



MAYER DE GROOT

MARKETING-RESEARCH AND CONSULT

Mayer de Groot · Im Rehwinkel 4 · D 65817 Eppstein · Tel. +49 (0)61 98 - 50 29 21 · Mobil 0178 - 33 684 65 · mdg@mayerdegroot.com



Dr. Ralf Mayer de Groot

Warum Benchmarks im Marketing falsch sind und zu falschen Entscheidungen führen.

Aktualisierte Fassung 2011, ursprünglich veröffentlicht in planung & Analyse 5 / 2000

Management Zusammenfassung

Die einfache Übertragung des bewährten Benchmark-Gedankens aus der Produktion auf das Marketing missachtet die Grundidee des Marketings. Das Marketing-Ziel ist es, einer Marke ein eigenständiges und möglichst gut differenziertes Image zu verschaffen. Zu diesem Zwecke werden oft über den faktischen Produktnutzen hinaus spezifische Zusatznutzen angeboten werden, die sich von den Wettbewerbern unterscheiden.

Kaum jemand scheint jedoch bisher zu bedenken, dass Benchmarks nur dort sinnvoll sind, wo auch Vergleiche Sinn machen. **Wenn es darum geht, eine Alleinstellung zu erzielen, also anders zu werden als die Konkurrenz!** – dann folgt daraus zwangsläufig, dass **die Konkurrenz der falsche (Vergleichs-) Maßstab ist.**

Benchmarks in der bisher genutzten Form führen zu gravierenden methodischen Fehlern. Sie verleiten beispielsweise dazu, eine Kommunikationsidee überwiegend von Nichtzielgruppen beurteilen zu lassen. Zudem werden oft Schlüssel-Kriterien für den Werbeerfolg missachtet und stattdessen „einfach“ zu vergleichenden Kriterien gemessen. Damit besteht aber die Gefahr, dass die Marktforschung dazu beiträgt, die getesteten Kommunikationsmittel in eine falsche Richtung zu „optimieren“, die der Test-Marke erheblich schaden kann.

Es werden daher alternative markenspezifische „Benchmarks“ vorgeschlagen, die mit den Grundsätzen des Marketings vereinbar sind.

Warum Benchmarks im Marketing falsch sind und zu falschen Entscheidungen führen.

Benchmarking ist in Unternehmen ein wichtiger Prozess der Selbstverbesserung. Es wurde von XEROX im Jahr 1979 entwickelt, um der in jener Zeit bedrohlichen japanischen Konkurrenz zu begegnen. Der damalige Präsident von XEROX, Charles Christ, schickte ein Mitarbeiterteam nach Japan mit dem Auftrag, Verfahren, Produkte und Material der Konkurrenz genauestens zu untersuchen. Das Ziel war Benchmarks zu entwickeln, an denen das eigene Unternehmen gemessen werden konnte.

Der damit eingeleitete ständige Vergleich wichtiger Kennzahlen des eigenen mit denen anderer Unternehmen und Organisationen wurde ein wichtiger Wegweiser zur Leistungssteigerung. Vergleichsmaßstäbe von vielen Benchmark-Untersuchungen sind Durchschnittswerte z.B. der Top 25 einer Branche in einzelnen Ländern oder länderübergreifend. Es gibt keinen Zweifel, dass Benchmarking

ein wichtiges Instrument in der Produktion ist. Es ist, um mit Johann Wolf zu sprechen „eine zeitlos gültige Methodik zur ständigen Verbesserung der erfolgsrelevanten Leistungsdaten im Unternehmen“.

Vermeintliche Vorteile von Benchmarking in Marketing und Marktforschung

Marketingausgaben - vor allem für die Kommunikation - sind in Unternehmen oft ein hohes Investment. Bei kaum einer anderen Investition waren aber die Erfolgchancen im Vorhinein schlechter abschätzbar. Die zukünftige Wirkung der Kommunikation wurde bis vor circa 20 Jahren – wenn überhaupt - nur „qualitativ“ überprüft. Es waren damals vor allem kleine, qualitative Marktforschungsinstitute, die Werbetests anboten und durchführten.

Gemessen wurde die kurzfristige Werbeerinnerung, das Verständnis der Werbebotschaft, Gefallen- und Missfallens-Elemente sowie das Testmarken-Image. Das Problem war, dass die Ergebnisse dieser Tests nicht objektiv zu beurteilen waren. Es war meist die subjektive Kompetenz und Erfahrung des Institutsmarktforschers entscheidend, ob die Werbekonzepte positiv oder negativ bewertet wurden.

Insofern erschien die Idee vielversprechend, das Benchmarking auch auf die Marktforschung, speziell auf Werbetests zu übertragen. Es verhiess die Chance, harte, nachprüfbare Vergleichsmesskriterien für die Kommunikation zu bekommen. Vor allem amerikanische Marktforschungsinstitute griffen diese Idee auf. Europäische Institute kauften Lizenzen oder entwickelten ähnliche Maßstäbe. Inzwischen werden auch für andere Fragestellungen Benchmarks entwickelt, etwa für Imageanalysen (Brand Equity Tests), Konzept- oder Packungstests. Vor allem aber wurden standardisierte Werbetests und immer spezifischere Benchmarks oder Normwerte entwickelt. Diese Verfahren dominierten bald den Markt. Denn sie brachten allen Seiten - vermeintlich - nur Vorteile.

