

Expertenbeitrag

Auf der Suche nach dem Return-on-Communication





Autor
Peter Bauer

peter.bauer@lukb.com

Marketingleiter Luzerner Kantonalbank. BWL Studium Universität Zürich, Nachdiplomstudium Northwestern University Chicago und London Business School. Peter Bauer hatte verschiedene Marketing- und Managementfunktionen bei: 3M (Schweiz) AG, Philips AG, Confetti AG, BBDO Business Communication, Zentrum für Unternehmensführung AG, zuletzt bei CREDIT SUISSE als Vice President, Leiter Marketing Retailbanking. Nebenamtlicher Dozent an Fachhochschulen und Inhaber der Firma Spyglass Marketeering GmbH.

Durch Tests suchen Direktmarketer nach zuverlässigen Entscheidungsgrundlagen, um den Erfolg von geplanten, breit angelegten Direct-Marketing-Kampagnen im Voraus berechnen zu können. Dass die bis auf Kommastellen gerechneten guten Testresultate in vielen Fällen dann doch nicht den erwarteten Erfolg liefern, wird oft erst nach einer erfolglosen Kampagne erkannt. Durch das Beachten einfacher Grundsätze lassen sich diese «Genauigkeitsfallen» vermeiden.

Return-on-Communication im Direct Marketing

Während man sich in der Finanzwelt auf den ROE (Return-on-Equity) konzentriert, verfügen die Marketer nicht über die allgemein gültige, zentrale Grösse – den ROC (Return-on-Communication). Eine einheitliche, branchenübergreifende Kennzahl, die eine klare Aussage über die Werbeeffizienz eines Unternehmens macht, fehlt. Brauchen Marketer keine derartige Optimierungsgrösse oder ist eine zentrale Zielgrösse im Marketing-Kommunikationsprozess ganz einfach nicht sinnvoll oder gar nicht berechenbar? Der Gesamtaufwand lässt sich zwar recht klar berechnen. Der durch die Marketing-Anstrengungen generierte Ertrag allerdings ist gerade bei qualitativen Zielen sehr schwierig zu quantifizieren. Man denke hier an Imagekampagnen, Sponsoring usw. Wesentlich einfacher ist es bei Direct-Marketing-Massnahmen. Durch die relativ einfache Messbarkeit der beim Empfänger ausgelösten Reaktionen lassen sich im DM klare Aussagen über die Wirksamkeit der durchgeführten Massnahmen machen. Dabei sind Tests ein geeignetes Instrument, um sich differenzierte Informationen zu verschaffen. Wesentlich schneller und günstiger als bei einer Vollerhebung mit der ganzen Grundgesamtheit.

Die zentralen Fragen

Welche Parameter testen wir?

Der Erfolg einer Direct-Response-Aktivität hängt nicht von einem Faktor alleine ab. Vielmehr ist es die kombinierte Wirkungsweise aller integrierten Parameter, die ein Werbemittel erfolgreich machen. Es macht wenig Sinn, verschiedene Ansprachen oder die Wirkung von Incentives zu testen, wenn Angebot und Zielgruppe nicht genauestens abgestimmt sind. Das Optimieren einzelner Parameter wie beispielsweise der grafischen Aufmachung hilft wenig, wenn man versucht, Säuglingsartikel an Rentner zu verkaufen. Bei sehr breit angelegten Mailings – diese gehören durch immer präzisere DataMining-Tools schon bald der Vergangenheit an – ist es sinnvoll, einzelne Parameter systematisch zu testen und mittels Wahrscheinlichkeitsrechnung die zu erwartenden Resultate zu berechnen. Doch auch das Testen kann recht aufwändig sein. Bei gross angelegten Mailingkampagnen sind deshalb genügend Mittel für das Durchführen von Tests zu budgetieren.

Die Erfahrungswerte liegen bei bis zu 15% der DM-Gesamtkosten.

Zur Definition der entscheidenden Einflussgrössen eines Mailings ist es unumgänglich, den Empfänger genauestens zu kennen, d.h. seine Wahrnehmung und sein Verhalten gründlich analysiert zu haben. Dazu bedarf es vorgängig gezielter Marktforschung.

