Version oder Print-Version gelesen?)
- Titelverwechslungen
- Erinnerungsfehler
- Anzahl der einzubeziehenden Titel / Sites
- Zeitskala

1. "Orts"verwechslungen (Online-Version oder Print-Version gelesen?)

In der nationalen Leserschaftsuntersuchung (NRS) und beim TGI wird in Großbritannien bei der Heftherkunft auch eine Kategorie Online vorgegeben. Dadurch kann ausgeschlossen werden, dass Leserschaften sich falsch zuordnen. Da es unser Ziel ist, Leserschaften aller Formate in der gleichen Untersuchung zu erfassen, sollten die Befragten an alle Kontakte Off- und Online denken. Die Heftherkunftsfrage ist eine Möglichkeit, hat aber

einige Nachteile:

- Die Online-Version einer Publikation kann einen leicht unterschiedlichen Titel haben. Um die Gesamtleserschaft erheben zu können, müssen die Vorlagen separat für Online- und Offline-Titel vorgelegt werden. Logos, die dann zweimal auftreten, können jedoch Verwirrung stiften und die Leserschaft verringern.
- Den Online-Editionen muss bei der Vorlage besondere Aufmerksamkeit gewidmet werden, damit sie nicht übersehen werden. Das kann jedoch die Erhebung der gedruckten Titel beeinflussen.
- Die Zeitskala ist bei Online-Titeln besonders wichtig. Eine unterschiedliche Skala für Off- und Online-Titel in einer Frage kann aber nur schwierig gehandhabt werden.

Wir sollten die beiden Kategorien getrennt behandeln. Damit kann der Befragte sich auf eine Version konzentrieren. Das heißt jedoch, in der Fragestellung sollte möglichst präzise formuliert werden, ob es sich um gedruckte Titel handelt oder um Online-Versionen.

2. Titel-Verwechslungen

Online-Editionen von Print-Titeln nutzen den gleichen oder sehr ähnlichen Namen wie die gedruckten Versionen. Wahrscheinlich besteht bei Online-Titeln hinsichtlich der Titel-Verwechslungen mindestens das gleiche Ausmaß an Verwechslungsfähigkeit wie bei Print-Titeln und es sollten für Untersuchungen die gleichen Maßnahmen ergriffen werden, wie z.B. die Gruppierung von verwechslungsfähigen Websites und die Vorlage von Logos.

Außerdem müssen für den Befragten die Grenzen der abgefragten Online-Marke mitgeteilt werden. Z.B. Vogue: in Großbritannien kann die gedruckte Vogue eindeutig definiert werden, aber die Befragten könnten auch das Lesen von Vogue-Websites meinen, die zu anderen internationalen Editionen gehören:

- style.com/vogue (die amerikanische Ausgabe)
- vogueparis.com
- vogue.de

Eine weitere Frage werfen die Websites auf, die mit Kundenmagazinen verbunden sind. Solche Titel wie Safeway-Magazine, Sainsbury Magazine, AA (Automobile Association) Magazine usw. haben keine eigene Website aber Safeway, Sainsbury und auch die Automobile Association haben Websites und bieten eine große Vielfalt an E-commerce-Service und Firmen-Informationen an. Ob die Nutzung dieser Sites für eine Off- und Online-Leserschafts-Untersuchung relevant ist, ist zu diskutieren und muss vom Forscher entschieden werden.

Eine Tatsache, die besonders Tageszeitungen betrifft, ist die Trennung von Werktags- und Sonntags-Titeln. Manche Verlage veröffentlichen auch zwei Websites, manche nutzen verschiedene Kanäle auf der gleichen Site und manche haben nur eine Website für beide Ausgaben.

Die Vorlagen für die Online-Titel müssen sehr genau sein und müssen auch die Adresse beinhalten sowie eventuell Ausschlusskriterien.

3. Erinnerungsfehler

Es ist eine Herausforderung für Recall-Untersuchungen die relevanten Online-Nutzungsvorgänge trotz der Erinnerungsschwächen und -fehler zu erheben. Sorgfältig ausgearbeitete Fragestellungen sind dabei sehr wichtig. Sie müssen den Befragten dazu bringen sich an alle möglichen Lesevorgänge und Leseorte zu erinnern. Es ist unrealistisch, Erinnerungen über lange Zeiträume zu erwarten. Die akzeptierte Maßeinheit, bei der es sich um einen brauchbaren Kompromiss handelt, ist "innerhalb des letzten Monats". Diese Definition sollte jedoch flexibel gehandhabt werden, wenn das Medium wächst.

4. Anzahl der einzubeziehenden Titel / Sites

Die Konkurrenzsituation beim Online-Medium ist vielfältiger als bei anderen Medien. So müssen z.B. Tageszeitungen zusätzlich mit Sites konkurrieren, wie BBC Online oder CNN.com. Der Druck alles einzubeziehen, was von Bedeutung sein könnte, führt zu Überlegungen, mit welchen Techniken der Befragte von der Titellanzahl entlastet werden könnte. Die naheliegendste Lösung ist die Einführung eines Filters, um herauszufinden, welche Titel bzw. Websites für den Befragten relevant sind. Das kann jedoch einen dramatischen Einfluß auf die Größe der Nutzerschaften für die Websites haben. Da wir an Websites interessiert waren, die mit Tageszeitungen und Zeitschriften verbunden sind, nahmen wir eine Kategorie, die theoretisch gut abzugrenzen ist: Nachrichten und Tagesereignisse. Wir testeten den Effekt eines solchen Filters auf die Nutzung von Websites im letzten Monat.

(Tabelle mit Ergebnissen)

Die Daten waren ohne Filter um 83% höher als mit Filter. Entweder verstehen die Befragten den Filter nicht oder sie verbinden diese Kategorie nicht mit den abgefragten Websites. Wenn wir nur die Sites analysieren, die mit gedruckten Tageszeitungen verbunden werden, dann schließen wir z.B. diejenigen aus, die für andere Zwecke als Nachrichten und Tagesereignisse genutzt werden. Aber auch dann drückt die Filterfrage die Nutzung nach unten. Außerdem gibt es Hinweise darauf, dass die Kategorien-Filterfrage niedrigere Ergebnisse liefert, wenn die Kategorien in einer Liste vorgelegt werden als wenn sie einzeln abgefragt werden.

Internet-Site-Kategorien, die im letzten Monat besucht wurden

Kategorien	Einzel abgefragt	Innerhalb einer Liste abgefragt
	n=673	n=762
	%	%
Nachrichten und Tagesereignisse	36	28
Finanzen und Bankgeschäfte	37	30
Stellenanzeigen bzw. Karrieresites	26	22
Eine Site, auf der man Infos über Produkte, Dienstleistungen oder Reisen erhält oder sie kaufen kann	63	56
Eine aktuelle Site für Frauen	18	11

Basis: Erwachsene ab 15 Jahren, die das Internet im letzten Monat genutzt haben
Quelle: BMRB, April 2001

5. Zeitskala

Die Abfrageperiode "im letzten Monat" ist bereits erwähnt worden. Sie repräsentiert die allgemein gebräuchlichste Definition. Bei Print-Titeln wird das Leserschaftsmodell durch die Erscheinungsintervalle geprägt. Bei Online-Sites gibt es kein Erscheinungsintervall. Die Inhalte werden im allgemeinen kontinuierlich bearbeitet und die Kontakt-Periode ist eine willkürliche Auswahl. Theoretisch wird die Auswahl leichter, wenn man die Zeiträume, für die Werbung gebucht werden kann, berücksichtigt. Aber die Verlage haben unterschiedliche Werbe-Packages und außerdem gilt unser Interesse auch Nicht-Werbe-Inhalten.

Was sind die Optionen?

- Man nimmt, wie jetzt, eine willkürliche Periode für alle Online-Publikationen (letzter Monat, letzte 7 Tage, gestern). Der Vorteil ist eine gleichmäßige Messung der Online-Titel und daher einfach für die Befragten und die Daten-Nutzer.
- Eine Alternative wäre, die Online-Perioden an die der gedruckten Ausgaben der Titel anzupassen. Das hätte zwar den Vorteil der Vergleichbarkeit, die Erscheinungsintervalle der gedruckten Ausgaben haben jedoch für die Online-Welt keine Bedeutung.

Aus diesen Gründen wird die feste Untersuchungsperiode bestehen bleiben. Um jedoch den Erinnerungsfehler beim Befragten so klein wie möglich zu halten, sollte das Zeitintervall so weit wie möglich reduziert werden.

6. Daten-Fusion und TGI.net

Bei der vorgeschlagenen Trennung der Fragen für Off- und Online-Leserschaften können zusätzlich Überlegungen zu Daten-Fusionen angestellt werden, da die verschiedenen Formate berücksichtigt werden können.

In Großbritannien werden die Befragten des TGI und des Internet-Monitors aus unserem wöchentlichen Omnibus rekrutiert, d.h. wir haben eine Überschneidung von Befragten, die an beiden Untersuchungen teilgenommen haben. Diese Daten-Überschneidungen nutzen wir zum modellieren der Internet-Monitor-Daten für die gesamte TGI-Internet-Nutzer-Stichprobe. Für Südamerika haben wir dafür Wiederholungsbefragungen bei den TGI-Befragten durchgeführt.

Zusammenfassung

Dieser Vortrag sollte einige praktische Herausforderungen für die Untersuchung von Off- und Online-Leserschaften in einer Studie aufzeigen. Bei allen Leserschaftsuntersuchungen muss der Forscher sich mit Titelverwechslungen und Grenzen der Erinnerungsfähigkeit eines Befragten beschäftigen. Die Vielfältigkeit des Internets verlangt die Definition von Grenzen und die Auswahl eines geeigneten Abfragezeitraums. Es gibt keine richtige oder falsche Entscheidung aber einige Prinzipien, denen zu folgen nützlich sein kann:

- Statt eine "Heftherkunftsfrage" für Online zu stellen, sollten Off- und Online-Titel separat abgefragt werden.
- Leserschaftsfragen für gedruckte Medien sollten das Lesen von Online-Editionen des Titels ausdrücklich ausschließen.
- Gute Forschungs-Praktiken, wie Titel-Gruppierungen und Vorlage von Logos, sind ebenfalls gut für Online-Medien.
- Grenzen für Online-Medien müssen explizit vorgegeben werden, wie z.B. der Ausschluss von ausländischen Websites.
- Fragen zur Online-Nutzung müssen den Befragten aktivieren, an alle möglichen Nutzungsvorgänge und Nutzungsorte zu denken.
- Das akzeptierte Zeitintervall der Haupt-Untersuchungen ist ein Monat. Diese Periode ist vornehmlich deshalb entstanden, um ausreichend Site-Nutzer zu finden. Ein kürzeres Zeitintervall würde selbstverständlich die Genauigkeit der Erinnerung fördern.
- Interessen-Kategorien-Filter reduzieren die angegebene Website-Nutzung und sollten deshalb vermieden werden.

Measuring Magazine Reading via the Internet: Testing the effect of number of titles and other questionnaire design issues

Erhebung des Zeitschriftenlesens via Internet: Testen des Effekts der Titellanzahl und anderer Fragebogen-Design-Merkmale

Jay Matlin, Knowledge Networks / Bruce Goerlich, Starcom MediaVest Group

Das Internet bietet zur Durchführung von Leserschaftsuntersuchungen viele Möglichkeiten. Die Befragten können die Fragen beantworten, wann es ihnen zeitlich paßt - vielleicht mit besserer Erinnerung und größerer Sorgfalt. Außerdem werden keine Interviewer mehr benötigt. Web-Surveys sind zudem erheblich schneller durchzuführen als postalische Untersuchungen oder face-to-face Untersuchungen.

Das Hauptproblem besteht darin, dass es keine externe Validität oder Projektionsmöglichkeit gibt. Es gibt keine Möglichkeit auf Bevölkerungsbasis Zufalls-Stichproben zu generieren. Knowledge Network hat ein Panel mit über 25.000 US-Haushalten per Telefon rekrutiert. Diesen Haushalten wurde der Internet-Zugang via TV ermöglicht; Knowledge Network zahlt die Kosten dafür. Dafür müssen die erwachsenen Haushaltsmitglieder jede Woche einen 10- bis 15-minütigen Fragebogen via Internet beantworten. Diese Stichprobe ist also genauso hochrechenbar, wie jede andere Telefon-Stichprobe (RDD) auch.

Zahlreiche Vorteile können damit auch für Leserschaftsforschungen genutzt werden. Wie nah sollte der Fragebogen an dem Standardfragebogen für Leserschaftsforschung angelehnt sein? Wieviele Titel können einbezogen werden? In welcher Reihenfolge? Wieviele sollten zeitgleich auf einer Bildschirmseite gezeigt werden? In welchem Umfang offeriert eine neue Datenerhebungsmethode die Strukturen der Datenerhebung zu überarbeiten, ohne Kompromisse hinsichtlich der Qualität und Glaubwürdigkeit eingehen zu müssen.

Ursprungsfragebogen

Der Ursprungsfragebogen folgte dem üblichen US-Standard: Ein 6-Monats-Filter mit dem Hinweis in der Fragestellung, man sollte an alle Orte denken, wo man eine Zeitschrift gelesen haben könnte. Die Zeitschriften-Logos werden auf dem Bildschirm gezeigt und zu jedem Logo muss angegeben werden, ob man den Titel in den letzten 6 Monaten gelesen hat oder nicht, oder ob man sich nicht sicher ist. Für alle gelesenen Titel wird dann die Frequenzfrage gestellt: Ob der Befragte üblicherweise 0, 1, 2, 3 oder 4 von 4 Ausgaben liest. Obwohl die Fragestellungen den US-Standardfragen sehr ähnlich sind, gab es doch eine Reihe von Unterschieden:

1. Es werden auf einer Bildschirmseite 5 bis 7 Logos gezeigt und nicht nur eins.
2. Der Befragte klickt auf kleine Buttons mit ja/nein/weiß nicht und legt die Titelkarten nicht auf ja-/nein-/weiß nicht-Felder einer Liste
3. Die Logos werden in alphabetischer Reihenfolge in Blöcken zusammengefaßt. Jede Bildschirmseite zeigt diese Logos eines Blocks. Die Blöcke werden in zufälliger Reihenfolge gezeigt.
4. Die Logos waren farbig.

Diese Ursprungsfassung war für einige interne und externe Kundenkreise unbefriedigend:

1. Unsere potenziellen Zeitschriften-Kunden waren im allgemeinen mehr daran interessiert, Käufer und Abonnenten zu identifizieren als regelmäßige Leser.
2. Es gab Druck mehr als 106 Titel zu erheben.
3. Die Analyse der Daten der ersten 10 Erhebungswochen im Herbst 2000 ergab, dass die Befragten ihr Leseverhalten zu überschätzen schienen.

Um die Validität der Methode zu überprüfen, verglichen wir die Ergebnisse mit den MRI-Daten. Da wir nicht nach dem letzten Lesen gefragt hatten, errechneten wir den LpA aus der Frequenzfrage.

