



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Albrecht, J.: Validitätsarten und ihre Überprüfung bei der Conjoint-Analyse. In: Bauer, S.; Herrmann, R.; Kuhlmann, F.: Märkte der Agrar- und Ernährungswirtschaft – Analyse, einzelwirtschaftliche Strategien, staatliche Einflussnahme. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Band 33, Münster-Hiltrup: Landwirtschaftsverlag (1997), S.299-311.

VALIDITÄTSARTEN UND IHRE ÜBERPRÜFUNG BEI DER CONJOINT-ANALYSE

von

J. ALBRECHT*

1 Einleitung

Die Conjoint-Analyse ist eine neuere Methode der Präferenzmessung, die zunehmend in der Wissenschaft und Praxis verwendet wird. Sie wird bei kommerziellen Anwendungen hauptsächlich im Rahmen der Neuproduktentwicklung, bei Preisentscheidungen, in Marktsegmentationen und zur Marktsegmentierung eingesetzt (vgl. WITTINK ET AL. 1994, S. 44f.). Um fundierte Entscheidungen auf der Grundlage der Ergebnisse einer Conjoint-Analyse treffen zu können, muß sichergestellt sein, daß die Conjoint-Analyse die Präferenzen der befragten Personen möglichst exakt erfaßt, d.h. daß die Ergebnisse der Conjoint-Analyse valide sind. Ob eine Conjoint-Analyse zu ausreichend validen Ergebnissen führt, läßt sich nur feststellen, wenn man die Validität überprüft. Die vernachlässigte Überprüfung der Validität in kommerziellen Conjoint-Analysen kommt in einer von WITTINK ET AL. (1994, S. 47) durchgeführten Umfrage zum Ausdruck, nach der in Europa nur in neun Prozent der erhobenen Studien die Validität überprüft wurde. Betrachtet man wissenschaftliche Untersuchungen zur Validität, erscheinen mehrere Studien der Conjoint-Analyse gegenüber anderen Methoden gute Validitätsergebnisse. Allerdings wurden in den wissenschaftlichen Untersuchungen fast immer Objekte benutzt, bei denen die Präferenzbildung aufgrund einer intensiven Auseinandersetzung mit dem Objekt erfolgt und fast ausschließlich Studierende der Wirtschaftswissenschaften als Testpersonen herangezogen, obwohl die kognitiven Anforderungen an die Testpersonen bei einer Conjoint-Analyse als hoch einzustufen sind. Es stellt sich die Frage, ob die Aussagen zur Validität auch für Conjoint-Analysen mit Objekten wie Nahrungsmittel gelten, bei denen die Präferenzbildung i.d.R. nicht aufgrund einer differenzierten Auseinandersetzung erfolgt und bei anders zusammengesetzten Stichproben.

Vor diesem Hintergrund werden im folgenden Beitrag die Möglichkeiten zur Überprüfung der Validität der Conjoint-Analyse für individuell ermittelte Präferenzen (vgl. GREEN, SRINIVASAN 1990, S. 13) theoretisch erläutert. Die empirische Anwendung erfolgt anhand einzelner Validitätskriterien am Beispiel einer Conjoint-Analyse für das Produkt Frischmilch mit einer Stichprobe aus Konsumenten.

2 Durchführung einer Conjoint-Analyse am Beispiel Frischmilch

Die Conjoint-Analyse zählt zu den dekompositionellen Ansätzen der multiattributiven Präferenzanalyse. Bei ihr gibt jede befragte Person ihre Präferenzen für mehrere Produkte an. Diese Produkte sind durch bestimmte, i.d.R. objektive Eigenschaften charakterisiert. Aus den Präferenzurteilen wird durch ein Schätzverfahren ermittelt, wieviel die einzelnen Eigenschaften aus der Sicht einer Person zur Gesamtpreferenz für ein Produkt beitragen.

Genauer läßt sich die Durchführung einer Conjoint-Analyse anhand von sieben Schritten beschreiben (vgl. GREEN, SRINIVASAN 1978, S. 104ff.), wobei die Einzelentscheidungen der

* Dipl.-Ing. agr. J. Albrecht, Institut für Agrarpolitik und Landwirtschaftliche Marktlehre, Fachgebiet Agrarmarktanalyse, Universität Hohenheim, 70593 Stuttgart

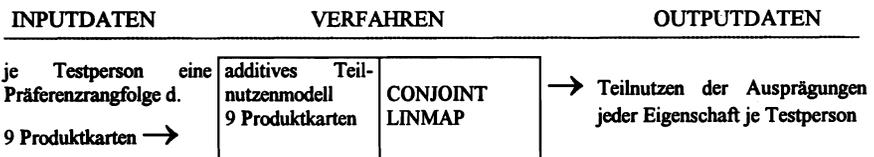
Schritte nicht unabhängig voneinander zu treffen sind. Zu Beginn sind die Eigenschaften und die jeweiligen Ausprägungen der Eigenschaften festzulegen. In der empirischen Untersuchung für das Produkt Frischmilch wurden die vier Eigenschaften Fettgehalt mit den Ausprägungen 1,5% und 3,5%, Verpackung (Glasflasche, Karton), Marke (Marke A, Marke B, Marke C) und Preis (1,09 DM, 1,39 DM, 1,79 DM) berücksichtigt. Im zweiten Schritt ist ein Präferenzmodell zu wählen, das den Zusammenhang zwischen der Gesamtpreferenz für ein Produkt und den Präferenzen für einzelne Eigenschaften abbildet. Für die empirische Untersuchung wurde das additive Teilnutzenmodell gewählt, das folgende Form aufweist:

$$(1) \quad U_{il} = \sum_{k=1}^K \sum_{j=1}^{J_k} x_{jkl} v_{ijk}$$

- $i = 1, \dots, I$ die betrachteten Testpersonen
- $j = 1, \dots, J_k$ die Ausprägungen von Eigenschaft k
- $k = 1, \dots, K$ die betrachteten Eigenschaften
- $l = 1, \dots, L$ die betrachteten Produkte (Produktprofile, Stimuli)
- U_{il} Gesamtpreferenz von Individuum i für das gesamte Produktprofil l
- x_{jkl} binäre Größe mit Wert 1, wenn Ausprägung j in Eigenschaft k im Produktprofil l vorkommt
- v_{ijk} Teilnutzen (Präferenzbeitrag) von Individuum i für Ausprägung j bei Eigenschaft k

