

## 7.4 Korrelationsforschung und korrelative Designs

### 7.4.1 Einführung

In der Korrelationsforschung werden Zusammenhänge zwischen zwei oder mehreren Merkmalen untersucht. Während die experimentelle Forschung darauf abzielt, Variationen der abhängigen Variable(n) durch die Manipulation von unabhängigen Variablen *selbst herzustellen*, geht es in der Korrelationsforschung um Zusammenhänge zwischen *bereits existierenden* Variationen zwischen Merkmalen von Individuen, sozialen Gruppen oder anderen interessierenden Merkmalsträgern (vgl. Cronbach, 1957): Gibt es z.B. einen Zusammenhang zwischen Intelligenz und Berufserfolg? Wie hängen Persönlichkeitsmerkmale mit der Bewältigung von Stress zusammen? Welche soziodemographischen (z.B. Alter, Geschlecht, Bildungsstand) und welche Persönlichkeitsmerkmale (z.B. Extraversion, Gewissenhaftigkeit, Impulsivität) hängen mit welchen Arten der Internetnutzung zusammen?

Experiment vs. Korrelation

Das Experiment gilt für die psychologische Forschung zwar als der „Königsweg“, weil mit dieser Methode kausale Beziehungen zwischen unabhängigen und abhängigen Variablen untersucht und aufgedeckt werden können. Die experimentelle Methode setzt allerdings voraus, dass sich unabhängige Variablen willkürlich variieren und Störvariablen möglichst vollständig kontrollieren lassen. Diese Bedingungen können bei vielen psychologischen und insbesondere sozialwissenschaftlichen Fragestellungen aus prinzipiellen, ökonomischen und ethischen Gründen nicht realisiert werden. Wenn z.B. die Bedeutung von Persönlichkeitsmerkmalen für Gesundheit und psychisches Wohlbefinden untersucht werden soll, dann können Persönlichkeitsmerkmale wie z.B. Extraversion, Gewissenhaftigkeit und emotionale Labilität nicht willkürlich variiert, d.h. hergestellt werden, da diese Eigenschaften als sogenannte Organismusvariablen bereits in einer bestimmten Ausprägung vorliegen. Ein Experimentator kann eine Person nicht gewissenhafter oder extravertierter machen als sie ist. Ein anderes Beispiel ist das Geschlecht eines Untersuchungsteilnehmers, das vom Experimentator natürlich auch nicht willkürlich verändert werden kann. Organismusvariablen können in einem Experiment zwar nicht systematisch variiert, aber selektiert werden, d.h. das Geschlecht einer Versuchsperson kann als zusätzlicher Faktor in einem mehrfaktoriellen Design berücksichtigt werden. Bei anderen Fragestellungen ist ein experimenteller Zugang aus ökonomischen oder ethischen Gründen nicht möglich. Ein weiterer Vorteil der Korrelationsforschung gegenüber experimentellen Designs ist, dass Zusammenhänge zwischen vielen Variablen untersucht werden können. In einem Experiment werden dagegen in der Regel nur wenige unabhängige und abhängige Variable berücksichtigt.

Grenzen der Kontrollierbarkeit

Korrelative Zusammenhänge dürfen nicht kausal interpretiert werden.

Verfahren zur Analyse von Zusammenhängen kommen in der sozialwissenschaftlichen Forschung (z.B. Soziologie, Politologie, Psychologie) aus den genannten Gründen häufiger zum Einsatz als die experimentelle Methode, sie haben allerdings einen entscheidenden Nachteil: *Korrelative Zusammenhänge dürfen nicht kausal interpretiert werden*; es ist lediglich möglich, durch bestimmte korrelative Designs (Längsschnittstudien und insbesondere cross-lagged panel designs, s.u.) und/oder inhaltliche Überlegungen die Anzahl kausaler Erklärungsalternativen einzuschränken bzw. zu falsifizieren. Jedoch hängt die Frage nach der kausalen Interpretierbarkeit korrelativer Zusammenhänge auch vom zugrundeliegenden Kausalitätskonzept ab (vgl. 4.1). Korrelative Zusammenhänge sind zudem nicht deterministisch, wie viele funktionale Zusammenhänge, die z.B. in der Physik ermittelt wurden, sondern lediglich stochastisch (zufallsabhängig) und damit nur mit einer gewissen Wahrscheinlichkeit zutreffend.

Um Missverständnisse zu vermeiden, sei gleich an dieser Stelle darauf hingewiesen, dass Korrelationsforschung nicht auf eine einzige statistische Prozedur beschränkt ist, etwa den Bravais-Pearson Korrelationskoeffizienten. Auch regressions- und faktorenanalytische Analyseverfahren (s.u.) zählen zur Korrelationsforschung, denn auch damit werden bestimmte Arten von Zusammenhängen ermittelt. Die statistischen Verfahren zur Untersuchung und Überprüfung von Zusammenhängen werden insbesondere im Modul 2, aber auch im Modul 6 und 7 vermittelt. In diesem Abschnitt wird lediglich ein inhaltlicher, weitgehend nicht-technischer Überblick zu verschiedenen korrelativen Designs bzw. Zusammenhangshypothesen gegeben.

### 7.4.2 Untersuchung bivariater Zusammenhangshypothesen

vermutete Assoziation zwischen zwei Merkmalen

Ungerichtete Zusammenhangshypothesen

Gerichtete Zusammenhangshypothesen

Bivariate Zusammenhangshypothesen betreffen *vermutete Assoziationen zwischen zwei Merkmalen*. Gibt es z.B. einen Zusammenhang zwischen dem Alter und der durchschnittlichen Internetnutzungszeit pro Woche? Diese Fragestellung lässt sich je nach Vorwissen (bzw. in anderen Fällen theoretisch begründet) als ungerichtete oder gerichtete Zusammenhangshypothese formulieren. *Ungerichtete Zusammenhangshypothesen* nehmen lediglich eine Assoziation zwischen zwei Merkmalen an und spezifizieren nicht, ob es sich dabei um einen negativen oder positiven Zusammenhang handelt, in unserem Beispiel: das Alter und die Internetnutzung hängen zusammen. *Gerichtete Zusammenhangshypothesen* spezifizieren dagegen die Richtung der Assoziation, z.B. Das Alter hängt negativ mit der durchschnittlichen Internetnutzungszeit zusammen oder anders formuliert: Höheres Alter geht mit niedrigerer Internetnutzung einher. Die Richtung des Zusammenhangs kann auch positiv formuliert werden, in unserem Fall würde die Hypothese dann lauten: Höheres Alter geht mit höherer Internetnutzung einher.

Um diese bivariate (ungerichtete oder gerichtete) Zusammenhangshypothese untersuchen zu können, müssen die beiden Merkmale an einer größeren, möglichst repräsentativen Stichprobe erhoben werden (die durchschnittliche Internetnut-

zungszeit lässt sich einfach, aber relativ ungenau per Selbsturteil einschätzen). Pro Person oder Merkmalsträger würden dann zwei Messwerte resultieren und für die gesamte Stichprobe zwei Messwerte-Reihen mit eindeutig einander zuordenbaren Messwerten. Mit Hilfe eines Korrelationskoeffizienten lässt sich nun ermitteln, in welchem Ausmaß die beiden Merkmale Alter und durchschnittliche Internetnutzungszeit pro Woche gemeinsam variieren bzw. kovariieren. Die Kovariation lässt sich als Art des Zusammenhangs beschreiben: Ein *positiver oder negativer linearer Zusammenhang* ist der Fall, wenn hohe Ausprägungen des einen Merkmals mit hohen bzw. niedrigeren Ausprägungen des anderen Merkmals assoziiert sind, z.B. je höher das Alter desto niedriger die durchschnittliche Internetnutzungszeit (negativer linearer Zusammenhang). Ein *nicht-linearer Zusammenhang* würde z.B. vorliegen, wenn die durchschnittliche Internetnutzungszeit bis zu einem bestimmten Alter abnimmt, im höheren Erwachsenenalter dann aber wieder zunimmt. Eine mögliche Erklärung für diesen fiktiven und konstruierten nicht-linearen Zusammenhang könnte sein, dass Personen im höheren Lebensalter zunehmend mehr das Internet als Kommunikationsmedium für sich entdecken und es besonders intensiv nutzen können, da ihnen viel freie Zeit zur Verfügung steht (sogenannte silver surfer). Ein empirisch gesicherter nicht-linearer Zusammenhang ist das *Yerkes-Dodson-Gesetz* (Yerkes & Dodson, 1908), nach dem die Produktivität bis zu einem mittleren Erregungsniveau kontinuierlich ansteigt, mit höheren Erregungsniveaus dann aber abnimmt (umgekehrt U-förmiger Zusammenhang, vgl. Abbildung 7-3).

