

Wolfgang Mayerhofer

# MAFO-Splitter



Dr. Wolfgang Mayerhofer, ao.Univ.Prof. am Institut für Werbewissenschaft und Marktforschung an der Wirtschaftsuniversität Wien.  
Wolfgang.Mayerhofer@wu-wien.ac.at

## Die Kunst der Fragebogenerstellung am Beispiel der Konstruktion von Skalen

Am Beispiel von drei experimentell durchgeführten Befragungen wird der empirische Nachweis erbracht, dass der Wertebereich der verwendeten Skala, die Verbalisierung der Endpole sowie die Anordnung von Beurteilungsobjekten einen Einfluss auf das Ergebnis der Befragung haben und Empfehlungen abgeleitet.

Zeit- und Kostendruck prägen unsere heutige Zeit. Dies betrifft die Gesellschaft in allen Lebenslagen und zieht sich durch Berufsleben, Alltagsleben und Freizeit und macht auch vor der Marktforschung nicht halt. Die Studie soll kostengünstig - um nicht zu sagen billig - sein, Entscheidungen mit weit reichenden Konsequenzen für Marken und Unternehmen absichern und die Ergebnisse am besten „gestern“ vorliegen. Es darf nicht verwundern, dass vereinzelt bereits der Begriff „quick and dirty“ fällt: „So genau will ich es gar nicht wissen, Hauptsache ich habe die Daten schnell!“ Bleibt zu hoffen, dass die mahnenden Stimmen, die darauf hinweisen, dass zwar jeder Fragen aber noch lange nicht jeder Markt- und Meinungsforschungsfragen stellen kann (Dürr 2002), sich gegen Do-it-yourself-Marktforschung stellen (Scheffler 2002) und nachdrücklich und anschaulich die Fallstricke von Frageformulierungen aufzeigen (Scheffler 2000), auch entsprechend Gehör finden. Die Beurteilung der Qualität von Marktforschung beginnt lange vor der Auswertung der Daten, will man nicht dem GIGO Phänomen (Garbage in - Garbage out) aufsitzen, wie es von Lachmann/Schroiff (1998) skizziert

wird. Die Fehler, die bei der Datenerhebung gemacht werden, können auch durch noch so aufwändige und anspruchsvolle Auswertungstechniken nicht korrigiert werden. Für die Beurteilung der Qualität von Marktforschungsleistungen ist es auch notwendig, sich nicht auf ISO-Zertifizierungen und Total-Quality-Management (TQM) Ansätze zu verlassen, sondern andere z.B. von Noelle-Neumann (1998) diskutierte Kriterien heranzuziehen.

Neben der richtigen Formulierung von Fragen ist die Verwendung von Skalen ein intensiv diskutiertes Thema in Forschung und Praxis. In diesem Zusammenhang muss zwischen der Messung von Konstrukten, wie dies am Beispiel einer Skala zur Messung der sozialen Verantwortung von Unternehmen und Marken von Grusch/Mertzbach (2006) demonstriert wird und der Zuordnung von Zahlen zu Objekten oder Eigenschaften meist an Hand von Ratingskalen unterschieden werden. Aber auch bei einfachen Ratingskalen gibt es eine Reihe von unterschiedlichen Gestaltungsmöglichkeiten (vgl. Berekoven/Eckert/Ellenrieder 2004, S. 77). Anlass für Diskussionen liefert einerseits das von manchen Auswertungsverfahren geforderte Ska-

