
Empirische Methoden zur Analyse kontextueller Faktoren in der Bildungsforschung

Michael Windzio und Janna Teltemann

1 Einleitung: Kontexte in der Bildungsforschung

In modernen Gesellschaften beginnt der Bildungsprozess im unmittelbaren Familienzusammenhang und ist darüber hinaus institutionell organisiert. Kinder und Heranwachsende lernen in vorschulischen Einrichtungen, in der Primar- und Sekundarstufe sowie in der berufsfachlichen Ausbildung und an Universitäten. Weil sie in diesen Institutionen Einflüssen der jeweiligen sozialen Kontexte ausgesetzt sind, muss die empirische Forschung über Methoden verfügen, die für die Analyse von Kontexteffekten geeignet sind. Diese Methoden sind relativ vielfältig und ihre Angemessenheit hängt von der jeweils verfolgten Fragestellung sowie von der Analyseebene ab, auf der die Kontexte des Bildungserwerbs verortet werden.

Auf der Ebene der Gesamtgesellschaft stellen die *Bildungsinstitutionen* Kontexte bereit, die den Bildungsverlauf des Individuums vorstrukturieren. Aus der Lebensverlaufsforschung ist bekannt, dass ein Wandel sowohl der gesellschaftlichen Institutionen auf der Makroebene als auch der Opportunitäten auf der Meso-Ebene – etwa in Form des Angebots von Bildungsgelegenheiten (z. B. mehr Gymnasien, mehr Schulhäuser in Nähe des Elternhauses, etc.) – zu einer Ausdifferenzierung der Lebensläufe zwischen Kohorten führen kann. Ein prominentes Beispiel ist die Bildungsexpansion, die die Lebensläufe jener Geburtskohorten nachhaltig beeinflusste, die sie in den sensiblen Phasen ihres Lebensverlaufs erlebten. Außerdem sind in bestimmten institutionellen Regimen (vgl. Allmendinger und Hinz 1998)

M. Windzio (✉) · J. Teltemann
Institut für empirische und angewandte Soziologie, Universität Bremen,
28359 Bremen, Deutschland
E-Mail: mwindzio@empas.uni-bremen.de

J. Teltemann
E-Mail: janna@uni-bremen.de

R. Becker, A. Schulze (Hrsg.), *Bildungskontexte*,
DOI 10.1007/978-3-531-18985-7_2, © Springer Fachmedien Wiesbaden 2013

Institutionen des Bildungssystems eng mit dem Arbeitsmarkt verzahnt, wodurch die Übergänge erleichtert werden. Allerdings sind auch die Übergangsprozesse von kurz- bis mittelfristigen Schwankungen der Arbeitsmarktlage beeinflusst. Mit einer Kombination aus *Ereignis-* und *Kontextanalyse* kann untersucht werden, ob z. B. konjunkturelle Schwankungen und ungünstige betriebsökologische Entwicklungen (vgl. Windzio 2001) in Verbindung mit hohen Arbeitslosenquoten eher ungünstige Kontextbedingungen für diesen Übergang darstellen. Aber auch Übergänge innerhalb des Bildungssystems – etwa Übergänge in die Weiterbildung (vgl. Becker 2000; Wingens et al. 2000) – können ereignisanalytisch in Abhängigkeit von sich wandelnden Kontextbedingungen untersucht werden.

Häufig wird in der Bildungsforschung von *institutionellen Effekten* gesprochen, auch wenn eigentlich die konkreten *Organisationen* – nämlich die Schulen, Universitäten oder Kindergärten – gemeint sind, in denen sich der Prozess des Bildungserwerbs im konkreten Alltag der betroffenen Akteure vollzieht. Eine auf die Ebene der Organisation bezogene Analyseverfahren von Kontexteffekten auf den Bildungserwerb ist die *Mehrebenenanalyse*. Mit diesem Verfahren lassen sich Eigenschaften der Organisation, die alle in eine jeweilige Organisation eingebetteten Akteure potentiell beeinflussen, modellieren. Innerhalb der Organisation lassen sich weitere Kontextebenen ausdifferenzieren, etwa die Schulklasse oder auch die Person selbst, wenn im Falle von Längsschnittdaten Messungen innerhalb der Person über die Zeit untersucht werden.

