



Dr. Ralf Mayer de Groot

## Lug und Betrug beim Recall?

*Aktualisierte Fassung, ursprünglich veröffentlicht in planung & Analyse 3 / 2001 und in marketing journal 1/ 2002, S. 50ff*

### Management Zusammenfassung

Der Recall steht im Pflichtenheft nahezu jeder Werbe-Pretest-Ausschreibung. Insofern überrascht es nicht, dass die meisten Werbe-Pretests behaupten, die Aufmerksamkeitswirkung einer Werbung zuverlässig prognostizieren zu können.

Die Erfahrung zeigt jedoch auf der Basis von mehreren hundert Studien, dass die meisten Pre-Tests nicht dazu in der Lage sind. **In Tracking-Studien werden meistens völlig andere Recallwerte ermittelt als von nahezu allen Werbe-Pretests vorhergesagt wurden.** Dies wird u. a. für die weit verbreiteten Werbe-Pretest-Verfahrensgruppen der „Portfolio-Recall-Tests“, der Recognition-Tests sowie Werbe-Pretests mit TV-Programmumfeld nachgewiesen. Nur wenige Institute waren bisher bereit, dies zuzugeben.

Es gilt – leider noch immer – die Aussage von Achenbaum auf der AMA Conference 1985: *„Die traditionellen Markt- und Werbetests entsprechen nicht den Anforderungen. Es macht keinen Sinn, Geld für Forschung auszugeben, die unzuverlässig ist. Markenartikler sollten die Anwendung von unzuverlässigen Marktforschungsmethoden mit geringer Prognosequalität vermeiden – auch wenn ihr Einsatz noch so wünschenswert erscheinen mag.“* – Ergänzend möchte man hinzufügen: *„Wer zu früh standardisiert, den bestraft das Leben“* – denn er nimmt nicht mehr am methodischen Fortschritt teil.“

**Trotz der großen Fortschritte in anderen wissenschaftlichen Bereichen „hat sich in den letzten 40 Jahren nichts in der Marktforschung verändert“ und „es gibt keine veröffentlichten Beweise (größerer Institute), dass Pretests überhaupt zuverlässig arbeiten“**, wie auf der weltweiten Admap Advertising Research Conference im Herbst 2004 festgestellt wurde.

**In den letzten 20 Jahren hat sich ein völlig neuer WerbePretest-Ansatz bewährt, der zuverlässig den Recall auf Basis von „faktoranalytisch geeichten“ Statements vorhersagen kann, wie zahlreiche Validierungen in diesem Artikel beweisen.**

Es ist aufgrund der Beweislast zu hoffen, dass Max Planck mit seiner skeptischen Beobachtung nicht recht behält: *„Eine neue wissenschaftliche Wahrheit pflügt sich nicht in der Weise durchzusetzen, dass ihre Gegner überzeugt werden und sich als belehrt erklären, sondern vielmehr dadurch, dass die Gegner langsam aussterben und die heranwachsende Generation von vorneherein mit der Wahrheit vertraut gemacht wird.“*

## Lug und Betrug beim Recall?

Unstrittig ist, dass der Recall ein wichtige Voraussetzung für Werbeerfolg ist. Auch wenn es inzwischen bewiesen ist, dass der **Recall in keiner direkten Korrelation zur Absatzwirkung steht**. (so bereits Koeppler, S. 69 ff.; Gibson 1983; Percy, Rossiter, 1992, S.11) Unsere Erfahrung zeigt allerdings umgekehrt, dass Werbung mit nachgewiesener Abverkaufs-Wirkung diese umso schneller erzielt, je höher der Recall ist. (Anmerkung: **64% der Werbung für etablierte Marken führender Markenartikler erzielt KEINEN positiven Abverkaufseffekt**. (Untersuchungsnachweise finden sich in planung & analyse 3/2000, Seite 38 ff.)

Es überrascht daher nicht, dass die meisten Werbe-Pretests behaupten, die Aufmerksamkeitswirkung einer Werbung zuverlässig prognostizieren zu können. Die Erfahrung zeigt jedoch, dass die meisten Pretests nicht dazu in der Lage sind. In Tracking-Studien werden meistens völlig andere Recallwerte ermittelt als von nahezu allen Werbe-Pretests vorhergesagt wurden. (Dies können Sie leicht selbst in Ihrem Unternehmen überprüfen). Insofern halten sich die meisten Institute offensichtlich an Mark Twain, der einmal gesagt hat: *„Die Wahrheit sei unser kostbarstes Gut und deshalb müsse man sparsam mit ihr umgehen.“*

Auf diesen Unterschied zwischen Recall-Vorhersage und Werbe-Erinnerungs-Realität sind wir bei den unterschiedlichsten Testverfahren nahezu immer gestoßen, zunächst bei Scholz & Friends und im Mars-Konzern sowie später bei zahlreichen Kunden in den unterschiedlichsten Branchen und Produktkategorien.

Viele Werbe-Pretests werden von großen Tracking-Instituten durchgeführt. Es stimmt zumindest nachdenklich, dass diese zwar oft von Validierungen sprechen, aber bisher nicht den Beweis angetreten haben, daß sie den Recall zuverlässig messen können.

In der Folge möchten wir die weit verbreitetsten WerbePretest-Verfahrensgruppen im Hinblick auf die Recall-Validität überprüfen. Dabei geht es ausdrücklich nicht darum, irgendjemanden schlecht zu machen, sondern Werbe-Fehl-Investitionen zu vermeiden.

