

Pyka, Sebastian; Furchheim, Pia

Working Paper

Experimentelle Marktforschung – Eine Einführung in die sozialwissenschaftliche Experimentalforschung

Chemnitz Economic Papers, No. 014

Provided in Cooperation with:

Chemnitz University of Technology, Faculty of Economics and Business Administration

Suggested Citation: Pyka, Sebastian; Furchheim, Pia (2017) : Experimentelle Marktforschung – Eine Einführung in die sozialwissenschaftliche Experimentalforschung, Chemnitz Economic Papers, No. 014, Chemnitz University of Technology, Faculty of Economics and Business Administration, Chemnitz

This Version is available at:

<https://hdl.handle.net/10419/170674>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



TECHNISCHE UNIVERSITÄT
CHEMNITZ

Faculty of Economics and
Business Administration



Experimentelle Marktforschung – Eine Einführung in die sozialwissenschaftliche Experimentalforschung

Sebastian Pyka
Pia Furchheim

Chemnitz University of Technology
Faculty of Economics and Business Administration
Thüringer Weg 7
09107 Chemnitz, Germany

Phone +49 (0)371 531 26000

Fax +49 (0371) 531 26019

<https://www.tu-chemnitz.de/wirtschaft/index.php.en>

wirtschaft@tu-chemnitz.de

Experimentelle Marktforschung – Eine Einführung in die sozialwissenschaftliche Experimentalforschung

Sebastian Pyka*, Pia Furchheim

Zusammenfassung:

Da experimentelle Versuchsanordnungen im Vergleich zu anderen Datenerhebungsverfahren Kausalitäten zwischen Untersuchungsvariablen am strengsten prüfen und zeitgleich validere Ergebnisse erzielen, steigt die Bedeutung von experimentellen Untersuchungen in der wirtschaftswissen-/betriebswirtschaftlichen Forschung.

Mit dem vorliegenden Beitrag erfolgt daher eine grundlegende Einführung in diese empirische Erhebungsmethodik. Dabei werden nicht nur die konstitutiven Merkmale von experimentellen Untersuchungen erläutert, sondern auch verschiedene experimentelle Versuchsanordnungen dargestellt und deren spezifische Vor- und Nachteile diskutiert, z.B. in Bezug auf die erforderliche Stichprobengröße oder die Wirkung und Eliminierung von Störvariablen. Darüber hinaus werden die Gütekriterien zur Beurteilung von experimentellen Untersuchungen thematisiert und aufgezeigt, wie diese auf einem hohen Niveau erfüllt werden können.

Dr. Sebastian Pyka*, Technische Universität Chemnitz, Professur für Marketing und Handelsbetriebslehre

Dr. Pia Furchheim, University of Lausanne (HEC), Department of Marketing

*Kontakt:

(+49) 0371-531-37953

sebastian.pyka@wirtschaft.tu-chemnitz.de

Inhaltsverzeichnis

<i>Tabellenverzeichnis</i>	II
1. Das Experiment als Verfahren der Datenerhebung	1
1.1 Definition und Einordnung des Experiments	1
1.2 Konstitutive Merkmale eines experimentellen Untersuchungsdesigns	2
2. Experimentelle Versuchsanordnungen	4
2.1 „Echte“, Quasi- und Non-Experimente	4
2.2 Klassifikation von Experimenten	4
2.3 Experimentelle Versuchspläne/-anordnungen	10
3. Gütekriterien der experimentellen Forschung	13
3.1 Kriterien zur Beurteilung der konzeptionellen Güte des Experiments	13
3.2 Kriterien zur Beurteilung der Güte der eingesetzten Messinstrumente	15
<i>Literaturverzeichnis</i>	III

Tabellenverzeichnis

Tabelle 1: Klassifikation von experimentellen Versuchsanordnungen..... 5

Tabelle 2: Zuordnung von Probanden bei einem Within Subject-Design und einem
Between Subject-Design 7

Tabelle 3: Übersicht über die verschiedenen experimentellen Versuchspläne 13

Tabelle 4: Gütekriterien zur Reliabilitäts-/Validitätsprüfung und deren Schwellenwerte..... 16

1. Das Experiment als Verfahren der Datenerhebung

1.1 Definition und Einordnung des Experiments

Die Durchführung von experimentellen Untersuchungen bietet sich insbesondere dann an, wenn gerichtete kausale Ursache-Wirkungs-Beziehungen zwischen Untersuchungsvariablen identifiziert und beurteilt werden sollen. Die bestehenden Kausalitäten zwischen den Untersuchungsvariablen sollen dabei tiefergehend verstanden werden, um genauere Vorhersagen bezüglich der Wirkungszusammenhänge treffen zu können (vgl. Haslam & McGarty 2004, S. 239). Das Experiment beschreibt die empirische Methode, bei der mindestens eine unabhängige Variable (= Ursache) bewusst durch den Experimentator variiert wird, um in Folge dieser Variation die Veränderung einer oder mehrerer abhängiger Variablen (= Wirkung) zu messen (vgl. Bröder 2011, S. 63; Churchill & Iacobucci 2004, S. 128f.). Während die unabhängige Variable(n) systematisch zu manipulieren ist/sind, wird die abhängige Variable, welche außerhalb der Kontrolle des Experimentators liegt, systematisch erfasst (vgl. Bröder 2011, S. 62; Harris 2010, S. 129f.; Reis & Gosling 2010, S 84).

Experimente stellen keine gesonderte Form der Datenerhebung dar, sondern erfolgen mittels Befragung und/oder (apparativer) Beobachtung. Experimentelle Untersuchungen können somit durch eine Vielzahl an Erhebungsverfahren ausgestaltet werden, die von einer einfachen computer-gestützten Befragung bis zu komplexen Kombinationen aus Befragung und Beobachtung reichen (vgl. Berekoven, Eckert & Ellenrieder 2009, S. 146). Im Vergleich zu anderen Datenerhebungsverfahren werden mit experimentellen Versuchsanordnungen aber Kausalitäten am strengsten geprüft und zeitgleich validere Ergebnisse erzielt, da die interessierende unabhängige Variable gezielt verändert wird, während zeitgleich alle anderen Größen, die die abhängige Variable beeinflussen, kontrolliert werden (vgl. Bröder 2011, S. 73; Harris 2010, S. 169; Reis & Gosling 2010, S. 84). Damit wird die Wirkung der interessierenden unabhängigen Variable auf die abhängige Variable isoliert erfasst, weshalb das Experiment auch als eine isolierte Variation zu bezeichnen ist (vgl. Harris 2010, S. 129f.; Reis & Gosling 2010, S 84).

1.2 Konstitutive Merkmale eines experimentellen Untersuchungsdesigns

Um die Kausalität zwischen den Untersuchungsvariablen zu überprüfen und damit tiefergehende Ursache-Wirkungs-Zusammenhänge aufzudecken, müssen bei einer experimentellen Versuchsanordnung verschiedene Voraussetzungen erfüllt sein. Diese Voraussetzungen sind mit den konstitutiven Merkmalen eines Experiments gleichzusetzen (vgl. Eid, Gollwitzer & Schmitt 2010, S. 55f.; Sedlmeier & Renkewitz 2013, S. 123):

1. Die unabhängige Variable (Ursache) und die abhängige Variable (Wirkung) müssen kovariieren.
2. Die unabhängige Variable (Ursache) tritt zeitlich vor der abhängigen Variable (Wirkung) auf.
3. Mögliche Alternativerklärungen für die Veränderung der abhängigen Variable (Wirkung) können ausgeschlossen werden.