Vorteile für Marktforschungsunternehmen:

Sie konnten stets den gleichen Fragebogen für alle Marken und Produktkategorien verwenden, unabhängig davon, ob es sich um Tiernahrung, Kosmetikprodukte oder Dienstleistungsangebote handelte. Der Aufwand hält sich also in Grenzen. War ein solcher Test etabliert, so wurden Diskussionen mit den Kunden weitgehend überflüssig, Fragebogenabstimmungen entfielen usw.

Die Interpretation der Daten war einfach, da nur ein Vergleich mit den Durchschnitts- oder Normwerten notwendig war. Selbst ein Anfänger in der Marktforschung hat eine gute Argumentationsplattform, wenn er auf hunderte von Vergleichstests verweisen kann. Die Forschung konnte standardisiert werden und wurde damit auch wirtschaftlicher. Vor allem große Institute profitierten von dieser Entwicklung, da sie viele Studien durchführen konnten, um dadurch über immer detailliertere Benchmarks verfügten.

Vermeintliche Vorteile für die Werbetreibenden

Die werbetreibenden Unternehmen begrüßten diese Entwicklung, weil sie Verfahren angeboten bekamen, die preisgünstig waren, vor allem, wenn es sich um Multi-Client-Studien handelte. Die Studien hatten ein klares Vergleichs-Ergebnis, das auch weniger werbe-erfahrene Manager oder Assistenten leicht nachvollziehen konnten.

Die Benchmarks machten die Werbetest-Ergebnisse vergleichbar für Länder und andere Produktkategorien. Damit wurden diese Tests zu einem international einsetzbaren Lenkungs- und Kontrollinstrumentarium. Entsprechend haben viele Unternehmen eine Vereinheitlichung und Standardisierung der Testverfahren durchgesetzt. Alle Tochtergesellschaften in den verschiedenen Ländern verwenden das gleiche Verfahren. Auch dies begünstigt große, internationale Institutsketten.

Vorteile für die Werbeagenturen

Den Werbeagenturen war klar, dass jede Werbung mit den gleichen Kriterien gemessen wurde, unabhängig von der Ausgangssituation der Test-Marke. Dennoch begrüßten die meisten Agenturen diese Entwicklung, da diese Beurteilungskriterien nur recht oberflächlich sind. Zudem erwarteten vor allem große Agenturen Globalisierungsvorteile in der Kommunikation.

Benchmarks widersprechen dem Grundgedanken des Marketings

Kaum jemand scheint jedoch zu bedenken, dass Benchmarks nur dort sinnvoll sind, wo auch Vergleiche Sinn machen. Die Kernzielsetzung des Marketings ist es, Kunden zu finden und zu binden. Zu diesem Zwecke soll eine Marke so positioniert und kommuniziert werden, dass sie eine **Alleinstellung** erzielt und die Austauschbarkeit zu anderen Marken reduziert wird. Bereits für Hans Domizlaff (1939) war „*das Ziel der Markentechnik die Sicherung einer Monopolstellung in der Psyche der Verbraucher*“. Ähnlich haben sich Ries und Ries (2000) geäußert: „*Das ultimative Ziel jedes Marken-Programms ist es, eine Kategorie zu dominieren.*“

Wenn es darum geht, eine Alleinstellung zu erzielen, also anders zu werden als die Konkurrenz! – dann folgt daraus zwangsläufig, dass **die Konkurrenz der falsche (Vergleichs-) Maßstab ist**. Lassen Sie uns in der Folge diesen Gedanken im Detail diskutieren.

Typische Benchmarks bei Werbepretests

Bei Werbepretests werden oft Benchmarks für folgende Kriterien verwendet

- ungestützte und gestützte Werbe-Erinnerung / Recall
- Einzigartigkeit der Werbung (nicht der Marke!)
- Neuigkeitswert der Werbebotschaft
- Verständlichkeit
- Gefallen der Werbung / Anteil der Likes zu Dislikes
- Darstellersympathie
- Bewertung der Musik.

Dies sind alles Dimensionen, von denen man annimmt, dass sie unabhängig von der Marke und Produktkategorie sind. Nun ist sicher richtig, dass jede Werbung Aufmerksamkeit erregen muss. Und auch die anderen Dimensionen sind meist erwünscht: Dennoch scheint ein Problem zu bestehen.

Die Erfahrung in der Praxis zeigt: Trotz der Benchmarks floppt Werbung.