Im dritten Schritt ist die Datenerhebungsmethode festzulegen. Es wurde die Vollprofilmethode (full-profile-Technik) angewandt, bei der die von den Testpersonen zu bewertenden Produktprofile durch alle berücksichtigten K Eigenschaften beschrieben werden. Diese Vorgehensweise kommt einer realen Entscheidungssituation näher als die Möglichkeit der vergleichenden Beurteilung einzelner oder weniger Eigenschaften (trade-off-Analyse). Der vierte Schritt, die Festlegung des Erhebungsdesigns, betrifft bei der Vollprofilmethode die beiden Fragen, wie die Ausprägungen der einzelnen Eigenschaften zu Produktprofilen zu kombinieren sind, damit eine unkorrelierte Schätzung der Parameter des gewählten Präferenzmodells möglich ist und wieviel Stimuli die Testpersonen beurteilen sollen. In der empirischen Untersuchung wurden aus den 36 möglichen Produktprofilen 9 Produktprofile entsprechend dem Basic-Plan 2 von ADDELMAN (1962, S. 36) gewählt. Schritt fünf, die Stimulipräsentation, kann in Form verbaler Beschreibungen, anhand von Abbildungen oder mittels realer Produkte erfolgen. In der empirischen Untersuchung wurden die 9 Produktprofile stichwortartig auf Papierkarten beschrieben. Schließlich ist in Schritt sechs eine Bewertungsskala zu wählen, mit der die Präferenzurteile der Befragten erfaßt werden. In der empirischen Untersuchung sollten die Befragten ihre Präferenzurteile durch Rangordnen der 9 Produktprofile zum Ausdruck bringen. Als Schätzverfahren für die Analyse ordinal skalierten Daten wurde in Schritt sieben das Verfahren CONJOINT LINMAP, die PC-Version von LINMAP (vgl. SRINIVASAN, SHOCKER 1973, S. 377ff.), benutzt. Schematisiert läßt sich das Prinzip der Conjoint-Analyse wie folgt darstellen:

Abbildung 1: Prinzip der Conjoint-Analyse



Die Erhebung der Präferenzurteile für Frischmilch fand in vier Verbrauchermärkten im Großraum Stuttgart statt. Es konnten insgesamt 601 Fragebögen ausgewertet werden. Die für die Validitätsuntersuchung notwendigen Zusätze zur bisher beschriebenen Durchführung werden nachfolgend an der entsprechenden Stelle beschrieben (vgl. Kapitel 5).

In Tabelle 1 sind die geschätzten durchschnittlichen Teilnutzen (TN) sowie die durchschnittlichen relativen Wichtigkeiten (rW) der Eigenschaften aufgeführt. Die relative Wichtigkeit einer Eigenschaft berechnet sich als Quotient der Differenz zwischen maximalem und minimalem Teilnutzen der Eigenschaft und der Summe dieser Differenzen über alle Eigenschaften (vgl. BACKHAUS ET AL. 1994, S. 521f.). Die über alle Probanden aggregierten Daten (vgl. Tab. 1) sollen lediglich einen ersten Eindruck vermitteln, sie sind bei heterogenen Stichproben für die Beschreibung der Konsumentenpräferenzen nur eingeschränkt sinnvoll.

Tabelle 1: Geschätzte Teilnutzen und relative Wichtigkeiten für die Gesamtstichprobe (n=601)

	Fettgehalt		Verpackung		Marke			Preis in DM (o. Pfand)		
	3,5%	1,5%	Glas	Karton	A	B	C	1,09	1,39	1,79
TN	8,66	-8,66	7,95	-7,95	4,12	-1,63	-2,49	11,01	2,49	-13,50
rW	26,92 %		24,71 %		10,28 %			38,09 %		

Quelle: eigene Berechnungen.

3 Validitätskategorien

In der Marktforschung dienen zahlreiche Meßinstrumente, unter anderem die Conjoint-Analyse, der Messung hypothetischer Konstrukte. Bei hypothetischen Konstrukten handelt es sich um Begriffe wie beispielsweise Nutzen oder Präferenz, die keinen unmittelbaren Bezug zu empirisch wahrnehmbaren Sachverhalten haben und deshalb schwierig zu messen sind. Die Gütekriterien, zu denen auch die Validität gehört, beschreiben unterschiedliche Aspekte der Eignung bzw. Güte eines Meßinstruments (vgl. z.B. BORTZ, DÖRING 1995, S. 180). Die Validität gilt als das zentrale Gütekriterium und im Vergleich zu den anderen Gütekriterien als am schwierigsten erfassbar und überprüfbar (vgl. z.B. BORTZ, DÖRING 1995, S. 185; BÖHLER 1992, S. 105). Die Validität beschreibt die inhaltliche Gültigkeit eines Meßinstruments. Sie kennzeichnet, wie gut ein Meßinstrument in der Lage ist, dasjenige hypothetische Konstrukt zu messen, das es messen soll und betrifft somit die Frage, ob das Meßinstrument tauglich und angemessen ist, das interessierende hypothetische Konstrukt tatsächlich zu erfassen (vgl. LIENERT 1969, S. 255). Es ist beispielsweise denkbar, daß ein Meßinstrument ein hypothetisches Konstrukt nur unvollständig erfaßt oder ein dem hypothetischen Konstrukt ähnliches hypothetisches Konstrukt mißt. Für die Validität sind in der Literatur eine Vielzahl unterschiedlichster Kategorisierungsansätze zu finden. (vgl. HOSSINGER 1982, S. 16ff., S. 35). Im folgenden sollen diejenigen Hauptkategorien der Validität unterschieden werden, die sich auch in der Marktforschung allmählich durchsetzen (vgl. HAMMANN, ERICHSON 1994, S. 77): Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität.

