

Parameterfreie Entmischung von Reaktionszeitdaten

Rainer W. Alexandrowicz

25. Oktober 2019

1 Überblick

Reaktionszeiten sind charakteristischerweise rechtsschief verteilt und können keine (sinnvollen) Werte kleiner als null annehmen. Deren Auswertung mit Modellen, die auf der Normalverteilung aufbauen, stellt daher einen unbefriedigenden Ansatz dar. Doch es wurden bereits eine Reihe von Modellierungsstrategien speziell für Reaktionszeitdaten entwickelt. Das Diffusionsmodell nach Ratcliff hat in den letzten Jahren – nicht zuletzt aufgrund zunehmender Verfügbarkeit geeigneter Schätzroutinen – große Bedeutung zur Reaktionszeitmodellierung erlangt.

Ein Problem, das bei Reaktionszeitdaten stets zu berücksichtigen ist, ist die Möglichkeit der Vermengung unterschiedlicher Antwortprozesse, die sich in charakteristischer Form auf die gemessene Zeit auswirken. Vor allem vor-schnelle Antworten sowie extrem verzögerte Antworten können als Hinweis auf das Vorliegen konkurrierender Prozesse verstanden werden. Aus methodischer Sicht liegt somit eine Mischverteilungsproblematik vor. Im vorliegenden Beitrag soll ein latenter Entmischungsansatz auf dieses Problem angewendet werden.

Im Rahmen einer Simulationsstudie werden Mischverteilungen auf Basis des Diffusionsmodells erzeugt und mit dem parameterfreien Entmischungsansatz nach [Benaglia, Chauveau, Hunter, und Young \(2009\)](#) analysiert.

Literatur

Benaglia, T., Chauveau, D., Hunter, D. R. & Young, D. S. (2009). `mixtools`: An **R** Package for Analyzing Finite Mixture Models. *Journal of Statistical Software*, 32, 1–29.

25. Workshop Angewandte Klassifikationsanalyse (AKA)

Arbeitstitel:

Wie der Vater so der Sohn, wie die Mutter so die Tochter? Die Klassifikation von Familien anhand beruflicher Interessen unter Berücksichtigung der stereotypen Ähnlichkeit.

Florian G. Hartmann¹ und Jutta von Maurice²

¹ Universität der Bundeswehr München

² Leibniz-Institut für Bildungsverläufe

In der folgenden Studie wird eine Analysestrategie präsentiert, mit der Familien unter Berücksichtigung von Geschlechts- und Kohorteneffekten anhand der beruflichen Interessen der Familienmitglieder in homogene Gruppen eingeteilt werden können. Dabei werden zunächst stereotyp-adjustierte Profilkorrelationen berechnet, um die dyadische Ähnlichkeit zwischen Eltern und ihren Kindern sowie zwischen Müttern und Vätern zu quantifizieren. Anschließend wird eine Clusteranalyse mit Doppelkreuzvalidierung basierend auf den innerfamilialen Interessenrelationen durchgeführt. Die Ergebnisse der Clusteranalyse werden mit jenen verglichen, die im Falle nicht adjustierter Profilkorrelationen resultieren. Für die Analysen stehen die beruflichen Interessen von 197 Familien-Triaden zur Verfügung, die vor dem Hintergrund der Theorie von Holland (1997) mit Hilfe des Allgemeinen Interessen-Struktur-Tests (Bergmann & Eder, 1992) erhoben wurden.

Werden stereotypadjustierte Profilkorrelationen berechnet, ergeben sich vier homogene Gruppen von Familien, die hinsichtlich der Ähnlichkeitsmuster mit den Begriffen ‚Familien-Cluster‘, ‚Eltern-Cluster‘, ‚Mutter-Cluster‘ und ‚Vater-Cluster‘ bezeichnet werden können. Basierend auf nicht adjustierten Profilkorrelationen ergeben sich vier Gruppen mit vergleichbaren Interessensrelationen, wobei jedoch nur 61.4% der klassifizierten Familien äquivalenten Gruppen angehören.

25. Workshop Angewandte Klassifikationsanalyse (AKA)

Arbeitstitel:

Konfigurationen und Latente Klassen von PISA Kompetenzen mit Kovariaten

Jörg-Henrik Heine¹ und Mark Stemmler²

¹Technische Universität München (TUM), School of Education, Zentrum für internationale
Bildungsvergleichsstudien (ZIB)

²Friedrich-Alexander-University Erlangen-Nürnberg (FAU)

Abstract:

Theorie

Zur anschaulich, inhaltlichen Interpretation des in PISA erreichten Kompetenzniveaus klassifiziert PISA die Leistungen der Schülerinnen und Schüler für jede der drei Domänen (*Naturwissenschaften*, *Lesen* und *Mathematik*) in distinkte, ordinal aufsteigende Kompetenzstufen (z. B. Heine, Sälzer, Borchert, Siberns & Mang, 2013; Heine, Mang, Borchert, Gomolka, Kröhme, Goldhammer & Sälzer, 2016). In Sekundäranalysen, basierend auf PISA Daten, wird das erreichte Kompetenzniveau typischerweise mit unterschiedlichen Variablen in Bezug gesetzt – entweder auf der Systemebene (z. B. Unterrichtsqualität – z. B. Schiepe-Tiska, Reiss, Obersteiner, Heine, Seidel & Prenzel, 2013) oder auf der Individualebene (z. B. Interesse und Motivation – vgl. Schiefele & Wild, 2001). Neben solchen Variablen erhebt PISA auch Informationen zur geographischen Herkunft der Schülerinnen und Schüler und deren Eltern. Diese Variable zum so genannten ‚*Migrationshintergrund*‘ hat sich, nicht zuletzt durch PISA, in der (Bildungs-)Wissenschaft, aber auch in anderen gesellschaftlichen Bereichen, als prävalente Differenzkategorie (vgl. Stošić, 2017) zur Beschreibung von Ungleichheiten festgesetzt.

Trotz der Popularität der Variable ‚*Migrationshintergrund*‘ in der sozialwissenschaftlichen Forschung muss aber die Frage nach deren inhaltlich, substanzwissenschaftlichen Bedeutung gestellt werden und deren Erklärungswert überprüft werden. So kann beispielsweise angezweifelt werden, dass der ‚*Migrationshintergrund*‘ als nationale, kulturelle Konstruktion speziell im bildungswissenschaftlichen Kontext überhaupt eine Differenz per se repräsentiert (vgl. Dannenbeck, Lösch & Eßer, 2001). Der vorliegende Beitrag untersucht ausgehend von der empirisch zunächst vorgefundenen Assoziation zwischen dem ‚*Migrationshintergrund*‘ und der PISA-Kompetenz, ob sich diese Verbindung über individuelle Merkmale zur Bildungsaffinität und sozialen Lage, auflösen lässt. Es soll somit überprüft werden, ob es sich bei der Variable ‚*Migrationshintergrund*‘, in Analogie zu dem von Christian Morgenstern verfassten Gedicht ‚*Die Westküsten*‘ (vgl. Morgenstern, 1975), um eine zwar gesellschaftlich prävalente, kulturell konstruierte Gruppenbezeichnung handelt, deren bildungswissenschaftlicher Erklärungswert aber möglicherweise begrenzt ist.

Methode

Die methodische Grundlage besteht in der vergleichenden Anwendung unterschiedlicher Modelle der Latent Class Analysis (LCA) und der Konfigurationsfrequenzanalyse (KFA), als Werkzeug zur Analyse multidimensionaler Kontingenztafeln in der personenorientierten Forschung (vgl. Stemmler, 2014). Getestet wird das KFA-Hauteffektmodell und ein erweitertes Modell mit kontinuierlichen Kovariaten. Für die LCA werden Modelle mit einer unterschiedlichen Anzahl latenter Klassen und zusätzlichen Informationen aus begleitenden Variablen oder Kovariaten gerechnet. Durchgeführt werden die Analysen mit dem R-Paket `poLCA` (Linzer & Lewis, 2011) und `confreq` (Heine, Alexandrowicz, & Stemmler, 2019), in dem die KFA nach dem log-linearen Modellierungsansatz implementiert ist (von Eye und Mun, 2013). Aus der Perspektive eines log-linearen Modellierungsansatzes stellen die Typen und Antitypen der KFA lokale Abweichungen vom formulierten Modell dar. Durch den Vergleich der Passung des Haupteffektmodells mit dem erweiterten Modell, lässt sich die Bedeutung der Kovariaten zur Vorhersage der beobachteten Häufigkeiten der Konfigurationen bewerten. Die inhaltliche Bedeutung der gefundenen latenten

Klassen aus der LCA werden mit den Ergebnissen aus der KFA qualifiziert und Validiert; bedeutsame latente Klassen erscheinen in der KFA als Typen.

