



# Faktoren- und Latent-Class-Analysen zur Auswertung von Teilnehmerbefragungen: Überblick am Beispiel des Online-Kurses COER13

Stand: 28.04.2014  
 Portalbereich: Aus der Praxis  
 Autoren/innen: Jens Jirschwitzka

## Inhaltsverzeichnis

1. Einleitung .....	1
2. Faktorenanalyse und Latent-Class-Analyse .....	1
3. Ein Anwendungsbeispiel .....	2
4. Fazit .....	6
Literaturverzeichnis .....	6

## 1. Einleitung

Der vorliegende Bericht widmet sich weiterführenden Auswertungsstrategien von Befragungsdaten, die im Rahmen des Online-Kurses COER13 (*Online Course zu Open Educational Resources 2013*, <http://www.coer13.de>) erhoben wurden. Dieser für alle Interessierten frei zugängliche Kurs zum Thema „Offenen Bildungsressourcen“ fand vom 8. April bis zum 1. Juli 2013 statt. Die Veranstaltung wurde von e-teaching.org (Simone Haug, Markus Schmidt und Anne Thillosen) in Kooperation mit folgenden Partnern durchgeführt: Patricia Arnold (Hochschule München), Martin Ebner (TU Graz, L3T), Andreas Link (Learning Agency Network), Johannes Moskaliuk (Universität Tübingen) und Sandra Schön (BIMS e.V., L3T).

Vor dem Hintergrund der während des COER13 durchgeführten Teilnehmerbefragungen werden die *Latent-Class-Analyse* (LCA) und spezielle Formen der *Faktorenanalyse* (FA) kurz erörtert und erstere mithilfe von COER13-Befragungsdaten illustriert. Zu den skizzierten Verfahren und den zugrundeliegenden testtheoretischen Modellen werden an entsprechenden Stellen Literaturhinweise gegeben.

## 2. Faktorenanalyse und Latent-Class-Analyse

Über die Betrachtung von Einzelvariablen, bivariaten Zusammenhängen oder Unterschieden zwischen manifesten Gruppen hinaus lassen sich auch bei Befragungen fast immer auch komplexere Zusammenhangsstrukturen analysieren. Dies kann sowohl unter eher variablenzentrierter Perspektive als auch unter eher personenzentrierter Perspektive geschehen. Erstere ist eher mit faktorenanalytischen Verfahren verbunden, während eine personenzentrierte Perspektive eher mit clusteranalytischen Verfahrensweisen verbunden ist (vgl. z. B. Wang & Wang, 2012).

Faktorenanalytische Verfahren zielen darauf ab, die korrelativen Zusammenhänge zwischen den manifesten (den beobachteten und messfehlerbehafteten) Indikatorvariablen durch eine diesen Variablen gemeinsam zugrundeliegende latente und als kontinuierlich angenommene Faktorvariable zu erklären. Ein Beispiel: Die Einstellung von Personen zu einem bestimmten Thema wäre eine solche „latente Variable“. Diese kann nicht direkt, sondern nur mithilfe von manifesten Variablen gemessen werden – etwa mehreren Fragen (Items) zu bestimmten Aspekten dieses Themas bzw. Einstellungsgegenstandes. An-



gestrebt wird dabei zumeist, dass die jeweiligen manifesten Variablen nur jeweils eine latente Variable „messen“ und dass sie dies möglichst zuverlässig und genau tun. Unterschieden werden in diesem Zusammenhang die erkundende *exploratory factor analysis* (EFA) und die hypothesenprüfende *confirmatory factor analysis* (CFA; vgl. z. B. Brown, 2006).<sup>1</sup> Obwohl faktorenanalytische Verfahren ursprünglich zur Anwendung bei kontinuierlichen manifesten Variablen bestimmt waren, wurden inzwischen Messmodelle für dichotome bzw. ordinale Indikatorvariablen entwickelt, mit denen nicht nur konfirmatorische sondern auch explorative Faktorenanalysen durchführbar sind (Muthén, 1984; Muthén & Muthén, 2012).<sup>2</sup>