Die erste Bedingung erfordert, dass ein Zusammenhang zwischen der unabhängigen Variable und der abhängigen Variable besteht. Dies bedeutet, dass eine Veränderung der unabhängigen Variable zu einer Veränderung der abhängigen Variable führen muss. Diese geforderte Kovariation zwischen den Variablen ist damit eine notwendige Bedingung für den kausalen Zusammenhang zwischen einer Ursache (unabhängige Variable) und einer Wirkung (abhängige Variable) (vgl. Sedlmeier & Renkewitz 2013, S. 123; Eid, Gollwitzer & Schmitt 2010, S. 55f.). Die unabhängige Variable muss hierbei mindestens zwei Ausprägungen aufweisen, um die Kovariation zwischen den Untersuchungsvariablen zu zeigen. Die unterschiedlichen Ausprägungen der unabhängigen Variable werden durch den Experimentator festgelegt (systematisch manipuliert), wobei jede Ausprägung als eine Experimentalbedingung zu verstehen ist (vgl. Bröder 2011, S. 74).

Neben der Kovariation zwischen den Untersuchungsvariablen ist die zeitliche Präzedenz der unabhängigen Variable eine weitere Bedingung des Experiments. Dies bedeutet, dass die Ursache zeitlich vor der Wirkung auftreten muss, bzw. die zeitliche Abfolge zwischen den Variablen bekannt sein muss. Diese Bedingung ist erforderlich, um die Richtung der kausalen Beziehungsstruktur zwischen den Variablen festzulegen (vgl. Eid, Gollwitzer & Schmitt 2010, S. 55f.). Durch die Manipulation der unabhängigen Variable legt der Experimentator die kausale Richtung des zu untersuchenden Zusammenhangs fest, wobei eine isolierte Betrachtung der unabhängigen Variable und der abhängigen Variable vorausgesetzt wird (vgl. Sedlmeier & Renkewitz 2013, S. 125).

Durch diese beiden konstitutiven Merkmale zeigt sich die zentrale Bedeutung von Kausalhypothesen, die die Grundlage von experimentellen Untersuchungen darstellen. Durch diese Hypothesen wird die kausale Beziehungsstruktur zwischen den Variablen festgelegt und damit determiniert, welche der Untersuchungsvariablen die unabhängige Variable (= Ursache, die zu manipulieren ist) und welche die abhängige Variable (= Wirkung, die in Folge der Manipulation zu messen ist) darstellt (vgl. Sedlmeier & Renkewitz 2013, S. 124f.).

An dieser Stelle ist allerdings zu beachten, dass ein kausaler/korrelativer Zusammenhang zwischen den Untersuchungsvariablen noch nicht bedeutet, dass die unabhängige Variable tatsächlich einen ursächlichen Einfluss auf die abhängige Variable besitzt. Deshalb sind bei einer experimentellen Versuchsanordnung Alternativerklärungen für die Veränderung der abhängigen Variable auszuschließen (vgl. Sedlmeier & Renkewitz 2013, S. 126). Mit dieser Bedingung geht einher, dass alle möglichen Einflussgrößen (mit Ausnahme der interessierenden unabhängigen Variable) und damit Störvariablen zu kontrollieren bzw. konstant zu halten sind, um damit deren Einflüsse auf die abhängige Variable zu eliminieren (vgl. Eid, Gollwitzer & Schmitt 2010, S. 55f.). Störvariablen, auch exogene Variablen genannt, sind personen-, bedingungs- und/oder situationsgebundene Variablen, die gemeinsam mit der unabhängigen Variable die Ausprägung der abhängigen Variable beeinflussen und den zu beobachtenden Effekt verzerren können (vgl. Bröder 2011, S. 64). Das Experiment muss demnach eine Versuchsanordnung darstellen, die es ermöglicht, die unabhängige Variable systematisch zu variieren und zeitgleich alle Störvariablen zu kontrollieren¹ (vgl. Bröder 2011, S. 64). In diesem Zusammenhang ist die zufallsbasierte Zuordnung der am Experiment teilnehmenden Personen zu den verschiedenen Experimentalbedingungen von zentraler Bedeutung. Die sogenannte Randomisierung basiert auf der Wahrscheinlichkeitstheorie und geht davon aus, dass durch eine zufallsbasierte Zuordnung der Probanden zu den Experimentalbedingungen eine Gleichverteilung der Merkmale erreicht wird und damit eine Balancierung der Störvariablen erfolgt, wodurch deren Wirkung auf die abhängige Variable eliminiert wird (vgl. z.B. Bröder 2011, S. 65, S. 68; Harris 2010, S. 159; Ryan 2007, S. 7). Die Randomisierung ist damit essentiell für experimentelle Designs und qualifiziert eine Untersuchung erst als Experiment. Gerade durch die zufallsbasierte Zuordnung der Probanden zu den Experimentalbedingungen kann gewährleistet werden, dass die verschiedenen Untersuchungsgruppen hinsichtlich der Ausprägungen der personengebundenen Störvariablen strukturgleich sind

¹ Eine ausführliche Darstellung der Möglichkeiten zur Konstanthaltung und Balancierung bzw. Kontrolle von Störvariablen findet sich zum Beispiel bei Sedlmeier & Renkewitz (2016, S. 129-140) oder bei Eid, Gollwitzer & Schmidt (2010, S. 58-60).

und damit vor der Durchführung des Experiments keine bzw. nur geringe Unterschiede bei der Ausprägung der abhängigen Variable vorliegen (vgl. Reis & Gosling 2010, S 84; Bortz & Döring 2006, S. 529).

2. Experimentelle Versuchsanordnungen

2.1 „Echte“, Quasi- und Non-Experimente

In der wissenschaftlichen und anwendungsorientierten Forschung ist zunächst zu unterscheiden, ob ein „echtes“ Experiment, ein Quasi-Experiment oder ein Non-Experiment vorliegt.

Ein „echtes“ Experiment zeichnet sich durch drei wesentliche Merkmale aus (vgl. Reis & Gosling 2010, S. 84):

- aktive Manipulation von einer oder mehreren unabhängigen Variablen,
- zufällige (randomisierte) Zuordnung der Teilnehmer zu den einzelnen Experimentalbedingungen,
- Kontrolle oder Eliminierung von Störvariablen.

Quasi-Experimente verfügen über ähnlich charakteristische Merkmale, jedoch fehlt hier die zufällige Zuordnung der Teilnehmer zu den einzelnen Experimentalbedingungen (vgl. Cook & Campbell 1979, S. 6). Stattdessen weisen die einzelnen Gruppen natürliche Unterschiede auf, die nicht auf die Manipulation der unabhängigen Variable zurückzuführen sind (vgl. Bortz & Döring 2010; S. 551; Cook & Campbell 1979, S. 6). Dadurch sind die Ergebnisse oft schwerer interpretierbar als bei „echten“, randomisierten Experimenten.