Die eingesetzten Daten stammen aus der PISA Erhebung in Deutschland im Jahre 2015

Zur Begründung der Konfigurationen sowie latenten Klassen werden die sechs- bis siebenstufigen kategorialen Variablen zu den PISA Kompetenzstufen sowie die dreistufige Variable zum ‚*Migrationshintergrund*‘ verwendet. Als kontinuierliche Kovariaten werden individuelle Merkmalsindikatoren zur sozialen Lage sowie Motivation und individuelles Interesse an Naturwissenschaften berücksichtigt.

Vorläufige Befunde

Die Analyse der Konfigurationen ergibt signifikante Typen und Antitypen im Haupteffektmodell, die zunächst auf eine Verbindung zwischen der PISA Kompetenzstufe und dem ‚*Migrationshintergrund*‘ zu belegen scheinen. Allerdings verschwinden diese Typen und Antitypen unter der Berücksichtigung der individuellen Kovariaten im erweiterten KFA-Modell.

Literatur

- Dannenbeck, C., Lösch, H., & Eßer, F. (2001). An-, Zu- und Ungehörigkeiten Jugendlicher: Herkunft als Auskunft? In H. Lutz & N. Wenning (Hrsg.), *Unterschiedlich verschieden: Differenz in der Erziehungswissenschaft* (S. 231–248). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-663-11705-6_13
- Heine, J.-H., Alexandrowicz, R. W., & Stemmler, M. (2019). confreq: Configural Frequencies Analysis Using Log-Linear Modeling (Version R package version 1.5.4-5). Abgerufen von <https://CRAN.R-project.org/package=confreq>
- Heine, J.-H., Mang, J., Borchert, L., Gomolka, J., Kröhme, U., Goldhammer, F., & Sälzer, C. (2016). Kompetenzmessung in PISA 2015. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme, & O. Köller (Hrsg.), *PISA 2015: Eine Studie zwischen Kontinuität und Innovation* (S. 383–430). Münster: Waxmann Verlag.
- Heine, J.-H., Sälzer, C., Borchert, L., Siberns, H., & Mang, J. (2013). Technische Grundlagen des fünften internationalen Vergleichs. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 309–346). Münster: Waxmann Verlag.
- Linzer, D. A., & Lewis, J. B. (2011). PolLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software*, 42(10), 1–29.
- Morgenstern, C. (1975). Die Westküsten. In *Christian Morgenstern: Ausgewählte Werke* (S. 264–266). Leipzig: Insel-Verlag.
- Schiefele, U., & Wild, K.-P. (2000). *Interesse und Lernmotivation: Untersuchungen zu Entwicklung, Förderung und Wir*
- Schiepe-Tiska, A., Reiss, K., Obersteiner, A., Heine, J.-H., Seidel, T., & Prenzel, M. (2013). Mathematikunterricht in Deutschland: Befunde aus PISA 2012. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 123–154). Abgerufen von http://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungen/PISA_2012_EBook_ISBN3001.pdf
- Stemmler, M. (2014). Person-centered methods: Configural Frequency Analysis (CFA) and Other Methods for the Analysis of Contingency Tables. Series: Springer Briefs in Statistics: New York: Springer Publishing Company.
- Stošić, P. (2017). Kinder mit ‚*Migrationshintergrund*‘. Reflexionen einer (erziehungs-) wissenschaftlichen Differenzkategorie. In I. Diehm, M. Kuhn, & C. Machold (Hrsg.), *Differenz-Ungleichheit-Erziehungswissenschaft: Verhältnisbestimmungen im (Inter-)Disziplinären* (S. 81–99). Wiesbaden: Springer VS.

25. Workshop Angewandte Klassifikationsanalyse (AKA)

Arbeitstitel:

Vertiefende Analysen zu differentiellen impliziten Antwortmodellen und Antworttendenzen

Jörg-Henrik Heine¹² und Christian Tarnai²

¹Technische Universität München (TUM), School of Education, Zentrum für internationale
Bildungsvergleichsstudien (ZIB)

²Universität der Bundeswehr, München, Neubiberg

Arbeitsskizze und Theorie

Die Beziehung zwischen Personen und Items in psychodiagnostischen Fragebogen kann entweder als *Dominanz* oder als *Nähe–Distanz*-Relation charakterisiert werden. Durch diese Beziehungen können zwei unterschiedliche Antwortprozesse theoretisch begründet werden, welche die Grundlage für eine ganze Reihe unterschiedlicher psychometrischer Mess- und Antwortmodelle bilden (z. B. Stokman & van Schuur, 1980; Heine, 2019 eingereichte Dissertation). Im Gegensatz zu den verbreiteten Modellen zur summativen Skalierung zur Abbildung einer *Dominanz*-Relation (z. B. PCM – Masters, 1982), welche eine monoton steigenden Item-Characteristic-Curve (ICC) aufweisen, haben Modelle zur Abbildung einer *Nähe–Distanz*-Relation eine eingipflige ICC und werden unter dem Begriff *Unfolding* zusammengefasst (z. B. Coombs, 1950; Schönemann, 1970).

Bei der Auswertung von Fragebogenverfahren nach einem summativen Skalierungsmodell, ergeben sich immer wieder Antwortmuster, welche von den eigentlich (vom Testentwickler oder Anwender) erwarteten Antwortmustern abweichen. Bei Fragebögen mit einem mehrstufigen Antwortformat zeigt sich oft eine unterschiedliche Tendenz zu entweder mittleren (middle response style – MRS) oder extremen (extreme response style – ERS) Antwortkategorien (z. B. Wetzel, Lüdtke, Zettler & Böhnke, 2016), welche sich typischerweise mit Mischverteilungsmodellen (z. B. Rost, Carstensen & von Davier, 1997) darstellen lassen.

Daneben zeigt sich bei der faktorenanalytischen Auswertung von Fragebogendaten oft ein artifizieller, zusätzlicher *Methodenfaktor* (z. B. Maraun & Rossi, 2001; Matschinger & Krebs, 1998; Post, van Duijn & van Baarsen, 2001; Schönemann, 1970; van Schuur & Kiers, 1994). Dieser *Methodenfaktor* kann nach Schönemann (1970) und Coombs (1967, S. 181-182) durch das Vorhandensein individueller Präferenzurteile als sogenannte *I–Skalen* (vgl. Coombs, 1950) in den Daten erklärt werden, welche sich aus einer (abweichenden) Interpretation der psychometrischen Skala im Rahmen eines Antwortprozesses nach der *Nähe–Distanz*-Relation ergeben.

In der Arbeit von Heine (2019 eingereichte Dissertation) werden in den Fragebogenskalen von drei Konstrukten (*Persönlichkeit – Big-Five Modell, berufliche Interessenorientierungen und Präferenzen des Musikgeschmacks*) in einer personenorientierten Analyse Personengruppen identifiziert, welche sich jeweils den beiden Antwortprozessen zuordnen lassen. Ferner wird gezeigt, dass eine systematische Beziehung zwischen der Antworttendenz im Rahmen eines Dominanz-Antwortmodells und der Passung zu beiden Antwortprozessen besteht. Diese Beziehung lässt sich über eine von Stark, Chernyshenko, Drasgow und Williams (2006) formulierte theoretische Betrachtung zur partiellen Identität der ICCs beider Antwortmodelle in Abhängigkeit der Differenz aus Merkmalsausprägung und Itemschwierigkeit, erklären. So ergeben sich bei Heine (2019 eingereichte Dissertation) für vier Skalen des Konstruktes beruflicher Interessenorientierungen signifikante Zusammenhänge zwischen einer Zuordnung zu beiden Antwortprozessen mit einer Antworttendenz zur Mitte. Die Antworttendenz zu Extremen ist mit der Nichtgültigkeit beider Antwortprozesse verbunden. Dieser Befund deckt sich im Wesentlichen mit der von Stark, et. all. (2006) formulierten theoretischen Betrachtung zur Identität der ICCs beider Antwortmodelle in Abhängigkeit der Differenz aus Merkmalsausprägung und Itemschwierigkeit.

Da sich die ICCs der beiden Antwortprozesse allerdings wenigstens im oberen Bereich der Merkmalsausprägung unterscheiden sollten (nicht so im unteren Bereich), ergibt sich ein Ansatzpunkt für weitere Analysen, welche sich auf diese Personengruppen konzentrieren.

