



Dr. Ralf Mayer de Groot

Fiese Benchmark-Falle: Flops über Flops!

Kennzahlen (KPI) gelten im Marketing als objektiv und dem Bauchgefühl überlegen. Aber Ihre uneingeschränkte Verwendung führt zu Misserfolgen. Denn sie widersprechen oft neuen wissenschaftlichen Erkenntnissen (u.a. der Hirnforschung) und sogar dem Grundgedanken des Marketings.

Copyright: Alle Rechte vorbehalten: Dr. Ralf Mayer de Groot 2012

Management Zusammenfassung

Trotz aller Erkenntnisse der Wissenschaften, trotz aller Studien und Pretests wird im Marketing immer noch viel Geld verbrannt. Die Liste der gescheiterten Marken-Existenzen ist lang. Die Auswirkungen der Orientierung an falschen Benchmarks lassen sich in wenigen, erschreckenden Zahlen zusammenfassen: 94 Prozent aller Neu-Einführungen werden keine großen Erfolge, 73 Prozent aller neuen Produkte scheitern als Flops, 84 Prozent aller Verkaufsförderungsaktionen kosten mehr als sie einbringen und 64 Prozent der Werbung für etablierte Marken erzielt keine Umsatzwirkung! – Obwohl die meisten Marketingmaßnahmen vorher getestet wurden und die Benchmarks erfüllten.

Die einfache Übertragung des Benchmark-Gedankens aus der Produktion auf das Marketing erweist sich als ein äußerst problematischer und teurer Irrweg. Die auch als Key Performance Indicator (KPI) bezeichneten Leistungskennzahlen führen oft zu Misserfolgen, da die zugrundeliegenden Annahmen und die Vorgehensweise bei ihrer Anwendung fehlerhaft sind:

1. Kennzahlen haben den Charme, dass sie den Eindruck entstehen lassen, man könne Marken durch eine Handvoll von klar definierten und einfach zu messenden Variablen effizient und effektiv führen. – **Relevante, aber schwieriger zu ermittelnde Variablen, werden häufig vernachlässigt – mit gravierenden Folgen.**
2. Benchmarks unterstellen implizit die Relevanz (leicht abfragbarer) rationaler Beurteilungs-Kriterien der Kunden. Aber nach den Erkenntnissen der Hirnforschung treffen wir Menschen **über 95 Prozent aller Entscheidungen unbewusst** und emotional – **nicht rational.**
3. Benchmarks gehen davon aus, dass positive Bewertungen das Gegenteil von negativen Beurteilungen sind. Aber das Gehirn ist anders aufgebaut. **Positive und negative Gefühle werden von unterschiedlichen Regelkreisen verarbeitet**, dem so genannten *Belohnungssystem* sowie dem *Bestrafungssystem*.
4. Zudem stehen viele einfach zu messende Kennzahlen **in keinem klar nachweisbaren Zusammenhang oder Wirkungs-Beitrag zu den Marketing-Endzielen**, insbesondere der Umsatzwirkung. Die Variablen sind oft nur bedingte Voraussetzungen für Erfolg, aber nicht deren Ursache.
5. Benchmarks werden in der Marktforschung ermittelt, indem die Durchschnittswerte verschiedener Test-Ergebnisse errechnet werden. Dabei werden oft alle gemessenen Dimensionen ohne Gewichtung berücksichtigt. Es wird implizit unterstellt, dass negative Urteile durch eine größere Anzahl positiver Bewertungen überkompensiert werden können. Die Nobelpreisträger Kahnemann und Tversky sowie weitere wissenschaftliche Erkenntnisse haben aber bewiesen, dass Menschen **Verluste viel stärker gewichten als Gewinne**. Dies hat große Auswirkung auf die Markenwahl.
6. Oft werden **Benchmarks für erwiesenermaßen unzuverlässige Kriterien** eingesetzt, die keine Korrelation zu den Werten in der Marktrealität erzielen. Solche Kennzahlen sind irreführend, wie z.B. Benchmarks für einfache Fragen nach der Kaufbereitschaft, unzuverlässige Recallwerte uvm.
7. Benchmarks sind nur dort sinnvoll, wo Vergleiche auch Sinn machen. Im Marketing geht es darum, eine **Alleinstellung** zu erzielen, also anders zu werden als die Konkurrenz. Die Orientierung an **Durchschnittswerten** verniedlicht Probleme, verdeckt die Chancen anders zu werden und führt geradewegs in **die erfolgloseste aller Markenstrategien, die Me-too-Marke**. Durchschnittsverbraucher mit durchschnittlichen Ansprüchen existieren längst nur noch als Fiktion der Statistik.