Sind beide Bedingungen (randomisierte Zuordnung der Teilnehmer zu den Experimentalbedingungen und Kontrolle/Elimination von Störvariablen) nicht erfüllt, kann nicht von einem Experiment gesprochen werden. Vielmehr handelt es sich dann um eine Korrelationsstudie, bei welcher keine Aussage über Ursache-Wirkungs-Zusammenhänge getroffen werden kann (vgl. Cook & Campbell 1979, S. 6).

2.2 Klassifikation von Experimenten

„Echte“ experimentelle Untersuchungen sind weiterhin anhand von verschiedenen Kriterien zu charakterisieren, insbesondere durch die Anzahl der zu berücksichtigenden unabhängigen Variablen und der damit einhergehenden Anzahl an Experimentalbedingungen, durch die Art der Zuordnung der Probanden zu diesen Experimentalbedingungen sowie durch die Experimentalsituation selbst. Zur Klassifikation von experimentellen Versuchsanordnungen dient

somit die Differenzierung zwischen ein- und mehrfaktoriellen Design, die Unterscheidung zwischen Within Subject-Design und Between Subject-Design sowie die Unterscheidung zwischen Feld- und Laborexperiment. In der folgenden Tabelle sind diese verschiedenen Klassifikationen von experimentellen Versuchsanordnungen zusammenfassend dargestellt.

Tabelle 1: Klassifikation von experimentellen Versuchsanordnungen

Klassifikation von Experimenten	Formen von Experimenten
nach der Anzahl der zu berücksichtigenden unabhängigen Variablen	<ul style="list-style-type: none">• Einfaktorielles Design• Mehrfaktorielles Design
nach der Art der Zuordnung der Probanden zu den Experimentalbedingungen	<ul style="list-style-type: none">• Within Subject-Design• Between Subject-Design
nach der Experimentalsituation	<ul style="list-style-type: none">• Laborexperiment• Feldexperiment

Quelle: eigene Darstellung

Ein- und mehrfaktorielles Design

Bei der Klassifikation von Experimenten nach dem ein- und mehrfaktoriellen Design wird die Anzahl der unabhängigen Variablen betrachtet, die in der experimentellen Versuchsanordnung unterschieden werden (vgl. Bröder 2011, S. 74). Als Faktor wird in diesem Zusammenhang eine unabhängige Variable bezeichnet, die mindestens zwei unterschiedliche Ausprägungen (sogenannte Faktorstufen) besitzen muss. Die Anzahl an Faktoren und Faktorstufen determiniert hierbei die Anzahl der Experimentalbedingungen, die in der Untersuchung zu berücksichtigen sind. Jede Faktorstufe (d.h. jede Ausprägung der unabhängigen Variable) ist mit einer Experimentalbedingung gleichzusetzen (vgl. Sedlmeier & Renkewitz 2013, S. 124; Bröder 2011, S. 74).

Dabei wird mit einem einfaktoriellen Design nur eine unabhängige Variable mit mindestens zwei verschiedenen Ausprägungen in der experimentellen Versuchsanordnung erfasst, so dass mindestens zwei Experimentalbedingungen zu unterscheiden sind (vgl. Bröder 2011, S. 79). Im Gegensatz dazu werden bei einem mehrfaktoriellen Design mindestens zwei unabhängige Variablen berücksichtigt. Bei der einfachsten Form eines mehrfaktoriellen Designs sind zwei unabhängige Variablen mit jeweils zwei verschiedenen Ausprägungen zu unterscheiden und dadurch vier Experimentalbedingungen gegeben (vgl. Sedlmeier & Renkewitz 2013, S. 160f.). Die Beschreibung von solchen mehrfaktoriellen Versuchsplänen folgt hierbei einer spezifischen Schreibweise. Erfasst werden dabei die Anzahl an unabhängigen Variablen sowie die Anzahl der verschiedenen Ausprägungen je unabhängiger Variable. Am Beispiel der einfachsten Form eines mehrfaktoriellen Designs wird von einem 2x2-Design gesprochen: 2 (unabhängige Variable 1 mit zwei Ausprägungen) x 2 (unabhängige Variable 2 mit zwei Ausprägungen)-Design (vgl. Sedlmeier & Renkewitz 2013, S. 160f.; Bröder 2011, S. 79). Die

Anzahl der Experimentalbedingungen ergibt sich nun durch die Multiplikation der Anzahl der Ausprägungen der unabhängigen Variablen (vgl. Bortz & Döring 2006, S. 537). Mehrfaktorielle Experimente können hierbei entweder durch eine Erhöhung der Anzahl der Ausprägungen je unabhängiger Variable (z.B. 2x3-Design oder 3x3-Design) oder durch die Berücksichtigung von weiteren unabhängigen Variablen (z.B. 2x2x2-Design) angereichert werden. Schlussfolgernd führt dies aber zu einer steigenden Anzahl an Experimentalbedingungen (vgl. Bortz & Döring 2006, S. 537).

Hinsichtlich der praktischen Anwendung zeigt sich, dass in der Regel Experimente mit mehr als drei unabhängigen Variablen nur selten durchgeführt werden. Dies begründet sich durch die Anzahl der Experimentalbedingungen, die nicht nur die Komplexität der Ergebnisauswertung und -interpretation beeinflusst, sondern auch die Durchführung von Experimenten erschwert, indem entweder eine größere Anzahl an Probanden benötigt wird oder die einzelnen Probanden vielzählige Experimentalbedingungen durchlaufen müssen (vgl. Sedlmeier & Renkewitz 2013, S. 161).

Within Subject- und Between Subject-Design

Neben der Anzahl der erfassten unabhängigen Variablen werden experimentelle Versuchsanordnungen auch anhand des Within Subject-Designs und des Between Subject-Designs klassifiziert (vgl. z.B. Sedlmeier & Renkewitz 2013, S. 144; Bröder 2011, S. 76; Kantowitz, Roediger, III & Elmes 2009, S. 236f.). Diese beiden grundsätzlichen, aber konträren Ansätze, legen die Art der Zuordnung der Probanden zu den verschiedenen Experimentalbedingungen fest. Dadurch bestimmen beide Verfahren die Variation der unabhängigen Stichproben und beeinflussen die Größe der benötigten Stichprobe sowie die Wirkung von Störvariablen (vgl. Harris 2010, S. 150f.).

Bei dem Within Subject-Design werden die Probanden allen Ausprägungen der unabhängigen Variable ausgesetzt (vgl. Tabelle 2). Alle Probanden durchlaufen somit alle Experimentalbedingungen, wobei die verschiedenen Experimentalbedingungen in der Regel in einer randomisierten Reihenfolge dem jeweiligen Probanden vorgelegt werden (vgl. Sedlmeier & Renkewitz 2013, S. 144). Hierbei wird die Ausprägung der abhängigen Variable bei allen Probanden in allen Bedingungen ermittelt. Durch einen Vergleich der abhängigen Variable in den einzelnen Experimentalbedingungen innerhalb derselben Person wird folglich die experimentelle Wirkung der unabhängigen Variable auf die abhängige Variable ersichtlich. Aufgrund dieses Vorgehens wird das Within Subject-Design auch als das intraindividuelle Design bezeichnet (vgl. z.B. Harris 2010, S. 156; Kantowitz, Roediger, III & Elmes 2009, S. 236).