Wahrnehmung des Empfängers und seine Entscheidungsbereitschaft

Es gilt also in erster Linie, die für den Empfänger relevanten Parameter zu finden und so zu verpacken, dass sie von ihm wahrgenommen werden und ihn zur Handlung motivieren. Der Empfänger durchschreitet bei der Informationsaufnahme drei Ebenen. Jede Ebene bildet die Grundlage für die Entscheidung, mit der Informationsaufnahme weiterzufahren oder auszusteigen. Aus jeder Ebene kann aber auch immer direkt die Handlung resultieren. Sind die Botschaften in einem Mailing entsprechend aufgebaut, wird der Empfänger klarer geführt, er versteht den Inhalt besser und hat damit eine bessere Grundlage zu entscheiden. Doch nicht nur die Struktur der Botschaften, sondern auch der dynamische Fluss, also die Einbettung und die grafische Unterstützung der textlichen Aussagen, ist für die richtige Wahrnehmung beim Empfänger verantwortlich. Er erleichtert oder erschwert es dem Empfänger, die Informationen zu finden und zu verarbeiten.

Trotz all der vielen Möglichkeiten, eine DM-Massnahme zu optimieren, ist es wichtig, sich vor jedem Test auf wenige zu testende Parameter zu konzentrieren. Denn nur das systematische Vorgehen, d.h. das Verändern einer Variablen bei gleichzeitig unveränderten Restvariablen, kann aussagefähige Resultate über die Response-Treiber bringen. Wird beispielsweise beabsichtigt, verschiedene Listen zu testen, so müssen im Testangebot Timing, Tonalität usw. unverändert bleiben. Neben diesen systematischen Tests einzelner Parameter werden zum Teil auch komplette, unterschiedliche Pakete gegeneinander getestet. Diese Resultate zeigen, welches der getesteten Pakete einen höheren Rücklauf generiert, hingegen liefern sie keine Angaben über die Gründe der besseren Performance.

Folgende Parameter, geordnet nach den aufgezeichneten Kommunikationsebenen, können getestet werden.

1. Stufe: Attraktivität

Ziel: die Aufmerksamkeit des Empfängers gewinnen, die gewünschte Handlung auslösen.

- Adressselektion (induktive oder deduktive Verfahren, Modelle, Listen)
- Angebot (die angebotene Marktleistung)
- Absender
- Personalisierung
- Inhalt (Titel)
- Gestaltung (Format, Farben, Layout)
- Timing (Datum, Saison, Dauer des Angebots)
- Frankatur
- Umschlag
- Antwortelement (Stufigkeit: Zu was sagt der Empfänger «ja»?)

2. Stufe: Relevanz

Ziel: das Interesse des Empfängers wecken, die gewünschte Handlung auslösen.

- Inhalt (allg. Überschriften, Beilagen)
- Angebot (Incentives)
- Antwortelement (Einfachheit, Response-Kanäle)

3. Stufe: Nutzenaffinität

Ziel: den Wunsch beim Empfänger wecken, die gewünschte Handlung auslösen.

- Inhalt (Argumentarium, Text)
- Angebot (Preis, Konditionen)
- Antwortelement (Einfachheit, Antwortumschlag)

4. Stufe: Dynamischer Fluss

Ziel: die Botschaft einfach und verständlich übermitteln.

- Gesamtangebot
- Stil (Tonalität)

Was messen wir?

Das gemeinsame Ziel aller Direct-Response-Medien ist das Auslösen einer Handlung beim Empfänger. Die Rücklaufquote und bei einstufigen Kontaktstrategien auch der Bestellwert sind deshalb die wichtigsten Testgrössen. Mit diesen Testresultaten lassen sich, abhängig von der Kontaktstrategie – also egal ob einstufige oder mehrstufige Mailings – die Kosten pro Interessenten resp. pro Kunden für die geplante Kampagne kalkulieren. Die Kosten eines über eine DM-Massnahme gewonnenen Kunden werden mit dem diskontierten erwarteten Ertrag – dem Barwert – verrechnet. Der so berechnete Nettobarwert oder NPV (Net-Present-Value) muss schliesslich die zentrale Entscheidungsgrösse sein, ob und wie die Grundgesamtheit bearbeitet wird. Ein positiver NPV zeigt, dass sich die Marketing-Kosten gegenüber dem generierten Ertrag rechtfertigen.