Inhaltsvalidität liegt vor, wenn ein Meßinstrument inhaltlich so ausgestaltet ist, daß eine erschöpfende Erfassung des hypothetischen Konstrukts erwartet werden kann. Die Inhaltsvalidität wird anhand subjektiver Einschätzungen beurteilt.

Kriteriumsvalidität eines Meßinstruments ist gegeben, wenn die mit dem Meßinstrument gewonnenen Meßwerte des hypothetischen Konstrukts mit Meßwerten eines sogenannten Krite-

riums übereinstimmen. Unter einem Kriterium versteht man entweder dasselbe hypothetische Konstrukt, das mit einem anderen Meßinstrument gemessen wurde oder einen empirisch beobachtbaren Sachverhalt, der mit dem hypothetischen Konstrukt logisch eng zusammenhängt. Wenn es mehrere Kriterien gibt, kann ein Meßinstrument gegenüber mehreren Kriterien validiert werden und folglich mehrere Kriterienvaliditäten haben. Die Auswahl geeigneter Kriterien hängt vom Untersuchungszweck ab.

Die Konstruktvalidität ist die umfassendste und anspruchsvollste Validitätskategorie. Bei der Konstruktvalidität geht es um die theoretische Klärung dessen, was genau das jeweilige Meßinstrument mißt (vgl. LIENERT 1969, S. 261). Die Konstruktvalidität soll Anhaltspunkte liefern, wie präzise das hypothetische Konstrukt mit einem Meßinstrument erfaßt wird.

Die Konstruktvalidierung eines Meßinstruments interessiert besonders aus theoretischer Sicht (vgl. HAMMANN, ERICHSON 1994, S. 78). Sie ist für die Überprüfung und Entwicklung von Theorien wesentlich (vgl. PETER 1981, S. 133). Kann man davon ausgehen, daß ein Meßinstrument ein theoretisches Konstrukt abbildet, dann treten aus praktischer Sicht die beiden anderen Validitätskategorien in den Vordergrund. Bei der praktischen Anwendung innerhalb der Marktforschung ist neben der Inhaltsvalidität die Kriteriumsvalidität eines Meßinstruments von vorrangigem Interesse (vgl. HAMMANN, ERICHSON 1994, S. 78; LIENERT 1969, S. 255).

4 Inhalts- und Kriteriumsvalidität bei Conjoint-Analysen

4.1 Inhaltsvalidität

Conjoint-Analysen werden hinsichtlich ihrer inhaltlichen Validität überprüft, wenn beurteilt wird, ob alle (wesentlichen) präferenzbeeinflussenden Eigenschaften des zu untersuchenden Objekts berücksichtigt werden und die einzelnen Ausprägungsstufen der Eigenschaften dem Untersuchungszweck entsprechend sinnvoll festgelegt sind.

4.2 Kriteriumsvalidität

Es wurde bereits erwähnt, daß sich häufig mehrere Kriterien zur Validierung heranziehen lassen. Um über die Kriterien einen Überblick zu erhalten, wird die Kriteriumsvalidität häufig in Validitätsarten eingeteilt (vgl. BORTZ, DÖRING 1995, S. 186; LIENERT 1969, S. 257ff.). Eine generelle Unterteilung in Validitätsarten ist für die Messung eines bestimmten hypothetischen Konstrukts mit einem bestimmten Meßinstrument jedoch häufig nicht befriedigend (vgl. HOSSINGER 1982, S. 39f.). Deshalb ist im folgenden eine Einteilung zu wählen, die für Conjoint-Analysen geeignet ist. Bei der Abgrenzung von Arten der Kriteriumsvalidität im Rahmen der Conjoint-Analyse bietet es sich an, folgende Aspekte zu berücksichtigen:

- Bei der Conjoint-Analyse ist ein bestimmtes Präferenzmodell auszuwählen, dessen Parameter aus den Präferenzurteilen der Testpersonen geschätzt werden. Bei der Validierung kann somit auf die Anpassung des Modells an die Ausgangsdatenbasis zurückgegriffen werden.
- Es können Kriterien zur Validierung herangezogen werden, bei denen es sich um beobachtbare und direkt meßbare objektive, empirische Sachverhalte handelt.
- Es stehen desweiteren andere Meßinstrumente (Meßverfahren) zur Verfügung, die dieselben oder einen Teil der hypothetischen Konstrukte messen, die auch die Conjoint-Analyse mißt.
- Außer diesen objektiven Kriterien können auch subjektive Kriterien in Form von subjektiven Einschätzungen bei der Validierung verwendet werden.

Unter Berücksichtigung dieser Aspekte lassen sich für die Conjoint-Analyse folgende Validitätsarten innerhalb der Kriteriumsvalidität unterscheiden: Face-Validität, interne Validität, Prognose-Validität und konvergierende Validität (vgl. hierzu GREEN, SRINIVASAN 1978, S. 115f.; MÜLLER-HAGEDORN et al. 1993, S. 127f.; SCHWEIKL 1985, S. 70ff.).

