



ARBEITSPAPIER  
WORKING PAPER

Nr. 45  
2007

Maurice Eschweiler  
Heiner Evanschitzky  
David Woisetschläger

**Laborexperimente in der Marketingwissenschaft -  
Bestandsaufnahme und Leitfaden bei  
varianzanalytischen Auswertungen**

Herausgeber:  
Förderkreis für Industriegütermarketing e.V.  
an der Westfälischen Wilhelms-Universität Münster  
c/o Prof. Dr. Dr. h.c. Klaus Backhaus  
Am Stadtgraben 13-15, 48143 Münster  
Tel.: 0251-83-22861  
Fax: 0251-83-22903

**Dr. Maurice Eschweiler**

c/o Marketing Centrum Münster  
Institut für Anlagen und Systemtechnologien  
Am Stadtgraben 13-15  
48143 Münster  
Tel.: 0251-83-29927  
Fax: 0251-83-22903  
e-mail: [maurice.eschweiler@uni-muenster.de](mailto:maurice.eschweiler@uni-muenster.de)

**Dr. Heiner Evanschitzky**

c/o Marketing Centrum Münster  
Lehrstuhl für Betriebswirtschaftslehre, insb. Distribution und Handel  
Am Stadtgraben 13-15  
48143 Münster  
Tel.: 0251-83-22036  
Fax: 0251-83-22032  
e-mail: [evanschitzky@ifhm.de](mailto:evanschitzky@ifhm.de)

**Dr. David Woissetschläger**

c/o Lehrstuhl für Marketing, Universität Dortmund  
Otto-Hahn-Str. 6  
44227 Dortmund  
Tel.: 0231-755-4611  
Fax: 0231-755-3271  
e-mail: [d.woissetschlaeger@wiso.uni-dortmund.de](mailto:d.woissetschlaeger@wiso.uni-dortmund.de)

**Danksagung:**

Die Autoren danken John Deighton (*Journal of Consumer Research*), Joel Huber (*Journal of Marketing Research*) und Roland T. Rust (*Journal of Marketing*) für ihre Unterstützung und ihre wertvollen Kommentare im Rahmen der Erstellung dieses Manuskripts.

## **Inhaltsverzeichnis**

1	Einleitung.....	1
2	Bedeutung von Experimenten in der Marketingforschung im internationalen Vergleich.....	2
3	Leitfaden für den Einsatz von Laborexperimenten mit varianzanalytischen Auswertungen.....	5
3.1	Vorbereitung eines Laborexperiments.....	5
3.1.1	Formulierung des Forschungsproblems.....	5
3.1.2	Festlegung des Forschungsdesigns.....	6
3.2	Durchführung eines Laborexperiments.....	8
3.3	Varianzanalytische Auswertung von Laborexperimenten.....	8
3.3.1	Manipulation Checks.....	8
3.3.2	Festlegung des problemadäquaten Typs der Varianzanalyse.....	9
3.3.3	Prüfung der Verfahrensprämissen.....	10
3.4	Dokumentation.....	15
4	Fazit.....	17



## 1 Einleitung

Neben deskriptiven und explorativen Studien ist das kausale Design eine dritte Grundform von Studien, die im Marketing große Bedeutung erlangt haben. Kausale Studien verfolgen das Ziel, Ursache-Wirkungs-Beziehungen zwischen Variablen zu analysieren. Demnach soll herausgefunden werden, ob eine Änderung einer unabhängigen Variablen ursächlich für eine Änderung einer abhängigen Variablen ist. Eine Möglichkeit zur empirischen Prüfung solcher Kausalzusammenhänge besteht in der Durchführung eines kontrollierten Experiments, bei dem der Forscher aktiv durch systematische Variation der unabhängigen Variablen in den untersuchten Prozess eingreift (*Crano/Brewer, 2002*). Experimente stellen die strengste Form der Überprüfung kausaler Hypothesen dar und sind primär dazu geeignet, Theorien bzw. die aus diesen resultierenden Hypothesen zu überprüfen. Nicht zuletzt aufgrund dieses großen wissenschaftlichen Potenzials stellt sich die Frage, welche Bedeutung experimentelle Studien inzwischen in der deutschsprachigen Marketingforschung gerade im Vergleich mit der internationalen Spitzenforschung im Marketing aufweisen. Ein erstes Ziel dieses Beitrags besteht daher darin, eine **Bestandsaufnahme zur Verbreitung von Experimenten in der deutschsprachigen Marketingwissenschaft** durchzuführen und durch Experteninterviews Gründe für eine im hiesigen Raum womöglich noch vorhandene Zurückhaltung gegenüber diesem Marketingforschungsinstrument zu identifizieren (vgl. Abschnitt 2).

Die überwiegende Anzahl bislang publizierter Experimente hat unter Laborbedingungen stattgefunden. Anders als beim Feldexperiment findet ein Laborexperiment unter künstlichen Bedingungen statt, und i.d.R. wissen die Versuchspersonen, dass sie Teil eines Experiments sind. Auch für die Zukunft ist aufgrund der Vorteile dieses Verfahrens (z.B. sehr gute Möglichkeit zur Kontrolle von Störgrößen oder Zeit- und Kostenvorteile) mit einer weiten Verbreitung zu rechnen. Das bei der Auswertung solcher Laborexperimente am häufigsten verwendete statistische Verfahren ist die Varianzanalyse. Grundsätzliche Aufgabe der Varianzanalyse ist es, den nicht-zufälligen bzw. systematischen Einfluss einer oder mehrerer nominal skalierten unabhängiger Variablen – auch als Faktoren bezeichnet – gemäß dem Prinzip der Streuungszerlegung auf eine oder mehrere metrisch skalierte abhängige Variable(n) zu untersuchen. Trotz der gerade im Marketing herausragenden Bedeutung von Laborexperimenten mit varianzanalytischen Auswertungen (*Leiberich 1995*) bestehen zuweilen auch heute noch – nicht nur bei Marketingforschern aus dem deutschsprachigen Raum – mehr oder weniger große Defizite bei der Anwendung dieser Marke-

tingforschungsart. Weit verbreitete Fehler bzw. Nachlässigkeiten in der einschlägigen Literatur bestehen z.B. darin, dass

- der konfirmatorische Charakter von (Labor-)Experimenten verkannt wird,
- die Faktorstufen falsch festgelegt werden,
- die Test- und Kontrollgruppen zu wenige Probanden aufweisen,
- wichtige Kovariablen nicht beachtet werden,
- keine Manipulation Checks durchgeführt werden,
- die eigentliche Datenerhebungsphase des Laborexperiments fehlerbehaftet ist,
- die umfangreichen Anwendungsvoraussetzungen der Varianzanalyse gar nicht geprüft werden,
- ein nicht Ziel führender Typ der Varianzanalyse gewählt wird,
- gravierende Fehler bei der Reihenfolge zu interpretierender Effekte gemacht
- oder das Vorgehen unzureichend dokumentiert wird.

Infolge solcher Fehler (es wird an dieser Stelle bewusst darauf verzichtet, fehlerbehaftete Quellen aufzuführen) sind gravierende inhaltliche Falschinterpretationen wahrscheinlich. Von daher besteht ein zweites Ziel dieses Beitrags darin, dem Leser möglichst umfassende **Leitlinien für die Anwendung von varianzanalytisch ausgerichteten Laborexperimenten** zu geben und so einen Beitrag für eine weitere Verbreitung korrekt durchgeführter (labor-) experimenteller Marketingforschung zu leisten (vgl. Abschnitt 3). Dabei soll der Fokus an den entsprechenden Stellen weniger auf statistischen bzw. methodischen Einzelheiten als vielmehr auf einem klaren Anwendungsbezug liegen.

