

Zu den Beziehungen zwischen Werbemittel-Beurteilungen und Kaufneigung

von Wolfgang und Felix Schaefer

Vorbemerkungen

Wann immer man in der Forschung inne hält, um einen Bericht über das darzulegen, was sich bisher ergeben hat, wird man sich der Tatsache der Vorläufigkeit bewußt, welche der Arbeit anhaftet. Während man noch an der Bestandsaufnahme arbeitet, ist die praktische Arbeit schon weiter; wollte man diese weiteren Schritte einbeziehen, käme man nie zum Ende. Dies ist in der sog. "kommerziellen" Forschung mit ihrer großen Geschwindigkeit wohl besonders ausgeprägt, besonders dann, wenn sich die beteiligten Forscher nicht mit der routinemäßigen Anwendung von standardisierten Verfahren begnügen, sondern bemüht sind, die Grenzen der Kenntnisse zum Horizont hin zu verschieben (diese Metapher wurde bewußt gewählt, denn der Horizont wandert bekanntlich immer unerreichbar vor uns her). Dies hat zur Folge, daß man das Erreichte bereits für überholt und kaum für Wert hält, vorgetragen zu werden, da schon in Kürze mehr zu sagen sein wird. Aber es ist für die Kommunikation in der Forschungsbranche erforderlich, in gewissen Abständen Zwischenbilanz zu ziehen. Dies soll der folgende Aufsatz leisten; er ist zweiteilig:

1. Zunächst zitieren wir andere relevante Untersuchungen. Selbstverständlich bauen wir auf zahlreichen und bedeutenden Tests und Analysen anderer Forscher auf.

Um den Zusammenhang erkennbar zu machen gehen wir auf jene Untersuchungen ein, die für unsere Arbeit größere Bedeutung hatten.

2. Dann berichten wir über ein von der Colgate-Palmolive-Marktforschung entwickeltes Verfahren zum Testen von Werbemitteln und die zusammenfassende Analyse von 53 Tests mit diesem Verfahren.

Werbemittelforschung von Politz

Historisch gesehen stand die von Alfred Politz Research entwickelte Methode am Anfang unserer Tätigkeit: *Informarkt*, unser Vorgängerinstitut, gehörte zur Politz'schen Firmengruppe; es und 1965 wir übernahmen die Methode.

Sie ist grundsätzlich anderer Art als jene Untersuchungen, über die wir anschließend referieren, und zwar verwendet sie das Verfahren von Test- und Kontrollgruppe:

In der **Testgruppe** wird das Werbemittel für das Testprodukt vorgeführt (im Rahmen von Werbemitteln für Marken aus anderen Produktfeldern).

In der strukturgleichen **Kontrollgruppe** werden nur die anderen Werbemittel vorgeführt.

Anschließend findet eine in beiden Gruppen identische Befragung statt. Statistisch signifikante Unterschiede in den Antworten der beiden Gruppen können somit auf die einzige Variable zurückgeführt werden, die verändert wurde: die Werbung für das Testprodukt.

leicht überarbeitete Version des Aufsatzes in: K. F. Holm (Hrsg.):
Werbewirkungsforschung ex ante und ex post, M+K Hansa Verlag, 1986

In weiteren strukturgleichen Gruppen können andere Varianten der Werbung für das Testprodukt oder für Wettbewerber getestet werden.

Normalerweise wurden folgende Meßwerte erhoben:

1. spontane Markenbekanntheit (nach Positionen),
2. Einstufung der generellen Produktqualität,
3. Zuordnung von relevanten Produkteigenschaften,
4. Ermittlung der Kaufbereitschaft (verbale Skala).

Diese Meßwerte entsprechen der von Politz so formulierten Aufgabe:

"Our interest should be directed to the ability of advertising to make more consumers aware of the brand, to persuade more consumers of the brand's merits, and to make more consumers susceptible to trying the brand."
(McCall's, 1962).

Wie erkennbar, werden mit diesem Verfahren Veränderungen zwischen "vorher" und "nachher" gemessen. Und diese Veränderungen werden indirekt ermittelt, das heißt, die Testteilnehmer wissen gar nicht, daß sie Teil eines (kontrollierten) Experiments sind.

Insofern unterscheidet sich dieses Verfahren von anderen:

- a) Es gab ("Schwerin-Test") und gibt ("Advantage") Verfahren, in denen die Veränderungen bei ein und denselben Testteilnehmern durch Vorher-nachher-Befragung ermittelt werden; dort wissen die Testteilnehmer also, worum es geht.
- b) Die meisten anderen Verfahren - der von uns verwendete "CAT" eingeschlossen - ermitteln nur die direkten Reaktionen zu Werbemittel und beworbener Marke. Hierbei kann man nur Korrelationen ermitteln, keine Beziehung zwischen Ursache und Wirkung.

Diese Unterschiede muß man kennen, wenn man die Ergebnisse unterschiedlicher Verfahrensweisen interpretiert.

Es gibt neuerdings eine Darstellung der Politz'schen Methode als reiner Werbemittel-Test (*Hardy 1990*). Das Verfahren hat er in mindestens 3 Media-Untersuchungen verwendet (*Curtis 1960, McCall's 1962, Readers' Digest 1966*). Und seine frühere Mitarbeiterin Shirley Young hat (*1972*) in einem lesenswerten Aufsatz eine Abwandlung des Verfahrens vorgestellt, dazu andere Verfahren kritisch gewürdigt.

Als Essenz seiner Werbemittelforschung hatte Alfred Politz 1955 formuliert:

"Subjective importance of your claim is most important. Then your claim must be believable (...). Then comes the third dimension in advertising: uniqueness.

It can be shown that the relation between the three terms should be multiplicative and not additive."

Bei der Bewertung dieser Aussagen ist immer zu bedenken, daß die Grundlagen seines Tests die Bewertung der beworbenen Marke durch die Verbraucher ist, nicht primär die Bewertung des Werbemittels (*s.a. Schaefer, 1965*).

Aufgrund dieser Politz'schen Forschungsrichtung standen für uns immer Kauf- bzw. Probiertneigung und Produktbeurteilung im Vordergrund; und als wesentliche Einflußfaktoren die Wichtigkeit/persönliche Relevanz der Aussage sowie deren Überzeugungskraft. Beides traf für andere Forscher nicht oder nur teilweise zu. Die Folgen kann man in den nachfolgenden Beschreibungen erkennen.

Analysen von Wells (1964 und 1971)

Über diese Untersuchungen und die Anwendung der Verfahrensweise in Deutschland hat Koeppler (1974, ab S.85) ziemlich ausführlich berichtet, so daß wir uns kurz fassen können. Die erste Untersuchung von Wells (1964) lieferte als wesentliche Beurteilungskriterien für Werbemittel 3 Hauptfaktoren:

Tab. 1	Attractiveness	Attraktivität
	Meaningfulness	Bedeutsamkeit für den Befragten
	Vitality	Vitalität (d.h. Anzeige ist "lebendig, frisch, neuartig")

Im Auftrag der Heinrich Bauer Stiftung hat das DIVO-Institut eine entsprechende Untersuchung in Deutschland durchgeführt (*Bauer Stiftung, 1967*). Die Eigenschaften zur Werbemittel-Beurteilung mußten eingedeutscht und in einer Vorstudie ausprobiert werden. Die Hauptstudie lieferte dann ebenfalls 3 Faktoren, wie in Tabelle 2 zu erkennen. Dabei kommt die deutsche Analyse zu 2 vergleichbaren Faktoren und zu einem anderen. Da die Eigenschaftswerte in 2 Sprachen nie ganz übereinstimmen, überdies die Auswahl begrenzt war (9 Wortpaare in Deutschland), ist dies nicht verwunderlich.

Tab. 2	1.	Subjektive Bedeutung der Anzeige (bedeutungsvoll/überzeugend/interessant vs. nicht schön/nicht klar/nicht bedeutungsvoll)
	2.	Dynamik (frisch/bewegt vs. alt/ruhig)
	3.	Subjektive Klarheit (ruhig/überzeugend/klar vs. nicht seriös/nicht glaubwürdig/nicht klar)

Wells und Mitarbeiter haben ein paar Jahre später ihre Basis für Beurteilungskriterien in TV-Tests verbreitert, indem sie von 1.000 Wörtern ausgingen, die sie in mehreren Schritten sinnvoll reduzierten. In der letzten Stufe blieben 73 Eigenschaften übrig, welche alle Bereiche und Faktoren abdeckten, die sich als relevant erwiesen.

