

# Statistische Modelle und Auswertungsverfahren

## Statistical Models and Evaluation Procedures

Jonas W. B. Lang & Martin Kersting

Die Arbeits- und Organisationspsychologie (AO-Psychologie) verfolgt in der Forschung und Praxis vorwiegend quantitative Forschungsstrategien. Theorien werden in messbare Variablen operationalisiert und die Zusammenhänge zwischen diesen Variablen werden analysiert. Dazu sind statistische Modelle und Auswertungsverfahren notwendig. Ziel des vorliegenden Beitrags ist es, einen Überblick über die für die AO-Psychologie gängigen Auswertungsverfahren zu geben.

### 1 Verfahren zur Analyse von Variablenzusammenhängen

#### 1.1 Effektstärke, p-Wert, Teststärke und Stichprobengröße

Die Stärke des Zusammenhangs zwischen Variablen lässt sich auf zwei unterschiedliche Weisen beschreiben: durch Effektstärken und durch p-Werte. Effektstärken beschreiben die Größe eines Zusammenhangs zwischen Variablen unabhängig von der Größe der Stichprobe, die der Berechnung zu Grunde liegt. Sie werden deswegen vor allem eingesetzt, um das Ausmaß und die Bedeutung von Zusammenhängen einzuschätzen und die Ergebnisse von Studien miteinander zu vergleichen.

P-Werte werden ungeachtet der Kritik (Kline, 2004) in der AO-Psychologie häufig verwendet. P-Werte geben an, wie wahrscheinlich das Zustandekommen des gegebenen Datensatzes ist, wenn die Null-Hypothese zutrifft. Die Null-Hypothese besagt, dass kein Zusammenhang zwischen den Variablen existiert. Im Gegensatz zu Effektstärken sind p-Werte abhängig von der Stichprobengröße. Ein kleiner und unbedeutender Zusammenhang kann daher in einer großen Stichprobe signifikant sein und umgekehrt. P-Werte dienen vor allem dem Testen von Hypothesen. Dabei wird vorab ein Wahrscheinlichkeitswert gewählt, an dem eine Hypothese angenommen oder zurückgewiesen wird. Dieser Wert wird als Signifikanzniveau oder Irrtumswahrscheinlichkeit  $\alpha$  bezeichnet und ist die Wahrscheinlichkeit, fälschlicherweise die Null-Hypothese zurückzuweisen (Fehler 1. Art). In der Regel wird in psychologischen Untersuchungen ein Signifikanzniveau von 5 % gewählt. Bei kleinen Stichproben oder wenn es gilt, den Fehler zweiter Art (fälschliche Annahme der Null-Hypothese) zu vermeiden, kann es jedoch sinnvoll sein, ein niedrigeres Signifikanzniveau, z. B. 10 % zu wählen („progressives Testen“). Wenn

viele Hypothesen überprüft werden oder die negativen Folgen des Fehlers erster Art sehr gravierend sind, kann es dagegen notwendig sein, ein höheres Niveau wie etwa 1 % zu wählen, um das Risiko, eine falsche Hypothese anzunehmen, zu begrenzen.

Wird ein hohes Signifikanzniveau bei einer kleinen Stichprobe verwendet, wird die Teststärke zu gering (Cohen, 1988). Die Teststärke ist die Wahrscheinlichkeit, eine Theorie auf Grund eines Datensatzes zu bestätigen, falls diese zutreffend ist. Die Teststärke ergibt sich direkt aus der Irrtumswahrscheinlichkeit  $\beta$  (Fehler 2. Art). Die Irrtumswahrscheinlichkeit  $\beta$  ist die Wahrscheinlichkeit, fälschlicherweise die Null-Hypothese nicht zu verwerfen, obwohl sie falsch ist. Die Teststärke entspricht  $1 - \beta$ . Wenn  $\beta$  niedrig ist, ist also die Teststärke hoch. Bereits bei der Planung von Untersuchungen muss man daher auf Grund der Theorie eine konkrete Vorstellung entwickeln, wie groß die zu erwartenden Effekte ausfallen. Der Umfang der Stichprobe, an der die Hypothesen überprüft werden, sollte auf Grund dieser Effektgrößen vorab bestimmt werden. Die Bedeutung der Teststärke wird in der Praxis vernachlässigt, wenn kleine Stichproben mit hohen Signifikanzniveaus kombiniert werden. Auf diese Weise werden viele gute psychologische Theorien zu Unrecht verworfen (Maxwell, 2004).