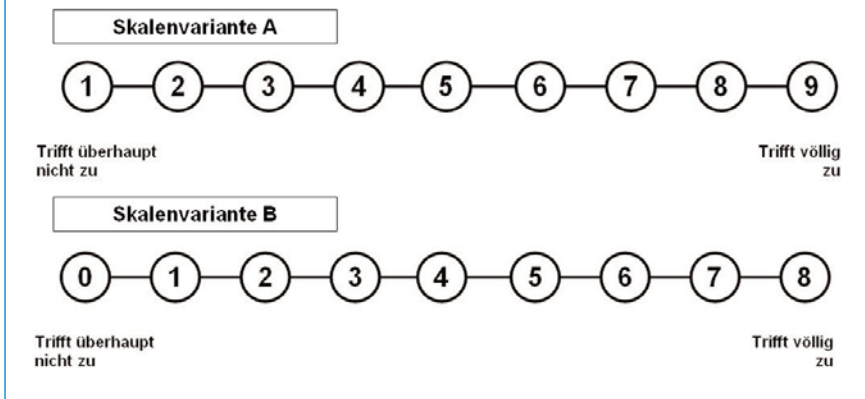
lenniveau (vgl. Weis/Steinmetz 2002, S. 124 und Hüttner 1997, S. 8f und S. 107ff) und andererseits die Anzahl der Skalenstufen, deren Verbalisierung sowie die Gerichtetheit der Skala (vgl. Malhotra/Birks 1999, S. 299). Am Institut für Werbewissenschaft und Marktforschung wurden von 2004 bis 2006 drei Studien durchgeführt, die neben umfangreichen inhaltlichen auch methodisch interessante Ergebnisse liefern und im Folgenden auszugsweise vorgestellt werden.

### Akzeptanz von Markenerweiterungen

Im Rahmen von Studie 1 wurden im Zeitraum Dezember 2003 bis Jänner 2004 insgesamt 330 Personen befragt. Die umfangreichen inhaltlichen Ergebnisse zur Akzeptanz von Markenerweiterungen in zahlreichen Produkt- und Dienstleistungsbereichen können bei Mahnik/Mayerhofer (2006) nachgelesen werden. Aufgabe des Skalenexperiments war die Überprüfung der Frage, ob die Verwendung einer 9-stufigen Skala von 1 bis 9 (Variante A) im Vergleich zu einer Skala von 0 bis 8 (Variante B) zu unterschiedlichen Ergebnissen führt (vgl. Mahnik 2004 und Abbildung 1).

Abb. 1

Skalenexperiment 1 mit unterschiedlichem Wertebereich



Zwei strukturgleiche Teilstichproben wurden mit den beiden unterschiedlichen Skalenvarianten konfrontiert. Die Ergebnisse zeigen bei einer Fülle an Erhebungstatbeständen (Kauf- und Nutzungsfrequenz, subjektiv empfundene Ähnlichkeit von Marken, Kaufwahrscheinlichkeit, Werbeintensität, Selbstimage der Befragten, Einstellungsmessung, usw.) einen durchgängigen Befund: Mit Variante B (0 - 8) werden in rund 16% aller durchgeführten Tests höhere Werte erzielt. Erklärt werden kann dies durch den Extremwert Null, der zu einem Fluchtverhalten und damit zu einer Tendenz zur anderen Seite der Skala führt. Dies gilt insbesondere für ältere Personen und Personen ohne Matura/Abitur. Darüber hinaus gibt es eine geschlechtsspezifische Sensitivität, die aber je nach Erhebungstatbestand unterschiedlich gerichtet ist. Eine Skala mit dem Endpunkt Null sollte deshalb nur dann eingesetzt werden, wenn der Wert „0“ tatsächlich für „kaufe ich nie oder verwende ich nie“ steht und nicht für die häufig anzutreffende Verbalisierung „trifft überhaupt nicht zu“ (Reinauer/Seewald 2005, S. 149).