Benchmarks können bei Teilzielen sinnvoll sein, aber bei den Marketing-Kernzielen sind sie auch insofern überholt, da es inzwischen möglich ist, die durch die Marketingmaßnahmen (Produkt, Packung / Design, Werbung, Promotion usw.) ausgelösten Umsatz- und Imageeffekte zuverlässig vorherzusagen. Zudem können diese vor ihrer Realisierung gegebenenfalls gezielt optimiert werden. Zahlreiche veröffentlichte Fallstudien beweisen: **Flops sind eigentlich überflüssig.**

1. Fast alle Verhaltens-Entscheidungen werden emotional gefällt, aber die meisten Benchmarks sind rational

Emotionen sind der kürzeste Weg ins Gehirn und werden am schnellsten verarbeitet. Wenn jemand die richtigen emotionalen Erlebnisse auslöst, hat er uns als Kunden meist schon gewonnen. Nach den Erkenntnissen der Hirnforschung und Behavioural Economics treffen wir über **95% aller Entscheidungen unterbewusst und emotional – nicht rational**.

In jeder Sekunde nehmen wir über unsere Sinnesorgane circa 11 Millionen Bits an Informationen auf. Bewusst verarbeiten kann unser Gehirn in dieser Zeit aber nur 40 bis maximal 50 Bits. Das entspricht einer kurzen Telefonnummer oder einem dreisilbigen Wort wie „Brasilien“. Eine Anzeige wird in Publikumszeitschriften durchschnittlich 1,7 Sekunden betrachtet und die Kaufentscheidung fällt im Handel in durchschnittlich 3,4 Sekunden, bei unreflektierten Routinekäufen sogar viel schneller. Von jeder Marketingbotschaft verarbeiten wir also nur einen winzigen Bruchteil bewusst.

Wer die richtigen Gefühle auslöst, kann seine Kunden glücklich machen. Aus Bedarf wird Sehnsucht nach Ihrem Marken-Angebot. Nicht Angst und Geiz beflügelt, sondern die Lust, sich Erlebnisse zu gönnen, gefühlte emotionale Defizite zu reduzieren, sich abzusetzen, sich und anderen etwas ganz besonders Gutes zu gönnen. Je genauer Sie die wahren Motive der Kunden treffen, desto größer ist die Absatzwirkung.

- „Der Duft, der Frauen provoziert“, statt „Geruchsvermeidung“, steigerte den Absatz von AXE vor einigen Jahren von 1,3 auf über 65 Millionen.
- Der Anblick einer hässlichen Unterhose neben der Aussage „Viele Männer verhüten, ohne es zu wollen“ steigerte die Absätze von Mey Herren-Unterwäsche.
- Die Frage „Ist es nicht ungerecht, dass Ihr Müll immer Mercedes fährt und Sie nicht?“ führte zu Warteschlangen bei Sixt.
- Mit „Liebe ist, wenn es Landleibe ist“ wuchs der Marktführer um 60 Prozent.
- Leitz besetzte mit „Alles im Griff“ erfolgreich den Kern-Vorteil für Büromaterial und erzielte zweistelliges Wachstum bei geringem Werbe-Budget. (Anzeigenabbildungen auf Seite 9)



Aber Erfolg und Misserfolg liegen gerade im emotionalen Bereich dicht beieinander. Nicht jede Emotion verkauft gleich gut und die Verhaltensrelevanz bestimmter Gefühle unterscheidet sich von einer Kategorie zu anderen deutlich. Erfahrungsgemäß entscheiden oft Kleinigkeiten in der emotionalen Ansprache über Erfolg oder Niederlage, wie zwei weitere Mey-Motive auf der folgenden Seite verdeutlichen.

In einem Test mit rein rationalen Benchmarks wird aber das, was AXE, Leitz, Mey, oder Sixt erfolgreich macht, nicht gemessen. Denn es passt in keine „Benchmark“, es ist nicht vergleichbar. In einer offenen Frage nach der Marken-Botschaft wird dies auch nur sehr unvollkommen oder gar nicht zum Ausdruck kommen. Hier ist oft mit einer Überforderung der Verbalisierungsfähigkeit der Testpersonen zu rechnen und es gibt auch eine Scheu, solche Dinge direkt auszusprechen.

Dieses Problem gilt für jede Marke. Der spezifische emotionale Nutzen einer Marke, also das, was oft ihre Einzigartigkeit und ihren Erfolg ausmacht, wird meist nicht gemessen. Weil es keiner Benchmark und keiner Standardisierung zugänglich ist.



2. Viele Benchmarks stehen in keinem eindeutigen Ursache-Wirkungs-Zusammenhang zu den Marketing-Endzielen

Marketing fassen wir als Investition auf, die sich nachweisbar rechnen soll. Insofern interessiert vor allem, was die jeweilige Marketingmaßnahme für die Marke tut. Nach unserem Verständnis ist Ziel jeder Marketingmaßnahme, die Wahrnehmung einer Marke zu verändern und zwar so zu verändern, dass das Verwendungsverhalten zu Gunsten der eigenen Marke positiv beeinflusst wird. Das Marketing-Kernziel ist es letztlich, die Umsätze und Gewinne zu steigern. Das ist die eigentliche Existenzberechtigung des Marketings. – Es mag mehrere Zwischenschritte dahin geben, aber Teilziele wie beispielsweise Aufmerksamkeit, Attraktivität, Glaubwürdigkeit usw. zum Beispiel bei Kommunikationsmaßnahmen oder Packungen können keine Endziele sein. Wenn eine Marketingmaßnahme keinen nachweisbar positiven Umsatzeffekt erzielt, ist sie meist eine Fehlinvestition.

