

Wolfgang Schaefer:

Ein aktueller Hinweis zum Thema „Prognosetests“¹

Prognose-Tests sind durch den Prozeß der Firma Beecham gegen das Institut Yankelovich auf eine verkehrte Art und Weise aktualisiert worden. Dies ist Anlaß genug, wenigstens ein Verfahren zu skizzieren, das seine Bewährungsprobe längst bestanden hat.

Aufgabe: Risiko-Minimierung

In der noch heute lesenswerten Arbeit über „Pretesting und Testmarkt“ hat Dr. Irninger einmal eine sehr anschauliche Darlegung gebracht, wann sich ein Testmarkt finanziell lohnt. Die von ihm zitierten Autoren waren Myers und Samli. Die Entscheidungsmethode geht davon aus, daß man den durch eine falsche Entscheidung entstehenden Verlust und die durch einen Testmarkt zusätzliche erhöhte Entscheidungssicherheit vorher abzuschätzen versucht. Im Beispiel in Tabelle 1 wurde unterstellt, daß der Mißerfolg 200.000 Dollar kostet (das Beispiel stammt aus dem Jahre 1969; heute wird es teurer).

Tabelle 1	Wahrscheinlichkeit einer korrekten Entscheidung %	Risiko einer falschen Entscheidung Dollar
1.) Entscheidung durch Münz-Werfen	50 : 50	100.000
2.) Verwendung vorhandener Informationen (ohne zusätzliche Marktforschung)	60 : 40	80.000
3.) Verwendung zusätzlicher Marktforschungs-Informationen (Testmarkt)	75 : 25	50.000

In diesem Beispiel sinkt das „Risiko“ von 40 % auf 25 %, wenn man einen Testmarkt o. ä. einsetzt. Wenn der Testmarkt 30.000 Dollar oder weniger kostet, lohnt er sich; kostet er mehr, lohnt er sich nicht mehr, - jedenfalls unter diesem einfachen Blickwinkel. Ein nationaler Reinfall hat ja auch noch andere Folgen, z. B. beim Handel.

Bei der Festsetzung der „Kosten“ ist selbstverständlich auch zu berücksichtigen, welche entgangenen Möglichkeiten auftreten, wenn als falsche Entscheidung die Nichteinführung beschlossen wird.

Jeder kann solche Rechnungen nach seinem Gusto anstellen. Wichtig in unserem Zusammenhang ist die Vorstellung der zitierten Autoren, daß die Sicherheit durch einen Testmarkt nicht im Bereich von 100 %, das Risiko nicht bei 0 % liegt. Sondern es bleibt ein durchaus beträchtliches „Restrisiko“.

Liegt die Größe des Restrisikos an der Unvollkommenheit der Testverfahren? Wird man es durch Verbesserung reduzieren können?

¹ veröffentlicht in planung+analyse, 1988

Aufgrund der Erfahrungen der letzten Jahre nehme ich an, daß das Restrisiko stückchenweise verringert, aber keineswegs verschwinden wird. Als Glaubensbekenntnis: unter 15-20 % wird man es nicht so bald drücken können.

Und irgend ein Rest wird m. E. grundsätzlich bleiben. Wenn ich C. F. v. Weizsäcker (1986) recht verstanden habe, ist die Zukunft grundsätzlich „offen“, weil alle Vorgänge und Entwicklungen einen „probabilistischen“ Charakter haben (wohingegen die Vergangenheit kausal determiniert ist).

Grundlage: Gesetzmäßigkeiten

Ich habe an anderer Stelle darauf hingewiesen, daß wir in der Marktforschung nur in dem Maße Fortschritte erzielen werden, wie es uns gelingt, Gesetzmäßigkeiten im Markt zu finden, und gezeigt, daß es sie gibt (Schaefer 1985). Einige grundlegende Erkenntnisse sind übrigens so neu nicht: siehe Ehrenberg (1966), Parfitt und Collins (1969).

Alle Gesetzmäßigkeiten im Markt haben einen statistischen Charakter. Sie bestehen zumeist nicht aus einfachen, sondern komplexen Beziehungen; sie sind sehr oft nicht linear, sondern entsprechen Wachstums- oder Sättigungskurven.

Test von Silk und Urban

Nach dem, was ich über Tests wie „Bases“ und „News“ weiß, beruht ihr Ansatz und Erfolg auf systematisch gesammelten und analysierten Daten und Tests und dabei gefundenen Gesetzmäßigkeiten. Dazu mag man sich anhand der vorhandenen Literatur näher informieren.

Die eigenen Erfahrungen beziehen sich auf den von den M.I.T.-Professoren Silk und Urban entwickelten Test (1978). Er erschien uns so intelligent, daß wir ihn adoptierten. Die Verfahren „Sensor“ und „Tesi“ beruhen auch auf dieser Entwicklung, so daß die nachfolgenden Ausführungen mutatis mutandis für diese gelten.