Wie testen wir?

Die Zuverlässigkeit des Tests hängt von den zu testenden Varianten ab sowie von der Grösse und Auswahl der Stichprobe. Das Bilden von zuverlässigen Stichproben ist eine sehr wichtige Voraussetzung für einen aussagefähigen Test. Sie setzt voraus, dass die Grundgesamtheit exakt definiert ist. Diese Definition muss auch eine klare räumliche und zeitliche Abgrenzung einschliessen.

Definition der Grundgesamtheit und Stichprobenauswahl

Ziel ist es, eine Stichprobe mit höchster Repräsentanz zu definieren. Einer Stichprobe also, welche die Grundgesamtheit möglichst genau repräsentiert. Verschiedene Möglichkeiten stehen offen. Neben den zufallsgesteuerten Stichprobenverfahren (Random Sampling, einfaches, geschichtetes und Klumpen-Stichprobenverfahren) existieren noch weitere Methoden, wie z.B. Quotenauswahl- und Konzentrationsverfahren. (Quellen: Marketing, Armin Seiler, OF, Zürich).

Stichprobengrösse

Nur eine genügend grosse Stichprobe bringt relevante Aussagen. Die Grösse der Stichprobe hängt allerdings nicht von der Anzahl Adressen, der Grundgesamtheit ab, die man anschreiben möchte, sondern von der erwarteten Rücklaufquote einer DMAktion. Um eine erwartete hohe Rücklaufquote zu testen, reicht eine kleine Stichprobe, währenddem bei einer erwarteten kleinen Rücklaufquote mit einer grossen Stichprobe gearbeitet werden muss, um dieselbe Qualität der Resultate zu erreichen.

Angenommen, wir arbeiten bei zwei zu testenden Listen mit Stichproben von je 400 Adressen. Die Rückläufe sind bei der einen Stichprobe 4,2% und bei der andern 5,3%. Die zweite Stichprobe liefert also die besseren Resultate, und der Schluss liegt nahe, die Liste Nr. 2 zu verwenden. Doch ist der Unterschied in den Resultaten signifikant oder einfach nur das Resultat eines Stichprobenfehlers? Immerhin haben wir ja nur eine kleine Auswahl aus der gesamten Liste getestet, und es gibt keine Gewähr für eine 100%-ige Repräsentanz der Stichprobe. Mit Hilfe des so genannten Vertrauensintervalls lässt sich berechnen, wie die Streuung der Werte in der Grundgesamtheit mit grosser Wahrscheinlichkeit (95%) liegen wird:

$$\text{Vertrauensintervall VB} = 2 \sqrt{\frac{P \times Q}{N}}$$

Formel zur Berechnung der Stichprobengrössen:

$$\text{Stichprobe N} = \frac{2^2 \times P \times Q}{\text{VB}^2}$$

P: der prozentuale Anteil positiver Ereignisse
(z.B. Rückläufe)

Q: der prozentuale Anteil negativer Ereignisse
(z.B. keine Reaktion)

N: Grösse der Stichprobe

Die Vertrauensintervalle zeigen in diesem Fall, dass bei Verwendung der gesamten Liste 1 der Rücklauf mit 95% Wahrscheinlichkeit zwischen 2,2% und 6,2% liegen wird. Bei Verwendung der gesamten Liste 2 würde der Rücklauf ebenfalls mit 95% Wahrscheinlichkeit zwischen 3,1% und 7,5% liegen. Die Testresultate haben also in diesem Beispiel kaum einen signifikanten Unterschied. Anders sieht es aus, wenn statt der je 400 Adressen je 8000 Adressen für die Tests verwendet werden.

Welche Schlüsse ziehen wir?