Tabelle 2: Zuordnung von Probanden bei einem Within Subject-Design und einem Between Subject-Design

Between Subject-Design mit 2 Bedingungen		Within Subject-Design mit zwei Bedingungen		
Bedingung 1	Bedingung 2	Teilnehmer	Bedingung 1	Bedingung 2
P02	P01	P01	2nd	1st
P04	P03	P02	1st	2nd
P09	P05	P03	1st	2nd
P10	P06	P04	2nd	1st
P11	P07	P05	2nd	1st
P13	P08	P06	1st	2nd
P14	P12	P07	2nd	1st
P19	P15	P08	1st	2nd
P21	P16	P09	2nd	1st
P24	P17	P10	1st	2nd
P25	P18	P11	2nd	1st
P26	P20	P12	2nd	1st
P28	P22	P13	1st	2nd
P29	P23	P14	2nd	1st
P30	P27	P15	1st	2nd
P32	P31	P16	1st	2nd
Anzahl benötigter Teilnehmer (P) = 32		Anzahl benötigter Teilnehmer (P) = 16		

Quelle: in Anlehnung an Harris (2008, S. 150f.)

Aus forschungsökonomischer Sicht ist das intraindividuelle Design vorteilhaft, da nur eine relativ geringe Anzahl an Probanden zur Durchführung des Experiments benötigt wird (vgl. Kantowitz, Roediger, III & Elmes 2009, S. 235f.). Ein weiterer essentieller Vorteil des Within Subject-Designs besteht darin, dass jeder Proband zeitgleich seine eigene Kontrollperson darstellt. Deshalb besitzen personengebundene Störvariablen keinen Einfluss auf die abhängige Variable und die bestehende Restvarianz² ist relativ gering (vgl. Bröder 2011, S. 76; Harris 2010, S. 156). Trotz dieser Vorteile sind beim Einsatz des intraindividuellen Designs aber verschiedene Effekte zu beachten, die zu Ergebnisverzerrungen führen können und damit gegen die Verwendung dieses Verfahrens sprechen. Besonders zu beachten sind der Carry Over-Effekt, der Positionseffekt, der Ermüdungseffekt und der Demand-Effekt (vgl. z.B. Bröder 2011, S. 77; Harris 2010, S. 157ff.). Ein Carry Over-Effekt entsteht, wenn die Reihenfolge, in welcher die Probanden die verschiedenen Experimentalbedingungen durchlaufen, die experimentelle Wirkung beeinflusst. Carry Over-Effekte liegen somit vor, wenn die Teilnahme an einer experimentellen Bedingung das (inhaltliche) Verhalten des Probanden in einer nachgelagerten Bedingung beeinflusst (vgl. Sedlmeier & Renkewitz 2013, S. 158). Eine weitere Möglichkeit, die beim Einsatz des Within Subject-Designs zu Ergebnisverzerrungen führen kann, zeigt sich durch Positionseffekte. Da den Probanden alle Experimentalbedingungen zugewiesen werden, besteht das Risiko, dass die Probanden ihr Verhalten anpassen, indem Frage-/Aufgabenstellungen wiederholt bearbeitet werden und somit im Zeitablauf des Experiments Lerneffekte auftreten, die sich durch eine größere Vertrautheit mit der Frage/Aufgabe oder dem Ablegen von vorhandener Nervosität kennzeichnen. Demnach

² Anteil an der Steuerung der abhängigen Variable, der nicht durch die unabhängige/n Variable/n erklärt werden kann.

ist zu erwarten, dass die Probanden im zeitlichen Ablauf des Experiments ihre Leistung verbessern, so dass Bedingungen begünstigt werden, die in der experimentellen Versuchsanordnung zu einem späteren Zeitpunkt durchlaufen werden (vgl. Sedlmeier & Renkewitz 2013, S. 152). Eine spezifische, negative Form des Positionseffekts ist der Ermüdungseffekt. Durch die Teilnahme der Probanden an allen Experimentalbedingungen besteht die Möglichkeit, dass sich die Ergebnisse der Probanden im Zeitverlauf des Experiments nicht verbessern, sondern aufgrund einer abnehmenden Konzentration, Aufmerksamkeit und Motivation verschlechtern (vgl. Bröder 2011, S. 77; Harris 2010, S. 157). Darüber hinaus besteht das Risiko, dass die Probanden die Unterschiede zwischen den verschiedenen Bedingungen erkennen und über den Untersuchungszweck spekulieren, was zu vermeintlich sozial erwünschten Antworten führen kann. Die dadurch entstehenden Ergebnisverzerrungen sind auf den Demand-Effekt zurückzuführen (vgl. Sedlmeier & Renkewitz 2013, S. 151; Bröder 2011, S. 77).

Sollte die Gefahr bestehen, dass durch diese Nachteile des Within Subject-Designs verzerrte Ergebnisse erzielt werden, ist das Between Subject-Design hinsichtlich der Zuordnung der Probanden zu den experimentellen Bedingungen zu bevorzugen (vgl. Sedlmeier & Renkewitz 2013, S. 151, S. 158).

Das Between Subject-Design wird auch als das interindividuelle Design bezeichnet, da jeder Proband per Zufall nur einer Experimentalbedingung zugeordnet wird, so dass jede Bedingung eine eigene Experimentalgruppe darstellt (vgl. Tabelle 2). In jeder dieser Gruppen wird die Ausprägung der abhängigen Variable ermittelt und durch Gruppenvergleiche die experimentelle Wirkung der unabhängigen Variable auf die abhängige Variable bestimmt (vgl. Bröder 2011, S. 76). Da alle Probanden nur einer Experimentalbedingung ausgesetzt werden, können demnach Carry Over-Effekte, Positions- und Ermüdungseffekte sowie Demand-Effekte³ weitestgehend ausgeschlossen werden (vgl. Harris 2010, S. 162). Während dies essentielle Vorteile des Between Subject-Designs gegenüber dem Within Subject-Design sind, müssen dennoch einige Nachteile dieses Verfahrens beachtet werden. Bei der Durchführung eines Experiments mit dem Between Subject-Design wird eine wesentlich größere Stichprobe benötigt, so dass dieses Verfahren im Vergleich zum Within Subject-Design weniger forschungsökonomisch ist (vgl. Kantowitz, Roediger, III & Elmes 2009, S. 235f.). Ein weiterer Nachteil des Between Subject-Designs zeigt sich durch die Möglichkeit des Gruppeneffektes. Dieser Effekt liegt vor, wenn die verschiedenen Experimentalgruppen unterschiedliche Ausprägungen von Merkmalen besitzen und somit hinsichtlich verschiedener Störvariablen nicht strukturgleich sind. In diesem Fall würden die Störvariablen die abhängige Variable beeinflussen (vgl. Harris 2010, S. 162). Daher besitzt die randomisierte Zuordnung der Probanden

³ Trotz randomisierter Versuchsanordnung ist es ratsam gezielte Fragen, die mögliche Demand-Effekte aufdecken können, am Ende der Studie einzubauen. Dies ist insofern wichtig, da es erforderlich ist, dass Probanden blind gegenüber der zu testenden Hypothesen und der verwendeten Manipulationen sind. Je nach experimentellem Design kann ein Forscher einfache Fragen wie „Was glauben Sie war der Zweck der Studie“ bis hin zu komplexeren „Funneled Debriefings“ (vgl. z.B. Chartrand & Bargh 1996, S. 268) einbauen.

zu den verschiedenen Experimentalbedingungen im Rahmen des Between Subject-Designs einen exponierten Stellenwert. Gerade durch die auf der Wahrscheinlichkeitstheorie basierende Zuordnung der Probanden zu den verschiedenen experimentellen Bedingungen wird gewährleistet, dass die verschiedenen Untersuchungsgruppen vor der Durchführung des Experiments hinsichtlich der Ausprägung der Störvariablen gleich sind und damit die Wirkung von Störvariablen auf die abhängige Variable ausbalanciert ist (vgl. z.B. Harris 2010, S. 159; Ryan 2007, S. 7).