Weniger prominent, aber aktuell von zunehmender Bedeutung, sind *Netzwerkmodelle*, bei denen die Fragestellung verfolgt wird, welche Faktoren sich auf die Chance eines Kontakts in einem Netzwerk auswirken. Insbesondere Ansätze, die auf sozialkapitaltheoretischen Überlegungen basieren, nehmen implizit Effekte der sozialen Einbettung in Netzwerke an, die aber häufig nicht mit Methoden der Netzwerkanalyse untersucht werden. Die sogenannten *p*-Modelle* (auch als *Exponential Random Graph Models* bezeichnet) lösen dabei ein ähnliches Problem, für das auch die Mehrebenenanalyse entwickelt wurde, nämlich die adäquate Behandlung des Problems der fehlenden statistischen Unabhängigkeit der Beobachtungen innerhalb der jeweiligen Kontexte. Während man in der Mehrebenenanalyse klar abgegrenzte Kontexte definieren kann – nämlich beispielsweise Schulen oder Schulklassen – ist das in sozialen Netzwerken in der Regel nicht der Fall. In der Schulforschung werden häufig soziale Netzwerke in Schulklassen untersucht (vgl. Baerveldt et al. 2004; Stubbe et al. 2007; Winkler et al. 2011), die somit relativ klar abgrenzbar sind. Jedoch lassen sich wiederum *innerhalb* dieser Schulklassen-Netzwerke eindeutig abgegrenzte Kontexte nur schwer feststellen. Dies ist der Grund, warum die Chancen auf eine Bindung zwischen zwei Knoten mit der Methode der *p*-Modelle* untersucht werden.

In der Ökonometrie entwickelten sich in den letzten Jahren neuere Methoden zum Testen von Kausalhypothesen. Auch hier geht es häufig – in Abhängigkeit vom Untersuchungsdesign – um bestimmte Kontexte, innerhalb oder zwischen denen die Akteure randomisiert entweder der Treatment- oder der Kontrollgruppe zugewiesen werden. Aber auch Instrument-Variablen-Modelle und Verfahren zur Kontrolle von Selektionsprozessen werden häufig zur Analyse von Kontexteffekten angewandt (vgl. Angrist und Pischke 2009).

In den folgenden Abschnitten werden die genannten Methoden und statistischen Verfahren dargestellt und diskutiert. Besprochen werden die Mehrebenenanalyse für hierarchische und nicht-hierarchische Kontexte, das Problem der Analyse von Peer-Effekten, die Ereignisanalyse der Effekte zeitveränderlicher institutioneller Rahmenbedingungen sowie Exponential-Random-Graph-Modelle für Kontakte in Netzwerken. Abschließend wird das Problem der Kausalattribution am Beispiel des Effektes der Klassengröße auf die Schulleistung diskutiert und die Möglichkeiten und Grenzen von Instrumentvariablen-Modellen aufgezeigt.

2 Mehrebenenanalyse

Die Mehrebenenanalyse ist ein spezielles regressionsanalytisches Verfahren, das zur Anwendung kommen muss, wenn Daten in einer *hierarchischen Struktur* vorliegen oder die Untersuchungseinheiten in anderer Weise innerhalb der Kontexte voneinander abhängig sind. Das Verfahren ist zunächst also nichts anderes als eine Generalisierung von Regressionsmodellen. Diese Regressionsmodelle können entweder lineare Modelle für stetig normalverteilte abhängige Variablen sein; sie können aber auch aus der großen Familie der generalisierten linearen Modelle stammen (vgl. Dunteman und Ho 2006; Hox 2010, Kap. 7).

2.1 Hierarchische Kontexte

Liegen statistische Abhängigkeiten der Beobachtungen aufgrund einer hierarchischen Datenstruktur vor, muss die Mehrebenenanalyse angewendet werden, weil eine wichtige Modellannahme der klassischen Regression, die Unabhängigkeit der Beobachtungen, nicht erfüllt ist. Nehmen wir das Beispiel der Schulforschung: Untersucht man den Leistungserwerb von Schulkindern, wird man feststellen, dass die abhängige Variable (z. B. Schulnote oder PISA-Testwert) zu einem erheblichen Teil von individuellen Merkmalen der Schülerinnen und Schüler bestimmt ist. Allerdings wirken zusätzlich zu diesen Individualmerkmalen auch Kontextmerk-

