

Wolfgang & Felix Schaefer:

## **Erfahrungen mit dem Prognose-Test von Silk & Urban<sup>1</sup>**

---

Die besondere Leistung der Professoren Alwin J. Silk und Glen L. Urban vom *Massachusetts Institute of Technology* bei der Entwicklung des Prognose-Verfahrens, welches sie selbst „ASSESSOR“- Tests genannt haben und das wir heute vorstellen möchten, bestand u. E. darin, daß sie eine Reihe an sich bekannter Konzepte und Untersuchungsbefunde zu einem Konsistenzen System vereinigten - das dann auch tatsächlich funktionierte.

### **1. ) Das zugrunde liegende Konzept**

#### **1.1.) Kauf und Wiederkauf**

Im wesentlichen durch die bekannte Arbeit von Parfitt und Collins (1969) wurde bewußt gemacht, daß der Marktanteil einer Marke durch Kauf und Wiederkauf bestimmt wird, - genauer: durch die kumulative Erstkauftrate (engl. „trial rate“) und die Wiederkauftrate der Erstkäufer („repeat rate“).

Parfitt und Collins fanden diese grundsätzliche Formel in Haushaltspanels. Daß im Kaufverhalten der Verbraucher Gesetzmäßigkeiten existieren, hatten Bird und Ehrenberg schon früher festgestellt (1966; auch Ehrenberg 1971).

Wenn es solche Gesetzmäßigkeiten gibt, dann sollte es möglich sein, sie in einer entsprechenden Versuchsanordnung im „Labor“ einzufangen:

- Man müßte unter einigermaßen realistischen Bedingungen im ersten Interview einen „Einkauf“ veranstalten; das wäre das Äquivalent der kumulativen **Erstkauftrate**.
- In einem 2.Interview wäre dann die Möglichkeit zum „Wiederkauf“ zu geben; damit hätte man die **Wiederkauftrate** ermittelt.

#### **1.2.) Präferenzen**

Es ist schon lange bekannt, daß gewisse Fragen und Skalen hoch mit dem Kaufverhalten korrelieren; Axelrod hatte 1968 über die verschiedenen Fragen und deren Qualitäten berichtet. Derartige Fragen und Skalen ermitteln „Markenpräferenzen“. Solche Präferenzen steuern das Kaufverhalten, hatte 1971 Juchems vorgetragen.

Wenn dem so ist, dann hätte man einen zweiten Ansatz für eine Prognose. Man müßte in einer ersten Befragungswelle die Präferenzen für existierende Marken messen und nach (werblicher) Vorstellung und Ausprobieren der neuen Marke die Präferenzen - unter Einschluß der neuen Marke - nochmals messen.

---

<sup>1</sup> Dieser unveröffentlichte Aufsatz stammt aus dem Jahre 1988; er basiert auf den Präsentationen, die wir bei Kunden vorgetragen haben und stellt eine Zusammenfassung aller Erkenntnisse, Erfahrungen und Praktiken dar; Teile hieraus wurden in anderem Zusammenhang, etwa beim 1. planung+analyse-Symposium 1985 oder später beim Berliner BVM-Kongress 1995 vorgetragen.

Von den in Betracht kommenden besten Verfahren entschieden sich Silk und Urban für den „Paarvergleich mit konstanter Summe“: alle in Betracht kommenden Marken sind paarweise zu vergleichen.

### 1.3.) „Relevant Set“

Diese Art von Paarvergleich ist ein sehr aufwendiges Verfahren, weswegen es auch nicht so gerne und oft verwendet wird. Denn die Anzahl der durchzuführenden Paarvergleiche (*anz*) steigt mit der Zahl der Marken (*n*) im Verhältnis

$$anz = n(n-1)/2$$

Bei 3 Marken sind das also 3 Paarvergleiche, bei 5 Marken 10, bei 7 Marken schon 21, bei 10 Marken 45 Paarvergleiche. Mehr als 2 Dutzend Paarvergleiche dürften in einem normalen Interview Probleme bereiten; - aber in den meisten Produktfeldern liegt die Zahl der aufgrund ihrer Marktbedeutung in Betracht zu ziehenden Marken (weit) über 10. Was also tun?

Die Lösung lag im Konzept des „evoked set“ bzw. „relevant set“ von Marken. Schon vor längerer Zeit war gefunden worden (Howard and Sheth 1967), daß die meisten Verbraucher weder auf 1 Marke fixiert sind, noch alle auf dem Markt befindlichen echt in Betracht ziehen. In sehr vielen Produktfeldern liegt der für die/ den einzelne(n) Verbraucher(in) „relevante Markensatz“ in der Größenordnung von 3-4 Marken.

Silk und Urban benutzten dieses Konzept und ließen Paarvergleiche für die Marken des „evoked set“ des einzelnen Verbrauchers durchführen. Wie man sah, sind das selten mehr als 7 Marken; also kann man getrost dort abschneiden.

### 1.4.) Vorstellung der neuen Marken

Im üblichen Werbemitteltests führt man das zu testende Werbemittel entweder *solo* vor, oder in einem Umfeld von Werbemitteln für *andere* Produkte.

Das mag für die normalen Zwecke des Werbemitteltests zweckmäßig sein (unsere Zweifel wachsen). Aber realistisch ist es nicht: In der Realität wird die Werbung für die eigene Marke in kurzer Zeit durch die Werbung der direkten Konkurrenten mehr oder weniger stark neutralisiert. Silk und Urban scheinen eine Vorstellung davon gehabt haben, denn sie beschlossen, in ihrem Test die Werbung für die *neue* Marke mit Werbung für *Konkurrenten* gleichzeitig vorzuführen.

### 1.5.) Verrechnungsweisen

Es erscheint uns wenig anschaulich, die notwendigen Verrechnungsweisen hier ohne Bezug zu den jeweiligen Aufgaben vorzustellen. Es mag die Feststellung genügen, daß es auch dafür bereits Ansätze gab. Sie mußten aufgespürt, geprüft und ggf. modifiziert werden. Wir kommen an den entsprechenden Stellen darauf zurück.

## 2.) Zur Aufgabenstellung des Tests

Im Folgenden sollen auf der Basis der vorangegangenen Darstellung des grundlegenden Konzeptes und der zugrunde liegenden Überlegungen nun die typische Aufgabenstellung, der

Testablauf, beispielhafte Verrechnungsweisen und einige typische wie interessante Ergebnisse solcher Prognosetests dargestellt und erläutert werden:

### 2.1.) Ziel des Tests

Der Anspruch dieses Testverfahrens ist es, eine mittelfristige Prognose der Marktchancen neuer oder als „Relaunch“ einzuführender Marken im Konsumgüter-Bereich liefern zu können. Diese Prognose bezieht sich ...

- A) auf den Käufer-Anteil nach Beendigung der Einführungsphase (das ist ein operationaler, weniger ein zeitlicher Bezug, s.u.);
- B) allerdings zu den Bedingungen des „status quo“ zum Zeitpunkt der Untersuchung.

Der Test ist entweder als Vortest oder als Ersatz für einen Testmarkt gedacht. Er soll darüber Auskunft geben ...

- |   |
|---|
| <ul style="list-style-type: none"><li>☺ ob der angestrebte Marktanteil erreichbar ist, also die Einführung gewagt werden kann;</li><li>☹ oder ob das angestrebte Ziel offenbar nicht erreicht werden kann;</li><li>☺ oder ob das zweifelhaft ist.</li></ul> |
|---|

Dabei vereinigt der Test mehrere Messungen in einem:

- |  |
|--|
| <ol style="list-style-type: none"><li>1. zwei (verschiedene) Schätzungen des Markt-Anteils</li><li>2. einen identifizierten Produkttest</li><li>3. einen Werbemitteltest</li><li>4. eine Messung des Effektes der Verteilung von Produktmustern.</li></ol> |
|--|

Dadurch werden diese vier wesentlichen Elemente im Zusammenhang überprüft, eventuelle Schwächen im Gesamtkonzept oder in Teilen davon können herausgefunden werden.