4.2.1 Face-Validität

Face-Validität zielt darauf ab, ob die Meßwerte plausibel sind und nicht gegen den „gesunden Menschenverstand“ verstoßen. Zur Beurteilung der Face-Validität sind a priori-Erwartungen bezüglich der Meßergebnisse aufzustellen, die anhand der Ergebnisse zu überprüfen sind. Die a priori-Erwartungen leitet der Marktforscher aus Vorüberlegungen ab. Bei den a priori-Erwartungen handelt es sich somit - wie auch bei der Inhaltsvalidität - um subjektive Kriterien. Bei der Face-Validierung ist allerdings zu beachten, daß eine Verletzung der a priori-Erwartungen außer durch mangelnde Face-Validität auch durch nicht zutreffende a priori-Erwartungen bedingt sein kann.

Für die Überprüfung der Face-Validität einer Conjoint-Analyse gibt es zwei Ansatzpunkte. Wird a priori ein bestimmter Verlauf der Teilnutzen der Ausprägungen bei einzelnen Eigenschaften erwartet, so sind diese Erwartungen bei diesen Eigenschaften anhand der Teilnutzenergebnisse zu prüfen. Es können aber auch a priori-Erwartungen hinsichtlich des Verlaufs der relativen Wichtigkeit der Eigenschaften bestehen, die sich anhand der Ergebnisse überprüfen lassen.

4.2.2 Interne Validität

Die interne Validität bezieht sich auf die Anpassung des Präferenzmodells an die Inputdaten. Interne Validität liegt vor, wenn die von den Probanden erhobenen Präferenzurteile (Inputdaten) mit den aus diesen Präferenzurteilen geschätzten Parametern des Präferenzmodells (Outputdaten) reproduziert werden können (vgl. Abb. 1). Die zur Schätzung der Parameter herangezogenen Produktprofile werden als calibration sample bezeichnet.

Gemessen wird die interne Validität durch Korrelation der Inputdaten mit den reproduzierten Präferenzurteilen. In Abhängigkeit vom Skalenniveau der Inputdaten ist ein entsprechender Korrelationskoeffizient zu verwenden. Bei ordinalem Skalenniveau wird der Rangkorrelationskoeffizient von Spearman (r_s) oder Kendall (τ) (vgl. SIEGEL 1956, S. 213ff.) verwendet, bei metrischem Skalenniveau der Korrelationskoeffizient von Pearson (vgl. z.B. AKAAH, KORGAONKAR 1983, S. 192).

Die interne Validität einer Conjoint-Analyse wird beeinträchtigt, falls das unterstellte Präferenzmodell ungeeignet ist, die Beurteilung (nicht den Prozeß der Urteilsbildung) der Versuchspersonen abzubilden, falls mangelnde Reliabilität vorliegt, also beispielsweise das Schätzverfahren unzulänglich ist oder die Probanden bei einem an sich geeigneten Präferenzmodell keine konsistenten Präferenzurteile abgeben.

4.2.3 Prognosevalidität

Die Prognosevalidität beschreibt die prognostische Leistungsfähigkeit einer Conjoint-Analyse. Dazu sind - im Gegensatz zur internen Validierung - Produktprofile heranzuziehen, die nicht zur Schätzung der Parameter des Präferenzmodells verwendet werden. Diese Produktprofile werden als holdout sample bezeichnet. Für die Überprüfung der Prognosevalidität ist es also notwendig, daß die Befragten zusätzlich zur Abgabe der für die Parameterschätzung benötigten Präferenzurteile weitere Produktprofile beurteilen. Prognosevalidität liegt vor, wenn die von den Probanden erhobenen Präferenzurteile zu den Produktprofilen des holdout sample mit ihren aus dem calibration sample geschätzten Parametern des Präferenzmodells prognostiziert werden können.

Zur Messung der Prognosevalidität werden hauptsächlich zwei Maße benutzt. Das eine Maß ist die Korrelation der erhobenen Präferenzurteile mit den prognostizierten Präferenzurteilen des holdout sample. In Abhängigkeit vom Skalenniveau der erhobenen Präferenzurteile der Validierungsobjekte ist ein entsprechender Korrelationskoeffizient anzuwenden (vgl. Kapitel 4.2.2). Beim zweiten Maß konzentriert sich das Interesse nicht auf die Präferenzwerte aller Produktprofile des holdout sample, sondern auf das Produktprofil, das eine Person am stärksten präferiert. Das Maß ist der Anteil der Personen, für deren am stärksten präferiertes Produkt des holdout sample auch die prognostizierte Präferenz am höchsten ist. Dieses Maß zielt darauf ab, ob die Parameter des Präferenzmodells einer Person dazu geeignet sind, aus einer vorgegebenen Menge von Produkten das Produkt zu bestimmen, das eine Person am stärksten präferiert. Geht man davon aus, daß eine Person das am stärksten präferierte Produkt auch auswählt bzw. kauft (first choice Annahme), wird die Bedeutung dieses Maßes für die Prognose des Kaufverhaltens klar.

Für die Prognosevalidität sind geringere Werte als für die interne Validität zu erwarten, da das holdout sample bei der Parameterschätzung des Präferenzmodells nicht berücksichtigt wird und sich deshalb die Ursachen, die zu einer Beeinträchtigung der internen Validität führen (vgl. Kapitel 4.2.2), bei der Prognosevalidität i.d.R. verstärkt auswirken.