In der hier angedachten erweiterten Analyse, sollen daher die Antwortmuster derjenigen Personen vertieft betrachtet werden, welche einerseits keinem der beiden Antwortprozesse zugeordnet werden konnten und gleichzeitig eine Antworttendenz zu extremen Antwortkategorien aufweisen.

Daten

Die Datengrundlage bilden zwei Stichproben aus dem Erhebungszeitraum 2007 – 2009 (Stichprobe I) und 2010 – 2011 (Stichprobe II) von Studierenden der Universität der Bundeswehr mit einem Umfang von $n = 1343$ antwortenden Personen. Die hier zu analysierenden abgeleiteten Indikatorvariablen stammen aus vorausgehenden Analysen (vgl. Heine, 2019 eingereichte Dissertation). Einerseits wurde hier die Antworttendenz zu entweder eher mittleren oder extremen Antwortkategorien bei mehrstufigen Antwortformaten identifiziert. Andererseits wurden dichotome Indikatorvariablen zum impliziten Antwortmodell bei der Beantwortung von insgesamt 15 Skalen in drei Konstrukten gebildet.

Methode

Die methodische Grundlage bildet die Anwendung unterschiedlicher Modelle der Latent Class Analysis (LCA) und der Konfigurationsfrequenzanalyse (KFA), als Werkzeug zur Analyse multidimensionaler Kontingenztabelle in der personenorientierten Forschung (vgl. Stemmler, 2014). Durchgeführt werden sollen die Analysen mit dem R-Paket `poLCA` (Linzer & Lewis, 2011) und `confreq` (Heine, Alexandrowicz, & Stemmler, 2019), in dem die KFA nach dem log-linearen Modellierungsansatz implementiert ist (von Eye und Mun, 2013). In den Analysen sollen die abgeleiteten Indikatorvariablen aus Heine (2019 eingereichte Dissertation), sowie die Antwortmuster der als „Extremkreuzer“ klassifizierten Personen betrachtet werden.

Literatur

- Coombs, C. H. (1950). Psychological scaling without a unit of measurement. *Psychological Review*, 57 (3), 145–158. doi: 10.1037/h0060984
- Coombs, C. H. (1967). *A theory of data* (2. Aufl.). New York: Wiley.
- Heine, J.-H. (2019 eingereichte Dissertation). *Untersuchungen zum Antwortverhalten und zu Modellen der Skalierung bei der Messung psychologischer Konstrukte* (Monographie). München, Neubiberg.
- Maraun, M. D. & Rossi, N. T. (2001). The extra-factor phenomenon revisited: Unidimensional unfolding as quadratic factor analysis. *Applied Psychological Measurement*, 25 (1), 77–87. doi:10.1177/01466216010251006
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47 (2), 149–174. doi: 10.1007/BF02296272
- Matschinger, H. & Krebs, D. (1998). Zum Problem der Abbildung eindimensional konzipierter Konstrukte bei entgegengesetzter Itempolung. *ZA-Information / Zentralarchiv für Empirische Sozialforschung*, 43, 81–110.
- Post, W. J., van Duijn, M. A. & van Baarsen, B. (2001). Single-peaked or monotone trachelines? On the choice of an IRT model for scaling data. In A. Boomsma, M. A. J. van Duijn & T. A. B. Snijders (Hrsg.), *Essays on Item Response Theory* (Bde. Lecture notes in statistics, 157, S. 391–413). New York: Springer.
- Rost, J., Carstensen, C. & von Davier, M. (1997). Applying the mixed Rasch model to personality questionnaires. In J. Rost & R. Langeheine (Hrsg.), *Applications of latent trait and latent class models in the social sciences* (S. 324–332). Münster: Waxmann Verlag.
- Schönemann, P. H. (1970). On metric multidimensional unfolding. *Psychometrika*, 35 (3), 349–366.
- Stark, S., Chernyshenko, O. S., Drasgow, F. & Williams, B. A. (2006). Examining assumptions about item responding in personality assessment: Should ideal point methods be considered for scale development and scoring? *Journal of Applied Psychology*, 91 (1), 25–39.
- Stokman, F. & van Schuur, W. H. (1980). Basic scaling. *Quality and Quantity*, 14 (1), 5–30. doi: 10.1007/BF00154792
- van Schuur, W. H. & Kiers, H. A. L. (1994). Why factor analysis often is the incorrect model for analyzing bipolar concepts, and what model to use instead. *Applied Psychological Measurement*, 18 (2), 97–110. doi:10.1177/014662169401800201
- Wetzel, E., Lüdtke, O., Zettler, I. & Böhnke, J. R. (2016). The stability of extreme response style and acquiescence over 8 years. *Assessment*, 23 (3), 279–291. doi: 10.1177/1073191115583714
- Stemmler, M. (2014). *Person-centered methods: Configural Frequency Analysis (CFA) and other methods for the analysis of contingency tables*. New York: Springer Publishing Company.
- Linzer, D. A. & Lewis, J. B. (2011). `poLCA`: An R package for polytomous variable latent class analysis. *Journal of Statistical Software*, 42 (10), 1–29.
- Heine, J.-H., Alexandrowicz, R. W. & Stemmler, M. (2019). `confreq`: Configural frequencies analysis using Log-Linear modeling (R package version 1.5.4-5). <https://CRAN.R-project.org/package=confreq>

Identifikation verschiedener Muster von Zufriedenheitseinschätzungen von Kindern in stationär-psychiatrischer Behandlung [Arbeitstitel]

Alternativ?: Wie schätzen Kinder ihre Zufriedenheit mit verschiedenen Aspekten der stationär-psychiatrischen Behandlung ein: eine Identifikation verschiedener Antwortmuster mit LCA

Ferdinand Keller (+ evtl. noch Mitautoren)

Behandlungszufriedenheit ist zu einem wichtigen Bestandteil guter stationärer Behandlung geworden. Ihre Erfassung dient dabei nicht nur den Therapieeinrichtungen, die sie als Qualitätsindikator und damit als Rückmeldung für ihr „Angebot“ nutzen können, sondern sie ist auch auf Seiten der Behandelten für die Motivation und Fortführung einer Behandlung bedeutsam. Während es für Jugendliche und für Eltern mit den FBB und den Best-Bögen schon Erhebungsinstrumente gibt (vgl. Keller, Fegert & Naumann, 2018), liegen für Kinder noch kaum Erhebungsinstrumente vor. Ein inzwischen häufig eingesetzter Fragebogen ist der Best-Bogen in der Kinderversion (Keller et al., 2004).

Er besteht aus 18 Items mit 5 Kategorien, die jeweils mit Text erläutert sind. Eine psychometrische Analyse von $n = 1641$ Kinderbögen aus neun Kliniken mittels „klassischer“ Methoden bringt eine gute interne Konsistenz der Gesamtskala und eine EFA verweist auf Lösungen mit 2 oder 3 Faktoren (PCA mit SPSS auf Basis paarweiser Korrelationen ($n = 1150$), EFA mit WLSMV (Mplus)). Auch Bifactor-Modelle sind aussagekräftig und weisen einen sehr guten Fit auf.

Insgesamt unterscheiden die Kinder gemäß diesen FA zwei Bereiche: Personen (Therapeuten und Betreuer) und Regeln/Umgebung. Bei drei Faktoren zeichnet sich eine Aufteilung in Betreuer/Umgebung und Regeln/Umgebung mit überlappenden (Doppel-)Ladungen auf den Faktoren 2 und 3 ab, was inhaltlich für das Erleben von Kindern nachvollziehbar wäre. In einem bifactor-Modell mit den zwei spezifischen Faktoren „Personen“ und „Regeln/Umgebung“ ergibt sich, dass die Umgebung („Zimmer schön“ u.ä.) hoch auf dem Generalfaktor lädt und keinen spezifischen Anteil mehr aufweist. Die verbleibenden Items zum Auskommen mit anderen Kindern und den Regelungen korrelieren negativ.

Angesichts der nicht ganz eindeutig interpretierbaren Ergebnisse der FA und vor dem Hintergrund, dass manche Kinder vielleicht die Items nicht ganz verstehen oder Antworttendenzen u.ä. aufweisen, z.B. fast immer die Mittelkategorie ankreuzen („Unscalables“), bietet es sich an mit einem personenorientierten Ansatz wie der LCA zu versuchen, Subgruppen zu unterscheiden und eventuell auch "unscalables" zu identifizieren und herauszunehmen. Dazu werden nur die ca. 10 wichtigsten Items des Fragebogens verwendet, um die LCA handhabbar und aussagekräftig zu halten. Weiterhin soll inhaltlich vertieft auf die Items zu „Personen“ (Therapeuten und Betreuer) und ihre differenzierte Wahrnehmung durch die Kinder eingegangen werden, z.B. die interessante Unterscheidung in „Problem verstanden“ und „Zeit für mich gehabt“.