Das bedeutet nicht, daß dieser Test ein Ersatz für Produkt- oder Werbemitteltests ist. Vielmehr sollten Produkt und Werbung vorher getestet und optimiert worden sein. Würde man auf sie verzichten, fände man zwar auch noch in diesem späten Stadium Fehler heraus. Dies wäre aber eine ziemliche Geldverschwendung, weil dergleichen viel früher und damit preiswerter hätten gefunden und ggf. beseitigt werden können.

### 3.) Testablauf

Nachfolgend schildern wir die Testdurchführung aufgrund unserer Praxis; sie mag partiell von denen anderer Testanwender abweichen, jedoch wohl nicht in grundsätzlicher Weise.

### 3.1. Zwei Befragungswellen

Wir verfügen<sup>2</sup> über 4 mobile Teststudios: umgebaute Wohnwagenanhänger, zu deren Umbau wir sachlich und finanziell beigetragen haben. Zwei von ihnen sind für den Prognosetests speziell geeignet, weil sie je 3 Räume haben: für das Interview, für die Vorführung von TV-Werbung, für das „Verkaufsregal“. Diese mobilen Studios werden in Einkaufsstrassen, Einkaufszentren und neben Verbrauchermärkten aufgestellt.

Dort werden von speziell geschulten Interviewerinnen mögliche Zielpersonen für den jeweiligen Test angesprochen; aufgrund der Fragen eines Kontaktinterviews wird festgestellt, ob sie für den Test in Betracht kommen (das Thema „Zielgruppe“ wird später behandelt). Im positiven Fall werden sie zu einem „Filmtest“ eingeladen (wir nehmen hier als Prototyp die TV-Werbung). Sagen sie zu, findet das 1. Interview statt.

Die Testteilnehmerinnen erhalten im Verlauf dieses Interviews Produkte zum Ausprobieren und es wird ihnen gesagt, daß wir dazu später ihre Meinung erfahren möchten; deshalb hätten wir gerne ihre Telefonnummer oder Adresse.

Nach angemessener Ausprobierzeit (2-4 Wochen, entsprechend Verwendungshäufigkeit des Testproduktes) findet ein 2. Interview statt, in der Regel telefonisch.

#### **Inhalt der Interviews:**

- 1) *Auswahl der Befragten* nach Quoten entsprechend der Zielgruppendefinition. Einladung zum Fernsehtest; Prämie für Mitwirkung (geringfügig höher im Wert als Preis der teuersten Marke).
- 2) **Erste Befragung**
  - A) Vor dem Vorführen der Werbefilme:
    - Fragen zur Marken-Bekanntheit
    - Feststellung des „relevant set“ von Marken der Befragten (einschließlich normalerweise benutzter und zuletzt gekaufter Marke)
    - Paarvergleich mit konstanter Summe bezogen auf alle Marken im „relevant set“
    - Fragen zur „Wichtigkeit“ bestimmter Produktmerkmale
    - Zuordnung von Eigenschaften zu jeweils zwei Marken (normalerweise benutzt und in letzter Zeit gekauft)
  - B) Vorführung von 5-6 Werbe-TV-Spots (neue Marke unter Konkurrenten)
  - C) Fragen zur Erinnerung und Beurteilung der Werbespots.
  - D) Im anderen (Verkaufs-)Raum:
    - Möglichkeit zum „Kauf“ einer der ausgestellten Marken mit Gutschein (im Angebot: neue Marke und alle wichtigen Mitbewerber) oder mehr (in bar zu zahlen),
    - oder gar nichts (dafür Geld für Gutschein).
    - Falls Testprodukt nicht gekauft wurde, Übergabe als Geschenk (= kostenloses Muster).
- 3) **Zweite Befragung**
  - Nach 2-4 Wochen Zeit zum Ausprobieren der neuen Marke erfolgt (telefonisches) Interview
  - mit Fragen zur Marken-Bekanntheit und -Verwendung (der neuen Marke),
  - Paarvergleich mit konstanter Summe für alle Marken im „evoked set“ incl. neuer Marke,
  - Zuordnung von Eigenschaften zur neuen Marke,
  - Fragen zum Wiederkauf (zweite Möglichkeit zum „Markenkauf“),
  - Fragen bzgl. der Qualität der neuen Marke.

<sup>2</sup> Diese Aussage bezieht sich auf die frühen 80er-Jahre; heute testen wir vorrangig in festen Studios; die Raumaufteilung ist aber beispielhaft auch in festen Studios vorzusehen

## 3.2.) Die Auswahl der Konkurrenzprodukte

### 3.2.1.) Das übliche Verfahren

Für die Ermittlung der Erstkaufrate, die in das Kauf-/Wiederkauf-Modell einfließen wird, wird im Studio ein Verkaufsregal aufzustellen sein, in dem die wesentlichen Produkte des infrage kommenden Produktfeldes plaziert sind.

Die Frage ist, welche Produkte bzw. Marken müssen im Regal stehen? Im Normalfall wählt man die Konkurrenzprodukte aus den wesentlichen Wettbewerbern in der jeweiligen Qualitäts- und Preisklasse aus, zu der auch das neue Produkt zu zählen ist. Im Prinzip sollten durch diese Auswahl mindestens 80% des jeweiligen Marktsegmentes repräsentiert werden. Falls es verschiedene Sorten gibt, sind diese (anteilig) einzusetzen. Gemeint sind mit „Sorten“ z. B.:

- bei Shampoos:           für trockenes/fettiges/normales Haar
- bei Deos:               die verschiedenen Duftrichtungen.

So einfach und einleuchtend diese Regel klingt: Es können hierbei gleichwohl schon ernste Probleme auftreten, und sachgerechte Lösungen müssen gefunden werden, um die Erstkaufraten-Messung zweifelsfrei zu gestalten. Z. B. standen wir bei einem Test einmal vor der Frage, ob man Zahncremes im „Spender“ neben die normalen Tuben ins Regal stellen sollte. Sie gehörten seinerzeit zunehmend zum normalen Konkurrenzumfeld, waren aber teurer. Die Überlegungen sahen wie folgt aus:

- Würde man solche Marken/Sorten herauslassen, hätte man die betreffenden Käuferanteile mißachtet. Der Einfluß auf die Prognose wäre nicht abschätzbar.
- Nähme man diese Marken/Sorten mit ihrem normalen Preis auf und erhöhte den Wert des Gutscheins, um den höheren Preis mit abzudecken, dann „verleitet“ man u.U. die Testteilnehmer(innen) dazu, Marken/Sorten zu nehmen, die ihnen sonst zu teuer sind.

Nach längeren Überlegungen entschieden Kunde und Institut, die teureren Spender einzubeziehen, aber den Gutscheinwert *nicht* zu erhöhen; wer auf den Spender unbedingt Wert legte, mußte ein wenig eigenes Geld zuzahlen. Das war keine perfekte Lösung, aber die Kaufsimulations-Ergebnisse sprachen für diesen plausiblen Ansatz.

### 3.2.2) Bei großen Preisunterschieden

In zwei anderen Fällen waren die Unterschiede in den Preisen und Packungsgrößen zwischen den in Betracht kommenden existierenden Marken und der neuen Marke extrem groß. Zum Beispiel hätte man zum Preis der einen Marke zwei Packungen einer zweiten bekommen können. Und es war offensichtlich, daß es nicht einfach nur die Packungsgrößen waren, die die Preisunterschiede bedingten.

