Bernd Chilian
Ruth Fleuchaus
Beate von Keitz

Werbetests:
Haben sie etwas mit dem Markterfolg zu tun?

Bernd Chilian
Diplom-Psychologe, studierte an der Friedrich Alexander Universität in Erlangen-Nürnberg. Von 1973 bis 1977 war er Studienleiter im Institut "Arbeitsgruppe für psychologische Marktanalysen" in Rückersdorf (heute PSYMA). 1977 trat Chilian bei Bestfoods Heilbronn als Studienleiter qualitative Markforschung ein. Seit 1985 ist er Leiter qualitative Marktforschung, seit 1995 Leiter Gesamt-Marktforschung Bestfoods.

Dr. Ruth Fleuchaus
ist studierte Önologin. Nach der Promotion am Institut für Marketing und Marktforschung war sie als Projektmanagerin bei der CMA im Bereich zentralregionales Marketing tätig. Seit 1998 arbeitet sie als Marktforscherin mit dem Schwerpunkt qualitative Marktforschung bei Bestfoods Deutschland.

Dr. Beate von Keitz
war nach dem Studium der Psychologie und der Betriebswirtschaft als wissenschaftliche Assistentin bei Prof. Kroeber-Riel tätig und promovierte bei ihm mit dem Thema Fernsehwerbung. Seit 1983 ist sie Inhaberin des Instituts für Kommunikations-Forschung Dr. von Keitz in Hamburg und Saarbrücken.

Zusammenfassung
Werbetests werden mit dem Anspruch durchgeführt, dass sie die richtige Werbe-Entscheidung unterstützen und dazu gute Kampagnen identifizieren und schwache Ansätze herausscreenen. Leistet der Werbetest dies tatsächlich? Die Autoren gehen dieser Frage anhand von empirischen Daten zu der laufenden Knorr-Kampagne nach. Lässt sich anhand von Daten belegen, dass die Spots, die im Werbetest gut arbeiten, auch im Markt überlegen sind, also zum Beispiel im Tracking gut arbeiten und vielleicht sogar die Abverkäufe stärker stimulieren als Spots mit schwächeren Testergebnissen? Die Antwort heißt "ja". Der Werbetest wird durch die Ergebnisse aus dem Markt bestätigt. Es gibt eine Reihe von Pretest-Ergebnissen, die das Abschneiden im Tracking vorhersagen. Es gibt sogar signifikante Korrelationen zwischen Pretest-Ergebnissen und Abverkäufen. Diese Ergebnisse stärken das Vertrauen in den Werbetest. Sie werden aber auch genutzt, um die Spots fundiert zu bewerten und um die Werbung noch zielorientierter zu gestalten.

Manuskript ursprünglich veröffentlicht in:
planung & analyse
Zeitschrift für Marktforschung und Marketing
www.planung-analyse.de

Werbetests haben eine zentrale Aufgabe. Sie sollen zeigen: *Arbeitet* die Werbung und kann sie geschaltet werden, oder funktioniert die Kampagne nicht wie gehofft, muss sie optimiert oder möglicherweise komplett verändert werden? Werbetests sollen also in erster Linie Hilfestellung bieten, die richtige Werbe-Entscheidung zu treffen. Gute Kampagnen, die einen positiven Markterfolg versprechen, sollten identifiziert werden und schwache, wenig erfolgversprechende Kampagnen herausgescreent werden.

Mit diesem Anspruch bieten Institute Werbetests an, und die Auftraggeber vertrauen diesem Angebot. Doch leisten Tests dies wirklich - schaffen sie es tatsächlich, gute Werbung von weniger guter zu unterscheiden? Aufbauend auf dem Gedanken der empirischen Forschung stände es Werbetests gut an, ihre Leistungsfähigkeit auch mit empirischen Daten und Fakten zu belegen.

Dass Werbetests ihre Prognoseleistung tatsächlich nachweisen können, wird vielfach nicht für möglich gehalten. Selbst wenn der Werbetest recht hat und die gute Kampagne herausfindet, so gibt es doch im Markt vielfache Einflüsse, die sich neben der Werbung ebenfalls auf die Absatzerfolge auswirken, seien es andere absatzpolitische Instrumente wie Verkaufsförderung und Preispolitik oder die Aktivitäten der Wettbewerber. Manchmal werden die Leistungen des Werbetests auch noch grundlegender angezweifelt. Wer kennt nicht den Vorwurf, dass allenfalls mittelmäßige Werbung beim Test durchkomme und der Test gerade kreative und herausragende Kampagnen zu Fall bringt?