Während in faktorenanalytischen Modellen kontinuierliche latente Variablen bzw. Faktoren angenommen werden, um die Zusammenhänge zwischen den manifesten Variablen zu erklären, wird diese Rolle in der *latent class analysis* (LCA) durch eine nominalskalierte latente Klassenvariable übernommen, mit welcher die Personen in Subgruppen eingeteilt werden (vgl. Geiser, 2011; Rost, 2004; Wang & Wang, 2012; mehr dazu weiter unten im Anwendungsbeispiel). Im engeren Sinne verweist der Begriff LCA auf Modelle, die sich auf kategoriale manifeste Variablen und deren stochastische Zusammenhänge beziehen (bspw. dichotome Items mit nur zwei Ausprägungen wie z. B. *stimmt* und *stimmt nicht*). Modelle mit kontinuierlichen manifesten Variablen werden hingegen als *latent profile analysis* (LPA) bezeichnet (Muthén, 2002), wobei aber auch Indikatorvariablen mit jeweils unterschiedlichem Skalenniveau in einem gemeinsamen Modell berücksichtigt werden können (Muthén & Muthén, 2012).

Zusammenfassend und vereinfachend lassen sich CFA und LCA wie folgt differenzieren: „CFA focuses on grouping items, and thus is a variable-centered approach; in contrast, LCA focuses on grouping respondents or cases based on the patterns of item responses, and thus is a person-centered approach“ (Wang & Wang, 2012, S. 291). An dieser Stelle sei darauf hingewiesen, dass die Latent-Class-Analyse nicht ohne weiteres mit den als *Clusteranalysen* im engeren Sinne bezeichneten Verfahren gleichgesetzt werden kann (für eine kurze Gegenüberstellung vgl. z. B. Rost, 2004, S. 156).

### 3. Ein Anwendungsbeispiel

Die erste COER13-Befragung wurde durch die Veranstalter zu Beginn des Kurses in Form einer Online-Umfrage durchgeführt. In der ersten Befragungswelle wurden unter anderem folgende Themenbereiche berücksichtigt: Vorerfahrungen mit ähnlichen Veranstaltungsformaten, Gründe für die Teilnahme sowie spezifische Nutzungspräferenzen in Bezug auf bestimmte Online-Werkzeuge. Für den letztgenannten Themenbereich lautete eine Teilfrage: „Welche webbasierten Werkzeuge nutzen Sie bereits (also vor dem Kurs) zum Abrufen von Neuigkeiten (also als Leser)?“. Eine zweite Teilfrage lautete: „Welche webbasierten Werkzeuge nutzen Sie bereits (also vor dem Kurs) um Beiträge zu publizieren

---

<sup>1</sup> Auf die in diesem Zusammenhang wichtige testtheoretische Unterscheidung verschiedener Messmodelle der *klassischen Testtheorie* und der *Item-Response-Theorie* soll an dieser Stelle nur verwiesen werden (z. B. Raykov & Marcoulides, 2011; Steyer & Eid, 2001), ebenso wie auf die Unterscheidung zwischen explorativer Faktorenanalyse und Hauptkomponentenanalyse (vgl. Schmitt, 2011).

<sup>2</sup> Die CFA-Modellparameter lassen sich je nach Modellspezifikation in die Parameter der Item-Response-Theorie überführen (vgl. z. B. Muthén & Muthén, 2012).



oder Beiträge zu kommentieren?“. Anschließend konnten diesbezüglich jeweils neun webbasierte Werkzeuge bewertet werden. Insbesondere in Bezug auf die zweite Teilfrage erscheint die Durchführung einer Latent-Class-Analyse nicht nur theoretisch sinnvoll, sondern auch praxisrelevant und im Rahmen des Erfahrungsberichts als geeignet zur Veranschaulichung des Verfahrens, so dass sich die folgenden Darstellungen nun einer solchen LCA widmen.

Als Eingangsvariablen fungierten die insgesamt neun Items zur Nutzung verschiedener Online-Werkzeuge zum Veröffentlichen oder Kommentieren von Beiträgen (vgl. *Abbildung 1* auf Seite 4). Das ursprüngliche Antwortformat umfasste vier Antwortmöglichkeiten: (a) *kenne ich nicht*, (b) *kenne ich, aber nutze ich nicht*, (c) *nutze ich selten* und (d) *nutze ich regelmäßig*. Zur Vereinfachung von Interpretation und nachfolgenden Darstellungen und nicht zuletzt, um Konvergenz- und Schätzprobleme bei den Modellschätzungen zu vermeiden, wurden die ersten drei Kategorien zusammengefasst, so dass 9 dichotome Variablen resultieren, die darüber Auskunft geben, ob jemand ein bestimmtes Online-Werkzeug regelmäßig zum Veröffentlichen und Kommentieren von Beiträgen nutzt oder nicht (vgl. Geiser, 2012, S. 239). An der Online-Befragung nahmen insgesamt 475 Personen teil, davon beantworteten 25 Personen jedoch keines dieser 9 Items, so dass insgesamt nur 450 Personen in der nachfolgenden Analyse berücksichtigt werden können (269 Frauen und 181 Männer).