male der Schule oder der Schulklassen auf die Schüler ein (vgl. Stanat 2006; Pong und Hao 2007). Beispielsweise sind alle Schüler einer Schulklasse gemeinsam dem Kontexteinfluss des jeweiligen Klassenlehrers ausgesetzt. Ein Klassenlehrer könnte bei seiner Beurteilung auf Stereotype zurückgreifen, die Kinder mit oder ohne Migrationshintergrund oder Kinder aus bildungsfernen oder akademischen Elternhäusern benachteiligen oder bevorzugen. Manche Lehrkräfte lösen bei bestimmten Gruppen möglicherweise in höherem Maße eine Bedrohung durch Stereotype (*stereotype threat*) aus als andere (vgl. Schmader et al. 2008). Liegen derartige Kontexteffekte vor, ist die Annahme der statistischen Unabhängigkeit der Untersuchungseinheiten, also der Schüler, verletzt. Je stärker die Kontexteffekte wirken, desto abhängiger sind die Beobachtungen innerhalb eines Kontexts voneinander, und desto gravierender ist auch die Unterschätzung der Standardfehler durch eine konventionelle Regression. Ist ein Kontexteffekt extrem stark ausgeprägt, dann ist die abhängige Variable bzw. deren Varianz in hohem Maße durch diesen Kontexteffekt bestimmt. Ist der Kontexteffekt nur sehr schwach ausgeprägt, ist die abhängige Variable fast ausschließlich durch die individuellen Merkmale der Schüler determiniert. In der Regel stellen wir bei Mehrebenenanalysen in der Schulforschung fest, dass die Kontexteffekte weder sehr gering noch übermäßig stark sind, sondern sich ca. 20 bis 40 % der Varianz durch Merkmale der Schulkontexte erklären lassen. Folglich besteht der erste Schritt der Mehrebenenanalyse darin, den Anteil der durch die jeweiligen Kontexte zu erklärenden Varianz zu quantifizieren. Je nachdem wie hoch der Anteil der durch die Kontextebene erklärbaren Varianz ist, würde man die Standardfehler der Regressionskoeffizienten unterschätzen, wenn man ein Einebenen-Modell – beispielsweise eine einfache OLS-Regression – anwenden würde. Man kann auf Basis der Intraklassenkorrelation die effektive Stichprobengröße berechnen, auf der bei gegebener Intraklassenkorrelation die Schätzung der Standardfehler basiert (Hox 2010, S. 241).

Vereinfacht dargestellt lässt sich das Grundprinzip der Mehrebenenanalyse in wenigen Gleichungen ausdrücken. Auf der Ebene der Individuen, z. B. der Schülerinnen und Schüler wird der Wert einer abhängigen Variablen durch individuelle Merkmale und deren Regressionsgewichte vorhergesagt:

$$y_{ij} = \beta_{0j} + \beta' \mathbf{x}_{ij} + e_{ij}$$

Der Term e_{ij} bezeichnet die Fehlerkomponente eines Schülers i in der Schulklasse j . Dieser Fehler auf der Individualebene variiert auch zwischen den Schülerinnen und Schülern innerhalb einer Klasse. Besonders an der Mehrebenenanalyse ist nun ein weiterer Fehlerterm u_{0j} , der nur zwischen den j Kontexten (hier also den Klassen), nicht aber zwischen Subjekten innerhalb der Kontexte variiert. Im Gegensatz

zur klassischen Regressionsgleichung für eine Ebene, die nur eine Regressionskonstante (Interzept) kennt, erlaubt die Mehrebenenanalyse, dass ebenso viele unterschiedliche Regressionskonstanten geschätzt werden, wie Kontexte vorliegen. Die Regressionskonstanten sind also *kontextspezifisch*. Die kontextspezifische Regressionskonstante ergibt sich aus der durchschnittlichen Regressionskonstanten über alle j Kontexte β_0 plus dem kontextspezifischen Abweichungswert u_{0j} . Der Term u_{0j} wird als *Zufallseffekt*, $\beta_0 + u_{0j}$ als *random intercept* bezeichnet. Die Abweichung eines spezifischen Interzepts vom Mittelwert aller Interzepts kann allein durch den Zufallseffekt u_{0j} bedingt sein und damit nur von unbeobachteten Kontexteinflüssen abhängen. Sie kann aber auch durch manifeste Kontextmerkmale \mathbf{Z}_j , die in den Daten gemessen wurden, vorhergesagt werden, wobei γ' das jeweilige Regressionsgewicht der Kontextmerkmale darstellt.