Abschließend ist anzumerken, dass keines dieser beiden Verfahren als besser zu beurteilen ist (vgl. Bröder 2011, S. 77 f.; Harris 2010, S. 150f.). So besitzt das Within Subject-Design die Vorteile, dass forschungsökonomischere und effektivere Ergebnisse erzielt werden, da eine geringere Anzahl an Probanden erforderlich ist und die Wirkung von Störvariablen eliminiert wird. Dennoch sind aber verschiedene Effekte zu berücksichtigen, die zu Ergebnisverzerrungen führen können. Während diese Effekte beim Between Subject-Design reduziert werden, ist hingegen eine größere Stichprobe erforderlich und es besteht die Möglichkeit, dass zusätzliche Störvariablen die abhängige Variable beeinflussen und damit die Ergebnisse verzerren. Daher ist sicherzustellen, dass die Probanden den verschiedenen Experimentalbedingungen randomisiert zugeordnet werden. Zusammenfassend ist damit vor der jeweiligen Zielstellung der Untersuchung und in Abhängigkeit des Aufbaus der experimentellen Versuchsanordnung die Zweckmäßigkeit beider Verfahren abzuwägen (vgl. Harris 2010, S. 150f.).

Feld- und Laborexperimente

Neben den bislang diskutierten Kriterien zur Klassifikation von Experimenten können diese auch anhand der Experimentalsituation charakterisiert und voneinander abgegrenzt werden. Zur Differenzierung dienen hierbei die Umwelt- und Kontextbedingungen der experimentellen Versuchsanordnung. Durch diese werden die Repräsentativität und die Generalisierbarkeit der Untersuchungsergebnisse sowie die Wirkung von Störvariablen beeinflusst (vgl. Bereikoven, Eckert & Ellenrieder 2009, S.149; Bortz & Döring 2006, S. 57). Hinsichtlich der Untersuchungssituation sind in der experimentellen Forschung Feld- und Laborexperimente voneinander abzugrenzen.

Laborexperimente sind mit einer künstlichen, planmäßig vereinfachten Erhebungssituation gleichzustellen, die speziell für den zugrundeliegenden Untersuchungszweck geschaffen wird (vgl. Bereikoven, Eckert & Ellenrieder 2009, S. 149). Diese künstliche Erhebungssituation erlaubt es, die interessierende unabhängige Variable systematisch zu variieren, während zeitgleich alle anderen Variablen, die einen Einfluss auf die zu messende abhängige Variable besitzen, kontrolliert oder konstant gehalten werden (vgl. Bortz & Döring 2006, S. 57). Laborexperimente ermöglichen somit eine bestmögliche isolierte Betrachtung von kausalen Wirkungsbeziehungen, da der Einfluss von insbesondere bedingungs- und situationsgebundenen Störvariablen durch die künstlich geschaffenen Umwelt- und Kontextbedingungen eliminiert

wird (vgl. Berekoven, Eckert & Ellenrieder 2009, S. 149; Bortz & Döring 2006, S. 57). Diese künstlichen Rahmenbedingungen der experimentellen Versuchsanordnung gehen aber mit einer geringen Realitätsnähe der Untersuchungssituation einher, da die Erhebung weitestgehend isoliert von natürlichen Umfeldbedingungen stattfindet. Damit besitzen Laborexperimente eine relativ geringe Repräsentativität und Generalisierbarkeit. Dennoch ist diese Form des Experiments insbesondere dann geeignet, wenn zwischen den Untersuchungsvariablen grundsätzliche Effekte zu identifizieren sind (vgl. Bortz & Döring 2006, S. 57).

Feldexperimente finden hingegen unter realitätsnahen Bedingungen statt, d.h. der Untersuchungsgegenstand wird nicht aus seiner natürlichen Umgebung herausgelöst (vgl. Berekoven, Eckert & Ellenrieder 2009, S. 149). Somit besitzen Feldexperimente im Vergleich zu Laborexperimenten eine höhere Repräsentativität und Generalisierbarkeit. Die Durchführung des Experiments unter natürlichen Umwelt- und Kontextbedingungen erschwert jedoch die Kontrolle/Konstanthaltung von Variablen, die neben der interessierenden unabhängigen Variable auch einen Einfluss auf die abhängige Variable besitzen. Da bei Feldexperimenten der Einfluss von bedingungs- und situationsgebundenen Störvariablen auf die abhängige Variable somit nicht vollständig eliminiert werden kann, ist eine isolierte Betrachtung der kausalen Wirkungsbeziehungen zwischen den Untersuchungsvariablen nur eingeschränkt möglich (vgl. Churchill & Iacobucci 2004, S. 293).

Eine trennscharfe Abgrenzung zwischen Labor- und Feldexperimenten ist häufig allerdings nicht gegeben, da beide Untersuchungsformen fließend ineinander übergehen (vgl. Bröder 2011, S. 70). Weiterhin ist keines dieser beiden Verfahren als grundsätzlich besser zu bewerten. Während Feldexperimente unter natürlichen Bedingungen stattfinden und somit eine höhere Repräsentativität besitzen, erlauben Laborexperimente eine bessere Kontrolle von Störvariablen (vgl. Bortz & Döring 2006, S. 57). Im Hinblick auf die spezifischen Vorteile beider Verfahren werden in der praktischen Anwendung deshalb häufig Kompromisslösungen aus Labor- und Feldexperimenten eingesetzt (vgl. Bortz & Döring 2006, S. 57). Die Zweckmäßigkeit beider Verfahren ist folglich vor der Zielstellung der jeweiligen Untersuchung zu diskutieren (vgl. Bröder 2011, S. 70).

2.3 Experimentelle Versuchspläne/-anordnungen

In der experimentellen Forschung werden Versuchspläne, welche den Aufbau bzw. die Struktur von experimentellen Untersuchungsanordnungen beschreiben, in Prä-experimentelle Pläne und experimentelle Pläne unterschieden (vgl. Campbell 1957). Die nachfolgende Tabelle illustriert die geläufigsten experimentellen Versuchspläne, welche im Folgenden kurz vorgestellt werden.