In einem ersten Schritt wurden mit dem Programm Mplus 7.11 (Muthén & Muthén, 2012) zunächst fünf Modelle mit jeweils unterschiedlicher Klassenanzahl berechnet: ein 1-Klassenmodell, ein 2-Klassenmodell, ein 3-Klassenmodell, ein 4-Klassenmodell und ein 5-Klassenmodell.<sup>3</sup> In Bezug auf die Entscheidung über die geeignete Klassenanzahl sind in *Tabelle 1* die Werte für die Informationskriterien *AIC*, *BIC* und adjustiertes *BIC* sowie die inferenzstatistischen Ergebnisse des modellvergleichenden *bootstrap likelihood ratio tests* (BLRT) dargestellt (vgl. z. B. Nylund, Asparouhov & Muthén, 2007).

*Tabelle 1: Modellvergleich verschiedener Klassenlösungen*

Modell	AIC	BIC	Adj. BIC	BLR-Diff.test ( $p$ -Wert)
1 Klasse	3678.94	3715.93	3687.36	
2 Klassen	3232.60	3310.68	3250.38	0.000
3 Klassen	3150.23	3269.40	3177.37	0.000
4 Klassen	3105.07	3265.33	3141.56	0.000
5 Klassen	3098.15	3299.50	3143.99	0.030

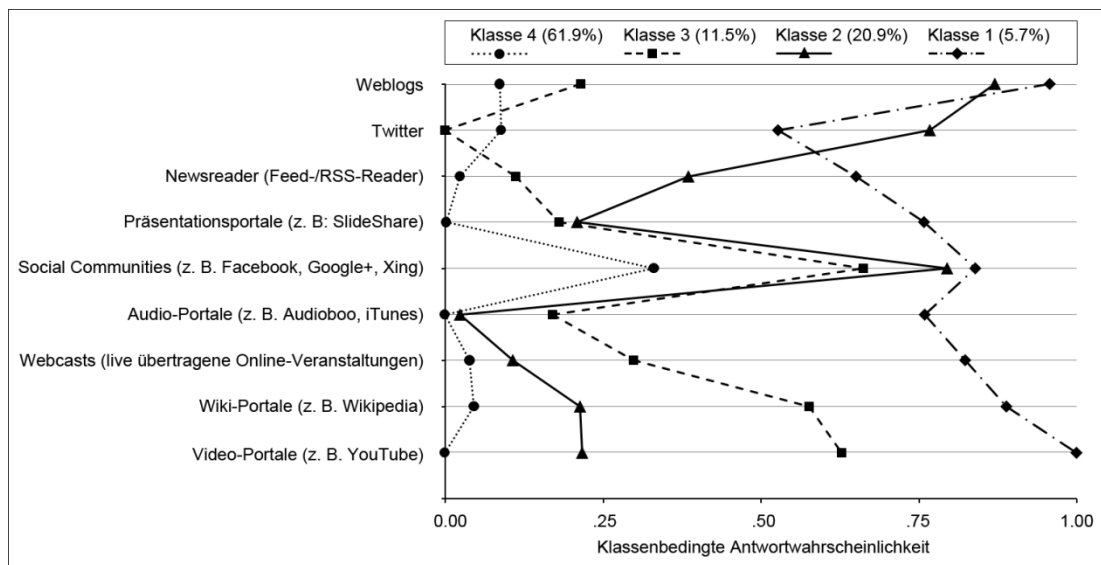
Der BLRT vergleicht das jeweilige  $k$ -Klassenmodell mit dem  $(k-1)$ -Klassenmodell, wobei ein signifikantes Ergebnis für das  $k$ -Klassenmodell spricht. Im vorliegenden Fall sprechen die vergleichsweise geringen Werte von BIC und adjustiertem BIC im Gegensatz zum BLRT jedoch für ein 4-Klassenmodell. In diesem Zusammenhang weist Geiser (2011, S. 267)

<sup>3</sup> Die Modellschätzungen erfolgten unter Verwendung des robusten Maximum-Likelihood-Schätzverfahrens MLR (vgl. Muthén & Muthén, 2012).



auch darauf hin, dass die Entscheidung zwischen verschiedenen Klassenmodellen immer auch unter inhaltlichen Gesichtspunkten der Interpretierbarkeit und nicht nur auf Basis deskriptiver bzw. inferenzstatistischer Kennwerte getroffen werden sollte. Überdies enthält das 5-Klassenmodell relativ viele *boundary estimates* – in diesem Fall liegen die geschätzten Wahrscheinlichkeiten an den Grenzen ihres möglichen Wertebereichs 0 und 1 (vgl. Geiser, 2011, S. 236). Eine Entscheidung zugunsten des 4-Klassenmodells wird außerdem zumindest durch die Werte von BIC und adjustiertem BIC gestützt.