Ermutigend war, dass 89 von 97 Zeitschriften, die in beiden Erhebungen enthalten waren, eine Korrelation zwischen unseren abgeleiteten LpA und dem MRI-Wert auf der Basis des letzten Lesens, aufwiesen. Wir fanden jedoch heraus, dass unsere abgeleiteten LpA im Durchschnitt 81% höher waren als die MRI-Leserschaften. Einige mögliche Erklärungen, die die Diskrepanz erklären können:

- Unsere Stichprobe: Unsere Stichprobe wird telefonisch über Random-Digit-Dialing rekrutiert. MRI nutzt eine Flächenstichprobe. Unsere Stichprobe spiegelt jedoch die US-Bevölkerung wieder und wird nach demographischen Merkmalen gewichtet.
- Vergleichbarkeit der Messungen: Vielleicht war der Vergleich von Leserschaften, die über die Frequenz errechnet wurden, mit Leserschaften, die nach der Frage nach dem letzten Lesen gebildet wurden, problematisch. Dieser Unterschied kann jedoch nicht allein die höheren LpA-Niveaus erklären. Die durchschnittliche Summe der Zeitschriften, die in den letzten 6 Monaten gelesen wurden, ist die gleiche, obwohl wir weniger Titel als MRI abgefragt haben.
- Anzahl der Titel: Es ist bekannt, dass man, je mehr Titel man abfragt, desto niedrigere Leserschaftsdaten erhält und umgekehrt. Die MRI-Befragten können anhand der Titelkarten sehen, wie viele Titel abgefragt werden sollen. Knowledge Network hat vorher nicht gesagt, wie viele Titel abgefragt werden. Vielleicht dachten die Befragten, dass es weniger seien als es dann tatsächlich waren.
- Anzahl der Titel pro Bildschirmseite: Jeder Effekt, der aus weniger Titeln insgesamt resultiert, kann durch die geringe Anzahl von Titeln, die pro Bildschirmseite gezeigt wurden (nur 5 bis 7), verstärkt werden.
- Datenerhebungsmethode als solches: Vielleicht produziert das Interviewen im Netz andere Leserschaftsdaten als wenn ein face-to-face Interview durch einen Interviewer durchgeführt wird. Auch andere Untersuchungen bestätigen dieses.

Forschungsfragen

Wir bemühten uns, diese Erklärungen zu überprüfen und einige Einschränkungen des Ursprungsfragebogens zu umgehen. Im letzten Frühjahr führten wir mit Starcom MediaVest Group ein Experiment durch, bei dem die Anzahl der Titel auf 200 erhöht wurde und Informationen zum Kauf bzw. Abonnement aufgenommen wurden. Dabei sollten folgende Fragen geklärt werden:

- Frage 1: Werden durch die Änderung des Formats des 6-Monats-Filters und durch die Angabe wieviele Zeitschriften abgefragt werden, die Überschätzungen reduziert? Bisher musste für jeden Titel der 6-Monats-Filter einzeln beantwortet werden, jetzt wurde der Filter für Gruppen abgefragt.
- Frage 2: In welcher Höhe reduzieren mehr abgefragte Titel die Leserschafts-Niveaus?
- Frage 3: Können die Leserschaften für die individuellen Titel verbessert werden, indem man eine Frage voranstellt, die danach fragt, wieviele Titel der Befragte liest?
- Frage 4: Ist es möglich, glaubwürdige Leserschaftsschätzungen zu erhalten, wenn man die Befragten, die in der letzten Zeit gelesenen Titel ungestützt aufzählen lässt?
- Frage 5a: Werden Überschätzungen durch eine selektive Abfrage der Zeitschriften-Kategorien, für die das Lesen angegeben wurde, reduziert?
- Frage 5b: Verbessert die Präsentation der Zeitschriften-Logos nach zusammengefassten Gattungen (wie bei der britischen NRS) die Leserschaftsangaben?
- Frage 6: Produziert die kalibrierte Frequenzfrage, die Frage nach dem letzten Lesen oder die Gestern-zum-ersten-Mal-gelesen-Frage die glaubwürdigsten Angaben bei dieser Methode?

Design des Experiments

(Es wurden verschiedene Fragebogen entwickelt, die in dem Vortrag ausführlich beschrieben werden. Die Zusammenfassung wird in der nachfolgenden Tabelle gezeigt.)

	Version A 1 und Version A 2 Kontroll- Versionen	Version B Vollständige Version	Version C 80 Titelverson	Version D Filter-Kategorien- System-Version
Zahl der Titel, die in dem 6-Monats-Filter enthalten sind	jeweils 100	200	80	max. 200, die Anzahl pro Befr. differiert
Fragen vor der Filterfragen	4 Einleitungsfragen	4 Einleitungsfragen	4 Einleitungsfragen +1 Frage über die Anzahl der ZS, die gelesen werden	4 Einleitungsfragen + Frage über das Lesen von 20 Gattungen in den letzten 6 Monaten
Offene Frage über andere ZS, die in den letzten 6 Mon. gelesen wurden	nicht gefragt	nicht gefragt	gefragt mit Platz für 10 Nennungen	nicht gefragt
Fragen zum Abonnement bzw. Kauf	gefragt für alle Titel, die durch den Filter kamen	nicht gefragt	gefragt für alle Titel, die durch den Filter kamen	nicht gefragt
Frage nach dem letzten Lesen und nach gestern gelesen	gefragt für alle Titel, die durch den Filter kamen	gefragt für alle Titel, die durch den Filter kamen	gefragt für alle Titel, die durch den Filter kamen	gefragt für alle Titel, die durch den Filter kamen
Zeitaufwand für die in letzter Zeit gelesenen Titel	gefragt für alle Titel, die in den letzten 7 Tagen gelesen wurden	gefragt für alle Titel, die in den letzten 7 Tagen gelesen wurden	nicht gefragt	nicht gefragt

Die Fragebögen wurden zwischen 25. Mai und 1. Juni 2001 an 6.700 Panel-Teilnehmer verschickt. Insgesamt antworteten 5.443 Befragte (81%) in zwei Wochen (einige antworteten noch danach). Stichprobengröße pro Version:

Version A insgesamt	1.661
Version A 1	818
Version A 2	843
Version B	1.116
Version C	1.327
Version D	1.339

Die Daten wurden demographisch gewichtet und ebenfalls nach der Reihenfolge in der die Titel in Version A bis C erschienen (alphabetische Reihenfolge erste 40 zuerst und letzte 40 zuerst).

(Tab. mit Ergebnissen über Reihenfolge)

Ergebnisse

Eine Tabelle über den 6-Monats-Filter und den LpA ist im Anhang enthalten. Nachfolgend werden die aufgeworfenen Fragen beantwortet:

Frage1: Übergang von Angaben für jeden Einzel-Titel zu Angaben von Gruppen von Titeln

Wie erwartet sanken die Leserschaften bei der Abfrage von Gruppen. Die durchschnittliche Anzahl der 99 Titel, die in Experiment- und Ursprungsversion enthalten waren, sank von 14,4 auf 8,8 beim 6-Monats-Filter. Der Durchschnitt des LpA, durch Frequenzangaben errechnet, sank von 6,39% auf 3,29%. Dadurch, dass die Befragten nicht gezwungen wurden, zu jedem Titel eine Antwort zu geben und dadurch, dass die Anzahl der Bildschirmseiten mit Logos von 18 auf 5 reduziert wurde, wurde das Interview für die Befragten leichter. Außerdem kann die Angabe, wieviele Logos gezeigt werden, verhindert haben, dass sie eine Vielzahl von Titeln angeben, die sie lesen, um gebildet zu erscheinen.

Fragen 2 und 3: Anzahl der einbezogenen Titel

Die Anzahl der Titel, die in der 6-Monats-Filter-Frage abgefragt werden, beeinflusst die Anzahl der Titel, die den Filter passieren. Je größer die Anzahl der Zeitschriften ist, die abgefragt werden, desto kleiner ist der Prozentsatz der Titel, die die Befragten angeben in den letzten 6 Monaten gelesen zu haben.

Version	Anzahl gezeigter Titel im 6-Monats-Filter	Durchschnittliche Anzahl Titel, die den Filter passierten	Hochgerechnete Zahl der Titel, die den Filter passiert hätten bei 200 Titeln*	Index vs. Kontrollversion
B (Vollständig)	200	11,5	11,5	95
A (Kontroll)	100 (Ø der 2 Kontroll-Versionen)	6,0	12,1	100
C (80 Titel & offene Frage)	80 - justiert an der größeren durchschn. Reichweite dieser 80 ZS in den and. 3 Versionen	5,3	13,3	111
C (80 Titel & offene Frage)	80 - unjustiert	6,1	15,2	126

* Durchschnittliche Anzahl in der Spalte links multipliziert mit 200 und geteilt durch die Anzahl der abgefragten Titel

Es scheint, dass die Frage, wieviele Titel man liest, vor der 6-Monats-Filter-Frage die Anzahl der Titel nicht reduziert. Die Anzahl der Titel hat keinen Einfluß auf das "Lese-Antwortkategorien-Verhältnis" der Frequenzfrage. Ein Vergleich der 80 Titel, die in allen drei Versionen enthalten sind, zeigt ein Verhältnis von .38 bis .40. Der Unterschied zwischen den Versionen mit 80 Titeln und den beiden Kontrollversionen mit je 100 Titeln beträgt auf der Basis des LpA (errechnet aus der Frequenz) im Durchschnitt eine 14% höhere Reichweite für die Versionen mit 80 Titeln. Zwischen 100 und 200 Titeln verringert sich die Differenz um weniger als ein Zehntel eines Prozentpunktes.

(Tab. mit Ergebnissen)

D.h. auf unserer Interview-Plattform scheint die Anzahl der Titel das Ergebnis zu beeinflussen, der Effekt ist jedoch bei über 100 Titeln sehr gering.

Fragen 3 und 4: Offene Fragen

Das Ergebnis der offenen Frage nach weiteren Titeln zeigt die gegenwärtige Situation in den USA. Über 1.100 Titel wurden dort von 1.327 Befragten dieser Versionen zusätzlich genannt. 62% der Befragten nannten mindestens einen Titel.

Wie erwartet, wurden die 120 Titel, die in Version C nicht vorgelegt wurden, auf einem viel niedrigerem Niveau genannt als in den anderen Versionen. Das Angaben-Verhältnis war jedoch für diese Titel höher als in den Versionen A und B, wo sie aufgeführt waren. Dieses höhere Angaben-Verhältnis glich jedoch die niedrigeren Nennungen nicht aus. Der LpA war also für alle 200 Titel in der "Offenen-Frage-Version" viel geringer als in den Versionen, wo alle 200 Titel durch Vorlage abgefragt wurden.

Auch die Anzahl der genannten Titel, die in den letzten 6 Monaten gelesen wurden, erreicht lediglich 8,1 bei der Version C, wenn alle offen genannten Titel mitgezählt werden, gegenüber 11,5 Titeln, wenn alle 200 vorgelegt werden, obwohl mehr als 900 Titel in der gestützten Version gar nicht vorkommen. Wenn diese 900 Titel abgezogen werden, sinkt der Durchschnitt auf 7,0. Wenn auf der Basis der beiden 80-Titel-Versionen auf die 200 Titel hochgerechnet wird, ist die Filterpassage 39% geringer als in den Versionen mit 200 Titeln. (Tab. mit Ergebnissen)

Für die 120 Titel, die nicht in der 80-Titel-Version vorgelegt wurden, die aber offen genannt werden konnten, waren die Angaben zu den letzten 6 Monaten 86% niedriger als in der 200-Titel-Version (bezogen auf diese 120 Titel). (Tab. mit Ergebnissen)

Fragen 5a und 5b: Filterung und Gruppierung nach Gattungen

Das Experiment, das Interview dadurch stromlinienförmiger zu machen, dass vor der 6-Monats-Filterfrage nach dem Lesen von Zeitschriften-Gattungen gefragt wird, scheint besser zu funktionieren als die offene Frage. Diese Filterfrage nach Gattungen kürzt die Anzahl der abzufragenden Titel enorm, selbst bei der Version in der alle 200 Titel enthalten sind. Im Durchschnitt wurden 75 Titel auf 8,5 Bildschirmseiten abgefragt. (Dieser Durchschnitt beinhaltet die 19 Titel, die allen Befragten gezeigt wurden.) Es ist daher nicht überraschend, dass die Anzahl der Titel, die den 6-Monats-Filter passierten, in dieser Version niedriger war als in der ungefilterten Version mit allen 200 Titeln. Allerdings erlaubt diese Methode Titel zu eliminieren, die für den Befragten uninteressant sind und nicht sein Verständnis von Zeitschriften-Gruppen widerspiegeln. Obwohl den Befragten in der Kategorien-Filter-Version 60% weniger Logos gezeigt wurden, gaben sie lediglich 7% weniger Zeitschriften für die letzten 6 Monate an, als die Befragten mit 200 gezeigten Titeln. Gegenüber den Befragten mit 100 Titeln waren es 11% weniger.

(Tab. mit Ergebnissen)

Die Kategorien-Filter-Methode produziert zudem geringere Antwortkategorien-Verhältnisse, d.h. der LpA ist geringer als in den Kontrollversionen.

(Tab. mit Ergebnissen)

Bis auf zwei Ausnahmen (Parade und USA Weekend) gilt das auch für die Erscheinungsintervalle.

(Tab. mit Ergebnissen)

Es kann jedoch festgestellt werden, dass einige der Gattungen von der Filterung profitieren und andere nicht (Aufzählung der 20 Gattungen).

Frage 6: Frage, auf der der LpA basieren sollte

Die beiden klarsten Ergebnisse aus diesem Experiment sagen, dass die Frage nach dem letzten Lesen die höchsten Reichweiten produziert und die Frage nach dem "Gestern-zum-ersten-Mal-gelesen", die niedrigsten.

(Tab. mit Ergebnissen)

Die "Gestern-Ergebnisse" sind so niedrig, dass sie unglaublich sind. Weniger als 15% geben an, irgendeine Zeitschrift gestern zum ersten Mal gelesen zu haben. Es gibt zwei Erklärungen dafür, warum die "Gestern-Frage" nicht funktionierte:

- Da die Interviews selbst ausgefüllt wurden, haben wir keine Kontrolle über den Tag an dem sie beantwortet werden. Meistens geschieht das am Wochenende und damit ist das Lesen am Freitag und Samstag überrepräsentiert.
- Die "Gestern-Frage" ist nicht verstanden worden.

Das Antwortkategorien-Verhältnis für die Frage nach dem letzten Lesen war höher als bei der kalibrierten Frequenzfrage. Insbesondere bei der Version mit nur 80 Titeln wurde das deutlich. Es scheint, dass bei weniger als 100 Titeln sowohl der LpA höher bei der Letztes-Lesen-Methode ist als auch die Anzahl der Titel, die den 6-Monats-Filter passieren. Die Reichweiten nach der Letztes-Lesen-Methode kommen auch näher an die MRI-Daten heran (Tab. mit Ergebnissen)

Folgerungen

Alle getesteten Versionen waren besser als unsere ursprüngliche Version. Auch die kalibrierten Frequenz-Ergebnisse, die damals 81% höher waren als der "Standard", waren in diesem Experiment 20% bis 25% niedriger. Es scheint, dass die Anzahl der Titel und die Art und Weise der Abfrage nicht entscheidend sind, sondern dass die kritischen Faktoren der 6-Monats-Filter (wo alle 20 auf dem Bildschirm abgebildeten Titel abgefragt werden und nicht nur 5 bis 7, bei denen jeder Titel einzeln beantwortet werden muss) und die Angabe, wieviele Titel abgefragt werden, sind.