In der letzten Untersuchungsstufe kamen 6 Faktoren heraus, die praktisch immer, sowie 2 weitere, die gelegentlich auftraten. Nachfolgend die Beschreibung der Wells'schen Faktoren mit den jeweils am höchsten ladenden Eigenschaften (die deutschen Übersetzungen sind mögliche Übertragungen - manchmal passen sie nur ungenau oder schlecht):

Tab. 3

1.	Humor jolly merry playful humorous amusing	Humor lustig fröhlich spielerisch humorvoll amüsant
2.	Vigor vigorous exiting energetic enthusiastic exilerated	Kraft kraftvoll aufregend energisch enthusiastisch hochgemut
3.	Sensuousness tender gentle soothing serene lovely	Gefühl zart sanft beruhigend klar und ruhig lieblich
4.	Uniqueness novel imaginative ingenious unique	Einzigartigkeit neuartig phantasievoll einfallsreich einzigartig
5.	Personal Relevance important to me meaningful to me worth remembering valuable	persönlich relevant wichtig für mich bedeutungsvoll für mich erinnerungswert wertvoll
6.	Irritation terrible stupid ridiculous irritating phony	Verärgerung schrecklich dumm lächerlich irritierend unecht
7.	Familiar saw before well-known familiar	bekannt schon mal gesehen allgemein gut bekannt vertraut
8.	Professional confident professional consistent in style	professionell ihrer Sache sicher professionell einheitlicher Stil

Diskussion

Bei dieser Analyse wurde nicht der Versuch gemacht, die Bedeutung der Eigenschaften bzw. Faktoren für den Kauf oder die Kaufneigung zu ermitteln. Das wäre über das selbstgesteckte Ziel der Autoren hinausgegangen. Damit begrenzten die Autoren allerdings die Aussagekraft ihrer Ergebnisse, was auch an einigen von ihnen vorgeführten Beispielen erkennbar ist. Somit ist man auch nicht in der Lage zu verallgemeinern, auf welche Eigenschaften man besonders achten sollte.

Mit dem Mangel der zu erklärenden Variablen hängt unmittelbar zusammen, daß die "Kaufneigung" oder "Probiierneigung" nicht als ein Faktor unter anderen auftritt. In der Liste der 254 Eigenschaften sind uns 5 Eigenschaften aufgefallen, die einen solchen Faktor hätten bilden können:

Tab. 4	1. empfehlenswert
	2. für mich
	3. gibt mir den Wunsch, es zu kaufen
	4. prospektiver Kauf
	5. würde Produkt empfehlen

Nur die zweite Eigenschaft war für die späteren Faktoren-Analysen brauchbar, die anderen nicht. Ausgerechnet die dritte Eigenschaft war ein ganz schwacher Begriff, der schon die zweite Analysestufe nicht überlebte.

Es ist aus zeitlicher und räumlicher Distanz nur zu vermuten, daß diese Untersuchungsreihe auf den theoretischen Faktor "Kauf-/Probiierneigung" so wenig abgestellt war und daß man relevante Eigenschaften frühzeitig eliminierte, weil zu wenige Testteilnehmer zu wenige TV-Filme als kaufanregend betrachteten.

In diesem Zusammenhang ist unsere eigene Auffassung vorzutragen, daß einzelne Worte schlechter sind als kurze, prägnante Beschreibungen. In relativ vielen Fällen weiß man bei der Verwendung von einzelnen Worten hinterher nicht mehr genau genug, auf was sie sich beziehen. Ein typischer Fall ist der Begriff "einzigartig": die dargestellte Marke kann einzigartig sein, aber auch die Art des Films. - Was ist gemeint ?

Die Autoren beschreiben selbst, daß Werbefilme für Kosmetik beim Faktor "Gefühl" hohe Werte zu bekommen pflegen. Ist dies nun primär die (produktspezifische) Gestaltung des Werbemittels oder das Produkt selbst, das diesen Eindruck bewirkt - oder beides ?

Von der Eindeutigkeit der Beschreibungen hängen selbstverständlich die Filmbeurteilungen ab; und diese bestimmen die "Schärfe" der Ergebnisse der Faktorenanalyse.

Wells et al. hatten mit Polaritäten wie beim Semantischen Differential gearbeitet. Für die zukünftige Arbeit schlagen die Autoren selbst die Vorgabe der Eigenschaften ohne Gegenpol vor, dafür mit einer 5-Punkte-Verbalskala. Aufgrund eigener und fremder Erfahrungen halten auch wir die Polaritäten für wenig empfehlenswert:

- zu oft hat man keinen oder keinen echten Gegenpol
- oder gleich 2 oder 3 ähnliche Gegenpole.

Beides tritt bei der unipolaren Verwendung nicht als Problem auf. (Die Art der Skala scheint dagegen praktisch keine große Rolle zu spielen).

Überprüfung der Skalen von Wells und Leavitt

Zinkham und Fornell haben 1985 eine Überprüfung der Skalen von Wells (1964) und Leavitt (1970) durchgeführt (die Skala von Leavitt war uns unbekannt).

Diese Überprüfung ist aus mehreren Gründen interessant:

1. Im Test wurde nicht nur überprüft, ob die Skalen messen, was sie zu messen vorgeben, sondern die Ergebnisse wurden auch mit zwei "Wirkungsmeßwerten" in Beziehung gesetzt.
2. Ein Ergebnis sagt etwas aus über möglicherweise unterschiedliche Wirkungsweisen von Anzeigen und TV-Spots.

Die Skala von Wells, die bekanntlich eine relativ weite Verwendung und Nachahmung gefunden hat, bezieht sich auf Anzeigen. Sie umfaßte 25 Eigenschaftspaare (mit Gegenpol) und soll 3 hauptsächliche Faktoren liefern:

Tab. 5	1. Attraktivität
	2. Bedeutsamkeit
	3. Vitalität

Die Skala von Leavitt bezog sich auf Fernseh-Werbefilme. Sie hatte 27 Eigenschaften (wobei die Art der Abfrage und die Skalierung nicht beschrieben wurde). Sie soll 4 hauptsächliche Faktoren liefern:

Tab. 6	1. lebhaft & amüsant
	2. persönliche Relevanz
	3. gefühlvoll
	4. gut bekannt

Die Überprüfung erfolgte:

- a) für die Batterie von Wells ...
 - mit 20 Anzeigen,
 - jeweils ein Folder mit redaktionellem Material,
 - 20 Testteilnehmer pro Anzeige;
 - insgesamt also 400 Beurteilungen.

b) für die Batterie von Leavitt ...

- mit 20 TV-Werbefilmen,
- jeweils einer im normalen TV-Umfeld vorgeführt,
- 20 Testteilnehmer pro Film;
- d.h. wieder insgesamt 400 Beurteilungen.

Nach der Exposition des Werbemittels wurde ein Fragebogen mit den jeweiligen Items (in rotierter Reihenfolge) ausgefüllt.

Dann wurde ein zweites Interview durchgeführt, in dem zwei Dinge ermittelt wurden:

1. Einstellung zur beworbenen Marke, indem die Befragten auf einer 8-Punkte-Skala angeben sollten, wie gut/schlecht die Marke ist;
2. die Kaufbereitschaft, indem die Befragten um eine Schätzung gebeten wurde, wie groß die Wahrscheinlichkeit ist, daß sie die beworbene Marke in der nahen Zukunft kaufen werden.

Ergebnisse

a) Messen sie, was sie sollen?

Es wurden Faktorenanalysen mit "Target Rotationen" (ohne Erklärung) durchgeführt. Die Kongruenz-Koeffizienten (Übereinstimmung zwischen den Lösungen und dem Ziel) waren unterschiedlich gut:

1. Wells:	CC = .93
2. Leavitt:	CC = .58

Die Korrelationen zwischen den Einzelitems und den Faktoren waren bei Wells hoch bis sehr hoch; mit wenigen positiven und negativen Abweichungen. Bei Leavitt sind die Korrelationen deutlich schlechter, außer beim ersten Faktor. Ferner gab es zahlreiche "falsche" Items, die mit einem Faktor höher korrelierten als die "richtigen" Items.

b) Beziehung zu Einstellung und Kaufneigung?

Die Korrelationen der Skalen zur generellen Einschätzung der Marke waren gleichartig mittelmäßig:

1. Wells:	r= .57
2. Leavitt:	r= .56

Die Korrelationen mit der Kaufneigung waren bei Wells gut, bei Leavitt mittelmäßig:

1. Wells:	r= .75
2. Leavitt:	r= .52

Aufgrund dieser Ergebnisse ist die Skala von Wells als eindeutig überlegen anzusehen.