In diesen Fällen haben wir uns dafür entschieden, das zum „Kauf“ anzubietende „Sortiment“ zu vergrößern: es wurden drei verschiedene Produktbereiche angeboten; aus jedem sollte eine Marke „gekauft“ werden, der Gutschein deckte das ab. Damit sollten die Preis- und Größenunterschiede in der eigentlich interessierenden Produktkategorie weitgehend relativiert oder neutralisiert werden.

In der Praxis zeigte sich, daß diese Versuchsanordnung tatsächlich dazu führte, daß keine artifizien preis- und/oder mengenbewußten Kaufverhaltensweisen durchschlugen.

### 3.3.) Stichproben

#### 3.3.1.) Umfang

Der Stichproben-Umfang wird mehr durch pragmatische als durch theoretische Überlegungen bestimmt. Es hat sich generell gezeigt, daß sich die Ergebnisse von Tests bei 150-200 Interviews zu stabilisieren pflegen<sup>3</sup>. In dieser Größenordnung sollte also die 2. Befragungswelle liegen; das verlangt, in der 1. Welle etwa 250 (auswertbare) Interviews durchzuführen.

In unserer Praxis lagen die Ausschöpfungsquoten der 2. Welle um 75 % bis 85 %, also bei 200 Interviews.

Sofern mehrere Alternativen zu testen sind (das können z. B. unterschiedliche Preise, Werbemittel, Ausstattungen, Produktvarianten sein), muß die Stichprobe vergrößert werden, - möglichst für jede Variante auf 250 Interviews in der 1. Welle.

#### 3.3.2.) Geographische Streuung

Wer sich ernsthaft mit dem Thema Stichproben befaßt, hat natürlich die große Besorgnis, daß ein Test in einer Region oder gar an einem Ort kaum repräsentative, also verallgemeinerungsfähige, hochrechenbare Ergebnisse bringen wird.

Die Praxis hat merkwürdigerweise gezeigt, daß bei diesem Test die Repräsentativität im üblichen strengen Sinn keine große Rolle spielt. Natürlich gibt es bestimmte Produkte, bei denen die Konsumgewohnheiten im Norden anders sind als im Westen oder Süden Deutschlands. Aber derartige Fälle sind normalerweise bekannt; in solchen Fällen haben auch wir die Tests in mehreren Regionen durchgeführt.

Generell haben wir den Eindruck gewonnen, daß mit diesem Testverfahren grundlegende Strukturen der Einstellungen, Präferenzen und Verhaltensweisen angesprochen und erfaßt werden, die von relativ oberflächlichen Unterschieden kaum tangiert werden. Das hat sich u.a. bei Gewichtungen zwecks Angleichungen bzw. Korrekturen von Merkmalen gezeigt, die wenig Effekt zeigten.

Es ist uns jedoch bewußt, daß in diesem Punkt große Wachsamkeit nötig ist sowie alle Bemühungen dahin gehen müssen, alle relevanten Gruppen im Test in genügendem Umfang repräsentiert zu haben. Dies bezieht sich vor allem auf Markenverwenderschaften.

#### 3.3.3.) Zielgruppen

Damit kommen wir zum Thema der in Betracht zu ziehenden Zielgruppe(n) für einen Test.

Die grundlegende Ausgangsbasis stellen die Verwender derjenigen Produktgruppe dar, in welche die neue Marke gehört. Es ist offensichtlich, daß Nichtraucher für einen Zigarettentest nicht infrage kommen. Bei anderen Produkten ist das nicht immer so sicher. Allerdings halten wir die Aussicht, daß eine neue Marke in der Lage ist, über die bisherigen Produktverwender hinaus neue Kunden zu gewinnen, für sehr gering.

Ganz anders ist die Situation, wenn eine neue Marke eine *Innovation* ist, deren Art also bisher gar nicht existierte, oder nicht in der betr. Version/Kombination. Dann gibt es regelmäßig umfangreiche Überlegungen und Erörterungen zwischen Auftraggeber und Institut über die Definition und Abgrenzung der zu betrachtenden Zielgruppe. Generell sollte man einen (billi-

---

<sup>3</sup> s.a. Daniel Starch: Measuring Advertising Effects & Readership,

geren) Konzept- oder Konzept-Produkt-Test vorab durchführen, um die Frage zu klären, wer alles zur Zielgruppe gehören könnte.

Das Problem wird durch den relativ großen Aufwand so bedeutsam. Theoretisch könnte man ja „Jedermann“ (oder doch sehr weite Kreise) einbeziehen und durch den Test selbst herausfinden, wer zur Zielgruppe gehört.

Bei einer Stichprobe von 250 Personen kann das im Hinblick auf eine (sehr) kleine Zielgruppe zu arg kleinen Zahlen und dementsprechend unsicheren Aussagen führen. Umgekehrt würde eine Ausweitung des Tests zu einer erblichen Vergrößerung des Aufwandes an Geld und Zeit führen, - beides ist meist knapp.

### 3.3.4.) Disproportionale Stichprobe

Eine praktikable Lösung bietet der disproportionale Ansatz. Wenn man z.B. vorab eine „Kernzielgruppe“ und einen „weitesten Kreis potentieller Käufer“ definieren kann, hat man die Möglichkeit, diese zu etwa gleichen Teilen in die Untersuchung einzubeziehen. Die Ergebnisse sind dann am Ende entsprechend der realen Verteilung der beiden Teilgruppen in der Grundgesamtheit zu gewichten.

Eine solche Aufteilung ist ggf. in Bezug auf eine schon existierende eigene Marke notwendig: Wenn man eine 2. eigene Marke in einem Produktfeld einführen will, möchte man das Ausmaß an „Kannibalisierung“ erfahren, d.h., wie viel die neue Marke der vorhandenen eigenen wegnehmen wird. Oder positiv formuliert: wieviel gewinnt man per Saldo durch die neue Marke dazu?

Der Test kann eine solche Auskunft liefern (s.u.). Aber die Auskunft ist nur dann ausreichend gesichert, wenn die Verwender der schon existierenden 1. Marke in genügender Zahl in der Stichprobe vertreten sind, um ihre Reaktionen separat analysieren zu können. Und das erfordert im Normalfall einen disproportionalen Stichprobenansatz.

## 4.) Prognose aufgrund Kauf- & Wiederkauf-Modell

Nachfolgend wird an einem konkreten, wenn auch etwas verfremdeten Beispiel dargestellt, wie die Prognose aufgrund des Kauf-/Wiederkauf-Modells zustande kommt. Es handelt sich hierbei um ein Konsumgut mit mittlerer Umschlaggeschwindigkeit in einem Markt, der ca. zwei Dutzend wesentliche Marken umfaßt.

### 4.1.) Erstkauftrate

Die „kumulierte Erstkauftrate“ („trial rate“) wird im 1. Interview geliefert:

Teilnehmerinnen am Test	253
es „kauften“ irgendeine der ausgestellten Marken	238 = 100 %
es „kauften“ die Testmarke	86 = 36 %

Dies bedeutet, daß unter den Bedingungen von ...