Es gibt eine Testmisere. Pretestsieger floppen oft und erfolgreiche Kampagnen haben schlechte Pretestergebnisse. – Die Erfahrung zeigt jedoch auf der Basis von mehreren hundert Studien, dass die meisten Pretests beispielsweise nicht in der Lage sind, die Werbeerinnerung vorherzusagen. **In Tracking-Studien werden meistens völlig andere Recallwerte ermittelt als von nahezu allen Werbe-Pretests vorhergesagt wurden.** Dies haben wir u. a. für die weit verbreiteten Werbe-Pretest-Verfahrensgruppen der „Portfolio-Recall-Tests“, der Recognition-Tests sowie Werbe-Pretests mit TV-Programmumfeld nachgewiesen. (Mayer de Groot u.a. 2001, 2002, 2010)

So stellte auch Dr. Schroiff, Vice-President Market Research Henkel KGaA, fest: „*Wir begegnen noch einem flächendeckenden Validitäts-Vakuum.*“ „*Selbst als Branchenführer bezeichnete Agenturen und Institute zeichnen sich durch eine konzeptionelle und methodische Orientierungslosigkeit beim Thema Werbe-Effizienz aus.*“

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 Ad-map Advertising Research Conference im Herbst 2004 festgestellt wurde.

Wenn aber die meisten Werbe-Pretests in ihren Vorhersagen keine zuverlässige Korrelation mit den in der Realität erzielen, dann helfen auch darauf basierende Benchmarks nicht weiter. 64 % der Werbung für etablierte Marken und 45% für neue Produkte erzielt keinen Absatz-Effekt – trotz aller vorherigen Benchmark-Tests, wie 830 internationale Werbewirkungs-Untersuchungen beweisen. (Studiennachweise bei Mayer de Groot 2000, 2011) – Und die Floprate bei neuen Produkten liegt weiterhin bei 73%, trotz steigender Ausgaben für Markt- und Werbeforschung.

Dennoch vermuten nur die wenigsten die Ursache für diese völlig unbefriedigende Situation in den Werbetests und den darin verwendeten Maßstäben, also den Benchmarks. Genau dies ist aber unseres Erachtens der Fall und wird in der Folge detaillierter diskutiert.

Benchmarks führen zu falschen Schlussfolgerungen für das Marketing

1. Die meisten Benchmarks eignen sich nicht zur Identifikation „guter“ Werbung

Benchmarks in der Marktforschung werden ermittelt, indem die Durchschnittswerte verschiedener Test-Ergebnisse errechnet werden. Dabei werden in der Regel alle gemessenen Dimensionen ohne Gewichtung berücksichtigt. Dies führt oft dazu, dass die Grenze zwischen „guter“ oder „schlechter“ Werbung bei 50% angesetzt wird. Aber dies ist nicht nachvollziehbar. Denn 64 % der Werbung für bestehende Marken erzielt keine positiven Absatzeffekte. Zudem ist in den Benchmarks auch ein hoher Anteil an „schlechter“ getesteten Werbe-Alternativen enthalten. Schon deshalb müsste die Grenze zwischen „gut“ und „schlecht“ bei 64% oder höher liegen. Jedoch nicht bei 50%.

Zudem ist unbekannt, welche Gewichtung die einzelnen Messkriterien haben und in welcher Beziehung sie zueinander stehen. So wäre es z.B. möglich, dass eine Verbesserung der Darstellersympathie die Aufmerksamkeitswirkung verschlechtert, oder das polarisierende Kommunikationselemente bei der Mehrheit der Befragten zu „Dislikes“ führen, diese jedoch für die eigentlich relevante Zielgruppe überzeugend sind.

Das zentrale Werbeziel ist es, dass Kaufverhalten zugunsten der beworbenen Marke zu verändern. Wenn dieses aber nicht oder nicht hinreichend exakt gemessen wird, etwa indem nicht-vergleichbare Messungen vorgenommen werden, entstehen natürlich Probleme. (z.B.: wenn in einem Gewinnkorb ein Aldi-Produkt durch eine teure Premium-Marke mit höheren Qualitätsassoziationen ersetzt wird, ergibt sich immer ein hoher „*Motivationsschub*“, der aber nicht unbedingt durch die Test-Werbung ausgelöst sein muss).

Zudem ist leicht nachvollziehbar, dass für große, etablierte Marken andere Maßstäbe anzulegen sind also für kleine oder gar neue Marken. Die Veränderungsraten sind hier ganz anders zu interpretieren. Die vermeintlich so einfachen Benchmark-Beurteilungskriterien sind also mit Sicherheit nicht „klar und eindeutig“. Diese Fehler sind jedoch, angesichts der folgenden Probleme, nur von nachgelagerter Bedeutung.

2. Beurteilung durch Nicht-Zielgruppen führt zu falschen Schlussfolgerungen

Sollen die Ergebnisse eines Tests mit einem anderen vergleichbar sein, dann setzt dies eine vergleichbare Stichprobenbasis voraus. Will man z.B. eine Benchmark für Joghurts entwickeln und nimmt einmal eine repräsentative Stichprobe, das andere Mal eine nur mit Käufern von Marken-Fruchtjoghurts, dann sind die Ergebnisse nicht vergleichbar. Das Gleiche gilt, wenn man für eine Werbung für einen probiotischen Joghurt nur besonders gesundheitsorientierte Verbraucher interviewt oder bei Sahnejoghurt-Werbung Ablehner von Sahnejoghurt von der Befragung ausschließt.

Da Benchmarks aber Vergleichbarkeit erfordern, werden für standardisierte Werbepretests überwiegend repräsentative oder quasi repräsentative Stichproben verwendet. Einige Institute weichen (auf Kundenwunsch) davon ab und lassen spezifische Stichproben zu. Dann ist jedoch die Verwendung von Benchmarks unzulässig. Es bleibt völlig offen, womit die Ergebnisse eines solchen Tests dann verglichen werden sollen. Was aber passiert, wenn jeweils die gleiche Stichprobe für alle Marken einer Produktkategorie verwendet wird?