Zwei Auswertungsbereiche:

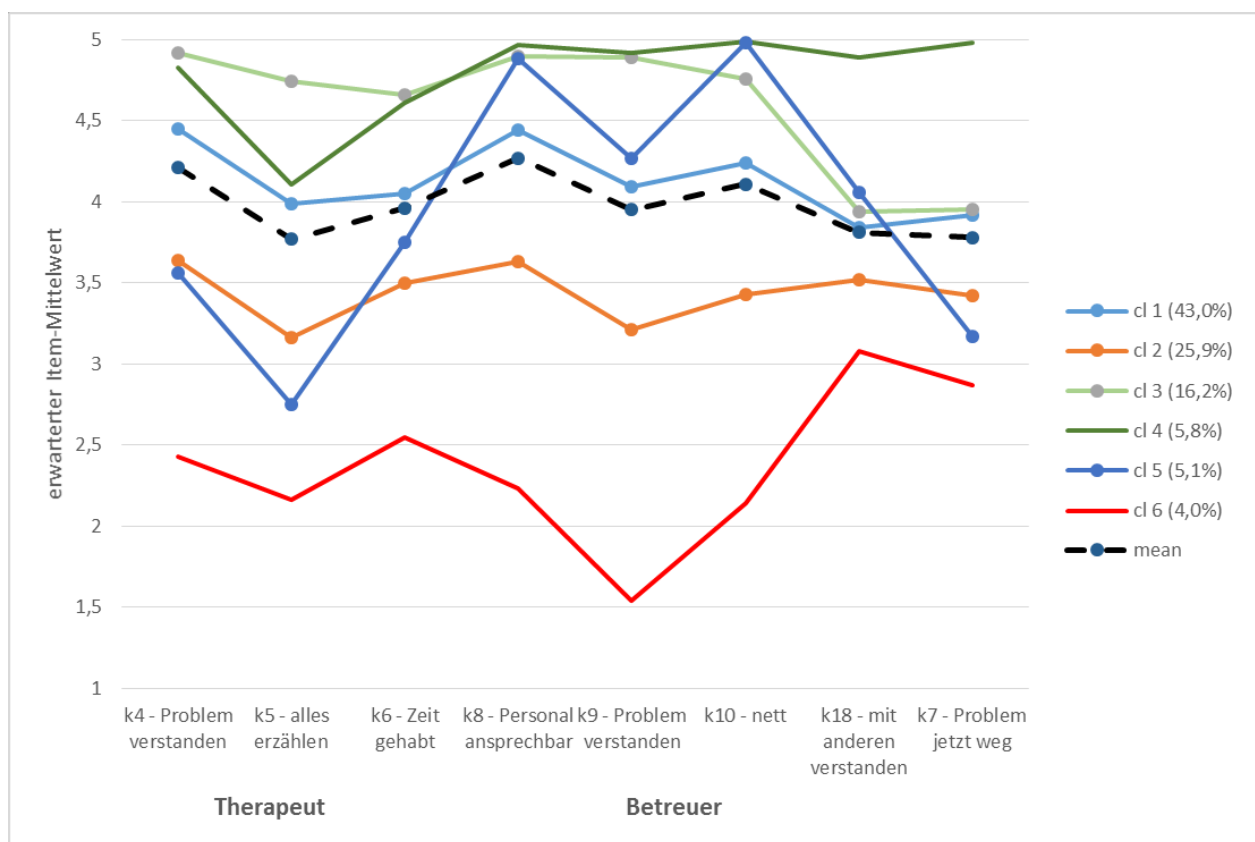
- 1) wichtige Items über den ganzen Fragebogen hinweg: Items x, x, ... [noch offen]
- 2) Items speziell zu Therapeut und zu Betreuer: Items 4-6 (Therapeut), 8-10 (Personal und Betreuer), 18 (mit anderen Kindern verstanden) sowie 7 (Problem jetzt weg), das den Einfluss von Therapieerfolg abbilden soll.

evtl. 3) gegenläufigen Effekt in bifactor-Modell untersuchen, d.h. nach Herausnehmen des Generalfaktors korrelieren zwei Bereiche negativ.
[unklar, ob das ein methodisches Artefakt sein könnte; und ob das überhaupt mit LCA evaluierbar ist!]

Ergebnis zu 2)

Gemäß Informationskriterien werden fünf (CAIC) bzw. sechs (BIC) latente Klassen als beste Lösung vorgeschlagen. Die Klassenprofile der 6-Klassenlösung sind in Abb. 1 dargestellt. Anhand des Gesamtmittels ist ersichtlich, dass die Kinder grundsätzlich in den meisten Items eine relativ hohe Behandlungszufriedenheit angeben. Die größte Klasse (cl 1) liegt sehr nahe beim Gesamtmittel, während die zweitgrößte Klasse (cl 2) darunter liegt, aber im wesentlichen dem Gesamttrend folgt. Die Klasse 3 ist sehr zufrieden mit Therapeuten und Betreuer, hat aber bezüglich

therapeutischem Erfolg („Problem jetzt weg“) keine besseren Werte. Dagegen ist bei Klasse 4 ihrer eigenen Einschätzung nach das Problem nun weg; auch diese Klasse zeichnet eine hohe Zufriedenheit aus, wobei sie trotzdem mit „Therapeut alles erzählen“ etwas unzufrieden sind (auf hohem Niveau). Die Klasse 5 ist sehr zufrieden mit den Betreuern, hat aber Probleme mit dem Therapeuten und (damit?) eine unterdurchschnittliche Erfolgseinschätzung. Die letzte Klasse schließlich ist mit allem unzufrieden und weist in allen Items die niedrigsten Werte auf; sie ist zwar klein, aber klar abgrenzbar und tritt auch in der 5-Klassenlösung als solche auf.



Literatur:

Keller, F., Fegert, J.M., & Naumann, A. (2018). Fragebögen zur Behandlungseinschätzung stationärer Therapie (BesT) in der Kinder- und Jugendpsychiatrie: Entwicklung und Validierung für Jugendliche und für Eltern. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 47, 186-197.

Keller, F., Schäfer, S., Konopka, L., Naumann, A. & Fegert, J.M. (2004). Behandlungszufriedenheit von Kindern in stationär-psychiatrischer Behandlung: Entwicklung und psychometrische Eigenschaften eines Fragebogens. *Krankenhauspsychiatrie*, 15, 3-8.

Entwicklung und Validierung eines Instruments zur Messung von allgemeiner Moralität

Heinz Leitgöb, Stefanie Eifler & Julia Weymeirsch

AKA Rothenberge 2019

In der kriminalsoziologischen bzw. kriminologischen Forschung wird Moralität als einem der handlungsleitenden Prinzipien zentrale Bedeutung beigemessen. Als Beispiel kann hierfür die derzeit in Europa verstärkt rezipierte *Situational Action Theory* (SAT) genannt werden. Ihr zufolge kann kriminelles Handeln von Individuen nur dann in Erwägung gezogen werden, wenn die internalisierten Moralvorstellungen einer Person die Perzeption einer kriminellen Handlungsoption nicht kategorisch ausschließen und somit der *moralische Filter* versagt (z. B. Wikström et al., 2012).

Die bislang im Rahmen der Konfrontation der Moralitätshypothese(n) mit empirischen Daten verwendete Messung von Moralität basiert vornehmlich auf sog. *thin conceptions* (Abend, 2012; Williams, 1985). Diese versuchen Moralität über Fragen nach der empfundenen *Richtigkeit* konkreter Handlungen abzubilden und lassen somit die *thick conceptions* von Moralität inhärenten Bedeutungszuschreibungen, die den Charakter der abgefragten Handlungen spezifizieren, vermissen. Aus diesem Grund wurde der Versuch unternommen, auf der Grundlage der Differenzierung in *ethics of being* und *ethics of doing* eine seins- und eine handlungsorientierte allgemeine Moralitätsskala zu entwickeln, die den Kriterien der *thick conceptions* genügt und sich zur Differenzierung der Itemschwierigkeiten am Grundprinzip des Stufenmodells der Moralentwicklung nach Kohlberg (1974, 1996) orientiert.

Zur Evaluation der psychometrischen Beschaffenheit der Skalen werden konfirmatorische Faktorenmodelle sowie *item response theory* Modelle auf zwei empirische Datensätze angewendet, die jeweils einfache Zufallsauswahlen der Wohnbevölkerung aus zwei größeren deutschen Städten repräsentieren. In diesem Zusammenhang wird ein Datensatz zur Entwicklung des Instruments herangezogen während der zweite der Validierung dient.