$$\beta_{0j} = \beta_0 + \gamma' \mathbf{Z}_j + u_{0j}$$

$$u_{0j} \approx N(0, \sigma_u^2)$$

Bei der Schätzung wird für u_{0j} als Zufallsverteilung eine näherungsweise Normalverteilung angenommen, deren Varianz σ^2 empirisch ermittelt wird. Wenn keine unabhängigen Variablen kontrolliert sind („leeres Modell“), resultieren hohe Werte von y aus Kontexten, die ein hohes mittleres Niveau der abhängigen Variablen aufweisen. Sind relevante Prädiktoren sowohl auf der Individualebene (\mathbf{x}_{ij}) als auch auf der Kontextebene (\mathbf{Z}_j , z. B. Klassengröße oder Schulmerkmale) kontrolliert, deuten hohe Werte von u_{0j} darauf hin, dass noch immer in den jeweiligen Kontexten aufgrund von unbeobachteten Eigenschaften der Kontexte hohe Werte von y festzustellen sind. Darum wird σ^2 von u_{0j} auch als *Residualvarianz* auf der Kontextebene bezeichnet. Somit lässt sich für die Regressionskonstanten eine Gleichung für die Kontextebene spezifizieren, bei der die Lage der jeweiligen Konstanten β_{0j} aus einer Regression $\gamma' \mathbf{Z}_j$ inklusive des Fehlerterms u_{0j} der Kontextebene resultiert. Ist die geschätzte Varianz von u_{0j} hoch, dann unterscheiden sich die Kontexte hinsichtlich y_{ij} auch sehr deutlich, d. h. die Kontexte haben einen relativ starken Einfluss. Ist die Varianz dagegen klein und insignifikant, liegen keine nennenswerten Kontexteinflüsse vor und man könnte auch eine konventionelle Ein-Ebenen-Regression schätzen. Das Maß der Intraklassenkorrelation zeigt an, wie hoch der Varianzanteil ist, der in einem jeweiligen Modell durch Merkmale der Kontextebene erklärt werden kann. Zur Berechnung wird die Residualvarianz der Kontextebene durch die Summe aus dieser und der Residualvarianz der Individualebene, also der Varianz der Werte von e_{ij} dividiert.

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Mit dem Verfahren der Mehrebenenanalyse wird also mindestens eine zusätzliche Fehlerkomponente u_{0j} auf der Ebene der Kontexte eingeführt. Auf diese Weise wird auch die durch die Kontexte bedingte Fehlervarianz vollständig in das Modell eingeführt und die Standardfehler korrekt geschätzt. Allerdings ist die Vermeidung des Ärgernisses einer möglichen Unterschätzung der Standardfehler nicht das einzige Ziel der Mehrebenenanalyse. Das Verfahren ist auch dazu geeignet, substantielle Fragestellungen zu beantworten, die sich nur aus der Perspektive dieser hierarchischen bzw. „geclusterten“ Daten ergeben. Man kann untersuchen, wie sich bestimmte Merkmale der Kontexte auf den Bildungserwerb und den Bildungsvorlauf der Individuen auswirken. Es ist auch möglich, durch einen weiteren Fehlerterm nicht nur die Interzepts, sondern auch die Steigungskoeffizienten spezifischer unabhängiger Variablen kontextspezifisch zu modellieren, indem wiederum auf den Mittelwert der Steigungen (*slopes*) einer unabhängigen Variablen x_1 ein kontextspezifischer Abweichungswert u_{1j} aufaddiert wird. Dies wird als *random slopes* bezeichnet, d. h. die Effektstärke hängt vom Kontext ab

$$\beta_{1j} = \beta_1 + u_{1j}$$

Man erhält ebenso viele Steigungskoeffizienten der Variable x_1 wie Kontexte vorliegen und erhält so eine Mehrebenen-Gleichung mit drei Fehlertermen:

$$y_{ij} = \beta_0 + \gamma' \mathbf{Z}_j + u_{0j} + (\beta_1 + u_{1j})x_{1ij} + \beta' \mathbf{x}_{ij} + e_{ij}$$

Man kann zudem untersuchen, wie die Kontexte Einflüsse der Individualebene auf Bildungserwerb und -verlauf moderieren, was als *cross-level-Interaktion* bezeichnet wird. Ist eine *cross-level-Interaktion* signifikant, wird zumindest ein Teil der durch unbeobachtete Kontextfaktoren bedingten Varianz der Steigungen von x_1 durch diese Interaktion erklärt. Es wird, mit anderen Worten, zumindest ein Teil der *nicht erklärten* Varianz von u_{1j} in *erklärte* Varianz überführt.