Attraktivität von Handelsunternehmen als Arbeitgeber

Im Rahmen von Studie 2 wurden im Mai 2005 insgesamt 485 face-to-face Interviews mit Befragten aus den drei Zielgruppen Lehrlinge, Management-Nachwuchs und Verkaufspersonal durchgeführt (Bukovics/Falter 2005). Um einen Einfluss der Gerichtetheit der Skala auf die Beurteilung wichtiger Erhebungstatbestände zu überprüfen, wurden jeweils zwei strukturgleichen Teilstichproben aus den beiden Zielgruppen Management-Nachwuchs und

Verkaufspersonal unterschiedliche Skalenvarianten vorgelegt. In Variante A war die niedrigste Skalenstufe 1 mit „trifft überhaupt nicht zu“ verbalisiert und die höchste Skalenstufe 6 mit „trifft völlig zu“. In Variante B war es genau umgekehrt: Die Bezeichnung von Skalenstufe 1 lautete „trifft völlig zu“ und jene von Stufe 6 „trifft überhaupt nicht zu“ (vgl. Bukovics/Falter 2005). Die Begründung für Variante A lautet, dass ein höherer Wert auch für eine bessere Beurteilung steht: das Unternehmen bzw. die Marke ist attraktiver, vertrauenserweckender oder die Kaufwahrscheinlichkeit höher. Die Begründung für Variante B bezieht sich auf die Schulnotenskala: Eine bessere Note (= Beurteilung) entspricht einem niedrigeren Werte, während eine Note 5 oder 6 für ein „Nicht genügend“ und damit für eine schlechtere Beurteilung steht. Zusammenfassend kann festgestellt werden, dass das Skalenniveau von Variante B (niedriger Wert = bessere Beurteilung)

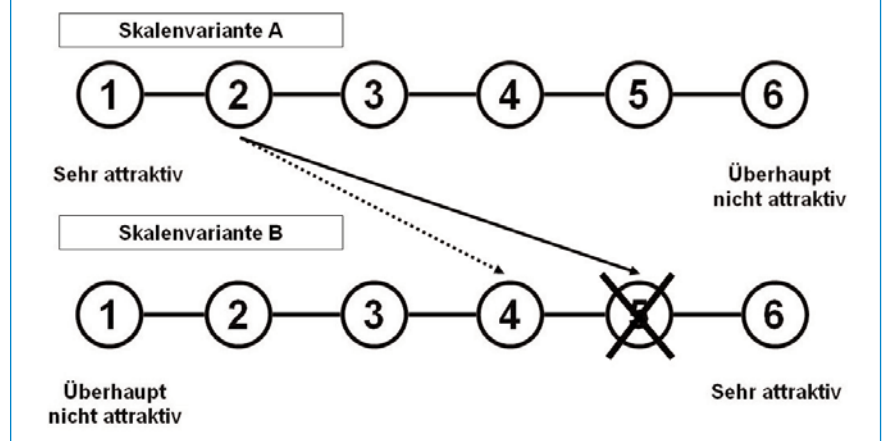
in rund 21% aller durchgeführten Tests zu höheren Werten und damit zu einer besseren Beurteilung führt. Die durchgehende Bevorzugung von Variante B zeigt sich in der Zielgruppe Management-Nachwuchs insbesondere bei jüngeren Männern und älteren Frauen. Für die Zielgruppe Verkaufspersonal gilt dies für alle Männer und für Frauen der höchsten Alterskategorie. Beide Befunde sprechen gegen Variante B und damit für eine Skala, deren niedrigster Endpunkt mit der geringeren Zustimmung verbalisiert ist (vgl. Reinauer/Seewald 2005, S.208f und Abbildung 2).

Images von Digitalkameramarken

Im Zusammenhang mit der Messung von Marken- oder Unternehmensimages ist es häufig zur Stützung der Erinnerungsleistung notwendig, die zur Beurteilung stehenden Objekte den Befragten auf einem Blatt vorzulegen. Dies kann durch einen neutralen Schriftzug oder, um ein facettenreicheres inneres Bild zu erzeugen, in Form von Markenlogos erfolgen. Falls es keine Gegenargumente gibt, wird man die Marken alphabetisch auf dem Blatt anordnen, da dies den Befragten keinen Anlass für das Hinterfragen der Reihenfolge der Marken liefert. Um Reihenfolgeeffekte zu vermeiden, die in Form von Primacy- und Recency-Effekten, also der Bevorzugung der ersten und letzten Elemente, mehrfach empirisch nachgewiesen wurden, geschieht die Vorlage meist nicht in Form von senkrechten Listen sondern durch waagrechte Anordnung der Marken auf einem Blatt. Dass auch diese Vorgangsweise zu unerwünschten Effekten führt, zeigen die Ergebnisse ei-

