

# Lehrerurteile im Lebensverlauf: Valide Prognose oder Self-Fulfilling Prophecy?

Von Klaus Birkelbach

**Zusammenfassung:** Im Längsschnitt einer Kohorte ehemaliger Gymnasiasten zwischen dem 15. und 43. Lebensjahr wird gezeigt, dass Lehrerurteile mit prognostischem Anspruch vermittelt über die Erfolgserwartungen der Schüler deren Entscheidungen bezüglich ihres weiteren Bildungswegs zu beeinflussen vermögen und dadurch zu einer Self-Fulfilling Prophecy im Sinne Mertons (1948) werden können. Ursprünglich unzutreffende Lehrerurteile können durch die Pfadabhängigkeit und Kumulativität des Ausbildungs- und Berufsverlaufes indirekt noch das Berufsprestige im 43. Lebensjahr beeinflussen. Die hier vorlegten Analysen belegen zugleich, wie sich die Längsschnittperspektive der Lebensverlaufsorschung fruchtbar mit der handlungstheoretischen Perspektive einer Wert-Erwartungstheorie verknüpfen lässt.

## Einleitung

Die herausragende Rolle schulischer Beurteilungen im Prozess gesellschaftlicher Statuszuweisung wird deutlich, wenn man sieht, dass die Schule die „erste und damit entscheidende zentrale Dirigierungsstelle für die künftige soziale Sicherheit, für den künftigen sozialen Rang und für das Ausmaß künftiger Konsummöglichkeiten“ ist, wie Schelsky (1957: 18) schon in den 1950-er Jahren formuliert hat. Aber nur wenn die Schule ihre Selektions- und Allokationsfunktionen auf der Basis inter-individueller Leistungsunterschiede, also nach meritokratischen Prinzipien wahrnimmt, erscheint die resultierende Ungleichheit vor dem Hintergrund eines normativen Modells von Leistungsgerechtigkeit legitim (z.B. Solga 2005; Hadjar / Becker 2009). Kritisch sind in diesem Zusammenhang die Bewertungen der Leistungen durch die Lehrer, die als Schulnoten, Zeugnisse, Abschlüsse und generalisierende Beurteilungen mit prognostischem Anspruch, wie die Übergangsempfehlungen am Ende der Grundschulzeit, die Basis für diesen Allokations- und Selektionsprozess bilden und so eine zentrale Weichenstellungsfunktion für den weiteren Lebensverlauf haben.

Aufgrund der Pfadabhängigkeit des Lebensverlaufs haben frühe Weichenstellungen weitreichende Folgen, denn sie erweitern oder reduzieren subjektiv und objektiv verfügbare Handlungs- und Entscheidungsspielräume. Dies gilt in besonderem Maße für den Ausbildungs- und Berufsverlauf, wo die institutionell vorgegebene Struktur des Bildungswesens eine Folge von Übergängen erzwingt, deren Richtung auf der Mikroebene individueller Akteure über die beruflichen Möglichkeiten, die Erreichbarkeit von sozialem Status und materiellen Lebenschancen bestimmt und dadurch auf der Makroebene die Strukturen sozialer Ungleichheit prägt. Die Übergänge und die Richtung der möglichen Sequenzen wird durch das institutionelle Gefüge des Bildungswesens vorgegeben, so dass individuelle Bildungsverläufe – und damit Lebensverläufe insgesamt – maßgeblich durch staatliche Institutionen vorgeprägt werden (Mayer / Müller 1989; Mayer 1990, 2004). An jedem Übergang sind im Rahmen der institutionellen Vorgaben Entscheidungen über den weiteren Weg zu treffen. Ein wichtiges Element der Vorgaben, die die Entscheidungs- und Handlungsräume der Betroffenen (bzw. der Eltern) je nach spezifischem Übergang in unterschiedlichem Maße beschränken (Hillmert 2004), stellen institutionell verankerte Akteure dar, die, wie z.B. Lehrer, Empfehlungen aussprechen und gegebenenfalls auch Entscheidungen bezüglich des jeweiligen Übergangs treffen und somit als „Gatekeeper“ (Heinz 1996: 59) fungieren.