Die Erfahrung der Autoren sieht anders aus. Immer wieder hat sich in der Vergangenheit bei einzelnen Kampagnen gezeigt, dass Werbung, die gut testet, auch im Markt gut arbeitet und zum Beispiel im Tracking überzeugt. Das systematische Testing, mit dem die laufende Knorr-Kampagne begleitet wird, bietet für die Marktforschung die Möglichkeit, sich systematisch mit der Frage "Stimmen Werbetest und Marktrealität überein?" auseinanderzusetzen. Wir haben anhand vorliegender empirischer Daten untersucht, ob sich Zusammenhänge zwischen Werbetest und Markterfolg empirisch nachweisen lassen.

Die Elemente für die Analyse

Knorr arbeitet seit 1999 mit einer neuen Kampagne, in der die einzelnen Submarken wie *Salatkrönung* oder *Kräuterlinge* im Rahmen einer Dachmarkenstrategie präsentiert werden. Die Spots sind jeweils ein bis zwei Monate on air. Das dahinterstehende Werbebudget erscheint zwar auf den ersten Blick groß; der Hauptwettbewerber gibt aber doppelt so viel für seine Werbung aus. Daraus ergibt sich für Knorr die Herausforderung, diesen Budgetnachteil durch qualitativ gute und besonders effizient arbeitende Werbung zu kompensieren.

Die Knorr-Werbung wird von der Marktforschung her einerseits durch Werbetests und andererseits durch ein Werbetracking begleitet. Weiter liegen zu den beworbenen Produkten die laufenden Abverkaufszahlen vor. In die Analyse können also drei Elemente einbezogen werden: der ACTIVITY Werbetest®, das Werbetracking und die Abverkaufszahlen. Zu den Elementen im Einzelnen:

1. Der ACTIVITY Werbetest®

Die Zielpersonen sehen den zu testenden Spot zunächst eingebettet in einen Werbeblock und in ein redaktionelles Umfeld. Die Aufmerksamkeit der Zuschauer wird vorab nicht auf Werbung oder auf die Testmarke gelenkt. Während des Werbeblocks wird die Aufmerksamkeit der Zuschauer auf psychophysiologischem Weg gemessen. Die Dimension der Aufmerksamkeit ist besonders im Hinblick auf die häusliche Wohnzimmersituation mit ihrer starken Medienkonkurrenz und der Möglichkeit, Sekundärbeschäftigungen nachzugehen, wichtig. Mit dem Aktivierungs-Test wird der elektrische Hautwiderstand kontinuierlich erfasst. Damit prüfen wir, ob der Spot die Fähigkeit besitzt, die Aufmerksamkeit der Zuschauer zu wecken und zu halten.

Nachdem die Zielpersonen den Testspot in der Rolle gesehen haben, erfolgt ein relativ ausführliches Interviewprogramm. Hier werden die Werbe- und die Markenerinnerung abgefragt, zusätzlich auch die erinnerte Botschaft, Werbeakzeptanz und spontane Artikelinteressen. Die Ergebnisse aus dieser Phase sind besonders wichtig, weil sie auf der spontanen Werbenutzung basieren, also ohne dass die Aufmerksamkeit durch die Marktforschung auf den Spot gelenkt worden wäre. In einem zweiten Interviewteil sehen die Zielpersonen den zu testenden Spot noch einmal, diesmal mit voller Aufmerksamkeit. In dieser Phase werden die verschiedenen Faktoren der Werbewirkung differenziert beleuchtet, und es werden auch diagnostische Fragen gestellt, z.B. zu den Darstellern, zur Musik und zur Food-Optik.

Für Knorr wurde auf der Basis der Instituts-Standardfragen ein festes Fragenraster entwickelt, das in allen Werbetests eingesetzt wird. Damit werden verschiedene Wirkungsdimensionen der Werbung wie Branding, Verständnis, Akzeptanz, Marken-Images und

Kaufabsichten abgedeckt. Ergänzend werden in jeden Test individuelle spot- und produktspezifische Fragen integriert. Der Test umfasst normalerweise 35 bis 45 Fragen. Jeder Spot wird mit mindestens 100 Zielpersonen getestet. Die Ergebnisse werden in einer Datenbank erfasst. In dieser Datenbank sind die geschlossenen Fragen enthalten, aber auch Ergebnisse zu zentralen offenen Fragen.