positiver oder negativer  
linearer Zusammenhang

nicht-linearer Zusam-  
menhang

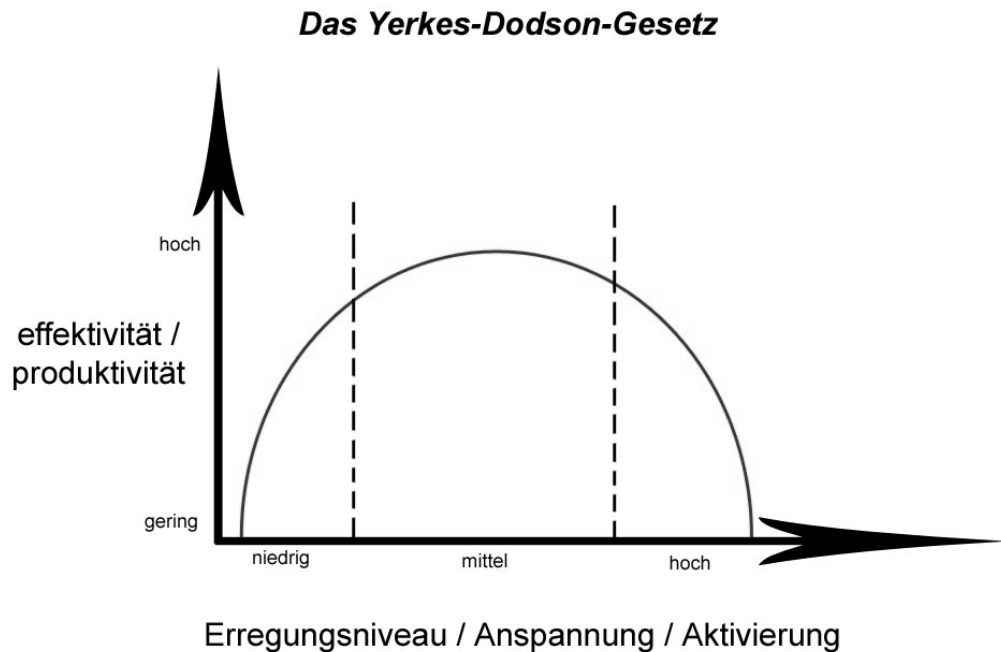


Abb. 7-3: Yerkes-Dodson-Gesetz vgl. die Grafik in Wikipedia

Richtung und Enge des  
Zusammenhangs

Neben der Art des Zusammenhangs (linear vs. nicht-linear) lässt sich noch die Richtung (positiv oder negativ) und die Intensität bzw. Enge bestimmen. Korrelationskoeffizienten können Werte zwischen + 1 und – 1 annehmen. Je höher der Koeffizient desto enger der (positive oder negative) Zusammenhang zwischen 2 Merkmalen. Wird ein Korrelationskoeffizient von + 1 oder – 1 ermittelt, dann liegt ein perfekter linearer und damit nicht mehr stochastischer, sondern deterministischer Zusammenhang vor. Die Ausprägungen des einen Merkmals lassen sich dann bei Kenntnis der Ausprägungen des anderen Merkmals perfekt vorhersagen (über eine lineare Regressionsgleichung, vgl. Modul 2 bzw. nächster Abschnitt). Derart hohe Korrelationen kommen empirisch in der Psychologie aus verschiedenen Gründen so gut wie nie vor, da Erleben und Verhalten immer mit multiplen Bedingungen assoziiert ist, die zudem in Wechselwirkung treten können (vgl. nächster Abschnitt). Abgesehen davon mindern Fehler bei der Messung der interessierenden Merkmale die Enge des Zusammenhangs (vgl. hierzu noch ausführlicher Modul 6, Kurs Testkonstruktion). Nach Konventionen von Cohen (1988) können Korrelationen um +/- .10 als schwache Zusammenhänge, um +/- .30 als mittlere Zusammenhänge und um +/- .50 als starke Zusammenhänge interpretiert werden. Eine Korrelation von +/- .50 zwischen zwei Merkmalen ist noch weit von einem perfekten Zusammenhang entfernt, gilt in der Psychologie aber wegen der multiplen Bedingtheit des Erlebens und Verhaltens schon als starker Zusammenhang.

Welcher Korrelationskoeffizient herangezogen werden kann, hängt vom Skalenniveau der erfassten Merkmale ab. In Tabelle 7-6 werden die Korrelationskoeffi-

zienten in Abhängigkeit vom Skalenniveau lediglich aufgelistet; genauere Informationen erhalten Sie im Modul 2. Unterstellt man in unserem Beispiel den Selbsteinschätzungen der Internetnutzungszeit Intervallskalen-Niveau (was sicher nicht unproblematisch ist), dann kann die Bravais-Pearson-Produkt-Moment-Korrelation berechnet werden.

Tab. 7-6: Bivariaten Korrelationsarten

| Merkmal y          | Merkmal x                  |  |                          |
|--------------------|----------------------------|--|--------------------------|
|                    | Intervallskala             | dichotomes Merkmal                               | Ordinalskala             |
| Intervallskala     | Produkt-Moment-Korrelation | Punktbiserale Korrelation<br>$\Phi$ -Koeffizient | Rangkorrelation          |
| dichotomes Merkmal |                            |  | biserale Rangkorrelation |
| Ordinalskala       | -                          | -  | Rangkorrelation          |
|                    | -                          |  |                          |

Bei dem angeführten Beispiel zum Lebensalter und dem Ausmaß der Internetnutzung handelt es sich um ein *korrelatives Querschnittsdesign*, bei dem die Erhebung der Merkmalsausprägungen nur zu einem bestimmten Messzeitpunkt erfolgt (Genaueres zu Querschnittsdesigns erfahren sie im Modul 4 zur Entwicklungspsychologie).

korrelatives Querschnittsdesign

In einem *Längsschnittsdesign* würde im bivariaten Fall ein und dasselbe Merkmal bei allen Personen einer Stichprobe zu zwei verschiedenen Messzeitpunkten erhoben werden. Dieses Design kommt in der Differentiellen Psychologie und Persönlichkeitsforschung zum Einsatz, wenn die zeitliche Stabilität von Eigenschaften, z.B. Gewissenhaftigkeit, ermittelt werden soll. Eigenschaften als habituelle Erlebens- und Verhaltenstendenzen müssen empirisch u.a. eine Stabilität über die Zeit aufweisen. Um diese zeitliche Stabilität bestimmen zu können, wird eine Retest-Korrelation zwischen den Ausprägungen eines Merkmals bei einer Stichprobe von Personen zum Zeitpunkt  $t_1$  und den Ausprägungen desselben Merkmals zum Zeitpunkt  $t_2$  ermittelt (die Ausprägungen des Merkmals Gewissenhaftigkeit werden z.B. durch einen geeigneten Fragebogen im Selbst- oder Fremdurteil erhoben). Die resultierende Korrelation betrifft die *differentielle Stabilität* des Merkmals, d.h. die zeitliche Stabilität interindividueller Unterschiede. Differentielle Stabilität meint das „Verharren“ von Personen auf denselben relativen Positionen innerhalb einer Gruppe über die Zeit.

Längsschnittsdesign

differentielle Stabilität

|                               |   |
|-------------------------------|---|
| nomothetischer Anwendungsfall | Wer zum Zeitpunkt $t_1$ einen Wert für Gewissenhaftigkeit aufwies, der im Vergleich zu den meisten anderen Probanden in der Stichprobe höher (niedriger) war, der oder die sollte auch zum Zeitpunkt $t_2$ einen höheren (niedrigeren) Wert aufweisen als die meisten anderen Probanden. Die hier beschriebene Restestkorrelation betrifft den <i>nomothetischen Anwendungsfall</i> , d.h. es wird festgestellt, ob die Rangordnung von <i>mehreren Personen</i> einer Stichprobe in Bezug auf <i>ein Merkmal</i> über die Zeit gleich bleibt oder sich verändert.  |
| idiographische Variante       | In der <i>idiographischen Variante</i> wird nicht über Personen sondern über <i>mehrere Merkmale</i> innerhalb <i>einer Person</i> korreliert (Q-Korrelation, vgl. Stephenson, 1952; Cronbach & Gleser, 1953). Es geht dabei also um die Stabilität bzw. Veränderung der relativen Positionen von Merkmalen innerhalb eines individuellen Variablen-Profiles. Dabei handelt es sich um eine Form von ipsativer Stabilität, die als ipsativ-differentiell bezeichnet werden könnte: „Ipsativ“, weil es um die Stabilität bei <i>einer Person</i> geht, und „differentiell“, weil wie im nomothetischen Anwendungsfall das „Verharren“ nicht von Personen, aber von Profilmerkmalen auf denselben Positionen über die Zeit bestimmt wird (vgl. Renner, 2002). |

### 7.4.3 Untersuchung multivariater Zusammenhangshypothesen

Unsere bisher formulierte bivariate Hypothese zum Zusammenhang zwischen Alter und Internetnutzung ist sehr einfach; zu einfach, denn Erleben und Verhalten hängt immer von vielen verschiedenen Merkmalen und Bedingungen ab, die u.U. miteinander interagieren können. Auch für die durchschnittliche Internetnutzungszeit pro Woche sind sicherlich mehr Bedingungen von Bedeutung als nur das Alter. Das Geschlecht, der Bildungsstand, der berufliche und sozioökonomische Status, der Wohnort (Stadt vs. Land und damit Zugang zu DSL oder nicht) und auch Persönlichkeitsmerkmale können mit der durchschnittlichen Internetnutzungszeit assoziiert sein. Auch die Internetnutzung selbst lässt sich vielfältiger beschreiben und erfassen als lediglich durch die durchschnittliche Nutzungszeit pro Woche: Welche Internetdienste werden wie lange genutzt? Soll zwischen beruflicher und privater Internetnutzung unterschieden werden?