Bei der *One-Shot Fallstudie* (Design 1) wird der Zusammenhang zwischen einer unabhängigen Variable, die bewusst verändert wird (Treatment), und einer abhängigen Variable betrach-

tet. Bei diesem Versuchsplan werden die Probanden dem Treatment (z.B. einer neuen Plakat-anzeige) ausgesetzt und infolgedessen das Ergebnis (z.B. die Beurteilung der Anzeige, Einstellung gegenüber der Marke) ermittelt, d.h. die interessierende abhängige Variable gemessen. Damit sind bei One-Shot Fallstudien die konstituierenden Merkmale eines Experiments (z.B. dass die unabhängige Variable mindestens zwei Ausprägungen aufweisen muss, randomisierte Zuweisung der Teilnehmer zu den Experimentalbedingungen) nicht erfüllt. Es können weiterhin keinerlei Vergleiche angestellt werden, da lediglich eine Messung der abhängigen Variable nach der Präsentation des Treatments erfolgt. Campbell (1957) nutzt diesen Versuchsplan jedoch als grundsätzlichen Ausgangspunkt zur Erklärung von experimentellen Versuchsanordnungen.

Der *Eingruppen-Pretest-Posttest-Plan* (Design 2) erweitert Design 1 um eine Pretest-Messung der interessierenden abhängigen Variable, welche damit einmal vor und einmal nach der Präsentation des Treatments gemessen wird (vgl. Bortz & Döring 2010, S. 558; Campbell 1957, S. 300f.). Trotz dieser Erweiterung ist dieser Versuchsplan dennoch kritisch zu erachten, da lediglich ein formaler Vergleich zwischen Vorher- und Nachhermessung der abhängigen Variable möglich ist. Ferner konstatiert Campbell (1957, S. 301f.), dass diesem Versuchsplan eine mangelnde interne Validität unterstellt werden muss, da verschiedene Effekte (z.B. Carry Over-Effekt, Demand-Effekt) auftreten können und unberücksichtigt bleiben, die zu Ergebnisverzerrungen führen können.

Eine weitere Möglichkeit zur Gestaltung von experimentellen Versuchsplänen ist der *Statische Gruppenvergleich* (Design 3) (vgl. Campbell 1957, S. 300), der den Quasi-Experimenten zu zuordnen ist und einen Gruppenvergleich ermöglicht. Dabei wird eine Gruppe, die einem Treatment ausgesetzt wurde, mit einer Gruppe verglichen, die dieses Treatment nicht erhalten hat. Aufgrund der fehlenden Randomisierung können jedoch keine Aussagen darüber getroffen werden, ob die Gruppen bereits von vornherein unterschiedlich waren. Der sog. Gruppeneffekt (bestehende Unterschiede zwischen den zu vergleichenden Gruppen) kann daher nicht vollständig ausgeschlossen werden, wodurch die Aussagekraft der gewonnenen Erkenntnisse eingeschränkt wird und die Ermittlung der experimentellen Wirkung erschwert wird (vgl. Berekoven, Eckert & Ellenrieder 2009, S. 151).

Als echtes Experiment gilt hingegen der *Pretest-Posttest-Kontrollgruppenplan* (Design 4). Im Rahmen dieses Versuchsplans werden sowohl bei der Experimentalgruppe als auch der Kontrollgruppe eine Vorher- und eine Nachher-Messungen durchgeführt, wobei lediglich die Experimentalgruppe dem Treatment ausgesetzt wird. Die Aufteilung der Stichprobe in Experimental- und Kontrollgruppe erfolgt dabei randomisiert. Campbell (1957) bescheinigt dieser Versuchsanordnung eine hohe interne sowie eine angemessene externe Validität. Jedoch merkt er an, dass aufgrund der Vorher-Messung Carry-Over Effekte, d.h. Lerneffekte, nicht kontrolliert werden können.

Eine Versuchsanordnung, mit welcher alle systematischen Verzerrungen bedingt durch den Gruppen- und dem Carry-Over Effekt kontrolliert werden können, ist der Solomon-Viergruppenplan (Design 5). Die Teilnehmer werden randomisiert vier Untersuchungsgruppen zugewiesen. Gruppe 1 wird dem Treatment ausgesetzt, wobei vor und nach dem Treatment eine Messung der interessierenden abhängigen Variable erfolgt. Gruppe 2 absolviert nur diese Vorher-Nachher-Messung ohne jedoch einem Treatment ausgesetzt zu sein. Im Gegensatz dazu finden bei den Gruppen 3 und 4 lediglich Nachher-Messungen statt, wobei Gruppe 3 dem Treatment ausgesetzt wird. Obwohl mit dieser experimentellen Versuchsanordnung zuverlässige und gültige Ergebnisse erzielt werden, ist kritisch anzumerken, dass dieses experimentelle Design sehr aufwendig ist und eine sehr große Stichprobe erfordert⁴.

Die am häufigsten verwendete experimentelle Versuchsanordnung ist der Posttest-Kontrollgruppenplan (Design 6), welcher auf Vorher-Messungen der interessierenden abhängigen Variable bei der Experimental- und Kontrollgruppe verzichtet. Dieses Verfahren wird von Campbell (1957, S. 311) favorisiert. Auch wenn sich Ähnlichkeiten zu Design 3 ergeben, liegt der zentrale Unterschied in der randomisierten Zuordnung der Teilnehmer zu den einzelnen Bedingungen (Experimentalgruppe, Kontrollgruppe).

⁴ Im Fall einer unabhängigen Variable mit drei Ausprägungen oder sogar einer zweiten und dritten unabhängigen Variable wäre ein experimenteller Versuchsaufbau nach dem Solomon-Viergruppenplan nicht mehr praktikierbar.

Tabelle 3: Übersicht über die verschiedenen experimentellen Versuchspläne

Bezeichnung	Versuchsanordnung
Prä-Experimente	
One-Shot Fallstudie	$X \rightarrow O$
Eingruppen-Pretest-Posttest-Plan	$O_1 \rightarrow X \rightarrow O_2$
Statischer Gruppenvergleich	$X \rightarrow O_1$ O_2
Echte Experimente	
Pretest-Posttest-Kontrollgruppenplan	$[R] O_1 \rightarrow X \rightarrow O_2$ $[R] O_3 \quad O_4$
Solomon-Viergruppenplan	$[R] O_1 \rightarrow X \rightarrow O_2$ $[R] O_3 \quad O_4$ $[R] \quad X \rightarrow O_5$ $[R] \quad O_6$
Posttest-Kontrollgruppenplan	$[R] X \rightarrow O_1$ $[R] \quad O_2$
Legende:	
X: Treatment, d.h. eine Gruppe wird einer experimentellen Variable oder einem Ereignis ausgesetzt.	
O: Beobachtung oder Messung (z.B. durch Verhaltensbeobachtung, Zuhören, Aufzeichnungen, Interview, Fragebögen, Zählungen, etc.).	
→: Die Pfeilrichtung impliziert die zeitliche Präzedenz.	
[R]: Randomisierte Zuordnung zu den Bedingungen.	