Es kann also durchaus vorkommen, dass bei zwei unterschiedlichen Resultaten nicht zwingend das bessere auch die bessere Zielgruppe (bzw. andere Elemente wie Clusters, Panels usw.) bedeutet. Mittels der Vertrauensintervalle können wesentlich präzisere Aussagen über die zu verwendende Grösse der Stichproben gemacht werden und auch darüber, ob ein quantitativer Unterschied der Stichprobenresultate signifikant ist oder nicht. Die Resultate der Tests werden auf das gesamte Volumen projiziert, um abzuschätzen, wie der Erfolg der gesamten Kampagne aussehen wird. Die Tabelle zeigt auf, mit welchen Werten man rechnen kann, wenn man eine 95%-ige Wahrscheinlichkeit berücksichtigt. Sie hilft, die «Genauigkeit» der Testresultate richtig zu interpretieren und die richtigen Schlüsse zu ziehen.

Research vs. Testing

Für Direktmarketer ist das Testen nichts Neues. Sämtliche der bereits aufgeführten Fragen werden mit unterschiedlichsten Testverfahren zu beantworten versucht. Oft aber reichen auch ausgeklügelte Tests nicht aus, um die Werbeeffizienz weiter zu optimieren. Was fehlt, sind vorangehende Fragestellungen, die nur durch Marktforschung beantwortet werden können. Trotzdem: Marktforschung kann das Testen nicht ersetzen. Doch durch die Kombination der beiden Instrumente können signifikant bessere Resultate der DM-Aktivitäten erzielt werden. Der Vollständigkeit halber sei hier nur auf die groben Zusammenhänge hingewiesen.

Qualitative Forschung

Die qualitative Forschung ist die erste Phase der Marktforschung. Sie versucht die Motive des Kunden zu ergründen. Dazu werden oft Kleingruppen interviewt, so genannte Fokusgruppen, oder Tiefeninterviews durchgeführt. Die qualitative Forschung liefert Erkenntnisse und kann auch dazu verwendet werden, Hypothesen abzuleiten. Für die Konzeption eines Mailings können daraus wichtige Hinweise gezogen werden:

- in der für den Kunden verständlichen Sprache kommunizieren
- auf die für den Kunden relevanten Handlungsmotive eingehen
- die Einstellung der Kunden gegenüber den verschiedenen Werbemitteln berücksichtigen
- Vorbehalte der Kunden und ihre Probleme im Umgang mit Direct-Response-Medien berücksichtigen
- neue Verwendungszwecke, Nutzen des Produkts erkennen
- die Rolle der Beeinflusser im Kaufprozess identifizieren und die relevanten Entscheidungsfaktoren berücksichtigen

Quantitative Forschung

Die quantitative Forschung folgt der qualitativen Forschung und versucht, durch strukturierte Prozesse die Hypothesen aus der qualitativen Forschung in der Breite zu erhärten. Sie kann u.a. dazu dienen, homogene Zielgruppen bezüglich Verhalten und Einstellungen zu definieren.

Fazit

Tests sind ein geeignetes Mittel, um die Wirkung von DM-Aktivitäten systematisch zu verbessern. Die integrale Betrachtungsweise eines Werbemittels darf ob all der zu testenden Parameter jedoch nicht in den Hintergrund treten. Durch die Einbettung eines Mailings in eine klassische Werbekampagne und gezielte Verkaufsförderung kann die Wirkung der für DM eingesetzten Budgetmittel wesentlich gesteigert werden.