Unser Experiment zeigt, dass die Durchführung von Leserschafts-Untersuchungen via Internet bei einer Zufallsstichprobe glaubwürdige Ergebnisse bringen können. Das beste Ergebnis wurde durch Einsatz der Frage nach dem Letzten-Lesen, mit vorgegebenen nicht gefilterten 100 bis 200 Titeln, erreicht. Weniger als 100 Titel könnten die Reichweiten inflationieren.

Der Gebrauch des neuen Mediums kann jedoch keine radikalen methodischen Änderungen bei der Datenerhebung nach sich ziehen. Die "Gestern-zum-ersten-Mal-gelesen"-Technik ist nach wie vor problematisch. Genauso unser Versuch zuerst alle Titel, die in den letzten 6 Monaten gelesen wurden auflisten zu lassen. Beim Einsatz von Gattungsfiltern und der Gruppierung der Titel sinken die Reichweiten. Für einige Gattungen, wie z.B. Inflight-Magazine oder Zeitungs-Supplements mag das jedoch besser sein als für andere. Es besteht für die Einteilung von Gattungen weiterer sorgfältiger Forschungsbedarf, damit die Vorstellungen der Forscher denen der Leser näher kommen. Wenn dieses der Fall ist, kann über diese Methode für zukünftige Entwicklungen weiter nachgedacht werden.

Außerdem haben wir festgestellt, dass wir mehr Informationen über das Leseverhalten von unseren Panel-Teilnehmern erhalten können als angenommen. Die Anzahl der Titel von ursprünglich 100 wurde auf 200 verdoppelt und außerdem wurden Zusatzinformationen zu Kauf und Abonnement aufgenommen sowie die Zeit, die für das Lesen einer in der letzten Zeit gelesenen Ausgabe aufgewendet wurde. Und das alles mit Reichweiten-Ergebnissen, die

nahe an dem "Standard" sind. Die durchschnittliche Dauer des "Kontroll-Interviews", das alle aufgezählten Informationen abfragte, betrug 12,5 Minuten.

(Anhang mit Abbildungen der Bildschirmseiten und Tabellen)

Modeling magazine effectiveness

Modellbildung zur Zeitschriften-Effektivität

Ellen Oppenheim & David Arpin, Magazine Publishers of America

Robert Wyman & Gerald Pollak, Media Marketing Assessment

Thomas Robinson, Precision Marketing Services

Die Magazine Publisher of America (MPA) (so etwas wie der VDZ, Verband Deutscher Zeitschriftenverleger in Deutschland) präsentierten 1999 in Florenz auf dem Readership Research Symposium Ergebnisse von drei Untersuchungen aus einer Serie von Werbewirkungsstudien. Der vorliegende Vortrag präsentiert die herausragendsten Ergebnisse der letzten MPA-Initiative, die Werbewirkung in Zeitschriften und wie Zeitschriften-Werbung im Media-Mix arbeitet, untersuchte. Es handelt sich um eine umfassende Analyse, die den umfangreichen Einfluss und die Wirkung auf Zeitschriften auf ansteigende Verkäufe bei Einsatz der neuesten Marketing-Mix-Modelle dokumentiert. Die Initiative zeigt klar auf, dass Zeitschriften-Werbung wirkt - alleine und in Kombination mit anderen Medien. Sie richtet sich an die Werbungtreibenden und ihre Agenturen in den USA, die diese Frage aufgeworfen haben.

Methode

Es wurden 186 Marken aus verschiedenen Produktbereichen analysiert. Die Datenbasis beinhaltete Daten zu Marketing-Mix-Modellen, die für individuelle Marken entwickelt wurden und die Print in Verbindung mit anderen Medien über sieben Jahre, von 1994 bis 2000, geschaltet haben.

(Tab. mit Aufstellung der Spendings)

Aufzählung der Produktbereiche und weiterer Charakteristika, die in die Analyse einbezogen wurden, wie z.B. Saisonalität, Preis, Verkaufszyklen, Konkurrenzumfeld usw.

(Media Marketing Assessments) Analytischer Anspruch

Das Ziel jeden Projektes bestand darin, die finanzielle Effizienz von jedem Marketing-Dollar durch die Kenntnis, welchen Verkaufseffekt jeder Dollar hat, zu maximieren. Die Antwort-Variable (Marken-Verkauf) wird als Vorhersagefunktion für alle signifikanten Einflüsse auf den Verkauf, modelliert. Vorhersagevariable sind Marketing-Anstrengungen, Trends, Konkurrenz-Aktivitäten und andere externe Faktoren. Das Modell nutzt allgemein statistische Ansprüche, um den Einfluss verschiedener Marketing-Aktivitäten auf den Verkauf darzustellen. Es handelt sich dabei um eine iterative Such-Routine, um die verschiedenen Transformationen, die auf Marketing- und Werbung-Variablen entfallen, zu erklären. Es können ebenfalls Nicht-Werbe-Variable herangezogen werden, die Tausende von möglichen Modell-Spezifikationen ermöglichen. Jedes Modell wird dann hinsichtlich der statistischen Zuverlässigkeit, der akkuraten Vorhersage und der Marketing-Sensibilität evaluiert. Das endgültige Modell führt zur Analyse des Marketing-Beitrags, Effektivität, Elastizität und Kosten-Effizienz.

Basis- versus Zuwachs-Volumen

Die Analyse zeigt den Verkaufs-"Anteil" der verschiedenen Marketing-Mix-Elemente. Für jede untersuchte Marke wurde ein bestimmter Prozentsatz der jährlichen Verkaufsanteile als "Basis"-Volumen bestimmt bzw. als Volumen, was ohne Marketing-Aktivitäten erzielt wird. Durchschnittlich repräsentiert das Basis-Volumen 75% des Gesamt-Verkaufs während 25% den Zuwachs ausmacht.

(Tab. mit Prozentanteilen)

Kalkulation der Effektivität

Der Anteil alleine erklärt die Effektivität nicht. Die Ausgaben, die getätigt wurden, um eine Verteilung zu erhalten, müssen berücksichtigt werden. Die Media-Spending-Daten, die in die Analyse eingingen, wurden von den jeweiligen Kunden zur Verfügung gestellt.

(Effektivität = (% der Anteile / Spendings) x 1.000.000)

Erste Ergebnisse

Eine ausführliche Datenanalyse wird noch vorbereitet. Es gibt jedoch ein paar Vorabergebnisse, die die Effizienz von Zeitschriften-Werbung im Vergleich zu anderen Marketing- und Media-Mix-Elementen aufzeigen.

Der Einfluss des zunehmenden Zeitschriften-Anteils an der Werbung

Bei der Analyse von 70 Marken mit gleicher Zielgruppe (Frauen 25-54 J.) wurde der Einfluss von zunehmenden Zeitschriften-TRPs (Target Rating Points) am Beitrag der Zeitschriften insgesamt am Verkauf gemessen. Im Durchschnitt zeigten die Marken, die den höchsten Print-Anteil hatten - bis zu 1.300 TRPs im Jahr, auch den höchsten Zeitschriften-Anteil an den Verkäufen. Das zeigt, dass eine Erhöhung des Werbevolumens in Zeitschriften auch eine Erhöhung der Verkäufe zur Folge hat.

(Graph. mit Ergebnissen)

Effektivität der Marketing-Elemente

Um die Effektivität insgesamt der einzelnen Marketing-Elemente darzustellen, wurde der Anteil der Elemente durch die aktuellen Spendings geteilt. Die Analyse zeigt, dass nach dem Handel, die Zeitschriften das zweiteffektivste Element im Marketing für die betrachteten Marken sind. Zeitschriften sind 50% effektiver als Fernsehen und 71% effektiver als Radio.

(Graph. mit Ergebnissen)

Der Einfluss des Werbebudgets auf die Marketing-Mix-Effektivität

Um den Einfluss des Werbebudgets zu untersuchen wurden die Marken in vier unterschiedliche Gruppen eingeteilt (Werbebudget: 1-10 Mio. \$, 21-30 Mio. \$, 31-50 Mio. \$, 51 Mio. \$ und mehr). Für jede Gruppe wurde ein Marketing-Effektivitäts-Index gebildet, der auf der durchschnittlichen Effektivität für die Marke in der Gruppe basiert.

- Für alle Budget-Levels waren Zeitschriften effektiver als Fernsehen oder Radio.
- Die Effektivität für Zeitschriften-Werbung war am höchsten in der Gruppe der Marken mit einem Marketing-Budget von 51 Mio. \$ und mehr.
- Zeitschriften waren effektiver als andere Marketing-Mix-Elemente für die Marken, die 31 Mio. \$ oder mehr ausgaben.

(Graph. mit Ergebnissen)

Der Einfluss des Zusammenspiels von Zeitschriften und Fernsehen (Flighting)

Es wurde auch der gemeinsame Einfluss von Zeitschriften- und Fernseh-Werbung analysiert. Es wurden drei Gruppen gebildet: keine (wöchentliche) Überschneidung von TV- und Zeitschriften-Werbung, minimale Überschneidung, hohe Überschneidung. Es zeigt sich, dass die zwei Medien zusammen arbeiten und einen Synergie-Effekt aufweisen. Am höchsten wird die Effektivität bei einer hohen Überschneidung. Zusätzlich kann festgestellt werden, dass unabhängig von der wöchentlichen Überschneidung der Werbung, Zeitschriften immer effektiver sind als TV.

(Graph. mit Ergebnissen)

Zusammenfassung

Zusätzliche Analysen werden weitere detaillierte Ergebnisse für jede der zwölf gemessenen Kategorien darstellen. Zur Zeit kann festgestellt werden, dass Zeitschriften-Werbung effektiv zu Verkäufen beiträgt und einen Synergie-Effekt auf andere Elemente im Marketing- und Media-Mix ausüben.

Synergy brought about through duplication of contact with newspaper and internet

Synergien durch Überschneidungs-Kontakte von Tageszeitungen und Internet

Shinichiro Okauchi, Dentsu Inc.

Zusammenfassung

Der Vortrag berührt einen Aspekt der Werbe-Effektivität der aus der Kombination zweier Medien resultiert: Tageszeitungs-Werbung und Internet-Websites. Der erste Vorteil sind Synergien beim Erreichen der Zielgruppe. Es ist möglich, Reichweiten mit Tageszeitungen bei Zielgruppen zu erreichen, die nicht erreicht werden können, wenn man nur das Internet belegt. Damit wird die Reichweite der Werbe-Kommunikation insgesamt erhöht. Außerdem führt das Lesen von Tageszeitungs-Werbung die Leser dazu, Websites zu kontaktieren. Der zweite Effekt ist eine verbesserte Kommunikationsqualität durch die Überschneidung von Tageszeitungs- und Internet-Kontakt mit Informationen zu einem einzigen Thema (nachfolgend Kontakt-Überschneidung genannt). Wir haben die Synergien, die durch Kontakt-Überschneidungen zwischen Tageszeitungen und Internet entstehen, beobachtet. Die Analyse basiert auf zehn Fallstudien zur Werbe-Kommunikation in Japan, die zur gleichen Zeit mit der gleichen Methode durchgeführt wurden. Die Ergebnisse zeigen die Möglichkeit auf, dass Kontakt-Überschneidungen tiefere Informationen bringen, die die Produkt-Informationen besser verarbeiten.

Einleitung

Das Jahr 1995 wird in Japan im allgemeinen als erstes Internet-Jahr bezeichnet. Seitdem hat sich das Internet rapide ausgebreitet. Es existieren bereits 25 Mio. Nutzer (bei einer Gesamtbevölkerung von 127 Mio.), die Haushaltsabdeckung erreichte bereits 30% (von ungefähr 47 Mio. Haushalten). Außerdem erreichte Werbung im Internet im Jahr 2000 ein Budget von US \$ 540 Mio. gegenüber US \$ 15 Mio. in 1996.

Inmitten aller Veränderungen gab es eine Debatte über die Verbindung zwischen Massenmedien und dem Internet. Auf dem Gebiet der Media-Planung für Pressemedien besteht die Notwendigkeit, die Synergien zwischen dem Internet und Pressemedien nachzuweisen. Da das Internet eine Infrastruktur mit extrem breiten Funktionen ist und nicht einfach "Medium" genannt werden kann, ist es notwendig, einen vielfältigen Ansatz zur Evaluierung des Internets auszuwählen. Dentsu führte von Februar bis März 2001 eine große Online-Untersuchung durch, die dem Ziel diente, den Kontakt von Internet-Nutzern mit Tageszeitungen und Internet zu untersuchen. Wir fokussierten uns auf 10 Fallstudien mit Werbe-Kommunikation, um über die Synergien, die die Kontakt-Überschneidungen bei den Nutzerschaften bringen, zu ermitteln.

Methode um Synergien zu verifizieren

Für die 10 Fallstudien wurden Firmen ausgewählt, die kurz vor der Untersuchung in Tageszeitungen geworben haben. Diese Firmen gaben in den Tageszeitungs-Anzeigen zu den Produkt- und Service-Informationen außerdem Informationen über ihre Websites an. Die

Untersuchung liefert Daten, die es uns erlauben die Unterschiede in der Wahrnehmung des Befragten für die Unternehmen, Produkte und Dienstleistungen, die von den unterschiedlichen Medien-Kontakten herrühren, zu analysieren. Die beobachteten drei Ausprägungen von Media-Kontakten waren: Kontakt mit der Tageszeitungs-Anzeige und der Website derselben Firma, Kontakt ausschließlich mit der Tageszeitungs-Anzeige und Kontakt ausschließlich mit der Website.

(Graph. der Analysegruppen)

Es wurden verhältnismäßig unbekannte, kleine Unternehmen in die 10 Fallstudien einbezogen. Dadurch war es notwendig, eine große Fallzahl realisieren zu können. Da Internet-Nutzer die Zielgruppe waren, bot es sich an, die Untersuchung über das Internet durchzuführen.