Literatur

- Abend, G. (2012). What the science of morality doesn't say about morality. *Philosophy of the Social Sciences*, 43:157–200.
- Kohlberg, L. (1974). *Zur kognitiven Entwicklung des Kindes*. Suhrkamp, Baden Baden.
- Kohlberg, L. (1996). *Die Psychologie der Moralentwicklung*. Suhrkamp, Frankfurt am Main.
- Wikström, P.-O. H., Oberwittler, D., Treiber, K., and Hardie, B. (2012). *Breaking Rules. The Social and Situational Dynamics of Young People's Urban Crime*. Oxford University Press, Oxford.
- Williams, B. A. O. (1985). *Ethics and the Limits of Philosophy*. Harvard University Press, Cambridge, MA.

Abstract für die AKA Tagung in Rothenberge
06.-08.11.2019

Hellfelddaten der Korruption: Konstruktion, Probleme und eine Analyse ihrer Ranginformation mit Hilfe einer Korrespondenzanalyse

Heinz Leitgöb, Peter Graeff & Jörg Blasius

Die prozessbasierte Generierung von Hellfelddaten der Kriminalität besitzt so viele Störquellen, dass eine Auswertung (und oftmals auch eine einfache Beschreibung) von Fall- oder Täterwerten nur selten vorgenommen wurde. Damit wurde diese Möglichkeit, deviante Phänomene empirisch zu untersuchen, bisher weitgehend vernachlässigt.

Anstatt die fehlerbehafteten metrischen Fall- oder Täterwerte zu verwenden, benutzen wir für unsere Analyse Rangdaten von 96 Raumordnungsregionen in Deutschland. Auf diese Weise werden nicht nur Ausreißer (und damit verzerrte Varianzen) eliminiert, sondern auch eine Möglichkeit geschaffen, die Daten mit Hilfe der Korrespondenzanalyse auszuwerten. Um auch metrische Daten auswerten zu können, müssen die Daten zuvor transformiert werden. Dabei muss berücksichtigt werden, dass alle Variablen und alle Untersuchungseinheiten (hier Gebiete) mit dem gleichen Gewicht in der Analyse eingehen. Eine Möglichkeit dies zu gewährleisten ist die Verwendung von Rangdaten.

Wir verwenden verschiedene Korruptionsdelikte für unsere Korrespondenzanalyse, die eine Visualisierung sowohl der Regionen als auch der Korruptionsvariablen ermöglicht. Die Ergebnisse zeigen zwei Dimensionen von Korruptionsdelikten auf – eine Dimension, die überwiegend die bürokratischen/administrativen Delikte umfasst, und eine Dimension, die überwiegend die geschäftlichen Korruptionsprozesse umfasst. Diese Resultate passen zu den Korruptionsbekämpfungsmaßnahmen, die in den verschiedenen deutschen Regionen (meist den Bundesländern) geleistet werden. Damit können Aussagen über die Bedingungen der Korruption in Deutschland getroffen werden, die auf der Basis der administrativen Rohdaten in dieser Weise ansonsten nicht möglich wären.

Entwicklungsverläufe von Jugend- und Erwachsenenkriminalität:
Mischverteilungsmodelle mit Daten der Studie „Kriminalität in
der modernen Stadt“ (CrimoC)

Jost Reinecke / Georg Kessler

Inhaltsverzeichnis

| | | |
|----------|--|----------|
| 1 | Einführung | 3 |
| 2 | Modelltypen | 3 |
| 2.1 | Das klassische Wachstumsmodell | 3 |
| 2.2 | Mischverteilungsmodelle und Verteilungsannahmen für Zählraten | 3 |
| 3 | Empirische Studie | 3 |
| 3.1 | Erhebungsdesign | 3 |
| 3.2 | Bisherige Verlaufsanalysen | 3 |
| 3.3 | Fragestellungen | 3 |
| 4 | Mischverteilungsmodelle | 3 |
| 4.1 | Modelle mit Jugenddelikten | 3 |
| 4.2 | Modelle mit Jugend- Erwachsenenendelikten | 3 |
| 5 | Vergleich der Entwicklungsverläufe | 3 |
| 6 | Zusammenfassung | 3 |

Abstract

Die kriminologische und kriminalsoziologische Längsschnittforschung wird methodisch stark geprägt durch die Anwendung von Wachstums- und Mischverteilungsmodellen, die es ermöglichen, unterschiedliche Entwicklungsverläufe delinquenten Verhaltens statistisch zu schätzen. Mit Daten der Längsschnittstudie „Kriminalität in der modernen Stadt“ (CrimoC, vgl. Boers & Reinecke, 2019) werden diese Verläufe sowohl für das Jugendalter als auch für das junge Erwachsenenalter ermittelt. Im Vordergrund steht die Frage, wie viel Gruppen bzw. latente Klassen benötigt werden, um die unbeobachtete Heterogenität des Delinquenzverlaufs abbilden zu können. Des Weiteren wird der Zusammenhang der Klassenlösungen für das Jugendalter und das junge Erwachsenenalter geprüft, d.h. inwieweit können gleiche Verlaufstypen identifiziert werden und/oder kann ein neuer Verlaufstyp im jungen Erwachsenenalter identifiziert werden. Die Problematik der Vergleichbarkeit bei unterschiedlicher Zusammensetzung der Deliktindizes wird ergänzend thematisiert. Die Ergebnisse werden den bisherigen Erkenntnissen der kriminologischen Längsschnittforschung gegenüber gestellt.

- 1 Einführung
- 2 Modelltypen
 - 2.1 Das klassische Wachstumsmodell
 - 2.2 Mischverteilungsmodelle und Verteilungsannahmen für Zähl­daten
- 3 Empirische Studie
 - 3.1 Erhebungsdesign
 - 3.2 Bisherige Verlaufsanalysen
 - 3.3 Fragestellungen
- 4 Mischverteilungsmodelle
 - 4.1 Modelle mit Jugenddelikten
 - 4.2 Modelle mit Jugend- Erwachsenen­delikten
- 5 Vergleich der Entwicklungsverläufe
- 6 Zusammenfassung

25. Workshop Angewandte Klassifikationsanalyse (AKA)

Arbeitstitel:

Ipsative und normative Messung beruflicher Interessen nach dem Holland-Modell

Christian Tarnai und Florian G. Hartmann
Universität der Bundeswehr München

Grundlage des Beitrags sind zwei Interessentests, die sich in ihrer Art der Erhebung unterscheiden:

Persönlichkeits- und Interessen-Test (PIT) von Mittenecker und Toman (1951, 1972) in einer revidierten Fassung nach Tarnai (1973): Die erfassten Interessenbereiche sind: Vorliebe für Land- oder Stadtleben, Handwerk, Wissenschaft, Verwaltung, Umgang mit Menschen in Geschäft und Wirtschaft, bildende Kunst, Sprache/Literatur, Musik, soziale Berufe. In den Interessen-Items werden in Form von Statements Tätigkeiten von zwei Interessenrichtungen gegenübergestellt. Die Probanden entscheiden sich für eine oder keine der beiden Tätigkeiten.

Allgemeiner Interessen-Struktur-Test (AIST) von Bergmann und Eder (1992). Die Entwicklung des AIST basiert auf dem Modell der sechs Interessenorientierungen von Holland (1973, 1997): Realistic, Investigative, Artistic, Social, Enterprising und Conventional. In 60 Items werden Tätigkeiten genannt, welche die Probanden auf einer fünfstufigen Skala nach der Stärke ihres Interesses einschätzen.

Das Holland-Modell ist die Vergleichsbasis der beiden Tests. Für den PIT werden die in den Statements gegenübergestellten Tätigkeiten von Experten den sechs Dimensionen des Holland-Modells zugeordnet.

Für den PIT sind die Summenscores der Dimensionen ipsative, für den AIST normative Messungen der Interessen. Entsprechendes gilt für die aus den Summenscores der beiden Tests abgeleiteten Drei-Buchstaben-Codes. Beim AIST kann die Bestimmung des Drei-Buchstaben-Codes bei gleichen Scores uneindeutig sein. Beim PIT besteht die Möglichkeit, den direkten Vergleich der Tätigkeiten zur Bestimmung der Rangreihe heranzuziehen. Darüber hinaus ergeben sich für PIT und AIST verschiedene Auswertungsmöglichkeiten, welche jeweils auf ihre Übereinstimmung mit dem Holland-Modell überprüft werden.

Die Bedeutung der Ergebnisse der beiden Tests wird im Rahmen der Untersuchung der Interessenähnlichkeit von Familien in Analogie zu Hartmann (2018) bestimmt. Die Passung zu gewählten Studienrichtungen und Berufen ist ein weiterer Aspekt des Vergleichs der Auswertungen beider Interessentests.