2.2 Nicht hierarchische Kontexte

Die obige Darstellung der Mehrebenenanalyse folgte dem klassischen Fall einer eindeutig hierarchischen Gliederung der Kontexte. Eindeutig hierarchisch sind

z. B. Daten der Schulforschung, bei denen individuelle Merkmale der Schülerinnen und Schüler die abhängige Variable darstellen (Ebene 1), diese Schüler in Klassen gruppiert sind (Ebene 2) und diese Klassen wiederum unterschiedlichen Schulen angehören (Ebene 3). In der empirischen Anwendung ergeben sich jedoch häufig komplexere Designs, bei denen eine hierarchische Anordnung der Kontexte nicht mehr möglich ist, wenngleich das Problem der statistischen Abhängigkeit der Beobachtungen innerhalb der Kontexte fortbesteht. So kann etwa die Schulleistung sowohl vom Kontext der *Schule* als auch vom Kontext der *Nachbarschaft*, in dem die Schülerinnen und Schüler wohnen, abhängig sein (vgl. Pong und Hao 2007), wobei der Einzugsbereich einer Schule unterschiedliche Nachbarschaften umfasst und Schüler einer Nachbarschaft in unterschiedlichen Schulbezirken zur Schule gehen. Derartige Datensituationen werden als *kreuzklassifiziert* (*cross-classified*) bezeichnet (Hox 2010, S. 171). Ein Kontext ist anderen nicht hierarchisch über- oder untergeordnet, sondern innerhalb eines Kontexts in der einen Dimension (z. B. Schule) überlappen sich mindestens zwei Kontextdimensionen einer anderen Dimension (z. B. mehrere Nachbarschaften, aus denen die Schule Jugendliche rekrutiert). In der Bildungsforschung sind diese Modelle in jüngerer Zeit durch international vergleichende Analysen der Kompetenz von Immigranten prominent geworden (vgl. van Tubergen und Kalmijn 2005; Levels und Dronkers 2008). Kreuzklassifizierte Mehrebenenmodelle stellen hier eine sehr unmittelbare methodische Umsetzung der wichtigen theoretischen Einsicht dar, dass der Integrationsprozess von Immigranten von drei unterschiedlichen Kontextdimensionen abhängig ist, was sich bereits in den frühen 1990er Jahren im Konzept der *modes of incorporation* andeutete (Portes und Zhou 1993, S. 84): erstens wirken Merkmale des *Herkunftskontexts* auf die Integrations- und Assimilationschancen, indem bspw. bestimmte kulturelle Ressourcen asiatischer Immigranten – wie etwa Konfuzianismus, Buddhismus oder Taoismus – eine strenge Arbeitsethik und Bedürfnisaufschub bewirken und damit dem Bildungserwerb zu Gute kommen (Zhou 1997, S. 994). Merkmale des *Aufnahmekontexts* bedingen zweitens z. B. durch die Art des Schulsystems oder der Arbeitsmarktlage die Integrationschancen. Schließlich resultieren *Community-Effekte* (auch: „Setting-Effekte“) aus der Interaktion von Aufnahme- und Herkunftskontext, indem z. B. die linguistische Nähe oder Distanz der Herkunftssprache zur Sprache des Aufnahmelandes die Integration beeinflusst (vgl. van Tubergen und Kalmijn 2005), oder indem die Integrationschancen von der relativen Gruppengröße der Immigranten im jeweiligen Aufnahmekontext abhängen (vgl. Esser 2008). Somit steht eine bestimmte Herkunftsgruppe als spezifische Community den Kontextbedingungen ihres Aufnahmelandes gegenüber. Levels und Dronkers (2008, S. 846) fanden entgegen ihrer Erwartungen einen positiven Effekt der relativen Größe der Zuwanderergruppe.