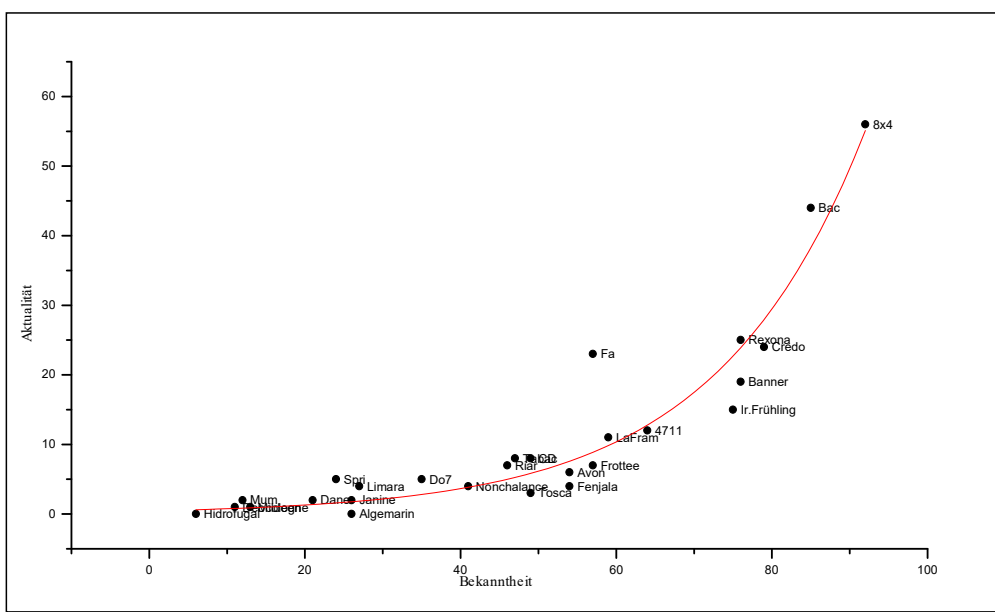
- ⇒ 100 % Distribution und
- ⇒ 100 % Markenbekanntheit

eine Erstkauftrate von 36 % erzielt wurde. In der Realität sind beide Bedingungen natürlich ganz anders. Hinsichtlich der Distribution hat der Auftraggeber mit großer Wahrscheinlichkeit eine ziemlich realistische Vorstellung davon, was er erzielen kann; vernünftigerweise wird man eine gewisse Spannweite angeben und Alternativen durchrechnen.

- Im vorliegenden Fall erwartete der betr. Auftraggeber, eine (gewichtete) Distribution von 85 % zu schaffen.

Desweiteren sind Annahmen zur zu erzielenden Markenbekanntheit einzuführen. Vermutlich gibt es Beziehungen zwischen Werbeaufwand (bzw. Anteil am Werbeaufwand in der betr. Produktgruppe) und Entwicklung der Markenbekanntheit für neue Marken. Leider sind u. W. keine veröffentlichten Daten hierüber in Deutschland vorhanden<sup>4</sup>. Aber man kann z. B. bei unseren „WSM-Werbetrend“-Untersuchungen mehrere Neueinführungen verfolgen und daraus ableiten, ungefähr in welcher Größenordnung die Markenbekanntheit liegen kann. In der folgenden Graphik ist ein solcher Verlauf beispielhaft dargestellt:

**Beziehung zwischen Bekanntheit & Aktualität**



Schaefer Marktforschung GmbH, Hamburg

Bei der hier in unserem Beispiel getesteten Marke wurde eine (gestützte) Markenbekanntheit von 70 % als mittelfristig erreichbar angenommen. Die o.a. Kauftrate ist mit den beiden, so gewonnenen Annahmen zu multiplizieren. Dies führt zur Vorstellung, daß die Erstkäufer-Penetration im Markt rd. 21% betragen müßte:

<sup>4</sup> Mittlerweile gibt es eine interessante Arbeit von ... veröffentlicht in..., die die Beziehung zwischen Werbeaufwendungen und Bekanntheit beschreibt und deren Nutzbarmachung für diese Art von Prognoseverfahren darlegt.



	Erstkaufrate im Studio	36 %
x	Distribution	70 %
x.	Bekanntheit	85 %
=	Erstkaufrate im Markt	21 %

#### 4.2.) Wiederkaufrate

Als nächstes wird die stabilisierte Wiederkaufrate benötigt. Sie setzt sich aus zwei Daten zusammen:

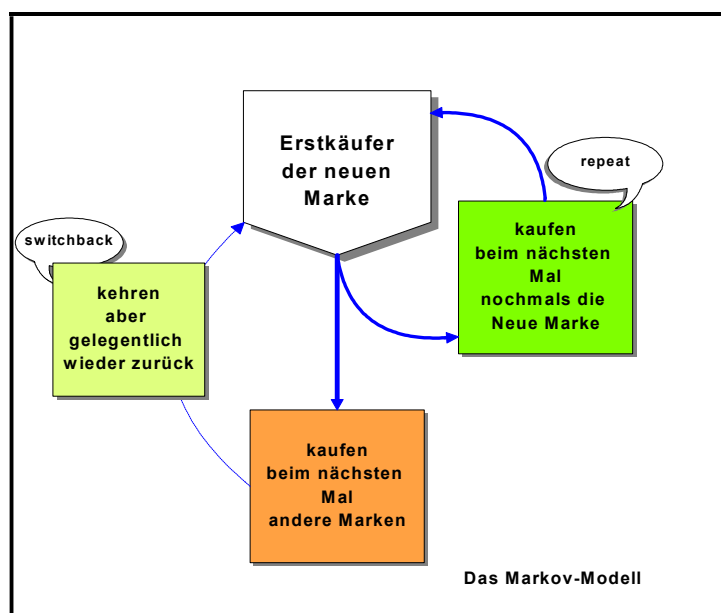
- aus der direkten Wiederkaufrate der Erstkäufer
- und einer „Rückkehr“-Wahrscheinlichkeit für Nichtwiederkäufer.

Wie erstere gewonnen wird, ist ohne weiteres einsichtig; das Ergebnis liefert das 2. Interview, in dem die Möglichkeit zum „Wiederkauf“ geboten wird. In diesem Fall betrug die Wiederkauftrate 73 %.

Ein Teil der Erstkäufer hat sich aber üblicherweise im 2. Interview sich nicht gleich wieder für einen weiteren „Kauf“ der Testmarke entschieden. Aber entsprechend des grundsätzlichen probabilistischen Modells, das der Vorstellung des Geschehens am Markt (festgemacht an der Idee des „relevant set“) zugrunde liegt, ist ein mehr oder minder großer Wechsel zwischen den Marken zu erwarten. Und damit auch eine - gelegentliche - Rückkehr („switchback“) zur Testmarke.

Bei einer etwaigen 3. Befragung würde man auch über deren Größenordnung eine Auskunft gewinnen. Eine solche 3. Welle wäre allerdings aufwendig und ist wegen der zu erwartenden Verringerung der Stichprobe prinzipiell nicht empfehlenswert und wird - außer in Sonderfällen, wenn es dazu wirklich keine Alternative gibt - zumeist auch nicht durchgeführt.

Die von Silk und Urban vorgeschlagene und eingeführte Alternative besteht darin, daß man sich diese Auskunft vom Präferenz-Modell „borgt“. Man kann dort (s.u.) für verschiedene Gruppen die Kaufwahrscheinlichkeiten errechnen, so auch für die „Nichtwiederkäufer“. Genau deren Präferenzen sind es aber, die als switchback-rate eingesetzt werden kann:



Im Test ergaben sich

- ⇒ als Wiederkauftrate „R“ 73 %
- ⇒ als „Rückkehr“-Rate „SB“ 16 %

Daraus ist mit Hilfe der Formel für die Übergangswahrscheinlichkeit bei einem 2-stufigen Markov-Prozesses die stabilisierte Wiederkauftrate wie folgt zu berechnen:

$\text{Wiederkauf} = \text{Switchback} / (1 + (\text{Switchback-Repeat}))$
--

Diese Wiederkauftrate beträgt demnach 37 %.