#### 1.2 Häufig angewandte Verfahren

Die in der AO-Psychologie gebräuchlichsten statistischen Modelle und Verfahren zur Analyse von Variablenzusammenhängen lassen sich danach ordnen, welche Art von Variablenbeziehungen sie aufdecken können. Dazu ist es notwendig, verschiedene Arten von Variablen zu unterscheiden:

1. **Kontinuierliche Variablen lassen sich von Dummy-Variablen unterscheiden.** Kontinuierliche Variablen zeichnen sich dadurch aus, dass sie intervallskaliert sind, d. h. sie können theoretisch jeden beliebigen Wert annehmen. Dummy-Variablen sind Variablen, die entweder den Wert 0 oder den Wert 1 annehmen können. Eine oder mehrere Dummy-Variablen werden zusammenfassend auch als kategoriale Variable oder im Kontext der experimentellen Psychologie als Faktor bezeichnet.
2. **Unterschieden werden außerdem abhängige Variablen, unabhängige Variablen und Kontrollvariablen.** Unabhängige Variablen werden auch als Prädiktoren und abhängige Variablen auch als Kriterien bezeichnet. In psychologischen Fragestellungen wird die Auswirkung von einer oder mehreren unabhängigen Variablen auf eine oder mehrere abhängige Variablen untersucht. Kontrollvariablen sollen Fehlervarianz in der Kriterienvariable erklären und auf diese Weise die Teststärke erhöhen. Auf der Modellebene entsprechen sie somit unabhängigen Variablen.

Tabelle 1 greift die beschriebenen Variablenklassifikationen auf und bietet einen daran orientierten Überblick über die in der AO-Psychologie häufig eingesetzten Verfahren. Die weiter dargestellten Verfahren setzen alle eine annähernde Gaußsche Normalverteilung der kontinuierlichen Variablen voraus (voraussetzungsärmere Verfahren werden im vorliegenden Beitrag aus Platzgründen nicht besprochen, siehe dazu z. B. Bortz, Lienert & Boehnke, 2000).

Die Tabelle informiert darüber, mit welchen Verfahren welche Art von Variablenzusammenhängen interpretiert werden kann. Außerdem werden die mit den Ver-

**Tabelle 1:** Häufig verwendete Verfahren zur Analyse von Variablenzusammenhängen

Verfahren	Unabhängige Variablen (UVs)/ Kontrollvariablen	Abhängige Variablen	Effektstärke-Maße	Effektstärke-Interpretation <sup>1</sup>		
				klein	mittel	groß
1. t-test für unabhängige Stichproben	Eine Dummy-Variablen	Eine kontinuierliche Variable	$d$	.20	.50	.80
2. Produkt-Moment-Korrelation	Eine kontinuierliche Variable oder eine Dummy-Variablen	Eine kontinuierliche Variable	$r$	.10	.30	.50
3. Multiple Regression	Mehrere Dummy-Variablen oder mehrere kontinuierliche Variablen	Eine kontinuierliche Variable	$R^2$ (Gesamtmodell)	.02	.13	.26
			$\beta$ (einzelne UVs)	.10	.30	.50
4. Logistische Regression	Mehrere Dummy-Variablen oder mehrere kontinuierliche Variablen	Eine Dummy-Variablen	Pseudo- $R^2$ (z. B. nach Nagelkerke für das Gesamtmodell)	.02	.13	.26
5. Varianz- und Kovarianzanalyse	Mehrere Dummy-Variablen (Varianzanalyse) oder mehrere Dummy-Variablen und mehrere kontinuierliche Variablen (Kovarianzanalyse)	Eine kontinuierliche Variable	$\eta^2$ (entspricht $R^2$ )	.02	.13	.26
6. Random Coefficient Models (auch hierarchisch lineare Modelle oder gemischte Effektmodelle)	Mehrere Dummy-Variablen oder mehrere kontinuierliche Variablen auf unterschiedlichen Organisationsebenen	Eine kontinuierliche Variable	Pseudo- $R^2$ (Gesamtmodell)	.02	.13	.26
			$\beta$ (einzelne UVs)	.10	.30	.50