## **2 Bedeutung von Experimenten in der Marketingforschung im internationalen Vergleich**

Um die Bedeutung von (varianzanalytisch ausgerichteten) Laborexperimenten in der Marketingforschung abschätzen zu können, wurde eine zweistufige empirische Untersuchung durchgeführt. Im ersten Schritt wurde die Häufigkeit von solchen Experimenten in den drei weltweit angesehensten Marketing-Journals, *Journal of Marketing* (JM), *Journal of Marketing Research* (JMR) und *Journal of Consumer Research* (JCR) über einen Zeitraum von 11 Jahren (1996-2006) ermittelt. Diese Ergebnisse wurden mit der Anzahl der im wichtigsten deutschen Marketing-Journal, *Marketing ZFP* (MZFP), erschienenen vergleichbaren Experimente verglichen. Dabei wurden insgesamt je 1.523 Artikel von zwei unabhängigen Bewertern (*Rater*) darauf hin untersucht, ob es sich bei dem Artikel um eine experimentelle Studie handelt [1].

	MZFP		JM		JMR		JCR	
	absolut	prozentual	absolut	prozentual	absolut	prozentual	absolut	prozentual
1996	5	31,25%	19	45,24%	18	31,03%	40	64,52%
1997	2	12,50%	3	6,25%	15	27,27%	46	70,77%
1998	5	19,23%	2	5,13%	12	29,27%	54	71,05%
1999	2	12,50%	8	25,81%	15	41,67%	33	73,33%
2000	2	9,52%	4	14,29%	14	37,84%	26	55,32%
2001	2	12,50%	2	8,00%	14	35,00%	26	56,52%
2002	6	31,58%	2	8,00%	14	35,90%	24	66,67%
2003	3	15,00%	5	16,67%	13	32,50%	18	75,00%
2004	2	10,53%	6	20,69%	11	28,95%	17	65,38%
2005	2	10,00%	1	4,35%	6	12,77%	19	57,58%
2006	0	0,00%	8	26,67%	7	20,00%	15	57,69%
Summe	31		60		139		318	
Gesamt	209	14,83%	362	16,57%	466	29,83%	486	65,43%

Tab. 1: Anzahl Beiträge mit laborexperimenteller Marketingforschung in Top-Journals

Die Auswertung zeigt zunächst, dass es sowohl zwischen dem deutschsprachigen und den internationalen Marketing-Journals als auch innerhalb der Gruppe der internationalen Journals große Unterschiede bzgl. der Anzahl der publizierten varianzanalytisch ausgerichteten Experimente gibt. Während im Durchschnitt etwa 40% aller in den internationalen Journals erschienen Studien von solch experimenteller Natur sind, liegt die MZFP mit knapp 15% deutlich dahinter. Auffällig ist, dass das JCR mit etwa 65% laborexperimenteller Studien deutlich an der Spitze liegt, während das eher auf *General Marketing* fokussierte JM unterdurchschnittlich viele solcher Studien publiziert.

In einem zweiten Schritt wurde eine halbstrukturierte Befragung der Herausgeber der drei internationalen Marketing-Journals durchgeführt [2]. Insgesamt wurden dabei die folgenden Grundfragen angesprochen: (1) Wie wird die Wichtigkeit der experimentellen Marketingforschung im Vergleich zu anderen Arten der empirischen Marktforschung eingeschätzt und welche Trends lassen sich erkennen? (2) Werden die bei Top-Journals eingereichten Studien formal gemäß etablierter Kriterien geprüft (z.B. Prämissentests) und wie werden Validität und Reliabilität überprüft? (3) Konfrontiert mit den Ergebnissen unserer Auszählung: Welche Gründe sprechen dafür, dass die Anzahl der publizierten Experimente in den USA auffällig höher ist als in Deutschland?

Insgesamt herrscht bei der Beantwortung der drei Fragen große Einigkeit zwischen den Herausgebern. Sie stellen fest, dass die **experimentelle Marketingforschung zur Überprüfung von Kausalzusammenhängen deutlich an Bedeutung gewinnen** wird. Dabei wird neben der bereits hohen Bedeutung von Laborexperimenten auch eine zunehmende Relevanz von Realexperimenten erwartet. Ebenso wird angemerkt, dass die Kombination

mehrerer Methoden in Zukunft an Bedeutung gewinnen wird. Diese Auffassung vertritt auch Winer (1999), der insbesondere die Kombination von Experimenten und Analyse von Verhaltensdaten als zukunftsweisend einschätzt.

Interessanterweise werden eingereichte Manuskripte offenbar nicht formal auf die Erfüllung von Prämissen geprüft, die zeigen, ob man u.a. varianzanalytische Verfahren überhaupt anwenden kann. Es obliegt den Gutachtern, ob ein Nachweis dessen im Verlauf des Begutachtungsprozesses von den Autoren gefordert wird. Man erkennt daran, dass auch im internationalen Bereich durchaus Verbesserungspotenzial bei der Durchführung und Auswertung von Experimenten besteht. Um die Validität der empirischen Ergebnisse zu überprüfen, wird in den internationalen Journals in der Regel verlangt, mehrere Studien in einem Paper durchzuführen. Ebenso wurde angemerkt, dass die Ablehnungsquote für experimentelle Studien in etwa genau so hoch sei, wie die aller übrigen Einreichungen.

Die Editoren der internationalen Journals sehen die offensichtliche Bedeutung experimenteller Studien für die drei Top Marketing-Journals im Vergleich zur Bedeutung in beispielsweise der MZFP im Wesentlichen darin begründet, dass bei der **Ausbildung US-amerikanischer Wissenschaftler bzw. Doktoranden großer Wert auf experimentelle Methoden** gelegt wird und dass die Konsumentenverhaltensforschung, die traditionell ein wesentliches Einsatzgebiet von Experimenten darstellt, generell einen sehr hohen Stellenwert aufweist. Es wird insbesondere in den USA durchaus in Kauf genommen, dass die praktische Relevanz der Ergebnisse weniger direkt erkennbar ist als die wissenschaftliche Rigorosität. Demgegenüber scheint bei der Ausbildung deutscher Nachwuchswissenschaftler mehr Wert auf die praktische Relevanz und die Anwendbarkeit der Forschungsergebnisse in der Marketingpraxis gelegt zu werden (vergleiche hierzu die Diskussion zwischen Rigor und Relevanz, z.B. in *Ahlert/Evanschitzky/Hesse 2005; Kieser/Nicolai, 2005; Varadarajan, 2003*).

Zusammenfassend kann gemutmaßt werden, dass die geringere Bedeutung von Experimenten in der deutschsprachigen Marketingforschung im Wesentlichen durch zwei Aspekte begründet ist: Erstens wurde (und wird) in der deutschsprachigen Marketingforschung größerer Wert auf die praktische Anwendbarkeit der Forschungsergebnisse gelegt. Von der Tradition her hat die Erforschung des Konsumentenverhaltens einen weniger zentralen Stellenwert in der Marketingforschung (*Gröppel-Klein/Weinberg 2000*). Allerdings ist für die Zukunft zu vermuten, dass sich die deutschsprachige Marketingforschung noch stärker an der angloamerikanischen Spitzenforschung orientieren wird [3]. Zweitens scheint in der deutschsprachigen Marketingforschung insgesamt das Wissen um die korrekte Durchfüh-

rung von Experimenten weniger stark in die Ausbildung der Nachwuchswissenschaftler integriert zu sein.

### **3 Leitfaden für den Einsatz von Laborexperimenten mit varianz-analytischen Auswertungen**

Vor dem Hintergrund des aufgeführten Nachholbedarfs soll im Folgenden ein Versuch unternommen werden, das Wissen um den korrekten Einsatz und die richtige Durchführung von Marktforschungsexperimenten – und hier im speziellen von Laborexperimenten – durch einen möglichst umfassenden Leitfaden auszuweiten. Dazu wird ein vierstufiges Vorgehen bei Experimenten vorgeschlagen, im Rahmen dessen die Phasen der Vorbereitung, Durchführung, Analyse und Dokumentation unterschieden werden.