2. Das Werbetracking

Für die beworbenen Produkte wird kontinuierlich ermittelt, inwieweit sich die Werbung in den Köpfen der Verbraucher verankert hat. Dazu werden 200 persönliche Einzelinterviews pro Monat über Icon durchgeführt. In der Analyse werden zwei zentrale Erfolgsvariablen betrachtet, nämlich die gestützte Werbeerinnerung in den letzten zehn Tagen sowie der Proven Recall, mit dem die richtig erinnerten Spotinhalte erfasst werden. Als *Erfolgsmaß* hat sich für unsere Analyse das Niveau, das die Kampagne während der Schaltzeit und in den beiden folgenden Wochen im Durchschnitt erreicht, bewährt.

3. Die Abverkaufszahlen

Hier werden die Abverkäufe im deutschen Einzelhandel erfasst. Wir betrachten für die Analyse den sogenannten *durchschnittlichen normierten Normalverkauf*. Das ist der durchschnittliche Absatz pro Verbrauchermarkt aus den Wochen, in denen keine Promotion stattgefunden hat. Damit wird die abverkaufsfördernde Wirkung einer eventuell gleichzeitig stattfindenden Promotion weitgehendst eliminiert.

Als Erfolgsmaß haben wir hier die Veränderung der Abverkäufe verwendet, genauer das Niveau, das die Kampagne während der ersten vier Wochen der Schaltung erreicht, im Vergleich mit den durchschnittlichen Abverkäufen in den vier Wochen vor Werbebeginn.

Wir haben mit unserer Analyse die Zusammenhänge zwischen den Variablen des ACTIVITY Werbetests® auf der einen Seite und den Erfolgsmaßen aus dem Markt, also den Tracking-Ergebnissen und den Abverkäufen, auf der anderen Seite für neun Spots untersucht. Dazu betrachten wir die Korrelationen in Form von Pearson-Korrelationskoeffizienten. Belegen die Daten Zusammenhänge zwischen Werbetest und Marktreaktion?

Die wichtigsten Ergebnisse

In Abbildung 1 auf der nächsten Seite sind die Maße aufgeführt, die einen Zusammenhang zeigen zwischen den Variablen aus dem Werbetest und den Variablen aus dem Tracking (gestützte Werbeerinnerung der letzten zehn Tage und Proven Recall) beziehungsweise den Abverkaufszahlen. In der Matrix sind alle Korrelationen aufgeführt, die die - für die Feldforschung bereits sehr akzeptable - Größenordnung von 0.5 oder mehr erreichen. Die schwarz markierten Zusammenhänge sind statistisch signifikant ($p < 0{,}05$ bei einer einseitigen Hypothesenformulierung). Die anderen Korrelationen erreichen mindestens einen Koeffizienten von 0.5, sind aber nicht im klassischen Sinn statistisch signifikant.

In den Spalten 2 - 4 betrachten wir die Korrelationen, die sich für die Testspots ergeben, unabhängig von ihrem Werbebudget. Ob sich ein Spot im Tracking durchsetzt und Abverkäufe stimuliert, ist natürlich nicht nur eine Frage der Werbequalität, sondern auch eine Frage des Werbebudgets. Hinter den Spots, die wir hier analysieren, stehen unterschiedlich hohe Budgets, so dass die Spots unterschiedlich stark von den Werbeausgaben profitieren. Dieser Effekt des Werbebudgets wurde in einem zweiten Schritt herausgerechnet, indem die Werte für die Marktreaktion mit dem jeweiligen Werbebudget gewichtet wurden. In den mit pro Geldeinheit gekennzeichneten Spalten (Spalten 5 bis 7) werden also die Wirkungen auf Tracking und Abverkauf pro eingesetzter Geldeinheit betrachtet.

Der Werbetest wird durch die Ergebnisse aus dem Markt bestätigt.