Anmerkung: <sup>1</sup> nach Cohen (1988)

fahren erzielten Effektstärke-Maße klassifiziert. Die Informationen der Tabelle können als Leitfaden für die Planung und Durchführung von Untersuchungen genutzt werden. Darüber hinaus sind die Informationen als Maßstab zur Beurteilung von empirischen Studien hilfreich.

### 1.3 Kausalität

Bei der Interpretation von Variablenzusammenhängen ist es wichtig, zwischen randomisiert zugewiesenen Variablen und attributiven Variablen zu unterscheiden. Randomisiert zugewiesene Variablen sind Variablen, deren Werte im Rahmen einer experimentellen Prozedur durch einen Zufallsprozess der Untersuchungseinheit zugewiesen werden. Durch eine Randomisierung ist es möglich, andere Einflussfaktoren zu kontrollieren. Dadurch können randomisiert zugewiesene Variablen kausal als Verursacher interpretiert werden. Randomisierte Variablen sind der Gruppe der unabhängigen Variablen zuzuordnen. In der Regel sind randomisiert zugewiesene Variablen Dummy-Variablen. Forschung mit randomisierten Variablen wird als experimentelle Forschung bezeichnet. Attributive Variablen sind Variablen, die unveränderliche Merkmale von Untersuchungseinheiten beschreiben (z. B. Geschlecht, Alter). Attributive Variablen werden in der nichtexperimentellen oder korrelativen Forschung eingesetzt und bergen immer die Gefahr eines kausalen Fehlschlusses. Als angewandte Disziplin arbeitet die AO-Psychologie vorwiegend mit solchen attributiven Variablen. Eine typische Form des kausalen Fehlschlusses tritt auf, wenn ein Zusammenhang zwischen Variablen existiert, der nicht durch einen tatsächlichen Wirkmechanismus zwischen den Variablen verursacht ist, sondern durch andere unbekannte Variablen zu Stande kommt, die auf beide Variablen wirken. Beispielsweise lässt ein hoher negativer Zusammenhang zwischen dem Grad der körperlichen Arbeit in einem Arbeitsverhältnis und dem erreichten Lebensalter nicht unbedingt darauf schließen, dass körperliche Arbeit zu einem frühen Tod führt. Stattdessen kann der Zusammenhang auch dadurch zu Stande kommen, dass Menschen, die körperliche Arbeit leisten, dazu neigen, andere Gesundheitsrisiken (z. B. hoher Alkoholkonsum) einzugehen, und dadurch kürzer leben. Eine andere Form eines kausalen Fehlschlusses tritt auf, wenn Variablenzusammenhänge in die falsche Richtung interpretiert werden (d. h. nicht A wirkt auf B, sondern B auf A). Eine Möglichkeit, diese Art von Fehlschlüssen aufzudecken ist es, zwei Variablen mehr als einmal zu erheben und dann über Kreuz ihren Einfluss aufeinander zu bestimmen. Ist ein Zusammenhang höher als der andere, so lässt dies Rückschlüsse über die Wirkrichtung zu. Diese Art des Designs wird als „Cross-Lagged Panel“ bezeichnet. Aussagen auf der Basis eines Cross-Lagged Panel sind aber nur möglich, wenn beide Variablen weitgehend die gleiche Stabilität aufweisen. Ein populäres Beispiel für die Notwendigkeit von Cross-Lagged Panel-Analysen zur Vermeidung von kausalen Fehlschlüssen ist der vielfach bestätigte Zusammenhang zwischen Arbeitszufriedenheit und Unternehmenserfolg. Im Gegensatz zu der häufig geäußerten Vermutung, dass Arbeitszufriedenheit

zu Unternehmenserfolg führt, legt eine Cross-Lagged Panel-Studie nahe, dass Unternehmenserfolg stärker die Arbeitszufriedenheit beeinflusst, als das umgekehrt der Fall ist (Schneider, Hanges, Smith & Salvaggio, 2003).