### 4.3.) Prognose

Das Prognose-Modell sieht vor, daß man Erstkauftrate mit der stabilisierten Wiederkauftrate multipliziert, also die obige 21% Erstkäufer-Penetration mit den 37%, die zuletzt herauskamen.

Die Prognose, die aus dem Kauf-/Wiederkauf-Modell unter den gegebenen Bedingungen folgt, beträgt also 8 % für einen zu erwartenden Käuferanteil.

## 5.) Prognose aufgrund Präferenzmessung

Im 1. Interview werden die Testteilnehmer(innen) gebeten, für alle Marken ihres individuellen relevanten Markensatzes Paarvergleiche durchzuführen, entsprechend ihrer persönlichen Bevorzugung. Silk und Urban haben 11 „Chips“ hierfür vorgegeben, mit deren Hilfe diese Paarvergleiche durchzuführen sind. Diese „Chips“ müssen jedoch nicht physisch existieren; man kann auch 11 Punkte verteilen, etwa 7:4 oder 3:8 oder 9:2 oder 5:6,- ganz so, wie es den Präferenzen für die beiden jeweils zu vergleichenden Marken entspricht. Dabei muß man sich offensichtlich immer zumindest ein „bißchen“ entscheiden, weil 5:6 oder 6:5 das mindeste ist, was an „Parity“ zugelassen ist.

Diese Präferenz-Noten-Verteilungen werden bei der Auswertung in Verhältniszahlen (=Präferenzen) umgewandelt.

### 5.1.) Ein Beispiel:

Präferenz-Noten-Verteilung zwischen drei Marken im „relevant set“ einer beliebigen Testperson:

A = 7	vs.	B = 4
B = 6	vs.	C = 5
A = 8	vs.	C = 3

Die Verhältnisse dieser Präferenznoten werden in eine Ratio-Werte-Matrix eingetragen. Für jede Marke wird daraus nach der Methode der kleinsten Quadrate die Präferenz als *geometrisches* (Zeilen-)Mittel aus dieser Ratio-Werte-Matrix<sup>5</sup> abgeleitet:

<sup>5</sup> Die Berechnung geht zurück auf Torgerson, W.S.: *Theory & Method of Scaling*, N.Y., 1958

Marke	Präferenz
A	$\sqrt[3]{(1*7/4*8/3)}$ = 1,67
B	$\sqrt[3]{(1*4/7*6/5)}$ = 0,88
C	$\sqrt[3]{(1*3/8*5/6)}$ = 0,68
	= <b>3,23</b>

Abschließend werden die Präferenzen zu Kaufwahrscheinlichkeiten verrechnet<sup>6</sup>:

$$P = \frac{\text{Präferenz für Marke A}^{\text{beta}}}{\text{Summe der Präferenzen für relevant-set Marken}}$$

## 5.2.) Justierung am letzten Kauf

Die auf diese Weise ermittelten und errechneten Markenpräferenzen sind abstrakte Werte. Sie müssen auf irgendeine Weise mit dem Kaufverhalten in Verbindung gebracht werden. Silk und Urban vertraten den sicherlich richtigen Standpunkt, daß der jeweils letzte getätigte Kaufakt besser als jede andere Auskunft der Testteilnehmer(innen) die Realität wiedergibt. Also haben sie die Markenpräferenzen an diesem letzten Kauf verankert<sup>7</sup>.

In diesem Fall kam unter Einsetzen der im Test ermittelten Werte und nach Berechnung der Wahrscheinlichkeiten, samt *beta*-Korrektur eine justierte Kaufwahrscheinlichkeit von **38 %** heraus.

Diese beruht auf den Bewertungen durch alle Testteilnehmerinnen. Sie alle hatten ja die Marke ausprobiert, also die Marke sozusagen in ihren „relevant set“ aufgenommen. In der Realität würde das nicht bei 100 % der Konsumenten geschehen, sondern bei einem (viel) kleineren Prozentsatz. Bei wie vielen?

## 5.3.) Berechnung des „relevant set“

Es bestehen generell hohe bis sehr hohe Korrelationen zwischen den Meßwerten ...

- (gestützte) Markenbekanntheit
- Markenaktualität (spontane Markenbekanntheit)
- Marken im „evoked set“/„relevant set“
- für den Kauf in Betracht gezogene Marken
- derzeit verwendete Marken.

Diese Korrelationen können dazu benutzt werden, um eine ziemlich gute Abschätzung zu liefern, wie viele Käufer(innen) diese neue Marke in Betracht ziehen werden.

<sup>6</sup> Basis hierfür sind die Überlegungen von Luce, R.D.: *Individual Choice Behavior*, N.Y.,1969

in der modifizierten Form nach McFadden, P.B.: *Conditional Logit Analysis of Qualitative Choice Behavior*, N.Y.,1970

<sup>7</sup> Die o. a. Gleichung mußte dafür mit einem Faktor „beta“ versehen werden. Zur Berechnung von *beta* dient die iterativ bestimmte Nullstelle der Ableitung der aus der obigen Formel gebildeten Likelihood-Funktion. (Herr Dipl.Math F.Wendt unterstützte uns dankenswerterweise bei der Entwicklung der hierfür notwendigen Programme).

Im konkreten Fall wurden die Beziehungen zwischen gestützter Markenbekanntheit und Markenverwendung benutzt, die eine (nichtlineare) Korrelation von  $r = .93$  aufwiesen, um auf den Umfang des „relevant set“ zu schließen. Dabei kam ein „evoked set“ von 26 % heraus. Setzt man diese Zahl mit der Kaufwahrscheinlichkeit in Beziehung, so erhält man:

$$38 \% \times 26 \% = 9,9 \%$$

Diese Prognose von 9,9 % liegt also um 1,9 % über der des anderen Modells. Die Differenz ist nicht signifikant. Gleichwohl stellt sich die Frage: kann man sie erklären?

Dazu einige Überlegungen: Die stabilisierte Wiederkauftrate des 1. Modells betrug 37 %, die Kaufwahrscheinlichkeit im 2. Modell betrug 38 %. Beide sind strukturell vergleichbar, und sie sind praktisch gleich, wie man sieht. Der Unterschied in den Prognosen kann also nur durch die Schätzungen der Markenbekanntheit und des „evoked set“ produziert worden sein.

Silk und Urban haben in ihrem 1. Bericht darauf hingewiesen, daß besondere Aktionen im Handel oder Werbemaßnahmen vom Kauf-Wiederkauf-Modell nicht erfaßt werden können; Die Markenbekanntheit ist eine ziemlich „statische“ Größe, sie reagiert erst relativ spät auf erhöhte Promotion-Maßnahmen. Aber im Präferenzmodell würden sie durch einen erhöhten „evoked set“ auffallen. Besonders sichtbar wird dies, wenn so etwas in der betr. Branche häufiger praktiziert wird. Dies konnte auch für das betreffende Produktfeld angenommen werden, so daß zumindest ein Teil der 1,9 % dadurch erklärt würde.

## 6.) Sonderauswertungen

### 6.1.) „Kannibalisierung“

Einen „freien Raum“ für eine neue Marke gibt es nicht, - von offensichtlichen Extremfällen einmal abgesehen. Also ist jede neue Marke darauf angewiesen, sich ihren Marktanteil von bestehenden Marken zu holen. Grundsätzlich ist jeder Hersteller sehr daran interessiert, zu erfahren, woher seine neue Marke ihre Käufer gewinnen kann. Zumeist hat er mehr oder weniger begründete Vorstellungen davon, oft auch Daten, die dazu Hinweise liefern.