In vielen Kommunikations- und Packungs-Tests, um beim gewählten Beispiel zu bleiben, ist die Image- und Umsatzwirkung nicht Hauptziel der Tests. Meistens werden diese zentralen Marketingziele „stiefmütterlich“ behandelt. Stattdessen stehen andere Informationen über das Werbemittel oder die Packung an sich im Mittelpunkt der Analyse wie Impact, Glaubwürdigkeit, Verständlichkeit, Attraktivität der Werbe-Exekution usw. und werden mit entsprechenden Benchmarks verglichen. Alles durchaus interessante Informationen, die auch im Advertising sowie Packaging Sales Effect Test zu Optimierungszwecken gemessen werden, aber kaum zentrale Wirkungsdimensionen.

Wird über den Einsatz von Marketingmaßnahmen anhand solcher Ergebnisse entschieden, so muss man sich darüber klar sein, dass es sich bei diesen Faktoren nur um Voraussetzungen für Kommunikationswirkung handelt, um Indikatoren, die im Vorhof des Erfolgs liegen. Über die wirkliche Effektivität sagen diese Voraussetzungen nicht viel aus. So kommt es, dass Kommunikations- oder Packungs-Exekutionen oft trotz hervorragender konventioneller Testergebnisse zum Flop werden.

3. Nachweisbar unzuverlässige Vorhersagewerte sind als Benchmarks wenig sinnvoll

Benchmarks oder Key Performance Indicator für Prognose-Messwerte, die später überhaupt keine Korrelation zu den in der Realität erreichten Werten erzielen, sind sinnlos und irreführend. – Es sei denn man möchte Marktforschungs-Kunden in falscher Sicherheit wiegen.

Gut fundierte Studien haben bereits vor längerer Zeit bewiesen, dass einfache Kaufintentions-Fragen das Kaufverhalten nur unzuverlässig vorhersagen können. So stellt auch Morwitz 2001 zusammenfassend fest: „Umfangreiche Untersuchungen haben etabliert, dass direkte Messungen (der Kaufbereitschaft) zu unzuverlässigen Vorhersagen führen. Intentionen sind keine perfekten Verhaltensmaßstäbe wegen der zufälligen und systematischen Fehler in dieser Messgröße.“

Dies ist seit langem bekannt und in seriösen Marktforschungskreisen unstrittig (z.B. Bird, Ehrenberg 1966; Juster 1966, Theil, Kosobud 1968, McNeil 1974, Morrison 1979, Kalwani, Silk 1982, Mayer de Groot 2000, 2002, 2004).

Institute, die noch mit Kaufbereitschaftsfragen zu Prognosezwecken arbeiten, verwenden daher meist unterschiedliche Gewichtung-Schemata. Morwitz nennt allein 6 alternative, teilweise komplexe Gewichtungsmodelle unterschiedlicher Institute, von denen sich keines als zuverlässiger erwies. – Grundsätzlich gilt, gravierende und systematische Messfehler können durch eine Gewichtung zwar eventuell reduziert, aber nicht beseitigt werden. Einfache Kaufbereitschaftsfragen sind und bleiben somit unzuverlässig. Sie sind für eine zuverlässige Prognose der Absatzwirkung nicht geeignet. – Dennoch setzen viele (Marktforschungs-) Unternehmen nach wie vor Benchmarks für einfache Kaufbereitschaftsfragen ein. Eine Art „russisches Roulett“ im Marketing mit entsprechend hoher Flop-Rate.

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 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 haben wir im Artikel „Lug und Betrug beim Recall?“ u. a. für die weit verbreiteten Werbe-Pretest-Verfahrensgruppen der „Portfolio-Recall-Tests“, der Recognition-Tests sowie Werbe-Pretests mit TV-Programmumfeld nachgewiesen. Übrigens, ohne nennenswerten Widerspruch zu erhalten.

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.

4. Kennzahlen können Fehlanreize verursachen

Es geht nicht darum, auf Kennzahlen im Marketing völlig zu verzichten. – Bei einigen Teilzielen wie Aufmerksamkeit, Attraktivität und Glaubwürdigkeit z.B. von Kommunikationsmaßnahmen können Vergleiche zu Erfahrungswerten durchaus sinnvoll sein. – Entscheidend aber ist, dass sich das Marketingmanagement intensive Gedanken über die Nebenwirkungen von Benchmarks macht. Dies gilt sowohl für den Einsatz einzelner Zahlen und umso mehr für den Einsatz von Kennzahl-Kombinationen.

Wenn man mit Marketing-Managern oder –Mitarbeitern spricht, dann nennt fast jede(r) mehrere Beispiele, bei denen die Erreichung einzelner Kennzahlen zu vollkommen kontraproduktiven Marketing-Entscheidungen geführt hat. Es wird gestöhnt über die absurde Vorgehensweise, aber resignierend als Spielregeln für den „ganz normalen Wahnsinn im Marketingalltag“ erlebt. Eine vergleichbare Märchensituation drängt sich auf: „Des Kaisers neue Kleider“. Jeder erkennt das Problem, aber keiner traut sich wirklich, die Sache beim Namen zu nennen.