|             |   |
|-------------|---|
| Prädiktoren | Wenn der Zusammenhang zwischen mehreren Merkmalen zu einem oder mehreren weiteren Merkmal(en) untersucht werden soll, dann wird formal zwischen <i>Prädiktoren</i> , also Vorhersage-Variablen, und <i>Kriterien</i> , die vorhergesagt werden sollen, unterschieden. Diese Unterscheidung macht besonders dann Sinn, wenn das oder die Kriterien zeitlich gesehen nach der Erfassung der Prädiktoren erhoben werden. In der psychologischen Internetforschung ist dieses methodische Prinzip u. a. in den Studien von Kraut und Mitarbeitern (Kraut, Lundmark, Kiesler, Mukhopadhyay & Scherlis, 1998) umgesetzt worden. Die Autoren erfassten zuerst mehrere Prädiktorvariablen bei einer Stichprobe von Personen aus der Unterschicht und stellten den Probanden dann kostenlos Rechner mit Internetzugang |
| Kriterien   |   |

zur Verfügung. Die Internetnutzung wurde online protokolliert, d.h. die Untersuchungsteilnehmer beobachteten über log-Dateien, wie oft, wie lange und welche Dienste die Probanden im Internet in Anspruch genommen haben. Diese Methode hat gegenüber den oben angedeuteten Selbsteinschätzungen der Internernutzung entscheidende Vorteile, lässt sich aber unter ethischen Gesichtspunkten kontrovers diskutieren (vgl. Abschnitt 10). Die Unterscheidung zwischen Prädiktoren und Kriterien entspricht der Unterscheidung zwischen unabhängigen und abhängigen Variablen im Experiment. In einigen Lehrbüchern (z.B. Cohen, Cohen, West & Aiken, 2003) wird auch in korrelativen Designs von UVs und AVs gesprochen. Wir bevorzugen allerdings die Begriffe Prädiktor und Kriterium, um nicht Gefahr zu laufen, die im Experiment mögliche Manipulation der UV und die Möglichkeit zur Prüfung von Kausalhypothesen stillschweigend auch für die Korrelationsforschung zu implizieren.

Im Folgenden werden die wichtigsten Varianten bzw. Designs zur Untersuchung multivariater korrelativer Zusammenhänge vorgestellt. Wir beginnen mit einem in der Forschungspraxis sehr häufigen Anwendungsfall, bei dem *mehrere* Prädiktoren zur Vorhersage *eines* Kriteriums herangezogen werden. Ein sich systematisch anschließender Fall, der hier nur erwähnt wird, betrifft den sogenannten *kanonischen Zusammenhang* eines Sets aus mehreren Prädiktoren mit einem Set aus mehreren Kriterien. Statistisch wird diese Form des Zusammenhangs mit Hilfe der kanonischen Korrelation ermittelt (vgl. Modul 2).<sup>4</sup>

kanonischer Zusammenhang

<sup>4</sup> Der Begriff „multivariat“ wird in Statistik- und Methodenlehrbüchern nicht ganz konsistent verwendet. Im Zusammenhang mit Experimenten und varianzanalytischen Methoden bedeutet „multivariat“, dass *mehrere abhängige* Variable untersucht werden. Die Berücksichtigung mehrerer unabhängiger Variablen wird dagegen mehrfaktoriell, aber immer noch univariat genannt. In der Korrelations- und Regressionsrechnung wird dagegen auch die multiple Regressionsanalyse (mehrere Prädiktoren, *ein* Kriterium) den multivariaten Methoden (vgl. Z.B. Bortz, 2005) bzw. den multivariaten Zusammenhangshypothesen (Bortz & Döring, 2006) zugeordnet. Wir folgen in diesem Abschnitt dieser Konvention, obwohl eigentlich nur kanonische Zusammenhänge multivariat bezeichnet werden könnten, wenn die Unterscheidung zwischen univariaten und multivariaten Methoden aus der experimentellen und varianzanalytischen Vorgehensweise beibehalten werden würde.

### 7.4.3.1 Multivariate Zusammenhänge 1: Mehrere Prädiktoren und eine Kriteriumsvariable

In der Psychologie werden häufig Zusammenhänge zwischen mehreren Prädiktoren und einem Kriterium untersucht. Zur Veranschaulichung greifen wir noch einmal auf unser Internetbeispiel zurück: Wie hängen das Geschlecht, der Bildungsstand, der berufliche und sozioökonomische Status, der Wohnort und Persönlichkeitsmerkmale mit der durchschnittlichen Internetnutzungszeit zusammen? Nachdem mit geeigneten Methoden Daten zu den genannten Prädiktoren und zum Kriterium erhoben wurden, lässt sich diese Zusammenhangshypothese mit Hilfe der *multiplen Regressionsrechnung* analysieren (vgl. Modul 2). Das Ergebnis einer multiplen Regressionsanalyse ist eine Gleichung zur Vorhersage der Kriteriumswerte und der multiple Korrelationskoeffizient, der den Zusammenhang zwischen den vorhergesagten und den tatsächlichen Kriteriumswerten quantifiziert. Warum werden eigentlich nicht mehrere bivariate Korrelationen berechnet, um den Zusammenhang zwischen den Prädiktoren und dem Kriterium zu bestimmen? Aus verschiedenen Gründen (vgl. dazu ausführlicher Modul 2):

multiple Regressionsrechnung

Vorteile der multiplen Regression

Identifikation substantieller Prädiktoren

Erstens lässt sich eine multiple Regressionsanalyse so durchführen, dass aus einer Vielzahl von Prädiktoren diejenigen bestimmt werden können, die zur Vorhersage des Kriteriums einen substantiellen Beitrag leisten. Anders ausgedrückt lassen sich Prädiktoren identifizieren, die im Kontext der multiplen Vorhersage redundant sind, obwohl sie eine bivariate Korrelation mit dem Kriterium aufweisen. Diese Redundanz kann resultieren, wenn die Prädiktoren untereinander korrelieren und damit ähnliche Merkmalsanteile erfassen, was z.B. für den Bildungsstand und den sozio-ökonomischen Status der Fall ist oder auch für Persönlichkeitsmerkmale wie soziale Ängstlichkeit und Selbstwertschätzung (deutliche negative Korrelation). Wenn die Prädiktoren zu hoch interkorrelieren, dann ist die Validität der Regressionsanalyse u.U. sogar gefährdet (Problem der Multikollinearität, vgl. Modul 2); deswegen sollten die Prädiktoren idealerweise linear unabhängig sein, was empirisch jedoch oft nicht der Fall ist. Ein Beispiel für den Einsatz der multiplen Regressionsanalyse zur Vorhersage der Internetnutzung finden Sie bei Wolfradt und Doll (2005). Die beiden Autoren haben verschiedene Persönlichkeitsmerkmale und das Geschlecht als Prädiktoren für verschiedene Arten der Internetnutzung untersucht. Dabei resultierte u.a., dass die instrumentell-unterhaltungsorientierte Internetnutzung (z.B. Musik und Filme downloaden, die Homepage von Freunden besuchen, online-shopping) bei den männlichen Befragten durch hohe Werte in Extraversion und Offenheit für neue Erfahrungen vorhergesagt werden kann.

Interaktionseffekte

Zweitens können *Interaktionseffekte* zwischen den unterschiedlichen Prädiktoren bestimmt und damit das gemeinsame Zusammenwirken zweier Merkmale auf das Kriterium untersucht werden. In der multiplen Regressionsanalyse werden solche Interaktionen im Rahmen von sogenannten Moderator-Hypothesen geprüft (moderierte Regression, vgl. Abschnitt 7.2.3.2).



Drittens können in einer multiplen Regressionsanalyse manchmal sogenannte *Suppressionseffekte* auftreten. Bei einem Suppressor handelt es sich um eine Prädiktorvariable, die den Vorhersagebeitrag einer (oder mehrerer) anderer Variablen erhöht, in dem sie für die Vorhersage irrelevante Varianzanteile unterdrückt. Ein Beispiel für einen replizierbaren Suppressionseffekt stammt von Paulhus, Robins, Trzesniewski und Tracy (2004). Die Autoren konnten zeigen, dass bei der Vorhersage von anti-sozialem Verhalten sogenannte reziproke bzw. kooperative Suppressionseffekte (vgl. Bortz, 2005, Cohen & Cohen, 1975) auftreten, wenn Narzissmus und Selbstwertschätzung gemeinsam als Prädiktoren in die Regressionsgleichung eingegeben werden. Narzissmus korreliert für sich genommen positiv mit antisozialem Verhalten; Selbstwertschätzung weist dagegen ein inkonsistentes Muster über mehrere Studien hinweg auf: in einigen Studien korreliert die Selbstwertschätzung gar nicht, in anderen leicht positiv und in wieder anderen Studien leicht negativ mit antisozialem Verhalten. Werden Narzissmus und Selbstwertschätzung gemeinsam in einer Regressionsanalyse zur Vorhersage von antisozialem Verhalten eingegeben, so zeigt sich ein konsistent negativer Zusammenhang zwischen Selbstwertschätzung und anti-sozialem Verhalten und ein positiverer Zusammenhang zwischen Narzissmus und antisozialem Verhalten als bei der bivariaten Korrelation. Wie lässt sich diese wechselseitige Verbesserung der Prädiktionsleistung erklären? Zunächst ist zu beachten, dass Narzissmus und Selbstwertschätzung schon auf theoretischer Ebene überlappen: Narzissmus wird als besonders hohe (grandiose), aber fragile Selbstwertschätzung expliziert, während Selbstwertschätzung neutraler als positive oder negative Bewertung der eigenen Person definiert wird. Bei der Erfassung dieser beiden Konstrukte durch einen Fragebogen zur Selbsteinschätzung entsteht nun die Schwierigkeit, dass „normale“ positive Selbstwertschätzung und narzisstische grandiose Selbstwertschätzung auf Itemebene schwer separat erfassbar sind. Anders ausgedrückt werden Personen mit hohen Narzissmusausprägungen auch solche Items bejahen, die zur Erfassung normaler positiver Selbstwertschätzung dienen, wie z.B. „Haben Sie eine positive Einstellung zu sich selbst?“. Andererseits mag es auch Narzissmus-Items geben, die von Personen mit normaler positiver Selbstwertschätzung bejaht werden können, wie z.B. „Ich will erfolgreich sein“. Die Erfassung des Merkmals Selbstwertschätzung ist also durch Narzissmus „kontaminiert“ und umgekehrt. Wenn nun Summenscores für Selbstwertschätzung alleine mit antisozialem Verhalten korreliert werden, dann können die mit der Selbstwertschätzung erfassten Narzissmus-Anteile dazu führen, dass manchmal sogar eine leicht positive Korrelation resultiert. Umgekehrt mindert der Anteil an normaler positiver Selbstwertschätzung in den Narzissmus-Scores die positive Assoziation zwischen Narzissmus und antisozialem Verhalten (Narzissten verhalten sich unempathisch, nutzen andere aus und reagieren leicht gereizt und aggressiv, wenn sie kritisiert werden). Erst wenn beide Merkmale gemeinsam in eine Regressionsgleichung eingegeben werden, tritt eine gegenseitige Unterdrückung (Suppression) der jeweils irrelevanten Varianzanteile auf, die zu einer Erhöhung des negativen bzw. positiven Zusammenhangs von Selbstwertschätzung bzw. Narzissmus mit antiso-