Quelle: eigene Darstellung in Anlehnung an Campbell (1957)

3. Gütekriterien der experimentellen Forschung

3.1 Kriterien zur Beurteilung der konzeptionellen Güte des Experiments

Auch bei Experimenten ist die Qualität und die Aussagekraft der erhobenen Daten sicherzustellen und damit verbunden die Zuverlässigkeit und die Gültigkeit der durchgeführten Messung zu beurteilen. Die Gütekriterien der Reliabilität und der Validität sind daher auch im Rahmen von experimentellen Versuchsanordnungen zu bestimmen (vgl. Sedlmeier & Renkewitz 2013, S. 674; Homburg & Giering 1996, S. 6), wobei die Validität eine exponierte Stellung einnimmt. Von besonderer Bedeutung sind hierbei die Beurteilung der internen Validität, der externen Validität und der Konstruktvalidität, da diese die konzeptionelle Güte der experimentellen Versuchsanordnung bzw. der Experimentalsituation festlegen und damit die Generalisierbarkeit sowie die Repräsentanz der erzielten Ergebnisse bestimmen (vgl. Cook & Campbell 1979, S. 37ff.).

Dabei liegt *externe Validität* vor, wenn die durch das Experiment erzielten Ergebnisse generalisierbar sind, d.h. von der Stichprobe Rückschlüsse auf die Grundgesamtheit gezogen werden

können (vgl. Eid, Gollwitzer & Schmitt 2010, S. 61; Berekoven, Eckert & Ellenrieder 2009, S. 82; Churchill & Iacobucci 2004, S. 281). Die externe Validität fordert die Repräsentanz der Untersuchungssituation und der Auskunftspersonen sowie die Verwendung von gültigen Messinstrumenten (vgl. Berekoven, Eckert & Ellenrieder 2009, S. 82). Extern valide Ergebnisse werden hinsichtlich der Umwelt- und Kontextbedingungen der Experimentalsituation insbesondere durch Feldexperimente erzielt, da diese im Vergleich zu Laborexperimenten unter realitätsnahen Bedingungen ablaufen (vgl. Bortz & Döring 2006, S. 57).

Als spezifische Form der externen Validität ist die Konstruktvalidität einzuordnen. Diese Validitätsart ist allerdings von der aus der klassischen Testtheorie bekannten Konstruktvalidität abzugrenzen (vgl. Bortz & Döring 2006, S. 504). So bezieht sich die Konstruktvalidität der experimentellen Forschung nicht auf die Beurteilung der materiellen Genauigkeit der Merkmalerfassung, sondern auf die Beurteilung der verschiedenen Experimentalbedingungen (vgl. Eid, Gollwitzer & Schmitt 2010, S. 60; Bortz & Döring 2006, S. 504). Konstruktvalide Ergebnisse liegen vor, wenn die verschiedenen Experimentalbedingungen sich hinsichtlich der Ausprägung der interessierenden unabhängigen Variable voneinander unterscheiden, aber hinsichtlich der Ausprägung aller anderen Merkmale zueinander identisch sind. Mit Ausnahme der systematisch zu manipulierenden Variable müssen demnach alle Experimentalgruppen zueinander strukturgleich sein. Da diese Forderung u.a. durch eine randomisierte Zuordnung der Probanden zu den Experimentalbedingungen erfüllt wird, ist die Randomisierung eine notwendige Voraussetzung der Konstruktvalidität (vgl. Eid, Gollwitzer & Schmitt 2010, S. 60; Bortz & Döring 2006, S. 504; Churchill & Iacobucci 2004, S. 282).

Neben der externen Validität ist die interne Validität bei experimentellen Untersuchungen zu beurteilen, die gegeben ist, wenn die Veränderung der abhängigen Variable ausschließlich auf die systematische Manipulation der unabhängigen Variable zurückzuführen ist (vgl. Berekoven, Eckert & Ellenrieder 2009, S. 82; Churchill & Iacobucci 2004, S. 279). Die interne Validität fordert damit, dass die Wirkung von allen anderen unabhängigen Variablen sowie von Störvariablen zu kontrollieren bzw. konstant zu halten ist. Diese Forderung erfüllen v.a. Laborexperimente, die charakteristisch unter künstlichen Bedingungen ablaufen und daher eine bessere Kontrolle von Störvariablen ermöglichen (vgl. Eid, Gollwitzer & Schmitt 2010, S. 55; Berekoven, Eckert & Ellenrieder 2009, S. 82). Bei der internen Validität ist darüber hinaus zu berücksichtigen, dass diese durch verschiedene Faktoren beeinträchtigt wird (vgl. Bortz & Döring 2006, S. 502f.; Campbell & Stanley 1963, S. 5f.):

- externe zeitliche Einflüsse: die erfassten Unterschiede könnten auf einen allgemeinen zeitlichen Wandel zurückgeführt werden und nicht alleinig auf die Versuchsbedingung.
- Reifungsprozess: die Probanden verändern unabhängig von der Untersuchung ihr Verhalten.
- Testübung: das eingesetzte Messinstrument beeinflusst das Verhalten der Probanden.

- mangelnde instrumentelle Reliabilität: das eingesetzte Messinstrument ist ungenau/fehlerhaft und damit nicht reliabel, so dass der Zufallsfehler zu Ergebnisverzerrungen führt.
- statistische Regressionseffekte: ohne eine randomisierte Zuordnung der Probanden zu den Experimentalbedingungen könnten die erfassten Veränderungen der abhängigen Variable statistisch bedingt sein.
- experimentelle Mortalität: sollten die Teilnehmer einer Untersuchungsgruppe eine geringe Bereitschaft besitzen, am Experiment teilzunehmen, kann dies die Aussagekraft der gewonnenen Ergebnisse verringern.

Zusammenfassend zeigt ein Vergleich der Inhalte und der Zielstellungen der externen und der internen Validität, dass beide Validitätsarten konfliktär zueinander sind und eine asymmetrische Beziehungen zueinander besitzen. Eine hohe externe Validität führt zu einer niedrigen internen Validität, während eine hohe interne Validität mit einer geringen externen Validität einhergeht. Trotz dieser konkurrierenden Beziehung sind bei experimentellen Untersuchungen dennoch beide Validitätsarten gleichermaßen zu berücksichtigen (vgl. Berekoven, Eckert & Ellenrieder 2009, S. 82; Bortz & Döring 2006, S. 502). Wilson, Aronson und Carlsmith (2010, S. 59) sprechen in diesem Zusammenhang von einem grundsätzlichen Dilemma. Es liegt damit im Ermessen des Forschers, welche Validität für seine Zwecke wichtiger ist (vgl. Churchill & Iacobucci 2004, S. 131). Ist es beispielsweise wichtiger, zunächst einen Effekt zu identifizieren und eindeutig zuordnen zu können, oder ist es wichtiger, das Ergebnis zu verallgemeinern und dafür andere Störvariablen, die ebenfalls für den Effekt verantwortlich sein könnten, in Kauf zu nehmen.

3.2 Kriterien zur Beurteilung der Güte der eingesetzten Messinstrumente

Sowohl die externe Validität als auch die interne Validität setzen die Verwendung von zuverlässigen und gültigen Messinstrumenten voraus. Daher ist neben der Beurteilung der konzeptionellen Güte der Experimentalsituation auch die psychometrische Güte der eingesetzten Skalen/Messinstrumente zu bestimmen (vgl. Campbell & Stanley 1963, S. 5f.). Hierzu dienen die Gütemaße der klassischen Testtheorie zur Beurteilung der formalen Genauigkeit (Reliabilität) und der materiellen Genauigkeit (Validität) der Merkmalerfassung. Die Prüfmaße und kritischen Schwellenwerte dieser beiden Gütekriterien sind in der nachfolgenden Tabelle zusammenfassend dargestellt.