Klassifikation sozialer Erwünschtheit im Zeitverlauf

Lena M. Verneuer - RWTH Aachen - lverneuer@soziologie.rwth-aachen.de
Maximilian Wächter - Universität Bielefeld - maximilian.waechter@uni-bielefeld.de

Ausgangspunkt und Forschungsfragen

Sozial erwünschte Antworttendenzen sind als Problem in der sozialwissenschaftlichen Forschung bekannt, konsentiert und werden mithilfe direkter oder aber indirekter Messinstrumente zu antizipieren versucht. Indirekten Messinstrumenten (z.B. Randomized Response Techniques; vgl. Chaudhuri et al., 2013) wird dabei zugesprochen, ebendiese Antworttendenzen vermeiden zu können. Direkte Messinstrumente in Form von Skalen (z.B. Stöber, 1999) werden vor allem zur Kontrolle und nachträglichen Identifizierung (Eifler et al., 2015) von verzerrten Antwortmustern eingesetzt.

In der Erforschung von sozialer Erwünschtheit bestehen jedoch trotz der weiten Anerkennung und Berücksichtigung einige Leerstellen, die sich bei Durchsicht der entsprechenden (Methoden-)Literatur sowohl in der konzeptionellen Fundierung und als auch den empirischen Analysen finden lassen. Aktuelle Publikationen befassen sich erstens selten damit, wie genau soziale Erwünschtheit theoretisch (soziologisch) fundiert werden kann. Zweitens reproduzieren sie empirisch ausgemachte Einflussfaktoren (Bildung, Geschlecht, etc.) mit einer weitestgehend querschnittlichen Datengrundlage ohne nähere Analysen bezüglich der Entwicklung dieser Antwortmuster.

Unsere geplante Artikel für den Sammelband konzentriert sich primär auf die zweite Leerstelle und will einen Beitrag zur fundierten empirischen Analyse von sozialer Erwünschtheit leisten: Zum aktuellen Zeitpunkt gibt es keine empirische Untersuchung, welche diese Entwicklung von sozialer Erwünschtheit über einen längeren Zeitraum verfolgt. Unser Artikel soll einen ersten Ansatz liefern, um diese Lücke zu füllen. Weiterhin sollen die gefundenen Unterschiede in der Entwicklung sozialer Erwünschtheit für die verschiedenen Befragten klassifiziert und somit weiter analysiert werden. Auf diese Weise soll der Effekt sozialer Erwünschtheit in zukünftigen Studien, welche auf der Befragung junger Personen basieren, besser verständlich und abschätzbar werden.

Zentral sind für unseren Beitrag die folgende Fragestellungen: Wie entwickelt sich die Tendenz zu sozialer Erwünschtheit im Verlauf des Jugend- und jungen Erwachsenenalters? Und wie kann dieser Verlauf erklärt werden?

Analysestrategie und Datengrundlage

Zur Beantwortung dieser Fragen wird ein dreiteiliges Forschungsprogramm verfolgt: Erstens wird die Entwicklung der Tendenz zu einem sozial erwünschten Antwortverhalten vom Jugend- zum jungen Erwachsenenalter analysiert. Zweitens sollen die gefundenen Entwicklungsverläufe sozialer Erwünschtheit mittels sozio-demographischer Faktoren näher erklärt werden. Zu diesem Zweck werden latente Wachstumsmodelle mit multiplen Gruppenvergleichen geschätzt. Drittens soll die unbeobachtete Heterogenität

in den Verläufen sozialer Erwünschtheit klassifiziert werden. Dazu werden Growth Mixture Models geschätzt und die dadurch entstehenden Klassen weiter analysiert.

Die verwendeten Daten stammen aus dem DFG-Projekt *Kriminalität in der modernen Stadt* (Crime in the Modern City; CrimoC), welches von Jost Reinecke (Universität Bielefeld) und Klaus Boers (WWU Münster) seit 2000 durchgeführt wird (u.a. Boers, Reinecke, et al., 2014; Boers and Reinecke, 2019). Bei diesem Projekt handelt es sich um eine Längsschnittstudie, welche thematisch auf das Phänomen der Jugendkriminalität ausgerichtet ist.

Innerhalb des Crimoc-Projektes wird eine Kurzform der von Stöber (1999) publizierten SES-17-Skala zur Erfassung sozialer Erwünschtheit verwendet. Von den 17 Items, welche die Skala von Stöber bilden, wurden 6 in die Skala des Projektes übernommen (s. Tabelle 1). Zum einen enthält sie Aussagen, welche eine Form von Verhalten präsentieren, welche zwar normativ erwartbar ist, für die es jedoch wahrscheinlich ist, dass die Befragten sie nicht in jeder angebrachten Situation zeigen (Übertreibung positiver Qualitäten, SE1). Zum anderen enthält sie Items, welche Verhaltensweisen beschreiben, die normativ negativ besetzt sind, für die es jedoch wahrscheinlich ist, dass Befragte sie dennoch manchmal ausüben (Minimierung negativer Qualitäten, SE2). Die Skalierung der Skala ist fünfstufig angelegt mit 1"trifft gar nicht zu" bis 5"trifft völlig zu".

Table 1: Dimension, Itembezeichnung und wording der SES-17-Skala von Stöber (1999) im CrimoC-Projekt

| Dimension | Bezeichnung | Wording |
|---|--------------------|--|
| | x0001 | <i>Ich akzeptiere alle anderen Meinungen, auch wenn sie mit meiner eigenen nicht übereinstimmen.</i> |
| Übertreibung positiver Qualitäten [SE1] | x0003 | <i>Ich zögere niemals, jemandem in einer Notlage beizustehen.</i> |
| | x0004 | <i>Wenn ich etwas versprochen habe, halte ich es ohne Wenn und Aber.</i> |
| | x0002 (recoded) | <i>Ich habe schon einmal jemanden ausgenutzt oder übers Ohr gehauen.</i> |
| Minimierung negativer Qualitäten [SE2] | x0005 (recoded) | <i>Ich habe schon einmal geliehene Sachen nicht zurückgegeben.</i> |
| | x0006 (recoded) | <i>Manchmal helfe ich nur, weil ich eine Gegenleistung erwarte.</i> |

Die Skala zur Sozialen Erwünschtheit wurde nicht durchgängig erhoben – unsere Analysen beziehen sich auf die folgenden Ausschnitte der Paneldaten (s. Tabelle 2): Die direkte Messung mittels einer Skala erfolgte in den Jahren 2002, 2003, 2004 und 2005 sowie erneut in 2013, 2015 und 2017. Damit stehen sieben verschiedene Wellen über einen Zeitraum von 15 Jahren zur Verfügung. Vier der Messzeitpunkte befinden sich im Jugendalter der Befragten. Drei decken das junge Erwachsenenalter ab.

Table 2: Erhebungszeitpunkte der SES-17-Skala von Stöber (1999) im CrimoC-Projekt

| | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2011 | 2013 | 2015 | 2017 |
|--------------|------|------|------|------|------|------|------|------|------|------|------|------|
| Zeitpunkt | t1 | t2 | t3 | t4 | t5 | t6 | t7 | t8 | t9 | t10 | t11 | t12 |
| Wellenkürzel | a | b | c | d | e | f | g | h | i | j | k | l |
| Alter | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 22 | 24 | 26 | 28 |

Unser Programm für Rothenberge

In unserem Vortrag in Rothenberge möchten wir einerseits unser Beitragskonzept (u.a. Aufbau, Argumentation) vorstellen; andererseits wird ein Fokus auf unsere aktuellen Analysen gelegt. Im letzten Jahr haben wir bereits Analysen über 6 Zeitpunkte vorgestellt (Paneldatensatz bis 2015), hatten allerdings Schwierigkeiten mit der Modellierung der Messzeitpunkte 2013 und 2015: Da mit 2 Zeitpunkten für das junge Erwachsenenalter eine separate Entwicklungsmodellierung nicht möglich war, haben wir zunächst die zeitliche Lücke (2006 – 2011) unberücksichtigt lassen müssen und konnten keine Befunde generieren, die Aussagen zur Stabilität von sozial erwünschten Antworttendenzen im jungen Erwachsenenalter zuließen. Mit dem aktuellen Paneldatensatz des CrimoC-Projektes liegen nun auch für den zweiten Abschnitt drei Messzeitpunkte vor, sodass sich adäquatere Modellierungen umsetzen lassen. Die neuen Ergebnisse möchten wir gerne vor dem Hintergrund unseres geplanten Beitrages diskutieren.