Interessant ist das Nebenergebnis betreffend der Beziehung zwischen den Meßwerten "generelle Einschätzung der Marke" und "Kaufbereitschaft":

1. bei Wells (Anzeigen)	r= .57
2. bei Leavitt (TV-Filme)	r= .44

Die Autoren schreiben zum Befund, daß die Einstellung als Vorhersagemessung für die Kaufbereitschaft bei gedruckten Medien besser ist als bei TV: Dieses Ergebnis sei konsistent mit Krugman's (1967) Theorie, daß TV-Werbung "low involvement" sei im Gegensatz zu Zeitschriften-Werbung mit "higher involvement".

Wir würden dieses Ergebnis eher mit einer späteren Arbeit von Krugman (1977) in Verbindung bringen:

- o Anzeigen sprechen eher die linke Gehirnhälfte an; diese ist stärker auf rationale, verbalisierte/analytische Verarbeitung abgestellt; das führt zu höherer Konsistenz von Einstellungen und Verhaltensweisen.
- o TV-Werbung spricht eher die rechte Gehirnhälfte an; diese verarbeitet eher nonverbal/holistisch; Verhaltensweisen zeigen schwächere Beziehungen zu Einstellungen.

Moldovan's Analyse

Stanley Moldovan, Leiter der Kreativforschung bei SSC & B/ Lintas, N.Y., hat 1984 eine ziemlich umfangreiche Sekundäranalyse von Hunderten von TV-Werbefilm-Tests vorgelegt. An anderer Stelle (*VJH 1985*) ist darüber berichtet worden, aber dies soll in etwas detaillierterer Weise nachfolgend nochmals geschehen. Denn erst dann, wenn man die Einzelheiten kennt, kann man ggf. für seine eigene Arbeit Konsequenzen ziehen:

Methode

Stichproben von Angehörigen der jeweiligen Hauptzielgruppe ("key target group") wird der Werbefilm einzeln vorgeführt; Kleine Quota-Stichproben unbekanntem Umfangs; persönliche Interviews.

Produktfelder

Folgende Produktfelder wurden vom Autor angegeben:

- o Haushalts-Produkte,
- o Wegwerfartikel für Kinderpflege,
- o Erkältungsmittel,
- o Körperpflegeprodukte (getrennt für Frauen und Männer),
- o Zahnpasta,
- o alkoholfreie Getränke,
- o Tiernahrung.

Datenverarbeitung

Moldovan hat zuerst eine Faktorenanalyse durchführen lassen, dann eine multiple Regressionsanalyse, wobei er die 5 Hauptfaktoren als erklärende Variablen für den Faktor "Persuasion" (besser: Kaufneigung) nahm.

Ergebnisse

Die Faktorenanalyse lieferte 6 Faktoren, für die in Tabelle 7 die Items mit den höchsten Ladungen genannt werden.

Tab. 7

Faktor 1: Kaufneigung

- o würde Marke (bestimmt) kaufen
- o würde gerne Gutscheine nehmen
- o besser als andere Marken
- o würde sie ausprobieren

Faktor 2: Glaubwürdigkeit

- o realistisch
- o glaubwürdig
- o vernünftig
- o offen/ehrlich
- o wie auch ich mich fühle
- o überzeugend

Faktor 3: Guter Geschmack

- o geschmackvoll
- o nicht abstoßend
- o erfreulich anzusehen
- o paßt zur Produktkategorie
- o nichts abgelehnt
- o mag Darsteller

Faktor 4: Persönlich angesprochen

- o richtig Spaß gemacht
- o gerne noch einmal sehen
- o sehr persönlich gehalten
- o Gefühl derselben Erfahrungen
- o Film erregte meine Aufmerksamkeit
- o spricht Leute wie mich an

Faktor 5: Anregend

- o bewegt /lebhaft
- o kurz und zur Sache
- o meine Aufmerksamkeit gefesselt
- o interessant

Faktor 6: Klarheit

- o gut dargestellt, um was es geht
- o klar
- o informativ
- o zeigte, was das Produkt anzubieten hat

Die Korrelationen zwischen Kaufneigung und den anderen Faktoren waren sehr hoch:

Tab. 8

	Korrelation mit "Kaufneigung" r =
o Glaubwürdigkeit	.85
o Guter Geschmack	.80
o Persönlich angesprochen	.78
o Anregend	.75
o Klarheit	.74

Bei der schrittweisen multiplen Regressionsanalyse zeigte sich, daß der Faktor "Glaubwürdigkeit" der entscheidende war und die anderen nur wenig mehr zur Erklärung der Varianz der Kaufneigung beitrugen:

Tab. 9	
	erkl. Varianz
o Glaubwürdigkeit	72%
o Glaubwürdigkeit + anregend	73%
o Glaubwürdigkeit + anregend + Klarheit	74%
o Glaubwürdigkeit + anregend + Klarheit + guter Geschmack	75%

Moldovan schreibt im Text, daß die entscheidende Bedeutung der "Glaubwürdigkeit" quer durch alle untersuchten Produktkategorien gehe. Er erwähnt ferner, daß man unterschiedliche Einflüsse der anderen Faktoren erhalte, wenn man die Glaubwürdigkeit innerhalb der Produktgruppen konstant hält; und dann schreibt er "That what is important sells" (was wichtig ist, das ist verkaufswirksam). Leider ist der Passus (absichtlich?) undeutlich, so daß nicht erkennbar ist, worauf er sich bezieht. Einen Faktor "Wichtigkeit" hat er nicht; das Item "Botschaft hatte persönliche Bedeutung" (= personal relevance) kam nirgendwo hoch.

Was bei der Moldovanschen Analyse fehlt, ist das Ergebnis einer multiplen Regressionsanalyse, welche eine Gewichtung der einzelnen Faktoren liefern würde.

Olson's Analyse

David Olson ist Forschungsgruppenleiter der Werbeagentur Leo Burnett, Chicago. Er hat zwei Berichte über die Wirkungsfaktoren von Werbemitteln veröffentlicht (*Olson 1984 und 1985*). Im 84er ESOMAR-Bericht skizzierte er auch einige Versuche anderer Forscher zur Validierung von Werbemitteltests. In einem früheren Aufsatz (*1982*) hatten er und seine Mitarbeiter Vergleiche der Werbemittel für neue und bestehende Marken angestellt.

Externes Kriterium

Olson hat seine neuesten Analysen ausschließlich auf neue Marken bezogen. Es ist, nach seinen Worten, extrem schwierig, im Markt ein objektives Maß der Verkaufswirksamkeit von Werbemitteln für etablierte Marken zu finden. Aber bei der neuen Marke kann man die Anzahl der "Erstkäufer" ("trial") als Maß der Werbewirksamkeit nehmen.

Die Bedeutung anderer Faktoren (Distribution, Preis etc.) war ihm klar; er hat sich bemüht, sie einigermaßen konstant zu halten. Das Maß "Erstkauf" bezog er aus Tracking Studies.

Gruppierung von Werbefilmen

Für die neu eingeführten Marken waren Tests der benutzten Werbemittel durchgeführt worden. Diese Werbemittel (man bekam 65 zusammen) konnten aufgrund der Erstkaufraten in 3 Gruppen eingeteilt werden: hohe, mittlere und niedrige Erstkaufrate.

Werbemitteltest-Eigenschaften

Mary Jane Schlinger hatte für Burnett eine Batterie von 52 Eigenschaften entwickelt (*Schlinger 1979*). Und eine Faktoren-Analyse hatte ergeben, daß diese Eigenschaften 7 grundlegende Dimensionen für die Werbung für neue Marken liefern (siehe Tabelle 14) mit deutschen Übersetzungen. Die Bezeichnungen und Übersetzungen beinhalten wieder die üblichen Probleme.

Der Faktor 7 ist die andere Seite der Medaille von "Uniqueness", hier also der Mangel daran.

Die Statements und Faktoren kommen einem mehr oder minder bekannt vor. Das deutet auf Grundstrukturen von Sprachen(n), Denken und Wirkungsweisen hin. Dies nährt die Hoffnung, demnächst zu einem gewissen Konsens zu gelangen.

Tab. 10

1. News

(How informative are the ads seen as being?)

1. Neuigkeit

(Als wie informativ werden die Werbemittel betrachtet?)

2. Relevance

(How personally meaningful is the information in the ad?)

2. Relevanz

(Wie bedeutsam ist die Information des Werbemittels für einen selbst?)