Das Herzstück dieses Prognosetests sind zwei Modelle, deren Realisierung zwei weitgehend unabhängige Schätzungen des zu erwartenden Käuferanteils liefern:

1. Kauf- und Wiederkauf-Modell
2. Präferenz-Modell.

Beim 1. Modell wird der von Parfitt und Collins nachgewiesene Sachverhalt benutzt, daß Erstkauf und Wiederkauf den Käuferanteil bestimmen.

Das 2. Modell geht von der Erkenntnis aus, daß es für die Käufer(innen) einen individuellen relevanten Markensatz mit (mehr bis minder) unterschiedlichen Präferenzen gibt, die das Verhalten steuern. Es ist für die Prognose notwendig, die Präferenzen *vor* und *nach* der Kenntnisnahme der neuen Marke zu messen.

Beide Modelle erfordern also zwei Wellen:

1. Welle:

Ermittlung der Präferenzen für die relevanten existierenden Marken, Vorführung von Werbung für die neue Marke und simulierter „Kauf“.

2. Welle:

Ermittlung der Präferenzen unter Einschluß der neuen Marke und Angebot zum „Wiederkauf“²

Die 1. Welle wird üblicherweise im Studio durchgeführt, die 2. telefonisch nach ausreichender Ausprobierzeit der Marke.

Der Test liefert zwei Prognosen mit unterschiedlichen Ansätzen und Rechenverfahren. Im Idealfall stimmen diese überein. Andernfalls muß man prüfen, woher die Diskrepanz kommt; sie ist ja kein Zufall, sondern hat Gründe. Die diagnostischen Teile des Tests (hier nicht ausgeführt) liefern in der Regel dazu die erforderlichen Auskünfte; ferner gibt es hilfreiche Erfahrungswerte.

Die Bestimmung der Zielgruppe(n) liegt primär beim Auftraggeber. Faßt er sie sehr weit, kann man mit dem Test bei entsprechender Anlage herausbekommen, ob sie enger zu fassen ist; faßt er sie zu eng, gilt verständlicherweise die Prognose nur für diesen Ausschnitt und erfaßt nicht außerhalb befindliche Käufer.

Der Test liefert die Prognosen für die untersuchte(n) Zielgruppe(n) bei 100 % Distribution sowie 100 % Markenbekanntheit.

Statt der 100 % Distribution muß in die Verrechnung die zu erwartende Distribution eingegeben werden. Hierzu liegen bei den Auftraggebern zumeist gute Erfahrungswerte vor; es können auch mehrere plausible Alternativen durchgerechnet werden.

Hinsichtlich der zu erwartenden Markenbekanntheit gibt es verhältnismäßig viele Beispiele, an denen man sich orientieren kann. In seiner normalen Version liefert der Test ebenfalls brauchbare Werte (für Bekanntheitswerte der etablierten Marken, mit deren Hilfe man Erwartungswerte ableiten kann). Auch hier können Alternativen durchgerechnet werden.

Aus dieser Darstellung wird (hoffentlich) deutlich: **Der eigentliche Test liefert Prognosen unter bestimmten Rahmenbedingungen.** Diese Rahmenbedingungen können einen großen Einfluß auf das Ergebnis haben; und die gehören zum Verantwortungsbereich des Auftraggebers.

Noch etwas ist in diesem Zusammenhang zu bedenken. Die Prognose - jede Prognose! gilt unter den Bedingungen:

1. des „Status quo“ zum Zeitpunkt des Tests;
2. des „ceteris paribus“.

Sofern sich im Markt, in der Konkurrenzsituation, in der Werbung der Konkurrenz, bzw. am Testprodukt, an seiner Verpackung oder Werbung etwas ändert, sind diese beiden Grundvoraussetzungen verletzt. Dann ist die Prognose nicht mehr gültig. (Sie kann zufällig trotzdem stimmen, so wie eine kaputte Uhr 2x am Tag die richtige Zeit zeigt.)

In Anbetracht dieser möglichen Einflußfaktoren ist es erstaunlich, wie gut die bisher dokumentierten Vergleiche zwischen Test und Marktrealität ausgefallen sind.

² (Dies ist eine stark verkürzte Darstellung der Verfahren, soweit dies für den gegenwärtigen Zweck erforderlich ist. Wer an Details interessiert ist, findet diese bei den Autoren, bei Urban und Hauser und bei Wind.)

Erfahrungswerte für dieses Verfahren

Ein Marketing-Direktor hat uns einmal gesagt, daß ihm die Erfolgsprognose nicht so wichtig sei wie das Aussondern von (wahrscheinlichen) Mißerfolgen. Und genau das passiert nach unseren Erfahrungen mit Testmärkten und diesem Prognose-Test: Die Zahl der damit ausgesonderten Mißerfolge ist weit höher als die der wahrscheinlichen Erfolge. Das aber hat zur Folge, daß die meisten getesteten Produkte gar nicht erst eingeführt werden. Dementsprechend gibt es nur relativ wenige Fälle, in denen das Testergebnis mit dem Ergebnis nach Einführung im Testmarkt und/oder im nationalen Markt verglichen werden kann.

Aber auch dann, wenn die neue Marke oder der Relaunch eingeführt wird, ist der Vergleich oft problematisch: aufgrund des Prognosetests oder aus anderen Gründen werden Veränderungen bzw. Verbesserungen vorgenommen, - sei es am Produkt, an der Verpackung oder an der Werbung.