#### 1.4 Moderatoren

AO-psychologische Theorien nehmen häufig an, dass die Stärke der Beziehung zwischen zwei Variablen von einer dritten Variable (Moderatorvariable) abhängt. Statistisch kann eine solche Hypothese überprüft werden, indem neben (a) dem additiven Effekt der interessierenden unabhängigen Variable auch (b) der additive Effekt der Moderatorvariable und zusätzlich noch (c) eine Interaktionsvariable der beiden Variablen als unabhängige Variablen in die Analyse eingeschlossen wird. Die Interaktionsvariable wird gebildet, indem die beiden anderen Variablen miteinander multipliziert werden. Wenn der Einfluss der Interaktionsvariable auf die abhängige Variable bedeutsam ist, dann bestätigt dies den Moderatoreffekt. Bei derartigen Analysen ist zu beachten, dass die Effekte der additiven unabhängigen Variablen vom Nullwert der jeweils anderen unabhängigen Variable abhängen. Wenn die Variablen kontinuierlich sind und ihr Nullwert unbedeutend ist (z. B. diese Ausprägung in der Praxis nicht vorkommt), können die aufgezeigten Effekte nicht sinnvoll inhaltlich interpretiert werden. Es ist daher in der Regel empfehlenswert, bei Interaktionsanalysen kontinuierliche unabhängige Variablen zu „zentrieren“.

##### Merke:

Beim Zentrieren wird bei allen unabhängigen Variablen jeweils der Mittelwert der Variablen von allen Werten abgezogen, so dass die additiven Effekte der unabhängigen Variablen dem mittleren Effekt der unabhängigen Variablen auf die abhängige Variable entsprechen (Cohen, Cohen, West & Aiken, 2003).

#### 1.5 Mediatoren

Mediatorvariablen sind für den Zusammenhang zwischen zwei Variablen verantwortlich. Bei einer vollen Mediation wirkt Variable A auf Variable B und Variable B auf Variable C. Variable B ist also der Mediator zwischen Variable A und Variable C und somit für deren Zusammenhang verantwortlich. Mediationshypothesen können entweder durch eine Serie von multiplen Regressionsanalysen (vgl. Tab. 1) oder durch die weiter unten erwähnten Strukturgleichungsmodelle überprüft werden.

#### 1.6 Mehrebenenperspektive

Häufig sind nicht nur einzelne Individuen, sondern Unternehmen, Organisationen oder Gruppen Gegenstand der AO-Psychologie. Solche Daten erfordern in der Regel eine Mehrebenenperspektive und entsprechende Methoden (Klein & Kozlowski,

2000), weil Unternehmen, Organisationen und Gruppen zumeist eine hierarchische Struktur aufweisen. Bei einer hierarchischen Struktur sind Untereinheiten in darüberliegende Einheiten eingebettet: Zum Beispiel gehören Individuen zu Gruppen, und Gruppen gehören zu Organisationen. Eine Folge von hierarchischen Strukturen ist häufig eine hierarchische Abhängigkeit in den Daten. Das bedeutet, dass Untereinheiten, die zu einer gemeinsamen Einheit gehören, sich ähnlicher sind als Untereinheiten, die zu unterschiedlichen Einheiten gehören. Beispielsweise haben zwei Arbeitnehmer einer Arbeitsgruppe in der Regel eine ähnlichere Arbeitszufriedenheit als zwei Arbeitnehmer aus unterschiedlichen Arbeitsgruppen.

##### Merke:

Statistisch bedeutet hierarchische Abhängigkeit, dass die Gruppenzugehörigkeit zu einer höheren Einheit einen Teil der Varianz in einer Variablen auf der darunter angeordneten Ebene erklärt. Wie hoch der Varianzanteil auf Grund der Gruppenzugehörigkeit ist, kann durch die Berechnung der Intraklassenkorrelation vom Typ 1 (ICC1) bestimmt werden. Die ICC1 variiert zwischen 0 und 1. In organisationspsychologischen Studien werden typischerweise ICCs um .10 beobachtet. Dies bedeutet, dass 10% der Varianz durch Kontextvariablen wie beispielsweise Gruppenklima erklärt werden (Bliese, 2000).