3. Purchase interest

3. Kaufinteresse

4. Stimulation

(How entertaining or interesting is the ad itself?)

4. Anregung

(Wie unterhaltend oder interessant ist das Werbemittel selbst?)

5. Empathy

(How involving is the commercial?)

5. Gefühl

(Wie stark involviert das Werbemittel den Betrachter?)

6. Confusion

6. Konfusion

7. Familiarity

(How much is the ad felt to be just like all other ads on television?)

7. Gewöhnlichkeit

(In welchem Maß wird das Werbemittel als genauso wie alle andere Werbung im Fernsehen angesehen?)

Wirkungsfaktoren

Zunächst hatte man die Mittelwerte für die 3 Gruppen von Werbemitteln ausgezählt; zur allgemeinen Freude und zum allgemeinen Erstaunen fand man heraus, daß die Werbemittel mit hoher Erstkauftrate durchweg für alle 7 Faktoren höhere Werte als die beiden anderen Gruppen aufwiesen.

Bereits die einzelnen Items zeigten diese Tendenz: 40 der 52 Items lieferten signifikante Abweichungen, und alle in der "richtigen" Richtung.

Im nächsten Schritt haben Olson und Mitarbeiter sich bemüht, die relative Bedeutung der einzelnen Faktoren herauszufinden. Sie verwendeten die Diskriminanz-Analyse, um zu erkennen, welche Kombinationen von Faktoren die Einstufung der einzelnen Fälle in die Kategorien hohe, mittlere oder niedrigere Erstkauftrate prognostizieren. Zwei Variablen lieferten 66% korrekte Prognosen; alle 7 Faktoren schafften etwas mehr, nämlich 75%.

Der erste Faktor der o. g. beiden Variablen, so Olson, war nicht sehr überraschend. Den zweiten Faktor nennt Olson "überraschend"; er bezieht sich auf die Gestaltung der Werbefilme.

Tab. 11

1. Relevanz

- Das Werbemittel zeigte mir, daß das Produkt gewisse Vorzüge hat.
- Das Produkt ist für mich wichtig.
- Das Werbemittel erinnert mich daran, daß ich mit dem, was ich jetzt verwende, unzufrieden bin und nach etwas anderem suche.
- Während des Werbefilms dachte ich daran, wie das Produkt für mich nützlich sein könnte.
- Das Werbemittel gab mir das Gefühl, daß das Produkt für mich richtig ist.

2. Anregung

- Die Personen im Werbemittel gewinnen meine Aufmerksamkeit.
- Der Enthusiasmus des Werbemittels ist ansteckend.
- Das Werbemittel ist amüsant.
- Das Werbemittel ist spielerisch.
- Es macht großen Spaß, dem Werbemittel zuzusehen und zuzuhören.
- Ich fand ihn clever und ziemlich unterhaltsam.

Olson beschreibt anschließend anschaulich, wie aufregend diese Ergebnisse den Forschern erschienen - ja, und sie dachten, sie müßten auch aufregend und interessant sein für die Kreativen in der Agentur. Einer von diesen faßte nach einer gedankenvollen Pause zusammen:

"Also, was Sie uns sagen ist, daß wir eine gute Strategie und auch eine gute Gestaltung haben sollten. Stimmt's ?"

An dieser Stelle müssen wir zugeben, daß dies die typische Situation des Werbeforschers vis-à-vis den Kreativen ist. Was unsereins für einen wesentlichen Schritt in der menschlichen Erkenntnis hält, betrachten unsere Partner im günstigsten Fall als die (wissenschaftliche) Elaboration des Offensichtlichen.

Die Burnett-Forscher sind daraufhin einen Schritt weiter gegangen: sie haben versucht, unter ihren Tests solche zu finden, wo für ein und dieselbe Marke zwei verschiedene Werbemittel produziert worden waren. Sie fanden zehn Paare. Diese haben sie informellen Inhaltsanalysen unterzogen und folgende, noch vage Andeutungen herausgelesen:

- o Falls das Produkt sich klar von Konkurrenten unterscheidet oder ihnen überlegen ist, werde im besseren Werbemittel dieser Sachverhalt kommuniziert werden.
- o Um die Aufgabe wirklich gut zu erfüllen, ist es notwendig, nicht nur den Unterschied zu erläutern. Die wirklich überragenden Werbemittel haben es auf die eine oder andere Weise verstanden, das Produktkonzept zu "dramatisieren".
- o In einigen Fällen schien es wichtig gewesen zu sein, daß die Zuschauer sich mit der gezeigten Situation oder den Darstellern identifizieren konnten.
- o Einige Werbefilme scheinen dadurch wirksam(er) zu sein, daß sie der möglichen Skepsis gegenüber den angepriesenen Produktbenefits direkt entgegentraten.

Diese und einige weitere Anmerkungen von Olson weisen auf die Notwendigkeit hin, Inhalt und Gestaltung von Werbemitteln nicht nur von den Befragten beurteilen zu lassen, sondern einer Contentanalyse zu unterwerfen. Die Kategorien einer solchen Analyse müssen sich an dem Denken und Fühlen der Kreativen orientieren, also an einer für den Marktforscher oft ziemlich fremden Welt.

Eigene Analyse *

1985 haben wir eine Reihe von statistischen Analysen mit dem Test *CAT* ("Competitive Advertising Test") durchgeführt.

Dieser Test wurde 1972 von Frau U. Agrell, Abt. Market Research der Colgate-Palmolive GmbH, Hamburg, entwickelt und seit 1976 zum Testen von 53 Werbemitteln bei der Schaefer Marktforschung eingesetzt. Im Laufe der Zeit ist er in gemeinsamer Arbeit erweitert worden, insbesondere in den letzten 2 Jahren. Der Beitrag von Marianne Drehmann auf dieser Tagung bezieht sich hierauf (*Drehmann 1985*).

Der Test läuft im Prinzip wie folgt ab:

1. Vorführung des Test-Films in einem Block mit Filmen für ganz andere Produkte und Marken (d.h. "forced exposure");
2. persönliches Interview:
 - o spontan erinnerte Werbung mit Markennennung;
 - o skalierte Messung, in welchem Maße bestimmte Produkteigenschaften wahrgenommen und dem Testprodukt zugeordnet werden;
 - o Beurteilung des Werbemittels anhand spezieller Beurteilungskriterien;
 - o Ermittlung der Kaufbereitschaft;
 - o "offene" Fragen zur Beurteilung des Werbemittels.

Für diesen Test gibt es keine absoluten Maßstäbe, sondern die Orientierung und Bewertung der Ergebnisse erfolgt durch statistische Vergleiche:

- eigenes Werbemittel vs. Konkurrent(en),
- oder: neues Werbemittel vs. bisheriges.

Da es sich um einen monadischen Test handelt, erfolgt der Vergleich erst in der Auswertungsphase durch Gegenüberstellung der Ergebnisse aus strukturgleichen Testgruppen.

Als Zielgruppe werden normalerweise Verwender(innen) der betr. Produktkategorie gezählt; manchmal werden sie begrenzt auf Verwender bestimmter Marken.

Die Stichprobe umfaßt in der Regel 100-150 Testteilnehmer(innen) pro Testgruppe. Die Testgruppen werden durch Quotenvorgabe, nachträgliche Quotenkontrolle und ggf. durch Gewichtung strukturgleich gehalten. Die ausgewählten Testteilnehmer(innen) werden durch die Interviewer(innen) in unsere in Einkaufsstraßen und vor Verbrauchermärkten plazierten (mobilen) Teststudios eingeladen.

*— den beteiligten Damen und Herren der Marktforschungs-Abteilungen von CP, MWI, Ted Bates und wird für die konstruktive Zusammenarbeit bei den Tests und deren Auswertung gedankt.

Beurteilung von vermittelter Produkt-Information

Im Mittelpunkt des Interviews steht einmal die Frage, ob die beabsichtigte Vorstellung von der Marke vom Werbemittel kommuniziert wird. Dementsprechend wird skaliert die Zuordnung von Produkteigenschaften erhoben. Am Beispiel von Deo-Mitteln z.B. bezüglich folgender Eigenschaften:

- o ist ein besonders mildes Deo
- o ist hautverträglich
- o hat eine langanhaltende desodorierende Wirkung
- o man kann sich auf seine Wirkung verlassen
- o hat einen angenehmen Duft
- o ist besonders wirksam gegen Körpergeruch
- o ist ein Deo, das meinen Ansprüchen genügt
- o ist genau das richtige Deo für meine Haut
- o ist etwas Besonderes
- o ist von hoher Qualität
- o etc.