In vielen Firmen wird beispielsweise in Produkt(geschmacks)-Tests angestrebt, bei mehr als der Hälfte der Befragten ein positives Resultat zu erreichen, obwohl Marktführer heutzutage in den meisten Kategorien selten Marktanteile von 25 Prozent oder mehr erreichen. So werden oft an sich viel versprechende Produkt-Ideen aufgeweicht. Beispielsweise wird ein schärferes oder saures Geschmackserlebnis, das ursprünglich eine eigentlich hinreichend große Marktminderheit überzeugte, immer weiter abgemildert. Im Ergebnis macht man es dann fast allen recht, aber so gut wie keinem richtig recht. Das Ergebnis: ein weiterer Flop.

Ein weiterer zu beachtender Aspekt, der seltener geschildert wird, ist die Verführung zur Manipulation von Leistungs-Kennzahlen. Der Dekan der London Business School Andrew Likierman bringt es auf den Punkt: *„In dem Moment, in dem man sich entscheidet, über eine Kennzahl zu steuern, lädt man Manager dazu ein, die Kennzahl zu manipulieren.“*

Gute Testergebnisse verkaufen sich leichter als schlechte. Dies führt oft dazu, dass die Grenze zum Beispiel zwischen „guter“ oder „schlechter“ Werbung von Marktforschungs-Instituten bei vielen Kriterien bei 50 Prozent angesetzt wird. Aber dies ist nicht nachvollziehbar. Denn 64 Prozent der Werbung für bestehende Marken erzielt keine positiven Absatzeffekte, wie 830 internationale Werbewirkungsstudien bewiesen haben. 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 Prozent oder höher liegen. Jedoch nicht bei 50 Prozent.

Noch erschreckenender: Große internationale Institute werben teilweise sogar für Ihre Werbe-Pretests damit, dass nur etwa ein Drittel der getesteten Werbung seine Ziele nicht erreicht. Die Chancen, in solchen Pretests gut abzuschneiden, sind somit drastisch besser als in der Realität – mit möglichen teuren Konsequenzen für die Kunden: - Beim Advertising \$ales Effect Test ist dies anders. Nur in etwa einem Drittel aller Tests wird eine ausreichend große Absatzwirkung erzielt.

Schon David Ogilvy kritisierte: *I sometimes wonder if there is a tacit conspiracy amongst clients, media and agencies to avoid putting advertising to such acid tests. Everyone involved has a vested interest in the myth that all advertising increases sales to some degree. It doesn't.*

5. Die implizite Linearitäts Annahme bei Benchmarks ist falsch und führt zu Flops

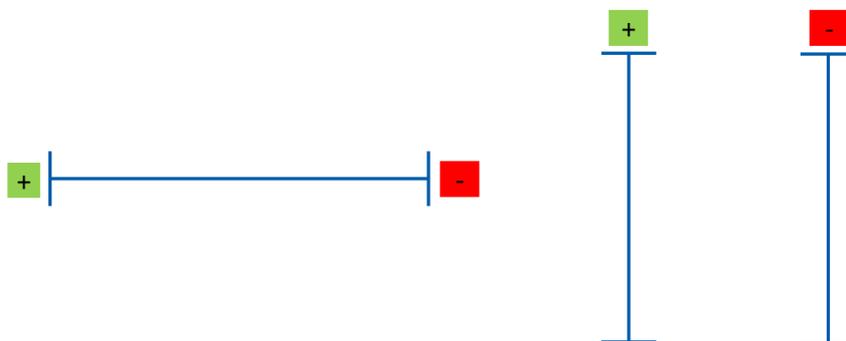
5.1 Positive und negative Gefühle werden im Gehirn von unterschiedlichen Regelkreisen verarbeitet

Konventionelle Tests und Benchmarks gehen implizit davon aus, dass negative Gefühle das Gegenteil von positiven Gefühlen sind. Aber das Gehirn ist anders gebaut. Positive und negative Gefühle werden von unterschiedlichen Regelkreisen verarbeitet.

Für positive Gefühle ist ein Schaltkreis zuständig, der im Allgemeinen unter der Bezeichnung **Belohnungssystem** zusammengefasst wird. (Hieran beteiligt sind u.a. das mediale Vorderhirnbündel und der Nucleus accumbens.) Hier wird die Motivation erzeugt sich einer Sache anzunähern, die sogenannte **Anreizmotivation**.

Die Motivation eine Sache zu vermeiden, weil sie erfahrungsgemäß an negative Gefühle gekoppelt ist, wird durch einen anderen Schaltkreis vermittelt, den man das **Bestrafungssystem** nennt. (Ein wesentliches biologisches Strukturelement dieses Schaltkreises ist die Amygdala.)