### 1. Die sogenannten „Portfolio-Recall-Tests“

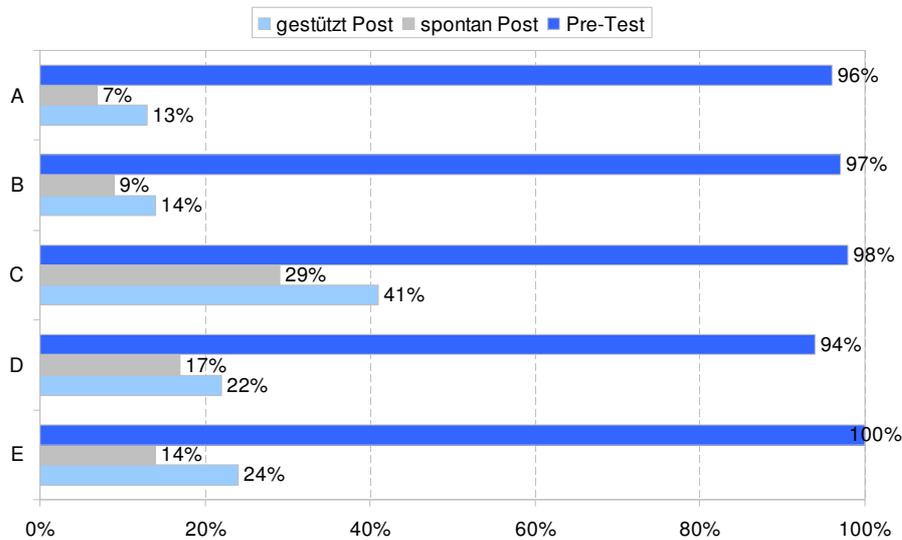
Diese Verfahrensgruppe wird unter den unterschiedlichsten Namen (z.B. „Folder- und Reel-Recall-Tests“) und häufig auch als MultiClient-Studie angeboten. Für sie ist u.a. kennzeichnend, dass

- verdeckt (ohne Bekanntgabe der Produktkategorie) rekrutiert wird.
- In einem Folder eine ungerade Anzahl von Anzeigen gezeigt wird (z.B. 11). Die mittlere Anzeige ist die Test-Anzeige. - Gleiches gilt für einen Test im TV-Block (= „Reel“-Test) oder Funk-Block.
- Kurze Zeit nach der Exposition die spontane und gestützte Werbeerinnerung erhoben wird.

Bei Mars-Tiernahrung lagen damals über 40 Tests dieser Verfahrensgruppe vor. In nahezu allen Fällen wurde ein Werbe-Recall bei unterschiedlichen Katzen- oder Hundenahrungs-Marken von über 85% vorhergesagt. – Auch Methodenexperimente mit bis zu 20 Umfeld-Werbespots konnten das Niveau nicht wesentlich drücken.

Das Problem: Im PreTest wurden völlig andere Recallwerte vorhergesagt als in der Marktrealität gemessen wurden. (Basis: Trackingstudien mit 1000 Befragten oder mehr pro Markt). Dies verdeutlicht Abbildung 1: Nach zwei Jahren Werbung zeigt sich – trotz hoher Budgets und hohem Share of Voice - bei fünf zufällig ausgewählten Tiernahrungsmarken-Kampagnen keine Korrelation oder Gesetzmäßigkeit in Form eines allgemein anwendbaren Multiplikators oder Divisors. Es liegen vielmehr „Welten“ zwischen den vorhergesagten und den tatsächlich realisierten Werbe-Erinnerungswerten.

## Beispiele für unzuverlässige Werbe-Pre-Tests Folder/Reel-Recall-Test - Tiernahrung -



Copyright by Mayer de Groot

48

## Warum scheitert die Werbeerinnerungsmessung beim Portfolio-Recall-Test?

Hilfreich waren bei unserer Ursachenforschung Publikationen von J.C. Maloney aus den Jahren 1961 und 1962 (!). Dieser beobachtete in 6 Anzeigentests für Bier, dass sich unabhängig von deutlichen Differenzen in der Anzeigengestaltung alle Bier-„Recall-Werte“ innerhalb einer Bandbreite von 8% lagen. – Auch bei den neun identischen Umfeldanzeigen ergaben sich nur geringe Unterschiede im Rahmen der statistischen Toleranzen. Die Testergebnisse finden Sie in der folgenden Abbildung.

## Beispiele für unzuverlässige Werbe-Pre-Tests Folder/ Reel-Recall-Test

Marken und Produkt-Recall							
Getestete Anzeigen	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Bandbreite höchster / niedrigster Wert %
	%	%	%	%	%	%	
Getestete Bieranzeigen	75	76	74	78	82	74	8
Kontroll-Anzeigen							
Wachs-Prod.	41	40	41	40	41	44	4
Silberbesteck	42	44	41	36	41	35	9
Auto	77	75	76	75	80	75	5
Tiefkühlkost	47	52	52	54	51	48	7
Käse	34	29	37	36	37	34	8
Zigaretten	48	50	54	51	55	44	11
Fernseher	43	42	49	51	51	43	9
Filmkamera	57	50	49	46	55	54	11
Erfrischungsgetränk	50	50	50	55	60	54	10

Copyright by Mayer de Groot

52

Maloney's Schlussfolgerung war, dass der sogenannte „Folder- oder Reel-Recall“-Test nicht die Aufmerksamkeitswirkung der Werbung misst, sondern das Interesse an der jeweils beworbenen Produktkategorie.

Diese Schlussfolgerung erklärt auch nachvollziehbar das hohe falsch vorhergesagte „Recall“-Niveau z.B. im Tier- und Babynahrungs-Markt. Diese beiden Kategorien weisen das höchste Produkt-Involvement aller bisher getesteten Kategorien auf. Denn in beiden Fällen handelt es sich um Wesen, die nur begrenzt mit der Zielgruppe kommunizieren können, die von den Befragten abhängig sind und bei denen diese nichts falsch machen möchten, weil man sie liebt.

Zur gleichen Schlussfolgerung ist auch Millward Brown (1991, S. 63) gelangt, eines der führenden Werbe-Tracking-Institute (in Großbritannien), nach hunderten von Pretests dieser Verfahrensgruppe: *„Wir waren in der einzigartigen Position, dass unsere Werbe-Pre-Testergebnisse nahezu immer in nachfolgenden Tracking-Studien „validiert“ wurden .... bei hunderten von TV-Spots. Aber trotz ständiger Veränderung der Erhebungsmethode ist es uns **nie gelungen, eine gute Korrelation der Vorhersage mit der Werbeerinnerung in der Realität zu erzielen, wie wir sie in Tracking-Studien messen.**“*

Festzuhalten bleibt: Die sogenannten „Folder-/Reel-Recall-Test“-Verfahren sind erwiesener Maßen nicht in der Lage, den Recall zuverlässig zu vorherzusagen. Sondern sie messen das Interesse an der Produktkategorie. – Sie können aufmerksamkeitsstarke von werbeerinnerungsschwachen Alternativen nicht (zuverlässig) unterscheiden.