Die Beurteilung erfolgt anhand einer 7er-Kästchen-Skala (1= niedrig; ... 7= hoch). _

Beurteilung der Gestaltung des Werbemittels

Zum anderen wird die Frage untersucht, ob das zu testende Werbemittel als Transporteur der Information geeignet ist. Hierzu haben wir seit 1979 eine standardisierte Itemliste verwendet:

- o Der Film sagt etwas aus, was für mich wichtig ist.

- o Der Film ist glaubwürdig.
- o Die Darstellung des Films ist überzeugend.
- o Was im Film gesagt wurde, ist überzeugend.
- o Aufgrund des Films würde ich ... probieren.

- o Der Film ist anders als andere ...-Werbung.
- o Der Film ist originell.
- o Der Film ist auffällig.
- o Der Film sagt mir etwas Neues über diese Marke.
- o Der Film vermittelt den Eindruck, daß die Marke ganz anders ist als andere ...-Marken.

- o Der Film informiert gut.
- o Der Film ist einprägsam.
- o Der Film ist verständlich.

- o Der Film ist sympathisch.

Die Reihenfolge der Items wurde hier wegen der besseren Übersichtlichkeit sortiert; im Fragebogen werden sie gemischt und in zwei rotierten Versionen aufgeführt.

Analyse der Ergebnisse für 53 Werbemittel

Die 1985 vorgenommene Sekundär-Analyse hatte folgende Ziele:

1. den Zusammenhang zwischen den bisher in den CATs gewählten Beurteilungs-Kriterien und der Kaufbereitschaft zu überprüfen;
2. ggf. Hinweise zu erhalten, wie diese Liste ergänzt oder verändert werden kann oder muß.
3. Daraus sollten, soweit möglich, Verallgemeinerungen hinsichtlich der Einflußfaktoren für die Kaufbereitschaft abgeleitet werden;
4. damit sollten Maßstäbe für die Bewertung einzelner Werbemittel gewonnen werden;
5. und es sollte versucht werden, den Kreativen Hilfestellungen für die Entwicklung wirksamerer Werbemittel an die Hand zu geben (von einer möglichen "Checkliste" bis zu konkreten Hinweisen).

Ein erstes Beispiel:

Nachfolgend ist in neutralisierter Form das Ergebnis eines CATs mit drei TV-Spots dargestellt. Zunächst ist dargestellt, inwieweit die drei Spots in der Lage waren, wesentliche Aussagen über das Testprodukt zu vermitteln. Es handelt sich um Mittelwerte der o.a. 7er-Skalen:

Tab. 12

Produkt Eigenschaft	Spot A	Spot B	Spot C
A	5.4*	5.4*	5.1
B	4.9*	4.9*	4.4
C	4.1	4.8	4.1
D	4.5	4.6*	4.2
E	4.8*	4.8*	4.4
F	5.0	5.1*	4.8
G	4.1	4.3*	3.8
H	4.6	4.7*	4.5
I	5.0	5.5*	5.3
K	4.8*	4.8*	4.6
Kaufbereitschaft	3.7*	3.4	3.3

—
* = relativ gut/besser

Spot A zeichnet sich dadurch aus, daß er insbesondere den Eindruck vermittelt, daß diese Marke die Eigenschaften (A), (B), (E) und (K) hat.

Spot B vermittelt den Eindruck, daß die Marke die Eigenschaften (B), (C) und (D) hat - und auch sonst für alle anderen Probleme recht gut geeignet erscheint.

Spot C erreicht bei kaum einem Item das Niveau der beiden anderen Spots.

Die Frage, die sich aus diesen Zahlen unmittelbar ergibt, lautet: Warum schlägt das durchweg gute Abschneiden des Spots B hinsichtlich dessen, was die Marke alles leistet, nicht auf die Kaufbereitschaft durch? Warum schneidet der Spot A bei der Kaufbereitschaft wesentlich besser ab ?

Und für das Marketing stellt sich die Frage: welchen TV-Spot soll man senden:

- den mit der höheren Kaufbereitschaft oder
- den, der die wichtigen Aussagen besser vermittelt?

Diese Ergebnisse allein liefern dazu keine sehr weitgehende Entscheidungshilfe. So wenden wir uns zunächst einmal dem zweiten Item-Satz zu:

	Spot A	Spot B	Spot C
o der Film ist glaubwürdig	4.3*	3.6 ---	--- 3.9
o Darstellung ist überzeugend	4.3*	3.6 ---	--- 3.9
o was gesagt wurde, ist überzeugend	4.3*	3.8	--- 4.0
o sagt etwas Wichtiges	4.0*	3.5	3.5
o informiert gut	4.9*	3.9 ---	--- 4.3
o ist verständlich	5.8*	5.2	5.4
o ist einprägsam	5.0*	4.6	4.4
o anders als andere Werbung	3.7	3.6	3.6
o ist originell	3.9	4.0	4.0
o ist auffällig	4.6	4.5	4.2
o sagt Neues über die Marke	3.7*	3.5	3.3
o vermittelt den Eindruck, die Marke ist anders	3.3*	3.2+	2.9
o ist sympathisch	4.5*	3.9	4.0
o Kaufbereitschaft	3.7*	3.4	3.3

* = relativ am besten / + = B besser als C / --- = C besser als B

Hier hat sich das Bild gewandelt: jetzt schneidet Spot A durchweg am besten ab; und dies korrespondiert mit der erzielten Kaufbereitschaft.

Spot B zeichnet sich durch Auffälligkeit aus, Spot C informiert besser als B und ist glaubwürdiger.

Aus diesen beiden Tabellen und den Einstufungen der 3 Spots lassen sich nun gewisse Hypothesen über den Einfluß der verschiedenen Beurteilungsmeßwerte auf die kaufanreiz-auslösende Wirkung der Werbespots ableiten:

- o Die Reaktion auf die Filmgestaltung hat eine stärkere Beziehung zur Kaufbereitschaft als die Produkteigenschaften.

Dies wird bei Spot B besonders deutlich: Hier haben die Testteilnehmerinnen durchaus verstanden, was der Film über das Produkt aussagen sollte - nur glaubwürdig war das nicht (mit 3.6 der niedrigste Wert).

- o Die Glaubwürdigkeit scheint damit ein wesentlicher Faktor zu sein.
- o Aus dem Vergleich der Spots B und C läßt sich ableiten, daß die Uniqueness des Spots B mehr zur Kaufbereitschaft beiträgt als der verständlich-informative Charakter von Spot C.

Berechnung des Einflusses der Zustimmung zum Film und zu seiner kommunikativen Leistung auf die Kaufbereitschaft

Nach dieser mehr heuristischen Betrachtung der Test-Ergebnisse haben wir versucht, diese Hypothesen zu überprüfen. Wir sind auf verschiedene Weisen vorgegangen, um den Einfluß einzelner Beurteilungs-Kriterien auf die Kaufbereitschaft zu berechnen:

a) durch einfache Kreuzauswertung

Hierzu haben wir die Kaufbereitschaft derjenigen Testteilnehmerinnen ausgewiesen, die bei der Beurteilung der Spots oder bei der Zuordnung von Eigenschaften zum Produkt "top boxes" (5-7) angekreuzt haben, also der jeweiligen Eigenschaft in hohem Maße zustimmten.

b) durch Korrelationsanalyse

Hiermit wird der Zusammenhang zwischen Zustimmung zu Einzel-Items und der Kaufbereitschaft ermittelt.

c) durch multiple Regressionsanalyse

Hier wird der verbundene Einfluß aller Beurteilungskriterien auf die Kaufbereitschaft gemessen und dargestellt.

Dabei haben wir mehrere getrennte Analysen durchgeführt:

- Film-Beurteilungskriterien allein;
- die Zuordnung von Produkt-Eigenschaften allein;
- alle Film- und Produkt-Beurteilungskriterien zusammen.

ad a)

Messung des Einflusses hoher Zustimmung zu einzelnen Items auf die Kaufbereitschaft

Wenn man diejenigen Testteilnehmerinnen aus den Tests herausfiltert, die dem jeweiligen Item in besonderes hohem Maße zugestimmt haben, kann man deren Kaufbereitschaft auszählen. Je höher diese ist, desto größer ist der Einfluß, den die Zustimmung zu diesem Item auf die Kaufbereitschaft hat.

Können einem Item viele/fast alle Testteilnehmerinnen zustimmen, liegt der zu erzielende Kaufbereitschaftswert naturgemäß nahe am Durchschnitt der gesamten Testgruppe. Das heißt, daß das Item nicht zwischen kaufgeneigten und weniger kaufgeneigten Personen trennen kann - und damit einen nur geringen Einfluß auf die Kaufbereitschaft hat.