References

- Boers, Klaus and Jost Reinecke, eds. (2019). *Delinquenz im Altersverlauf. Erkenntnisse der Längsschnittstudie 'Kriminalität in der modernen Stadt'*. Vol. 20. Kriminologie und Kriminalsoziologie. Münster, New York: Waxmann.
- Boers, Klaus, Jost Reinecke, Christina Bentrup, Andreas Daniel, Kristina-Maria Kanz, Philipp Schulte, Daniel Seddig, Maïke Theimann, Lena Verneuer, and Christian Walburg (2014). "Vom Jugend- zum frühen Erwachsenenalter: Delinquenzverläufe und Erklärungszusammenhänge in der Verlaufsstudie „Kriminalität in der modernen Stadt“". In: *Monatsschrift für Kriminologie und Strafrechtsreform* 97.3, pp. 183–202.
- Chaudhuri, Arijit and Tasos C. Christofides (2013). *Indirect Questioning in Sample Surveys*. Heidelberg: Springer.
- Eifler, Stefanie, Daniela Pollich, and Jost Reinecke (2015). "Die Identifikation von sozialer Erwünschtheit bei der Anwendung von Vignetten in Mischverteilungsmodellen". In: *Empirische Forschung über Kriminalität*. Ed. by Stefanie Eifler and Daniela Pollich. SpringerLink : Bücher. Wiesbaden: Springer VS, pp. 217–247.
- Stöber, Joachim (1999). "Die Soziale-Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität." In: *Diagnostica* 45.4, pp. 173–177.

Rainer Watermann: Latente Markov Modelle: Erweiterungen und Anwendungen auf die Entwicklung der intrinsischen Lernmotivation am Übergang von der Grundschule in die weiterführende Schule

1. Einordnung des Beitrags in Verfahren der Klassifikationsanalyse

Latente Markov-Modelle (LMM; Langeheine, 1994; Langeheine & van de Pol, 2002) ermöglichen die longitudinale Analyse kategorialer latenter Variablen (d.h. latente Klassen, Typen oder Zustände). Viele Anwendungsmöglichkeiten in den Sozialwissenschaften sind denkbar. Kaplan beispielsweise untersuchte die Veränderung und Stabilität in der Lesekompetenz bei Schülerinnen und Schülern (Kaplan, 2008). Nylund (2007) hat sich in ihrer Dissertation damit befasst, welche Typen von Jugendlichen mit unterschiedlichen Gewalterfahrungen durch peers identifizierbar sind und wie sich die Zugehörigkeit zu diesen Typen über die Zeit verändert.

LMM basieren auf zwei unterschiedlichen Modellklassen: auf der latenten Klassenanalyse (LCA, McCutcheon, 1987) und einem autoregressiven Modell. Die LCA dient als Messmodell zur Identifikation der latenten Klassen (d.h. Gruppen, Typen oder Zuständen). Das autoregressive Modell beschreibt das Ausmaß des Wechsels zwischen den latenten Klassen bzw. der Stabilität über die Zeit. Latente Markov-Modelle definieren somit simultan die latente Klassenvariable über die Zeit (Messmodell) wie auch die Stabilität und Veränderung in der latenten Klassenvariable (Strukturmodell).

In den letzten Jahren wurden Erweiterungen von LMM diskutiert und umgesetzt. Diese betreffen die Berücksichtigung von manifesten Kovariaten, latenten Kovariaten zur Analyse unbeobachteter Heterogenität (z.B. mover vs. stayer), autoregressiven Effekten höherer Ordnung sowie Beziehungen zwischen den Entwicklungsmustern und distalen Outcomevariablen (vgl. zusammenfassend Nylund, 2007; Kaplan, 2008).

2. Ziel des Beitrags, Theorie und Fragestellung

Im vorliegenden Beitrag sollen latente Markov-Modelle sowie einige der genannten Erweiterungen auf die Entwicklung der intrinsischen Lernmotivation (z.B. Freude, Neugier beim Lernen) im frühen Jugendalter angewendet werden. In den meisten Arbeiten werden Merkmale wie die intrinsische Lernmotivation als kontinuierliche Merkmale aufgefasst. Entsprechend kommen in der Regel longitudinale Modelle für kontinuierliche Variablen (z.B. latente Wachstumskurvenmodelle; Meredith & Tisak, 1990) oder autoregressive Modelle zur Anwendung (Watermann et al, 2007). Ich werde in diesem Beitrag

die intrinsische Lernmotivation als eine kategoriale latente Variable behandeln. Es ist davon auszugehen, dass die latenten Klassen geordnet sind, d.h. sich ordinal auf dem latenten Kontinuum anordnen. Die folgenden drei inhaltlichen Hypothesen stehen im Vordergrund:

- 1) Zum einen sollte der Anteil von Personen mit hoch oder höher ausgeprägter Lernmotivation im Zeitverlauf abnehmen, insbesondere nach dem Wechsel von der Grund- in die weiterführende Schule. Eccles et al. (1993) sehen die Ursache hierfür in der fehlenden Passung zwischen den Lehr-/Lernumwelten in der Sekundarstufe I und den psychologischen Grundbedürfnissen der Lernenden nach Autonomie und sozialer Wertschätzung (Developmental-Stage-Environment-Fit-Effekt, DSEFE). Zudem sei der deutliche Anstieg des kompetitiven Klassenklimas oftmals inkompatibel zum individuellen Bedürfnis nach Kompetenzerleben.
- 2) Zum anderen sollte diese Entwicklung von einem weiteren Mechanismus überlagert werden: In Anlehnung an den Big-Fish-Little-Pond-Effekt (BFLPE; Marsh, 1987) sollte mit dem Wechsel in eine leistungsstärkere Lernumgebung (Gymnasium) ebenfalls eine Reduzierung des Anteils von Personen mit hoher oder höherer Lernmotivation verbunden sein.
- 3) Personen, die im Zeitverlauf eine hohe oder höhere intrinsische Lernmotivation aufrecht erhalten, sollten bessere Schulleistungen am Ende des Untersuchungszeitraums (Klasse 7) aufweisen.

3. Datengrundlage

Die Daten stammen aus der Übergangsstudie (Maaz et al., 2010), einem Kooperationsprojekt zwischen der FU Berlin und dem MPIB Berlin. Für N=1769 Schülerinnen und Schüler liegen zu vier Messzeitpunkten (Ende Klasse 4, Mitte Klasse 5, Mitte Klasse 6, Mitte Klasse 7) Messungen zu deren intrinsischer Lernmotivation vor. Diese wurde anhand von vier Items erfasst (Lernen macht mir Spaß; Ich bin immer ganz neugierig, wenn ich etwas Neues lernen kann; Neues zu lernen, ist oft aufregend; Manchmal bin ich richtig gierig, weiter lernen zu können). Die Antwortstufen waren ordinal mit den Ausprägungen „1=stimmt gar nicht“, „2=stimmt eher nicht“, „3=stimmt eher“ und „4=stimmt genau“. Die Items wurden dichotomisiert, so dass der Wert 1 für fehlende Zustimmung (stimmt gar nicht/stimmt eher nicht) und der Wert 2 für Zustimmung (stimmt eher/stimmt genau) zu einem Item steht. Tabelle 1 gibt die relativen Häufigkeiten nach Messzeitpunkt wieder. Als zeitinvariante Kovariate liegt die besuchte Schulform in der

Sekundarstufe I vor. Auch diese Variable wurde dichotomisiert (0=andere Schulform, 1=Gymnasium). Als weitere zeitinvariante Kovariate wurden die kognitiven Grundfähigkeiten berücksichtigt. Das distale Outcome-Maß bildete die Durchschnittsnote aus Mathematik, Deutsch und erster Fremdsprache im Halbjahreszeugnis der 7. Klasse.

Tabelle 1. Relative Häufigkeiten der Zustimmung zu den Items der intrinsischen Lernmotivation nach Messzeitpunkt

| | T1 | T2 | T3 | T4 |
|--|------|------|------|------|
| | K. 4 | K. 5 | K. 6 | K. 7 |
| Lernen macht mit Spaß | .726 | .648 | .565 | .469 |
| Ich bin immer ganz neugierig, wenn ich etwas Neues lernen kann | .820 | .744 | .584 | .518 |
| Neues zu lernen, ist oft aufregend | .806 | .726 | .600 | .552 |
| Manchmal bin ich richtig gierig, weiter lernen zu können | .503 | .463 | .359 | .303 |

4. Vorgehen und Ergebnisse

Sämtliche Analysen wurden mit Mplus (Version 8.3) durchgeführt. Das Vorgehen orientierte sich an den Empfehlungen von Nylund (2007).