Die Daten zeigen signifikante Zusammenhänge zwischen den Variablen aus dem Werbetest und der Marktreaktion. Es gibt eine Reihe von Pretest-Ergebnissen, die das Abschneiden im Tracking vorhersagen können. Und es gibt sogar signifikante Korrelationen zwischen den Pretest-Ergebnissen und den Abverkäufen. Die Befürchtung bzw. der Vorwurf, dass Werbetest und Werbeerfolg nichts miteinander zu tun haben, bestätigt sich somit nicht. Die Variablen aus dem Werbetest erlauben einen Rückschluss auf den Erfolg, den der Spot im Markt haben wird. Die Spots, die besser testen, arbeiten auch im Markt besser, sowohl in der Werbeerinnerung als auch in den Abverkäufen.

Tracking und Abverkauf	Gestützte Werbe-Erinnerung	Proven Recall	Abverkauf	Gestützte Werbe-Erinnerung pro Geldeinheit	Proven Recall pro Geldeinheit	Abverkauf pro Geldeinheit
ACTIVITY Werbetest®						
Branding						
Spot-Erinnerung	□	□	■			■
Marken-Recall gestützt	□	□		□		□
Marken-Recall ungestützt			□			■
Submarken-Recall				■	□	
Aufmerksamkeitswert						
Aktivierung	■			■	■	
Benefit 1						
erinnertes Verständnis: Geschmack		■		■	■	■
Verständnis nach Wiedervorführung: Geschmack	□	■		■	■	■
Likes zur Produktdarstellung		□		□		
Bewertung der Produktdarstellung	□	□	■			□
Benefit 2						
erinnertes Verständnis: Convenience	□	■	■			
Relevanz						
Spot verdeutlicht Vorteile	■	■	■	□	□	■
Spot enthält etwas Neues	■	■		■	□	□
Kaufabsichten						
spontanes Artikel-Interesse (Top 2)				■		
Sonstiges						
Text-Erinnerung	□	■	■			□
nicht: unkompliziert	□			■	■	□

Abbildung 1: Welche Pretest-Maße sind besonders wichtig für die Resonanz im Markt?

Der Zusammenhang zwischen den Ergebnissen des Werbetests und den Ergebnissen aus Tracking (gestützte Werbeerinnerung letzte 10 Tage und Proven Recall) und Abverkäufen. Bei den mit "pro Geldeinheit" gekennzeichneten Spalten wurden die Ergebnisse für den jeweiligen Spot mit der Höhe des Werbebudgets gewichtet. Alle markierten Zellen beinhalten Korrelationen von 0.5 und höher. Die schwarz ausgefüllten Korrelationen sind statistisch signifikant ($p < 0.05$ einseitig)