Die Überprüfung der Prognosevalidität kann mittels hypothetischer Produktprofile, realer Produktprofile und anhand der tatsächlichen Produktwahl erfolgen. Hypothetische Stimuli sind auf dieselben Eigenschaften wie die Stimuli des calibration sample reduziert. Ihre Präsentation erfolgt normalerweise auf die gleiche Weise wie beim calibration sample, bei der Conjoint-Analyse für Frischmilch in Form verbaler Produktbeschreibungen auf Karten. Bei realen Produkten erfolgt keine Einschränkung der Eigenschaften wie bei den hypothetischen Produktprofilen. Die Probanden entscheiden somit selbst, welche Eigenschaften sie bei der Präferenzbeurteilung berücksichtigen. Die Präsentation realer Produkt erfolgt i.d.R. in Form von echten Produkten bzw. Prototypen oder als Photos, Prospekte etc. Die tatsächliche Produktwahl bezieht sich auf die Wahl realer Produkte in einer tatsächlichen Entscheidungssituation „außerhalb“ der Untersuchung. Bei der Conjoint-Analyse für Frischmilch handelt es sich bei der tatsächlichen Produktwahl um die normalerweise gekaufte Frischmilch.

4.2.4 Konvergierende Validität

Die konvergierende Validität kann überprüft werden, wenn mindestens zwei Meßinstrumente verfügbar sind, die dasselbe hypothetische Konstrukt messen. Konvergierende Validität liegt vor, wenn diese Meßinstrumente zu übereinstimmenden Ergebnissen führen. Die Überprüfung der konvergierenden Validität kann bei Conjoint-Analysen an den Teilnutzen der Eigenschaftsausprägungen, an den relativen Wichtigkeiten der Eigenschaften und an den Gesamtnutzen von

Produktprofilen ansetzen. Das bedeutet, daß zur Überprüfung der konvergierenden Validität Meßinstrumente in Frage kommen, die zu mindestens einem dieser drei Ergebnisebenen Meßwerte liefern. Ergebnisse zu allen drei Bereichen erhält man beim Vergleich unterschiedlicher Varianten der Conjoint-Analyse, wie z.B. der trade-off-Analyse und der full-profile-Technik sowie durch einen Vergleich der Conjoint-Analyse mit kognitiven Einstellungsmodellen (vgl. AKAAH, KORGAONKAR 1983, S. 187ff.; MÜLLER-HAGEDORN et al. 1993, S. 127f.).

Als Maß zur Darstellung, inwieweit zwei Verfahren zu übereinstimmenden Ergebnissen kommen, werden ebenfalls häufig Korrelationskoeffizienten verwendet. Entsprechend der drei Ergebnisebenen läßt sich die konvergierende Validität mittels Pearson'schem Korrelationskoeffizienten und/oder den Rangkorrelationskoeffizienten für Teilnutzen, relative Wichtigkeiten und Gesamtnutzen prüfen (vgl. z.B. SHELUGA et al. 1979, S. 171f.). Bei der Interpretation der Ergebnisse zur Feststellung der konvergierenden Validität ist zu berücksichtigen, daß der Grad der Übereinstimmung auch durch die Ähnlichkeit der Verfahren beeinflußt wird.

5 Überprüfung einzelner Validitätsarten der Conjoint-Analyse für Frischmilch

5.1 Überprüfung der internen Validität

Bei der Conjoint-Analyse für Frischmilch hatte jede Person die 9 Produktprofile des calibration sample entsprechend ihrer Präferenz in eine Rangordnung zu bringen. Aus dieser Rangordnung wurden für jeden Befragten seine Parameter (Teilnutzen) des Präferenzmodells geschätzt (vgl. Kapitel 2). Damit die interne Validität überprüft werden kann, wird zunächst für jeden Probanden die Rangfolge dieser 9 Stimuli anhand seiner Teilnutzen berechnet. Die Überprüfung der internen Validität erfolgt durch den Vergleich der angegebenen Präferenzrangfolge (Inputdaten) und der berechneten Präferenzrangfolge dieser Stimuli für jeden Probanden. Als Maß wird der Rangkorrelationskoeffizient von Kendall zwischen den beiden Präferenzrangfolgen jeder Testperson berechnet. Zur Charakterisierung der Verteilung von Kendalls τ wird hier und im folgenden Kapitel die Häufigkeitsverteilung, das 5 % Quantil, der Median, das 95 % Quantil sowie für den Vergleich mit anderen Untersuchungen der Mittelwert von τ berechnet (vgl. Tab. 2).

Tabelle 2: Häufigkeitsverteilung von τ zur Beurteilung der internen Validität

Klassen von τ	Anzahl Testpersonen	
	absolut	%
0,401-0,6	7	1,2
0,601-0,8	29	4,8
0,801-1	565	94,0
Σ	601	100,0

Quelle: eigene Berechnungen.

Betrachtet man die Häufigkeitsverteilung von τ , so ergeben sich für die meisten Testpersonen hohe Werte, was auf eine hohe interne Validität schließen läßt. Die niedrigen Werte für einige Personen können dadurch zustande kommen, daß das unterstellte Präferenzmodell für diese Personen ungeeignet ist, daß sie die Rangordnungsaufgabe nicht verstanden haben oder damit überfordert sind (vgl. Kapitel 4.2.2).

Tabelle 3: Statistische Verteilungsmaße von τ zur Beurteilung der internen Validität

	τ
5 % Quantil	0,778
Median	1,0
95 % Quantil	1,0
Mittelwert	0,954

Quelle: eigene Berechnungen.

In Conjoint-Analysen, die in etwa mit der vorliegenden Untersuchung vergleichbar sind, ermittelten TROEGEL und ROTHSPRACH (1996, S. 530, S. 534) für das Produkt Käse ein durchschnittliches Kendalls τ von 1, für Brot von 0,944, SATTLER (1991, S. 173) für Konfitüre von 0,98 und WEISENFELD (1989, S. 121) Werte um 0,9 für Fleischsalat. Ob bei Nahrungsmitteln Konsumenten (ALBRECHT, TROEGEL und ROTHSPRACH, SATTLER) oder Studierende (WEISENFELD) als Stichprobe herangezogen werden, macht bei den verglichenen Untersuchungen bezüglich der internen Validität offenbar keinen großen Unterschied. Der Vergleich obiger Conjoint-Analysen für Nahrungsmittel läßt außerdem vermuten, daß bei Anwendung des additiven Teilnutzenmodells für die interne Validität durchschnittliche Kendalls τ Werte von über 0,9 erzielbar sind.