In einem ersten Schritt wurden getrennt für jeden Messzeitpunkt latente Klassenmodelle ohne Kovariaten spezifiziert. Hierbei waren jeweils ein 2-Klassen- bzw. 3-Klassen-Modell am besten mit den Daten vereinbar. Aus inhaltlichen Gründen fiel die Entscheidung zugunsten des 2-Klassen-Modells aus. Inhaltlich ergab sich eine Klasse mit niedriger Lernmotivation (hohe Antwortwahrscheinlichkeiten für die fehlende Zustimmung bei allen Items) und eine Klasse mit hoher Lernmotivation (hohe Antwortwahrscheinlichkeiten für die Zustimmung bei allen Items). An den Klassengrößen war erkennbar, dass der Anteil von Schülerinnen und Schülern mit einer hohen intrinsischen Lernmotivation erwartungsgemäß im Zeitverlauf abgenommen hat.

In einem zweiten Schritt wurden Wechsel zwischen den Klassen deskriptiv untersucht. Personen wurden jeweils der Klasse zugeordnet, für die sie die höchste Zugehörigkeitswahrscheinlichkeit besaßen. Dabei zeigten sich die höchsten und auch konstan-

ten Wechsleranteile zwischen den Messzeitpunkten t2 (Klasse 5) und t3 (Klasse 6) bzw. t3 (Klasse 6) und t4 (Klasse 7), wobei die Anteile von Personen, die von der Klasse mit hoher Lernmotivation in die Klasse mit niedriger Lernmotivation gewechselt haben, größer waren als umgekehrt.

In einem dritten Schritt wurde das latente Markov-Modell mit zeitkonstanten Antwortwahrscheinlichkeiten spezifiziert und mit einem Modell mit frei geschätzten Antwortwahrscheinlichkeiten verglichen. Das Modell mit zeitkonstanten Antwortwahrscheinlichkeiten ließ sich besser anpassen als das Modell mit den frei geschätzten. Alle weiteren Modelle wurden mit dieser Restriktion durchgeführt.

In einem fünften Schritt wurden Modelle gerechnet, in denen den Übergangswahrscheinlichkeiten jeweils Restriktionen aufgegeben wurden. Ein Modell mit zeitkonstanten Übergangswahrscheinlichkeiten (stationäres LMM) wies schlechtere Fitmaße auf als das Modell mit unrestringierten, d.h. frei geschätzten, Übergangswahrscheinlichkeiten. Dies war aufgrund der deskriptiven Analysen (aus Schritt 2) zu erwarten gewesen. Ein Modell mit zeitkonstanten Übergangswahrscheinlichkeiten nur zwischen den Messzeitpunkten t2 (Klasse 5) und t3 (Klasse 6) bzw. t3 (Klasse 6) und t4 (Klasse 7) erwies sich dagegen als das beste der drei Modelle. Während zwischen t1 (Klasse 4) und t2 (Klasse 5) die Wechsel in beide Richtungen ausgewogen waren, fielen zwischen t2 (Klasse 5) und t3 (Klasse 6) sowie zwischen t3 (Klasse 6) und t4 (Klasse 7) die Übergangswahrscheinlichkeiten von high zu low deutlich höher aus als von low zu high (vgl. Tab. 2).

Tabelle 2. Übergangswahrscheinlichkeiten im LMM mit partiell zeitkonstanten Übergangswahrscheinlichkeiten (zeitkonstante Übergangswahrscheinlichkeiten zwischen t2 und t3 bzw. t3 und t4)

| | | T2 | |
|----|------|-------|-------|
| | | High | Low |
| T1 | High | 0.765 | 0.235 |
| | Low | 0.235 | 0.765 |

| | | T3 | |
|----|------|------|------|
| | | High | Low |
| T2 | High | .689 | .311 |
| | Low | .109 | .891 |

| | | T4 | |
|----|------|------|------|
| | | High | Low |
| T3 | High | .689 | .311 |
| | Low | .109 | .891 |

In einem sechsten Schritt wurden die messzeitpunktspezifischen Effekte der Kovariate Schulform geschätzt. Sobald man Kovariaten im Modell mit berücksichtigt, macht die Annahme zeitkonstanter Übergangswahrscheinlichkeiten keinen Sinn mehr (Nylund, 2007). Deshalb wurden die Restriktionen in den Übergangswahrscheinlichkeiten aufgehoben und die messzeitpunktspezifischen Effekte der Kovariaten auf die Klassenwahrscheinlichkeiten geschätzt. Signifikante Effekte deuteten auf Unterschiede in den Übergangswahrscheinlichkeiten nach Schulform hin. Erwartungsgemäß waren zu t1 (Klasse 4) zukünftige Gymnasiasten, häufiger in der Klasse mit hoher Lernmotivation vertreten als Personen, die nicht-gymnasiale Schulkarrieren einschlugen. Zu t2 (Klasse 5) drehte sich der logistische Regressionskoeffizient jedoch um: Gymnasiasten wechselten häufiger in die Klasse mit niedriger intrinsischer Lernmotivation, was der Annahme eines BFLPE am Übergang entspricht. Zu den späteren Zeitpunkten gab es keine signifikanten Effekte der Schulform.

In einem siebten Schritt wurde die Robustheit der Effekte auch unter Kontrolle der kognitiven Grundfähigkeiten geprüft. Auch diese gingen als zeitinvariante Kovariate in das Modell ein. Die Effekte blieben durchgehend unverändert.

Abschließend sollen die Schulleistungen zu t4 (Klasse 7) als distale Outcome-Variable in das Modell eingeführt werden. Dieses Modell muss noch spezifiziert werden.

5. Diskussion

Im vorliegenden Beitrag wurde die Entwicklung der intrinsischen Lernmotivation von Klasse 4 bis Klasse 7 mithilfe von latenten Markov-Modellen untersucht. Auf Basis eines Modells mit zeitkonstanten Messmodellen ließ sich zeigen, dass die Abnahme der intrinsischen Lernmotivation nicht unmittelbar nach dem Übergang in die weiterführende Schule erfolgt, sondern erst zwischen Klasse 5 und 6 einsetzt. Somit konnte der DSEFE bestätigt werden. Die Ergebnisse sprechen allerdings gegen einen linearen Trend. In einem Modell mit der Schulform als Kovariate wurde deutlich, dass der institutionelle BFLPE früher einsetzt als der DSEFE. Die mit dem Wechsel in eine leistungsstärkere Klassengemeinschaft einsetzenden sozialen Aufwärtsvergleiche haben psychologische Kosten und unterminieren die intrinsische Lernmotivation. Auf der anderen Seite ist davon auszugehen, dass mit dem Wechsel in eine leistungsschwächere Klassengemeinschaft auch die Gelegenheiten für soziale Abwärtsvergleiche zunahmten, was sich günstig auf die intrinsische Lernmotivation ausgewirkt haben sollte. Ich überlege noch, ob ich diesen Effekt in einem Mehr-Gruppen-LMM zeige. Die Ergebnisse und Diskussion der dritten Hypothese stehen noch aus.

6. Literatur

Eccles, J., Midgley, C., Wigfield, A., Buchanan, C.M., Reuman, D., Flanagan, C., Mac Iver, D. (1993). Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, 48, 90-101.

Kaplan, D. (2008). An overview of Markov Chain Methods for the study of stage-sequential developmental processes. *Developmental Psychology*, 44, 457-467.

Langeheine, R. (1994). Latent variables Markov models. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis* (pp. 373-395). Newbury Park, CA: Sage.

Langeheine, R., & Van de Pol, F. (2002). Latent Markov chains. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 304-341). Cambridge, United Kingdom: Cambridge University Press.

Maaz, K., Baumert, J., Gresch C. & McElvany, N. (2010) (Hrsg.). *Der Übergang von der Grundschule in die weiterführende Schule - Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*. Bonn, Berlin: BMBF.

Marsh, H. W. (1987). The big-fish-little-pond-effect on academic self-concept. *Journal of Educational Psychology*, 79, 280-295.

McCutcheon, A. L. (1987). *Latent class analysis*. Nuwbury Park, London, New Delhi: Sage.

Meredith, W. & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55, 107–122.

Nylund, K. L. (2007). *Latent transition analysis: Modeling extensions and an application to peer victimization*. Los Angeles: University of California.

Watermann, R., Klingebiel, F. & Kurtz, T. (2010). Die motivationale Bewältigung des Grundschulübergangs aus Schüler- und Elternsicht. In K. Maaz, J. Baumert, C., Gresch & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule - Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*, (S. 357-384). Bonn, Berlin: BMBF.