#### **3.1 Vorbereitung eines Laborexperiments**

##### **3.1.1 Formulierung des Forschungsproblems**

Die wichtigsten Ziele eines wissenschaftlichen Vorhabens sind der Erkenntnisgewinn und/oder die Gestaltungsempfehlung. Während beim Erkenntnisgewinn die Orientierung an der Wahrheit vorliegt, orientiert sich das Gestaltungsziel an der Maxime der Nützlichkeit. Grundsätzlich lässt sich konstatieren, dass bei Laborexperimenten der Fokus eher auf dem Erkenntnisziel liegt, da Laborsituationen mehr oder weniger stark von denen in der Realität abweichen. In jedem Fall **muss der Forscher eine theoretisch wohl begründete Vermutung** haben, welche Zusammenhänge zwischen Variablen bestehen. Diese Zusammenhänge werden in der Regel in Form von Hypothesen formuliert. Ziel eines Laborexperiments ist es demnach, theoretisch vermutete Zusammenhänge zwischen Variablen zu überprüfen. Dies geschieht dadurch, dass (mindestens) eine Gruppe von Probanden (Testgruppe) mit einer Manipulation (der experimentellen Situation) konfrontiert wird, während eine andere Gruppe (Kontrollgruppe) dieser Manipulation nicht ausgesetzt wird. Dies bedeutet auch, dass bereits im Zuge der Formulierung des Forschungsproblems unabhängige und abhängige Variablen eindeutig festgelegt sind. Dabei ist zu konstatieren, dass häufig weniger die direkten Effekte der unabhängigen Variablen auf, die abhängigen Variablen von Interesse sind, da diese Effekte zumeist aus konzeptioneller Sicht kaum kontrovers – weil intuitiv naheliegend – sind. Vielmehr gilt das Augenmerk der Marketingforschung in vielen Fällen den Interaktionseffekten, d.h. der Frage, durch welche Faktoren andere Faktoren in ihrer Wirkung moderiert, d.h. verstärkt oder abgeschwächt werden. Einem Expe-

riment müssen daher stets intensive konzeptionelle Vorüberlegungen über diese Zusammenhänge vorhergehen.

### 3.1.2 Festlegung des Forschungsdesigns

Stehen die Forschungsfragen in Form eines Sets theoriebasierter Hypothesen fest, so sind im Forschungsdesign zunächst die abhängigen Variablen zu operationalisieren und anschließend die Stimuli zur Operationalisierung der Faktoren (unabhängige Variablen) und die jeweiligen Faktorstufen, der etwaige Einbezug von Kovariablen sowie die Auswahl und Größe der Stichprobe festzulegen.

Häufig besteht bezüglich der genauen **Ausprägungen der Faktoren** Unsicherheit, da diese Ausprägungen so festgelegt werden müssen, dass sie auch in beabsichtigter Weise wahrgenommen werden. Für eine erste Überprüfung der unterschiedlichen Wahrnehmung der jeweiligen Faktorstufen sollten daher **Pretests** durchgeführt werden. Diese sind auch aus forschungsökonomischen Gründen zu empfehlen, da die Stichprobe i.d.R. überschaubar ist und in Konsequenz eine größere Sicherheit darüber gewonnen wird, dass die Faktorstufen im Experiment voneinander diskriminant sind. Die Probandenauswahl für solch einen Pretest sollte von der Struktur her mit der Auswahl im Rahmen des eigentlichen Experiments möglichst identisch sein, um Verzerrungen aufgrund einer unterschiedlichen Stichprobenzusammensetzung zu vermeiden. Um ein Lernen der Probanden zu verhindern, müssen die Probanden, die am Pretest teilgenommen haben, vom eigentlichen Experiment ausgeschlossen werden. Als Ergebnis des Pretests werden diejenigen Faktorstufen aus der weiteren Untersuchung ausgeschlossen, die von den Probanden als nicht unterschiedlich wahrgenommen werden. Eine Entscheidung für oder gegen die jeweilige Faktorstufe hat dann vor dem Hintergrund der verfolgten Untersuchungsziele zu erfolgen.

Der Einbezug von **Kovariablen** empfiehlt sich immer dann, wenn die vor einer Untersuchung angetroffenen A-Priori-Unterschiede zwischen den Probanden bei den abhängigen Variablen das Untersuchungsergebnis nicht beeinflussen sollen. Der zentrale Vorteil besteht darin, dass die ggf. vor der Untersuchung (weiterhin) bestehenden Unterschiede bei den Kovariablen nun herauspartialisiert werden können, d.h., der Einfluss der Kovariablen auf die abhängigen Variablen kann neutralisiert werden (*Herrmann/Seilheimer 2000*).

Nachdem durch die Festlegung der Stimuli sowie der abhängigen Variablen und Kovariablen das Design des Experimentes fixiert ist, muss in einem letzten Schritt die **Auswahl und Größe der Stichprobe** festgelegt werden. Bei der Auswahl der Probanden muss darauf geachtet werden, dass eine größtmögliche Homogenität über alle Zellen erreicht wird, da sonst die auf die unterschiedlichen Gruppenzusammensetzungen in den Zellen zurück-

gehende zusätzliche (und nicht kontrollierte) Varianz die Analyse beeinflusst. Daher werden in vielen experimentellen Studien vorwiegend Probanden mit gleichem Hintergrund und ähnlicher Biographie ausgewählt. Wird von der Heterogenität der Probanden ein Einfluss auf die abhängigen Variablen erwartet, so bietet sich die Berücksichtigung der Heterogenität als Kovariable oder alternativ als unabhängige Variable (und somit als zusätzliche Manipulation) an.

Die absolute Mindestgröße der Stichprobe zur Erlangung von belastbaren Ergebnissen wird in der Literatur mit einer Zellengröße angegeben, welche die Anzahl der abhängigen Variablen zumindest übersteigt, empfohlen wird jedoch mindestens eine Zellgröße von 20 Probanden (Hair et al. 2006). In diesem Zusammenhang ist es ratsam, die Effektgröße (vgl. Abschnitt 3.3.4) zu antizipieren, um abhängig davon die geeignete Zellgröße zu planen. Insbesondere kleine Effekte erfordern mit zunehmender Anzahl an abhängigen Variablen eine deutlich größere Anzahl an Probanden pro Zelle, um überhaupt statistisch signifikante Ergebnisse erzielen zu können. In Tab. 2 sind beispielhaft Empfehlungen für Zellgrößen in Abhängigkeit der erwarteten Effektstärke (vgl. Abschnitt 3.3.4 für diesbezüglich konkrete Werte bei verhaltenswissenschaftlich ausgerichteten Experimenten) und Anzahl der abhängigen Variablen aufgeführt (ausführlich bei Läuter 1978). Dabei ist jedoch zu beachten, dass insbesondere bei sehr großen Zellen (z.B. 400 Probanden und mehr) selbst sehr kleine Gruppenunterschiede statistisch signifikant werden. Es sollte demnach nicht primäres Ziel des Forschers sein, über alle Maßen viele Probanden für sein Experiment zu gewinnen. Liegt dennoch solch ein Fall vor, ist es wichtig, die dann fast zwangsläufig signifikanten Effekte auf eine „praktische Signifikanz“ zu prüfen.

Effektstärke	Anzahl der Gruppen											
	3				4				5			
	Anzahl der abhängigen Variablen											
	2	4	6	8	2	4	6	8	2	4	6	8
sehr groß	13	16	18	21	14	18	21	23	16	21	24	27
groß	26	33	38	42	29	37	44	46	34	44	52	58
mittel	44	56	66	72	50	64	74	84	60	76	90	100
klein	98	125	145	160	115	145	165	185	135	170	200	230

Tab. 2: Empfohlene Zellgrößen

Von der Größe des Samples und der Gruppenanzahl hängt schließlich auch die maximal einzubeziehende Anzahl an Kovariablen ab (Hair et al. 2006):

$$\text{Maximale Anzahl an Kovariablen} = (0,1 \times \text{Stichprobengröße}) - (\text{Anzahl der Gruppen} - 1)$$

## 3.2 Durchführung eines Laborexperiments

Bei der Durchführung von Laborexperimenten sind zunächst Überlegungen zur **Kontrolle von Störvariablen** anzustellen, so dass eine möglichst hohe interne Validität gewährleistet werden kann. Etwaige Fehler führen dazu, dass die Unterschiede zwischen den Gruppen nicht mehr mit hoher Sicherheit auf die erfolgten Manipulationen zurückgeführt werden können. Im Einzelnen ist darauf zu achten, dass die Zuteilung der Probanden auf die Gruppen zur Sicherstellung der Unabhängigkeit jedes einzelnen Probanden von anderen Versuchsteilnehmern randomisiert erfolgt (*Glaser 1978*) und der Einfluss der Forscher möglichst gering ausfällt und konstant gehalten wird (*Venkatesan 1967*). Ebenso muss der Effekt der Untersuchungssituation (Hawthorne Effekt) bspw. durch eine Kontrollgruppe berücksichtigt werden. Darüber hinaus müssen alle Variablen reliabel – bspw. in Form etablierter Skalen – gemessen werden (*Homburg/Giering 1996*) sowie Attritionseffekte in den Gruppen bspw. durch Vergleich der Missing Values kontrolliert werden.