Hierarchische Strukturen und hierarchische Abhängigkeiten erfordern, dass AO-Psychologen die Ebenen spezifizieren, auf denen ein hypothetisch angenommener Prozess abläuft. Dadurch wird die zu analysierende Variable festgelegt. Eine Variable auf einer höheren Ebene kann ein vollkommen anderes Konstrukt indizieren als die gleiche Variable auf der Individualebene. Wenn die interessierende Variable nicht auf der konzeptionellen Ebene gemessen wird, die sie beschreibt, sondern auf einer niedrigeren Ebene, so müssen die Einzelurteile aggregiert werden. Ein Beispiel: Der Gruppenerfolg wird durch alle einzelnen Mitglieder der Gruppe bewertet, muss aber aggregiert werden, um als Indikator für den Erfolg der Gruppe gelten zu können.

##### Merke:

Statistische Analysen auf der Basis von aggregierten Variablen sind nur dann sinnvoll, wenn die Urteile der Gruppe hinreichend übereinstimmen und hinreichend messgenau sind. Um dies zu prüfen, wird dabei neben der Intraklassenkorrelation vom Typ 1 (ICC1) als Indikator für die hierarchische Abhängigkeit der Daten auch die Intraklassenkorrelation vom Typ 2 (ICC2) herangezogen, die angibt, wie reliabel der Mittelwert des Gruppenurteils ist. Die Übereinstimmung der Urteile wird in der Regel auf der Basis von Gruppenübereinstimmungskoeffizienten analysiert, von denen  $r_{wg(j)}$  der am häufigsten verwendete ist. Sowohl ICC2 wie auch  $r_{wg(j)}$  sollten über .70 sein, um Analysen auf der Basis einer aggregierten Variable durchführen zu können (Bliese, 2000).

Wenn Mehrebenenstrukturen vorliegen, macht das unter gewissen Bedingungen den Einsatz von „Random Coefficient“-Modellen (RCM, auch als hierarchical linear models, HLM, bekannt) notwendig, um fehlerhafte Schlussfolgerungen zu vermeiden. (Eine andere Möglichkeit sind Erweiterungen der weiter unten dargestellten Strukturgleichungsmodelle). RCM-Modelle kontrollieren dabei die hierarchische Abhängigkeit in den Variablen. Sie müssen vor allem dann eingesetzt werden, wenn der Zusammenhang zwischen Variablen auf unterschiedlichen Ebenen untersucht werden soll. Werden solche Fragestellungen mit klassischen Verfahren untersucht, treten erhebliche Fehler 1. Art und damit falsche Schlussfolgerungen auf. Wird dagegen der Zusammenhang zwischen Variablen untersucht, die alle auf der obersten Organisationsebene liegen, so ist der Einsatz von RCM nicht notwendig. Stattdessen können klassische Verfahren wie die multiple Regression verwendet werden. Soll der Zusammenhang zwischen Variablen untersucht werden, die alle auf einer unteren Ebene liegen, so ist es sinnvoll, Abhängigkeiten von darüberliegenden Organisationsebenen zu kontrollieren. Dadurch erhöht sich die Teststärke (Minimierung des Fehlers 2. Art) der Untersuchung, da die Kontrolle der hierarchischen Abhängigkeiten Fehlervarianz im Kriterium unterdrückt.

## 2 Faktorenanalyse

Wenn ein AO-Psychologe eine große Menge von Variablen analysiert, die eng miteinander in Zusammenhang stehen, ist die Faktorenanalyse ein nützliches Hilfsmittel zur Datenreduktion. Die Faktorenanalyse erlaubt es, die vielen, einander ähnlichen Variablen durch wenige, einander unähnlichere Variablen (Faktoren) zu beschreiben. Sie analysiert also Variablenstrukturen. Faktoren werden auch als latente Dimensionen bezeichnet. Dahinter steht der Gedanke, dass beobachteten Variablen eine gemeinsame, nicht direkt beobachtbare Variable zu Grunde liegt. Die Anwendung der Faktorenanalyse ist aus zwei Gesichtspunkten heraus sinnvoll. Erstens macht die Reduktion von Variablen nachfolgende statistische Analysen mit diesen Variablen leichter interpretierbar. Zweitens sind bedeutungsvolle Aussagen über die „Natur“ der untersuchten Variablen möglich, die auch für sich genommen bereits interessante Erkenntnisse ermöglichen. Man unterscheidet die konfirmatorische von der exploratorischen Faktorenanalyse (CFA und EFA). Beide Verfahren gehen von einer Korrelationsmatrix der interessierenden Variablen aus.