Tabelle 4: Gütekriterien zur Reliabilitäts-/Validitätsprüfung und deren Schwellenwerte

Gütemaß	Schwellenwert	Quelle
Gütekriterien zur Reliabilitätsprüfung		
Voraussetzung: Eindimensionalität des Messinstruments		
Cronbachs Alpha	$\geq 0,7$	vgl. Nunnally & Bernstein (1994, S. 252)
Item-to-Total-Korrelation	$\geq 0,3$	vgl. Bearden, Netemeyer & Teel (1989, S. 475)
Indikatorreliabilität	Faktorladung $\geq 0,7$	vgl. Carmines & Zeller (1979, S. 27)
	Signifikanz der Faktorladung	vgl. Homburg & Giering (1996, S. 16)
Konstruktreliaibilität	$\geq 0,7$	vgl. Bagozzi & Yi (1988, S. 80)
durchschnittlich extrahierte Varianz	$\geq 0,5$	vgl. Fornell & Larcker (1981, S. 46)
Gütekriterien zur Validitätsprüfung		
Inhaltsvalidität	<ul style="list-style-type: none"> Messinstrument bildet alle Bedeutungsinhalte des Konstruktes ab 	vgl. Churchill (1979, S. 69)
	<ul style="list-style-type: none"> Eindimensionalität des Messinstruments 	vgl. Krafft, Götz & Liehr-Gobbers (2005, S. 73)
Konvergenzvalidität	<ul style="list-style-type: none"> Faktorladungen $\geq 0,7$ 	vgl. Carmines & Zeller (1979, S. 27)
	<ul style="list-style-type: none"> Konstruktreliaibilität $\geq 0,7$ 	vgl. Bagozzi & Yi (1988, S. 80)
	<ul style="list-style-type: none"> durchschnittlich extrahierte Varianz $\geq 0,5$ 	vgl. Fornell & Larcker (1981, S. 46)
Diskriminanzvalidität	Fornell-Larcker-Kriterium	vgl. Fornell & Larcker (1981, S. 46)
nomologische Validität	theoretische Begründung der Zusammenhänge zwischen den Konstrukten	vgl. Churchill (1979, S. 69)

Quelle: eigene Darstellung

Literaturverzeichnis

- Bagozzi, Richard P.; Yi, Youjiae (1988), On the Evaluation of Structural Equation Models, in: Journal of the Academy of Marketing Science, Vol. 16 (1), S. 74–94.
- Bearden, William O.; Netemeyer, Richard G.; Teel, Jesse E. (1989), Measurement of Consumer Susceptibility to Interpersonal Influence, in: Journal of Consumer Research, Vol. 15 (4), S. 473–481.
- Bereikoven, Ludwig; Eckert, Werner; Ellenrieder, Peter (2009), Marktforschung: Methodische Grundlagen und praktische Anwendungen, 12. Aufl., Wiesbaden.
- Bortz, Jürgen; Döring, Nicola (2006), Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, 4. Aufl., Heidelberg.
- Bröder, Arndt (2011), Versuchsplanung und experimentelles Praktikum, 1. Aufl., Götting.
- Campbell, Donald T. (1957), Factors Relevant to the Validity of Experiments in Social Settings, in: Psychological Bulletin, Vol. 54 (4), S. 297-312.
- Campbell, Donald T.; Stanley, Julian C. (1963), Experimental and Quasi-Experimental Designs for Research, Boston.
- Carmines, Edward G.; Zeller, Richard A. (1979), Reliability and Validity Assessment, Thousand Oaks.
- Chartrand, Tanya L.; Bargh, John A. (1996), Automatic Activation of Impression Formation and Memorizing Goals: Nonconscious Goal Priming Reproduces Effects of Explicit Task Instruction, in: Journal of Personality and Social Psychology, Vol. 71, 464–478.
- Churchill, Gilbert A. (1979), A Paradigm for Developing better Measures of Marketing Constructs, in: Journal of Marketing Research, Vol. 16 (1), S. 64–73.
- Churchill, Gilbert A.; Iacobucci, Dawn (2004), Marketing Research. Methodological Foundations, 9. Aufl., Mason, Ohio.
- Cook, Thomas D.; Campbell, Donald T. (1979), Quasi-Experimentation. Design & Analysis Issues for Field Settings, Boston.
- Eid, Michael; Gollwitzer, Mario; Schmitt, Manfred (2010), Statistik und Forschungsmethoden, 1. Aufl., Weinheim.
- Fornell, Claes; Larcker, David F. (1981), Evaluating Structural Equation Models with Unobservable Variables and Measurement Error, in: Journal of Marketing Research, Vol. 18 (1), S. 39–50.
- Harris, Peter (2010), Designing and Reporting Experiments in Psychology, 3. Aufl., New York.
- Harris, Peter (2008), Designing and Reporting Experiments in Psychology, 2. Aufl., New York.

- Haslam, Alexander S.; McGarty, Craig (2004), Experimental Design and Causality in Social Psychological Research, in: *Sansone, Carol; Morf, Carolyn C.; Panter, Abigail T.* (Hrsg.), *Handbook of Methods in Social Psychology*, Thousand Oaks, S. 237–264.
- Homburg, Christian; Giering, Annette (1996), Konzeptualisierung und Operationalisierung komplexer Konstrukte. Ein Leitfaden für die Marketingforschung, in: *Marketing: Zeitschrift für Forschung und Praxis*, Vol. 18 (1), S. 3–24.
- Kantowitz, Barry H.; Roediger, Henry L., III; Elmes, David G. (2009), *Experimental Psychology*, 9. Aufl., Stamford.
- Krafft, Manfred; Götz, Oliver; Liehr-Gobbers, Kerstin (2005), Die Validierung von Strukturgleichungsmodellen mit Hilfe des Partial-Least-Squares (PLS)-Ansatzes, in: *Bliemel, Friedhelm; Eggert, Andreas; Fassott, Georg; Henseler, Jörg* (Hrsg.), *Handbuch PLS-Modellierung. Methode, Anwendung, Praxisbeispiele*, 1. Aufl., Stuttgart, S. 71–86.
- Nunnally, Jum C.; Bernstein, Ira H. (1994), *Psychometric Theory*, 3. Aufl., New York.
- Reis, Harry T; Gosling, Samuel D. (2010), Social Psychology Methods Outside the Laboratory, in: *Fiske, Susan T; Gilbert, Daniel T; Gardner; Lindzey* (Hrsg.), *Handbook of Social Psychology*, 5. Aufl., New Jersey, S. 82–114.
- Ryan, Thomas P. (2007), *Modern Experimental Design*, 1. Aufl., New Jersey.
- Sedlmeier, Peter; Renkewitz, Frank (2013), *Forschungsmethoden und Statistik. Ein Lehrbuch für Psychologen und Sozialwissenschaftler*, 2. Aufl., Hallbergmoos.
- Wilson, Timothy D.; Aronson, Elliot; Carlsmith, Kevin (2010), The Art of Laboratory Experimentation, in: *Fiske, Susan T.; Gilbert, Daniel T.; Lindzey, Gardner* (Hrsg.), *Handbook of Social Psychology*, 51–81.