Damit können Sie rechnen: Kurzfristig denkende Werbeagenturen werden Ihnen diese Art von Werbe-Pretests immer dann empfehlen, wenn sie wissen, dass Ihre Produkt-Kategorie ein relativ hohes generelles Produkt-Interesse auslöst. Und der Widerstand ein solches „schönes“ Standard-Testverfahren abzuschaffen, ist groß. Denn der Vorteil für die Agenturen ist klar: Jede „auch noch so wenig kreative“ Idee testet gut.

## 2. Recognition-Test

In vielen Untersuchungen ist festgestellt worden, dass die Recognition-Methode unzuverlässig ist. Hierzu seien nur zwei frühe Falsifizierungs-Beispiele aus dem Jahr 1961 genannt:

- W.R. Simmons fand heraus, dass sich die Recognitionwerte von Befragten, die die Testausgabe NICHT gelesen hatten, fast genauso hoch war, wie bei den Lesern.
- E. Marder und M. David erhielten (hohe) „Wiedererkennungswerte“ für Anzeigen, die die Befragten unmöglich gesehen haben konnten.

Koeppler hat sich ausführlich mit dieser Methode auseinandergesetzt und viele weitere Untersuchungs-Nachweise für die Unzuverlässigkeit der Methode (S. 31 ff.) angeführt. Um es kurz zu machen, sei seine Schlussfolgerung zitiert: *„Das Fazit, das sich aus den verschiedenen Untersuchungen der Recognition-Methode ziehen lässt, ist, dass eine erhebliche Unsicherheit darüber besteht, was diese Methode eigentlich misst.- Sowohl in den USA als auch in Europa begegnet man der Recognition-Methode mit ziemlicher Skepsis.“*

Die eindeutigen Erkenntnisse liegen inzwischen seit 40 Jahren (!!!) vor, dass die beiden zuerst diskutierten Methoden erwiesener Maßen nicht in der Lage sind, den wirklichen Recall zu messen. Dennoch werden sowohl „Portfolio-Recall-Tests“ als auch „Recognition-Tests“ nach wie vor häufig angewendet. Der Grund ist Eigeninteresse, denn beide Verfahren ergeben in der Regel „schöne hohe Werte.“

Dies stellte bereits David Ogilvy fest : *„Ich habe mich oft gefragt, ob es eine taktische Verschwörung zwischen Klienten, Medien und Agenturen gibt, die verhindert, dass Werbung sich in „knallharten“ Tests beweisen muss. Jeder Involvierte hat ein großes Interesse daran, dass der Mythos weiterbesteht, dass jede Art von Werbung die Abverkäufe steigert. Dies trifft nicht zu.“*

## Exkurs zu den Copy-Tests von Medien

Dennoch erfreut sich der Recognition-Test insbesondere bei Medienvertretern im Anzeigen- oder Werbezeiten-Verkauf nach wie vor großer Beliebtheit.

Vor kurzem wurden wir wieder einmal aufgefordert, ein Copy-Test-Wettbewerbsangebot für eine Spezialzeitschrift abzugeben. Zu diesem Zwecke erhielten wir den Fragebogen des Vorjahrs mit – sage und schreibe – 457 überwiegend geschlossenen Fragen. In einem Telefoninterview von maximal 20 Minuten sollten Entscheider auf höherer Managementebene befragt werden.

Interessehalber haben wir einige Testinterviews durchgeführt. Selbst bei optimaler Filterauswahl und extrem schnellen Antworten betrug die Befragungszeit mindestens 30 Minuten. Insofern musste mit einer hohen Abbruch- und Fälschungsquote gerechnet werden.

Interessant war auch die Art der Fragen. Bei Vorlage der Testzeitschrift wurden beispielsweise folgende Fragen gestellt:

*Sehen Sie sich bitte Seite 2 mit der Anzeige von ... (Marke) an.*

- *Welche der folgenden Teile der Anzeige sind Ihnen beim Lesen der Zeitschrift aufgefallen bzw. haben Sie beachtet?*
- *Die Schlagzeile „XYZ“?*
- *Das Marken-Logo oben?*
- *Die Produktabbildung unten rechts?*
- *usw.*

Medien-Copy-Test dieser Art erzielen natürlich Werte von über 90%. Aber wem ist damit geholfen? Den Medienvertretern nicht und den Agenturen auch nicht, weil die Werbetreibenden diese Antwortniveaus – zu recht - nicht glauben.

Dabei besteht bei den Werbetreibenden ein großes Interesse an seriösen Test- und Vergleichsdaten sowie gezielten Optimierungshinweisen, wie einige seriöse Verlags-Copytests zeigen. - Und selbst wenn ein Kunde vorübergehend eine suboptimale Kampagne stoppen sollte, um später umso erfolgreicher zu werben, kann dies für die Kundenbindung und das mittelfristig eingesetzte Werbebudget nur hilfreich sein. - Wer aber kurzfristig denkt und seine eigenen Interessen über die legitimen Interessen seiner (potenziellen Neu-) Kunden stellt, kann nicht erwarten, als seriöser Gesprächs-Partner in Betracht zu kommen.

Insofern überrascht es nicht, dass unseriöse Copy-Test von Medien weitgehend ungenutzt bei den Werbetreibenden in den Papierkorb wandern. Schade um das falsch investierte Geld.

### 3. Physische Aufmerksamkeits-Messung

Physische und semi-physische Methoden können angewendet werden, um körperliche Phänomene der Testpersonen zu messen. Diese werden teilweise als Hinweise für Aufmerksamkeit, Wahrnehmung oder Stimulierung gewertet. Bekannte Beispiele sind EDG (Hautwiderstandsmessung), EEG (Gehirnaktivität) und EMG (Muskelanspannung).