Werden im Experiment fiktionale Szenarien präsentiert, die unter Umständen sogar real existierenden Personen oder Unternehmen bestimmte Verhaltensweisen unterstellen, ist nach dem „Code of Conduct“ (bspw. wie im Code of Ethics and Conduct der British Psychological Society dargelegt, o. V. 2006) im Anschluss an das Experiment eine Aufklärung der Probanden über den fiktionalen Charakter der Untersuchung sicherzustellen, um so ethischen Ansprüchen genügen zu können.

Da die mit der varianzanalytischen Auswertung verbundenen Prämissen in der Forschungspraxis regelmäßig verletzt werden, sollte schon bei der Durchführung von Laborexperimenten darauf geachtet werden, dass **jede Zelle** in etwa **gleichgroß besetzt** ist (vgl. Abschnitt 3.3.3).

## 3.3 Varianzanalytische Auswertung von Laborexperimenten

### 3.3.1 Manipulation Checks

Bevor untersucht wird, *wie* sich die unabhängigen Variablen unter Berücksichtigung etwaiger moderierender Variablen auf die abhängigen Variablen auswirken, muss grundsätzlich sichergestellt werden, **ob die unabhängigen Variablen überhaupt in vorgesehener Weise in der Stichprobe realisiert** sind (*Perdue/Summers 1986*). Diese Kontrolle entspricht dem sog. Manipulation Check und versichert, dass die auf den abhängigen Variablen festgestellten Merkmalsausprägungen auch auf die unterschiedlichen Faktorstufen der unabhängigen Variablen zurückzuführen sind (*Bortz/Döring 2006*). Dazu bedarf es geeig-

neten inhaltlicher Skalen, anhand derer die Manipulation untersucht werden kann und sich die bei den Untersuchungsgruppen gefundenen Ausprägungen ins Verhältnis setzen lassen. Manipulation Checks sind somit über den Pretest hinaus als Instrumente zur Sicherung der Qualität der empirischen Analyse zu betrachten. Eine Prüfung ist insbesondere dann erforderlich, wenn es sich bei den unabhängigen Variablen nicht um dichotome Merkmale, sondern um Variablen handelt, deren Manipulation individuell nicht unterschiedlich wahrgenommen werden darf. Liegt z.B. der Fall vor, dass eine vom Versuchsleiter als Manipulation angedachte hohe und niedrige Differenz zwischen einem Externen Referenz- und einem Verkaufspreis gar nicht unterschiedlich hoch empfunden wird, wären etwaige Gruppenunterschiede bei den Wirkungskriterien nicht auf diese Manipulation zurückzuführen und weitere diesbezügliche Analysen unzulässig. Bei dichotomen Merkmalen (z.B. vergleichende und nicht-vergleichende Werbung) ist eine Manipulationsprüfung nicht Ziel führend, da der Versuchsleiter keinen Fehler bei der Manipulation der unabhängigen Variablen begehen kann. Auch wenn beim Nachfrager Wahrnehmungsverzerrungen vorliegen sollten, kann eine vergleichende Werbung eben keine nicht-vergleichende Werbung sein. Etwaige Wahrnehmungstäuschungen können in diesem Fall in Kauf genommen werden.

### 3.3.2 Festlegung des problemadäquaten Typs der Varianzanalyse

Je nach Zahl der einbezogenen unabhängigen und abhängigen Variablen, lassen sich verschiedene Typen der Varianzanalyse unterscheiden. Bei lediglich einer abhängigen Variablen ist eine **ANOVA** (ANalysis Of VAriance) oder auch univariate Varianzanalyse durchzuführen. Bei zwei oder mehr abhängigen Variablen kann es sinnvoll sein, eine multivariate Varianzanalyse (**MANOVA** = Multivariate ANalysis Of VAriance) anzuwenden. Der Vorteil einer MANOVA gegenüber einer ANOVA besteht darin, dass hier die Unterschiede des *Vektors der Mittelwerte* (Variate) aller abhängigen Variablen in den durch die Faktorausprägungen gebildeten Gruppen betrachtet werden, während die ANOVA die Unterschiede der *Mittelwerte* der abhängigen Variablen in den entsprechenden Gruppen untersucht (Hair et al. 2006). Mit anderen Worten können bei der MANOVA die Gruppenunterschiede simultan über mehrere abhängige Variablen analysiert werden. Der Einsatz einer MANOVA ist im Vergleich zum Einsatz mehrerer unabhängiger ANOVAs immer dann Ziel führend, wenn die abhängigen Variablen miteinander korreliert sind, da nur dann auch die Wahrscheinlichkeit, einen  $\alpha$ -Fehler (eine wahre Nullhypothese wird fälschlicherweise abgelehnt) zu begehen, kontrolliert werden kann. Sind die abhängigen Variablen nicht miteinander korreliert, führen beide Vorgehensweisen zu gleichlautenden Ergebnissen (Huberty/Morris 1989; Huberty 1994).

Eine Möglichkeit, die durch die Faktoren nicht erklärte Varianz zu reduzieren und weitere die abhängigen Variablen womöglich beeinflussende Variablen zu ermitteln, besteht wie bereits beschrieben darin, Kontrollvariablen oder kurz Kovariablen in die Untersuchung zu integrieren (Bortz 2004). Die Varianzanalyse (MANOVA bzw. ANOVA) wird in diesem Fall auf sog. Kovarianzanalysen (**MANCOVA** = Multivariate ANalysis of COVariance bzw. **ANCOVA**) erweitert. Verfahrenstechnisch verbirgt sich hinter der Kovarianzanalyse eine Kombination von varianz- mit regressionsanalytischen Techniken. Indem mithilfe einer vorgeschalteten Regressionsanalyse zunächst der Einfluss der Kontrollvariablen auf die abhängigen Variablen dargestellt wird, gelingt es, die nicht erklärte Varianz zu verkleinern und damit die interessierenden Faktorwirkungen im Rahmen der Varianzanalyse deutlicher (d.h. in der Regel stärker) zu machen.

### 3.3.3 Prüfung der Verfahrensprämissen

Vor der Anwendung einer (multivariaten) (Ko-)Varianzanalyse ((M)AN(C)OVA) ist es notwendig, die verschiedenen Anwendungsvoraussetzungen dieser Methode zu testen. Es erscheint sinnvoll, diese Prämissenprüfung zweistufig durchzuführen, indem zunächst die Voraussetzungen einer multivariaten Varianzanalyse (MANOVA) und – falls Kovariablen integriert werden – im Anschluss daran die besonderen Prämissen einer Kovarianzanalyse geprüft werden. Durch die Prüfung der Voraussetzungen einer MAN(C)OVA werden implizit auch die Prämissen der Follow-Up-AN(C)OVAs getestet.

#### 3.3.3.1 Prämissen der multivariaten Varianzanalyse (MANOVA)

Weil die Ergebnisse der MANOVA **sensitiv gegenüber Ausreißern** sind (Tabachnik/Fidell 2006), sollten Probanden, die auf offenen Skalen offensichtlich unsinnige Werte angegeben haben, eliminiert werden. Eine solche Einschätzung unterliegt stets in hohem Maße einem gewissen Maß an Subjektivität, was jedoch nicht behoben werden kann.