Marken richten sich meist nicht an alle Konsumenten, sondern an bestimmte Zielgruppen. Und unterschiedliche Marken konzentrieren sich meist auf unterschiedliche (und unterschiedlich große) Zielgruppen. Nicht weil man nicht alle Konsumenten als Käufer haben wollte, sondern weil man sich auf die vielversprechendste Zielgruppe konzentrieren will und aufgrund des Wettbewerbsdruckes muss. Insofern richtet man sich an die Verbraucher, denen man ein besser auf das Bedürfnisprofil dieser Zielgruppe angepasstes Angebot machen kann und bei denen man eine echte Chance hat, sie für

sich zu gewinnen. Nicht alle Konsumenten wollen beispielsweise ein Cabrio, Spinat oder ein isotonisches Getränk.

Meistens macht die Zielgruppe einer Marke weniger als 50% der Konsumenten einer Warengruppe aus. Die meisten Marktführerprodukte haben inzwischen einen Marktanteil von 30 % oder weniger. Das bedeutet aber, dass in einer repräsentativen Pretest-Stichprobe deutlich mehr als 50% der Befragten Nicht-Zielgruppe für das Produkt/ die Marke sind. Die Beurteilung der Werbung wird dann **überwiegend** von dieser **Nichtzielgruppe** vorgenommen.

Eine Werbung überwiegend von Nichtzielgruppen beurteilen zu lassen, die für den Markenerfolg absolut irrelevant sind, kann zu falschen Schlussfolgerungen führen: Bei einem Werbe- Pretest kann der Befragte nicht „weg-zappen“, wenn er überhaupt kein Produktinteresse hat. Was bleibt ihm anderes übrig, als die Werbung als ein Stück Unterhaltung anzusehen und auf Nachfrage auch noch mehr Unterhaltung zu fordern. Dies kann dann bedeuten, dass eine Werbung wegen fehlenden Unterhaltungswerts für die Nichtzielgruppe „durchfällt“, obgleich sie in der Zielgruppe hoch wirksam ist und umgekehrt.

Wird die Werbung – im Sinne einer Optimierung - an der irrelevanten Befragten-Mehrheit ausgerichtet, dann wird die Kommunikation unter Umständen in eine völlig falsche Richtung gelenkt. Das ist dann Werbung, die sich an den Bedürfnissen der Zielgruppen anderer Marken orientiert und in Grenzfällen sogar die eigentliche Zielgruppe lächerlich macht zur Freude und Unterhaltung der Nichtzielgruppen. Man denke z.B. an die „tierische“ „Nichts ist unmöglich-Kampagne“ von Toyota, die Camel-Werbung oder den AXE-Eunuchen-Spot. Solche unterhaltsame und witzige Werbung erreicht in repräsentativen Umfragen hohe Attraktivitätswerte. Sie ist aufmerksamkeitsstark und sie wird gern gemocht, weil sie die - für die Nichtzielgruppe - lästigen Werbeunterbrechungen nicht nur erträglich, sondern amüsant gestaltet. Dies ist erfreulich, aber doch nicht das Ziel, das ein Werbetreibender anstrebt. In allen drei angesprochenen Beispielen blieb der erhoffte Absatzsteigerungseffekt aus. Noch schlimmer: im Falle von Camel und dem AXE-Eunuchen-Spot führte die Werbung zu deutlichen Absatzrückgängen.

3. Immer feinere Benchmarks führen zu immer falscheren Maßstäben

Ursprünglich wurden nur sehr grobe Benchmarks verwendet. Mit der wachsenden Zahl an Studien war es möglich, immer „feinere“ Benchmarks zu verwenden. Es ist leicht nachvollziehbar, dass Werbung für Lebensmittel, Kosmetik oder Autos nicht gut vergleichbar sind. Im Lauf der Jahre wurde dann z.B. Benchmarks für Kleinwagen entwickelt. Das Problem bleibt natürlich, dass auch Kleinwagen heterogen sind, z.B. gibt es teure und preisgünstige. Einigt man sich auf Benchmarks für eher hochpreisige Kleinwagen, ist das Problem immer noch nicht ausgestanden. Nehmen wir nur an, ein Kleinwagen sei eher als Zweitauto für Frauen konzipiert, ein anderer eher als sportliches Modell für „Yuppies“. Dann stellt sich wieder die Frage, ob das Urteil der „Mutties“ für eine Kommunikation, die sich auf „Yuppies“ richtet, wirklich relevant ist.

Denn in Wahrheit gibt es für eine Marke, die gekauft werden will gar keine andere Chance, als ein wenig anders zu sein. Sie muss andere Vorteile oder „Nutzen“ versprechen. – Warum sollte ein Konsument zu einer anderen Marke wechseln, wenn diese ihm keine zusätzlichen Vorteile verspricht? - Damit versucht die Testmarke zwangsläufig aber auch andere Zielgruppen zu überzeugen als konkurrierende Marken.