Im vorliegenden Beitrag sollen Zusammenhänge zwischen einer Einschätzung der Eignung für ein späteres Studium durch die Klassenlehrer im 10. Schuljahr und den Erfolgen im Ausbildungs- und Berufsverlauf im Längsschnitt einer Kohorte ehemaliger Gymnasiasten bis zum 43. Lebensjahr untersucht werden. Konkret wird gefragt, inwieweit ein qualitatives Lehrerurteil über die Studieneignung mit den Übergängen auf dem Weg in ein Studium, dem Studien-erfolg und dem beruflichen Erfolg im Sinne einer validen Prognose zusammenhängt, oder ob es die Entscheidungen und den langfristigen Erfolg in Form einer Self-Fulfilling Prophecy auch zu beeinflussen vermag. In Kapitel 2 werden Untersuchungen vorgestellt, die sich mit der Frage der Validität von Leistungsbeurteilungen auseinandersetzen, aber wie solche Beurteilungen individuelle Bildungsentscheidungen beeinflussen und welche langfristigen Folgen dies im Lebensverlauf hat, wurde m.W. zumindest für Deutschland noch nicht untersucht. Wenn Hauser (2002: 56) im Zusammenhang mit Blick auf die in den USA üblichen Tests als Kriterien für Zugänge an den Übergängen im Bildungswesen fordert: „As sociologists, we ought also to take a longer view and start thinking now about how to measure, analyze, and assess the long term consequences of test use for life chances“, dann gilt dies auch für Beurteilungen durch die Lehrer.

Die hier untersuchten Lehrerurteile sind nicht unmittelbar mit Übergangs- oder Zugangsberechtigung verknüpft; es handelt sich um eine Prognose des Klassenlehrers bezüglich der Eignung seiner Schüler für ein Studium, man könnte auch sagen: eine Einstellung gegenüber der Leistungsfähigkeit der einzelnen Schülerinnen und Schüler, die auf einem ganzen Bündel unterschiedlicher Beobachtungen und Erwartungen der Lehrer beruhen mag (Brophy / Good 1974: 32ff.), aber von der aufgrund des meritokratischen Anspruchs der Schulen erwartet wird, sich primär an den Kriterien Leistung und Leistungsfähigkeit zu orientieren. Wenn sich im weiteren Lebensverlauf Zusammenhänge mit diesem Urteil zeigen, dann scheint dies zunächst ein Indiz der Qualität und Validität des Lehrerurteils zu sein. Die Validität ist aber eine empirisch zu beantwortende Frage, die nicht allein am Kriterium des Eintreffens der Prognose überprüft (predictive validity) werden kann, sondern auch an konkurrierenden Kriterien (concurrent validity), wie der kognitiven Leistungsfähigkeit. Fällt beides auseinander, bestehen also Zusammenhänge zwischen dem Urteil und den Ereignissen im weiteren Ausbildungs- und Berufsverlauf, obwohl Lehrerurteil und die im Rahmen eines Tests nachgewiesene kognitive Leistungsfähigkeit des Schülers nicht übereinstimmen, dann ist dies ein Hinweis auf Mechanismen, wie sie Merton (1948) in seinem vielzitierten Artikel zur Self-Fulfilling Prophecy herausgearbeitet hat.

Im Folgenden wird zunächst ein kurzer Überblick über Untersuchungen zur Güte von Lehrerurteilen in Form von Schulnoten und Übergangsempfehlungen gegeben und Ergebnisse von Studien zur längerfristigen prognostischen Validität von Lehrerurteilen vorgestellt. Im Anschluss daran wird das Problem einer Unterscheidung zwischen einer validen Prognose von Lehrerurteilen und einer Self-Fulfilling Prophecy aufgegriffen und diskutiert, wie invalide Urteile als Self-Fulfilling Prophecy den Lebensverlauf in die prognostizierte Richtung zu lenken vermögen und dadurch das Leitbild eines leistungsbezogenen Statuszuweisungsprozesses verletzen. Dabei wird eine Erklärung vorgeschlagen, die sich an Essers Modell einer soziologischen Erklärung (u.a. 1993, 2006) und deren Übertragung durch Huinink und Schröder (2008) auf das Lebensverlaufsparadigma anlehnt. Zur empirischen Überprüfung der These, dass Lehrerurteile auch als Self-Fulfilling Prophecy den weiteren Lebensverlauf zu beeinflussen vermögen, wird ein einfaches Modell entwickelt, bei dem die Lehrerurteile der kognitiven Leistungsfähigkeit der Schüler gegenüber gestellt und in multivariaten Regressionsmodellen die Einflüsse auf die Entscheidung zum Abitur und dem Eintritt in ein Studium, den Studienabschluss und den beruflichen Erfolg im 43. Lebensjahr untersucht werden.