Der zentrale Vorzug der Durchführung einer MANOVA gegenüber mehreren ANOVAs besteht darin, dass Korrelationen zwischen den abhängigen Variablen berücksichtigt werden können. Die Berechtigung einer MANOVA ist demnach an die Voraussetzung geknüpft, dass die abhängigen Variablen auch tatsächlich miteinander korreliert sind. In diesem Zusammenhang ist allerdings darauf zu achten, dass die **Korrelation der abhängigen Variablen** nicht zu stark ausfällt, da ansonsten die Prämissenverletzung der Multikollinearität und infolgedessen Verzerrungen in der Teststatistik der MANOVA drohen. Zur Prüfung von Multikollinearität bietet es sich an, für jede abhängige Variable die Toleranz zu ermitteln (Tabachnik/Fidell 2006). Fickel (2001) geht davon aus, dass das Ausmaß der

Multikollinearität ab einem Toleranzwert von weniger als 0,1 problematisch wird (*Fickel 2001; van Doorn 2004*).

Eine weitere wichtige Anwendungsvoraussetzung der MANOVA besteht darin, dass alle abhängigen Variablen eine zellenweise **multivariate Normalverteilung** aufweisen sollten. Multivariate Normalverteilung impliziert eine univariate Normalverteilung aller abhängigen Variablen, eine univariate Normalverteilung aller Linearkombinationen der abhängigen Variablen und eine multivariate Normalverteilung aller Teilmengen der abhängigen Variablen (*Stevens 2002*). Eine Verletzung dieser Prämisse kann ebenfalls zu einer Verzerrung der Teststatistiken bzw. des Signifikanzniveaus führen und somit falsche Ergebnisse liefern. Leider ist in den gängigen Statistik-Software-Programmen noch kein Test auf multivariate Normalverteilung enthalten, so dass nur eine behelfsmäßige Prüfung dieser Prämisse über einen zellenweisen Kolmogorov-Smirnov-Test auf univariate Normalverteilung erfolgen kann. Allerdings stellt eine univariate Normalverteilung lediglich eine notwendige, jedoch keine hinreichende Bedingung für eine multivariate Normalverteilung dar (*Bray/Maxwell 1985*).

Die Prämisse der **Varianzhomogenität** bedeutet, dass sich die Varianzen innerhalb der Stichproben bzw. Gruppen nicht signifikant unterscheiden dürfen. Inhaltlich bedeutet dies, dass sich die über den kontrollierten Effekt einer unabhängigen Variable hinausgehenden, nicht kontrollierten Einflussgrößen auf die abhängigen Variablen bis auf zufällige Schwankungen gleich auswirken (*Backhaus et al. 2006*). Eine geeignete Möglichkeit, diese Prämisse zu prüfen, besteht in der Anwendung des Levene-Tests. Dieser hat den Vorteil, nicht von der Voraussetzung einer Normalverteilung abzuhängen. Er prüft die Nullhypothese, dass die Varianz der abhängigen Variablen über alle Gruppen gleich ist (*Janssen/Laatz 2005*). In der Literatur wird mithilfe von Monte-Carlo-Simulationen darauf hingewiesen, dass die beiden festgestellten Prämissenverletzungen der Nicht-Normalverteilung und Varianzheterogenität geheilt werden können, wenn zwei spezifische Bedingungen erfüllt sind. Erstens ist darauf zu achten, dass die **Stichproben ausreichend groß** sind (vgl. Abschnitt 3.1.2; *Bray/Maxwell 1985*). Zweitens erweist sich eine MANOVA dann als besonders robust, wenn die Zellen mit gleich vielen Probanden besetzt sind. Um eine solche **Gleichbesetzung der Zellen** (*Perreault/Darden 1975*) für den Fall ungleich großer Zellen methodisch sauber zu erreichen, schlägt Glaser (1978) vor, aus den zu hoch besetzten Zellen Daten nach Zufall zu eliminieren. Auf diese Weise ist sichergestellt, dass eine MANOVA trotz der angesprochenen Prämissenverletzungen angewendet werden darf.

### 3.3.3.2 Prämissen der Kovarianzanalyse

Der Einsatz einer Kovarianzanalyse ist in einer Untersuchung, bei der der hypothesenprüfenden Varianzanalyse eine Regressionsanalyse zur Neutralisierung des Einflusses von ex ante definierten Kovariablen vorgeschaltet wird, ebenfalls an spezifische Voraussetzungen geknüpft [4]. So dürfen die Kovariablen durch die **experimentelle Anordnung** nicht beeinflusst werden (*Diehl* 1983). In diesem Fall wäre zwar eine Adjustierung der Werte in den abhängigen Variablen weiter möglich, aber inhaltlich unsinnig. Darüber hinaus müssen die Kovariablen auf **intervallskaliertem Niveau** vorliegen.

Eine effektive Bereinigung des Datensatzes um systematische Fehler bzw. eine nicht-zufällige Reduktion der Fehlervarianz ist des Weiteren nur möglich, wenn die abhängigen Variablen mit den Kovariablen korrelieren. Es empfiehlt sich daher, die Signifikanz der **Korrelation der abhängigen Variablen mit den Kovariablen** auf ein Niveau von  $p < 0,05$  zu prüfen (*Bortz* 2004).

Darüber hinaus ist darauf zu achten, dass **zwischen Kovariable und unabhängiger Variable (Faktoren) kein substanzieller Zusammenhang** besteht, da ansonsten die Effekte der unabhängigen auf die abhängigen Variablen reduziert werden und damit die Gefahr von Fehlinterpretationen besteht. Es muss demnach geprüft werden, inwiefern Kovariable(n) und Faktoren interagierend auf die abhängigen Variablen wirken. Dazu bietet es sich an, eine multiple Regressionsanalyse durchzuführen, in der die Wirkung der Kovariable(n) und der Faktoren als gemeinsame Prädiktoren auf die jeweilige abhängige Variable geprüft wird. Da die Faktoren lediglich auf nominalem Niveau vorliegen, müssen diese mit einer Dummy-Kodierung versehen werden. Von besonderem Interesse ist dabei, ob sich die multiplikativ zu bildenden Interaktionsterme von Faktoren und Kovariable(n) signifikant von Null unterscheiden. Sollte dies der Fall sein, läge entsprechend eine prämissenverletzende Interaktion von Kovariable und Faktor vor, so dass die Kovariable folglich nicht in das Modell aufgenommen werden darf.

Als weitere Voraussetzung für die Anwendung einer Kovarianzanalyse wird in der Literatur angeführt, dass die **Steigungen der vorgeschalteten Regressionen** in den einzelnen Gruppen **homogen** sein sollten (*Tabachnik/Fidell* 2006; *Litz* 2000). Allerdings führt Bortz (2004) eine Reihe von Studien auf, die belegen, dass eine Verletzung dieser Prämisse zumindest bei gleichgroßen Zellen zu keinen Verzerrungen der Teststatistiken hinsichtlich des  $\alpha$ -Fehlers und der Testschärfe führt (*Bortz* 2004; *Dretzke/Levin/Serlin* 1982). In diesem Sinne sieht Levy (1980) eine Kovarianzanalyse nur dann kontraindiziert, wenn die Regressionen heterogen sind, keine Multi-Normalverteilung vorliegt und die Zellen ungleich be-

setzt sind. Sind in einer Studie alle Gruppen gleich groß, erscheint es weiterhin gerechtfertigt, zur Hypothesenprüfung die MANCOVA als Verfahren einzusetzen. Allerdings ist für den Fall einer Verletzung der Homogenität der Regressionsteigungen auf eine inhaltliche Interpretation der vorgeschalteten Regressionen zu verzichten. Die Kovariable erfüllt dann den ausschließlichen Zweck, die Fehlervarianz in den einzelnen Zellen zu reduzieren.

Sämtliche Prämissen mit zugehöriger Prüfungsmethode und Ansätzen zur Heilung von Prämissenverletzungen sind im Überblick Tabelle 3 zu entnehmen.