Glücklicherweise sind in der ersten Zeit nach der Entwicklung des Tests zwecks Überprüfung Einführungen im Testmarkt erfolgt. Somit konnte umgehend nachgewiesen werden, daß der Test so funktioniert, wie er soll. Die Tabelle 2 zeigt die Ergebnisse aus dem Bericht von Silk und Urban (1978).

	Prognose	Marktanteil
Deo	13.3	10.4
Antacid	9.6	10.5
Waschmittel	1.8	1.8
Haushaltsreiniger	12.0	12.5
Shampoo	3.0	3.2
Spülmittel	9.3	8.5
Schmerzmittel	3.0	2.0
Fruchtsaftgetränk	4.9	5.0
Cerealien	6.0	4.4

Jahre später, nach hunderten von Tests, wurde der Versuch unternommen, weitere Vergleiche anzustellen (Urban und Katz, 1983). Aufgrund der geschilderten Umstände, aber auch aufgrund der Zurückhaltung mancher Auftraggeber standen nur 44 Vergleiche zur Verfügung. Die Einzelergebnisse wurden nicht einzeln mit Angabe der Produktkategorie vorgestellt. Statt dessen gibt es Punkte in einer Korrelationsgraphik. Die lineare Korrelation ist hoch ($r = .95$). Die meisten Punkte liegen in der Diagonalen oder nahebei. Die maximale Abweichung liegt bei 16,4 % vs. 22 %, das wären 5,6 %.

Die Berechnungen der Autoren ergaben, daß die Fehlerspanne der Testprognosen knapp der Fehlerspanne der benutzten Stichproben entsprach. Das fanden sie selbst erstaunlich, denn man würde ja neben der Schwankungsbreite aufgrund der Stichprobe auch noch einen Meßfehler erwarten. Möglicherweise führt die Verwendung von zwei unterschiedlichen Prognoseverfahren in diesem Test zu seiner Genauigkeit, meinen die Autoren.

Relative Sicherheit

Die beobachtete Genauigkeit ist eine statistische Genauigkeit: In 10 von 100 Fällen kann die Abweichung des Testergebnisses vom Markterfolg mäßig groß sein, in 5 von 100 Fällen ziemlich und in einem von 100 Fällen sehr groß. Die beiden Autoren führen die oben zitierten

Schätzungen weiter, um auszusagen, wieviel mehr Sicherheit durch einen Test *und* einen Testmarkt gewonnen werden können. Auch sie kommen selbstverständlich nicht auf eine 100 %-ige Sicherheit. Aber die Berechnungen zeigen, daß das Risiko deutlich vermindert wird.

Eine Vergrößerung der Entscheidungssicherheit kann übrigens neuerdings schon **vor** einem Test gekauft werden. Höger hat kürzlich über die Ergebnisse der systematischen Auswertung vieler solcher Prognosetests berichtet (1987). Demnach gibt es (erwartungsgemäß) einige wenige entscheidende Merkmale, die den Erfolg eines neuen Produktes ausmachen. Setzt man plausible Erwartungswerte und allgemeine Daten zum betreffenden Markt in die von Höger vorgestellten Formeln ein, kann man relativ leicht berechnen, wie groß die Wahrscheinlichkeit ist, daß man einen „Hit“ oder einen „Flop“ an der Hand hat.

* * *

Quellen:

- Ehrenberg, A.S.C.: „Regularities of Behaviour“ (1966), in Ehrenberg and Pyatt (ed.) „Consumer Behaviour“, Penguin Books, England, 1971.
- Höger, A.: „Detector - Marktanteilsprognose mit Hilfe von Erfahrungen aus Datenbanken“, in K.-F. Holm (Hrsg.) „Produktforschung“, M+K Hansa-Verlag, Mölln 1987.
- Irniger, Jacques: „Pretesting und Testmarkt / Instrumente der Marketing-Forschung für die Prognose des Absatzvolumens neuer Produkte des Massenkonsums“, Verlag Peter Lang, Frankfurt a. M., 1972.
- Parfitt, J.H.,
und B.J.H. Collins: „Use of Consumer Panels for Brand Share Prediction“, Journal of Marketing Research, May 1968.
- Schaefer, Wolfgang: „Zu den Beziehungen zwischen Werbewirkungen und Kauf bzw. Verwendung“, in K.-F. Holm (Hrsg.) „Werbewirkungsforschung ohne Wirkung?“, Bd. 2, M+K Hansa-Verlag, Mölln, 1985.
- Silk, A.,
und G.L. Urban: „Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology“, Journal of Marketing Research, May 1978.
- Urban, G.L.,
und G.M. Katz: „Pre-Test-Market Models: Validation and Managerial Implications“, Journal of Marketing Research, August 1983.
- Weizsäcker, C.F. von: „Aufbau der Physik“, Hanser, München, 1985.
- Wind, Y.J.: „Product Policy: Concepts, Methods and Strategy“, Addison-Wesley Publ. Reading, London, 1982.