Im Folgenden wird deshalb das 4-Klassenmodell etwas genauer betrachtet. Die geschätzten Klassengrößen betragen im 4-Klassenmodell für die 450 Personen in absoluten und relativen Häufigkeiten: 25.63 Personen bzw. 5.7% (Klasse 1), 94.20 Personen bzw. 20.9% (Klasse 2), 51.72 Personen bzw. 11.5% (Klasse 3) und 278.45 Personen bzw. 61.9% (Klasse 4). Somit kann Klasse 4 als die größte latente Klasse gelten. Die mittleren Klassenzuordnungswahrscheinlichkeiten liegen mit .85 (Klasse 1), .89 (Klasse 2), .90 (Klasse 3) und .92 (Klasse 4) ebenfalls in einem akzeptablen Bereich (vgl. Geiser, 2011). Die für jedes der neun Items geschätzten klassenspezifischen Antwortwahrscheinlichkeiten sind in *Abbildung 1* dargestellt. Die Schätzungen für diese bedingten Wahrscheinlichkeiten beziehen sich jeweils auf die Kategorienausprägung „nutze ich regelmäßig“ (um Beiträge zu veröffentlichen oder zu kommentieren).



*Abbildung 1:* Klassenbedingte Antwortwahrscheinlichkeiten für das 4-Klassenmodell. Die bedingten Wahrscheinlichkeiten beziehen sich auf die Kategorienausprägung „nutze ich regelmäßig“ (um Beiträge zu veröffentlichen oder zu kommentieren).

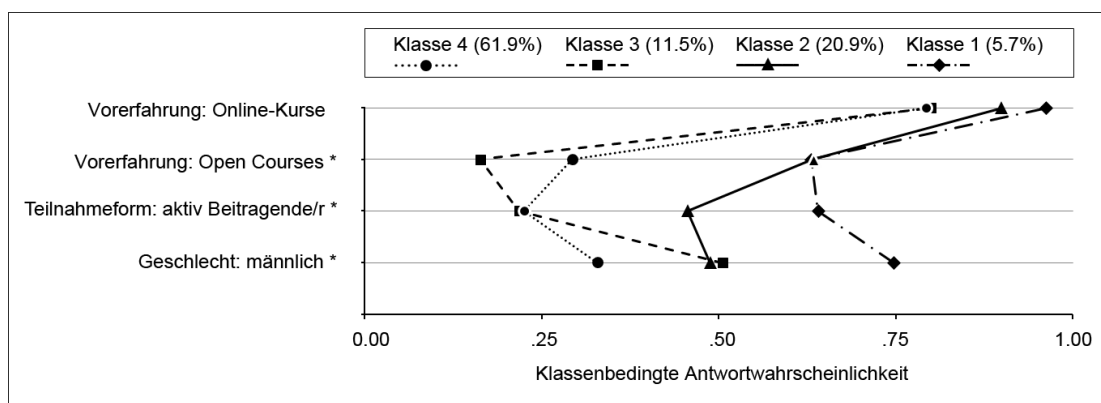
Es lässt sich zunächst feststellen, dass die bedingten Wahrscheinlichkeiten in der größten Klasse (Nr. 4) nur geringe Werte aufweisen. Die in den Items genannten Online-Werkzeuge werden in dieser Klasse also kaum genutzt – zumindest nicht regelmäßig, um Beiträge zu veröffentlichen oder zu kommentieren. Eine Ausnahme bilden jedoch die sozialen Netzwerke, die, wie *Abbildung 1* zeigt, in allen Klassen vergleichsweise häufig zum Veröffentlichen oder Kommentieren von Beiträgen genutzt werden. Im Kontrast zu den geringen Antwortwahrscheinlichkeiten in Klasse 4 zeigen sich in der kleinsten Klasse (Nr. 1) fast



durchgehend relativ hohe bedingte Antwortwahrscheinlichkeiten. Auch der Eindruck einer Verringerung bei dem auf Twitter bezogenen Item relativiert sich bei einem Blick auf die geschätzten Wahrscheinlichkeiten für die regelmäßige Nutzung von Twitter in den Klassen 3 und 4.