(Abb. Bildschirmseite mit Fragestellung -

Fragestellung: Bitte sehen Sie sich einmal die oben abgebildete Anzeige an, die kürzlich in einer Tageszeitung erschienen ist. Haben Sie diese Anzeige 1. ganz bestimmt gelesen, 2. vielleicht gelesen, 3. nicht gelesen? - Die Summe von 1+2 ergibt die Anzeigen-Leserschaft)

Übersicht über die Online-Untersuchung

A. Untersuchungs-Inhalte

1. Untersuchungsgebiet: Die größeren Hauptstadtgebiete (Tokyo und sechs nahegelegene Präfekturen: Knagawa, Saitama, Chiba, Gumma, Tochigi, Ibaraki)
2. Grundgesamtheit: Männer und Frauen im Alter von 15 bis 59 Jahren
3. Stichprobengröße: Phase A: 1.890 eingesammelte Fragebögen, 25,2% Ausschöpfungsquote; Phase B: 1.617 eingesammelte Fragebögen, 21,6% Ausschöpfungsquote
4. Forschungsansatz: Benachrichtigung per e-mail und Fragebogen zum Selbstauffüllen auf der Website (Zufalls-Auswahl von registrierten Teilnehmern, die an einer Befragung auf der Website teilgenommen haben)
5. Termine: Phase A: 24.02.-04.03.2001; Phase B: 10.03.-18.03.2001
6. Institut: Dentsu Research

B. Haupt-Befragungsthemen in dieser Untersuchung

1. Umstände, die zum Kontakt mit der Tageszeitungs-Werbung geführt haben
 - Die Tageszeitungs-Anzeige wurde auf dem Bildschirm gezeigt
 - Maßnahmen, die vom Befragten ergriffen wurden, nachdem er die Tageszeitungs-Anzeige gesehen hat
2. Umstände, die zum Kontakt mit der Website geführt haben
3. Wahrnehmung der Unternehmen, Produkte und Dienstleistungen
 - (Produkt-)Marken-Awareness
 - Verständnis der Produktleistung
 - zukünftige Nutzungsabsicht
 - Unternehmens-Image (freundlich, glaubwürdig, international usw.)
4. Basis-Informationen (demographisch)

C. In die Untersuchung einbezogene Werbung

Fallstudie	A	B	C	D	E	F	G	H	I	J
Produktkategorie	PKW	PKW	PKW	Online-Sicherheits-DL	Online-Sicherheits-DL	PC	PC	PC	Unternehmenswerbung	Unternehmenswerbung
Untersuchungs-Phase	A	A	A	B	B	B	B	B	A	A
Tageszeitung	5 Morgen-TZ	3 Morgen-TZ	1 Morgen-TZ	1 Morgen-TZ	1 Morgen-TZ	1 Morgen-TZ	2 Morgen-TZ	3 Morgen-TZ	2 Morgen-TZ	1 Morgen-TZ
Größe der Anzeige	15-spaltig sw	5-spaltig 4c	15-spaltig sw	15-spaltig sw	7-spaltig sw	15-spaltig sw	15-spaltig 4c	15-spaltig sw	15-spaltig sw und 15-spaltig 4c	15-spaltig 4c

D. Befragten-Profile

(Abb. von drei Graphiken mit den Befragten-Profilen nach Alter und Geschlecht, Beruf, Wichtigste Informationsquelle)

Ergebnisse

Reichweiten-Synergien

Die Möglichkeit die Netto Zielgruppen-Reichweiten zu erhöhen

Die Zahl der Befragten nach Media-Kontakten in den einzelnen Fallstudien ist ausreichend um sie analysieren zu können.

(Tab. mit Fallzahlen (zw. 31 und 462))

Die Gruppe, die ausschließlich Kontakt mit der Tageszeitungs-Anzeige hatte, zeigt, dass Tageszeitungs-Anzeigen einen großen Anteil an der Reichweite haben.

Fallstudie

A	B	C	D	E	F	G	H	I	J	Durchschnitt
27,1%	20,1%	18,9%	26,2%	57,4%	19,8%	36,0%	29,3%	11,6%	12,4%	25,9%

(Jeweilige Fallstudien-Gruppe = 100%)

Website-Kontakte, die durch die Tageszeitungs-Anzeige initiiert wurden

Das Lesen von Tageszeitungs-Anzeigen führt zum Zugang zu der Website. Die Werte für die PC-Anzeigen (Fallstudie F / G / H) sind höher als der Durchschnitt.

Fallstudie

A	B	C	D	E	F	G	H	I	J	Durchschnitt
4,6%	4,1%	10,0%	6,9%	6,8%	13,8%	14,6%	11,1%	6,7%	7,1%	8,6%

(Jeweilige Fallstudien-Gruppe = 100%)

Synergien zur Kommunikations-Verbesserung

Die Ergebnisse wurden nach den Gruppen 1. (Produkt-)Marken-Awareness, 2. Verständnis der Produktleistung, 3. zukünftige Nutzungsabsicht, 4. Unternehmens-Image für die einzelnen Media-Nutzungsgruppen, ausgewertet. Für die Analyse definierten wir einen Index für den Überschneidungs-Effekt, wie folgt:

$$\text{Synergie Index} = (\text{Kontakt-Überschneidungs-Gruppe} * 1 \text{ Score}) / (\text{Kontakt ausschließlich mit Website-Gruppe} * 2 \text{ Score})$$

*1 = Befragte, die Kontakt mit Tageszeitungsanzeige und Website hatten

*2 = Befragte, die ausschließlich Kontakt mit der Website hatten

1. (Produkt-)Marken-Erinnerung

(Graph. mit Ergebnissen)

Fallstudie	TZ- und Website- Kontakt	Nur TZ-Kontakt	Nur Website-Kontakt
	%	%	%
A PKW	99,7	97,4	99,7
B PKW	67,4	49,7	52,5
C PKW	100,0	96,8	96,1
D Online-Sicherheit	87,0	69,8	88,1
E Online-Sicherheit	90,3	69,3	77,3
F PC	56,1	19,4	29,8
G PC	98,0	75,6	85,8
H-1 PC	84,2	55,1	73,6
H-2 PC	68,9	46,3	57,7

Die Marken-Awareness der Kontakt-Überschneidungsgruppe ist 1,2 mal so hoch wie die der Gruppe, die ausschließlich Website-Kontakte hatte.

(Tab. mit Synergie-Index)

2. Verständnis der Produktleistung

(Graph. mit Ergebnissen)

Das Verständnis der Produktleistung der Kontakt-Überschneidungsgruppe war 1,65 mal so hoch, wie das der Gruppe, die ausschließlich Website-Kontakte hatte.

(Tab. mit Synergie-Index)

3. Zukünftige Nutzungsabsicht

(Graph. mit Ergebnissen)

Die zukünftige Nutzungsabsicht der Kontakt-Überschneidungsgruppe ist 1,75 mal so hoch, wie die der Gruppe, die ausschließlich Website-Kontakte hatte.

(Tab. mit Synergie-Index)

4. Unternehmens-Image (Graph. mit Ergebnissen)

Das Ergebnis ist ähnlich, wie bei den anderen Variablen. Der Durchschnitt liegt in der Kontakt-Überschneidungsgruppe 1,37 mal so hoch, wie in der Gruppe, die ausschließlich Website-Kontakte hatte. Nur einige ganz wenige Zellen weisen einen Wert unter 1,0 auf.

(Tab. mit Synergie-Index)

Zusammenfassung

Bis jetzt gab es einige Berichte, die die Synergien zwischen Fernseh-Werbung und Printmedien-Werbung beinhalten.

(Graph. mit Ergebnissen)

Durch die Analysen, die in diesem Vortrag vorgestellt werden, können wir einen ähnlichen Effekt für das Internet (als Massenmedium) und Tageszeitungen bestätigen. Die Ergebnisse zeigen, dass Kontakt-Überschneidungen tiefergehende Informationen erreichen und die Kommunikationsqualität erhöht wird:

- Erhöhung der Marken-Erinnerung
- Erhöhung des Verständnisses der Produktleistung
- Erhöhung der zukünftigen Nutzungsabsicht
- Verstärkung eines positiven Unternehmens-Images

Kontakt mit Websites ist eine aktive Beschäftigung, die sich von der mit Tageszeitungen unterscheidet. Ein Umstand, der es leichter macht zu verstehen, warum die Werte für ausschließlichen Website-Kontakt höher sind als für ausschließlichen Tageszeitungs-Kontakt. Aufmerksamkeit verlangt der Umstand, dass bei Kontakt-Überschneidungen die Ergebnisse deutlich höher sind als wenn ausschließlich Website-Kontakt bestand - ein klarer Hinweis auf die Unterstützungsfunktion der Tageszeitungs-Werbung bei Internet-Nutzern.

Es ist notwendig, die Untersuchung fortzuführen, um den vielschichtigen Einfluß zwischen Tageszeitungen und Internet herauszufinden. Diese Untersuchung beleuchtet nur einen partiellen Effekt dieser Verbindung.

Measuring the qualitative value of magazines

Erhebung des qualitativen Wertes von Zeitschriften

Cynthia Evans, The Media Edge/Y&R/WPP

Erhebung des Kommunikationswertes von Publikums-Zeitschriften bei Nutzung von Ergebnissen von TME Media Brand Asset Valuator™, dem Media-Set des Young & Rubicam Brand Asset Valuator™.

Situation

Lesequalitäts-Untersuchungen sind hinsichtlich des Inhaltes und der Umstände, die beleuchtet werden, unterschiedlich. In den USA wird traditionell die Heftherkunft, Lesedauer, Lesevorgänge, % der gelesenen Seiten, etwas ausprobiert nach dem Lesen, erhoben. Andere Studien berichten über andere Vergleiche. Starch bietet Ergebnisse von Copy-Tests und der Erinnerung und Wiedererkennung von Werbung an.

Die Felder psychologischer Werbe- und Redaktionsforschung, Verbraucherverhaltensforschung und Markenforschung werden mit der Durchführung von Bewertungen von Umständen und Beziehungen abgedeckt. Die Ergebnisse zeigen im allgemeinen, dass Zeitschriften Kommunikations-Synergien zwischen Botschaft und Medium aufweisen:

- Im Bewusstsein der Verbraucher spielt sich effektive Kommunikation ab.
- Dem Media-Umfeld angepasste Botschaften nutzen das Image (für sich).
- Werbe-Erinnerung wird durch die Assoziation mit Medien aufgebaut, die hinsichtlich der Eigenschaften und des Images der Marke auf gleicher Linie sind.
- Jeder Media-Einkauf ist eine Möglichkeit imagegeleitete Dimensionen zu erhalten.

Der Young & Rubicam Brand Asset Valuator ist eine Untersuchung der Images und der Eigenschaften, die Verbraucher mit Marken verbinden. Die Studie wird in mehr als 30 Ländern durchgeführt und misst mehr als 50 Items pro Marke. 1999 wurden 106 Zeitschriftentitel einbezogen und 115 für das erste Quartal 2001. Eine Beschreibung der Untersuchung findet sich im Anhang.

Dieser Vortrag beschreibt die Aktivitäten von The Media Edge zur Evaluierung des "qualitativen" Beitrags von Zeitschriften zur Zählbarkeit von Medien.

Harmonisierung über Grenzen

Die herausforderndsten Elemente der Qualität sind die Worte um sie zu beschreiben. Das beste ist, linguistische, kulturelle und antropologische Probleme zu vergessen und die von den Befragten gegebenen Antworten zu akzeptieren, so wie sie sind.

Es scheint zwischen den Geschlechtern einen Unterschied bei dem Verständnis für "modisch" und "unterhaltsam" zu geben.

Prozentsatz der Befragten, die das Image der Zeitschrift zuordnen

	Männer 18 J. - 49 J. %		Frauen 18 J. - 49 J. %
<hr/>			
modisch			
Vogue	36	People Magazine	40
Elle	35	Vogue	39
Cosmopolitan Magazine	31	Elle	38
Bride's	30	GQ	37
GQ	28	Glamour	36
Glamour	27	Modern Bride	36
Redbook (Magazine)	24	Cosmopolitan Magazine	35
O (The Oprah Magazine)	23	Martha Stewart Living	34
Southern Living	23	Bride's	30
Mademoiselle	21	House Beautiful	29
<hr/>			
unterhaltsam			
Nick Jr.	59	Nick Jr.	59
Sesame Street Parents	40	People Magazine	45
Sports Illustrated	35	Rolling Stone	38
This Old House	33	Sesame Street Parents	37
American Baby	32	Entertainment Weekly	37
Rolling Stone	31	Walking	32
Hot Rod	31	Sports Illustrated	31
Family Fun	30	Family Fun	30
Automobile (Magazine)	29	This Old House	30
Road and Track	29	Better Homes & Gardens	29

Beim Alter scheinen die Gruppen bei Image- und Eigenschaftsstatements enger zusammen zu liegen. Beim Einkommen (50.000 \$ +) und in der Gruppe Frauen mit Kindern im Haushalt zeigen sich jedoch Unterschiede.

(Graph. Abb.)

Die vier Pfeiler von Young & Rubicam: Spezialisierung, Bedeutung, Wertschätzung, Kenntnis bilden wichtige Blöcke beim Marken-Portfolio-Management. Die Ergebnisse für Frauen-Zielgruppen variieren. Für Verlage ist das eine Marken-Aufbau-Übung, d.h. größere Kommunikationspotentiale für Werbungtreibende.

(Graph. Abb.)

Das folgende Beispiel zeigt Unterschiede zwischen den Ländern, die trotz größerer potentieller Einsparmöglichkeiten für globale Media-Investments das Sprichwort bestätigen: "Think globally, act locally".

Top Ten Medien nach Ländern

Venezuela	%	Mexiko	%	Kolumbien	%	Brasilien	%	Argentinien	%
RCTV	98	Vogue	99	Caracol	99	Claudia	88	Telefe	97
Diario de Caracas	96	Vanidades	99	Semana	96	Quatro Rodas	97	Clarín	94
Cosmopolitan	94	Eres	98	Jes	95	Time	87	El Gráfico	93
Venevisión	94	DirecTV	98	RCN	93	Elle	84	Canal 13	92
Reader's Digest	93	Sky TV	97	Crome	93	Criativa	83	HBO	92
El Nacional	91	Somo	97	El Tiempo	92	O Globo	82	Vogue	88
HBO	91	TV Azteca	97	RTI	92	Wall St. Journal	81	Para Ti	88
Televen	89	Cosmopolitan	97	CBS	91	Playboy	79	Time	87
Vogue	87	Cablevision	96	CNN	91	Veja	78	Reader's Digest	87
Playboy	85	TV y Novelas	95	OTI	91	Reader's Digest	78	Gente	83
El Universal	84	Elle	95	Playboy	89	BBC	75	Playboy	82

Anzahl der erhobenen Medien = Venezuela 24, Mexiko 60, Kolumbien 30, Brasilien 23, Argentinien 29

Quelle: Erwachsenen-Stichprobe, Y&R BAV, 1998/1999 Feldarbeit

Was Verbraucher über Media-Marken hinsichtlich Image- und Eigenschafts-Profilen denken, differenziert nach Geschlechtern. Die Vorstellung von Verbrauchern differenziert zwischen den Ländern genauso, wie die Mechanismen der Medienwirtschaft. Wenn jedoch eine Eigenschaft nicht zugeordnet wird, kann nicht das Gegenteil angenommen werden (gut/schlecht, mehr/weniger, größer/kleiner). Wenn 0% der Verbraucher denken eine Zeitschrift sei nicht modern, kann das heißen, dass sie meinen die Zeitschrift sei nicht modern aber auch, dass sie etwas anderes ist.

(Tab. mit Beispiel "unterhaltsam" und "intelligent")

Es gibt verschiedene Notwendigkeiten, die mit Medien befriedigt werden. Diese können in zwei Gruppen eingeteilt werden: Information und Kultur. Das erste beinhaltet Hilfestellung, Analyse, Aufklärung und Selbst-Erhöhung. Das zweite beinhaltet Rituale, Versäumnisse, Entspannung, Unterhaltung und Wirklichkeitsflucht.