zialem Verhalten führt. Würde man nur bivariate Korrelationen berücksichtigen, dann könnten solche Suppressionseffekte nicht aufgedeckt werden.

### 7.4.3.2 Multivariate Zusammenhänge 2: Untersuchung von Moderator- und Mediatorhypothesen

#### Interaktionseffekte

Erleben und Verhalten ist multideterminiert und deshalb kovariiert ein Kriterium wie die Internetnutzung mit vielen unterschiedlichen Prädiktoren. Im vorherigen Abschnitt wurde bereits angedeutet, dass *Interaktionseffekte* zwischen den Prädiktoren bestimmt und damit das gemeinsame Zusammenwirken von in der Regel zwei Merkmalen auf das Kriterium untersucht werden kann. In der multiplen Regressionsanalyse werden solche Interaktionen im Rahmen von Moderator-Hypothesen sogenannten *Moderator-Hypothesen* untersucht. Ein Moderator ist eine qualitative (z. B. Geschlecht, Schichtzugehörigkeit) oder quantitative (z.B. Ausprägungen eines Persönlichkeitsmerkmals) Variable, die die Richtung und/oder die Enge des Zusammenhangs zwischen einer Prädiktor-Variable und einer Kriteriumsvariable beeinflusst (Baron & Kenny, 1986). Anders ausgedrückt liegt ein Moderatoreffekt dann vor, wenn die Beziehung zwischen zwei Variablen von der Ausprägung einer dritten Variable abhängt (vgl. Abbildung 7-4).

Als Beispiel kann noch einmal der oben bereits skizzierte Befund von Wolfradt und Doll (2005) herangezogen werden: Der Zusammenhang zwischen Extraversi-on bzw. Offenheit für Erfahrung und der unterhaltungsorientierten Internetnutzung gilt nämlich nur für Männer, nicht aber für Frauen. Anders ausgedrückt moderiert das Geschlecht den Zusammenhang zwischen den genannten Variablen. Nur wenn die Geschlechtsausprägung „männlich“ vorliegt, findet sich ein positiver Zusammenhang.

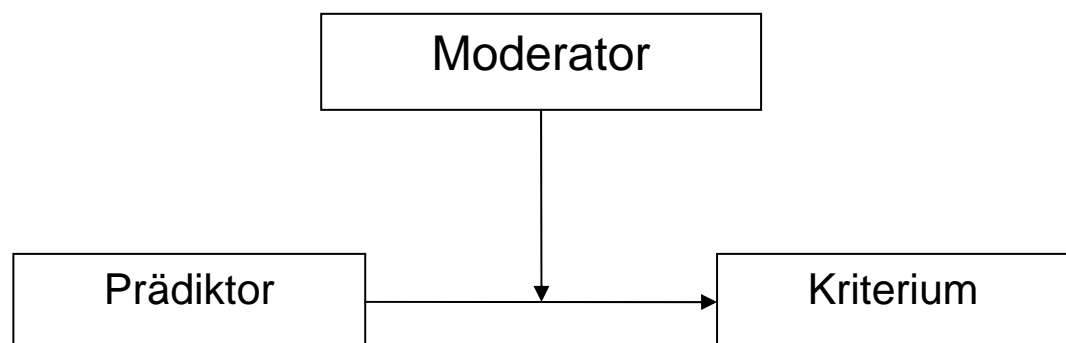


Abb. 7-4: Moderatoreffekt

*Statistisch-technischer Exkurs:* Die nachfolgenden Angaben werden der Vollständigkeit halber angeführt, müssen aber im Zusammenhang mit Modul 2 (Kurs Statistik II) gelesen werden, um verstanden werden zu können. Eine Moderator-Hypothese lässt sich in einem experimentellen Design prüfen, indem die Interaktion zwischen zwei Faktoren geprüft wird (vgl. Modul 2). Im korrelativen Ansatz erfolgt die Überprüfung einer Moderator-Hypothese in einer hierarchischen, moderierten Regression (vgl. Modul 2). Hierarchisch bedeutet, dass in drei Schritten zunächst der Prädiktor, dann zusätzlich der Moderator und im dritten Schritt auch noch ein Produktterm aus dem Prädiktor und Moderator eingegeben werden. Leistet der Produktterm, mit dem der Moderatoreffekt geprüft wird, einen zusätzlichen signifikanten Beitrag zur Vorhersage des Kriteriums, dann kann die Moderatorwirkung in sogenannten simple-slope-Analysen genauer untersucht werden. Dabei werden Regressionsgeraden des Kriteriums auf den Prädiktor bei hohen, durchschnittlichen und niedrigen Ausprägungen eines Moderators bestimmt und es wird getestet, ob die jeweilige Steigung von Null verschieden ist.

Statistisch-technischer  
Exkurs

Während eine Moderator-Variable spezifiziert, unter welchen Bedingungen welche Art des Zusammenhangs zwischen zwei Variablen besteht, erklärt eine *Mediator-Variable* den Prozess oder „Mechanismus“, durch den eine Prädiktorvariable eine Kriteriumsvariable „beeinflusst“ (MacKinnon, Fairchild & Fritz, 2007). Anders ausgedrückt ist ein Mediator eine Variable, die erklärt, *warum* ein Zusammenhang zwischen zwei anderen Variablen besteht. Damit verbunden ist die Idee einer Abfolge von Beziehungen, bei der eine Prädiktorvariable eine Mediatorvariable beeinflusst (vgl. Pfad a in Abbildung 7-5), die dann wiederum eine Kriteriumsvariable (Pfad b) beeinflusst. Durch die Mediation wird der Zusammenhang zwischen Prädiktor und Kriterium (Pfad c) erklärt.

Mediator-Variable

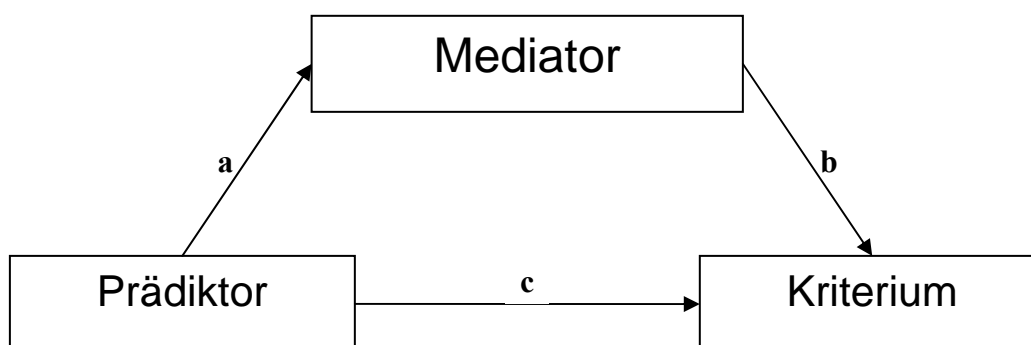


Abb. 7-5: Mediatoreffekt

In der Psychologie lässt sich die Idee der Mediation mit dem sehr verbreiteten S-O-R-Modell, das erstmals von Woodworth (1928) formuliert und später insbesondere im kognitionspsychologischen Paradigma aufgegriffen und ausgearbeitet

wurde veranschaulichen. Demnach resultiert die Reaktion (R) auf einen externen Reiz bzw. Stimulus (S) aufgrund zahlreicher intervenierender Verarbeitungsprozesse innerhalb eines aktiven Organismus (O).

In der persönlichkeitspsychologisch orientierten Internetforschung haben Hertel, Schroer, Batinic, Konradt und Naumann (2005, in press) einen Mediator für den Zusammenhang zwischen emotionaler Labilität (Neurotizismus) und der Präferenz für asynchrone Medien (z.B. E-Mail) identifizieren können, der insbesondere bei ambivalenten Kommunikationsanlässen auftritt, also z.B. wenn Konflikte oder Meinungsverschiedenheiten ausgetragen werden sollen. Die Bedingung Kommunikationsanlass ambivalent vs. nicht ambivalent moderiert also den genannten Zusammenhang! Warum aber bevorzugen emotional labile Personen bei ambivalenten Kommunikationsanlässen eher asynchrone Medien? Emotional labile Personen sind u.a. durch soziale Ängstlichkeit gekennzeichnet und werden deshalb in schwierigen sozialen Interaktionen besonders versuchen, ihre Unsicherheit zu reduzieren und die Kontrolle über die Situation zu erhöhen. Dies ist mit e-mail-basierter Kommunikation eher möglich, da hierbei kein direkter Austausch von Angesicht zu Angesicht stattfindet und genauer und länger überlegt werden kann, was zum Ausdruck gebracht werden soll und wie. Die soziale Angst und die damit verbundene Tendenz zur Reduktion von Unsicherheit fungieren demnach als Mediatoren des Zusammenhangs zwischen emotionaler Labilität und der Präferenz für asynchrone Medien bei ambivalenten Kommunikationsanlässen (vgl. Abbildung 7-6).