	Prämisse	Prüfungsmethode	Verletzung heilbar über
MANOVA	Keine Ausreißer	Plausibilitätsprüfung der Einträge bei offenen Skalen	Eliminierung
	Randomisierte Zuordnung zu Gruppen	(ex ante festgelegt)	
	Gruppengröße zumindest >20		
	Korrelation zwischen abhängigen Variablen	Signifikanzprüfung über Pearson's R	Anwendung mehrerer unabhängiger ANOVAs
	Keine Multikollinearität der abhängigen Variablen	Prüfung der Toleranz	
	Multivariate Normalverteilung	Kolmogorov-Smirnov-Test	Gleichbesetzung der Zellen
	Varianzhomogenität	Levene-Test	Gleichbesetzung der Zellen
Kovarianzanalyse	Keine Beeinflussung der Kovariablen durch experimentelle Anordnung	Plausibilität	
	Kovariable auf intervallskaliertem Datenniveau	(ex ante festgelegt)	
	Korrelation der Kovariablen mit abhängigen Variablen	Signifikanzprüfung über Pearson's R	
	Keine Interaktion zwischen Kovariable und Faktor	Multiple Regressionsanalyse	
	Homogenität der Regressionskoeffizienten	Regressionsanalyse	Gleichbesetzung der Zellen

Tab. 3: Prämissenprüfung der (M)AN(C)OVA im Überblick

### 3.3.3.3 Vorgehensweise bei der Interpretation multivariater (Ko-)Varianzanalysen

Die Nullhypothese in der MAN(C)OVA lautet, dass die Vektoren der Mittelwerte aller abhängigen Variablen keine signifikanten Unterschiede in der Grundgesamtheit aufweisen, d.h., dass die Mittelwerte innerhalb der einzelnen Variablen über alle Gruppen gleich sind (Stevens 2002). Die zugehörige Alternativhypothese besagt entsprechend, dass es Unterschiede zwischen mindestens zwei Gruppen bei mindestens einer abhängigen Variable gibt. Um diese Hypothesen testen zu können, stehen dem Forscher insgesamt vier Signifikanztests – auch als **Omnibustests** bezeichnet – zur Verfügung: Wilks Lambda, Pillai-

Bartlett's  $V$ , die größte charakteristische Wurzel nach Roy und Hotelling-Lawley's Spurkriterium (*Bray/Maxwell* 1985). Die Tests unterscheiden sich zum Teil nur geringfügig. Wenn ein Faktor nur zwei Stufen aufweist, sind Wilks Lambda, Pillai-Bartlett's  $V$  und das Hotelling-Lawley's Spurkriterium sogar identisch. In der anwendungsorientierten Literatur hat sich **Wilks Lambda** aufgrund seiner vergleichsweise robusten Eigenschaften als gebräuchlichster Test durchgesetzt.

Bei der Interpretation der Ergebnisse ist zwischen sog. Haupt- und Interaktionseffekten zu unterscheiden. Eine Interaktion tritt auf, wenn die Wirkung eines Faktors auf eine abhängige Variable von der Ausprägung eines anderen Faktors beeinflusst wird. Somit erschwert sich die Interpretation eines Haupteffekts eines bestimmten Faktors, da sich der Einfluss dieses Faktors nur dann adäquat beschreiben lässt, wenn gleichzeitig Bezug auf die jeweilige Stufe der anderen unabhängigen Variable genommen wird. Signifikante **Interaktionseffekte** haben damit bei der Interpretation immer **Vorrang vor** signifikanten **Haupteffekten**. Praktisch bedeutet dies in solchen Fällen, dass gemäß Jaccard (1998) aus einer Hypothese heraus stets eine „focal independent variable“ sowie eine „moderator variable“ festgelegt und daran anschließend der Datensatz nach Maßgabe der Faktorstufen der „moderate variable“ gesplittet werden muss. Auf diese Weise kann auf Basis reduzierter Datensätze für jede Faktorstufe die jeweilige Hypothese über einen Wirkungsunterschied zwischen Gruppen der „focal independent variable“ geprüft werden (*Jaccard* 1998) [5].

Zeigt der Omnibustest eine Signifikanz für die durchgeführte MAN(C)OVA an, entsteht das Interpretationsproblem, dass anhand dieses Tests keine Informationen darüber gewonnen werden können, bei welchen abhängigen Variablen durch welchen Faktor Unterschiede hervorgerufen werden. Genau diese Informationen liefern die in einem nächsten Schritt anzuwendenden **Follow-Up-AN(C)OVAs** [6]. Die Schlüsselstatistik ist hier ein **F-Test**, der prüft, ob die Unterschiede in den Mittelwerten der Gruppen groß genug sind, so dass diese nicht zufällig aufgetreten sein können.

Sowohl bei der AN(C)OVA als auch bei der MAN(C)OVA kann die **Stärke eines Faktoreffekts** mithilfe der Maßzahl des Eta-Quadrats ( $\eta^2$ ) angegeben werden. Diese Größe sagt aus, wie viel Prozent der Varianz einer abhängigen Variablen durch einen Faktor erklärt werden kann, und nimmt entsprechend Werte zwischen Null und Eins an (*Malhotra* 1996; *Peterson/Albaum/Beltramini* 1985). Es sei darauf hingewiesen, dass von den  $\eta^2$ -Werten der ANCOVAs nicht auf die entsprechende Größe bei der MANCOVA geschlossen werden kann, da im letzteren Fall Vektoren von Mittelwerten die abhängigen Variablen repräsentieren, während dies bei der ANCOVA durch die Mittelwerte selbst getan wird. Das  $\eta^2$  ist im multivariaten Fall tendenziell größer als im univariaten Fall (*Tabachnik/Fidell* 2006). Je

größer der Wert, desto stärker der Einfluss des Faktors auf die abhängigen Variablen. Cohen (1988) macht darauf aufmerksam, dass sehr hohe  $\eta^2$ -Werte in verhaltenswissenschaftlich ausgerichteten Experimenten aufgrund der Natur dieser Versuche per se nicht zu erwarten sind. Er legt als inzwischen in der Literatur anerkannte Richtwerte das  $\eta^2$  für einen kleinen Effekt bei 1%, für einen mittleren Effekt bei 5,9% und für einen starken Effekt bei 13,8% fest (Cohen 1988; Novak 1995).

Anhand der Ergebnisse der Follow-Up-ANCOVAs können die aufgestellten Hypothesen jedoch noch nicht endgültig geprüft werden, da es dazu noch Angaben darüber bedarf, welche Gruppen sich mit welcher Effektrichtung voneinander unterscheiden. Dies kann in einem abschließenden Schritt der hier anzuwendenden Vorgehensweise entweder durch sog. **Post-Hoc-Tests** oder **A-Priori-Kontraste** geleistet werden. Diese unterscheiden sich dadurch, dass bei Post-Hoc-Tests multiple paarweise Vergleiche zwischen allen Gruppen und bei A-Priori-Kontrasten paarweise Vergleiche zwischen im Vorhinein definierten Gruppen durchgeführt werden (Bortz 2004).

Bei Post-Hoc-Tests ist zunächst auf Folgendes zu achten: Sind mehr als zwei Gruppen miteinander zu vergleichen, entsteht analog zum Zusammenhang zwischen der MANCOVA und den Follow-Up-ANCOVAs das Problem, dass die Wahrscheinlichkeit, einen Fehler 1. Art zu machen, kontrolliert werden muss. Daher sollte an dieser Stelle auf multiple t-Tests verzichtet werden und stattdessen – bei gegebener Varianzhomogenität der Vergleichsgruppen – auf die Scheffé-Prozedur oder – bei Varianzheterogenität – auf den Games-Howell-Test zurückgegriffen werden. Diese sind in der Lage, die Wahrscheinlichkeit eines  $\alpha$ -Fehlers auf ein vorgegebenes Signifikanzniveau zu minimieren. Werden hingegen nur zwei Gruppen miteinander verglichen, könnte prinzipiell ein t-Test herangezogen werden. Da dieser jedoch an die Prämisse der Normalverteilung und Varianzhomogenität in den Vergleichsgruppen gebunden ist, sollte bei Verletzung der in dieser Hinsicht robustere Brown-Forsythe-Test zur Anwendung kommen (Janssen/Laatz 2005). Die vollständige Vorgehensweise bei der Hypothesenprüfung mithilfe der MAN(C)OVA ist in Abb. 1 weiter unten aufgeführt.