Neben der Kenntnis von den Kaufbereitschaftswerten, die, in diesem Sinne berechnet, der folgenden Tabelle zugrundeliegen, ist es also wichtig zu wissen, wie viele Testteilnehmerinnen dem jeweiligen Item zustimmten. Wir haben diese Anteile vor die Kaufbereitschafts-Mittelwerte gesetzt:

Tab. 14

Einfluß hoher Zustimmung zu einzelnen Film-Items auf das Niveau der Kaufbereitschaft

Zustimmung auf den Stufen 5-7 zu den Items ...	Zu- stimmung in %	Spot A	Spot B	Spot C
o der Film ist glaubwürdig	37-45	5.5	5.1	4.7
o Darstellung überzeugend	40-49	5.2	5.2	4.7
o was gesagt wurde, ist überzeugend	41-51	5.2	5.1	4.4
o sagt etwas Wichtiges	32-46	4.9	5.4	5.1
o informiert gut	38-65	4.7	5.1	4.1
o ist verständlich	66-81	4.2	4.0	3.7
o ist einprägsam	51-67	4.5	4.1	4.3
o anders als andere Werbung	36-39	4.8	4.5	4.6
o ist originell	43-48	5.0	4.3	4.1
o ist auffällig	45-55	4.5	4.2	4.5
o sagt Neues über ...	33-41	4.8	5.1	4.5
o vermittelt den Eindruck, die Marke ist anders	23-34	5.2	5.2	5.2
o ist sympathisch	46-57	5.1	4.6	4.5

Wir sehen, daß Eigenschaften wie

- o "sagt etwas, was für mich wichtig ist"
- o oder: "ist glaubwürdig/überzeugend",

sofern ihnen hinsichtlich eines Filmes zugestimmt wird, einen höheren Einfluß auf die Kaufbereitschaft haben als z.B.

- o "auffällig" und "originell"
- o oder: "verständlich" und "einprägsam."

Es wird auch deutlich, daß gerade der Spot B, der die Informationen des Films so gut 'rüberbrachte, im Gegensatz zu Spot A hinsichtlich der wesentlichen Items viel zu wenig Zustimmung erhielt.

Und Spot C mangelte es an "Auffälligkeit" und der Fähigkeit, die Marke als "anders" darzustellen. Wären hier bessere Zustimmungswerte erzielt worden, hätte dieser Film möglicherweise besser abschneiden können. Seine Stärke lag bei der "Verständlichkeit" - mit 3.7 war es das am wenigsten diskriminierende Item. So etwas "erwartet" man eben von der Werbung.

Auf die gleiche Weise ist in Tabelle 15 der Einfluß guter Produkt-Bewertung auf die Kaufbereitschaft dargestellt:

Tab. 15					
Einfluß hoher Zustimmung zu einzelnen Produkt-Items auf das Niveau der Kaufbereitschaft					
Zustimmung auf den Stufen 5-7 zu den Items ...		Zu- stim- mung in %	Spot A	Spot B	Spot C
			liefert eine Kaufbereitschaft von ...		
o der Film ist glaubwürdig		37-45	5.5	5.1	4.7
Produkt-Eigenschaft	A	66-77	4.1	3.7	4.0
	B	57-65	4.2	4.0	3.9
	C	47-63	3.9	3.8	3.9
	D	48-57	4.2	3.7	4.9
	E	52-63	4.4	3.9	4.2
	F	62-68	4.2	4.1	4.0
	G	49-62	4.8	4.2	4.8
	H	37-50	4.3	4.0	4.3
	I	59-78	4.2	3.7	3.7
	K	57-62	4.3	3.8	3.9

Man sieht, daß der Einfluß der Produkt-Bewertung auf die Kaufbereitschaft geringer ist als jener der Film-Bewertung.

Und interessant ist, daß diejenigen Items, die das relativ größte Gewicht beim Zustandekommen der Kaufbereitschaft haben, solche sind, die dem Produkt generell zugeschrieben werden - und in der Werbung nicht explizit herausgestellt werden müssen. Es handelt sich um die Items F und G, die etwa aussagen, daß ein Shampoo die Haare gründlich wäscht oder daß ein Putzmittel gründlich säubert.

Wenn die Testteilnehmerinnen also - warum auch immer - glauben, daß eine Marke ihren Grundnutzen erfüllt, würden sie diese auch für den Kauf in Betracht ziehen, und zwar unabhängig von der Werbung, die sie gerade gesehen haben.

ad b)
Korrelations-Analysen

Neben diesen "High-Scorer"-Analysen haben wir die Korrelationen zwischen den Einzel-Items und der Kaufbereitschaft berechnet.

Es zeigt sich das gleiche Bild wie vorher; die Filmbeurteilungs-Kriterien korrelieren stärker mit der Kaufbereitschaft als die Beurteilungen der Produktleistung, wie in Tabellen 16 und 17 dargestellt:

	Spot A r=	Spot B r=	Spot C r=
o der Film ist glaubwürdig	.81	.76	.68
o Darstellung ist überzeugend	.73	.77	.67
o was gesagt wurde, ist überzeugend	.69	.75	.69
o sagt etwas Wichtiges	.61	.71	.76
o informiert gut	.77	.73	.51
o ist verständlich	.50	.50	.58
o ist einprägsam	.53	.39	.41
o anders als andere Werbung	.39	.47	.57
o ist originell	.66	.46	.51
o ist auffällig	.45	.52	.64
o sagt Neues über die Marke	.49	.67	.53
o vermittelt den Eindruck, die Marke ist anders	.57	.68	.64
o ist sympathisch	.75	.65	.60

Tab. 17

Korrelationen zwischen den Produkt-Eigenschaften
und der Kaufbereitschaft

	Spot A	Spot B	Spot C
	r=	r=	r=
Produkt-Eigenschaft A	.30	.27	.48
B	.30	.40	.36
C	.18	.30	.34
D	.31	.24	.46
E	.38	.42	.50
F	.36	.45	.52
G	.49	.47	.53
H	.52	.48	.62
I	.26	.26	.39
K	.25	.27	.43

Und auch eine multiple Korrelation bestätigt, was wir als Hypothese formuliert haben: die Wirkung des Films an sich ist für die kaufanreiz-auslösende Wirkung entscheidender als die vermittelte Vorstellung konkreter Produktleistungen:

Tab. 18

Multiple Korrelation zwischen
Eigenschaften und Kaufbereitschaft

	Spot A	Spot B	Spot C
	r=	r=	r=
Film-Eigenschaften vs. Kaufbereitschaft	.89	.88	.87
Produkt-Eigenschaften vs. Kaufbereitschaft	.62	.65	.71

ad c)

Regressions-Analysen

Mit Hilfe einer multiplen Regressions-Analyse kann man die Gewichte einzelner Beurteilungskriterien beim Zustandekommen der Kaufbereitschaft berechnen.

Man erhält dazu eine Formel der Art:

$$KB = a \times \text{Item1} + b \times \text{Item2} + c \times \text{Item3} + \dots + \text{Konstante}$$

KB ist die Kaufbereitschaft und die Regressionskoeffizienten a, b, c, ... geben das Gewicht an, das dem jeweiligen Item beim Zustandekommen der Kaufbereitschaft zukommt.

Da es eine Reihe von Items gibt, deren Bedeutungen redundant sind, korrelieren diese untereinander; und die Effekte beim Zustandekommen der Kaufbereitschaft überlagern sich.

Deshalb wurde zunächst eine Faktoren-Analyse durchgeführt, um von einander unabhängige Beurteilungs-Faktoren zu erhalten. Mit diesen Faktoren wurde die anschließend beschriebene multiple Regressions-Analyse berechnet.

In die Faktoren-Analyse wurden alle Items außer "Kaufbereitschaft" und "Proberneigung" einbezogen; denn diese sollten ja später als zu erklärende Variablen dienen. Außerdem wurden die Beurteilungen aller drei Spots zu einer Datei zusammengefaßt (n=292).