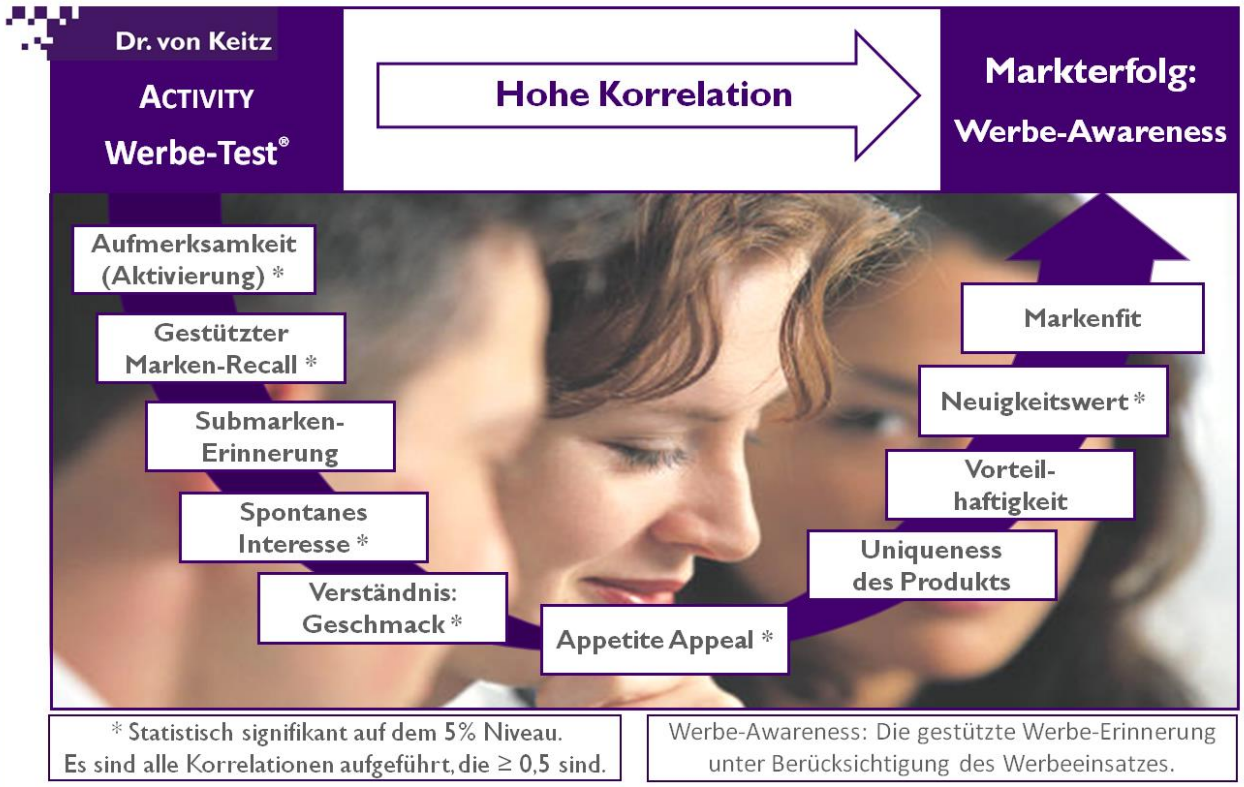

Abbildung 2: Korrelation von ACTIVITY Test Ergebnissen und Werbe-Awareness

Werbetests messen Werbequalität.

Der Zusammenhang zwischen Pretest und Markterfolg wird deutlicher, und es gibt insgesamt mehr Korrelationen, wenn wir den Effekt des Mediabudgets herausrechnen (also die Werte pro eingesetzter Werbegeldeinheit betrachten). Dieses Ergebnis macht deutlich, dass wir mit dem Pretest die *Werbequalität* erfassen und nicht die mit einem höheren Schaltaufwand *erkaufte* Werbewirkung.

Dies gilt nicht für den Proven Recall. Wenn wir das Mediabudget herausrechnen, werden die Zusammenhänge für diese Variable eher schwächer. Anders ausgedrückt: um die Details des Spots im Gedächtnis zu verankern, brauchen wir also neben guter Werbung auch einen entsprechenden Werbedruck.

Gute Werbung ist durch mehrere Erfolgsfaktoren gekennzeichnet.

Die Korrelationskoeffizienten, die sich zwischen Werbetest und Markterfolg in der Studie ergeben, erreichen Werte bis 0.8, sind also positiv, aber nicht hundertprozentig. Schränkt das den Wert der Ergebnisse ein? Aus unserer Sicht nicht; denn wir konnten ja nicht mit einem kontrollierten experimentellen Design arbeiten. Die Einflüsse aus der Marktrealität, beispielsweise die Wettbewerbsaktivitäten, spielen als sogenannte Störfaktoren mit.

Es gibt aber einen noch wichtigeren Grund, weshalb die Korrelationen nicht perfekt sein können: Es gibt verschiedene Variablen, die mit den Marktdaten korrelieren, und damit auch mehrere Wirkungsdimensionen, die gemeinsam zum Markterfolg beitragen. Der Spot muss auf verschiedenen Dimensionen *funktionieren*, damit er im Markt optimal arbeitet. Der Markterfolg ist damit auch nicht eindimensional zu messen. Er ist von mehreren Faktoren abhängig, die im Werbetest berücksichtigt werden müssen und deren Resultate sich in ihrem Informationswert ergänzen.

Zu den zentralen Erfolgsfaktoren gehören Branding, Aufmerksamkeit und die schnelle Kommunikation relevanter Benefits.