Bei der CFA wird vor der Datenanalyse auf theoretischer Basis eine Faktorenstruktur postuliert. Diese ordnet die einzelnen Indikatoren latent ihnen zu Grunde liegenden Faktoren zu. Darüber hinaus wird vorab spezifiziert, ob die Faktoren korreliert oder unabhängig voneinander sind. Auf der Basis dieser theoretischen Setzungen wird versucht, eine Matrix zu bestimmen, die der beobachteten Matrix möglichst genau entspricht. Je genauer dies gelingt, desto höher ist der „Fit“ des

Modells auf die Daten. Als Indikatoren für den Fit des Modells werden in der Regel Indizes herangezogen, die möglichst stichprobenunabhängig und präzise beschreiben sollen, wie gut das angenommene Modell auf die Daten passt.

### Wichtige Indizes

Die wichtigsten Fit-Indizes sind der „comparative fit index“ (CFI), der „root-mean-square error of approximation“ (RMSEA) und das „standardized root-mean-square residual“ (SRMR). Diese drei Indizes geben unterschiedliche Aspekte des Modellfits wieder. Als gut gelten in der Regel Modelle mit  $SRMR \leq .08$ ,  $RMSEA \leq .06$  und  $CFI \geq .90$  (manche Autoren empfehlen auch  $CFI \geq .95$ ). Eine Diskussion der Vor- und Nachteile verschiedener Fit-Indizes findet sich bei Beauducel und Witmann (2005).

Bei der EFA wird im Gegensatz zur CFA vor der Datenanalyse keine Faktorenstruktur spezifiziert. Stattdessen wird anhand von empirischen Kriterien (z. B. Guttman-Kaiser-Kriterium, Scree-Test, Parallelanalyse) entschieden, wie viele Faktoren einer Matrix zu Grunde liegen. Im nächsten Schritt wird dann versucht, eine Faktorenlösung zu generieren, bei der die einzelnen Indikatoren möglichst genau einzelnen Faktoren zugeordnet werden können. Dies geschieht durch Faktorenrotationsverfahren. Es gibt zwei unterschiedliche Typen von Faktorenrotationsverfahren: Schiefwinklige (oblique) und orthogonale Verfahren. Bei den obliquen Verfahren korrelieren die erzeugten Faktoren miteinander, während die durch die orthogonalen Verfahren extrahierten Faktoren unkorreliert sind.

## 3 Strukturgleichungsmodelle

Strukturgleichungsmodelle sind eine Verallgemeinerung der CFA. Wie bei der CFA werden zunächst Variablen ihnen zu Grunde liegenden latenten Variablen oder Faktoren zugeordnet. Diese Zuordnung wird auch als Messmodell bezeichnet, da es beschreibt, wie die einzelnen Konstrukte empirisch operationalisiert werden. Das Messmodell entspricht einem CFA-Modell. Im nächsten Schritt wird dann spezifiziert, wie die einzelnen Variablen aufeinander wirken. Diese Verknüpfungen zwischen den einzelnen Variablen werden zusammenfassend als Strukturmodell bezeichnet. Ein großer Vorteil von Strukturgleichungsmodellen ist die Möglichkeit, auch komplexe Fragestellungen mit mehreren abhängigen und unabhängigen Variablen in einem Modell zu repräsentieren und gemeinsam zu testen. Strukturgleichungsmodelle stellen somit eine Integration der Verfahren zur Analyse von Variablenzusammenhängen und der Faktorenanalyse dar. Ein Strukturgleichungsmodell von Colquitt (2001) ist in Abbildung 1 als Beispiel dargestellt.