### 3.4 Dokumentation

Im Sinne der Transparenz von wissenschaftlichen Erkenntnisgewinnen und entgegen der Auffassung angloamerikanischer Wissenschaftler (vgl. die Angaben der Journal-Herausgeber in Abschnitt 2), wird hier die Meinung vertreten, dass bei laborexperimenteller Marketingforschung nicht nur die **Ergebnisse der Hypothesenprüfung**, sondern auch sämtliche weiteren **prozessorientierten Schritte** (z.B. die Prämissenprüfungen) **vollstän-**

**dig dokumentiert** werden sollten. Nur dann ist nachvollziehbar, ob die Untersuchungen auch belastbare, methodisch richtige Ergebnisse hervorgebracht haben. Die geforderte Vollständigkeit bedingt auch, dass bei der Darstellung der varianzanalytischen Auswertung über **Signifikanztests und Gruppenunterschiede** hinaus ebenfalls die in der Marketing-Literatur oft vernachlässigte **Teststärke** und somit ein weiterer Nachweis für die Belastbarkeit von Befunden angegeben wird (Sawyer/Ball 1981). Die Teststärke (auch als Macht oder Power bezeichnet) macht Angaben über die Wahrscheinlichkeit, einen  $\beta$ -Fehler (man behält die Nullhypothese bei, obwohl sie in Wirklichkeit falsch ist) zu begehen. Die Literatur fordert hier ohne Ausschlusscharakter ein Niveau von mind. 0,8 (Hair et al. 2006).

Abbildung 1 fasst den gesamten Prozess zusammen, den man bei der Anwendung varianzanalytisch ausgerichteter Laborexperimente beachten sollte.

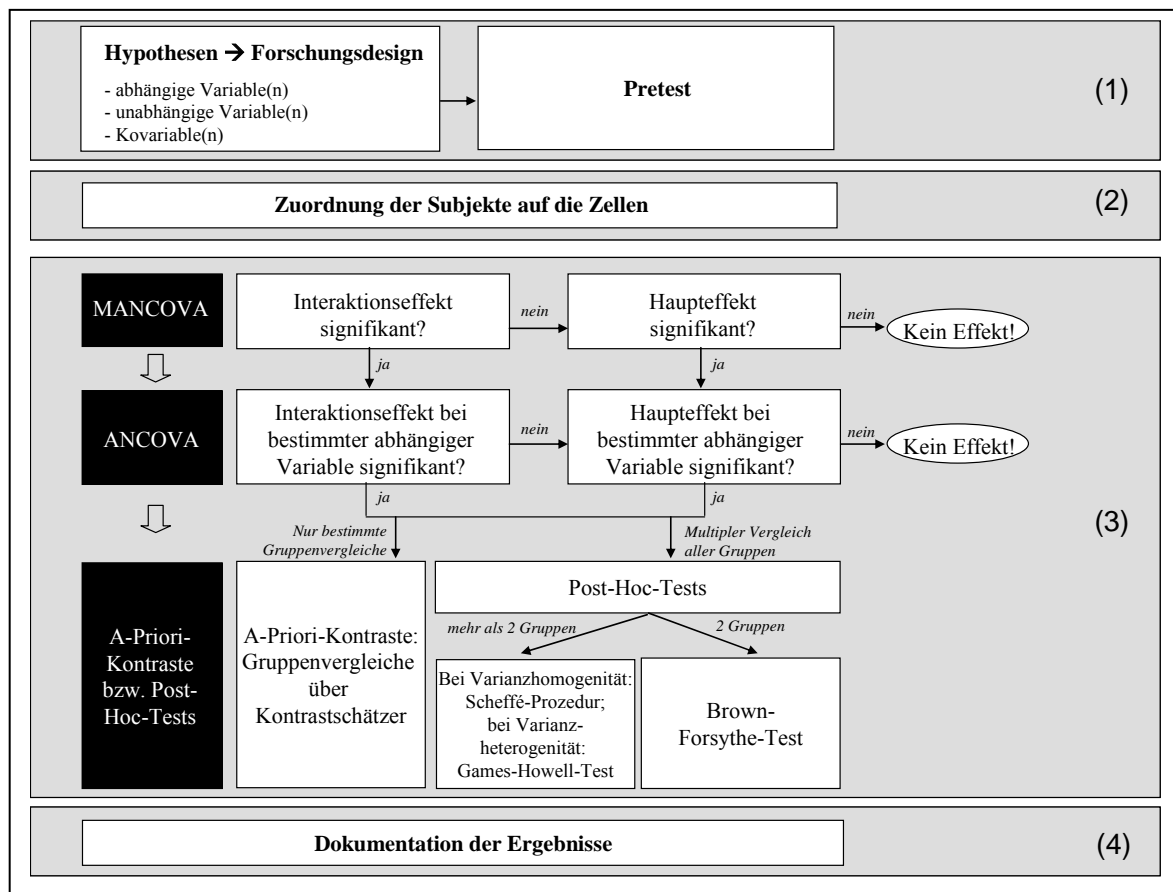


Abb. 1: Anwendung der Varianzanalyse bei Laborexperimenten im Überblick

## **4 Fazit**

Der Beitrag hat gezeigt, dass die experimentelle Marketingforschung im deutschsprachigen Raum zwar an Beliebtheit gewinnt, zumindest im Bereich der Top-Journals jedoch noch nicht soweit verbreitet ist wie im angloamerikanischen Bereich. Aufgrund des in Experteninterviews identifizierten diesbezüglichen Ausbildungsvorsprungs angloamerikanischer Marketingforscher wurde hier ein Leitfaden zur ordnungsgemäßen Anwendung vornehmlich varianzanalytisch ausgerichteter Laborexperimente vorgestellt. Im Zuge dessen hat sich gezeigt, dass eine Vielzahl an Einzelaspekten zu beachten ist, die in bisherigen, einschlägigen Publikationen nur sehr selten vollständig und korrekt beachtet werden. Vor diesem Hintergrund wäre es zu wünschen, dass der Leitfaden einen Beitrag zur Qualitätssteigerung laborexperimenteller Marketingforschung leisten kann.

## Endnoten

- [1] Die Inter-Rater Reliabilität belief sich auf über 99%, d.h., bei nur 7 von 1523 Studien lag bei der Einordnung einer Studie als Experiment ein Unterschied zwischen den beiden Ratern vor. Dieser konnte im Anschluss durch Diskussion jedoch gelöst werden.
- [2] Die Ergebnisse basieren auf je einer Email und einem Telefonat mit den drei Herausgebern Roland Rust (JM), Joel Huber (JMR) und John Deighton (JCR).
- [3] Dies erscheint schon deshalb wahrscheinlich, weil von Nachwuchswissenschaftlern zunehmend gefordert wird, in angloamerikanischen Top-Journals zu publizieren.
- [4] Es sei darauf hingewiesen, dass die bei einer „herkömmlichen“ Regressionsanalyse zu prüfenden Modellprämissen an dieser Stelle nur eine untergeordnete Rolle spielen. So tritt das Problem der Autokorrelation vor allem bei – in diesem Fall nicht relevanten – Zeitreihen auf. Multikollinearität ist lediglich dann zu prüfen, wenn mehr als eine unabhängige Variable in das Modell integriert wird. Hier liegt der Fokus jedoch einzig auf dem Konstrukt des Preiswissens als erklärender Variable. Hinsichtlich der Homoskedastizität der Störgrößen kann in dieser Untersuchung anhand einer visuellen Inspektion der Plots der Residuen gegen die jeweilige abhängige Variable kein Hinweis auf eine Verletzung dieser Prämisse festgestellt werden. Auch für einen nicht-linearen Zusammenhang des Preiswissens mit den abhängigen Variablen gibt es keine Anhaltspunkte ( vgl. zu Einzelheiten dieser Prämissen *Backhaus et al.* 2006).
- [5] Jaccard weist in diesem Zusammenhang darauf hin, dass die Bezeichnung der Interaktionseffekte eher irreführend ist, da bei der Interpretation nicht Interaktionen im engeren Sinne, sondern sog. „Simple Main Effects“ geprüft werden (*Jaccard* 1998).
- [6] Wenn bei der MAN(C)OVA keine Signifikanz auftritt, gibt es keinen Grund, Follow-Up-AN(C)OVAs anzuwenden. Sollten solche AN(C)OVAs entgegen der Aussagen der MAN(C)OVA überraschend signifikante Ergebnisse bei einem Faktor erbringen, ist das in der Regel darauf zurückzuführen, dass die Beziehungen zwischen den abhängigen Variablen nicht mehr berücksichtigt werden. Generell gilt in solchen Fällen, dass ein Modell niedrigerer Komplexität durch Überlagerung von Effekten ein falsches Bild von Zusammenhängen ergeben kann und entsprechend auf ein solches Modell zu verzichten ist. Vor diesem Hintergrund ist auch zu verstehen, warum die MAN(C)OVA als Protektionstest für die AN(C)OVA bezeichnet wird.