Welche Variablen korrelieren im Einzelnen mit dem Markterfolg? Das sind insbesondere das Branding, die Aufmerksamkeitsstärke des Spots, die Kommunikation von Benefits, die Relevanz und die Kaufabsichten bzw. Artikelinteressen. Dabei spielen eher solche Variablen eine Rolle, die wir im Testprogramm in der ersten Phase

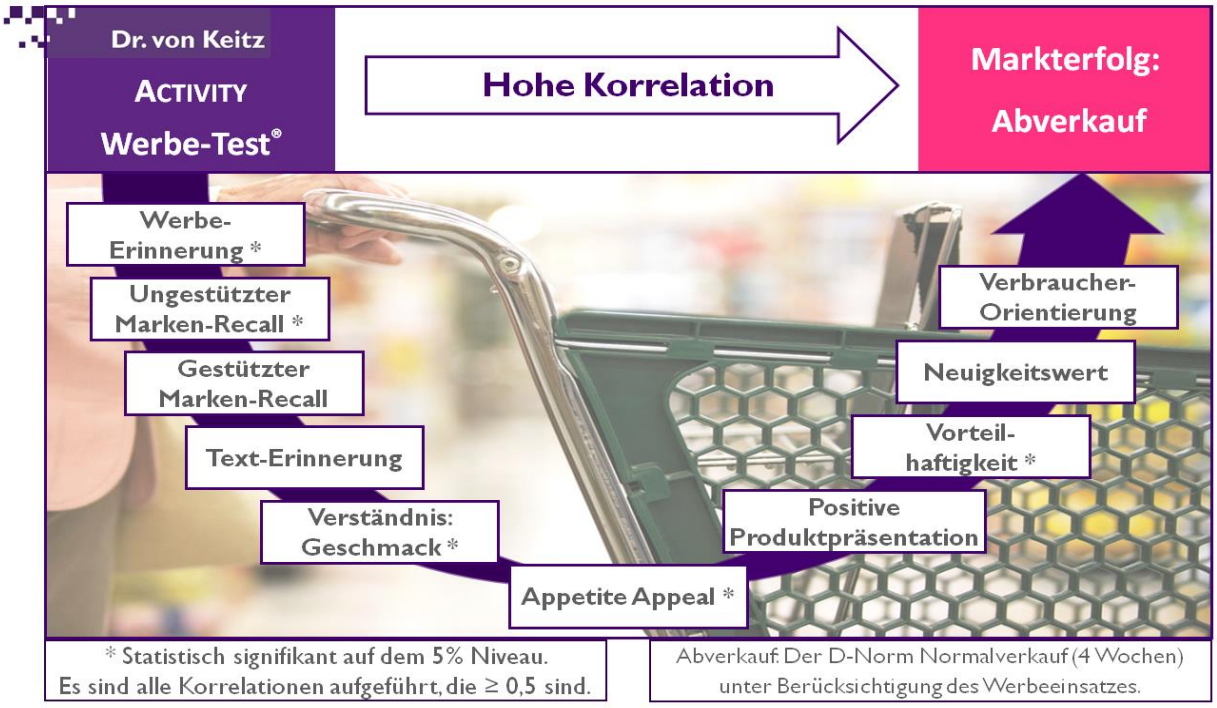

Abbildung 3: Korrelation von ACTIVITY Test Ergebnissen und Abverkauf

erfassen und die auf der spontanen Werbenutzung basieren. Dieses Ergebnis entspricht anderen Daten, wie wir sie in Werbetests mit Response-Elementen finden: Motive, die im Markt stärkere Responses generieren, sind im Test besonders in der Phase der spontanen Kommunikation überlegen.

Der Aktivierungs-Test mit der Aufmerksamkeitsmessung ist eine sinnvolle Ergänzung im Werbetest.

Auch das physiologische Wirkungsmaß, die über den Hautwiderstand gemessene Aktivierung, zeigt einen sehr guten Zusammenhang zur Werbeerinnerung im Tracking. Das bedeutet: Spots, die die Aufmerksamkeit besser halten, setzen sich im Markt besser durch als dezente Gestaltungen, bei denen die Aufmerksamkeit der Zuschauer abflaut. Mit der Aktivierung erfassen wir im Werbetest einen *Multiplikator*, der dann im heimischen Wohnzimmer wichtig für die Werbewirkung wird (vgl. dazu auch von Keitz, 1999). Im Wohnzimmer kann man sich von der Werbung abwenden und Sekundärbeschäftigungen nachgehen. Dieses Risiko besteht für den Spot besonders dann, wenn die Aktivierungskurve während des Spotverlaufs nach unten geht und anzeigt, dass die Aufmerksamkeit des Zuschauers abnimmt.

Die Werbe-Akzeptanz gehört nicht zu den relevanten Erfolgsfaktoren im Markt!