Es ist OK 0% bei "modern" zu erhalten, wenn man "intelligente Information" oder "Analyse" oder "Aufklärung" zugeordnet bekommt. Obwohl es wahrscheinlich für den Verkauf von Werbeplätzen hilft, wenn man in seiner Gattung bei "modern" an erster Stelle steht. Ein Hauptergebnis der Studie war, dass bei den Zeitschriften niemand kulturelle Versäumnisse auswählte, wie das oftmals beim Fernsehen geschieht.

Gleichgewichtige Nutzung mit Währungserhebungen

Im täglichen Leben gibt es zwei Arten, wie Image und Eigenschaften genutzt werden können - die Marketing-Anwendungen und die Media-Planungs-Anwendungen.

Marketing-Anwendungen

Brand Asset Valuator
Diagnostisches Werkzeug
Wo bin ich? Wo möchte ich sein?
Marken-Objektivität
Greifbare Vorteile
Wie Marken aufgebaut werden

Media-Anwendungen

Media BAV
Beschreibungswerkzeug
Wie komme ich dahin?
Agentur-Objektivität
- Erreichung der Markenziele
Zeitschriften-Objektivität
- Dokumentation der Qualität der Leser

Die Aufgabe ist, ein Media-Portfolio auszuwählen, das die beworbene Marke unterstützt. Ein nützlicher Rahmen um Image- und Eigenschafts-Kenntnisse zu integrieren war es, beides, Produkt- und Media-Marken, in den Prozess einzubinden.

(Graph. Abb.)

Die Nutzung von Image- und Eigenschafts-Daten ist das Äquivalent für Medien, von der Bevorzugung von Größen- und Temperatur-Unterschieden. Es kommt manchmal zu fühlbaren Dissonanzen, wenn diese Daten in den tatsächlichen Planungsprozess einbezogen werden. Image-Zielgruppen werden von größeren Bevölkerungs-Segmenten gebildet als Zielgruppen über Reichweitenanalysen. Geschlecht, Einkommen und Länder beeinflussen feinere Aufteilungen bei kombinierten Definitionen, wie z.B. Frauen im Alter von 25 bis 54 J. mit Kindern und einem Haushalts-Einkommen von \$ 50.000,- +.

Die Dissonanz wird dadurch gefördert, dass die Nutzung qualitativer Daten in der Media-Planung und beim Media-Einkauf ein höheres Analyseniveau benötigt. Die beste Methode es durchzusetzen ist, es möglichst einfach zu halten. Dazu wurde eine Darstellung mit Quadranten übernommen, bei der der "gute" Media-Wert in der oberen rechten Ecke dargestellt wird.

(Graph. der theoretischen Struktur und eines praktischen Beispiels)

Der Prozess beginnt meistens mit einer Analyse der Planungsgruppe welche Marken-Werte die Hauptattribute bilden, um damit kongruente Medien zu finden. Oder die Werbung-treibenden haben eigene Image- und Eigenschaften-Studien, woraus die Profile gebildet werden können.

(Graph. zwei Beispiele)

Theoretische Perspektive

Das Konzept basiert auf psychologischen Prinzipien. Darüber gibt es so viele Diskussionen, wie Worte benutzt werden, um sie zu definieren. Eine Herausforderung bevor die Quantifizierung beginnt, ist es, die semantischen Begriffe festzulegen.

Das ausgedehnte Archiv qualitativer Forschung deckt fünf Gebiete ab:

Involvement (Engagement / Beziehung)

- Lesedauer, % der aufgeschlagenen Seiten, Heftherkunft, etwas ausprobiert nach dem Lesen

Attention (Aufmerksamkeit)

- Wiedererkennung, Erinnerung, gleichzeitige Aktivitäten, Leseintensität

Feeling (Gefühl)

- Image, Eigenschaften, Sympathie, Wertschätzung, Treue, Zufriedenheit

Need measures (Notwendigkeiten)

- Zweck der Medien im Leben der Personen

Human factor (Menschlichkeit)

- Größe, Position, Farbigkeit und Tastbarkeit, physische Elemente der Lesegelegenheiten

Es gibt verschiedene Studien, die diese Aspekte untersuchen. Die qualitative Denkschiene muss wichtig sein und hat wahrscheinlich einen Platz im Media-Planungs- und Einkaufs-Prozess oder mindestens im Bereich der anerkannten Media-Forschung. Wenn Daten und Methoden verfügbar sind, um die Media-Entscheidung zu fördern, sollte man sie nutzen.

Schlussbemerkungen

Es gibt einige wundervolle Studien und Rahmen um über Lesequalität nachzudenken. Die Image- und Eigenschafts-Methode, die hier beschrieben wird, bietet eine Option um theoretische Grundlagen der Verbindungen zwischen Verbrauchern und Zeitschriften nutzen zu können. Es ist schwer zu glauben, dass Ergebnisse über Lesequalitäten so wenig genutzt werden, obwohl es so viele verfügbare Daten gibt.

Dass Media-Budgets jetzt auf der Basis seriöser ökonomischer Erhebungen des Return on Investment vergeben werden können, ist eine Herausforderung. Das verursacht einen Verteilungsprozess in dem Leser pro Nummer oder die Frage nach dem letzten Lesen schlecht vorbereitet sind, um das bieten zu können - ungeachtet der neueren Entwicklungen in der Leserschafts-Kumulationsforschung.

Langfristig wird dieses Defizit eventuell durch den Einsatz von People-Metern verschwinden. Bis dahin bietet Zeitschriften-Forschung die Möglichkeit, einen der größten Hauptnutzen darzustellen: Die Umstände in denen das Lesen stattfindet.

Quality Rating Point - The new currency

Quality Rating Point - Die neue Währung

Morten Kromann-Larsen & Rolf Randrup, Taylor Nelson Sofres / Gallup

Dieses Paper ist eine direkte Weiterentwicklung von Vorträgen auf dem World Wide - Readership Symposien 1997 und 1999, in denen die Entwicklung der nationalen dänischen Leseranalyse in ein dynamisches Instrument beschrieben wurde. Der Hintergrund ist, dass Print-Medien der TV-Konkurrenz, die viel detailliertere und aktuellere Ergebnisse anbieten kann, etwas entgegengesetzt sein müssen. Im Vergleich zu den GRPs (Gross Rating Points = Brutto-Reichweite) des Fernsehens muss die traditionelle Print-Forschung, die Leserschaftsdaten der vergangenen sechs oder zwölf Monaten einsammelt, mehr als nur kosmetisch aufgewertet werden.

Um den Trend von Print zu TV aufzuhalten, wurden in der National Readership Survey, Index Danmark / Gallup für Tageszeitungen einige Erhebungen aufgenommen:

- 1) Erhebung von Leserschaften für Tageszeitungs-Sektionen (-Büchern)
- 2) Erhebung der Frequenzen für das Lesen der Tageszeitungs-Sektionen und Integration in die NRS
- 3) Die Berichtsperiode wurde von sechs Monate in monatlich geändert
- 4) Einführung von Quality Rating Points, einer neuen Währung für die Mediaplanung bei Nutzung von Monatsdaten, Wochentagsdaten, Sektionen und Inhalt.

1. Erhebung der Tageszeitungs-Sektionen-Leserschaft

In Dänemark sind alle Tageszeitungen und Sonntagszeitungen in Sektionen aufgeteilt. Diese können unabhängige Einheiten sein, die ihre eigene Leserschaft haben. Diese Entwicklung zeigt, dass die traditionellen Leserschaftswerte nicht mehr verwendet werden konnten, sondern für jede Sektion, für die Planung, eigene Werte zur Verfügung stehen mussten.

Die Frequenzfragen waren dabei ein Problem. Für große Sektionen mit eigenem Namen kann die Abfrage funktionieren, bei anderen entsteht eine extreme Überschätzung. Deshalb wurden für das Lesen von Sektionen diskriminierende, beschreibende Variable entwickelt. Mit einer CHAID-Analyse werden die Ergebnisse dieser separaten Erhebung in die NRS transferiert. Damit können dann Zielgruppen-Planungen durchgeführt werden.

2. Produktentwicklung der National Readership Survey

Obwohl diese Aktivitäten bereits große Schritte waren, um den Anforderungen des Marktes gerecht werden zu können, sind die angebotenen Daten des Fernsehens erheblich reichhaltiger. Zunächst war es dafür notwendig, das Berichtsintervall zu ändern und aktueller zu machen. Außerdem wurde auch die Reichweite für jede Seite angeboten und nicht nur für Sektionen. Das letzte Angebot war die Einbeziehung von Quality Rating Points. Das Modell dafür beinhaltete die NRS-Daten, Leserschaftsdaten aufgesplittet nach Wochentagen und Monaten und das Leseverhalten nach Seiten innerhalb der Sektionen. Daraus folgt, dass

separate Untersuchungen zum Leseverhalten pro Seite in die NRS eingebracht werden müssen.

3. Monatliche Leserschafts-Daten

Um monatliche aktuelle Reichweiten ausweisen zu können, gleichzeitig aber auch die Frequenzfrage beibehalten zu können, werden die Leser pro Nummer-Reichweiten, die aktuell erhoben werden, mit den älteren Leser pro Nummer-Reichweiten und den Frequenzen verrechnet:

Beispiel: "Offizielle" Reichweite nach Frequenzdaten, 2. Halbjahr 2000	450.000	
Reichweite nach der Frage nach dem letzten Lesen, 2. Halbjahr 2000	470.000	
Reichweite nach der Frage nach dem letzten Lesen, Januar 2001	485.000	
Justierte "offizielle" Reichweite, Januar 2001	$450.000 \times \frac{485.000}{470.000}$	= 464.000

(Tab. mit monatlichen Reichweiten von dänischen Tageszeitungen)

4. Reichweiten-Daten für Wochentage

Die ideale "Waffe" gegen die TV-Kanäle wäre ein "Print-People-Meter", das die Leserschaften für jeden Tag misst. Leserschaften nach Wochentagen können gebildet werden, aber nicht innerhalb jeden Monats. Es werden die Daten der 6-Monats-Zeiträume benutzt und auf Wochentage heruntergebrochen. Dazu wird jedoch eine einwandfreie Tages-Stichprobe benötigt. Damit wird zwar das normale durchschnittliche Verhalten erhoben und nicht das jetzt aktuelle, es wird jedoch angenommen, dass die Verteilung über die Wochentage ziemlich stabil ist, so dass die Daten in die Monats-Daten transferiert werden können.

5. Datenerhebung für Quality Rating Points

Die noch ausstehende Komponente für die QRP-Berechnung ist die Leserschaft pro Seite. Dieser Teil wird durch eine große Leseverhaltensuntersuchung mit mindestens 20.000 Interviews erhoben, die das Lesen von Einzelseiten erfasst. Die Feldarbeit wurde von Oktober 2000 bis Februar 2001 durchgeführt. Alle 14 Tageszeitungen nahmen an dieser Untersuchung teil. Die Interviewzahl pro Tageszeitung betrug 2.000 für nationale Tageszeitungen und 1.000 für regionale Tageszeitungen. Die Erhebung fand an allen Wochentagen statt und dauerte mindestens acht Wochen.

Bei der Datenerhebung wurde für jeden Wochentag ein neuer Fragebogen benötigt, der den aktuellen Inhalt der Tageszeitung beinhaltet. Dadurch waren face-to-face Interviews nicht möglich, sondern es wurden CATI-Interviews durchgeführt. Dadurch werden allerdings sowohl Leser, die die Zeitung beim Zeitpunkt des Anrufes nicht mehr vorliegen haben

ausgeschlossen, als auch Leser, die die Zeitung außer Haus lesen. Durchschnittlich lesen 15% die Tageszeitung nicht zu Hause sondern am Arbeitsplatz. Durch die Kenntnis der demographischen Struktur von Zuhause- und Außer-Haus-Lesern, ist es möglich, die Reichweiten pro Seite zu gewichten. Das ist nicht befriedigend, so dass in Zukunft auch die Leserschaften am Arbeitsplatz einbezogen werden müssen.

Die Stichprobe der regionalen Tageszeitungen basieren auf Abonnenten-Listen, da 90% der Auflage abonniert ist. Die Leser der nationalen Tageszeitungen werden per Telefon-Screening ermittelt.

Das Interview beinhaltet für jede Seite:

Seite nicht gesehen

Seite gesehen oder einen Blick auf die Seite geworfen

Seite teilweise gelesen

Alles oder fast alles der Seite gelesen

6. Datenaufbereitung und Berichterstattung

Mit der QRP-Untersuchung sollen Leserschaftsdaten pro Seite oder von mehreren Seiten dargestellt werden. Das ist ein sehr komplizierter Prozess. Auch qualitative Aspekte müssen dabei berücksichtigt werden.

Grundsätzliche Berechnungsregeln

Die Daten werden an die Strukturen der NRS gewichtet. Für jede Tageszeitung werden die Anzahl der Leser pro Seite berechnet, ohne diejenigen Seiten, die nicht in der Gesamtauflage enthalten sind. Danach werden die Daten von einer befragtenorientierten Datenbasis in eine seitenorientierte Datenbasis transponiert. Für alle Tageszeitungen werden die Prozentwerte für alle Einzelseiten und für Gruppen von Seiten in der Untersuchung dargestellt und auf die Gesamtleserschaft pro Seite indiziert. Diese Indizes werden als Quality Rating Points dargestellt und die zukünftige Währung im dänischen Print-Markt sein.

Aber was definiert einen Leser?

Wie die Fragestellung zeigt, gibt es mehrere Abfragekategorien pro Seite. Seite nicht gelesen heißt, dass es sich um einen Nicht-Leser dieser Seite handelt.

Leserschaft pro Seite in drei dänischen Tageszeitungen

		A		B		C	
		%	Ind.	%	Ind.	%	Ind.
Nachrichten-Sektion	Leserschaft gesamt	74	100	75	100	76	100
	Seite gesehen oder einen Blick auf die Seite geworfen	32	44	36	48	33	43
	Seite teilweise gelesen	28	38	24	32	28	37
	Alles oder fast alles auf der Seite gelesen	14	18	15	20	15	20
Alle Sektionen	Leserschaft gesamt	52	100	56	100	54	100
	Seite gesehen oder einen Blick auf die Seite geworfen	25	47	27	48	26	49
	Seite teilweise gelesen	17	33	17	30	17	31
	Alles oder fast alles auf der Seite gelesen	10	20	12	22	11	20

Es ist klar, dass das Lesen der Nachrichten-Sektion insgesamt hauptsächlich mit fast 50% von der Ausprägung "Seite gesehen oder einen Blick darauf geworfen" resultiert. Dieses Phänomen ist über alle Seiten identisch. Insgesamt ist das Leserschaftsniveau vielleicht geringer, aber der Anteil der niedrigen Leserschaftsausprägung ist nahezu identisch.