Das Thema wird besonders brisant, wenn der betr. Hersteller bereits mit einer eigenen Marke vertreten ist, - oder sogar mit mehreren. Eine gewisse Kannibalisierung ist immer zu erwarten; aber vielleicht wird die eigene bestehende Marke ganz besonders stark „angeknabbert“. Evtl. lohnt sich der ganze Aufwand der Neueinführung nicht, falls nämlich die eigene Marke so stark in Mitleidenschaft gerät, daß per Saldo wenig Zuwachs stattfindet. Es kann allerdings immer einmal spezielle Gründe geben, warum trotzdem eine weitere Marke eingeführt wird.

Der Test liefert grundsätzlich eine Auskunft darüber, welche Marken wie stark durch die neue Marke in Mitleidenschaft gezogen wird, vorausgesetzt, die Teilbasen sind groß genug. Meist sind sie es nicht. Dann wird man einen disproportionalen Ansatz wählen, um die interessierende(n) Marke(n) individuell betrachten zu können.

Im folgenden Fall aus der Praxis war der betr. Hersteller bereits 2 Marken auf dem Markt vertreten und hielt damit einen Marktanteil von 64 %. Trotzdem war entschieden worden, eine weitere Marke zu lancieren, um ein bestimmtes (weiteres) Segment zu besetzen. Vorsichtshalber gab man einen Prognose-Test in Auftrag; und man wünschte Auskünfte über den Grad an Kannibalisierung und den zu erreichenden Nettomarktanteil für alle eigenen Marken nach Einführung der neuen.

Die Wünsche waren verständlich. Unsere Voraussage, daß um die 2/3 des Anteils der geplanten Marke von den eigenen bestehenden Marken kommen würden, stieß auf Betretenheit und Unglauben. Woher wir das wüßten? Unsere Gegenfrage: Woher solle der Marktanteil denn sonst kommen?

Wie sich herausstellte, hatten wir uns verschätzt, das Ergebnis war noch schlimmer: Es zeigte sich eine Kannibalisierung von rund 80 %, mit wenig Nettozuwachs für den Kunden, wie folgende Tabelle zeigt. Die neue Marke wurde nicht eingeführt.

Anteil der vorhandenen eigenen Marke	64 %
Prognose für neue Marke	10 %
davon Anteil von vorhandenen eigenen Marken	8 %
zusätzlicher Marktanteil	2 %

## 6.2.) Verteilung kostenloser Muster

Jene Testteilnehmerinnen, die im 1. Interview kein Produkt oder aber ein anderes als das Testprodukt „kaufen“, erhalten das Testprodukt (zusätzlich) als *kostenlose Probe*. Da im zweiten Interview auch von diesen Teilnehmerinnen die Kaufbereitschaft und die Präferenzen für das neue Produkt ermittelt werden, kann man errechnen, was man mit einer „Musterverteilung“ erzielen kann.

Nachfolgend ist dies an einem Beispiel dargestellt. Hierbei sind weitere Annahmen bzw. Erfahrungswerte einzusetzen, nämlich:

- a) An wie viele Haushalte sollen Muster verteilt werden? Hier sind zwei Möglichkeiten eingesetzt worden: 50 % und 75 %.
- b) Ferner muß man eine Schätzung einsetzen, welcher Prozentsatz der bemusterten Haushalte das Muster auch tatsächlich erhalten und verwendet wird..

|| Nach unseren Erfahrungen<sup>8</sup> sollte man hier eher pessimistische Annahmen einsetzen. Wir empfehlen 60-80 %.

Ferner ist davon auszugehen, daß jede Musterverteilung die sowieso schon durch *Werbung* zum Kauf der neuen Marke angeregten Personen genau so erreicht wie die Nichtkäufer. Das heißt, die kumulierten Erstkäufer (n diesem Fall 8,5 %) sind also vom Bemusterungs-Erfolg abzuziehen.

<sup>8</sup> Es gibt eine Reihe von Untersuchungen, die wir zum Thema „Musterverteilungsaktionen“ durchgeführt haben und bei denen wir ermittelt haben: a) wie viele Proben tatsächlich verteilt wurden, b) wie viele tatsächlich in den Briefkästen landeten, c) wie viele davon dann bemerkt, d) tatsächlich erinnert / genutzt wurden.

		<b>Plan A</b>		<b>Plan B</b>	
zu bemusternde Haushalte		<b>50 %</b>		<b>75 %</b>	
x	Muster erhalten und benutzt	60	80	60	80
=	Musterverwender	30	40	45	60
x	Nicht-Erstkäufer (100-8,5 % =)	91,5	91,5	91,5	91,5
=	zusätzliche Haushalte, die Muster benutzten	27,5	36,6	41,2	54,9
x	Erstkauftrate	53	53	53	53
x	Distribution	70	70	70	70
x	Wiederkauftrate	34	34	34	34
=	zusätzliche Käufer durch Musterverteilung	3,5	4,6	5,2	6,9

## 7.) Alternative Berechnungen

### 7.1.) Distribution

Der Auftraggeber hat normalerweise Erfahrungen, wieviel *Distribution* er erzielen kann. Im Zweifelsfall kann er Informationen über Distributionsniveaus vergleichbarer Marken bei den entsprechenden Marktforschungsinstituten kaufen<sup>9</sup>.

Die Annahmen hinsichtlich der erzielbaren (gewichteten) Distribution werden also der zu erwartenden Realität ziemlich nahe kommen. Trotzdem tritt ein verständliches Unbehagen auf, ob man denn mit einer immer noch vorhandenen Unsicherheit zu einer brauchbaren Prognose kommen kann.

Nun werden die Distributionsannahmen nicht extrem sein, - also nicht: von 10 % bis 90 %. Sondern sie liegen erfahrungsgemäß in einem verhältnismäßig engen Bereich von max. +/- 10 %. Dementsprechend können die Auswirkungen gar nicht sehr groß sein. Wenn man das erste Beispiel variiert, so kommen folgende Werte heraus:

<b>Annahme der (gewichteten) Distribution</b>				
	60 %	70 %	80 %	90 %
Prognose:	6,8 %	7,9 %	9,1 %	10,2 %

<sup>9</sup> Handelspanel = Nielsen

## 7.2.) Markenbekanntheit

Grundsätzlich gilt, daß eine Marke bekannt sein muß, um gekauft zu werden. Was „Bekanntheit“ ist, ist schwer genau zu definieren; wahrscheinlich kommt eine „bewußte Wahrnehmung“ dem Sachverhalt am ehesten nahe.

Für unsere Zwecke spielt es keine Rolle, woher die Bekanntheit kommt: von der Werbung, von der Wahrnehmung der Marke im Laden, durch Musterverteilung, oder wie sonst. Die Frage erhebt sich, welche Markenbekanntheit zu erwarten ist. Es gibt zahlreiche „Tracking Studies“, in denen u.a. die gestützte Markenbekanntheit ermittelt wird. Wir selbst hatten in der Zeit von 1980-1987 unter dem Namen „WSM-Werbetrends“ Umfragen dieser Art durchgeführt und somit Daten aus ca. 50 Produktfeldern mit über 1.000 Marken gesammelt gehabt. Aufgrund solcher Datenbestände kann man den Spielraum für neue Marken recht gut eingrenzen.