5.2 Überprüfung der Prognosevalidität

Die Prognosevalidität wird anhand hypothetischer Produktprofile, realer Produktprofile und anhand des tatsächlich gekauften Produkts überprüft. Zunächst wird auf die hypothetischen und realen Produktprofile eingegangen. In der Befragung hatten die Probanden zusätzlich zu den Stimuli des calibration sample ein holdout sample aus 5 hypothetischen Stimuli und ein weiteres holdout sample aus 5 realen Stimuli entsprechend ihrer Präferenz rangzuordnen. Die beiden holdout samples waren in ihren Eigenschaftsausprägungen identisch, sie unterschieden sich aber in ihrer Präsentationsform. Die 5 hypothetischen Stimuli wurden stichwortartig auf Karten beschrieben (Karten), die 5 realen Produktprofile bestanden aus echten Produkten (reale Produkte). Die Fragen zu den beiden holdout samples wurden im Fragebogen durch Zwischenfragen voneinander getrennt, wodurch erreicht wurde, daß fast keine Testperson merkte, daß es sich beide Male um dieselben fünf Produkte in unterschiedlicher Präsentationsform handelte. Damit die Prognosevalidität überprüft werden kann, wird für jeden Probanden die Rangfolge der 5 Stimuli anhand seiner Teilnutzen prognostiziert. Die Überprüfung der Prognosevalidität erfolgt, indem für jeden Befragten seine angegebene Präferenzrangfolge der Karten bzw. realen Produkte mit seiner prognostizierten Rangfolge verglichen wird.

Als erstes Maß wird der Rangkorrelationskoeffizient von Kendall berechnet. Die Häufigkeitsverteilung des für jeden Probanden berechneten Korrelationskoeffizienten zwischen der angegebenen und prognostizierten Präferenzrangfolge ist in Tabelle 4 dargestellt, die Maße zur Beurteilung der Häufigkeitsverteilung in Tabelle 5. Wie zu erwarten ist, fallen die Korrelationen für die holdout Stimuli niedriger aus als die zur Beurteilung der internen Validität berechnete Korrelation (vgl. Tab. 4 u. 2). Ursachen für den Abfall der Korrelationswerte kann z.B. sein, daß die Probanden bei den Rangordnungsaufgaben keine konsistenten Präferenzurteile abgegeben haben.

Tabelle 4: Häufigkeitsverteilung von τ zur Beurteilung der Prognosevalidität

Klassen von τ	Karten		reale Produkte	
	Anzahl Testpersonen		absolut	%
	absolut	%		
$\leq 0,4$	144	24,0	142	23,6
0,401-0,6	106	17,6	130	21,6
0,601-0,8	184	30,6	176	29,3
0,801-1	167	27,8	153	25,5
Σ	601	100,0	601	100,0

Quelle: eigene Berechnungen.

Tabelle 5: Statistische Verteilungsmaße von τ zur Beurteilung der Prognosevalidität

	τ Karten	τ reale Produkte
5 % Quantil	0,115	0,000
Median	0,800	0,800
95 % Quantil	1,000	1,000
Mittelwert	0,680	0,665

Quelle: eigene Berechnungen.

Zwischen den Karten und den realen Produkten ist ein relativ geringer Abfall der Korrelationen zu erkennen. Vergleicht man den Median von $\tau = 0,8$ für die Produktpräsentation mit dem Median von $\tau = 0,6$, den TSCHUELIN (1992, S. 144) für die Prognosevalidität von ebenfalls eher realistischen Produkten, nämlich von auf Katalogseiten abgebildeten Kreuzfahrten, ermittelt hat, ist die Prognosevalidität für dieses Kriterium bei der Frischmilchuntersuchung als hoch anzusehen. Allerdings ist der Vergleich beider Untersuchungen durch die unterschiedlichen Produkte und die Unterschiede in der Anzahl der Stimuli und Eigenschaften nur eingeschränkt möglich.

Als weiteres Maß wurde der Anteil der Personen berechnet, deren Karte bzw. reales Produkt in ihrer angegebenen Präferenzrangfolge auf dem ersten Rangplatz (höchste Präferenz) auch in ihrer prognostizierten Rangfolge auf dem ersten Rangplatz liegt. In Tabelle 6 ist die Häufigkeitsverteilung der prognostizierten Rangplätze für die von den Befragten am stärksten präferierten Produktprofile dargestellt. Bei den Karten wurde für 70,5 % der Probanden dasjenige der 5 Stimuli auf den ersten Rangplatz prognostiziert, das sie auch in ihrer angegebenen Präferenzrangfolge am stärksten präferiert haben. Für 19,3 % der Probanden wurde das am stärksten präferierte Produktprofil auf den zweiten Rangplatz prognostiziert etc. Vergleicht man den Anteil von 70,5 % mit 69,8 % korrekt prognostizierter erster Rangplätze bei den realen Stimuli, so zeigt sich ebenso wie bei den Korrelationskoeffizienten (vgl. Tab. 4 u. 5) ein sehr geringer Unterschied. Dieser Sachverhalt läßt sich - zumindest bei Produkten, die dem Konsumenten bekannt sind - so erklären, daß die Verwendung von verbal beschriebenen Produktprofilen als calibration sample durchaus geeignet ist, die Präferenz für reale Produkte zu erfassen. In der bereits oben angesprochenen Untersuchung von TSCHUELIN (1992, S. 147) konnte in 57,3% der Fälle der erste Rangplatz korrekt prognostiziert werden.