Für diese Zielgruppen sind dann aber oft auch andere Dinge in der Kommunikation wichtig, sowohl im Inhalt als auch in der Tonalität. Wenn jetzt aber die Kriterien, die für die Konkurrenten gelten, für die Beurteilung der eigenen Marke herangezogen werden, dann steigt die Gefahr, die eigene Eigenständigkeit aufzugeben und ein „me too“ zu werden. Dann wird man zwar von den Käufern der Konkurrenzmarken gemocht, aber eben nur an zweiter, dritter oder vierter Stelle. Gekauft wird aber hauptsächlich die Marke an der ersten Stelle in der individuellen Markenpräferenz steht. „Fast gekauft“ zu werden, reicht aber nicht, um im Markt zu bestehen. Benchmarks erhöhen somit das Risiko, ein „me too“ zu werden.

Und „der Verbraucher ist nicht „dumm“, sondern Ihre Frau/Mann“ (David Ogilvy 1981). Bei als austauschbar erlebten Produkten wird der Preis zum wichtigsten Entscheidungskriterium im Markt. In dieser Hinsicht ist der Konsument ist konsequent.

4. Das Fehlen zentraler Beurteilungskriterien führt zu falschen Schlussfolgerungen

In Benchmarks können nur Dimensionen eingehen, die für alle Marken und Produktkategorien gleich sind. Sind das aber die zentralen Erfolgskriterien einer Marke?

Die Erfahrung zeigt, der Erfolg von Marken-Positionierungen und deren Kommunikation stellt immer ein Grenzproblem dar: Nur wenn Ihre Marke ein emotionales (und faktisches) Bedürfnis sehr präzise trifft, besteht eine echte Chance auf durchschlagenden Erfolg. Selbst mit großem Budget und Werten, die die Benchmarks deutlich übertreffen, erzielen Sie kaum Wirkung, falls Sie die falsche Botschaft kommunizieren. Um nur zwei veröffentlichte Fallstudien zu nennen, die zudem beweisen, dass vermeintlich kleine (falsche) Details oder Exekutions-Elemente signifikanten Einfluss auf den bewirkten Absatz-Effekt haben können:

	<p>Zunächst erzielte der Dymo LabelWriter, ein innovatives Etikettendruck-Gerät ein Ergebnis im Saarland-Testmarkt, dass mit 38% Zielerreichungsgrad weit unter den Vorgaben lag – trotz intensiver TV-Werbung. Die Analyse ergab, dass die Schlüssel-Vorteile und Kaufmotive für das innovative Produkt nicht hinreichend gut kommuniziert wurden und dass die Werbebotschaft zu komplex war! Auf Basis der ermittelten konkreten Optimierungshinweise wurde das existierende Filmmaterial anders geschnitten und die Erläuterungstexte total überarbeitet. Die Verbesserungen im Detail erzielten folgende Resultate: Das Absatzvolumen wurde um das Fünfzehnfache gesteigert und die Absatzvorhersage von circa 40.000 Geräten während des 1. Werbeflights um 201 Maschinen oder 0,5% übertroffen. (Lübbe, Kappelt, Mayer de Groot 2006)</p>
	<p>Bei Iglo 4Sterne Menüs sagte ein internationales vor einigen Jahren hervorragende Werte voraus weit über den Instituts-Benchmarks sowie einen der höchsten Absatzeffekte, den sie je gemessen hatten (unter den Top 5 weltweit). Obwohl daraufhin ein hohes Mediabudget investiert wurde (€22 Millionen in 18 Monaten) blieb die Absatzentwicklung weit hinter den Erwartungen zurück und die Marke drohte vom Handel ausgelistet zu werden. Der WerbeWirkungs-Test hingegen ermittelten große Schwächen in der Absatzwirkung der Werbung trotz hoher Aufmerksamkeit und Sympathie für die grundsätzliche Werbeidee. Die Ursache für die geringe Wirkung waren ablenkende „Vampir Effekte“ und dass die wichtigsten Kauf-Motivationen nicht präzise genug angesprochen wurden. Kleine, aber wesentliche Details wurden daraufhin gezielt verändert. Das Resultat: Die Marktanteile wurden verdoppelt. In nur 3 Monaten wurde Iglo 4Sterne Menü Marktführer! In diesen 3 Monaten wurde weniger als 10% des vorherigen Budgets investiert. (nur € 2 Millionen laut Nielsen Media Research) – Das Iglo 4Sterne Beispiel beweist wie wichtig eine zuverlässige Messung der zentralen Beurteilungskriterien und ggf. gezielte Optimierung vorab sein kann, der durch die Werbung ausgelösten Wirkungseffekte. (Mayer de Groot u.a. 2004)</p>

In einem „benchmark-orientierten“ Werbetest wird aber das, was Dymo LabelWriter und Iglo 4Sterne Menüs erfolgreich gemacht hat, nicht gemessen. Denn es passt in keine „Benchmark“, es ist nicht vergleichbar. In einer offenen Frage nach der Werbe-Botschaft wird dies auch nur sehr unvollkommen oder gar nicht zum Ausdruck kommen. Hier ist eine Überforderung der Verbalisierungsfähigkeit zu erwarten und es gibt auch eine Scheu, solche Dinge direkt auszusprechen.