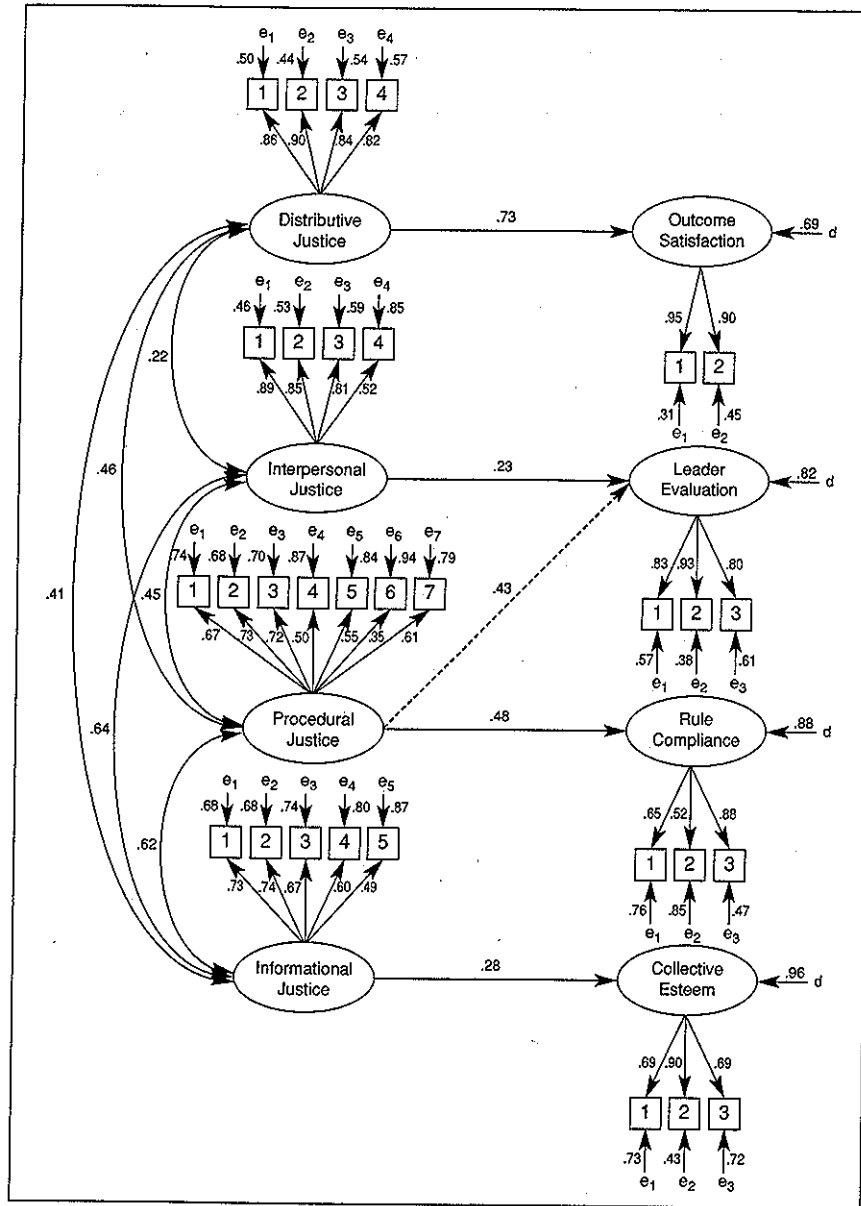


Abbildung 1: Strukturgleichungsmodell aus Colquitt (2001) zum Zusammenhang zwischen vier unterschiedlichen Dimensionen der organisationalen Gerechtigkeit mit verschiedenen organisationalen Kriterien. Der Fit des Modells war zufrieden stellend: CFI = .90, RMSEA = .06.