Tabelle 6: Häufigkeitsverteilung der prognostizierten Rangplätze der an die erste Stelle sortierten Validierungsobjekte (Karten, reale Produkte) und der tatsächlich gekauften Frischmilch

Prognostizierter Rangplatz	Karten Anzahl Testpersonen*		Reale Produkte		Gekaufte Produkte**	
	absolut	%	absolut	%	absolut	%
1	423,5	70,5	419,5	69,8	67	63,2
2	116	19,3	117	19,5	23	21,6
3	40	6,7	45	7,5	6	5,7
4	19,5	3,2	16,5	2,7	6	5,7
5	2	0,3	3	0,5	2	1,9
6	-	-	-	-	2	1,9
Σ	601	100,0	601	100,0	106	100,0

* Wurden für zwei Produktprofile mittlere Ränge prognostiziert, so wurde die Testperson je zur Hälfte auf die benachbarten ganzen Ränge aufgeteilt.
** Für die Käufer von zwei und mehr Produkten wurde ausschließlich der höchste prognostizierte Rangplatz eines Produkts berücksichtigt.

Quelle: eigene Berechnungen.

Die Überprüfung der Prognosevalidität anhand der tatsächlich von den Testpersonen gekauften Frischmilch wird für einen der vier Verbrauchermärkte vorgenommen. Die Testpersonen wurden deshalb in der Befragung gebeten, die Frischmilch zu nennen, die sie normalerweise kaufen. Durch die in die Conjoint-Analyse einbezogenen Eigenschaftsausprägungen konnten 6 der 8 in dem Verbrauchermarkt angebotenen Frischmilcharten mit dem Präferenzmodell abgebildet werden. In diesem Verbrauchermarkt kauften 106 befragte Konsumenten mindestens eine Frischmilch aus diesen 6 Produkten. Für jeden Probanden wurde der prognostizierte Rangplatz seiner gekauften Frischmilch unter den 6 möglichen Frischmilchprodukten berechnet. Der vierten Spalte von Tabelle 6 ist die Verteilung der prognostizierten Rangplätze zu entnehmen, die sich für die gekaufte Frischmilch ergibt. Für 63,2 % der Konsumenten wird für dasjenige Produkt aus den 6 angebotenen Produkten, das sie kaufen, die höchste Gesamtpreferenz berechnet. Ausgehend von der first choice Annahme kann also für 63,2 % der Konsumenten aus den sechs mit der Conjoint-Analyse abbildbaren Produkten die Frischmilch prognostiziert werden, die sie kaufen. Im Vergleich mit den korrekt prognostizierten ersten Rangplätzen der Karten und realen Produkte zeigt sich - wie erwartet - ein geringerer Wert. Ursachen für die geringere Prognosevalidität sind u.a., daß einige Konsumenten beim Kauf weitere Eigenschaften wie z.B. das Mindesthaltbarkeitsdatum berücksichtigen, die in der Conjoint-Analyse nicht berücksichtigt wurden, und daß die Konsumenten in der Befragung anders bzw. konsistenter als beim Einkauf reagieren.

Um einen Anhaltspunkt zu erhalten, wie die 63,2 % korrekt prognostizierter Produkte zu beurteilen sind, kann als Referenzgröße das Zufallsmodell herangezogen werden. Unter Anwendung des Zufallsmodells ist zu erwarten, daß für 21,4 % der Konsumenten zufällig die Frischmilch ermittelt wird, die sie kaufen, d.h. mit der Conjoint-Analyse lassen sich zusätzlich 41,8 % der Frischmilchkäufe korrekt prognostizieren. Für den Vergleich mit anderen Conjoint-Analysen, bei denen die tatsächliche Produktwahl ebenfalls erfaßt wurde, kann ein von SRINIVASAN (1988, S. 304) vorgeschlagenes Maß verwendet werden. Dieses Maß β gibt an, welcher Prozentsatz der falschen Prognosen mit dem Zufallsmodell durch die Conjoint-Analyse zusätzlich erklärt wird. Für die Frischmilchuntersuchung erhält man ein β von 53,4 %. Für eine Conjoint-Analyse zur Wahl von Arbeitsstellen durch Studenten von WITTINK und MONT-

GOMERY läßt sich ein β von 49,9 % berechnen (vgl. SRINIVASAN 1988, S. 304). Für die bereits oben zitierte Untersuchung von TSCHEULIN (1992, S.150) läßt sich ein β von 63,8 % berechnen. Der Vergleich zeigt, daß diese Art der Prognosevalidität der Frischmilchuntersuchung in einem Bereich liegt, der auch von anderen Conjoint-Analysen für Objekte erreicht wird, mit denen sich die Probanden intensiver auseinandersetzen als mit Frischmilch.

6 Zusammenfassung

Der Beitrag beschäftigt sich mit der Frage, ob die durch die Conjoint-Analyse geschätzten individuellen Präferenzen der Befragten mit ihren tatsächlichen Präferenzen übereinstimmen. Diese Frage betrifft die Überprüfung der Validität der geschätzten individuellen Präferenzen. Zunächst wird der Begriff Validität definiert und dargestellt, wie unterschiedliche Validitätsarten bei der Conjoint-Analyse überprüft werden können. Zwei wichtige Validitätsarten sind die interne Validität und die Prognosevalidität. Diese beiden Validitätsarten werden für eine Conjoint-Analyse für das Produkt Frischmilch überprüft. Die interne Validität, die sich auf den Vergleich der geschätzten Präferenzen mit den der Schätzung zugrunde liegenden Präferenzen bezieht, ist zufriedenstellend und entspricht den Werten der internen Validität anderer Conjoint-Analysen für Nahrungsmittel. Die Prognosevalidität, die sich auf die Fähigkeit der Conjoint-Analyse bezieht, das tatsächliche Wahlverhalten zu prognostizieren, wurde mit drei unterschiedlichen Ansätzen gemessen. Anhand der drei Ansätze ist es möglich, auf Ursachen für die Abweichung der geschätzten Präferenzen vom tatsächlichen Wahlverhalten der Befragten zu schließen. Die Werte der ermittelten Prognosevalidität und der Prognosevalidität anderer Conjoint-Analysen für Nichtnahrungsmittel sind ähnlich. Da die Conjoint-Analyse vor allem ein Modell zur Prognose des Wahlverhaltens ist, ist es notwendig, neben der internen Validität auch die Prognosevalidität zu überprüfen.