Diese physischen Meßmethoden, die nur in Experimental-Situationen im Teststudio durchgeführt werden können, gibt es bereits seit längerem. Sie werden durch ihren direkten Ansatz und die Möglichkeit charakterisiert, Abwehrreaktionen auszuschließen. Meist bleibt jedoch unklar, was eigentlich gemessen wird. Wenn das Herz der Testperson schneller schlägt, sie schwitzt oder eine Pupillenreaktion zeigt: Ist dies ein Hinweis auf eine gute oder eine schlechte Werbung? Daher fällt es schwer die Ergebnisse zuverlässig zu interpretieren.

Insofern könnte eine Verbalisierung der Reaktion durch die Testperson plausibel erscheinen. Es ist jedoch nicht sicher, welche mentalen Vorgänge die Testperson verbalisiert und ob sie diese ausreichend in Worte fassen kann. Darüber hinaus besteht die Gefahr, dass Vorgänge künstlich rationalisiert werden, die bei einer natürlichen Reizverarbeitung anders, zum Beispiel emotional, verarbeitet wurden. Insofern überrascht es nicht, dass es dieser Verfahrensgruppe bisher nicht gelungen ist, eine weite Verbreitung zu erzielen.

#### 4. Werbe-Pretests mit TV-Programmumfeld im Studio

Beispiel für unzuverlässige Werbe-Pre-Tests

**Werbetest im TV-Umfeld im Studio**



In diesem häufig angewendeten Untersuchungsdesign werden die Testpersonen unter dem Vorwand einer Beurteilung von TV-Programmen eingeladen, das Werbeblöcke einschließt. Die prinzipielle Testanlage zeigt die Abbildung. Circa 20 Minuten nach Vorführung der Werbung wird die Werbe-Erinnerung erhoben. Ursächlich für diese Zeitspanne war die Theorie, dass es zwanzig Minuten dauert, um vom Kurzzeit-Gedächtnis ins Langzeit- Gedächtnis abzuspeichern. Die Erwartung war, dass es so gelingen könnte, die langfristige Werbe-Erinnerung zuverlässig vorherzusagen.

Die Praxis hat jedoch gezeigt, dass diese Verfahrensgruppe NICHT in der Lage ist die langfristige Werbe-Erinnerung vorherzusagen. Viele große Markenartikler, die früher diese weitverbreitete Pre-test-Verfahrensgruppe eingesetzt haben, konnten **in keinem einzigen Fall eine vernünftige Korrelation** mit den später in der Realität gemessenen Werten feststellen (in einem **Extremfall nach über 50**, im anderen nach sogar **über 60 Studien**).

Die folgende Tabelle zeigt vier Ergebnisbeispiele, die in einer Neugeschäfts-Präsentation gezeigt wurden. Die Annahme erscheint daher plausibel, dass es sich, um die vermeintlich besten „Validierungen“ dieser Verfahrensgruppe handelt.

Bei objektiver Bewertung zeigen sich jedoch gravierende Unterschiede – bis zum Dreifachen der vorhergesagten Werte - trotz erheblichem Werbedruck. Eine vernünftige Korrelation oder ein durchgängig anwendbarer Multiplikator oder Divisor ist hier wie in den anderen - über 110 Fällen - nicht erkennbar. Insofern ist auch der Verweis auf branchenspezifische Benchmarks – als „argumentative Notlösung“ – nicht zielführend. (Zur generellen Kritik an Benchmarks im Marketing: planung & analyse 5/2000, Seite 16 ff.)

<b>Keine eindeutige Korrelation von Pretest-Recall mit Tracking-Resultaten bei Werbe-Pretests im TV-Programmumfeld im Studio</b>			
Neueinführungen: Schnelldrehende Konsumgüter			
	<b>Pretest-Recall-Prognose</b>	<b>Werbe-Tracking Gestützte Werbeerinnerung</b>	<b>Werbe-Budget pro Monat</b>
<b>Marke A</b>	58 %	20%	0,8 Mio. €
<b>Marke B</b>	55 %	20 %	0,9 Mio. €
<b>Marke C</b>	35 %	18 %	1.4 Mio. €
<b>Marke D</b>	15 %	5 %	1.4 Mio. €

Festzuhalten bleibt, dass diese Verfahrensgruppe leider auch nicht in der Lage ist den Recall zuverlässig vorherzusagen. So hat beispielsweise auch Henkel folgende Erkenntnis veröffentlicht:

*„Bei Advantage wird kurzfristige Erinnerung gemessen. Unsere Ergebnisse zeigen, dass oft **nur eine schwache Korrelation** zwischen der kurzfristigen und der langfristigen Erinnerungsleistung existiert.“* (Dr. Klaus Morwind, Geschäftsführer Henkel KGaA, in GWA: So wirkt Werbung im Marketing-Mix, 1998, S.61)

Diese Feststellung dürfte für alle Werbe-Pretests mit ähnlicher methodischer Untersuchungsanlage gelten.

## **Woran scheitern Werbe-Pretests mit TV-Programmumfeld ?**

Interessanter Weise liefert einer der Anbieter dieser PreTest-Verfahrensgruppe selbst die ursächliche Erklärung, woran diese Tests scheitern. MGM Mediagruppe München bot zusammen mit einem Institut den Test „ad pro“ an. Das Untersuchungsdesign von „ad pro“ stimmt weitgehend mit dem von „ad plus“ überein. Der Unterschied besteht vor allem in der gezielten Variation unterschiedlicher TV-Programmumfelder. Auf der Titelseite der „ad pro-Programmumfeld und Werbewirkung“-Broschüre steht : *„Wir wissen, dass die Effizienz eines Spots erheblich von dem Umfeld abhängt, in dem der Spot gesendet wird. Testen Sie mit ad pro, in welches Umfeld Ihr Spot gehört.“*

Frau Habermeier von MGM hat in einem Ergebnisbericht folgende Erkenntnisse veröffentlicht: *„Die Kreation beeinflusst die Wirkung am stärksten: alle Indikatoren sind davon betroffen...Der zweite untersuchte Faktor, **das Programmumfeld, beeinflusst die Wirkung der Spots ebenfalls in allen gemessenen Dimensionen signifikant...**“*

Zahlreiche eigene und fremde Studien bestätigen die Programmumfeld- Auswirkung auch für die gemessene Werbe-Erinnerung. (Burke (1975, 1976, 1980), Clancy & Kveskin, Lord & Burnkrant, Murphy u.a., Park & McClung, Soldow & Principe). Ein und derselbe Werbespot kann in unterschiedlichen TV-Sendungsumfeldern einen völlig anderes Recall- und Werbewirkungsniveau erzielen. (Deshalb machen wir auch Media-Platzierungs-Research (Mayer de Groot, Fritzen 2008).