## 4 Metaanalyse

Forschungszusammenfassungen und Überblicke sind für Forscher und Praktiker ein wichtiges Instrument, um psychologische Theorien voranzubringen. Gerade dann, wenn empirische Untersuchungen zu unterschiedlichen Aussagen darüber kommen, wie stark Variablen zusammenhängen, sind metaanalytische Verfahren hilfreich. Bei einer Metaanalyse werden (a) möglichst viele empirische Studien analysiert, in denen der Zusammenhang untersucht wurde, (b) die darin berichteten Effektstärken oder p-Werte zusammengetragen und in eine gemeinsame Metrik gebracht, (c) möglicherweise korrigiert (vor allem gegen Unreliabilität oder hierarchische Abhängigkeit, wenn mehrere Effektstärken aus einer Studie vorliegen) und (d) gemittelt. Die daraus resultierende mittlere Effektstärke lässt dann eine fundierte und zusammenfassende Aussage über die Stärke des Effektes zu. Darüber hinaus können Moderatorvariablen identifiziert werden. Metaanalysen haben sich als nützlich erwiesen. Mittlerweile liegen für eine große Anzahl von Fragestellungen in der AO-Psychologie entsprechende Studien vor. Trotz der Vorteile von Metaanalysen können natürlich auch diese Analysen zu Fehlschlüssen führen und wurden in der Vergangenheit kontrovers diskutiert (siehe Schulze, Holling & Böhning, 2003).

## 5 Ausblick

Im vorliegenden Kapitel wurde versucht, die für die aktuellen Fragen der AO-Psychologie wichtigsten statistischen Auswertungsverfahren kurz vorzustellen. Auf eine detaillierte Darstellung der einzelnen Verfahren und die Darstellung vieler seltener verwendeter Verfahren musste auf Grund des begrenzten Umfangs des Beitrags verzichtet werden. An dieser Stelle sei dazu auf die einschlägige weiterführende Literatur verwiesen (z. B. Holling & Schmitz, in Druck). Die Entwicklung von statistischen Modellen und Auswertungsverfahren erfolgt häufig als Reaktion auf drängende theoretische und praktische Fragen. Viele der vorgestellten Verfahren wurden erst in den letzten zwei Jahrzehnten populär. Mit ihrer Hilfe ist es der AO-Psychologie und benachbarten Disziplinen gelungen, aufgeworfene Fragen zu beantworten. Die Mehrebenenanalyse und die Metaanalyse sind zwei prominente Beispiele. Da sich die AO-psychologischen Theorien ständig weiterentwickeln, wird auch die Zukunft die Optimierung vorhandener und Entwicklung neuer statistischer Prozeduren erfordern, um neue Theorien zu prüfen.

## Weiterführende Literatur

Beauducel, A. & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA-based on data with slightly distorted simple structure. *Structural Equation Modeling, 12*, 41-77.

- Bortz, J., Lienert, G. & Boehnke, K. (2000). *Verteilungsfreie Methoden in der Biostatistik*. Berlin: Springer.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3<sup>rd</sup> ed.). Mahwah, NJ: Erlbaum.
- Holling, H. & Schmitz, B. (in Druck). *Handbuch der psychologischen Methoden und Evaluation*. Göttingen: Hogrefe.
- Klein, K. J. & Kozlowski, S. W. J. (Eds.). (2000). *Multilevel theory, research, and methods in organizations*. San Francisco, CA: Jossey-Bass.
- Kline, R. B. (2004). *Beyond significance testing*. Washington, DC: APA.
- Schulze, R., Holling, H. & Böhning, D. (Eds.). (2003). *Meta-analysis: New developments and applications in medical and social sciences*. Göttingen: Hogrefe.

### Literatur

- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349–381). San Francisco, CA: Jossey-Bass.
- Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, 3, 386–400.
- Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychological Methods*, 9, 147–163.
- Schneider, B., Hanges, P. J., Smith, D. B. & Salvaggio, A. N. (2003). Which comes first: Employee attitudes or organizational, financial and market performance? *Journal of Applied Psychology*, 88, 836–851.

## HANDBUCH DER PSYCHOLOGIE

# Handbuch der Arbeits- und Organisations- psychologie

herausgegeben von

Heinz Schuler und Karlheinz Sonntag

(2007)

HOGREFE



GÖTTINGEN · BERN · WIEN · PARIS · OXFORD · PRAG  
TORONTO · CAMBRIDGE, MA · AMSTERDAM · KÖPENHAGEN