Unternehmen, die ihre DM-Aktivitäten systematisch testen und die Resultate konsequent erfassen, prägen in ihren Marketingabteilungen ein Bewusstsein für den sorgfältigen Umgang mit den generierten Reaktionen. Ist man sich der hohen Kosten bewusst, die ein Interessent generiert hat, werden die Nachfassprozesse für die wertvollen Interessenten-Daten sorgfältiger definiert. Dies ist ein weiterer Faktor für die Wirkung von DM Aktivitäten, der nicht durch Tests verbessert werden kann. Mailings werden nicht nur auf ihren Rücklauf getestet, sondern auch auf ihre Wirkung (Erinnerungswert, Bekanntheit des Angebots usw.). Denn für den Marketer sind nicht nur die Reagierer interessant. Auch die Empfänger, die sich mit dem Mailing auseinandergesetzt haben, schliesslich aber nicht reagieren, sind gerade bei mehrstufigen Kontaktstrategien wichtig. Mit der steigenden Bedeutung von CRM werden auch die Direct-Response-Medien immer stärker als Dialogmedien genutzt. So steht nicht mehr nur der Direktabschluss im Vordergrund, sondern oft das Ziel, den Kunden zum Dialog anzuregen.

Konsequentes Testen sollte einen Lernprozess in Gang setzen. Mit Hilfe von Ranglisten (z.B. die erfolgreichsten und die erfolglosesten Mailings) wird in den Marketing-Abteilungen ein Wettbewerb geschaffen, der die Ressourcen besser kanalisiert und zu besseren Resultaten führt. Und das ist schliesslich das Ziel jedes Direktmarketers – auch wenn damit kein einheitlicher ROC geschaffen werden kann.

Ausblick

Incentives und Loyalität

Die Kreativität für wirkungsvolle Incentives scheint keine Grenzen zu kennen. Mal sind es innovative, mal unvollständige, noch nicht gebrauchsfertige Incentives, die grossen Anreiz zur Reaktion geben sollen. Die immer aufwändigeren Mittel werden getestet und schliesslich zur Steigerung der Response-Quote auch eingesetzt. Zweifellos zeigen diese Belohnungselemente ihre Wirkung. Doch gute Resultate in den Response-Tests dürfen nicht dazu verleiten, die langfristige Loyalitätswirkung bei den Kunden zu vernachlässigen. Kostspielige, attraktive und originelle Incentives, die einerseits die Response-Quote signifikant verbessern, können sich zur Akquisition von loyalen Neukunden als ungeeignet erweisen, da sie Kunden mit opportunistischem Verhalten generieren, welche bei der nächsten Gelegenheit – es lockt eine noch attraktivere Belohnung des Konkurrenten – wieder abspringen. Bei der Beurteilung der Testresultate ist also stets zu bedenken, dass (bei der Neukundenwerbung) der hohe Response zwar den Kampagnenerfolg «sichert», dass aber nur der Response neuer treuer Kunden den Unternehmenserfolg nachhaltig positiv beeinflusst. Hier versagen konventionelle Testgrössen. Es bräuchte eine weitere Entscheidungsgrösse: den Wert des erworbenen Kunden bezüglich Potenzial und Loyalität.

Das Individuum als Segment

Das One-to-One-Marketing – oder die Segmentierung des Markts in Kleinstsegmente von je einem Individuum – stellt Anforderungen, denen herkömmliche Test- und Marketing-Methoden nicht mehr genügen. Hier müssen Prognosemodelle zur Verfügung gestellt werden, die bereits bei der Adressselektion einsetzen. Solche induktive Verfahren, wie sie im Data Mining verwendet werden, liefern in der Regel sehr präzise Listen, die zwar nicht mehr sehr umfangreich, dafür aber qualitativ hoch stehend sind. Zur Verbesserung der Werbeeffizienz werden in diesen Fällen die verschiedenen Prognosemodelle (neuronale Netze usw.) gegeneinander getestet werden müssen.

Die 6 häufigsten Fehler

- zu ungenaue Kenntnis über die Gewohnheiten und das Verhalten der Zielgruppe (mangelnde Marktforschung)
- zu wenig genau definierte Grundgesamtheit
- falsch dimensionierte Stichprobe
- falsch ausgewählte Stichprobe
- zu viele Parameter, die getestet werden sollen
- vernachlässigte Loyalitätswirkung beim Kunden.

Weitere Informationen unter www.post.ch/directpoint.

Die Schweizerische Post
PostMail
Direct Marketing
Viktoriastrasse 21
3030 Bern

Telefon 0848 888 888
E-Mail directpoint@post.ch
Internet www.post.ch/directpoint