Es kann also davon ausgegangen werden, dass Personen, die sich Klasse 1 zuordnen lassen, regelmäßig eine eher große Bandbreite von Online-Werkzeugen zum Veröffentlichen oder Kommentieren von Beiträgen nutzen. Demgegenüber scheint sich die regelmäßige Nutzung in Klasse 2 – zusätzlich zur Nutzung sozialer Netzwerke – in hohem Maße auf Weblogs und Twitter zu fokussieren. In Klasse 3 ist die regelmäßige Nutzung von Online-Werkzeugen hingegen vor allem – und auch hier wieder parallel zu sozialen Netzwerken – bei Video- und Wiki-Portalen wahrscheinlicher als die Nutzung anderer Online-Werkzeuge.

Eine Möglichkeit zur weiterführenden Beschreibung von Klassen besteht darin, zusätzliche Variablen zu berücksichtigen und auf mögliche Unterschiede zwischen den latenten Klassen zu überprüfen. Durch spezielle und z. B. in Mplus implementierte Schätzverfahren sind diese Vergleiche möglich, ohne dass die Personen in einem unsicherheitsbehafteten Zwischenschritt anhand geschätzter bedingter Zuordnungswahrscheinlichkeiten zunächst einer der latenten Klassen zugeordnet werden müssen. Auch ohne diesen Zwischenschritt ist es möglich, die extrahierten Klassen hinsichtlich kategorialer oder kontinuierlicher Drittvariablen miteinander zu vergleichen (vgl. Muthén & Muthén, 2012). Abschließend wird deshalb in *Abbildung 2* veranschaulicht, in Bezug auf welche Variablen sich die vier latenten Klassen u. a. noch unterscheiden. Bei allen vier Items lag ein dichotomes Antwortformat vor. Die geschätzten klassenspezifischen Wahrscheinlichkeiten beziehen sich jeweils auf die Bejahung der in *Abbildung 2* genannten Ausprägungen.



*Abbildung 2:* Klassenbedingte Antwortwahrscheinlichkeiten für das 4-Klassenmodell. Die geschätzten Wahrscheinlichkeiten beziehen sich auf die Bejahung der jeweils genannten Ausprägungen. Bei der Teilnahmeform bzw. der Art der Teilnahme handelt es sich um die *geplante* Form der Teilnahme. \*Globaltest auf mindestens einen Unterschied ist signifikant ( $p$  jeweils kleiner als .01).

Wie sich zeigt, unterscheiden sich die vier latenten Klassen nicht in Bezug auf die *Vorerfahrung mit Online-Kursen*, die in allen vier Klassen in relativ hohem Maße vorhanden zu sein scheint. Unterschiede zeigen sich jedoch in Bezug auf die *Vorerfahrung mit Open Courses*. Auch nach Adjustierung für multiples Testen (Holm, 1979) zeigen sich diesbezüglich signifikante Unterschiede (a) *zwischen den Klassen 1 und 3*,  $\chi^2(1, N = 450) = 6.53$ ,  $p_{\text{adj.}} =$



.044, (b) zwischen den Klassen 2 und 3,  $\chi^2(1, N = 450) = 25.08, p_{\text{adj.}} < .001$ , sowie (c) zwischen den Klassen 2 und 4,  $\chi^2(1, N = 450) = 23.38, p_{\text{adj.}} < .001$ . Mit Vorerfahrungen in Bezug auf Open Courses kann somit vergleichsweise häufig in den latenten Klassen 1 und 2 gerechnet werden. Dieses Befundmuster wird durch die Ergebnisse zur geplanten *Form der Teilnahme als aktiv Beitragende/r* gestützt. Hier erweisen sich (ebenfalls nach Holm-Adjustierung) folgende Unterschiede als signifikant: (a) zwischen Klasse 1 und 3,  $\chi^2(1, N = 450) = 9.45, p_{\text{adj.}} = .008$ , (b) zwischen Klasse 1 und 4,  $\chi^2(1, N = 450) = 15.80, p_{\text{adj.}} < .001$ , sowie (c) zwischen Klasse 2 und 4,  $\chi^2(1, N = 450) = 11.98, p_{\text{adj.}} = .005$ .