An diesen Ergebnis beeinflussenden Umfeld-Effekt scheitert somit eine zuverlässige Werbeerinnerungs-Vorhersage, da unterschiedliche Personen und Zielgruppen unterschiedliche Programme mit unterschiedlicher Intensität oder Involvement nutzen. – Insofern kann es auch kein Standard-Umfeld geben, an dem die Befragten „durchschnittlich“ interessiert sind. (Barwise & Ehrenberg, S.25 ff.)

Mayer und Schuhmann haben darüber hinaus nachgewiesen, dass auch Positions-Effekte im Werbeblock sowie die Attraktivität der anderen Spots die Recall-Ergebnisse beeinflussen. Festzuhalten bleibt, dass die Verfahrensgruppe, die die Werbewirkung im TV-Programmumfeld zu messen versucht, leider auch nicht in der Lage ist den Recall zuverlässig vorherzusagen.

## **5. Zuverlässige Recall-Vorhersage auf Basis von „faktoranalytisch geeichten“ Statements**

Die Entwicklung und Überprüfung des neuartigen Recall-Prognoseverfahrens dauerte über 20 Jahre und erfolgte in ständiger Wechselwirkung mit zahlreichen Werbe-Pre- und -Post-Tests. Es ist gelungen, einen neuartigen Weg zu finden, der den Recall zuverlässig misst und es ermöglicht, in der gleichen Studie Imageveränderungen zu erfassen sowie alle anderen relevanten Werbewirkungseffekte. (Anmerkung: Zuvor war es nur möglich, entweder Recall oder Imageveränderungen in einer Studie zu messen - aber nicht beides)

Die Recall-Messung erfolgt bei dem Advertising \$ales Effect Test auf Basis von „faktoranalytisch geeichten“ Statements. Die Befragten werden offen rekrutiert und es wird ihnen die Werbung ohne

Umfeld gezeigt. Der Testansatz geht davon aus, dass die Befragten selbst – bei einer skalierten Abfrage - in der Lage sind zu beurteilen, wie gut sie eine Werbung in Erinnerung behalten können.

In Grundlagenstudien wurden ursprünglich knapp 600 Statements gesammelt, die zur Werbebeurteilung verwendet wurden. Das Ziel war es, diese Vielzahl auf möglichst wenige Faktoren zu reduzieren. Zu diesem Zwecke wurde gearbeitet mit mehreren, unabhängigen Stichproben, verschiedenen Sätzen von Beurteilungskriterien und unterschiedlichen Beurteilungskriterien.

In mehreren Faktorenanalysen ergaben sich praktisch immer dieselben Grunddimensionen. In diesem Zusammenhang ist es interessant, darauf hinzuweisen, dass sich 10 Jahre später nahezu die gleichen Faktorenlösungen ergaben. Der Umfang und die Stabilität der Ergebnisse machten die Grundlagenstudien daher zu einer vielversprechenden Basis für die Entwicklung einer neuen Gattung von Copytests.

Es ist unsere Überzeugung, dass ein Werbe-Pretest prinzipiell nur so gut ist, wie seine Fähigkeit, die späteren Resultate im Markt vorherzusagen. Entscheidend war daher die Frage, ob es diesem Werbe-Pretestansatz gelingt, die späteren Recall-Resultate im Markt mit hinreichender Zuverlässigkeit vorherzusagen. Die Praxis hat gezeigt, dass die Prognosezuverlässigkeit des Advertising \$ales Effect Test hoch ist - auch beim Recall. Hierfür spricht u.a. die **hohe Übereinstimmung mit allen** bisher vorliegenden Resultaten aus Werbe-Tracking-Studien) **mit den Pre-Testergebnissen** des Advertising \$ales Effect Test (= Validität). Im Falle von Pretest-**Wiederholungen** werden - im Rahmen der statistischen Toleranzen - **die gleichen Ergebnisse erzielt** (= Reliabilität). - Die Erfüllung dieser beiden Zuverlässigkeitskriterien sollte eigentlich selbstverständlich sein. Bedauerlicherweise ist dies, wie gezeigt, bei den meisten „Werbe-Pre-Tests“ nicht der Fall.