Affektbilanz mit zwei Skalen



- In der Alltagssicht stellt man sich negative Gefühle häufig als das Gegenteil von positiven Gefühlen vor.
- Das Gehirn ist aber anders gebaut.
- Positive und negative Gefühle werden von unterschiedlichen Regelkreisen verarbeitet. (zusammenfassend: Storch, Kuhl 2011)

Damasio nennt die Tatsache, dass das Hirn in seiner ursprünglichen Form nur die zwei Kategorien „gut“ oder „schlecht“ kennt, Dualität. LeDoux (2001) konnte nachweisen, dass es ein emotionales Erfahrungsgedächtnis gibt, mit dessen Hilfe sich Organismen extrem schnell und unterhalb der Bewusstseinschwelle gefühlsmäßig orientieren können.

Das Unbewusste greift bei seinen Bewertungen einerseits auf ererbte Muster (Instinkte) und andererseits auf das Extensionsgedächtnis zurück. Hier werden alle wichtigen Situationen abgespeichert, die das Individuum erlebt hat. So ergibt sich für jeden von uns ein persönliches, das ganze Leben umfassende Gedächtnissystem. Wichtige Erfahrungen werden hier gespeichert und zusätzlich mit einer Bewertung versehen, einem sogenannten somatischen Marker. Das Prinzip der Bewertung ist einfach; „Gute Erfahrung, wieder machen“ oder „Schlecht gewesen, bleiben lassen“.

Dieses System gilt als ein evolutionär entstandenes, erfahrungsbasiertes Überlebenssystem, das Menschen mit den Tieren gemeinsam haben. Es ermöglicht einem Organismus, sich an veränderte Umwelten anzupassen und aufgrund gespeicherten Wissens Hypothesen über bestmögliche zukünftige Verhaltensweisen zu bilden. Weil dieses System dem Überleben dient, sind somatische Marker sehr schnell verfügbar. Sie lassen sich innerhalb von 200 Millisekunden nachweisen (Ferguson, Porter; 2009).

Planungsvorgänge ohne körperliche und gefühlsmäßige Begleiterscheinungen bleiben hingegen „intellektuelles Geplänkel im Kopf“, das typischerweise nicht zum Vollzug einer Handlung führt.

5.2 Menschen gewichten Verluste viel stärker als Gewinne.

Benchmarks werden in der Marktforschung ermittelt, indem die Durchschnittswerte verschiedener Test-Ergebnisse errechnet werden. Dabei werden oft alle gemessenen Dimensionen ohne Gewichtung berücksichtigt. Es wird implizit unterstellt, dass negative Urteile durch eine größere Anzahl positiver Bewertungen überkompensiert werden können.

Die Psychologen und Wirtschafts-Nobelpreisträger Kahnemann und Tyversky sind aber durch Versuche zu der Auffassung gelangt, dass klassische Annahmen der Statistik über die Linearität von positiven und negativen Erwartungswerten nur begrenzt der Realität entsprechen. Eine Aussage der Prospekt Theorie beinhaltet die Loss Aversion, nach der Kunden Verluste und Risiken stärker scheuen, als sie Gewinne schätzen. *„Menschen hassen Verluste.... Vereinfacht ausgedrückt, wenn man etwas verliert, wird dies doppelt so stark negativ bewertet, als wenn man das Gleiche erhält.“ Anders ausgedrückt: Menschen sind „Verlust-Vermeider“.* (Thaler, Sunstein 2009).

Hardie, Johnson und Fader haben empirisch belegt, dass negative Abweichungen von derzeit verwendeten Produkten auf starke Ablehnung stoßen. Die Analyse bestätigt eindrucksvoll das Konzept der *Loss Aversion* bei der Markenwahl. - Ergänzend geht die Regret Theory von der Befürchtung aus, fehlerhafte Entscheidungen zu treffen und Wahlmöglichkeiten mit negativen Konsequenzen zu vermeiden, da diese zu Bedauern führen.

Operationalisiert, bedeutet die Loss Aversion, dass neue Angebot vorteilhaft und frei von wahrgenommenen relevanten Nachteilen erlebt werden müssen, um einen nachhaltigen Markenwechsel zu bewirken. - Dies wird in den meisten Benchmarks nicht berücksichtigt und verursacht hohe Flopraten. Die erste Abbildung auf der nächsten Seite veranschaulicht die unterschiedlichen Ergebnisse der Erfolgsvorhersage von konventionellen Benchmarks im Vergleich zur \$ales Effect Test Marktsimulation.

Konventionelle Benchmark im Vergleich mit dem Sales Effect Markenwahl-Kriterium, das der Loss Aversion entspricht.

Testperson	1	2	3	4		Konventionelle Benchmarks
Geschmack	-	+	+	+	→	++ (= Kauf)
gesund	+	-	+	+	→	++ (= Kauf)
attraktiv	+	+	-	+	→	++ (= Kauf)
seinen Preis wert	+	+	+	-	→	++ (= Kauf)
	↓	↓	↓	↓		
Loss Aversion-Kriterium	-	-	-	-		
erfüllt?	Nein = kein Kauf					


Benchmarks unterstellen implizit die Linearität der Messwerte und sind leichter zu erfüllen. Das Sales Effect Markenwahl-Kriterium ist härter und entspricht der Loss Aversion.