Die Faktoren-Analyse wurde nach Extraktion von 4 Faktoren abgebrochen; die Eigenwerte betragen:

- für Faktor 1	7.55
- " " 2	1.05
- " " 3	.78
- " " 4	.72

Diese 4-Faktoren-Lösung erklärte 77.9 % der Varianz, die sich wie folgt verteilte:

- Faktor 1	31.3 %
- " 2	20.4 %
- " 3	14.7 %
- " 4	11.4 %

Die Faktoren werden durch ihre Korrelationen zu den Ausgangs-Items charakterisiert. Diese Korrelationen zeigt die Ladungsmatrix in Tabelle 19:

Tab. 19

	Faktor 1	Faktor 2	Faktor 3	Faktor 4
o der Film ist glaubwürdig	.85	.25	.18	.20
o Darstellung ist überzeugend	.81	.32	.23	.20
o was gesagt wurde, ist überzeugend	.80	.31	.26	.14
o sagt etwas Wichtiges	.65	.25	.48	.14
o informiert gut	.69	.13	.30	.39
o ist verständlich	.31	.08	.18	.87
o ist einprägsam	.26	.59	.15	.54
o anders als andere Werbung	.15	.77	.32	-.01
o ist originell	.34	.79	.08	.16
o ist auffällig	.40	.65	.23	.11
o sagt Neues über die Marke	.24	.19	.83	.27
o vermittelt den Eindruck, die Marke ist anders	.40	.35	.71	.06
o ist sympathisch	.68	.38	.15	.25

Man kann diese 4 Faktoren wie folgt benennen:

Faktor 1: *"Glaubwürdigkeit und Überzeugung"*

Faktor 2: *"Uniqueness des Films"*

Faktor 3: *"Uniqueness der Marke"*

Faktor 4: *"Verständlichkeit"*

Die Gewichte dieser Faktoren zur Erklärung der Kaufbereitschaft zeigt das Ergebnis der multiplen Regressionsanalyse:

$$\begin{aligned} \text{Kauf-} \\ \text{bereit} \\ \text{schaft} &= 1.31 \times \text{Faktor 1} \\ &+ 0.69 \times \text{Faktor 2} \\ &+ 0.71 \times \text{Faktor 3} \\ &+ 0.38 \times \text{Faktor 4} \\ &+ \text{Konstante} \end{aligned}$$

Danach war Faktor 1 der "Wichtigste": Ist man in der Lage, einen Spot hinsichtlich dieses Faktors um eine Skalen-Stufe zu verbessern, steigt die Kaufbereitschaft überproportional (im Mittel um 1.31). Spot A war hier besonders gut und hinsichtlich der anderen Faktoren nicht schlecht - von daher erzielte er die relativ beste Kaufbereitschaft.

Ein Film, der hinsichtlich der "Uniqueness" seiner Machart herausragt, kann diesen Vorteil in knapp 0.7 Kaufbereitschafts-Skalenstufen umsetzen (so geschehen mit Spot B; kontraproduktiv kam hinzu, daß dieser Spot wenig "glaubwürdig" war).

Und ein Film, der lediglich "verständlicher" ist als andere, zieht praktisch keinen Vorteil daraus (s. Spot C).

Dieses Ergebnis verwundert uns nicht. Es deckt sich weitgehend mit der oben zitierten, 30 Jahre alten Erkenntnis von Politz:

1. I m p o r t a n c e
2. B e l i e v a b i l i t y
3. U n i q u e n e s s

Sind die Erkenntnisse generalisierbar?

Nachdem wir diesen Test in dieser Weise Stück für Stück analysiert hatten, stellte sich die Frage, ob man diese Ergebnisse auch in anderen CATs wiederfindet. Dazu wurde zunächst eine Reihe anderer Tests mit ganz unterschiedlichen Werbemitteln untersucht (Beispiele in Tabelle 20).

Zunächst ist noch einmal das Beispiel der drei ausführlich analysierten Spots dargestellt, danach die Ergebnisse eines Tests mit einem Werbemittel für eine etablierte Marke, für das die bestehende Werbung leicht zu korrigieren war, und als drittes Beispiel ein Test mit Werbung für eine Neueinführung.

Tab. 20

Beispiel 1: (unser bisheriges Beispiel)

<i>Faktor 1</i>	<i>Faktor 2</i>	<i>Faktor 3</i>	<i>Faktor 4</i>
-glaubwürdig	-originell	-sagt Neues	-verständlich
-Darstellung	-anders als	-Marke ist	-(einprägsam)
überzeugend	andere	anders	
-Aussage	Werbung		
überzeugend	-auffällig		
-informiert			
gut			
-sympathisch			
-wichtig			

Beispiel 2: (etablierte Marke)

<i>Faktor 1</i>	<i>Faktor 2</i>	<i>Faktor 3</i>	<i>Faktor 4</i>
-glaubwürdig	-originell	-sagt Neues	-verständlich
-Aussage	-anders als		
überzeugend	andere		
-Darstellung	Werbung		
überzeugend	-auffällig		
-informiert	-(einprägsam)		
gut			
-wichtig			
-(sympathisch)			

Beispiel 3: (neue Marke)

<i>Faktor 1</i>	<i>Faktor 2</i>	<i>Faktor 3</i>	<i>Faktor 4</i>
-wichtig	-anders als	-Marke ist	-(einprägsam)
-Aussage	andere	anders	-(informiert
überzeugend	Werbung	-auffällig	gut)
-(Darstellung	-(originell)	-(sagt Neues)	
überzeugend)		-(einprägsam)	
-(informiert		-(originell)	
gut)			
-wichtig			
-(sympathisch)			

In alle drei Fällen ergab sich eine weitgehend ähnliche Faktoren-Struktur. Einzelne Items bewegen sich in den drei Beispielen zwischen den Faktoren hin und her. Diese Abweichungen sind leicht zu erklären und bei der Einzel-Analyse der Spots wichtig:

- o Beim neuen Produkt befindet sich unter dem Faktor 3 ("Uniqueness des Produkts") auch "auffällig" und "originell". Hier wird deutlich, daß es sich um eine echte Novität handelte - kein x-tes "me-too-Produkt".
- o Im Fall des etablierten Produktes findet man unter dem Faktor "Uniqueness des Films" auch das Item "einprägsam". Das heißt, daß sich auffällige und originelle Werbung auch leichter einprägt.

Die grundsätzliche Ähnlichkeit der Faktoren hat uns dann dazu bewogen, einmal alle Tests, die wir seit 1979 durchgeführt hatten, in eine generelle Faktoren-Analyse einzubeziehen. Sie basiert im wesentlichen auf Tests von TV-Spots und einigen wenigen Anzeigen, insgesamt 53 Werbemitteln.

Wegen der Fülle der unterschiedlichen Spots und der damit gegebenen Breite des zugrundeliegenden Datenmaterials war es interessant, ob sich die oben gefundene Faktoren-Struktur generalisieren lassen würde.

Daneben war herauszufinden, wie sich andere Items neben den 13, die in die vorangegangenen Beispiele eingegangen waren, unter diese Faktoren einfügen würden.

Im Laufe der Zeit war immer wieder einmal mit verschiedenen Items experimentiert worden; es wurden neue erfunden und nur für bestimmte Filme verwendet, dann aber wieder verworfen. So kommt es, daß die in die "generelle" Faktoren-Analyse einbezogenen Items zwar zahlreicher und vielfältiger, aber nicht in allen Tests enthalten waren. Deshalb mußte in zwei Schritten vorgegangen werden:

- o Zunächst wurden die Korrelationen zwischen allen Items für die Fälle berechnet, in denen diese Items Bestandteil eines Tests waren. Die Korrelationen beruhen somit manchmal auf 5.500 Fällen, manchmal aber auch nur auf 300.
- o Die so errechnete Korrelationsmatrix enthielt einige "Lücken", da es vorkommt, daß einzelne Itempaare nicht gleichzeitig in mindestens einem Test zur Beurteilung herangezogen worden waren.
- o Zur Schließung dieser "Lücken" wurde ein spezielles Programm verwendet, das die nicht direkt berechenbaren Korrelationen zwischen zwei Items X und Y durch Vergleich der Korrelationen mit anderen, untereinander korrelierender Items abschätzt (das entsprechende Programm hierzu wurde durch Mind GmbH, Hamburg, entwickelt).