Abb. 2

Skalenexperiment 2 mit unterschiedlicher Gerichtetheit



ner Studie. Im November und Dezember 2005 wurden 100 persönliche Interviews mit Besitzern von Digitalkameras durchgeführt. Neben zahlreichen anderen Erhebungstatbeständen wurde auch das Image von sieben Digitalkameramarken anhand von Wörtern (verbal) und Bildern (nonverbal) erhoben. Zu diesem Zweck wurden den Befragten insgesamt 24 Wortreize und 28 Bilder vorgelegt, die nominal auf eine Marke, auf mehrere Marken oder auf keine Marke zugeordnet werden konnten. Obwohl sich die nonverbale Imagemessung in zahlreichen Anwendungsbereichen bewährt hat, so z.B. in der Länderimageforschung (Schweiger 1992) und zur Erhebung von Images in zahlreichen Produktgruppen (Mayerhofer 1995), erweist sich die Verwendung von Bildern in der Produktgruppe Kameras und Fotoapparate erneut als nicht unproblematisch (Slach/Mayerhofer 1993). Dies äußert sich in geringen Zuordnungen mancher Bilder auf die vorgelegten Marken und/oder hohen Zuordnungen auf die beiden Ausweichkategorien „weiß nicht“ oder „trifft auf keine Marken zu“. Um so schwer wiegender ist das Ergebnis des folgenden Experiments zu werten. Zwei strukturgleichen Teilstichproben wurden die sieben Marken in unterschiedlichen Anordnungen vorgelegt: In Variante A waren die Marken alphabetisch aufsteigend (von Canon bis Sony) und in Variante B absteigend (von Sony bis Canon) angeordnet (vgl. Wimmer 2006, S. 281 und Abbildung 3).

Für die starke Marke Sony und die eher schwächere Marke Pentax wurde überprüft, ob es durch die unterschiedliche Platzierung zu Unterschieden in der Zuordnung von Wort- und Bildreizen kommt. Auf beide Marken werden in der „vorne“ Platzierung signifikant häufiger negative Items wie z.B. das Wort Reparaturanfällig oder das Bild mit Pensionisten als Markenverwender und in der „hinten“ Platzierung signifikant mehr positive Items (Erotik, Sympathische Marke bzw. das gelungene Foto eines Hafens) zugeordnet. (vgl. Wimmer 2006, S. 197).

Es gibt somit einen signifikanten Zusammenhang zwischen der Reihenfolge der Anordnung der Markenlogos und der Markenbeurteilung. Die Beurteilung der Marke fällt besser aus, wenn der Markenname bzw. das Markenlogo weiter hinten angeordnet ist.

### Resümee

In Studie 1 werden in rund 16% aller durchgeführten Tests höhere Werte mit einer Skala von 0 bis 8 im Vergleich zu einer Skala von 1 bis 9 erzielt. Erklärt werden kann dies durch den Extremwert Null, der zu einem Fluchtverhalten und damit zu einer Tendenz zur anderen Seite der Skala führt. Eine Skala mit dem Endpunkt Null sollte deshalb nur dann eingesetzt werden, wenn der Wert „0“ tatsächlich für „kaufe ich nie oder verwende ich nie“ steht und nicht für die häufig anzutreffende Verbalisierung „trifft überhaupt nicht zu“. In Studie 2 sprechen alle Befunde für eine Skala, deren niedrigster Endpunkt mit der geringeren Zustimmung verbalisiert ist und gegen eine Skala, die sich an Schulnoten orientiert und niedrigere Werte demnach eine höhere Zustimmung zum Ausdruck bringen. Studie 3 belegt einen Einfluss der Anordnung von Marken einer Bildvorlage auf deren Beurteilung. Der letzte Befund spricht für die planmäßige Variation der Beurteilungsobjekte auf der vorgelegten Skala oder eine Randomisierung der Anordnung, wie dies bei Befragungen mit Computerunterstützung (CAPI) oder Online Befragungen leicht möglich ist.