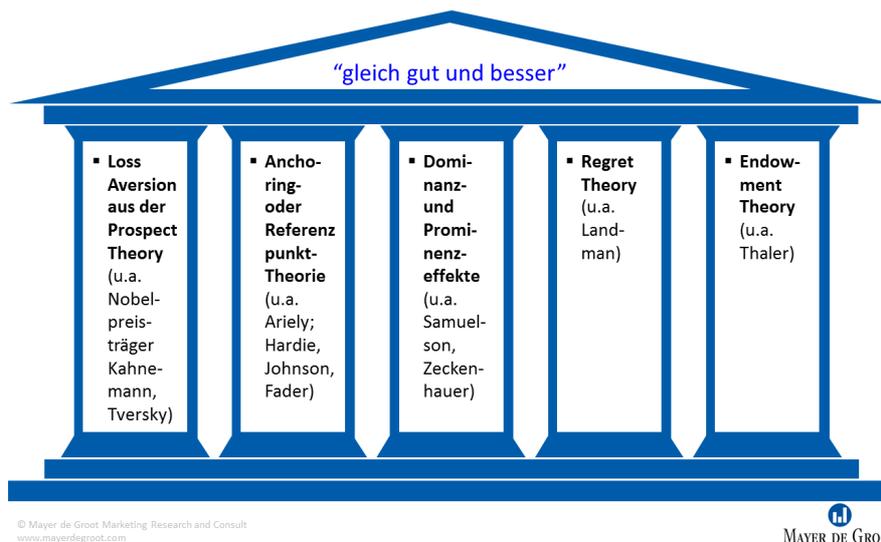
© Mayer de Groot Marketing Research and Consult
www.mayerdegroot.com


MAYER DE GROOT
MARKETING RESEARCH AND CONSULT

Das Markenwahl-Kriterium des Sales Effect Test-Systems

Das Markenwahl-Kriterium des Sales Effect Test-Systems (Konzept-, Produkt-, Packungs-, Kommunikations-Wirkungstests sowie Marktsimulation) entspricht der Loss Aversion, Regret Theorie sowie zahlreichen weiteren wissenschaftlichen Erkenntnissen, (die die Abbildung unten nennt). Es lautet: Kunden wechseln ihre aktuelle Hauptmarke langfristig nur dann, wenn die Alternative zumindest besser in einem kaufrelevanten Kriterium erlebt und zumindest gleich gut in allen anderen kaufrelevanten Dimensionen wahrgenommen wird.

Das Markenwahlkriterium der Sales Effect Marktsimulation hat eine solides theoretisches Fundament:



Dieses verwendete "harte" Marktsimulationskriterium hat sich in knapp eintausend Anwendungen in verschiedensten Märkten und Kategorien im In- und Ausland bewährt. In den Fällen, in denen die neuen Angebote (nahezu) unverändert eingeführt wurden, lag die **Prognose** in der Regel **innerhalb von 0,5 Marktanteilsunkten der aktuell erzielten Werte**. In den bereits - gemeinsam mit Kunden - publizierten Fallstudien gab es nur eine Differenz von den vorhergesagten Marktanteilen zu den in der Marktrealität erzielten Werten von:

Unsere Überzeugung: Eine Methode ist nur so gut, wie ihre nachweisbare Prognosezuverlässigkeit im Markt. - Prognose-Differenzen zur Marktrealität

0,1 %			David John Shaw, Leiter Marktforschung/Media, Inbev Germany, <u>planung & analyse 2/2004</u>	 
0,1 % 0,2 %			Jörg Schäffken, Iglo Marketing Direktor Kongress 2008, Markenartikel 1/2009 Iglo del Mar (neue Produkte): 0,1 % Iglo Rahm-Spinat (neue Werbung): 0,2 %	
0,3 %			Wolfgang Reese, Geschäftsführer Chefaro, <u>absatzwirtschaft 7 / 2004</u>	
0,5 %			Rita Lübbe, Vice President Marketing Central Region Esselte Leitz, u.a. in <u>planung & analyse 6/2005</u>	

© Mayer de Groot Marketing Research and Consult
www.mayerdegroot.com

- 0,1 % bei Beck's Gold. (Shaw, Mayer de Groot 2004)
- 0,1% bei Milka Tender (Mayer de Groot 2011)
- 0,1 % bei Iglo del Mar (Schäffken, 1. Kongress für mehr Markenerfolg 11/08)
- 0,2 % bei Iglo Spinat (Schäffken, 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 (Mayer de Groot 2010)
- 0,2 % bei Bärensnack (Mayer de Groot 2010)
- 0,4 % bei NIVEA Soft in Deutschland. (von Dassel u.a. 2001)
- 0,5 % bei Dymo LabelWriter. (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.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.
- Selbst bei der Vorhersage der Wirkung von Promotions in nur 5 Aktionswochenen lassen sich relativ marktnahe Prognosewerte erzielen. So betrug die Abweichung der verkauften Aktionsware in der WD-40 Volks-Profispray Bildaktion nur 0,9% (Gill, Mayer de Groot 2011)

Eine Methode ist nur so gut, wie ihre nachweisbare Prognosezuverlässigkeit im Markt

NIVEA Soft Marktanteil

Prognose-Differenz:
nur 0,2% in Italien +
Nur 0,4% in Deutsch.





nach 1 Jahr

- „Die Prognose-Übereinstimmungen mit den realen Markt-Ergebnissen von NIVEA Soft gehören **definitiv mit zu den besten Resultaten, die jemals** von Beiersdorf mit unterschiedlichen Markt-Simulationsverfahren erzielt werden konnten.“
- Die Abweichung der Vorhersagewerte zu den in Italien und Deutschland erzielten Werte lagen jeweils unter einem halben Prozentpunkt.“



Klaus M. Wecker, Beiersdorf-Marktforschungsleiter Consumer von 1971 bis 2001 (in planung & analyse 2001)

Die Prognose-Übereinstimmungen gehören definitiv mit zu den besten, die jemals von Beiersdorf mit unterschiedlichen Markt-Simulationsverfahren erzielt werden konnten.