Analog der Prinzipien für den Leser pro Nummer könnte man argumentieren, dass alle Leserschaftsgruppen zur Definition des Lesers herangezogen werden sollten. Die Seite wurde aufgeschlagen und ob etwas gelesen wurde, hängt vom Inhalt ab. Die Konsequenz ist ein ziemlich hohes Leserschaftsniveau pro Seite. In Zukunft kann die Definition eventuell geändert oder auch gewichtet werden, je nachdem, wieviel gelesen wurde.

Seiten-Werte

In Diagrammen werden nachfolgend die Ergebnisse für Einzelseiten der dänischen nationalen Tageszeitungen gegenübergestellt. Auffällig ist dabei ein deutlicher Abfall auf den Seiten drei und fünf, was damit begründet wird, dass es sich um die ersten ganzseitigen Anzeigen handelt, die erfahrungsgemäß immer niedriger bewertet werden als Seiten, die auch Redaktion enthalten.

(Diagramm)

In einem zweiten Diagramm wird die Lesekurve für die Heft-Viertel dargestellt, die auf ähnlichem Niveau für alle drei Titel leicht absinkt.

(Diagramm)

Redaktioneller Inhalt

Fast 60.000 Seiten wurden inhaltlich in ca. 40 Kategorien vercodet. Damit können die Seiten-Ergebnisse nach allen inhaltlichen Kategorien dargestellt werden. Diese wurden auf ca. 20 begrenzt.

(Tab. mit Indizes)

Auffällig ist, dass die Daten für die gleichen Sektionen in allen drei nationalen Tageszeitungen ähnliche Indizes aufweisen.

7. Schlussbemerkung

Quality Rating Points werden in Dänemark als neue Währung eingeführt werden. Die Danish Newspaper Association hat beschlossen, die Feldarbeit Ende 2000 und Anfang 2001 durchzuführen. Viele Tests mit den QRP sind durchgeführt worden. Zusammen mit der Einführung der QRP wurde eine web-basierte Software, die Kampagnen-Simulationen vornehmen kann und die Buchung einfach macht, installiert.

Was noch fehlt, ist die Implementierung von Monats- und Wochentags-Reichweiten in die QRP-Formel. Außerdem sollen QRPs in die NRS aufgenommen werden.

Using the 24-hour day as the common denominator for cross-media comparisons: The case of Multi-Media Mentor

Nutzung des 24-Stunden-Tages als gemeinsamen Nenner für Cross-Media-Vergleiche: Der Multi-Media Mentor

Scott C. McDonald, Ph.D., Condé Nast Publications

Cross-Media Planungs-Werkzeuge

Media-Mix und Cross-Media ist jetzt groß in Mode. Die großen Agenturen entwickeln Media-Mix-Optimierungsprogramme und man kann kaum eine Ausgabe von Advertising Age oder anderen Fachzeitschriften aufschlagen, ohne einen Beitrag über Cross-Media-Plattformen zu finden. Es ist unbestritten, dass ein breiterer Media-Mix geeignet ist, Reichweite und Kontakt kostengünstiger aufzubauen. Fernsehen ist dabei allerdings meistens am teuersten. In den USA stiegen die TV-Preise zwischen 1998 und 2000 um 50% an.

Media-Mix-Planung durchzuführen ist jedoch nicht so einfach. Die Währungen sind unterschiedlich und nicht direkt vergleichbar. In den USA ist auch keine geeignete zusammengefaßte Datenbasis erhältlich. Außerdem wird der Media-Markt zunehmend fragmentierter. Es gibt immer mehr spezialisierte TV- und Radio-Sender, Nischen-Zeitschriften und ein riesiges Online-Medium. Es ist schwieriger kleine Medien (geringe Nutzerschaft) abzufragen als große (große Nutzerschaft). Und so ist wirkliche Media-Mix-Planung mehr zum scheitern verurteilt, als 10 Jahre zuvor. Außerdem nutzen die Verbraucher die Medien auch noch gleichzeitig, so dass es noch einmal schwieriger wird.

MultiMedia Mentor

Statistical Research (SRI) hat kürzlich in den USA ein Produkt vorgestellt, um die Anforderungen an Media-Mix-Planungen zu erfüllen. Das Produkt heißt MultiMedia Mentor. Dabei wird festgehalten, wie die Leute verschiedene Medien nutzen. Allerdings liegen die Wurzeln von SRI in der TV-Forschung und bisher wurden TV, Radio und neue Technologien untersucht, aber nur wenig Print-Medien. Dieses soll sich aber bald ändern. Pilot-Studien wurden im Jahr 2000 durchgeführt. Alle Daten werden telefonisch erhoben. Grundgesamtheit ist die Bevölkerung ab 12 Jahren. Zur Zeit werden 5.000 Interviews pro Jahr in einem rollierenden System durchgeführt.

Die Befragten werden nach der Nutzung von TV, Radio, Online / Internet, Tageszeitungen und Zeitschriften gefragt. Dabei werden sechs Tagesabschnitte einbezogen (6-9 h, 9-12 h, 12-15 h, 15-18 h, 18-21 h, 21-24 h). Wenn eine Media-Nutzung zu der entsprechenden Zeit angegeben wird, wird dafür ausführlich weitergefragt: wie lange jedes Medium genutzt wurde, ob das Medium einzeln oder zusammen mit andere Medien genutzt wurde, ob Nebentätigkeiten gemacht wurden, ob es absichtlich oder zufällig genutzt wurde, ob es zu Hause oder außer Haus war, usw. Das Interview schließt mit ausführlichen demographischen Ermittlungen und Produktnutzungs-Informationen. Im Durchschnitt dauert das Interview 19 Minuten; Personen mit niedriger Media-Nutzung benötigen nur 7 Minuten und Personen mit hoher Medien-Nutzung bis zu 30 Minuten.

Mentor bietet für diese Daten auch eine Planungssoftware an, die das Werbebudget auf die fünf Medien verteilt - abhängig von den Parametern, die der Planer vergibt. Wenn z.B. ein Medium nicht belegt werden soll, weil es nicht zur Kampagne passt, kann "0" eingegeben werden oder man kann für ein Medium einen vorher bestimmten Prozentsatz des Budgets eingeben. Außerdem können die erhobenen Variablen, wie z.B. separate Nutzung des Mediums für eine Gewichtung der verschiedenen Medien genutzt werden. Es ist eine Optimierung nach Reichweite oder nach Kosten vorgesehen.

Beispiel: Frauen mit Computer-Affinität, 18-49 J., 4-Wo.-Budget 1.500.000 \$, höchste Reichweite

	TRPs	%	Dollars
TV Tagsüber	58	23	420.978
Radio	57	23	284.589
Zeitschriften	94	37	569.190
Online	42	17	224.917
Gesamt	251		1.499.678
Erreichte 4-Wochen-Reichweite: 76%			

Mentor empfiehlt keine Einzel-Titel oder Sender sondern schlägt eine Diversifikation des Geldes über die Medien hinweg vor. Das gilt vor allen Dingen, wenn man über die Kosten optimiert, da TV so teuer ist.

Beispiel: Kredit-Karten-Nutzer (3mal und mehr die Woche), 18-64 J., Ziel: 65% Reichweite in 4 Wochen zu niedrigsten Kosten

	TRPs	%	Dollars
TV Frühmorgens	19	13	417.121
TV Primetime	28	19	846.080
Radio	60	41	634.296
Tageszeitungen	2	2	71.637
Zeitschriften	11	8	154.742
Online	27	18	286.560
Geringste Kosten			2.410.436

Die geringen Radio- und Online-Kosten geben den beiden Medien in diesem Plan mehr Gewicht als es sonst der Fall hätte sein können. Beide Medien sind jedoch national schwierig zu belegen, so hätte man auch im Vorfeld Limits eingeben können.

Beispiel: Flugreisende (3 oder mehr Flüge im letzten Jahr), 18-64 J.,
 Budget: 4 Wochen \$ 2.000.000
 Ziel: größtmögliche Reichweite

	TRPs	%	Dollars
TV Frühmorgens	15	10	348.508
TV Primetime	12	8	376.308
Radio	60	40	661.081
Zeitschriften	31	21	322.072
Online	31	21	244.092
Gesamt	150	100	1.952.061
4-Wochen-Reichweite: 67%			

Die jetzige Media-Planung denkt nicht an eine derartige breite Streuung der Dollars über die Medien. Meistens werden von vornherein Aufteilungen vorgegeben - oder "Planung" besteht nur darin, dass meiste Geld auf TV zu setzen und ein bißchen, was übrig ist, je nach Vorlieben für Print, Radio oder Online zu nutzen.

Ich erwähnte, dass Mentor auch die Nutzungsvariablen für die Medien einbeziehen kann und diese gewichtet werden können. Kontakte, die erzielt werden, können unterschiedliche sein: von Medium zu Medium, aber auch innerhalb der Medien. Manche Werbe-Botschaften werden besser abends aufgenommen, manche sind mehr für Print geeignet und manche mehr für elektronische Medien. Manche Verbraucher sind dem Medium gegenüber aufmerksamer und manche sind weniger aufmerksam. SRI fand heraus, dass Print-Medien weniger Zeit in Anspruch nehmen aber diese Zeit dann nur Print-Medien gewidmet wird. Radio nimmt mehr Zeit in Anspruch, läuft aber meistens nebenher, mit wenig Aufmerksamkeit daran. Medien-Nutzung, die absichtlich erfolgt, ist einflußreicher als Medien-Nutzung, die zufällig, d.h. unbeabsichtigt erfolgt.

Wenn die Leute alleine sind, ist ihre Aufmerksamkeit gegenüber dem Medium größer, als wenn sie zusammen mit anderen Personen sind (besonders, wenn die anderen Kinder sind, die Aufmerksamkeit verlangen). Medien-Konsum zu Hause ist intensiver als Medien-Konsum außer Haus. Einzel-Medien-Nutzung bringt mehr Aufmerksamkeit als wenn man parallel mehrere Medien nutzt oder Tätigkeiten nebenbei erledigt.

(Graph. mit Ergebnissen)

Die dargestellten Ergebnisse überraschen nicht. Print-Medien werden eher einzeln genutzt, TV und Radio zusammen mit anderen Medien. Internet wird meistens exklusiv genutzt. Internet und Zeitschriften werden am ehesten alleine genutzt, ohne das Beisein anderer Personen. TV und Zeitschriften werden am meisten zu Hause genutzt, usw.

Mentor stellt aktuelle Informationen über die unterschiedlichen Umstände, unter denen Kontakte mit den Medien stattfinden, zur Verfügung. Die Gewichtung der Wertigkeit dieser Kontakte obliegt jedoch dem Planer.

MultiMedia Mentor im Zusammenhang

Mentor benutzt Zeit als gemeinsamen Nenner für den Cross-Media-Vergleich. Es wird nach der Media-Nutzung "gestern" gefragt. Insofern gibt es immer einige Vergleiche zur "Gestern-zum-ersten-Mal-gelesen-Methode" (FRY). Der SRI-Ansatz stellt keinen exzessiven Anspruch an die Erinnerungsleistung der Befragten und er belästigt die Befragten nicht mit dem Ausfüllen von Tagebüchern. Das Zeitbudget für Media-Nutzung ist ein fairer und logischer gemeinsamer Nenner für Cross-Media-Vergleiche, besonders dann, wenn auch noch die erwähnten Nutzungsvariablen erhoben werden.

Die Methode lässt jedoch das Niveau der Print-Nutzung im Vergleich zu TV, Radio und Online sehr armselig aussehen. Das ist eine Reflektion der Realität. Man darf jedoch nicht vergessen, auf welche Art und Weise manche Zeitschriften ihre Leserschaft über lange Zeiträume einsammeln. MRIs jüngste Kumulations-Untersuchung sagt, dass wöchentliche Nachrichtenmagazine vier Wochen benötigen, um 90% ihrer Gesamtleerschaft zu erreichen. Titel mit hohen Leser-pro-Exemplar-Werten kumulieren ihre Leserschaft sogar noch langsamer. "People" erreicht nur 46% der Leserschaft in der ersten Woche, wo sie erhältlich ist. Frauen-Zeitschriften und Titel, wie Readers Digest und National Geographic sind sogar noch statischer in ihrem Leserschafts-Aufbau.

Im MultiMedia Mentor wird nicht nach Einzel-Titeln gefragt und es ist keine Reichweiten-Untersuchung. Es wird einfach nur die Zeit festgehalten, die für die fünf Medien pro Tag (gestern) aufgewendet wird. Das Planungsinstrument nutzt die zeitbasierten Daten ähnlich wie Reichweiten und die Beispiele haben gezeigt, dass Planer nach Reichweiten optimieren können. Und Kosten und Reichweiten sind die beiden wirklichen Nenner, nach denen Planer ihre Mediapläne aufbauen wollen.

Und Mentor berichtet Print mit Target Rating Points (TRPs). Das bringt Print von Einschaltungen in mehr planungsrelevante Einheiten, wie sie auch von den anderen Medien genutzt werden.

Time related measures of relative 'effect' - An essential ingredient for Multi-Media campaign evaluation

Zeitbezogene Erhebungen des 'relativen Effekts' - Eine wesentliche Voraussetzung zur Berechnung von Media-Mix-Kampagnen

Peter Masson, Bucknull and Masson/Peter Callius, Research International

Marketing-Direktoren müssen ihre Budgets immer mehr rechtfertigen. Sie stehen einer sich rapide verändernden Medialandschaft gegenüber. Die Fragmentierung und die wachsende Zahl an Medien bedeutet, dass die Zeit, die einem einzelnen Medium gewidmet wird, abnimmt. Selbstverständlich gab es schon immer die Möglichkeit, die Effektivität von Media-Mix-Kampagnen über aggregierte Trackings herauszufinden. Nachträgliche Auswertungen helfen uns zwar beim Verständnis, sind aber schwierig in den Planungsprozess zu integrieren. Strategische Planungen der verschiedenen Media-Gruppen waren schon immer möglich. Auf der taktischen Ebene mussten wir uns aber auf die Evaluation und Optimierung einzelner Mediagattungen verlassen. Es ist augenscheinlich, dass über Media-Mix-Wirkung mehr herausgefunden werden muss. Media-Mix-Entscheidungen werden oftmals nicht durch Forschungsergebnisse gestützt. Werbungtreibende drängen darauf, mehr über Media-Mix-Planungen zu verstehen und den ROI ihrer integrierten Kommunikation zu evaluieren.

Politische und kommerzielle Interessen sowie methodische Fragen haben die Entwicklung bisher behindert. Die Interessen von Werbungtreibenden, Planern, Forschern, Werbeagenturen und den Medien sind sehr verschieden.

Es gibt verschiedene Alternativen: die Annahme der Unabhängigkeit der Medien, Single-Source-Ansätze und Fusionen. Bei der Annahme der Unabhängigkeit gehen wir davon aus, dass die Werbewirkung einer Mix-Kampagne durch Aggregation von Einzelanalysen der Medien berechnet werden kann. Fusionen werden immer beliebter. Durch die Zurverfügungstellung von Roh-Datensätzen gibt es interessante Möglichkeiten, die weiter verfolgt werden sollten. Der Single-Source-Ansatz birgt eigene Probleme. Der Vorteil, möglichst viele Daten von einer Person zu ermitteln, liegt auf der Hand, ist aber schwierig durchzuführen, da die Interviews zu lang werden.