Die in der darauffolgenden Faktoren-Analyse extrahierten Faktoren lassen sich anhand der Ladungen beschreiben, wie in Tabelle 21 dargestellt:

	Ladung
Faktor 1: Glaubwürdigkeit	
- ist glaubwürdig	.85
- was gesagt wird, ist überzeugend	.75
- informiert gut	.71
- Darstellung ist überzeugend	.68
- was gesagt wird, ist glaubwürdig	.66
- was gesagt wird, ist wichtig	.55
Faktor 2a: Uniqueness der Werbung	
- Werbung ist anders	.81
- ist auffällig	.77
- ist originell	.68
Faktor 2b: Uniqueness der Marke	
- sagt Neues über das Produkt	.85
- Marke ist anders	.55
Faktor 3a: Verständlichkeit	
- ist verständlich	.83
- ist einprägsam	.59
Faktor 3b: Seriosität	
- ist seriös	.84
- ist sachlich	.73
Faktor 3c: Realistik	
- ist wirklichkeitsnah	.58
Faktor 4a: Sympathie	
- ist sympathisch	.71
- gefällt insgesamt	.55
- ist unterhaltsam	.50
- (ist interessant)	.49
- (Stimmung spricht an)	.45
Faktor 4b: Langeweile	
- ist langweilig	.97

Obwohl alle 8 Faktoren unabhängig sind, haben wir sie noch einmal semantisch zu Gruppen zusammengefasst. Die ersten beiden Faktoren fanden wir auch schon in den vorher beschriebenen Faktoren-Analysen.

Der bisherige Faktor "Verständlichkeit" wird jetzt noch ergänzt durch "Sachlichkeit und Seriosität" und "Wirklichkeitsnähe".

Neu hinzugekommen ist der Faktor "Sympathie" (4a+b), der die emotionale Ansprache, die durch den Spot ausgelöst wird, näher kennzeichnet.

Dieser neue Faktor verschiebt die Reihenfolge der "Wichtigkeit" dieser Faktoren hinsichtlich der Erklärung der kaufanreiz-auslösenden Wirkung von Werbung:

- o die Glaubwürdigkeit/Überzeugung/Wichtigkeit bleibt die Nr.1;
- o dann folgt die emotionale Ansprache
- o und danach erst die Uniqueness:

$$\begin{aligned} \text{Kauf-} \\ \text{bereit} \\ \text{schaft} &= 0.99 \times \text{Faktor 1} \\ &+ 0.45 \times \text{Faktor 2a} \\ &+ 0.52 \times \text{Faktor 2b} \\ &+ 0.30 \times \text{Faktor 3a} \\ &+ 0.49 \times \text{Faktor 3b} \\ &+ 0.33 \times \text{Faktor 3c} \\ &+ 0.57 \times \text{Faktor 4a} \\ &- 0.12 \times \text{Faktor 4b} \\ &+ \text{Konstante} \end{aligned}$$

Zwischenbilanz

- o Man sieht, daß ein grundsätzlicher Zusammenhang zwischen den in den CATs gewählten Beurteilungskriterien und der Kaufbereitschaft besteht; dieser ist besonders eng zwischen Filmbeurteilung und Kaufbereitschaft.
- o Desgleichen sieht man, daß die bisher verwendete Liste von Items mindestens um den Faktor "emotionale Ansprache" ergänzt werden muß, um allen Typen von Werbe-Spots gerecht zu werden.
- o Es ist möglich, Verallgemeinerungen hinsichtlich der Einflußfaktoren für die Kaufbereitschaft abzuleiten. Dies wird zukünftig vermutlich noch besser gelingen in dem Maße, wie es möglich ist, die Liste der relevanten Faktoren zu erweitern.

Bis hierhin gilt:

- wichtigster Faktor ist "Glaubwürdigkeit/
Überzeugung/persönliche Wichtigkeit",
- danach kommt "emotionale Ansprache",
- danach "Uniqueness"

...in dem Maße, wie die Regressions-Analyse diesen Faktoren "Gewichte" zugemessen hat.

- o Im selben Sinne ist es auch möglich, Vergleichsmaßstäbe zu setzen. In Abb.1 wurden zu diesem Zweck einmal alle 53 Werbemittel hinsichtlich der Faktoren 1 und 4a positioniert. Jeder neue Werbefilm läßt sich in diese Grafik einpassen, so daß man erkennen kann, ob man mit den erzielten Ergebnissen das gewünschte Niveau an "Glaubwürdigkeit", "emotionaler Ansprache", "Uniqueness" oder "Sachlichkeit" erreicht hat.

Ausblick*

In letzter Zeit läßt sich ein Trend zu solchen Werbemitteln beobachten, die stärker die emotionale Ansprache betonen, die stärker erlebnisorientiert sind.

Die Frage ist, ob auch solche TV-Spots noch mit Hilfe des CATs in seiner bis hierher vorgestellten Art adäquat bewertet werden können.

Für die Tests mit diesen neueren Entwicklungen schlagen wir vor, die Itemliste nochmals zu erweitern (u.a. mit Hilfe der Moldovan'schen Items), um die stark auf emotionale Wirkung und Unterhaltung zielende Gestaltung der Filme messen zu können.

Außerdem wird man versuchen müssen, bestimmte Gestaltungs-Elemente wie Farben, Musik, Stimmungen etc. in die Liste der Beurteilungskriterien aufzunehmen, um den Kreativen konkrete Anhaltspunkte für deren Arbeit liefern zu können.

Hierzu gibt es noch keine "fertigen" Itemlisten.

Man wird möglicherweise erst in ein bis zwei Jahren darüber berichten können, wenn mehrere Tests dieser Art vorliegen und wenn stabile Faktoren-Lösungen gefunden werden konnten.

*Dieser Ausblick bezieht sich auf den Erkenntnisstand des Jahres 1986.

Literatur:

Bauer Stiftung: "Die Affizierbarkeit verschiedener Kategorien von Zeitschriftenlesern durch Anzeigen", Hamburg 1967

Drehmann, Marianne: "Zwischen Kreativität und Relevanz: Chancen und Risiken ungewöhnlicher Kampagnen" in: K.F. Holm (Hrsg.): "Werbewirkungsforschung ex ante und ex post", Band 2, M+K Hansa Verlag, Mölln 1986

Hardy, Hugh (ed.): "The Politz Papers - Science and Truth in Marketing Research", American Marketing Ass., Chicago 1990

Koeppler, Karlfritz, u.a.: "Werbewirkungen definiert und gemessen", Heinrich Bauer Stiftung, Hamburg 1974

Krugman, Herbert (1967): "The Measurement of Advertising Involvement", Public Opinion Quarterly, 30/S.583-596

Krugman, Herbert (1977): "Memory without Recall, Exposure without Perception", JAR, August 1977

Leavitt, C.: "A Multidimensional Set of Rating Scales in Advertising", JMR, November 1975

Maloney, John C.: "Is Advertising Believability Really Important ?", JoM, October 1963

McCall's: "A Measurement of Advertising Effectiveness", New York, 1962

Moldovan, Stanley E.: "Copy Factors Related to Persuasion Scores", JAR, Dec.84/Jan.85; (deutsch: Vierteljahreshefte für Mediaplanung 4/1985)

Olson, David, Mary J. Schlinger & Charles Young (1982): "How Consumers React to New Product Ads", JAR, June/July 1982

Olson, David (1984): "Validation of Copy-Testing Measures Based on In-Market Performance: An Analysis of New-Product Ads", ESOMAR/EMAC-Seminar, Kopenhagen 1984

Olson, David (1985): "The Characteristics of High-Trial New-Product Advertising", JAR, Oct./Nov. 1985

Politz, Alfred: "Politz on Copy: Make the Sales Points Stick Out", Printers' Ink; nach: Lucas & Britt: "Measuring Advertising Effectiveness"; deutsch: "Messung der Werbewirkung", S.151, Giradet 1966, Neuabdruck in: Hardy (1990)

Schaefer, Wolfgang (1965): "Should Advertising be Believable or Convincing ?", ESOMAR-Seminar, München 1965

Schaefer, Wolfgang (1985): "Zu den Beziehungen zwischen Werbewirkungen und Kauf bzw. Verwendung", in K.F. Holm (Hrsg.): "Werbewirkungsforschung ohne Wirkung ?", Band 2, M+K Hansa Verlag, Mölln 1985

Schlinger, Mary J.: "A Profile of Responses to Commercials", JAR, April 1979

Wells, William D. (1964a): "Recognition, Recall, and Rating Scales", JAR, September 1964

Wells, William D. (1964b): "EQ, Son of EQ and the Reaction Profile", JoM, October 1964

Wells, William D., Clark Leavitt & Maren McConville (1971): "A Reaction Profile for TV Commercials", JAR, December 1971

Young, Shirley: "Copy Testing Without Magic Numbers", JAR, February 1972

Zinkhan, George M. & Claes Fornell: "A Test of Two Consumer Response Scales in Advertising", JMR, November 1985

—
JAR = Journal of Advertising Research

JoM = Journal of Marketing

JMR = Journal of Marketing Research

—