Unzuverlässige Benchmarks Pretests können Produkte gefährden

Drei Fallstudien sollen zum Abschluss exemplarisch an Beispielen von Werbepretests das Risiko unzuverlässiger Benchmarks oder Key Performance Indicator verdeutlichen.

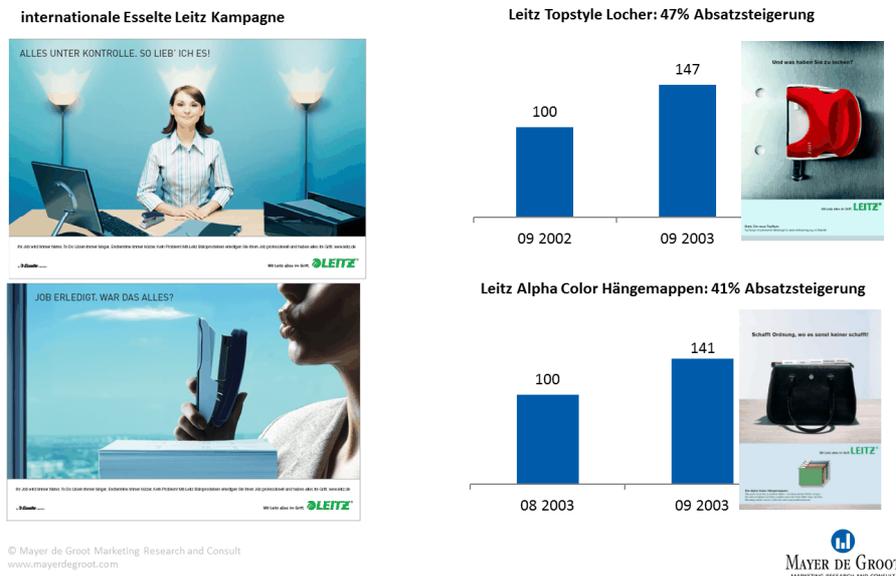
Esselte-Leitz: unzuverlässige internationale Pretests und Benchmarks führten zum Flop Deutschland erzielte hingegen mit einem anderen Werbe-Pretest einen ausgezeichneten Erfolg.

In 2003 wurde eine internationale Esselte-Leitz Werbe-Kampagne entwickelt. Sie zeigte Sekretärinnen, die Dank der hervorragenden Qualität der Esselte-Leitz Produkte bereits ihre Aufgabe erledigt haben. Ein standardisierter konventioneller Pretest sagte international positive Image- und Absatz-Effekte vorher.

In Deutschland kam aber unser "harter" Pretest zu deutlich anderen Ergebnissen und zu einer unbefriedigenden Absatz-Wirkungs-Prognose. Daher wurde in Deutschland eine andere Werbe-Kampagne eingesetzt, die auf den Erkenntnissen aus vorangegangenen qualitativen Analysen (projektiven Techniken) aufbaute. Sie erzielte gute Umsatz-Wachstums-Impulse für neue sowie etablierte Produkte und verbesserte das Leitz-Image signifikant, wie vorhergesagt. Der Leitz Topstyle Locher wurde sogar als "Produkt des Jahres 2004" ausgezeichnet.

Die internationale Esselte-Leitz Werbe-Kampagne erwies sich (wie in Deutschland vorhergesagt) auch weltweit als Flop. Die unbefriedigenden internationalen Umsatzresultate führten zur Veränderung des internationalen Management Teams.

Die internationale Esselte Leitz Kampagne erzielte unbefriedigende Pretest Resultate. Die deutsche Exekution erzielte hohe Image- und Absatz-Effekte.