Die geplante Teilnahme als *aktiv Beitragende/r* (im Gegensatz zu *Zuhörer/in* bzw. *Leser/in*) ist somit vor allem in den Klassen 1 und 2 in vergleichsweise hohem Ausmaß wahrscheinlich. Insgesamt unterstreichen diese Ergebnisse das Befundmuster aus *Abbildung 1*. Die in *Abbildung 2* dargestellten Ergebnisse verweisen zusätzlich jedoch auch auf Unterschiede in den klassenspezifischen Geschlechterverteilungen. Während Klasse 4 als der zugleich größten Klasse mehr Frauen angehören, kehrt sich dieses Muster in Klasse 1 um, wobei sich zumindest dieser Unterschied (auch nach Holm-Adjustierung) als signifikant erweist,  $\chi^2(1, N = 450) = 7.92, p_{\text{adj.}} = .030$ . Voreilige Interpretationen dieses sicher interessanten Befundes sollen an dieser Stelle jedoch vermieden werden. Um als robustes Befundmuster zu gelten, müssten sich diese Ergebnisse auch für ähnliche Online-Kurse zeigen. Außerdem sollte nicht zuletzt auch das tatsächliche Verhalten während des Kursverlaufs betrachtet werden, welches in den bisherigen Darstellungen aufgrund des Befragungszeitpunktes (zu Beginn des Kurses) nicht berücksichtigt werden konnte.

#### 4. Fazit

Die bisher vorgestellten Methoden- und Ergebnisdarstellungen könnten durchaus noch weiter vertieft werden – beispielsweise durch faktorenanalytische Verfahren. Ferner sei zudem auf die Kombination faktorenanalytischer Verfahren mit Latent-Class-Analysen hingewiesen (vgl. hierzu auch Lubke & Muthén, 2005).

An dieser Stelle sei als Fazit zunächst aber erst einmal festgehalten, dass sich der Ansatz der LCA auch im Kontext des hier vorgestellten COER13-Beispiels als durchaus nützlich und vielversprechendes Verfahren erweist. Eine erste Botschaft könnte an dieser Stelle lauten, dass sich bereits (bzw. zumindest) zu Beginn des Kurses Hinweise darauf finden lassen, dass es „die typische Teilnehmerin“ oder „den typischen Teilnehmer“ nicht gibt. Vielmehr haben die hier vorgestellten Ergebnisse gezeigt, dass von recht unterschiedlichen latenten Teilnehmerklassen ausgegangen werden kann, in denen bestimmte Vorerfahrungen, spezifisches Vorwissen, spezifische Interessen, Präferenzen und wohl auch spezifische Erwartungen mehr oder weniger wahrscheinlich sind. Die Berücksichtigung solcher Befundmuster kann in der Praxis hilfreich sein, gerade um der individuellen Teilnehmerin und dem individuellen Teilnehmer geeignete Angebote, Ideen, Möglichkeiten und Strukturen zur Verfügung zu stellen, die weder unter- noch überfordern, sondern ansprechen, unterstützen und zur Beteiligung anregen.

#### Literaturverzeichnis

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.



- Geiser, C. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70.
- Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, 10(1), 21-39.
- Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49(1), 115-132.
- Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, 29(1), 81-117.
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide*. Seventh Edition [Programmhandbuch]. Los Angeles, CA: Muthén & Muthén.
- Nylund, K. L., Asparouhov, T. & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14(4), 535-569.
- Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York: Routledge.
- Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion* (2. Aufl.). Göttingen: Huber.
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29(4), 304-321.
- Steyer, R., & Eid, M. (2001). *Messen und Testen* (2. Aufl.). Berlin: Springer.
- Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. New York: Wiley.

## Autor



Name: Jens Jirschitzka

E-Mail: [jjirschitzka@iwm-kmrc.de](mailto:jjirschitzka@iwm-kmrc.de)

### Kurzbiografie:

Jens Jirschitzka ist Diplom-Psychologe und war von 2008 bis 2010 wissenschaftlicher Mitarbeiter der Abteilung Kommunikationspsychologie an der Friedrich-Schiller-Universität in Jena, danach Promotionsstipendiat an der dortigen interdisziplinären Doktorandenschule „Laboratorium Aufklärung“. Seit Februar 2014 ist Jens Jirschitzka wissenschaftlicher Mitarbeiter am Leibniz-Institut für Wissensmedien (IWM) in Tübingen.