Summary

The paper deals with the question whether the estimated individual preferences of the interviewees in the conjoint analysis correspond with their actual preferences. This question covers the verification of the validity of the estimated individual preferences. First of all the term validity is defined and the ways in which different forms of validity can be tested in the conjoint analysis are presented. The internal validity and the predictive validity are two important validity forms. For a conjoint analysis of fresh milk these two validity forms are tested. The internal validity, which shows the comparison of the estimated preferences versus the stated preferences that are used for the estimation, is satisfactory and coincides with the values of internal validity from other conjoint analyses concerning food. The predictive validity, which shows the ability of conjoint analysis to predict the actual choice behaviour, is measured using three different approaches. These three approaches allow to identify the reasons for the difference between the estimated preferences and the actual choice behaviour of the interviewees. The results of the predictive validity here and the predictive validity of other conjoint analyses that deal with non-food products are similar. Above all the conjoint analysis is a model for predicting choice, therefore it is important to test the predictive validity as well as the internal validity.

Literaturverzeichnis

- ADDELMAN, S. (1962): *Orthogonal main-effect plans for asymmetrical factorial experiments*. In: *Technometrics* 4, No. 1, Feb. 1962, S. 21-46
- AKAAH, J.P.; KORGAONKAR, P.K. (1983): *An empirical comparison of the predictive validity of self-explicated, Huber-hybrid, traditional conjoint and hybrid conjoint models*. In: *JoMR* 20, May 1983, S. 187-197
- BACKHAUS, K.; ERICHSON, B.; PLINKE, W.; WEIBER, R. (1994): *Multivariate Analysemethoden*. Eine anwendungsorientierte Einführung. 7. Aufl., Berlin u.a.O.
- BÖHLER, H.(1992): *Marktforschung*. 2. Aufl., Stuttgart u.a.O.
- BORTZ, J.; DÖRING, N. (1995): *Forschungsmethoden und Evaluation*. 2. Aufl., Berlin u.a.O.
- GREEN, P.E.; SRINIVASAN, V.(1978): *Conjoint analysis in consumer research: issues and outlook*. In: *Journal of Consumer Research (JoCR)* 5, Sept. 1978, S. 103-123
- GREEN, P.E.; SRINIVASAN, V. (1990): *Conjoint analysis in marketing: new developments with implications for research and practice*. In: *Journal of Marketing (JoM)*, 54, Oct. 1990, S. 3-19
- HAMMANN, P.; ERICHSON, B. (1994): *Marktforschung*. 3. Aufl., Stuttgart u.a.O.
- HOSSINGER, H.-P. (1982): *Pretests in der Marktforschung*. Würzburg, Wien
- LIENERT, G.A. (1969): *Testaufbau und Testanalyse*. 3. Aufl., Weinheim u.a.O.
- MÜLLER-HAGEDORN, L.; SEWING, E.; TOPOROWSKI, W. (1993): *Zur Validität von Conjoint-Analysen*. In: *zfbf* 45, H. 2, S. 123-148
- PETER, J.P. (1981): *Construct validity: a review of basic issues and marketing practices*. In: *Journal of Marketing Research (JoMR)* 18, May 1981, S. 133-145
- SATTLER, M. (1991): *Herkunfts- und Gütezeichen im Kaufentscheidungsprozeß*. Stuttgart
- SCHWEIKL, M. (1985): *Computergestützte Präferenzanalyse mit individuell wichtigen Produktmerkmalen*. Schriften zum Marketing, Bd. 17, Berlin
- SHELUGA, D.A.; JACCARD, J.; JACOBY, J. (1979): *Preference, search and choice: an integrative approach*. In: *JoCR* 6, Sept. 1979, S. 166-176
- SIEGEL, S. (1956): *Nonparametric statistics for the behavioral sciences*. New York u.a.O.
- SRINIVASAN, V. (1988): *A conjunctive-compensatory approach to the self-explication of multiattributed preferences*. In: *Decision Sciences* 19, No. 2, Spring 1988, S. 295-305

- SRINIVASAN, V.; SHOCKER, A.D.(1973): *Linear programming techniques for multidimensional analysis of preferences*. In: Psychometrika 38, No. 3, Sept. 1973, S. 337-369
- TROEGEL, T.; ROTHSPRACH, M. (1996): Die Conjoint-Analyse als dekompositionelles Verfahren zur Nutzenmessung und Präferenzanalyse. In: Kirschke, M.; Odening, M.; Schade, G. (Hrsg.): *Agrarstrukturentwicklung und Agrarpolitik*, Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Bd. 32, S. 527-537
- WEISENFELD, U. (1989): *Die Einflüsse von Verfahrensvariationen und der Art des Kaufentscheidungsprozesses auf die Reliabilität der Ergebnisse bei der Conjoint Analyse*. Berlin 1989
- WITTINK, D.R.; VRIENS, M.; BURHENNE, W. (1994): *Commercial use of conjoint analysis in Europe: results and critical reflections*. In: International Journal of Research in Marketing 7, 11, S.41-52