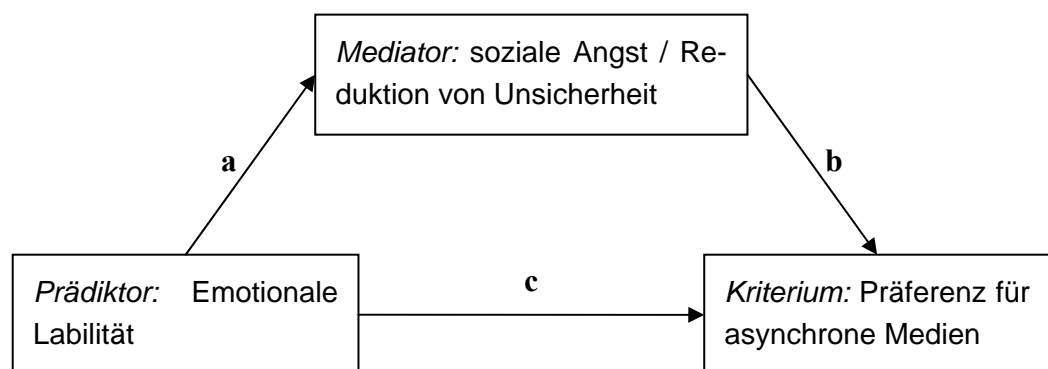


Abb. 7-6: Beispielhafter Mediatoreffekt

*Statistisch-technischer Exkurs:* Die nachfolgenden Angaben werden der Vollständigkeit halber angeführt, müssen aber im Zusammenhang mit Modul 2 (Kurs Statistik II) gelesen werden, um verstanden werden zu können. Ein relativ einfaches Vorgehen, mit dem sich eine Mediator-Hypothese statistisch prüfen lässt, haben Baron und Kenny (1986, p. 1176) vorgeschlagen (es gibt noch andere, weitaus komplexere Verfahren). Demnach müssen folgende Bedingungen erfüllt sein,

damit eine Mediatorhypothese angenommen wird: (1) der Prädiktor ist mit dem angenommenen Mediator korreliert (Pfad a in Abb. 7.2-4), (2) der potenzielle Mediator ist mit dem Kriterium korreliert (Pfad b), (3) wenn die Pfade (a) und (b) kontrolliert werden, verringert sich die Korrelation zwischen Prädiktor und Kriterium und wird bei vollständiger Mediation Null. Diese Bedingungen können im Rahmen von drei Regressionsgleichungen geprüft werden: (1) Regression des Mediators auf den Prädiktor, (2) Regression des Kriteriums auf den Prädiktor und (3) Regression des Kriteriums auf den Prädiktor und den Mediator. Mediation liegt vor, wenn der Effekt des Prädiktors auf das Kriterium in Gleichung 3 geringer ist als in Gleichung 2. Bei vollständiger Mediation ist der Effekt des Prädiktors in Gleichung 3 Null. Bei diesem Vorgehen werden u.a. sogenannte

Partialregressionskoeffizienten bestimmt (vgl. Modul 2). Das Konzept der Mediation ist zudem in *partiellen Zusammenhangshypothesen* impliziert (vgl. Bortz & Döring, 2006, S. 510f), die besagen, dass der Zusammenhang zwischen zwei Variablen X und Y unbedeutend wird, wenn man den „Einfluss“ einer dritten Variablen Z „ausschaltet“. Anders ausgedrückt wird der Zusammenhang zwischen X und Y um den Einfluss von Z bereinigt. Das statistische Verfahren zur Bestimmung und Prüfung solcher partiellen Zusammenhangshypothesen ist die Partialkorrelation (vgl. Modul 2 und einige Beispiele bei Bortz & Döring, 2006, S. 510f).

partiellen Zusammenhangshypothesen

Die *Unterscheidung zwischen Moderatoren und Mediatoren* spielt auch in der psychologischen Stressforschung eine große Rolle (vgl. zusammenfassend Matthews, Deary & Whiteman, 2003, p. 255ff). So kann der Zusammenhang zwischen kritischen Lebensereignissen (z.B. Tod einer nahestehenden Person) und Depression durch soziale Unterstützung moderiert werden. Bei hoher sozialer Unterstützung ist der Zusammenhang niedrig, bei geringer sozialer Unterstützung kann der Zusammenhang dagegen hoch ausfallen. Der in vielen Studien nachgewiesene Zusammenhang zwischen dem Persönlichkeitsmerkmal Neurotizismus (emotionale Labilität) und Stress-Symptomen (z.B. Angst, Niedergeschlagenheit, physiologische Symptome) wird mit mediierenden dysfunktionalen Bewertungs- und Bewältigungsprozessen erklärt. So neigen emotional labile Personen dazu potentielle Stressoren, wie z.B. eine Prüfung, als bedrohlicher wahrzunehmen als emotional stabile Personen, zudem schätzen sie ihre Bewältigungs- und Leistungsfähigkeiten niedrig ein.

Unterscheidung zwischen Moderatoren und Mediatoren in der psychologischen Stressforschung

In der empirisch-psychologischen Forschung werden Moderatoren typischerweise dann gesucht, wenn zwischen eine Prädiktor- und einer Kriteriumsvariable unerwartet niedrige oder inkonsistente Zusammenhänge resultierten, d.h. wenn ein Zusammenhang in unterschiedlichen Stichproben bzw. Populationen unterschiedlich ausfällt. Mediatoren werden dagegen insbesondere dann analysiert, wenn deutliche und gesicherte Zusammenhänge zwischen zwei Variablen bestehen (Baron & Kenny, 1986, p. 1178).

Forschungskontexte für Moderatoren und Mediatoren

### 7.4.3.3 Multivariate Zusammenhänge 3: Faktorielle Zusammenhänge

Im Abschnitt 7 zu den Methoden der Datenerhebung wurde bei der Beschreibung der Fragebogen-Methode darauf hingewiesen, dass einzelne Fragen in der Regel nicht für sich ausgewertet und interpretiert, sondern die Antworten auf eine größere Zahl von Fragen oder Feststellungen nach bestimmten statistischen Prinzipien in einem einzigen Messwert zusammengefasst werden. Wie kommt man nun darauf, bestimmte Items (Fragen, Feststellungen) zu einem Gesamtwert zusammenzufassen? Unter welchen Bedingungen ist es gerechtfertigt, mehrere Items zu einem Gesamt-Score zu kombinieren? Es erscheint - zunächst unabhängig von statistischen Erwägungen - sinnvoll, solche Items zusammenzufassen, die irgendwie dasselbe messen. Um statistisch feststellen zu können, was „irgendwie dasselbe“ sein könnte, ist die Durchführung einer *explorativen Faktorenanalyse* indiziert. Die mit einer Faktorenanalyse verbundene Hypothese ist, dass sich die wechselseitigen Zusammenhänge vieler beobachteter bzw. gemessener Variablen durch wenige, in der Regel voneinander unabhängige (orthogonale) *latente Faktoren* bzw. Dimensionen erklären lassen. Viele miteinander korrelierende Variablen werden somit auf wenige, voneinander unabhängige latente Faktoren reduziert - allerdings kann auch statistisch in einer sogenannten obliquen Rotation zugelassen werden, dass die Faktoren miteinander korrelieren.

explorative  
Faktorenanalyse

latente Faktoren

Die Faktorenanalyse als statistisches Verfahren werden sie genauer im Kurs über Testkonstruktion im Modul 6 kennen lernen. An dieser Stelle geht es nur um prinzipielle inhaltliche Ideen des Verfahrens. Die Reduktion einer großen Datenmenge, z.B. der Einschätzungen auf und Korrelationen zwischen 50 oder noch mehr Items auf wenige, z.B. 3 – 4 latente Faktoren ist mit mindestens 2 Vorteilen verbunden: (1) *Wenige Dimensionen können leichter weiterverarbeitet und kommuniziert werden.* Stellen Sie sich vor, Sie würden die Einschätzungen auf den 50 Items als einzelne Prädiktoren zur Vorhersage des Kriteriums Internetnutzung heranziehen. Ein solches Vorgehen macht wenig Sinn, insbesondere wenn viele der 50 Items redundant sind, d.h. jeweils hoch untereinander und mit dem Kriterium korrelieren. (2) Ein noch wichtigeres Argument für die Identifikation von Faktoren ist die damit verbundene Möglichkeit, Messungen zu aggregieren, d.h. zu einem Wert zusammenzufassen. Nach dem *Aggregationsprinzip* liefert die Summe mehrerer (miteinander korrelierender) Messungen eine stabilere und repräsentativere Schätzung eines Merkmals als eine einzelne Messung (Rushton, Brainerd & Pressley, 1983). Die höhere Messgenauigkeit mehrerer Messungen hängt damit zusammen, dass einzelne Messungen immer fehlerbehaftet sind. Wenn viele Messungen miteinander kombiniert werden, mitteln sich die Fehler dagegen aus. Eine *Analogie* aus dem nicht psychologischen Bereich macht dieses Prinzip vielleicht deutlicher: Wenn Sie die Länge eines Tisches möglichst genau messen wollen, dann ist es günstiger anstelle einer einzigen Messung 10 Messungen durchzuführen und einen Mittelwert zu bilden. Wenn Prädiktoren und auch Kriterien genau gemessen werden können, dann lassen sich auch Zusammenhänge zwischen den gemessenen Merkmalen besser identifizieren. Werden Merkmale sehr ungenau

Wenige Dimensionen  
können leichter weiter-  
verarbeitet und kommu-  
niziert werden.