Die technische Umsetzung eines *cross-classified*-Modells erfolgt in der Analysesoftware MLwiN durch ein datentechnisches Verfahren. Wir betrachten nun Schüler und Schülerinnen, die in Schulklassen und Nachbarschaften gruppiert sind: Zunächst werden wie in einem Zwei-Ebenen-Modell die Schüler und Schülerinnen als Ebene 1 und die Schulklassen als Ebene 2 angenommen. Für die Abbildung der Kreuzklassifikation wird eine künstliche Ebene 3 eingeführt, die nur aus einer (!) Einheit besteht, der also zunächst alle Schüler und Schülerinnen gleichermaßen angehören. Auf der Ebene 1 der Schüler und Schülerinnen werden nun ebenso viele Dummy-Variablen gebildet wie Kontexte auf der kreuzklassifizierten Ebene der Nachbarschaft vorliegen (z. B. 190). Die Besonderheit besteht darin, dass die fixen Effekte dieser *Dummies* durch einen *constraint* – also durch eine *a priori Festlegung* eines geschätzten Parameters – auf den Wert Null gesetzt und damit ausgeschlossen werden. Sie können aber Zufallseffekte v_{jk} im Sinne von *random slopes* auf der künstlichen Ebene 3 aufweisen. Sodann werden die Varianzen der v_{jk} wiederum durch einen *constraint* auf identische Werte fixiert – d. h. $var(v_{1k}) = var(v_{2k}) = var(v_{3k}) \dots = var(v_{jk})$ – und keine Kovarianzen zwischen den *random slopes* der Dummies zugelassen ($cov(v_{1k}, v_{2k}) = 0$) (Hox 2010, S. 172; Rabe-Hesketh und Skrondal 2008, S. 474).

Im folgenden Analysebeispiel untersuchen wir anhand einer großen deutschlandweiten Schülerbefragung (vgl. Baier et al. 2006) den Einfluss von Kontextmerkmalen der Schulklasse und der Nachbarschaft (grob definiert durch Postleitzahlbezirke) auf das Schulschwänzen. Diese Form des niedrigschwelligen Schulabsentismus tritt insgesamt bei knapp 38 % aller Schülerinnen und Schüler der 9. Jahrgangsstufe auf.

Tabelle 1 zeigt Ergebnisse kreuzklassifizierter binärer logistischer Regressionsmodelle auf die Chancenverhältnisse (*Odds Ratios*), dass im letzten Jahr *mindestens eine Schulstunde* geschwänzt wurde. Modell 1 zeigt ein Modell nur mit Kontexteffekten. Modell 2 enthält Individualfaktoren und Kontextfaktoren – inklusive den Schultyp.

Richten wir unser Augenmerk auf Modell 2: Auffällig ist zunächst, dass Mädchen sogar häufiger schwänzen als Jungen. Auch Jugendliche in ökonomisch prekären Familienlagen tendieren eher zum Schwänzen, ebenso jene, die zu Hause niemals Elterngewalt erlebt haben, sowie jene, die viele Geschwister haben (signifikant nur auf dem 10-Prozent-Niveau). Hat der oder die Jugendliche kein eigenes Zimmer, ist das Risiko dagegen um den Faktor 0,8 reduziert. Das Risiko des Schwänzens ist erhöht bei Jugendlichen, die nicht bei beiden leiblichen Eltern zusammenleben (Faktor 1,3) oder die zu geringer Selbstkontrolle und damit zu einer stark ausgeprägten Risikosuche (vgl. Grasmick et al. 1993) neigen (Faktor 1,5). Einen nur auf dem 10-Prozent-Niveau signifikanten Effekt weist der Faktor „Computerspiel-

Tab. 1 Einflussfaktoren auf den Schulabsentismus – logistische kreuzklassifizierte Mehrebenenmodelle (Kontexte Schulklasse und Nachbarschaft, 9. Jahrgang): Odds Ratios