Auffällig ist, dass in der Liste der Erfolgsfaktoren hier keine Akzeptanzparameter erscheinen, z.B. Gefallen, Likes, Dislikes etc. Wenn die Verbraucher einen Spot besser beurteilen, so bringt das noch keinen Vorteil im Markt. Dieses Ergebnis muss vor dem Hintergrund bewertet werden, dass die Spots der laufenden Knorr-Kampagne ohnehin viel Sympathie und Gefallen bei den Verbrauchern finden. Inhaltlich gesehen bedeutet dieses Ergebnis für die Spotgestaltung, dass es wichtiger ist, sich auf Branding und Kommunikationsleistung zu konzentrieren, statt einen Film zu machen, der einfach "gefällt".

Die Umsetzung der Ergebnisse

Welche Schlussfolgerungen haben wir aus der Analyse gezogen? Wie können wir die Ergebnisse nutzbringend einsetzen?

1. Vertrauen in den Test

Die Untersuchung liefert Belege dazu, dass der Werbepretest die Reaktionen im Markt zu prognostizieren vermag und dass er valide arbeitet. Dies ist nicht eine Frage des Glaubens, sondern ein Ergebnis empirischer Daten. Das Vertrauen in die Testergebnisse ist nicht nur auf Seiten der Marktforschung, sondern auch des Mar-

ketings gestiegen, und die Verbindlichkeit für Produktmanagement und Agentur ist damit höher. Man ist sich bewusst, dass die Testergebnisse ernst genommen werden müssen. Die Resultate werden sich hinterher auch im Markt zeigen.

2. Zielorientiertes Benchmarking

Die Variablen, die den Werbeerfolg dieser Kampagne bestimmen, werden in Form einer Checkliste aufbereitet. Diese Faktoren werden von der Marktforschung in besonderer Weise berücksichtigt, um die Testspots zu bewerten und eine Schaltung zu empfehlen oder von einer Schaltung abzuraten bzw. gezielt an kritischen Punkten eine Optimierung zu empfehlen.

Die Ergebnisse werden in Form von konkreten Benchmarks genutzt. Als Benchmarks werden keine allgemeinen Durchschnittswerte verwendet, sondern die Test-Ergebnisse, die die Spitzen-Spots (also die Spots, die sowohl im Tracking als auch im Abverkauf besonders gut arbeiten) erreicht haben. Ein Spot, der einen guten Markterfolg verspricht, sollte diese Werte erreichen.

3. Checkliste für die Kreation

Die Erfolgsfaktoren sind für die Marktforschung wichtig. Gleichzeitig stellen sie aber auch Zieldimensionen dar, die bei der Konzeption und Gestaltung neuer Spots aus der Kampagne besonders berücksichtigt werden müssen.
Was sagt unsere Analyse über das Thema Werbetest und Markterfolg aus?

Zunächst einmal verstehen wir unsere Arbeit als einen empirischen Beitrag zu dieser Thematik. Die vorliegenden Befunde zeigen sich bei bestimmten Produkten, einem bestimmten Kampagnentypus und einem bestimmten Werbebudget und bei überwiegend kurzen Flights. Möglicherweise würden sich Wirkungseffekte, die wir mit dem langen Befragungsprogramm, also relativ spät im Interview, erfassen, dann noch klarer zeigen, wenn die einzelnen Spots länger und mit einem höheren Werbedruck geschaltet würden. Natürlich werden wir unsere Datenbasis für die Analyse noch erweitern. Die neuesten Spots, für die jetzt die Ergebnisse des ACTIVITY Kommunikationstests® und des Trackings sowie die Abverkaufszahlen vorliegen, bestätigen das bisherige Bild: Was nicht gut testet, wird auch später im Markt nicht gut arbeiten. Die Spots, die im Werbetest gut abschneiden, leisten auch im Markt einen guten Job.

Literatur:

Von Keitz, Beate: Aktivierende TV-Spots setzen sich besser durch. Empirische Ergebnisse zur TV-Wirkung, in: planung & analyse, Heft 6/1998

Abdruck mit freundlicher Genehmigung des Deutschen Fachverlages, planung & analyse, Mainzer Landstraße 251, 60326 Frankfurt am Main, Telefon 069-7595-2019, Fax 069-7595-2017, redaktion@planung-analyse.de, www.planung-analyse.de