## Literaturverzeichnis

- Ahlert, D./Evanschitzky, H./Hesse, J. (2005):* Zwischen Relevance und Rigor: Der erfolgreiche Einsatz der Erfolgsforschung in Wissenschaft und Praxis, in: *Wirtschaftswissenschaftliches Studium (WiSt)*, Vol. 34, No. 7, 362-367.
- Backhaus, K./Erichson, B./Plinke, W./Weiber, R. (2006):* Multivariate Analysemethoden - Eine anwendungsorientierte Einführung, 11. Auflage, Berlin.
- Bortz, J. (2004):* Statistik für Sozialwissenschaftler, 6. Auflage, Berlin.
- Bortz, J./Döring, N. (2006):* Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, 4. Auflage, Berlin.
- Bray, J.H./Maxwell, S.E. (1985),* Multivariate Analysis of Variance, Beverly Hills.
- Cohen, J. (1988):* Statistical Power Analysis for the Behavioral Sciences, 2nd edition, Hillsdale.
- Crano, W.D./Brewer, M.B. (2002):* Principles and Methods of Social Research, 2nd Edition, Mahwah, NJ.
- Diehl, J.M. (1983):* Varianzanalyse, 4. Auflage, Frankfurt am Main.
- Dretzke, B.J./ Levin, J.R./ Serlin, R.C. (1982):* Testing for Regression Homogeneity Under Variance Heterogeneity, in: *Psychological Bulletin*, Vol. 91, No. 2, S. 376-383.
- Fickel, N. (2001):* Sequenzialregression, Berlin.
- Glanville, R. (1999):* Researching Design and Designing Research, in: *Design Issues*, Vol. 15, No. 2, S. 80-91.
- Glaser, W.R. (1978):* Varianzanalyse, Stuttgart.
- Gröppel-Klein, A./Weinberg, P. (2000):* Die Konsumentenforschung im Marketing – Stärken und Schwächen aus Erfahrungssicht, in: Backhaus, K. (Hrsg.), *Deutschsprachige Marketingforschung – Bestandsaufnahme und Perspektiven*, Stuttgart, S. 79-95.
- Hair, J.F./Black, W.C./Babin, B.J./Anderson, R.E./Tatham, R.L. (2006):* Multivariate Data Analysis, 6. Auflage, Upper Saddle River.
- Herrmann, A./Seilheimer, C. (2000):* Varianz- und Kovarianzanalyse, in: Herrmann, A./Homburg, C. (Hrsg.), *Marktforschung - Methoden, Anwendungen, Praxisbeispiele*, 2. Auflage, Wiesbaden, S. 267-294.
- Homburg, C./Giering, A. (1996),* Konzeptualisierung und Operationalisierung komplexer Konstrukte – Ein Leitfaden für die Marketingforschung, in: *Marketing ZFP*, 18. Jg., Nr. 1, S. 5-24.
- Huberty, C.J. (1994):* Why Multivariable Analysis?, in: *Educational and Psychological Measurement*, Vol. 54, No. 3, S. 620-627.

- Huberty, C.J./Morris, J.D.* (1989): Multivariate Analysis versus Multiple Univariate Analysis, in: *Psychological Bulletin*, Vol. 105, S. 302-308.
- Jaccard, J.* (1998): *Interaction Effects in Factorial Analysis of Variance*, Thousand Oaks.
- Janssen, J./Laatz, W.* (2005): *Statistische Datenanalyse mit SPSS für Windows- Eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests*, 5. Auflage, Berlin.
- Kieser, A./Nicolai, A.T.* (2005): Success Factor Research. Overcoming the Trade-Off Between Rigor and Relevance?, in: *Journal of Management Inquiry*, Vol. 14, Nr. 3, S. 275-279.
- Läuter, J.* (1978): Sample Size Requirements for the T2 Test of MANOVA (Tables for One-Way Classification), in: *Biometrical Journal*, Vol. 20, S. 389-406.
- Leiberich, P.* (1995): Varianzanalyse, in: Tietz, B./ Köhler, R./ Zentes, J. (Hrsg.), *Handwörterbuch des Marketing*, 2. Auflage, Stuttgart, Sp. 2507-2523.
- Levy, K.L.* (1980): A Monte Carlo Study of Analysis of Covariance under Violations of the Assumptions of Normality and Equal Regression Slopes, in: *Educational and Psychological Measurement*, Vol. 40, S. 835-840.
- Litz, H.P.* (2000): *Multivariate Statistische Methoden - und ihre Anwendung in den Wirtschafts- und Sozialwissenschaften*, München.
- Malhotra, N.K.* (1996): *Marketing Research*, 2nd Edition, Englewood Cliffs.
- Novak, T.P.* (1995): MANOVAMAP: Graphical Representation of MANOVA in Marketing Research, in: *Journal of Marketing Research*, Vol. 32, No. 8, S. 357-374.
- o. V.* (2006): Code of Ethics and Conduct – March 2006, The British Psychological Society, [http://www.bps.org.uk/downloadfile.cfm?file\\_uuid=5084A882-1143-DFD0-7E6C-F1938A65C242&ext=pdf](http://www.bps.org.uk/downloadfile.cfm?file_uuid=5084A882-1143-DFD0-7E6C-F1938A65C242&ext=pdf), Zugriff am 5. April 2007.
- Perdue, B.C./Summers, J.O.* (1986): Checking the Success of Manipulations in Marketing Experiments, in: *Journal of Marketing Research*, Vol. 23, No. 4, S. 317-326.
- Perreault, Jr., W.D./Darden, W.R.* (1975): Unequal Cell Sizes in Marketing Experiments: Use of the General Linear Hypothesis, in: *Journal of Marketing Research*, Vol. 12, No. 3, S. 333-342.
- Peterson, R.A./Albaum, G./ Beltramini, R.F.* (1985): A Meta-Analysis of Effect Sizes in Consumer Behavior Experiments, in: *The Journal of Consumer Research*, Vol. 12, No. 1, S. 97-103.
- Sawyer, A.G./Ball, A.D.* (1981): Statistical Power and Effect Size in Marketing Research, in: *Journal of Marketing Research*, Vol. 18, No.3, S. 275-290.

- Stevens, J.P.* (2002): *Applied Multivariate Statistics for the Social Sciences*, 4th Edition, Mahwah.
- Tabachnik, B.G./Fidell, L.S.* (2006): *Using Multivariate Statistics*, 5th Edition, Boston.
- van Doorn, J.* (2004): *Zufriedenheitsdynamik - Eine Panelanalyse bei industriellen Dienstleistungen*, Wiesbaden.
- Varadarajan, P.R.* (2003): Musings on Relevance and Rigor of Scholarly Research in Marketing., in: *The Journal of the Academy of Marketing Science*, Vol. 31, No. 4, S. 368-376.
- Venkatesan, M.* (1967): Laboratory Experiments in Marketing: The Experimenter Effect, in: *Journal of Marketing Research*, Vol. 4, No. 2, S. 142-146.
- Winer, R.S.* (1999): Experimentation in the 21st Century: The Importance of External Validity, in: *The Journal of the Academy of Marketing Science*, Vol. 27, No. 3, S. 349-358.