Leider fehlen in Deutschland veröffentlichte Sekundäranalysen zum Thema: wieviel Bekanntheit kann man bei einem Werbeetat von DM „x“ Mio. erwarten? Bzw., wieviel Prozent des Werbedrucks in der betr. Produktgruppe muß man erreichen, um ein bestimmtes Bekanntheitsniveau zu erzielen? In Abwesenheit solcher nützlicher Verallgemeinerungen ist trotzdem der Spielraum plausibler Annahmen nicht groß - genau wie bei der Distribution. Und ebenso sind die Auswirkungen nicht extrem, wie folgende Darstellung zeigt.

	angenommene Bekanntheit			
	60 %	70 %	80 %	90 %
Prognose:	5,6 %	7,0 %	7,9 %	8,4 %

Auch die Kombination der Annahmen zu Distribution und Bekanntheit führt nicht zu extremen Schwankungen. Aber man muß bei diesen Berechnungen und Betrachtungen berücksichtigen, daß einerseits die unterschiedlichen Annahmen mit unterschiedlichen Wahrscheinlichkeiten behaftet sind. Andererseits können Distribution und Bekanntheit nicht extrem auseinanderklaffen, schon weil Distribution selbst Bekanntheit schafft.

Nachfolgend eine Matrix welche die Folgen der Verknüpfung obiger Annahmen darstellt:

	angenommene Distribution			
	60 %	70 %	80 %	90 %
<b>angenommene Bekanntheit:</b>				
60 %	4,8 %	5,6 %	6,4 %	7,2 %
70 %	6,0 %	7,0 %	8,0 %	9,0 %
80 %	6,8 %	7,9 %	9,1 %	10,2 %
90 %	7,2 %	8,4 %	9,6 %	10,8 %

## 8.) Diagnostische Daten

Die in einem solchen Prognose-Tests automatisch anfallende Kombination von einer Reihe von Ermittlungen & Bewertungen liefert eine Fülle von diagnostischen Daten und damit fast zwangsläufig zusätzliche Hilfestellungen für die Verbesserung der Werbung auf der einen Seite oder des Produktes auf der anderen. Zum Teil haben wir darüber schon berichtet<sup>10</sup>, aber an dieser Stelle lohnt es sich, wesentliche Möglichkeiten nochmals im Zusammenhang darzustellen.

### 8.1.) Werbemittel-Test

Nach der Vorführung der Werbemittel für die neue Marke und ihrer Konkurrenten werden üblicherweise 4 Ermittlungen vorgenommen:

#### 1.) Erinnerung an die Marke

Zunächst wird erfragt, an welche Werbemittel man sich erinnert, mit Nennung der Marken. Entsprechend der seit langem vorliegenden Erkenntnisse (Axelrod 1968), ist die Nennung einer Marke an **1. Stelle** von Bedeutung; dementsprechend wird sie von uns extra ausgewiesen. Die neuesten hierzu publizierten Erkenntnisse gehen in die gleiche Richtung (Höger 1987, Blanchard und Schneersohn 1987).

Darüber hinaus kann dieser Meßwert mit anderen Daten verknüpft werden und somit eine Abschätzung der einmal zu erreichenden Markenbekanntheit führen. Die Kette von Regressionsgleichungen, die hierzu verwendet werden kann reicht von „Erinnerung an die explizit genannten Marken“ über „Werbeaktualität“, bezogen auf das etablierte Markenumfeld, die zugehörigen „Aktualitätswerte“, die bei Aufrechterhaltung der Werbemaßnahmen zu erreichen sind - und letztendlich das Niveau an „Bekanntheit“, das erzielt werden kann, wenn alles dies über einen längeren Zeitraum aufrecht erhalten werden kann.

#### 2.) Erinnerung an Inhalte

Ein zweiter Fragenkomplex ermittelt für alle erinnerten Werbemittel, welche Inhalte und Darstellungsdetails im Gedächtnis hängen geblieben sind. Vorrangiges Ziel dieser Ermittlung ist, ob die vom Markenkonzzept her wichtigen Aussagen richtig aufgenommen werden und in welchem Ausmaß. Dies kann besonders dann wichtige Erkenntnisse liefern, wenn die im Test ermittelte Erstkauftrate hinter den Erwartungen zurückbleibt. Unter Umständen ist dann nämlich genau das passiert, was Rosser Reves, dem „Erfinder“ des USP, in einem bemerkenswerten Interview in Advertising Age (April 1965) „Misdirection“ genannt hat - eine, wie er sagte, typische Todsünde vieler interessanter, lustiger, ansprechender - und vermutlich preisgekrönter Werbespots, Anzeigen oder Plakate.<sup>11</sup>

#### 3.) Beurteilung der Werbung „an sich“

Die Werbemittelforschung der letzten Dekade hat nachgewiesen, daß es einige wesentliche Faktoren in der Beurteilung von Werbemitteln gibt, die für den Werbeerfolg besonders verantwortlich sind (Zusammenfassung bei Schaefer, 1986).

<sup>10</sup> Wolfgang Schaefer: Zu den Beziehungen zwischen Werbung und Verwendung bzw. Kauf; 1.pl+a-Symposium, 1985

<sup>11</sup> in unserem Vortrag zum Thema „Erfahrungen mit Prognosetests“, gehalten 1995 auf dem 40. Kongress der deutschen Marktforschung in Berlin, greifen wir dies Beispiel nochmals auf - siehe auch weiter hinten in dieser Aufsatzsammlung



Entsprechend dieser Erkenntnisse haben wir unsere eigene Liste von Werbemittelbeschreibungen in den Test eingebaut, die sich im wesentlichen auf die folgenden Faktoren beziehen:

- Wichtigkeit der Aussage(n)
- Überzeugungskraft der Werbung
- sympathische Darbietung
- Besonderheit(en) der Marke
- Originalität/Auffälligkeit der Werbung
- Aussage realistisch
- Verständlichkeit/Prägnanz der Werbung

#### **4.) Beurteilung der neuen Marke - aufgrund der Werbung**

Die erste Beurteilung der neuen Marke erfolgt üblicherweise nach dem Vorführen der Werbemittel; mithin hat diese Ermittlung nichts mit dem Produkt an sich, sondern nur mit seiner werblichen Darstellung zu tun: welche Erwartungen an Qualitäten und Produktleistungen ruft diese hervor. Mit „werblicher Darstellung“ ist hier zwar primär das Werbemittel gemeint; jedoch spielen Markenname und Packungsgestaltung u.U. ebenfalls eine recht erhebliche Rolle.

Im zweiten Interview wird dann üblicherweise die Produktbeurteilung anhand der gleichen Beschreibungen nochmals abgefragt.

Durch Vergleich der durch die Werbung geweckten und im ersten Interview ermittelten Erwartungen und die dann im zweiten Interview nach der Produktprobierphase erfaßten Bewertungen der tatsächlichen Produktleistungen erkennt man, wie weit beides übereinstimmt - oder ob überzogene Erwartungen oder nicht erwartete Leistungen das Erst- oder Wiederkaufbereitschafts-Ergebnis nachhaltig und u.U. in eine negative Richtung gehende Weise beeinflußt haben.

## **8.2.) Produkt-Beurteilung**

Im Zusammenhang mit der Produktbeurteilung werden zumeist 3 Fragen behandelt:

### **1.) Wichtigkeit von Eigenschaften**

Entsprechend der begründeten Forderung, daß eine Produktaussage zunächst einmal von Bedeutung (wichtig/relevant) für die Zielgruppe sein muß, wird manchmal eine Liste relevanter Produkteigenschaften vorgegeben und ermittelt, für wie **wichtig** die einzelnen angesehen werden.