Dieses Problem gilt für jede Marke. Der spezifische „Zusatznutzen“ der Marke, also das, was ihren Erfolg ausmacht, wird nicht gemessen. Weil es keiner Benchmark und keiner Standardisierung zugänglich ist. Und es scheint auch nicht zu interessieren, weil immer nur auf die Konkurrenz geschaut wird und geprüft wird, ob „man selbst“ besser ist als die Konkurrenz. **Die Konkurrenz ist jedoch der falsche Maßstab, wenn es darum geht, anders zu werden als die Konkurrenz! – um eine Alleinstellung zu erreichen.**

Fasst man dies zusammen, so lässt sich feststellen, dass die einfache Übertragung des Benchmark-Gedankens auf das Marketing dazu geführt hat, dass in Werbetests zu einem großen Teil die falschen Personen, nämlich Nicht-Zielgruppen befragt werden. Darüber hinaus werden die zentralen Erfolgskriterien nicht erhoben und falsche Maßstäbe angelegt, aus denen dann notgedrungen falsche Schlussfolgerungen gezogen werden. Das klingt vielleicht überspitzt, aber manchmal muss man Dinge deutlich sagen, um auch Wirkung zu erzielen.

Sinnvolle markenspezifischen Benchmarks

1. Werbe-Absatz-Effekt

Das Kernziel der Werbung ist es, die Abverkäufe zu steigern. Das ist die eigentliche Existenzberechtigung der Werbung. Es mag mehrere Zwischenschritte dahin geben, aber Teilziele wie Markenbekanntheit, -sympathie, -vertrauen usw. können keine Endziele sein.

Dieses Hauptziel der Kommunikation, also die Absatz-Wirkung, wird in der Marktforschungswirklichkeit häufig gar nicht oder schlecht gemessen.

Der Advertising \$ales Effect Test ist erwiesenermaßen in der Lage, den Werbe-Abverkaufseffekt zuverlässig vorab zu messen. Zu diesem Zwecke werden bewährte Testmarktsimulationskriterien eingesetzt.

In den Fällen, in denen die Produkte mit (nahezu) unverändertem Marketing-Mix eingeführt wurden, lag die **Prognose** in der Regel **innerhalb von 0,5 Marktanteilsunkten der aktuell erzielten Werte**. In den bereits - gemeinsam mit unseren Kunden - publizierten Fallstudien gab es nur eine Differenz von den vorhergesagten Marktanteilen zu den in der Marktrealität erzielten Werten von:

- 0,1% bei Beck's Gold. In Hektolitern ausgedrückt betrug der Prognoseunterschied nur 1.719 hl bei einem Absatzvolumen von über 310.000 hl im ersten Jahr. (Shaw, Mayer de Groot 2004)
- 0,1% bei Iglo del Mar (Schäffken, Vortrag auf dem 1. Kongress für mehr Markenerfolg 11/08)
- 0,2% bei Iglo Spinat (Schäffken, Vortrag auf dem 1. Kongress für mehr Markenerfolg 11/08)
- 0,2 % bei NIVEA Soft in Italien. (von Dassel, Wecker, Mayer de Groot 2001)
- 0,2 % bei Sheba
- 0,4 % bei NIVEA Soft in Deutschland. Auch die prognostizierte Kannibalisierungsrage von NIVEA Creme wurde auf 0,8 % genau vorhergesagt. (von Dassel u.a. 2001)
- 0,5 % bei Dymo LabelWriter. – Nach einem Testmarktflop (38% Zielerreichungsgrad konnte das Produkt gezielt zum großen Erfolg geführt werden. (Lübbe, Mayer de Groot u.a. 2006)
- Ähnlich gute Übereinstimmungen wurden in zahlreichen weiteren Studien erzielt, von denen einige publiziert wurden: Efasit (2009), Enzym Lefax N (Kleij 1996), Hasseröder (Lennatz 2005), ibutop (Reese u.a. 2004), Lefax (Kleij 1996), Leitz (Lübbe u.a. 2003/2004), Niederegger (Strait u.a. 2006), Perfect Draft (Lennatz 2005/2006), Vorwerk (Weber u.a. 2006).
- Zudem sind mehrere anonyme Beispiele (ohne Markennennung) veröffentlicht worden.

Fasst man Werbung als wirtschaftliche Investition auf, so ergibt sich als erste sinnvolle Benchmark die Forderung, dass sich die Werbekosten (zumindest mittelfristig) auf Grund zusätzlich bewirkter Absatz-Effekte finanziell selbst tragen. – Diese Benchmark wird aber von 64 % der Werbung für etablierte Marken und von 45% für neue Marken bisher nicht erfüllt!

2. Positive Imageveränderungen

Das Image seiner Marken ist meistens das größte Kapital eines Unternehmens. Als zweite markenspezifische Benchmark ergibt sich daher die Forderung, dass das Markenimage in kaufrelevanten Dimensionen verbessert und keinesfalls verschlechtert werden sollte.

Die Image-Auswirkungen der getesteten Werbung werden in „benchmark-orientierten“ Werbepretests fast regelmäßig vernachlässigt. Einfach deshalb, weil diese Kriterien eben nicht für unterschiedliche Märkte und Produktkategorien identisch sind, sondern weil hier produkt- und markenspezifische Dimensionen gemessen werden müssen.