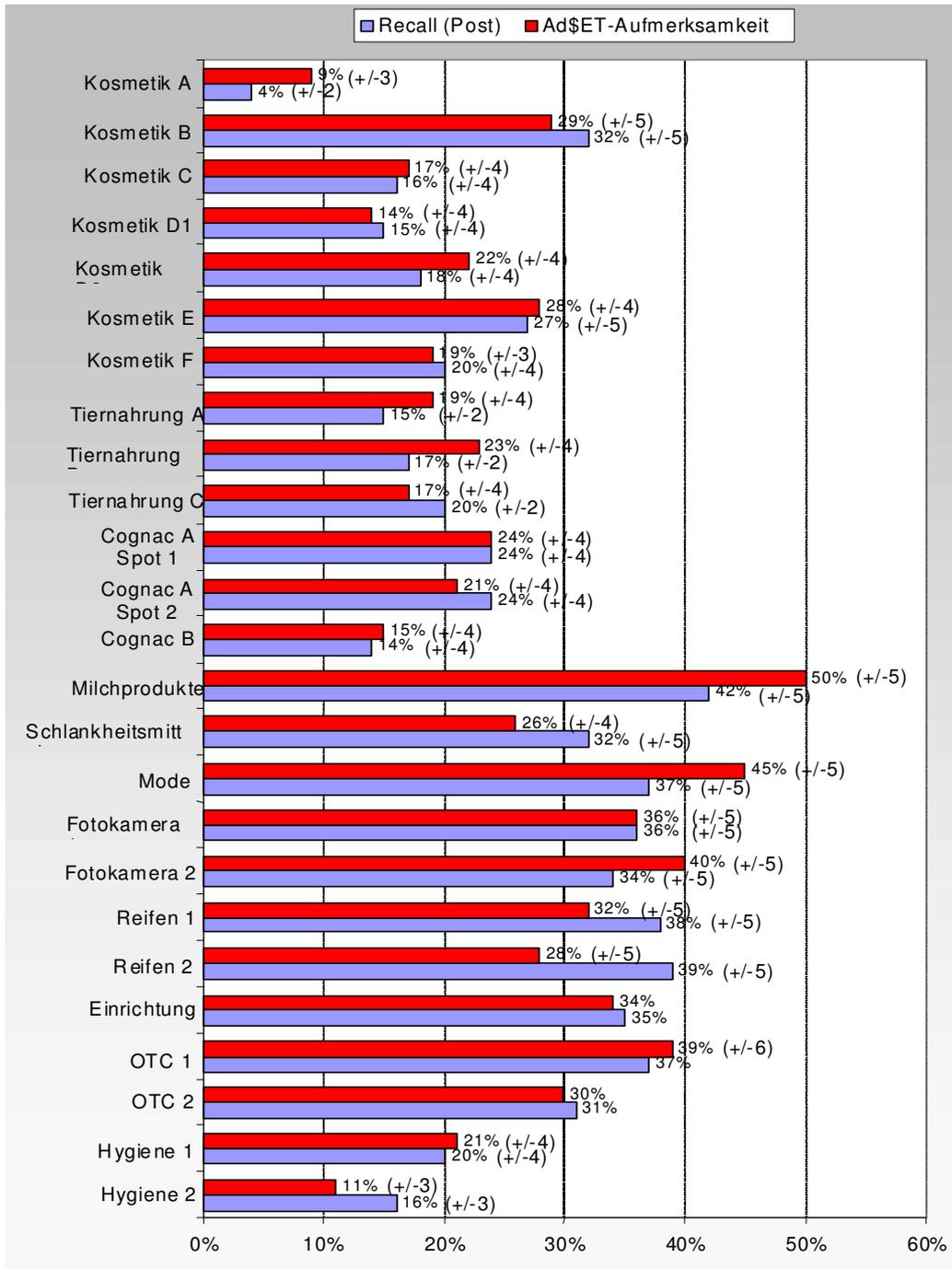
Die Abbildung auf der nächsten Seite zeigt exemplarisch anhand von 25 Beispielen, dass sich der Advertising \$ales Effect Test in der Recall-Vorhersage in den unterschiedlichsten Produkt- und Dienstleistungskategorien bewährt hat. Dies gilt sowohl für die Prognose von kleinen als auch großen Recall-Werten. Auch bei sehr bekannten Marken und Produktkategorien mit hohem Produkt-Involvement (z.B. Tiernahrung) wird der Recall zuverlässig vorhergesagt. Im Rahmen der statistischen Toleranzen, die hinter den jeweiligen Messergebnissen angegeben sind, zeigt sich immer eine Übereinstimmung.

Sicherheitshalber sei angemerkt, dass es in Abhängigkeit von der Höhe des eingesetzten Mediabudgets als auch der verfolgten Schaltstrategie einige Zeit (meist 1 bis 2 Jahre) dauern kann, bis die Ergebnisse validiert werden können.

Durch einen neuartigen und zuverlässigen Weg der Recall-Vorhersage ist der Advertising \$ales Effect Test vermutlich das erste Werbetestverfahren, das in der Lage ist, alle relevanten **Werbewirkungs**-Effekte sowie **Werbemittel**-Beurteilungskriterien in einer Studie zu messen. Zudem ist der Advertising \$ales Effect Test eines der am sorgfältigsten validierten Werbe-Pretestverfahren weltweit.

<b>Zusätzlich</b> durch den Ad \$ales Effect Werbewirkungstest analysiert	Meistens (auch) durch traditionelle Werbetests abgedeckt
<ul style="list-style-type: none"> <li>▪ symbolische Besetzung (Analyse der unterbewussten Werbewirkung)</li> <li>▪ ausgelöste Absatz-Wirkung</li> <li>▪ Trennung von Potenzial und Nicht-Potenzial bei der Interpretation</li> <li>▪ emotionale und faktische Imageshift-Analyse (im Vorher- / Nachher-Vergleich)</li> <li>▪ ggf. Rücktransfer-Analyse</li> <li>▪ ggf. Wear out-Analyse</li> <li>▪ Kombinierbarkeit mit Mediaplatzierungs-Research</li> </ul>	<ul style="list-style-type: none"> <li>▪ Recall / Impact</li> <li>▪ Verständnis der Werbebotschaft</li> <li>▪ Eigenständigkeit der Werbeexekution</li> <li>▪ Irritation / Glaubwürdigkeit</li> <li>▪ Attraktivität</li> <li>▪ Likes / Dislikes</li> <li>▪ Brand / Category Fit</li> <li>▪ Identifikation mit der Zielgruppe</li> </ul>

## Recall : hohe Korrelation von Pre-/Post-Test-Ergebnissen





### **Dr. Ralf Mayer de Groot**

führt seit 2007 die MAYER DE GROOT Marketing-Research and Consult sowie ein Netzwerk an ausgewählten Kooperationspartnern. Er war zuvor 7 Jahre Vorstand /COO bei Konzept & Analyse und ist Autor zahlreicher Fachveröffentlichungen. Strategische Markenführung und Marktforschung kennt er aus allen Perspektiven: aus über 14 Jahren Marketingforschung (Mars, Target Group, Konzept & Analyse), aus 7 Jahren Agenturen (Scholz & Friends, BJP, HMS & Carat), aus 7 Jahren Marketing (Mars Tier-nahrung in Deutschland und Großbritannien) und aus 20 Jahren Marke-tingstrategie-Beratung. [www.mayerdegroot.com](http://www.mayerdegroot.com)