Baileys Minis: unzuverlässige internationale Pretests und Benchmarks führten zum Flop In Deutschland wurde Baileys Minis hingegen als „Produkt des Jahres 2005“ ausgezeichnet

In 2005 wurde Baileys Minis weltweit eingeführt. Ein sehr kreativer Werbefilm erreichte international in konventionellen Pretests gute Testergebnisse. Er zeigte attraktive Frauen und Männer, die schwelbende Baileys Tropfen in einem schwerelosen Raum verfolgen, um sie zu trinken. (Sehen Sie hierzu Abbildung auf der folgenden Seite)

Trotz hoher ausgelöster Aufmerksamkeit und Attraktivität erzielte dieser Werbefilm im Pretest in Deutschland allerdings keine nennenswerte Absatz-Wirkung. Deshalb wurde eine deutsche Alternative entwickelt. Im "Mini Bar" Werbefilm nimmt ein attraktiver Mann einen Baileys Mini, neben mehreren Hühner-Eiern (zum Größenvergleich) platziert, während einer Party aus dem Kühlschrank, um ihn zu genießen. Dieser Spot kommunizierte die kleiner Packungsgröße von Baileys Mini deutlich besser

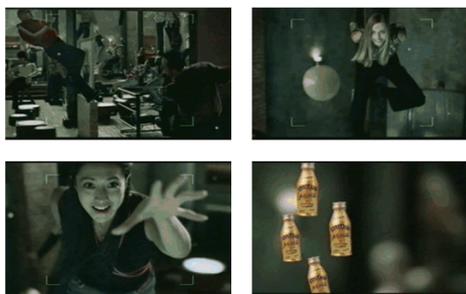
und überzeugte neue, vor allem männliche Zielgruppen. Die im Pretest vorhergesagte hohe Absatzwirkung bestätigte sich in der Marktrealität. Baileys Mini wurde in Deutschland Dank der Werbung ein großer Erfolg und sogar als "Produkt des Jahres 2005" ausgezeichnet.

Trotzdem wurde dieses Produkt später weltweit eingestellt, denn der kreative internationale Werbefilm erzielte (wie in Deutschland vorhergesagt) auch weltweit keine ausreichende Umsatz-Wirkung.

Iglo 4Sterne Menue: in drei Monaten vom Werbe-Flop zum Marktführer

Hohe Aufmerksamkeit erzielten kreative TV-Spots mit dem ersten homosexuellen Werbe-Paar eines großen Herstellers - nicht nur bei Kreativen und der Fachpresse. Ein internationales Institut sagte sogar in einem traditionellen Werbe-Pretest einen der höchsten jemals gemessenen Absatzeffekte voraus. (unter den Top Five weltweit für Lebensmittel!). Unilever investierte deshalb über 20 Millionen € in 18 Monaten in Werbung. Dennoch blieb aber der erwartete Erfolg zunächst aus!

Baileys: Der internationale Spot verkaufte nicht, aber die deutsche Exekution.
 Iglo 4Sterne Menue: Vom Werbe-Flop zum Marktführer in 3 Monaten



International Bailey's TV Spot „Float“



Deutscher Bailey's TV Spot „Mini-Bar“

www.mayerdegroot.com



Innovative qualitative und quantitative Werbe-Analysen kamen aber zu einem anderen Ergebnis. Es mussten relevante Details an der Exekution der attraktiven Basisidee verbessert werden. Das Ergebnis gezielter, kleiner, aber wesentlicher Korrekturen: Innerhalb von nur 3 Monaten verdoppelte Iglo 4Sterne Menue die Marktanteile und wurde Marktführer! Übrigens, mit einem Zehntel des vorher eingesetzten Budgets.

Diese Fallstudie beweist erneut, wie wichtig es ist, die richtigen Emotionen zu kommunizieren und wie gefährlich es sein kann, Werbe-Pretests mit dubiosen, rein rationalen Benchmarks einzusetzen, die auf veralteten Annahmen wie dem "homo oeconomicus" oder "AIDA" basieren.

Es dürfte wenig überraschen, dass Mayer de Groot auch nach dem Verkauf von Iglo von den neuen Eigentümern mit der Überprüfung der Iglo-Relaunch-Kampagne, der Werbung für die neue Seafood Range sowie die McFischstäbchen-Kooperation mit McDonald's ausgewählt wurde. Wieder erwiesen sich die Marktanteils-Prognosen als zuverlässig: Bei Iglo-Seafood wurde die Abweichung auf 0,1 Prozent genau vorhergesagt, beim Kernsortiment Spinat während der Werbepériode auf 0,2 Prozent.

Bei McFischstäbchen sind solche konkreten Prognosezuverlässigkeitsvergleiche leider nicht möglich. Denn das gemeinsame Kooperationsangebot wurde – wie vorhergesagt - schneller verkauft, als geplant. Welche von uns zusätzlich realisierbaren prognostizierten Umsatzchancen möglich gewesen wären, ist und bleibt deshalb eher hypothetisch.



Dr. Ralf Mayer de Groot

steigert als Marketingforscher und -Berater seit Jahren die Markterfolge von neuen und etablierten Angeboten mit innovativen Methoden und Erkenntnissen. Seit 2007 führt er 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 und Dozent an Universitäten. Strategische Markenführung und Marktforschung kennt er aus allen Perspektiven: aus über 14 Jahren Marketingforschung (MARS Gruppe, Target Group, Konzept & Analyse), aus 7 Jahren Agenturen (Scholz & Friends, BJP, HMS & Carat), aus 7 Jahren Marketing (MARS Gruppe im In- und Ausland) und aus 20 Jahren Marketingstrategie-Beratung.

www.mayerdegroot.com oder **06198 50 29 21**

Literaturhinweise erhalten Sie unter mdg@mayerdegroot.com

Ausführlichere und weitere Fallstudien finden Sie unter www.mayerdegroot.com