Aggregationsprinzip

gemessen, dann kann es sein, dass sie nur sehr niedrig oder gar nicht interkorrelieren, obwohl ein Zusammenhang besteht, der bei höherer Messgenauigkeit aufgedeckt werden könnte. Allerdings kann die ausschließliche Orientierung an der Optimierung der Messgenauigkeit zu einer zu starken Homogenisierung eines Messinstruments führen; es wird dann ein nur sehr schmaler Merkmalsbereich erfasst und die Korrelationen mit anderen Variablen fallen dann auch wieder niedriger aus (sogenanntes *Reliabilitäts-Validitäts-Dilemma*, vgl. hierzu im Detail den Kurs Testkonstruktion im Modul 6).

Reliabilitäts-Validitäts-Dilemma

Wie „funktioniert“ eine Faktorenanalyse? Im Folgenden soll eine ungefähre eher inhaltliche und formelfreie Beschreibung versucht werden, die darauf abzielt, dass Sie bestimmte Begriffe verstehen können, die Sie in der Originalliteratur häufig antreffen werden.

Wie „funktioniert“ eine Faktorenanalyse?

In der Faktorenanalyse werden Variablen anhand ihrer korrelativen Beziehungen in voneinander unabhängige Gruppen klassifiziert. Bestimmte Indexzahlen, sogenannte *Ladungen*, geben an, wie gut eine Variable zu einer Variablengruppe bzw. einem Faktor passt. Vereinfacht ausgedrückt sind Ladungen die Korrelationen einer einzelnen Variable (eines Items) mit einem Faktor. Je höher eine Variable/ein Item auf einem Faktor lädt, umso besser passt sie zu bzw. repräsentiert sie diesen Faktor. Diejenigen Items, die am höchsten auf einem Faktor laden, werden *Markiervariablen* genannt und zur Interpretation des Faktors herangezogen, also um inhaltlich auszudrücken, was ein Faktor erfasst.

Ladungen

Markiervariablen

Was genau ist eigentlich ein Faktor? Bei einer großen Menge von Items lassen sich Partialkorrelationen berechnen. Diese geben an, inwieweit die Korrelation zweier Variablen durch eine dritte gestiftet wird, was immer dann der Fall ist, wenn nach dem Herauspartialisieren die Korrelation zwischen den beiden Variablen praktisch unbedeutend wird. Ausgehend von den Korrelationen zwischen den gemessenen Variablen wird eine synthetische, latente Variable konstruiert, die mit allen Variablen so hoch wie möglich korreliert. Diese synthetische, latente Variable ist der Faktor. Wird der Faktor aus den Variablen herauspartialisiert, ergeben sich weitere Partialkorrelationen, die diejenigen Variablenzusammenhänge erfassen, die nicht durch den Faktor erklärt werden können. Zur Klärung dieser Restkorrelationen werden weitere, in der Regel voneinander unabhängige (orthogonale) Faktoren extrahiert. Das Ergebnis einer Faktorenanalyse sind wechselseitig voneinander unabhängige Faktoren, die die Zusammenhänge zwischen den jeweiligen Variablen erklären. Die inhaltliche Interpretation dieser Faktoren ist offen!

Wichtig ist die *Unterscheidung zwischen explorativen und konfirmatorischen Faktorenanalysen*. Eine explorative Faktorenanalyse ist ein heuristisches Verfahren, das einem Variablengeflecht eine Ordnung unterlegt, mit der sich die Variableninterkorrelationen erklären lassen. Insbesondere geht es um die Ermittlung der Anzahl der Faktoren. Es gibt jedoch nicht nur eine einzige Ordnung, die die

Unterscheidung von explorativer und konfirmatorischer Faktorenanalyse

Merkmalszusammenhänge erklärt. Deshalb muss der Forscher dasjenige Ordnungssystem herauszufinden, das sich nicht nur statistisch, sondern auch theoretisch am besten begründen lässt. Ist innerhalb einer gegebenen Menge von Variablen/Items mittels exploratorischer Faktorenanalyse eine Faktorenstruktur ermittelt worden, so kann in weiteren Studien geprüft werden, ob sich diese Faktorenstruktur auch in anderen Stichproben replizieren und bestätigen lässt. In einer *konfirmatorischen Faktorenanalyse* wird also die Hypothese geprüft, dass eine bestimmte Faktorenstruktur auch in weiteren Stichproben mit den gegebenen Variablen/Items gilt bzw. zutrifft.

#### 7.4.4 Eingrenzung und Falsifikation von Kausalmodellen in der Korrelationsforschung

In der Einführung wurde bereits betont, dass korrelative Zusammenhänge nicht kausal interpretiert werden dürfen. Es besteht jedoch die Möglichkeit, durch bestimmte korrelative Designs (Längsschnittstudien und insbesondere cross-lagged panel designs, s.u.) und/oder inhaltliche Überlegungen die Anzahl kausaler Erklärungsalternativen einzuschränken bzw. zu falsifizieren (vgl. Bortz & Döring, 2006). Letztendlich aber hängt die Frage nach der kausalen Interpretierbarkeit korrelativer Zusammenhänge auch vom zugrundeliegenden Kausalitätskonzept ab (vgl. z.B. Hodapp, 1984, Kapitel 2). Im Folgenden wird im Abschnitt 4.1 die Frage der kausalen Interpretierbarkeit von Korrelationen genauer erörtert. Abschnitt 4.2 geht dann auf korrelative Längsschnittstudien ein und stellt ein Design vor, mit dem besonders gut deutlich gemacht werden kann, wie sich die Anzahl kausaler Erklärungsalternativen einschränken lässt. Im Abschnitt 4.3 schließlich wird auf zwei Verfahren eingegangen, die die Untersuchung komplexer Zusammenhangsstrukturen, je nach Autor wird auch von „Kausalstrukturen“ gesprochen, erlauben. Auch mit diesen Verfahren (Pfadanalyse und Strukturgleichungsmodelle) werden keine kausalen Zusammenhänge bestätigt. Es wird lediglich geprüft, ob und wie gut empirische Zusammenhänge zu einem theoretischen Modell passen. Genauer gesagt, geht es auch bei diesen komplexeren Modellen um Falsifikation, also darum, ob ein theoretisches Modell aufgrund der empirischen Zusammenhänge verworfen werden muss oder nicht.

##### 7.4.4.1 Korrelation und Kausalität

Wenn zwischen zwei Variablen X und Y keine positive oder negative Korrelation festgestellt werden kann, dann ist – vorausgesetzt die Nullkorrelation ist nicht messfehlerbedingt – die Hypothese eines Kausalzusammenhangs zwischen X und Y falsifiziert. Was aber bedeutet es, wenn zwei Variablen, z.B. übermäßiger Alkoholkonsum (x) und die Lebenserwartung (y) negativ korrelieren? Bortz und Döring (2006, S. 518) unterscheiden sechs mögliche Kausalmodelle, die für die Interpretation dieses Zusammenhangs infrage kommen:



Modell a: Übermäßiger Alkoholkonsum reduziert die Lebenserwartung (x beeinflusst y).

Modell b: Eine geringe Lebenserwartung verursacht erhöhten Alkoholkonsum (y beeinflusst x)

Modell c: Übermäßiger Alkoholkonsum und eine geringe Lebenserwartung beeinflussen sich wechselseitig<sup>5</sup>.

Modell d: Durch erhöhten Alkoholkonsum wird man arbeitsunfähig und damit arm. Armut (z) bedingt schlechte Ernährung, die das Leben verkürzt (x beeinflusst eine dritte Variable z, die ihrerseits y beeinflusst; z mediiert also den Zusammenhang zwischen y und y).

Modell e: Eine angeborene „Ich-Schwäche“ (z) erhöht die Anfälligkeit für lebensbedrohende Krankheiten und für Alkohol (x und y werden durch eine dritte Variable z beeinflusst).

Modell f: Stress (w) verursacht Trinken und Rauchen (z). Lebensverkürzend wirkt aber nur das Rauchen (eine vierte Variable w beeinflusst y über z indirekt und x direkt).

Bortz und Döring (2006) betonen selbst, dass diese Kausalmodelle unterschiedlich glaubwürdig und plausibel sind, was jedoch nicht aus der Korrelation selbst resultiert. In bestimmten Fällen lassen sich bestimmte Kausalinterpretationen von Korrelationen aber inhaltlich ausräumen. Wenn z.B. untersucht werden soll, ob die Jahreszeit, in der eine Person geboren wurde, mit dem späteren subjektiven Glücksempfinden zusammenhängt (vgl. Chotai & Wiseman, 2005), dann wird man logischerweise das kausale Erklärungsmodell „späteres Wohlbefinden beeinflusst früheren Geburtsmonat“ ausschließen. Man wird in diesem Fall aber auch keinen direkten Kausaleinfluss, sondern eher eine Mediation durch eine andere Variable vermuten (im zitierten Beispiel werden Persönlichkeitsmerkmale und Neurotransmitter als mögliche Mediatoren diskutiert).