In Schweden haben wir uns für den Single-Source-Ansatz entschieden. Die schwedische Media-Studie Orvesto Consumer ist eine zweistufige postalische Untersuchung. Es werden in einem Random-Sample ca. 40.000 Interviews jährlich durchgeführt. In der ersten Stufe werden Print-, allgemeine Fernseh- und Radio-Daten erhoben. In der zweiten Stufe werden in einem 7-Tage-Tagebuch mit 1/4-Std. Radio- und Fernseh-Daten beim gleichen Befragten erhoben. Die Ausschöpfungsquote beträgt ca. 65% in der ersten Stufe und ca. 70% in der zweiten Stufe.

Nachdem die Untersuchungsmethode feststand mussten einige Annahmen getroffen werden. Dafür war ein solides Media-Planungs-Tool notwendig. In unserem Fall SESAME.

Werbewert

Die (Werbe-)Kontakte, die für die verschiedenen Mediagattungen benutzt werden, sind unterschiedlich. Die Chance von "offenen Augen und Ohren" bei dem Kontakt mit Werbung werden bei Print durch den Kontakt mit "irgendeinem Teil der Zeitschrift oder der Tageszeitung der Werbung beinhaltet" ausgedrückt; beim Fernsehen heißt es "vor dem eingeschalteten Fernseher bei dem Sender während einer Pause mit Werbung anwesend gewesen".

Um den Wert der Kontaktchancen anzugleichen, muss der Planer in der Lage sein, die Kontaktwahrscheinlichkeiten zu gewichten. Die Angabe 15 Minuten ferngesehen in dem Orvesto TV-Tagebuch wird modelliert zum durchschnittlichen Rating für 1 Minute. Das durchschnittliche 1-Minuten-Rating kann aber nicht mit Print-Reichweiten verglichen werden. Der Planer kann in diesem Fall einen Seiten-Faktor für gesehene Seiten hinzuaddieren, um die Anzeigen-Kontaktchancen zu berücksichtigen. Diese Modifikationen sollten in Abhängigkeit der Lesefrequenz vorgenommen werden, da hohe Werte für gelesene Seiten mit der Leseintensität und der Leser-Blatt-Bindung korrelieren. Entscheidend ist, dass diese Gewichtung lediglich die Kontaktchance für die Werbung beinhaltet und nicht den Kontakt selber.

Berechnung eines Plans mit Media-Mix-Kontakten

In jeder Kampagnen-Analyse ist der erste Schritt, die vorhandenen quantitativen Mediaforschungs-Daten zu nutzen, um damit einen Plan zu erstellen, mit dem die Kontaktwahrscheinlichkeiten für Zielgruppen dargestellt werden. Dabei werden Frequenz-Verteilungen zugrunde gelegt. Diese repräsentieren jedoch keine realistische Situation. Besonders wenn sie für Optimierungen genutzt werden, führen solche einfachen Frequenz-Verteilungen zu unkorrekten Media-Entscheidungen.

In SESAME stehen zwei verschiedene individuelle Modelle für TV und Print zur Verfügung. Da die beiden Frequenz-Verteilungen auf der individuellen Informations-Ebene berechnet werden, ist es in Ordnung damit akkurate Frequenz-Verteilungen für Media-Mix zu errechnen. Diese individuelle Informations-Ebene erlaubt zudem, Wirkungskurven für unterschiedliche Zeiten und Medien und/oder für unterschiedliche Marktsegmente im Plan darzustellen.

(Abb. Media-Mix-Plan mit SESAME produziert:
zweidimensionale Frequenz-Verteilungen nach Zeiten)

SESAME erstellt einen detaillierten Plan der Zielgruppen, die nur mit Print, nur mit TV und mit beiden erreicht werden, so z.B. mit 5 TV-Kontakten, 2 Print-Kontakten und welchen Anteil 1 TV-Kontakt und 6 Print-Kontakte erhält. Jetzt geht es darum, diesen Kontakt-Plan von SESAME zu interpretieren. Das einfachste Maß ist, dass ein (Print- oder TV-)Plan gegen einen anderen gesetzt wird.

Die Matrix - Relative Media-Effekte & Zielgruppen-Wirkung

Im folgenden wird ein Beispiel gezeigt, wie eine Wirkungskurve aufgebaut werden kann und für einen Kontaktplan (Frequenz-Verteilungsplan) angewendet werden kann. Die Annahme ist, dass 25% (der Zielgruppe) mit einer Kontaktchance die Möglichkeit aufgreift und

den Werbeinhalt "lernt", dass 50% mit zwei Kontaktchancen dieses tut, 75% mit drei Kontaktchancen und 100% mit vier Kontaktchancen. Mit diesen Annahmen haben wir eine Wirkungskurve definiert. Das kann genutzt werden, um verschiedene Pläne zu evaluieren, die jeweils mit der Plan-Reichweite, die über die Frequenz-Verteilungen erreicht werden, gewichtet werden. Beispiel:

Kontaktchancen (OTS)	'Kontakt'		% der Zielgruppe auf jeder OTS-Ebene (Frequenz-Verteilung, FD)			
	% die die Kontaktchance wahrgenommen haben		Plan A		Plan B	
			%	E.R.	%	E.R.
Keiner	0		50 =	0	40 =	0
1	25		10 =	2,5	30 =	7,5
2	50		12 =	6,0	20 =	10,0
3	75		20 =	15,0	8 =	6,0
4	100		8 =	8,0	4 =	4,0
Effektive Reichweite (E.R.)	Gesamt			31,5		27,5

(Weiteres Beispiel mit anderen Verteilungen)

Um jedoch einen Plan mit Print- und TV-Mix evaluieren zu können, wird eine Reihe von Wirkungskurven für jede Medienkombination auf jeder Wirkungsebene benötigt. Das kann manuell nicht erstellt werden. SESAME errechnet dieses für den Planer.

(2 Abb. mit Beispiel-Kurven)

Außerdem lässt sich das Ganze auch für unterschiedliche Bevölkerungssegmente darstellen. Dabei spielt die Definition der Zielgruppe eine große Rolle. Wie weit oder eng die Marketing-Strategie gezogen wird, wird die Wirkungskurve beeinflussen (genauso, wie die Budget-Voraussetzungen, die Kreativität der Werbung und die Media-Auswahl).

(Beispiel mit Nutzern zweier Marken)

Synergie-Effekte

Ein wichtiger Grund für Media-Mix sind Synergie-Effekte. Medien sind in ihrem Kommunikationseffekt nicht unabhängig. Kontakte mit Werbung in TV können Kontakte mit der Werbung in Print beeinflussen. Synergie-Effekte müssen daher Teil des Modells sein. Wir können eine Reihe von möglichen Wegen identifizieren, wie die räumliche Positionierungsnähe von Kontakten mit Medien mit verschiedenen Kommunikations-Möglichkeiten auch Kommunikations-Synergien erbringen. (Aufzählung verschiedener Synergien: Wirkung-Synergie, Inhalts-Synergie, Synergien zur Erinnerungsverstärkung, Stimulierungs-Synergien)

In SESAME kann der Planer einen Synergie-Faktor eingeben, der die Media-Mix-Kontakte erhöht.

Zeitliche Verbreitung der Kontakte

Die Zeitverzögerung zwischen den verschiedenen Kontakten beinhaltet Erinnerungslücken und muss ebenfalls berücksichtigt werden. Bis jetzt wurde der Faktor Zeit ignoriert. Media-Mix-Kampagnen beinhalten generell das Nebeneinanderstellen der Medien im Zeitverlauf. Wirkung (sowohl lernen als auch vergessen) beruht auf der zeitlichen Streuung der Kontakte und Erhebungen sollten das berücksichtigen. Analyse-Systeme müssen dem Planer dieses ermöglichen. In einer Plan-Evaluation wird sowohl eine "Lernkurve" als auch eine "Vergessenskurve", abhängig von der zeitlichen Streuung der Kontaktchancen, benötigt werden.

(Abb. Graphiken Zielske's Experiment)

Zielske führte 1959 ein Experiment über das Lernen und Vergessen von Werbebotschaften bei verschiedenen zeitlichen Streuungen durch. Zwei vergleichbaren Stichproben wurden 13 Anzeigen per Post zugeschickt. Einmal im Abstand von einer Woche und einmal im Abstand von vier Wochen. Dabei wurde festgestellt, dass bei der Vier-Wochen-Gruppe das "Wissen" erheblich langsamer zunimmt und auch in der Zwischenzeit zwischen den Kontakten erheblich zurückgeht. Ungefähr 65% des "Gelernten" wurde zwischen zwei Kontakten wieder vergessen. Diejenigen, die wöchentlich die Werbung erhielten, vergaßen ca. 50% innerhalb von vier Wochen nach der letzten Zusendung.

Die Ebbinghaus-Verhaltenskurve (aus den frühen 1900er-Jahren) dokumentiert den rapiden Verlust an Erinnerung.

Ebbinghaus: 40% Verlust in 20 Minuten
55% Verlust in 60 Minuten
66% Verlust in 24 Stunden
75% Verlust in 7 Tagen
80% Verlust in 30 Tagen

Es ist anzunehmen, dass dieser Verlust auch auf die fehlende Relevanz des Gelernten zurückzuführen ist (Zufallszahlenreihen und Unsinnswörter / -sätze). Bei der Zielske-Kurve handelte es sich hingegen um einen Durchschnitt von Personen, für die die Werbebotschaft hohe und geringe Relevanz hatte. Eine Kurve für diejenigen, für die die Werbebotschaft eine hohe Relevanz hat, würde wahrscheinlich schneller ansteigen und das Behalten der Informationen würde länger andauern. Unsere Hypothese lautet daher, dass die Segmentierung der Zielgruppe in Wirkungsgruppen essentieller Bestandteil der Plan-Evaluation ist. Die Frage ist, welche zeitlichen Abstände gewählt werden müssen. Wir schlagen vor, die Zielske Wirkungskurve für die wöchentliche Aussendung zu übernehmen.

Im SESAME-Modell wird die "Effektive Reichweite" auf einer täglichen Basis evaluiert und die Vergessenskurve für die folgenden Tage angewandt.

Tägliche "Effektive Reichweite" abgestuft nach Tagen (Vergessenskurve)

Tag	1	2	3	4	5	6	7	8	Gesamt (8 Tage)
Tag 1	5,4	5,1	4,9	4,6	4,3	3,8	3,2	2,7	
Tag 2		2,1	2,0	1,9	1,7	1,5	1,3	1,0	
Tag 3			6,2	5,9	5,3	4,3	3,7	3,1	
Tag 4				1,3	1,2	1,1	1,0	0,8	
Zeitliche Effektive Reichweite (TER)	5,4	7,2	13,1	13,7	12,5	10,7	9,2	7,6	79,4

Diese Evaluation der täglichen Kontakt-Frequenz-Verteilung muss innerhalb der verschiedenen Wirkungsgruppen der Medien durchgeführt werden. Die zusammengefasste Plan-Vorstellung wird durch die "Zeitliche Effektive Reichweite" (TER = Timed Effective Reach) aller Plan-Komponenten dargestellt.

Die Kontakt-Zeit und -Streuung dieser "Hits" steht für TV und Radio sowie für einige Tageszeitungs-Forschungsergebnisse zur Verfügung, jedoch nicht für Zeitschriften. Interessante Forschungen zu diesem Thema sind in Belgien und den USA durchgeführt worden. Es kann ein wesentlicher Nachteil für die Zeitschriften sein, wenn sie diese Forschung nicht aufgreifen und nicht in der Lage sind, zeitliche Verteilungen der Anzeigen-"Hits" zu liefern.

Zusammenfassung

Dieser Analyse folgend ist klar, dass eine Evaluation von Media-Mix-Plänen folgendes voraussetzt:

1. Die Möglichkeit, den Plan mit der Verteilung der Kontaktmöglichkeiten für jedes Medium präzise darzustellen, wobei eine Dimension die Zeit ist und wo der Planer in der Lage ist, Kontaktwahrscheinlichkeiten zu modifizieren.
2. Die Nutzung von Wirkungskurven, die den Anteil der Bevölkerung ausdrücken, die tatsächlich Kontakt mit der Werbung haben. Dieses für jede Kontaktwahrscheinlichkeitsebene.
3. Multiple Wirkungskurven, die den tatsächlichen Kontakt für jedes Medium in Abhängigkeit der Kontaktchance der Medien darstellen (z.B. 2 x TV und 3 x Print vs. 4 x TV und 1 x Print usw.)
4. Verschiedene Wirkungskurven, die die unterschiedlichen tatsächlichen Kontakte für die individuelle Aufnahme der Werbebotschaft in Abhängigkeit des Media-Mixes, der die Kontaktwahrscheinlichkeit liefert, darstellen.
5. Einen Faktor, mit dem der tatsächliche Kontakt in Abhängigkeit der Media-Synergien, bei Einsatz verschiedener Media-Kombinationen, justiert werden kann.
6. Erinnerung daran, dass beides, der tatsächliche Kontakt und das Behalten der Inhalte, abhängig ist von dem Zeitablauf, und dass es eine Lern- und eine Vergessens-Kurve gibt.

Nachfolgend wird in dem Vortrag ein Beispiel für eine Planung mit der Nutzung der Orvesto Consumer Untersuchung und dem SESAME Media-Mix-Planungsprogramm für den schwedischen Markt dargestellt.

Multi-Media Optimizers: The new couture?

Media-Mix Optimierungsprogramme: Die neue Mode?

Andrew Green, OMD USA

Media-Planer haben Print immer in den Planungsprozess integriert. Die Frage ist, wie sie das getan haben. Haben Sie Print genügend berücksichtigt? Hat Print einen fairen Anteil am Werbebudget (was immer das sein mag) erhalten? In den nächsten paar Jahren wird der US Media-Planungs- und -Einkaufs-Prozess von einer neuen Generation von Optimierungsprogrammen, die durch die Fusion mehrerer Untersuchungen ermöglicht wurden, beeinflusst.

TV-Optimierung

Die Mode der TV-Optimierung begann 1997 als P&G ankündigte ihre 1 Milliarde \$ TV-Budget überprüfen und mehr taktisch planen zu wollen. Das hatte zwar keine Auswirkungen auf die Rolle von Print bei P&G, aber es lenkte in den Unternehmen den Blick auf das im allgemeinen verwendete hauptsächliche Werbemedium. Um die Effektivität für das zur Verfügung stehende Budget zu maximieren, wurde über taktische Planungen der Prozess zur Auswahl der richtigen Tageszeit, des Sender-Mixes und der Programm-Genres für TV-Kampagnen, definiert. Interessanterweise ist in den USA im allgemeinen derjenige der Print plant auch derjenige, der Print einkauft - anders als beim Fernsehen.