### Literatur

**Berekoven L./Eckert W./Ellenrieder P. (2004):** Marktforschung - Methodische Grundlagen und praktische Anwendung, 10. Auflage, Gabler Verlag, Wiesbaden.  
**Bukovics M.-Th./Falter T. (2005):** Die Attraktivität von Handelsunternehmen als potentielle Arbeitgeber bei verschiedenen Zielgruppen, Diplomarbeit an der Wirtschaftsuniversität Wien.

**Dürr K. (2002):** Kann denn Liebe Sünde sein?, in: planung & analyse 4/2002, S. 4.

**Grusch L./Mertzbach M. (2006):** Der Einfluss von Corporate Social Responsibility (CSR) auf die Einstellung zu Unternehmen und Marken, Diplomarbeit an der Wirtschaftsuniversität Wien.

**Hüttner M. (1997):** Grundzüge der Marktforschung, 5. Auflage, R. Oldenbourg Verlag, München-Wien.

**Lachmann U./Schroiff H.-W. (1998):** Marktforschungs-Qualität: ein Begriff der Buchhaltung oder der Geisteshaltung?, in: planung & analyse 2/1998, S. 14-18.

**Mahnik N. (2004):** Einflussfaktoren auf die Akzeptanz von Markenerweiterungen, Diplomarbeit an der Wirtschaftsuniversität Wien.

**Mahnik, N./Mayerhofer W. (2006):** Erfolgsfaktoren von Markenerweiterungen, duv Verlag, Gabler Edition Wissenschaft, Wiesbaden.

**Malhotra N.K./Birks D.F. (2000):** Marketing Research - An Applied Approach, European Edition, Pearson Education Ltd, Harlow.

**Mayerhofer W. (1995):** Imagetransfer - Die Nutzung von Erlebniswelten für die Positionierung von Ländern, Produktgruppen und Marken, Service Fachverlag, Wien.

**Noelle-Neumann E. (1998):** Qualitätskriterien in der Marktforschung, in: planung & analyse, 2/1998, S. 10-13.

**Reinauer B./Seewald O.J. (2005):** Skalen in der Marktforschung - Konstruktion und Ansätze zur Überprüfung von Validität und Reliabilität, Diplomarbeit an der Wirtschaftsuniversität Wien.

**Scheffler H. (2000):** Vergessen Sie nicht die richtige Frageformulierung, in: planung & analyse, 5/2000, S. 10.

**Scheffler H. (2002):** Forschungs-Murks: Do-it-yourself-Marktforschung, in: planung&analyse, 4/2002, S. 9-10.

**Schweiger G. (1992):** Österreichs Image in der Welt - Ein Vergleich mit Deutschland und der Schweiz, Service Fachverlag, Wien.

**Slach M./Mayerhofer W. (1993):** Nonverbale Imagemessung bei Markenartikeln am Beispiel Fotoapparate, in: Werbeforschung und Praxis, 4/1993, S. 137-144.

**Wimmer H. (2006):** Positionierung von Digitalkameras für den Absatzmarkt Österreich anhand verbaler und nonverbaler Indikatoren unter dem Aspekt der digitalen Revolution, Diplomarbeit an der Wirtschaftsuniversität Wien.

Abb. 3

### Blatt mit Markenlogos zur Imagemessung