	(1)	(2)
Mädchen		2,183***
Anzahl der Geschwister		1,045 +
Eltern arbeitslos/Sozialhilfe		1,162 +
Lebt nicht mit beiden Eltern		1,329***
Eltern mit geringer Bildung		0,900
Nie elterliche Gewalt		1,109 +
Kein eigenes Zimmer		0,801*
SK: Risikosuche		1,514***
Computerspielhäufigkeit (Faktor)		1,058+
4 + delinquente Freunde		2,180***
Schulangst (Faktor)		1,033
Deutsch		Referenz
Deutsch 1.5		1,226*
Türkisch		1,366**
Russisch		1,252 +
Jugoslawisch		1,527*
Polnisch		1,484**
Südeuropäisch		1,257
Andere		1,453**
<i>Kontext Schulklasse</i>		
% Eltern geringe Bildung in Klasse	1,000	1,000
% Eltern arbeitslos/Sozialhilfe in Klasse	1,007 +	1,005
% kein eigenes Zimmer in Klasse	1,004	1,006
% 4 + delinquente Freunde in Klasse	1,019***	1,011***
Hauptschule		Referenz
IHR/R/REG		0,645***
Integrierte Gesamtschule		1,123
Gymnasium		1,134
Privat		1,255
<i>Kontext Nachbarschaft</i>		
% Eltern geringe Bildung in Klasse	1,002	1,005
% Eltern arbeitslos/Sozialhilfe in Klasse	1,007	1,008
% kein eigenes Zimmer in Klasse	0,994	0,990*
% 4 + delinquente Freunde in Klasse	1,012**	1,000
N (level 2) Klassen	748	694
N (level 2) Nachbarschaften	277	190
N (level 1) Schüler/-innen	11.036	10.129

Tab. 1 Fortsetzung

	(1)	(2)
var(Klassen)	0,396***	0,435***
var(Nachbarschaften)	0,045***	0,049***
R ² (McKelvey & Zavoina)	0,042	0,190

+ $p < 0,1$; * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

häufigkeit“ auf (vgl. Baier 2005), während vier oder mehr delinquente Freunde das Risiko des Schwänzens um das 2,18-fache erhöhen. Zudem scheinen Jugendliche mit Migrationshintergrund stärker zum Schwänzern zu neigen.

Als Kontextfaktor auf der Ebene der Schulklasse erhöht der Anteil der Individuen mit 4 oder mehr delinquenten Freunden höchst signifikant das Schwänzrisiko – obwohl diese Variable ebenfalls als Prädiktor auf der Individualebene kontrolliert wurde. Im Vergleich zur Hauptschule ist das Risiko nur in der Kategorie IHR/R/REG („Integrierten- Haupt- und Realschule, Realschule, Regelschule“) reduziert, während sich die anderen Schulformen nicht signifikant von der Hauptschule unterscheiden. Betrachten wir die Schulformunterschiede jedoch als Bruttoeffekte, d. h. ohne Kontrolle weiterer Faktoren (Modell hier nicht gezeigt), zeigen sich in IHR/R/REG und Gymnasium deutlich geringe Risiken als in der Hauptschule.

Als einziger Effekt des Nachbarschaftskontexts reduziert ein hoher Anteil Jugendlicher, die zu Hause kein eigenes Zimmer haben, das Risiko des Schwänzens. Als alleiniger Bruttoeffekt hat dieses Merkmal dagegen einen insignifikanten positiven Effekt (OR = 1,004, $p = 0,189$). Aufgrund der Kontrolle individueller und kontextueller Faktoren deutet sich also an, dass benachteiligte Wohnbedingungen sowohl auf der individuellen Ebene als auch auf der Kontextebene das Risiko des Schwänzens reduzieren. Auf der Individualebene kann dies eventuell dadurch erklärt werden, dass das Schwänzern die Jugendlichen den beengten häuslichen Verhältnissen aussetzen würde – zumindest dann, wenn sie sich zu Hause aufhalten, was aber sicher nicht in allen Fällen realistisch ist.

Die durch die fixen Effekte erklärte Varianz lässt sich zwar durch eine Variante des McKelvey und Zavoina-R² abbilden (vgl. Snijders und Bosker 2012). Jedoch sollte die Modellauswahl auf den Maßen AIC oder BIC basieren, die eine hohe Anzahl der in den Modellen geschätzten Parameter bestrafen (Hox 2010, S. 50).

Eine weitere Generalisierung der Mehrebenenanalyse stellen so genannte *multiple membership*-Modelle dar (Snijders und Bosker 2012, S. 210 ff.), bei denen eine Untersuchungseinheit nicht mehr nur einem Kontext, sondern mehreren Kontexten angehört. Derartige Datensituationen treten häufig – aber nicht nur – in Längsschnittdaten auf, wenn die Untersuchungseinheiten die Kontexte wech-

Bildungskontexte

Strukturelle Voraussetzungen und Ursachen ungleicher
Bildungschancen

Becker, R.; Schulze, A. (Hrsg.)

2013, X, 606 S. 45 Abb., Softcover

ISBN: 978-3-531-18226-1