### **Verzeichnis der verwendeten Werbewirkungsstudien / Literatur:**

- Aaker, D. A.; Bruzzone, D.E.: Causes of irritation in advertising, Journal of Marketing, Frühj. 1985, S. 47 ff.
- Aaker, D. A.; Bruzzone, D.E.: Viewer Perception of Prime Time TV-Advertising, in: JAR, Oct.1981, S. 15 ff.
- Armstrong, J.S. (ed): Principles of forecasting, Boston 2001
- Ariely, D.: Predictably irrational, New York 2008
- Barwise, P., Ehrenberg, A.: Television and its audience, London 1990
- Bird, M.; Ehrenberg, A.: Intention to buy and claimed brand usage, Operations Research Quarterly, March 1966, S. 27 ff.
- Bönisch, J.; Mayer de Groot, R. ; Scharf, T.: Nach 20 Jahren gezielt und schnell zur Marktführerschaft: Langnese Cremissimo", in: planung & analyse, 3/2002
- Bönisch, J.; Mayer de Groot, R. ; Scharf, T.: Langneses eiskaltes Marken-Manöver, in: absatz-wirtschaft, Sonderausgabe (Nov.) 2002, S. 10 ff.
- Burke Marketing Research, Inc.: Day after recall television commercial testing, 1975, 1976 & 1980
- Caples, J.: Tested Advertising Methods, Prentice Hall 1975
- Clancy, K.J., Kveskin: TV commercial recall correlates, Journal of Advertising Research, 2/1971, S. 18 ff.
- Dassel, C. von; Wecker, K.M.; Mayer de Groot, R.: "Durch strategische Marktforschung zum Erfolg: NIVEA und NIVEA Soft", in: planung & analyse, 2/2001; dies.: "Harte Fakten führen sanfte Creme zur Spitze", in: Horizont, 19/2001; dies.: "NIVEA SOFT – Sanft zur Haut und zur Dachmarke", in: Absatzwirtschaft, 10/2001
- Farr, A.: Persuasion Shift Testing. In: Admap, Jan 1993.
- Feldwick, P.: A brief guided tour through the copy-testing jungle, a paper from the 1996 Advertising Effectiveness Symposium. An ASI Conference.
- Franzen, G.: Brands & Advertising Effectiveness influences brand equity, Henley-on-Thames 1999
- Franzen, G.: Advertising: How advertising effectiveness, Henley-on-Thames 1994
- Frenzel, K.; Müller, M.; Sottong, H.: Storytelling, München 2006
- Garber, T.: Ein Bier wie flüssiges Gold, in absatzwirtschaft Sonderausgabe zum Marken Award 2004
- Gibson, L.D.: If the answer is copy testing, the answer is ..."Not Recall", Journal of Advertising Research, 1/1983
- Gigerenzer, G.: Bauchentscheidungen, München 2007
- GWA: So wirkt Werbung im Marketing-Mix, 1998
- Habermeier, J.: Umfeldqualitäten systematisch messen – Programmumfeld und Werbewirkung, Media Spectrum 5/1997, s. 8 ff.

- Haimerl, E., Mayer de Groot, R.: Benchmarking in der Marktforschung. Von den Falschen das Falsche lernen, um das Falsche zu tun!, *planung & analyse*, 5/2000, S. 16 ff.
- Haley & Baldinger: The ARF copy research validity project, in: *Journal of Advertising Research*, 31/2 1991.
- Hase, M.: Auf Anhieb erfolgreich. Der Büroartikelhersteller Leitz führt die Trendwende beim Umsatz auf seine erste Endverbraucherkampagne zurück, *w&v* 33/2004 p. 35
- Hollis, N.: The link between TV-ad awareness and sales, *Journal of the Market Research Society*, 36 (1) (1995), pp41-55
- Hollis, N.: Television advertising awareness and sales, a paper from the 1994 European Advertising Effectiveness Symposium. An ASI Conference.
- Jones, J. P.: When ads work, *The German Version*, 8 (1995)
- Juster, F.T.: Consumer buying intentions and purchase probability, *Journal of American Statistical Association*, 1966, S. 658 ff.
- Kast, B.: *Wie der Bauch dem Kopf beim Denken hilft*, Frankfurt a.M. 2007
- Koeppler, K.: *Werbewirkungen definiert und gemessen*, Velbert 1974
- Koschnick, W. J. (Hrsg.): *Focus-Jahrbuch 2007*, Schwerpunkt: Neuroökonomie, Neuromarketing und Neuromarktforschung, München 2007
- Kues, J.; Michel, A.; Scharf, T.: Jules Mumm: Fruchtig, frech und erfolgreich, in: *planung & analyse*, 1/2003, pp. 36
- Lord, K.R., Burnkrant, R.E.: Television program elaboration effects on commercial processing, *Advances in Consumer Research*, 1988, S. 213 ff.
- Lodish, L.M.: How TV advertising works: a meta-analysis of 389 real world split-cable TV advertising experiments, *Journal of Marketing Research*, May (1995), S. ff., vgl. auch ders. u.a.: in *Journal of Advertising Research*, 17, May (1995)
- Lodish, L.M., Lubetkin, B.: General truths? Nine key findings from IRI test data. *Admap*, February 1992.
- Lübbe, R.; Heine, T.: Mit der Kraft der Marke. Starke Marke als Lösungsanbieter effizienter Büroarbeit, *boss* Juli 2003, S. 40 ff.
- Lübbe, R.; Kappelt, J.; Mayer de Groot, R.: Präzises Innovations-Marketing, *planung & analyse* 6/2005, S. 43 ff.
- Lübbe, R.; Mayer de Groot, R.; Hoffmann, S.: Alles im Griff. Leitz: Eine Marke startet durch, *Markenartikel* 5/2004, S. 24 ff.
- Lübbe, R.; Mayer de Groot, R.; Hoffmann, S.: Werbung mit voller Absatz-Wirkung, *planung & analyse* 5/2004, pp. 69
- Maloney, J.C.: Portfolio-Tests – are they here to stay?, *Journal of Marketing*, 2/1961, S. 32 ff.
- Maloney, J.C.: More "why" about Portfolio-Tests, *Journal of Marketing*, 3/1962, S. 76
- Marder, E.: *The Laws of Choice: Predicting Customer Behaviour*. New York (1997)
- Mayer, H., Schumann, G.: Positions- und Umwelteffekte bei TV-Spots, *Jahrbuch der Absatz- und Verbraucherforschung*, 1981, S. 291 ff.
- Mayer de Groot, R.: Welche Werbung wirkt, *MARKANT HandelsMagazin*, 10 / 2010, S. 20 f.
- Mayer de Groot, R.: Das Geheimnis erfolgreicher Testimonialwerbung, *planung & analyse*, 5/2007
- Mayer de Groot, R.: Marketing: Radikal umdenken oder abschaffen, *Markenartikel* 1-2 2007, S. 42 ff.
- Mayer de Groot, R.: *Marken-Diversifikation und -Tragfähigkeit*, Eppstein/Nürnberg 2003
- Mayer de Groot, R.: Was ist das Geheimnis der Siegermarken?, *Mediagipfel St.Gallen, Das Kompendium 2003*, St. Gallen 2003, S. 7 ff.
- Mayer de Groot, R.: Werbewirkungstests, in: Pflaum, D. u.a. (Hrsg.): "Lexikon der Werbung", 7. Aufl., München 2002, S. 588 ff.
- Mayer de Groot, R. Werbe-Abverkaufswirkung pretesten! Auf jeden Fall! , in: *planung & analyse*, 3/2000, S. 38 ff.
- Mayer (de Groot), R.: Das Marketing-Mix zielgruppenadäquat optimieren, *planung und analyse* 2/1992 S. 5 ff.