Will man wissen, wie die Werbung das Markenimage (bewusst und unterbewusst) **beeinflusst** – und dies ist eigentlich ein Standard-Anliegen fast aller Werbemitteltests – **so muss man zwei Messungen durchführen:**

- eine Nullmessung vor Konfrontation mit der Werbung zur Feststellung der Ausgangssituation (Ermittlung des bereits bestehenden Images) und
- eine Zweitmessung nach Konfrontation mit der Werbung zur Feststellung der Werbewirkung (Ermittlung des Imageshifts bzw. der Imagebeeinflussung).

Aber die meisten konventionellen Werbepretests nehmen nur eine Messung vor. Dies nennt bereits Rehorn in seinem Grundsatzwerk „Was leisten Pretests?“ ausdrücklich als Beispiel für einen häufigen Verfahrensfehler: *„Es soll die Kommunikationsleistung für ein auf dem Markt befindliches Produkt getestet werden. Es interessiert insbesondere, ob und wie die Werbung das Image und die Abverkäufe des Produktes beeinflussen wird. In den meisten Fällen wird für eine solche Fragestellung **nur eine Messung** durchgeführt, wenn auch gewöhnlich mit einer ganzen Batterie von Fragen und allen möglichen Fragetechniken. **Das Ergebnis ist trotzdem invalide**; denn in ihm drückt sich nicht nur der von dem Werbemittel ausgehende Einfluss auf das Markenimage aus, sondern auch das bereits bestehende, vor dem Test bereits vorhandene, von der präsentierten Werbung ganz unabhängige Markenbild. Das Testergebnis ist also ein Konglomerat von bereits bestehendem Image und Imagebeeinflussung aufgrund der zu testenden Werbung; dabei kann nicht getrennt werden, welcher Faktor das Ergebnis in welchem Ausmaß beeinflusste.“*

Um die Wirkung der Werbung (also die bewirkte Veränderung!) zu ermitteln, muss man das Ausgangsniveau bei jeder Testperson kennen! Also eine Vorhermessung vornehmen. Wie will man sonst feststellen, dass die Werbung Ihr Marken-Image verbessert – und nicht eventuell sogar verschlechtert hat? – Aus diesem Grunde messen wir das Image der beworbenen Marke im experimentellen Design mit einer Vorher- / Nachher-Messung.

Der Status der individuellen Hauptmarke und der beworbenen Marke vor der Werbung werden in der Vorhermessung ermittelt und als „Messlatte“ für die Bewertung der Werbemaßnahme herangezogen. Im zweiten Schritt wird die Werbung gezeigt. Anschließend wird eine Nachhermessung anhand der gleichen Kriterien vorgenommen. Durch intrapersonale Vergleiche wird in der EDV die durch die Werbung ausgelöste emotionale und faktische Veränderung (= Imageshift) analysiert.

Der Image-Vorher/Nachher-Vergleich erfolgt beim Advertising Sales Effect Test für jeden einzelnen Befragten separat, um unterschiedliche Werbereaktionscluster erkennen zu können und nicht nur auf Basis einer Gesamtdurchschnittsbetrachtung aller Antworten.

3. Markenspezifische Benchmarks für Werbemittel-Beurteilungskriterien

Wenn ein Werbemittel erfolgreich ist, ist die Frage nach Likes oder Dislikes, nach Verständnis oder Attraktivität eigentlich eher nachgelagert. Dies ist aber anders, wenn eine Werbung keinen Absatz-Erfolg verspricht. In diesem Fall benötigt man analytische Informationen und dabei kann ein mangelndes Verständnis, unzureichende Likes, fehlende Darstellersympathie oder Glaubwürdigkeit beispielsweise Ursache der unzureichenden Wirkung sein.

Aus diesem Grunde messen wir neben der Werbewirkung auch alle relevanten Beurteilungskriterien des Werbemittels.

<p>Werbewirkung: Zusätzlich durch den Advertising \$ales Effect Test analysiert</p>	<p>Beurteilung des Werbemittels Meistens (auch) durch traditionelle Werbe-pretests abgedeckt</p>
<ul style="list-style-type: none"> ▪ Symbolische Besetzung (Analyse der unterbewussten Werbewirkung) ▪ ausgelöste Absatz-Wirkung ▪ emotionale und faktische Imageshift- Analyse (im Vorher- / Nachher-Vergleich) ▪ ggf. Rücktransfer-Analyse ▪ ggf. Wear out-Analyse ▪ Kombinierbarkeit mit Mediaplatzierungs-Research 	<ul style="list-style-type: none"> ▪ Recall / Impact ▪ Verständnis der Werbebotschaft ▪ Eigenständigkeit der Werbeexekution ▪ Irritation / Glaubwürdigkeit ▪ Attraktivität ▪ Likes / Dislikes ▪ Brand / Category Fit ▪ richtige Zielgruppen-Ansprache

Bei den Werbemittelbeurteilungskriterien kann es sinnvoll sein, starke Abweichungen vom „Üblichen“ zu erkennen und dann analytisch zu nutzen. Aus diesem Grunde verfügen wir auch über Erfahrungswerte. Allerdings geben wir neben dem Durchschnitt auch die besten, die Top 30% und schlechtesten erreichten Werte an, um Werbetreibenden zusätzliche Orientierungshilfen zu bieten.