Wir sind grundsätzlich der Meinung, daß es sehr spät ist, diesen Sachverhalt erst im Rahmen eines Prognosetests zu ermitteln; das gehört u. E. in den sehr frühen Bereich der Konzept- und Produktforschung. Außerdem ist die Fragestellung nicht ohne eigene Probleme. So ist z.B. bekannt, daß der Geschmack einer Zahnpasta von ausschlaggebender Bedeutung für die Akzeptanz ist, aber bei der Abfrage der *subjektiven Wichtigkeit* kommt diese Eigenschaft nicht an prominenter Stelle zum Vorschein.

Sofern aber überhaupt keine Daten zur Wichtigkeit von Produkteigenschaften vorliegen, kann es zweckmäßig sein, eine entsprechende Frage in den Prognose-Test aufzunehmen. Denn wenn eine unbefriedigende Erstkaufrate herauskommt, dann kann einer der Gründe dafür sein, daß für die Verbraucher wichtige Eigenschaften nicht genügend oder gar nicht vermittelt worden sind. Und dies sollte man wissen, um es ggf. in einem 2. Anlauf zu berücksichtigen.

Generell tendieren wir allerdings dazu, dies „indirekt“ zu messen, indem man die Bewertungen, die sich im Zuge der Frage nach den „erwarteten Produktleistungen“ ergeben, faktorenanalytisch verdichtet und dann per Regressionsanalyse auf die Kaufbereitschaft bzw. im Falle des hier betrachteten Prognosetests auf die Präferenzmessungen bezieht.

Dabei kommt dann nicht heraus, was „man“ so ganz generell bzgl. einer bestimmten Produktkategorie als wichtig erachtet, sondern was im ganz konkreten Fall die Kaufneigung bestimmte. Führt man diese Analyse mit Hilfe der Produktbewertung, die man nach dem Ausprobieren ermittelt hat durch, dann ersieht man die Faktoren, die die Wiederkaufabsicht bestimmen. Interessant kann gleichwohl sein, wie die per direkter Abfrage ermittelte „Wichtigkeits-Rangfolge“ aussieht - und die, welche sich aus den „indirekt“ messenden Regressionsanalysen ergibt.

## **2.) Beurteilung etablierter Marken**

Wie die als hauptsächliche Konkurrenten angesehenen Marken beurteilt werden, sollte man eigentlich ebenfalls schon in einer früheren Entwicklungsphase der neuen Marke erforscht haben (also etwa im Rahmen von U+A-Studies). Ohne eine solche Kenntnis kann man Produktgestaltung und Werbung eigentlich gar nicht gezielt auf den konkreten Markt hin entwickeln.

Aber es kommt (gar nicht selten) vor, daß diese Kenntnis nicht vorhanden - oder zumindest nicht auf einer breiten quantitativen Basis ermittelt worden ist. Dann besteht natürlich die Möglichkeit, noch schnell im ersten Interview, aber *vor* der Vorführung von Werbung, das Vorstellungsbild aller aktuell am Markt erhältlichen Konkurrenzprodukte zu ermitteln.

Erstellt man auf diese Basis geeignete „Mappings“ (mit Hilfe von Faktoren- u/o Korrespondenz-Analysen) dann erhält man eine recht gute Basis, um die später ermittelten Erwartungen und realiter gemessenen Produktleistungen einzuordnen und kann ablesen, ob sich die neue Marke genau dort positioniert, wo man sie plazierte haben wollte.

## **3.) Beurteilung der neuen Marke**

Die Bewertung der neuen Marke in der zweiten (Produkttest-) Welle des Tests schließt dementsprechend den diagnostischen Teil ab. Wie mit Blick auf die vorher durchgeführten Bewertungen formuliert wurde, gilt auch hier: Üblicherweise liegen Bewertungen aus Produkttests vor, wenn man sich für die Durchführung eines Prognosetests entschließt, wenn auch vielleicht nur in Form von Blindtests.

Insofern kann die im Rahmen eines solchen Prognosetests durchgeführte Produktbewertung neue Aufschlüsse liefern: Wie stellt sich die Produktleistung

unter dem Einfluß des Markenimages dar? Wie sieht die Bewertung der Produktleistung im Vergleich zu den von der Werbung geweckten Erwartungen aus?

Auch hierzu gibt es ein Beispiel: Wir hatten einmal ein Waschmittel getestet, das als Zusatznutzen eine „eingebaute“ Weichspüler-Wirkung aufwies. Die Einführungswerbung zeigt ein junges Mädchen in der Straßenbahn in einem bezaubernden luftigen Sommerkleidchen.

Als Message kam an: Hier wird für ein Feinwaschmittel geworben. Sicherlich macht dies weich, sanft, luftig - aber an der *Waschkraft* gab es erhebliche Zweifel. In der Folge ergab sich eine so niedrige Erstkauftrate, daß der Auftraggeber schon an einen Abbruch des Tests und die Einstellung des Projektes dachte.

Im Produkttest zeigte sich dann, daß das Produkt gegen alle Erwartungen (der Befragten) den Ansprüchen an ein Vollwaschmittel vollauf genügen konnte. Dies zeigte, daß der Produkttestteil, auch wenn er erst so spät in der Einführungsphase zum Zuge kommt, durchaus seine Berechtigung hat: als diagnostisches Instrument und als Hilfe bei der Abstimmung von werblichem Konzept einerseits und Produktleistung andererseits. Von Ehrenberg wissen wir, daß Werbung allein kein Produkt attraktiv machen kann. Die ausgelobten Eigenschaften müssen sich bestätigen, dann kann ein Anstoß in Richtung auf einen Nachkauf erfolgen, sich nach und nach eine stabile Wiederkauftrate aufbauen und ein Erfolg des Produktes erwartet werden.

---

#### **Danksagungen**

Den Anstoß zur Beschäftigung mit diesem Prognose-Test verdanken wir Marianne Drehmann und Friedo Goetzky, Marktforscher der Firma Colgate-Palmolive GmbH, Hamburg. Sie und ihre geschätzte Firma haben uns bei der Entwicklung der notwendigen Computer-Programme fachlich und finanziell unterstützt, in zahlreichen Gesprächen anlässlich konkreter Aufgabenstellungen durch fachkundige Ratschläge das Verständnis des Tests und seiner Voraussetzungen vergrößern helfen.

Bei der Entwicklung der Programme hat uns Hans-Peter Lorenz durch Einführung in APL und großzügig berechnete Computerzeit geholfen. Unser Freund Friedrich Wendt hat uns einige sehr grundsätzliche Auskünfte über relevante Verfahrensweisen der mathematischen Statistik und einschlägige Fachliteratur gegeben, so daß wir an entscheidenden Stellen vorangekommen sind.

Bei der Literaturzusammenstellung haben uns geholfen: Armin Höger, Walter Lulay und Jörg Stoffels.

Ihnen allen sei an dieser Stelle sehr herzlich gedankt.

#### **Quellenhinweise**

Die modernen Prognosetests sind selbstverständlich nicht aus dem „Nichts“ aufgetaucht, sondern beruhen auf z. T. schon viele Jahre zurückliegenden Ideen, Konzepten, Verfahrensweisen und Ergebnissen zahlreicher Forscher. Es schien uns nützlich, auf dergleichen Quellen schon in diesem Aufsatz hinzuweisen: Einerseits, um die Historie und Entwicklung zu dokumentieren, was auch zu gebührender Anerkennung der betr. Forscher führt; andererseits soll damit die Beschäftigung mit dieser so überaus interessanten und wichtigen Materie erleichtert und gefördert werden. So haben wir mittlerweile eine Literatursammlung mit ca. 150 Titeln angelegt, die wir Interessierten gerne zur Verfügung stellen. Über Hinweise auf weitere wesentliche Veröffentlichungen würden wir uns sehr freuen, um sie ggf. in einer Neuauflage angemessen zu berücksichtigen.