Mayer (de Groot), R. (1992): Welche neue Marke setzt sich durch?, planung und analyse 2/1992 S. 5 ff.

Mayer de Groot, R.; Fritzen, T.: Mehr Umsatz aus Media herausholen, in : Markenartikel 1 / 2 2008

Mayer de Groot, R.; Haimerl, E.; Zerzer, M.: Von Werbe-Flop auf Marktführer in 3 Monaten – dank modernster WerbeWirkungs-Forschung, in: planung & analyse, 1/2004 S. 61 ff.

Mayer de Groot, R.; Pallek, T. ; Haimerl, E. : Lug und Betrug beim Recall?, planung & analyse, 3/2001; dies.: Lug und Betrug beim Recall?, marketingjournal, 1/2002, S. 50 ff.

Mayer de Groot, R.; Salomon, J.: Innovativ, traditionell oder lieber schräg? Wie unser Unterbewusstsein die Werbewirkung beeinflusst, Markenartikel 9/2007

Mayer de Groot, R. u.a.: The Wrigley´s Extra success story: How to defend successfully a 90% segment share, planung & analyse market research, international issue 2005, S. 6 ff.

Mayer de Groot, R.; u.a.: Media auf der Spur: Wie viel zusätzliche Absatz- und Werbe-Wirkung holen Sie wirklich aus Ihrem Tracking raus?, Media & marketing 9/2005, S. 55 ff.

McDonald, C. 'How frequently should you advertise?', Admap, July / August (1996) S. 22 ff.

Millward Brown International Plc: How advertising affects the sales of packaged goods brands, 1991

McNeil, J.: Federal programs to measure consumer purchase expectations, 1946 – 1973, Journal of Consumer Research, 1974, S. 1 ff.

Morrison, D.G.: Purchase intentions and purchase behaviour, Journal of Marketing, 1979, S. 65 ff.

Morwitz, V.G.: Methods for forecasting from intention data, in: Armstrong, J.S. (ed): Principles of forecasting, Boston 2001, S. 33 ff.

Murphy, J.H., Cunningham, I.C.M., Wilcox, G.B.: The impact of program environment on recall of numerous television commercials, Journal of Advertising, 8/1979, S. 17 ff.

Park, C.W., McClung, G.W.:The effects of TV-program involvement on involvement with commercials, Advances in Consumer Research, 1986, S. 544 ff.

Percy, L., Rossister, J.R.: Measuring advertising effectiveness: copy testing (pre-testing) ads and tracking (post-testing) the campaign, working paper, no. 92-020, Kensington, Univ. of New South Wales, Australia

Reese, W.; Fritzen, T,; Mayer de Groot, R.: Werbebudget runter – Umsatz rauf (Media Efficiency Best Practice), in: absatzwirtschaft 7/2004, S. 92 ff.

Rehorn, J.: Was leisten Pretests?, Düsseldorf 1983

Rosenberg, K.E.; Blair, M.H.: Observations: The Long and Short of Persuasive Advertising. In: Journal of Advertising Research, July / August 1994.

Schmitt, H.J.: Durch Marktforschung zum Erfolg, in: planung & analyse, 5/1999

Shaw, D. J.; Mayer de Groot, R.; Scharf, T.: Goldrichtig. Fallbeispiel: Mit Präzision stößt Beck´s auf Gold, in: Markenartikel 3/2004, S. 38 ff.

Shaw, D. J.; Schipke, A.; Mayer de Groot, R.: Beck´s Gold segelt auf Erfolgskurs, in: planung & analyse 2/2004, S. 20 ff.

Soldow, G., Principe, V.: Response of commercials as a function of program content, Journal of Advertising Research, 2/1981, S. 59 ff.

Theil, H.; Kosobud, R.F.: How informative are consumer buying intention surveys?, Review of Economics and Statistics, 1968, S. 207 ff.

Vakratsas, D.; Ambler, T.: How advertising works: What do we really know?, Journal of Marketing, 1/1999

Weber, T.; Mayer de Groot, R.; Fritzen, T.: Vorwerk-Teppiche fliegen zu neuen Höhen, in marketingjournal 7/2006, S. 30 ff.

Wells, W.D.: EQ, son of EQ and reaction profile, Journal of Marketing, Okt. 1964, S. 45 ff.

Wells, W.D., Leavitt, C.; McConville, M: A Reaction Profile for TV-Commercials, in: Journal of Advertising Research, Dec.1971, S.11 ff

Zyman, S.: The end of advertising as we know it, Hoboken, New Jersey, USA 2002