Aus Großbritannien wurden zwei Systeme importiert: X*Pert und SuperMidas. Der Media-Optimierungsprozess begann in den USA bereits 1961 und wurde ein paar Jahre später in einem Buch veröffentlicht. An den damaligen Anforderungen hat sich kaum etwas verändert. Der Durchbruch kam als Nielsen TV-Daten anbot, die die durchschnittliche Minute pro Viertelstunde beinhalteten. Aus vielerlei Gründen, die an dem TV-Einkauf in den USA und an technischen Gründen der Daten liegen, stehen TV-Optimierungen nicht mehr im Fokus des Interesses.

Die OMD TV-Optimierungs-Untersuchung: 1999 versus 1996

Im letzten Jahr beleuchtete OMD die Erfahrungen, die die größten TV-Werbung-treibenden mit dem Gebrauch von Optimierungsprogrammen gemacht haben. Es wurden zwei Hypothesen getestet:

1. Werbungtreibende würden wöchentliche Reichweiten heruntersetzen und Wochen zu den jährlichen Kampagnen hinzu addieren.
2. Geld und Kampagnen Ratings würden über eine große Anzahl von Sendern und Tageszeiten gestreut.

Die Untersuchung

Um die Hypothesen zu prüfen wurde eine Analyse von CMRs Werbeausgaben-Datenbank von 1996 bis 1999 durchgeführt. Es wurden 76 große Marken, die im nationalen Fernsehen geworben haben, ausgewählt. Die Marken umfassten 21% des TV-Budgets. Vier Kriterien wurden überprüft:

- Anzahl der aktiven Wochen (von 52)
- Durchschnittliche wöchentliche Haushalts-Ratings
- Anzahl der verschiedenen Sender, die während des Jahres belegt wurden
- % der Haushalts-Ratings, die zur Primetime belegt wurden

Keine Änderung?

Im Vergleich zu 1996 wurde 1999 fast das ganze Jahr hindurch geworben. Optimierungsprogramme können jedoch aufzeigen, wie die wöchentlichen Reichweiten maximiert werden können. Es wurden keine Anzeichen gefunden, dass Werbungtreibende wöchentliche Reichweiten herabgesetzt haben, um eine größere Kontinuität zu erreichen. Unsere Hypothese der größeren Streuung wurde bestätigt. Es wurden 1999 mehr Sender belegt als zuvor. (1996: Ø 18, 1999: Ø 23). Die Belegung der Primetime war unterschiedlich. 40 der 76 Marken erhöhten den Primetime-Anteil und 36 verringerten ihn.

Was bedeutet das für Print?

Es gibt nur wenige objektive Informationen über den Einsatz von TV-Optimierungsprogrammen. Einige nutzen sie nur für Angebot-Pitches, einige gar nicht. Oftmals werden die Programme als zu umständlich und zu langsam empfunden. Es gibt kein System, das nach der Plug and Play-Methode arbeitet. Media-Planung ist mehr als Zahlenverwaltung. Subjektive Kriterien können leicht in eine Optimierungsentscheidung einfließen.

Media-Mix Optimierung

Planer müssen Entscheidungen, wie sie den besten Return on Marketing Investment erhalten, wie sie Ressourcen aufteilen und welche Medien sie auswählen, treffen. Sie müssen zwischen GRPs, Kosten pro Tausend, Reichweite, Qualität usw. entscheiden. Oftmals gibt es Konflikte zwischen den einzelnen Parametern, die der Planer beseitigen muss. Ein Optimierungsprogramm kann nicht für schlechte oder unlogische Media-Entscheidungen verantwortlich gemacht werden. Das ist die alleinige Verantwortung des Planers. Die Besonderheit eines Media-Mix-Optimierungsprogramms ist die Aufteilung der Investitionen zwischen den traditionellen Medien: Print, TV, Radio. Das Ziel ist eine Optimierung nach Reichweiten.

Single-Source-Untersuchungen

Diese Entscheidung muss täglich getroffen werden. MRI z.B. deckt alle drei Medien ab, obwohl es hauptsächlich eine Print-Untersuchung ist. SRIs MultiMedia Mentor ist ein weiteres Beispiel für Single-Source-Forschung, die Planer zur Maximierung der Reichweiten für Zielgruppen führen kann.

Der Hauptkritikpunkt ist die fehlende Tiefe der Media-Nutzungsdaten. SRIs Telefon-Untersuchung fragt die Media-Nutzung im allgemeinen für Tageszeiten ab, jedoch keine Einzel-Titel, Sender, Programme.

Jüngste Arbeiten bei MRI zeigen, dass eine Verbindung zu Niensens Meter-Daten möglich ist. Außerdem hat die Untersuchung den zusätzlichen Vorteil, Produktdaten zur Verfügung stellen zu können.

Daten-Fusion

Daten-Fusion wird in Europa in einigen Ländern seit ein paar Jahren durchgeführt. Dabei werden die Gewohnheiten und Charakteristika von Befragten auf andere Befragte, die ähnlich sind, übertragen. Auf diese Weise werden die Daten von Nielsen, MRI und RADAR fusioniert. Dabei gibt es viele Fallstricke. Die verschiedenen mathematischen Techniken erbringen unterschiedliche Ergebnisse. Es kann zu merkwürdigen Ergebnissen kommen, wenn nach der Fusion z.B. Hundefutterkäufer gar keinen Hund besitzen.

Multi-Basing

Telmar hat einen anderen Ansatz genutzt um Elemente von zwei Untersuchungen zu verbinden. Die Idee ist, dass Demographie und TV-Nutzung genutzt werden können, um Nielsen und MRI zu verbinden. Spezifische Produkt-Nutzer-Gruppen von MRI können mit Nielsen-Reichweiten-Daten dargestellt werden, ohne die Untersuchungen verbinden zu müssen.

Fusion-on-the-fly

OMDs bevorzugte Methode, die in mehreren europäischen Ländern genutzt wird und jetzt in den USA getestet wird. Es werden hochentwickelte mathematische Berechnungen genutzt, die separate Frequenz-Verteilungen verschiedener Media-Untersuchungen und spezieller Zielgruppen verbinden.

Wird Print maximiert werden?

Wie auch bei der TV-Optimierung ist mehr zu tun, als die Reichweiten zu gegebenen Budgets zu maximieren. Es ist auch eine Frage des Media-Vergleichs: Kann eine Vierfarbseite zu einem 30-Sekunden TV- oder Radio-Spot vergleichbar sein? Kann die Kontaktwahrscheinlichkeit eines Mediums direkt mit einem anderen verglichen werden? Was ist mit Programm- und Redaktions-Umfeldern? Media-Mix Optimierungsprogramme müssen mit allen Beteiligten diskutiert, angeschaut, debattiert und kritisiert werden. Verlage müssen aufpassen, dass ihr Anteil an dem Prozess genügend berücksichtigt wird. MRI und Nielsen können darüber streiten, wie Nutzerschaften verglichen werden können. Die Medien müssen aufpassen, dass sie keine Anteilsverluste am Share of Mind hinnehmen müssen.

(3 Tab. im Anhang)

Evaluating TV and Print sales effects

Bewertung von Fernseh- und Print-Verkaufs-Effekten

Bruce Goerlich, Starcom MediaVest Group

Planungs-Agenturen setzen heute auf der Jagd nach Reichweite und Effektivität mehr fragmentierte Media-Pläne ein: TV, Promotions, Print, Internet usw. Außerdem wächst die Akzeptanz von Market-Mix-Modellen, um die Effektivität des eingesetzten Mixes zu evaluieren. Allerdings sind die in den letzten Jahren herausgefundenen Ergebnisse nicht sehr bekannt und auch nicht sehr hilfreich. Es wird zuviel nicht Interessierendes dargestellt. Außerdem wird zuviel Geld für traditionelle Media-Forschung ausgegeben, da neue Methoden nur sehr langwierig akzeptiert werden.

Was muss mit Market-Mix-Modellen entschieden werden?

Wie sollten wir den Verkaufs-Effekt von zwei verschiedenen Medien, die in der gleichen Kampagne genutzt wurden, evaluieren? Dazu ist es notwendig, alte Daten über Media-Kosten und -Leistungen, über Verkäufe der Marken und über andere Faktoren, die den Verkauf beeinflussen (Konkurrenzverhalten, Preise, Verkaufsfördermaßnahmen usw.), zu nutzen. Die meisten Forscher teilen die Meinung, dass die Regressionsanalyse für diese komplexe Situation genutzt werden kann. Aber dann beginnen die Meinungsunterschiede: Sollten wir Marktanteile oder Kilogramm als Variable einsetzen? Welche Maßeinheiten gelten? Ausgegebene Dollar? Zielgruppen-Reichweiten? Reichweite? usw. (Aufzählung weiterer Fragen)

Beispiel:

Für das Beispiel standen für zwei Jahre wöchentliche Daten zur Verfügung. Der Media-Split für diese Marke betrug 82% TV und 18% Print. Für TV kauften wir 1.450 Zielgruppen-Ratings pro Jahr mit einigen Höhepunkten gut über das Jahr verteilt. Der Einsatz von Print nahm im zweiten Jahr ab. Es handelte sich um eine Niedrig-Preis-Marke (78% unter dem Produktdurchschnitt), die zunehmender Konkurrenz ausgesetzt war und deshalb Marktanteile verlor.

(Chart mit den Basisdaten)

Zunächst wurde das Marken-Volumen modelliert. Wir kreierte Adstocks mit verschiedenen Half-Lines (halben Impact-Sätzen). (Adstock ist der kumulierte verspätete Einfluss (Impact) von Werbung auf den Verkauf). Die Händler-Kosten werden indiziert. In der Regressionsanalyse nahmen wir für jedes Medium die Hälfte der Zeit und erhöhten r^2 auf 0.96.

	Halber Impact Satz/Woche	Koeffizient	+/-Wert
Unser relativer Preis		-0.1310	-16,3
Unser TV	4	0.0076	4,7
Unser Print	6	0.0336	6,0
Konkurrenz-Werbung	13	-0.0148	-20,2

Der ungewöhnlich hohe r^2 -Wert sagt nicht, dass wir eine exzellente Erklärung haben. Er warnt uns jedoch, dass der gleichzeitige Anteilsabfall und die Erhöhung der Konkurrenzwerbung unterschiedliches anzeigen. Die verschiedenen halben Impact-Sätze machen Sinn. Print hat einen längeren Lebenszyklus als TV. Wir erhöhten in einem nächsten Schritt r^2 auf 0.97. Wir definierten ein "Repetition"-(Wiederholungs-)Parameter R und wir setzten die Adstocks wieder ein mit 1- ROTS. (OTS= Opportunity to see (Reichweite))
(Zwei Kurven mit Ergebnissen)

Für diese Marke fanden wir heraus, dass für TV-Adstocks $R=0.4$ am besten ist und für Print $R=0.3$. Für solche niedrigen Werte kommt eine Wiederholung nicht zur Geltung. Für TV hat der Werbedruck 60% bei 100 Ratings pro Woche erreicht und 84% bei 200.
(Darstellung eines weiteren Beispiels mit anderen Parametern)

Brand Library™

SMG bietet diese Quelle an, um über ökonomische Analysen der Daten zu erfahren, wie Werbung wirkt. Neben den Daten für Marken beinhaltet die Library Interview-Daten von Kunden und Agenturen, die 29 Statements zur Markensituation ausfüllen mußten. Ich habe 83 Fälle ausgewählt, in denen wir TV-Effekte gemessen haben und 17 mit Print. In unserer Stichprobe scheint Print für größere Marken eingesetzt gewesen zu sein und TV für teurere Marken. Print ist definitiv der Junior-Partner mit 15% der Gesamt-Spendings. Außerdem ist der wöchentliche Adstock bei Print viel niedriger als bei TV.

(Chart: Verteilung der Adstocks)

Die Verteilung der Half-Lifes beträgt bei TV 12 Wochen und bei Print 13 Wochen, d.h. ungefähr das Gleiche. Print-Effekte erscheinen erst verspätet durch den Zeitraum zwischen Veröffentlichung und Lesen, im Durchschnitt mehr als eine Woche. Das zeigt, dass der Effekt einer Print-Anzeige schneller verfällt als der eines TV-Spots. Das ist entgegengesetzt dessen, was wir erwarteten. Wir wissen, dass Print-Kontakte sich langsamer als TV-Kontakte aufbauen. Das Ergebnis, das wir sehen zeigt jedoch, die Marken-Situation wird mehr vom Half-Life bestimmt als die eingesetzten Medien.

(Chart: Verteilung der Half-Lives)

TV hat einen größeren linearen Effekt ($R=0.6$) als Print ($R=0.4$). Das zeigt, dass TV am besten Reichweite aufbauen kann. Im allgemeinen sollten Marken Print besser nutzen, um wiederholte Kontakte zu generieren. Die individuelle Marken-Situation wird mehr vom Wiederholungs-Faktor beeinflusst.

(Chart: Verteilung des Wiederholungs-Faktors)

Was erzählen die Daten über die TV- und Print-Verteilung?
Das Ergebnis ist:

	Elastizität	Verkaufsanstieg für 100 zusätzliche Rating- Equivalents jährlich
TV	0.18	0.85
Print	0.09	3.53

Von der Elastizitätszahl könnte man annehmen, TV ist doppelt so effektiv wie Print. Das wäre falsch. Der Grund dafür ist, dass die Spendings für TV höher sind und eine Elastizitätsmessung lassen sie daher effektiver erscheinen. Ein Prozent mehr Spendings ist beim TV viel mehr als bei Print.

Wenn wir die zusätzlichen Spendings standardisieren, bekommen wir wahre Werte für den Zusatznutzen für jeden Extra-Dollar. Wir erhalten viermal so viel Nutzen von Print wie von TV.

(Chart: Verteilung von 100 TRP Impact Faktor)

Zum Schluss wurde auf einer Skala von 1 bis 10 die Verkaufs-Effektivität jeder Kampagne eingetragen. TV erhielt im Durchschnitt 5.6, Print 6.5.

Schlussfolgerung

Es gibt viele Entscheidungen zu fällen, wenn Marketing-Mix-Modelle eingesetzt werden sollen und Media-Daten zu evaluieren. Die Auswahl, die normalerweise getroffen wird, wurde in dem Beispiel gezeigt.

Folgende Ergebnisse sind wichtig:

Für jeden eingesetzten Dollar war Print im allgemeinen verkaufseffektiver als TV

Print arbeitet am besten als ein "Wiederholungs-(Mehrfach-)Medium"


Kontrolle der verspäteten Kontakte für Print; Print und TV haben in etwa den gleichen zeitlichen Verlauf beim Verkaufseinfluss.

Nachsatz1 Vakant

Worldwide Readership
Research Symposium 10

Deutsche Übersetzung von
ausgewählten Session Papers

IMPRESSUM

Herausgeber © ARBEITSGEMEINSCHAFT MEDIA-ANALYSE E.V. (
MEDIA-MICRO-CENSUS GmbH (MMC)
Auftragsforschung und Verlag
60487 Frankfurt am Main

1. Auflage 2002

Zitate und Teilveröffentlichungen
nur mit genauer Quellenangabe gestattet

Gestaltung Günter Hofmann

Herstellung PETERMANN GZW
Grafisches Zentrum Wetterau

ISBN 3-932247-17-5

Nachsatz1 Vakant

Nachsatz4 Vakant