#### 7.4.4.2 Längsschnitt- und Cross-lagged-panel designs

Ähnlich wie in dem soeben angeführten Beispiel können auch in Längsschnittstudien bestimmte Kausalhypothesen a priori ausgeschlossen werden. So liegen z.B. zahlreiche Studien vor, in denen der Zusammenhang zwischen der Abiturnote und

<sup>5</sup> Bei Bortz und Döring (2006, S. 518) ist Modell c anders formuliert: „Erhöhter Alkoholkonsum macht depressiv und verdunkelt damit die Lebensperspektive. Diese Lebensunlust lässt erneut zur Flasche greifen.“ Diese Formulierung ist aber eher für Modell d passend, da mit der Depression ein Mediator eingeführt wird!

dem Studienerfolg untersucht wurde. Da die Abiturnote bereits vor entsprechenden Studienleistungen vorliegt, können die Studienleistungen natürlich nicht die Abiturnote beeinflussen. Streng genommen ist in diesem und auch in dem davor angeführten Beispiel ein Längsschnittdesign nicht notwendig, da sich weder die Abiturnote noch die Jahreszeit der Geburt im Nachhinein noch ändern können. Die relevanten Variablen können also auch in einem Querschnittsdesign erhoben werden. Anders verhält es sich dagegen bei Variablen, die sich im Zeitverlauf verändern können, so etwa die Einstellung eines Lehrers gegenüber seinen Schülerinnen und Schülern. Wenn z.B. untersucht werden soll, ob diese Variable mit den Noten der Schüler zusammenhängt, dann muss die Lehrereinstellung am besten gleich zu Beginn des Schuljahres erfasst werden.

#### Cross-lagged panel design

Die Möglichkeit, unterschiedliche Kausalmodelle in korrelativen Längsschnittstudien auszuschließen, wurde von Campbell (1963) in einem eigenen Versuchsplan „verfeinert“, dem sogenannten *Cross-lagged panel design*. In einem Cross-lagged panel design werden mindestens 2 Variablen zu zwei verschiedenen Messzeitpunkten erhoben (vgl. Abb. 7-7). Es lassen sich dann zwei synchrone Korrelationen (A), zwei Autokorrelationen (B) und zwei zeitverzögerte Kreuzkorrelationen (C) berechnen. Die Hypothese, nach der X einen kausalen Einfluss auf Y ausübt, nicht aber umgekehrt, ist plausibler, wenn C1 höher ausfällt als C2. In manchen Fällen kann auch eine höhere synchrone Korrelation zum zweiten Messzeitpunkt (A2) eine der beiden konkurrierenden Kausalhypothesen stützen. So wird in einem Beispiel bei Bortz und Döring (2006, S. 519f) das Kausalmodell „Das Einkommen beeinflusst die Bildung“ gegenüber dem Alternativmodell „Die Bildung beeinflusst das Einkommen“ auch dadurch gestützt, dass A2 höher ausfällt als A1. Die Begründung dafür lautet, dass sich die Bildung zum Messzeitpunkt  $t_1$  bei Personen mit 25 Jahren noch nicht so stark auf das Einkommen auswirken konnte (A1) wie zum Zeitpunkt  $t_2$ , zu dem die Probanden 50 Jahre alt waren.

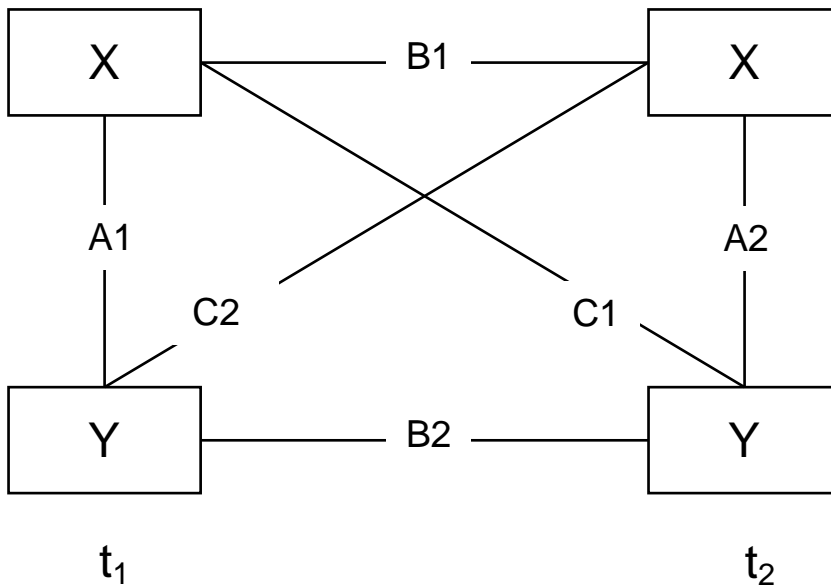


Abb. 7-7: Cross-lagged panel design

#### 7.4.4.3 Pfadanalysen und Strukturgleichungsmodelle

Die Pfadanalyse erweitert die multiple Korrelations- und Regressionsrechnung, indem Zusammenhänge zwischen mehreren unabhängigen und abhängigen Variablen gleichzeitig berücksichtigt werden. In Theorien bzw. theoretischen Modellen werden in der Regel Zusammenhänge zwischen mehr als zwei oder drei Variablen postuliert und es liegen auch Annahmen über Kausalbeziehungen zwischen den Variablen vor. Solche Zusammenhänge können in ein Pfadmodell „übersetzt“ und einer empirisch-statistischen Prüfung unterzogen werden. Sie haben im Abschnitt 3.2 bereits das Mediatorkonzept kennengelernt und in einem statistischen Exkurs gelesen, dass drei Regressionsgleichungen erforderlich sind, um zu prüfen, ob eine Variable den Zusammenhang zwischen zwei anderen Variablen mediiert. In der Pfadanalyse und auch in Strukturgleichungsmodellen können die mediierten Variablen ebenfalls als abhängige Variable aufgefasst werden, wenn *theoretisch* angenommen wird, dass eine unabhängige Variable die mediierte Variable „kausal“ beeinflusst. Abbildung 7-8 zeigt das Pfaddiagramm eines einfachen pfadanalytischen Modells.

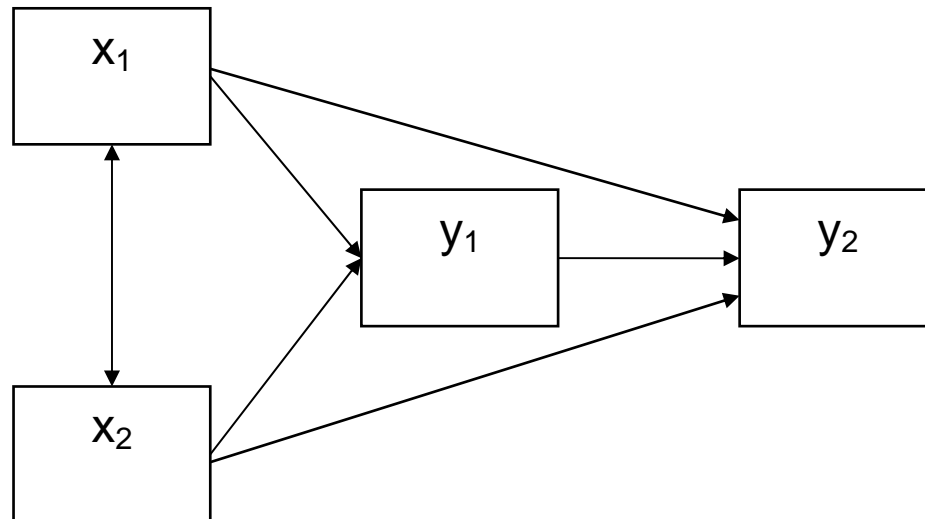


Abb. 7-8: Ein einfaches pfadanalytisches Modell

direkte Effekte

In einem Pfadmodell wird zwischen direkten und indirekten Effekten unterschieden. *Direkte Effekte* sind unvermittelte Zusammenhänge zwischen einer unabhängigen und einer abhängigen Variable. In Abbildung 7.2-6 symbolisieren z.B. die Pfeile von  $x_1$  zu  $y_1$  sowie  $x_1$  zu  $y_2$  direkte Effekte. Dagegen werden die Pfade von  $x_1$  bzw.  $x_2$  über  $y_1$  zu  $y_2$  als *indirekte Effekte* bezeichnet. Indirekte Effekte sind also Zusammenhänge zwischen zwei Variablen, die von einer (oder mehreren) anderen Variablen mediiert werden. Der Doppelpfeil zwischen  $x_1$  und  $x_2$  symbolisiert einen Zusammenhang zwischen diesen beiden Variablen, der vor dem Hintergrund einer Theorie als nicht-kausal angenommen wird.

indirekte Effekte

Wenn wir den oben von Hertel et al. (2005, in press) zitierten Mediationseffekt in dieses Pfadmodell übertragen, so wäre die emotionale Labilität  $x_1$ , die soziale Angst/Reduktion von Unsicherheit  $y_1$  und die Präferenz für asynchrone Medien  $y_2$ .  $x_2$  könnte eine weitere unabhängige Variable sein, z.B. Extraversion. Ein anderes inhaltliches Beispiel für ein Pfadmodell finden Sie bei Bortz und Döring (2006, S. 520).