Die Erfahrung zeigt, dass es unmöglich ist, in allen Werbemittel-Beurteilungskriterien durchgängig „Bestwerte“ zu erzielen. Die bisher jemals gemessenen Höchst-Werte verdeutlichen zudem, dass der theoretische Maximalwert von 100% in keinem Kriterium realisiert wurde und mit an Sicherheit grenzender Wahrscheinlichkeit auch niemals wird. Gleiches gilt für einen theoretischen Minimalwert von Null. – Durch das Ausweisen dieser Werte soll vor allem unrealistischen Zielen vorgebeugt sowie eine Ergebnis-Interpretationshilfe angeboten werden.

Erfahrungsgemäß empfiehlt es sich, dass der werbetreibende Kunde vor Beginn der Kommunikations-Entwicklung sowie dem Pretest zusammen mit der Agentur Werbezielsetzungen vereinbart, die die spezifische Markensituation berücksichtigen. Der Werbe-Pretest sollte dann überprüfen, ob und inwieweit die Zielsetzungen im Einzelnen erreicht werden konnten.

Eine reine Orientierung an Durchschnittswerten im Sinne einer Benchmark widerspricht dem Grundgedanken des Marketings. Jede Marke sollte versuchen, möglichst einzigartig zu sein und aufzutreten. Allerdings zeigt die Erfahrung, dass eine Werbung, die auf allen Werbemittel-Beurteilungskriterien durchschnittlich ist, im Absatz eher eine unterdurchschnittliche Wirkung erzielt.

Wichtig erscheint der Hinweis, dass in einem Advertising \$ales Effect Test erfahrungsgemäß für maximal **4 Werbevorlagen** die Werbemittel-Beurteilungs-Kriterien zuverlässig erhoben werden können. Online ist eine Wirkungs-Prognose für maximal drei unterschiedliche Testalternativen möglich – bedingt durch die kürzere Interviewzeit von 30 Minuten. Insofern können markenspezifische Benchmarks für die Werbemittel-Beurteilungs-Kriterien schnell und kostengünstig ermittelt werden.

Getrennte Analyse von Zielgruppe und Nicht-Zielgruppe

Sehr viel wichtiger als die Erfahrungswerte ist jedoch, dass alle diese analytischen Informationen getrennt für die eigentliche Zielgruppe und die Nichtzielgruppen ausgewertet werden. Nicht der Anteil der Likes oder die Glaubwürdigkeit in der Gesamtstichprobe ist von Interesse, sondern deren Ausprägung in der Zielgruppe. Viele Likes in der Nichtzielgruppe zu erhalten kann sogar häufig als Hinweis auf Fehler in der Kommunikation gewertet werden.

Die Rekrutierung erfolgt beim Advertising \$ales Effect Test meist in erweiterten Zielgruppen-Definition, um Wachstumseffekte erfassen zu können. Da man die eigentliche Zielgruppe, bei der eine spezifische Werbung wirkt, im Vorhinein nicht immer exakt bestimmen kann, wird sie im Werbe-pretest selbst definiert.

Die eigentlich relevante Zielgruppe sind diejenigen Personen, die neu für eine Marke gewonnen werden können sowie jene, die die Marke bereits vor der Werbeexposition bevorzugen. Personen, die kein Interesse an der Test-Marke haben und somit nicht als Potential anzusehen sind, werden getrennt analysiert.

Kritik oder mangelndes Verständnis von Nichtzielgruppen in diesem Sinne ist meist irrelevant: Wir müssen uns zunächst darum bemühen, unsere Zielgruppen für uns zu gewinnen und nicht für Nichtzielgruppen besonders attraktiv zu sein. Allein daraus ist ersichtlich, dass viele positive Werte bezogen auf Benchmarks eine eher nachgelagerte Bedeutung haben.

Resumee

Zielsetzung von Benchmarks war, die Risiken der Investition in „Kommunikations-Massnahmen“ abzuschätzen. Die einfache Übertragung des Benchmark-Gedankens aus der Produktion auf das Marketing war – wie gezeigt wurde, ein äußerst problematischer Irrweg. So berechtigt das Anliegen war, auf dem eingeschlagenen Weg war es nicht zu lösen. - Das Eingeständnis von Irrtümern macht es immer leichter, neue Wege zu beschreiten.

Benchmarks sind auch insofern überholt, da es inzwischen möglich ist, die durch die Werbung ausgelösten Abverkaufs- und Imageeffekte zuverlässig vorherzusagen. Die wichtigste „Benchmark“ für Werbung lautet nach wie vor: *„It is not creative - unless it sells!“*



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 Tiernahrung in Deutschland und Großbritannien) und aus 20 Jahren Marketingstrategie-Beratung. 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
- Bird, M.; Ehrenberg, A.: Intention to buy and claimed brand usage, Operations Research Quaterly, 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; dies: Langneses eiskaltes Marken-Manöver, in: absatzwirtschaft, 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., Kweskin: 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
- 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., dies.: 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