*Strukturgleichungsmodelle:* Im Abschnitt 2.4 wurde ausgeführt, dass eine Theorie gemäß der Aussagenkonzeption theoretische Begriffe oder Konstrukte und Beobachtungsbegriffe enthält. Ängstlichkeit z.B. ist ein hypothetisches Konstrukt, das nicht direkt beobachtbar ist; es ist *latent*, d.h. es „verbirgt“ sich sozusagen „hinter“ den beobachtbaren Phänomenen. Beobachtbare, manifeste Phänomene bzw. Indikatoren lassen sich aber über Zuordnungsregeln mit den theoretischen Begriffen verknüpfen. Im Abschnitt 3.3 haben wir mit der Faktorenanalyse ein statistisches Verfahren in Grundzügen kennen gelernt, das die Ableitung von latenten Variablen, eben den sogenannten Faktoren, aus den korrelativen Beziehungen zwischen beobachteten, manifesten Variablen (z.B. Selbst- oder Fremdeinschätzungen in einem Fragebogen, beobachtete Häufigkeiten bestimmter Verhaltensweisen) gestattet. Wenn wir nun dieses faktorenanalytische Prinzip der Unterscheidung zwischen manifesten und latenten Variablen mit den regressions- und

pfadanalytischen Modellen verbinden, dann entsteht eine neue Klasse von Verfahren, die sogenannten linearen Strukturgleichungsmodelle. Der Begriff „Strukturgleichungsmodelle“ ist die generische Bezeichnung für eine Gruppe von Verfahren (vgl. z.B. Raykov & Marcoulides, 2006) zu der z.B. auch die konfirmatorische Faktorenanalyse gezählt wird. Eine Pfadanalyse basiert lediglich auf manifesten, d.h. beobachteten Variablen. Statistisch gesehen bedeutet diese Einschränkung, dass kein Modell zur Messung von latenten Variablen in die Pfadanalyse integriert ist. Ein Strukturgleichungsmodell umfasst dagegen ein Messmodell und ein Strukturmodell. Das *Messmodell* spezifiziert die Beziehungen zwischen den beobachteten Variablen und den hypothetischen Konstrukten. Das *Strukturmodell* beschreibt die Zusammenhänge zwischen den theoretischen Konstrukten. Im Prinzip ist ein Strukturgleichungsmodell also eine um ein Messmodell erweiterte Pfadanalyse. Das Messmodell hat den Vorteil, dass Fehler bei der Messung der beobachteten Variablen explizit berücksichtigt und die Zusammenhänge zwischen den latenten Variablen deshalb genauer geschätzt werden können. Abbildung 7-9 veranschaulicht die Unterscheidung zwischen dem Strukturmodell und drei Messmodellen in einem Strukturgleichungsmodell mit drei latenten Variablen. Die Abbildung wurde in Anlehnung an Nachtigall, Kroehne, Funke & Steyer (2003, p. 5) erstellt. Der gut lesbare englische Artikel von Nachtigall et al. zu den Vor- und Nachteilen von Strukturgleichungsmodellen ist online verfügbar unter [http://www.dgps.de/fachgruppen/methoden/mpr-online/issue20/art1/mpr127\\_11.pdf](http://www.dgps.de/fachgruppen/methoden/mpr-online/issue20/art1/mpr127_11.pdf).

Messmodell

Strukturmodell

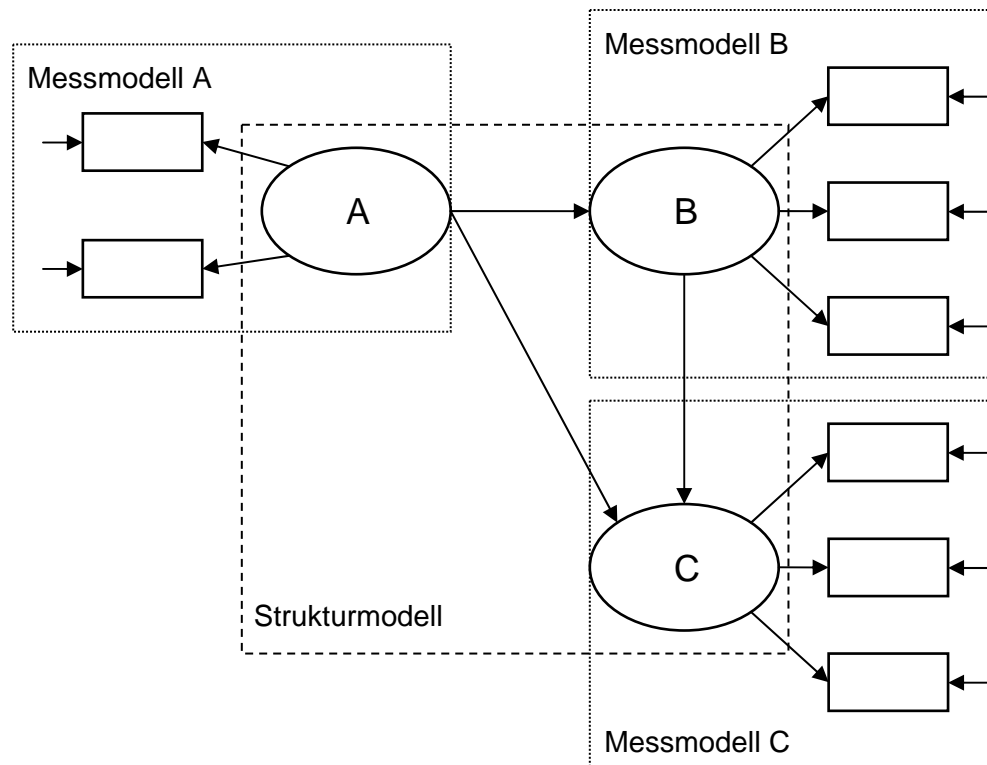


Abb. 7-9: Unterscheidung zwischen Strukturmodell und drei Messmodellen in einem Strukturgleichungsmodell mit drei latenten Variablen (in Anlehnung an Nachtigall et al., 2003, p. 5). A, B und C repräsentieren die latenten Variablen. Die leeren Rechtecke stehen für mögliche beobachtete Variablen, die „freien“ Pfeile symbolisieren die jeweiligen Messfehler zu den beobachteten Variablen. Die Pfeile innerhalb des Strukturmodells stehen für die theoretisch angenommenen unidirektionalen Pfade zwischen den latenten Variablen.

### Literaturempfehlungen

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.

Bortz, J. & Döring, N. (2006) *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Berlin: Springer.

Cohen, J., Cohen, P., West, S., & Aiken, L. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.

Nachtigall, C, Kröhne, U., Funke, F. & Steyer, R. (2003). (Why) Should We Use SEM? Pros and Cons of Structural Equation Modeling. *MPR-Online*, 8, 1-22.

## 7.5 Einzelfallstudien und Komparationsforschung – Einführung zur Vorlesung

In den bisher dargestellten Versuchsplänen war fast immer die Untersuchung von mehr oder weniger großen Gruppen von Personen impliziert. So zielt eine echte experimentelle Untersuchung darauf ab, Unterschiede zwischen der Experimental- und der Kontrollgruppe im Hinblick auf eine a priori aufgestellte Hypothese zu ermitteln – es geht aber nicht um Unterschiede zwischen einzelnen Personen der Experimental- und Kontrollgruppe! Angenommen wir würden eine experimentelle Untersuchung durchführen, in der geprüft werden soll, ob ein neues Trainingsprogramm (Intervention, treatment bzw. unabhängige Variable) die sozialen Kompetenzen (abhängige Variable) der Trainingsteilnehmer bedeutsam steigert. Nehmen wir weiter an, eine Experimentalgruppe würde dieses Trainingsprogramm durchführen und eine Kontrollgruppe würde ein Buch über soziale Kompetenzen lesen. Wir erheben ferner mit geeigneten Messinstrumenten (z.B. Fragebogen zur Selbst- und Fremdeinschätzung, Verhaltensbeobachtung) Daten zur sozialen Kompetenz vor und nach den Interventionen (Training vs. Buch), werten diese Daten mit einem statistischen Verfahren aus (in diesem Fall wäre eine zweifaktorielle Varianzanalyse mit Messwiederholungen geeignet) und stellen fest, dass sich die sozialen Kompetenzen in der EG signifikant mehr verbessert haben als in der KG. Was bedeutet dieses Ergebnis dann genau? Es bedeutet, dass sich die sozialen Kompetenzen im Durchschnitt über alle Probanden der Experimentalgruppe deutlicher verbessert haben als im Durchschnitt über alle Probanden der Kontrollgruppe. Es bedeutet aber *nicht* unbedingt, dass *jede einzelne* Person in der EG sozial kompetenter geworden ist als *jede einzelne* Person in der KG. Eine solche Konstellation ist zwar empirisch möglich, in vielen Interventionsstudien werden aber einzelne Personen innerhalb der EG und der KG mehr oder weniger stark bzw. gar nicht von einem treatment profitieren und es kann sogar vorkommen, dass ein Training bei einzelnen Probanden einen hypothesen*in*konformen Effekt hat. In unserem Beispiel würde die soziale Kompetenz einzelner Probanden nach dem Training dann niedriger ausfallen als vorher. Im Mittel können sich sehr positive, hypothesenkonforme Veränderungen bei bestimmten Probanden, gar keine Veränderungen bei anderen Probanden und Verschlechterungen bei wieder anderen Probanden dann aber soweit ausgleichen, dass die durchschnittliche, auf die gesamte Experimentalgruppe bezogene Veränderung bedeutsamer ausfällt als in der Kontrollgruppe, in der natürlich ähnliche Konstellationen bei einzelnen Probanden auftreten können wie sie für die EG beschrieben wurden.