## Untersuchungen zu Lehrerurteilen

### *Schulnoten und Lehrerempfehlungen beim Übergang in die Sekundarstufe*

Was Schulnoten wirklich messen, wird spätestens seit dem von Ingenkamp (1971) herausgegebenen Reader „Die Fragwürdigkeit der Zensurengebung“ kritisch betrachtet (z.B. Ziegenspeck 1999; Tent 2001; Trautwein et al. 2008). Eine Vielzahl von Untersuchungen zeigt, dass Lehrer sich bei der Zensurengebung an einem klasseninternen Bezugssystem orientieren (Ingenkamp 1971: 156ff., 1989: 59; Ziegenspeck 1999: 142 f.; Schrader / Helmke 2001: 50). Darüber hinaus werden klassenstufen-, schulform- und länderspezifische Unterschiede in der Benotungspraxis berichtet (Ziegenspeck 1999: 143 f.). In verschiedenen Untersuchungen konnten schicht- und geschlechtsspezifische (zusammenfassend: Ziegenspeck 1999: 209ff.) sowie bei den einzelnen Lehrern liegende subjektive Unterschiede der Bewertungsmaßstäbe nachgewiesen werden (Ingenkamp 1971: 57ff.; Ziegenspeck 1999: 173ff.). Die Folge ist, dass Schüler der gleichen Schulform und Klassenstufe für objektiv gleiche Leistungen unterschiedlich bewertet werden.

Ähnlich fragwürdig wie Schulnoten, die konkrete Leistungen bewerten, erscheinen qualitative Lehrerurteile, die eine Prognose zukünftiger Erfolge im Bildungswesen beinhalten. Das gilt auch für die Lehrerempfehlungen am Ende der Grundschule zur Schulformwahl. Aktuelle Studien zeigen, dass die Lehrerempfehlungen von Kompetenzen, die davon unabhängig in Tests ermittelt wurden, zumindest teilweise entkoppelt sind (Lehmann et al. 1997; Bos / Pietsch 2004; Bos et al. 2004; Arnold et al. 2007). Zwar unterscheidet sich die durchschnittliche Höhe der gemessenen Kompetenzen zwischen den Kindern mit Hauptschul-, Realschul- und Gymnasialempfehlung erwartungsgemäß, aber es existiert ein breiter Bereich mit fast der Hälfte der Schüler, wo jede Empfehlung möglich ist (Bos et al. 2004: 197; Arnold et al. 2007). Zugleich werden Zusammenhänge zwischen den Übergangsempfehlungen und der sozialen Herkunft nachgewiesen (Ditton / Krüskens 2006; Pietsch / Stubbe 2007). Bei der Beurteilung von Schülerinnen und Schülern und den darauf basierenden Schullaufbahneempfehlungen orientieren sich Lehrer in starkem Maße an den Schulnoten (Bos et al. 2004: 204ff.; Kristen 2006; Arnold et al. 2007: 283). Wenn Noten die Grundlage derartiger Beurteilungen mit einem in die Zukunft weisenden prognostischen Anspruch darstellen, dann haben solche Beurteilungen die gleichen Probleme wie Noten. Doch trotz aller Reliabilitäts- und Validitätsprobleme sind Noten vor allem Messungen manifester Leistungen, während in prognostische Beurteilungen der Lehrer weitere Kriterien einfließen. So können Becker und Birkelbach (2010, 2011) zeigen, dass solche Beurteilungen nicht nur auf der kognitiven Leistungsfähigkeit und den Noten der Schüler basieren, sondern auch durch die soziale Herkunft der Schüler beeinflusst werden. Es bestehen auch Zusammenhänge zwischen der sozialen Herkunft und den Noten, aber diese sind deutlich schwächer

### *Längerfristige Prognosen*

Der Übergang von der Grundschule in die weiterführende Schule ist eine besonders bedeutsame Weichenstellung im Ausbildungs- und Berufsverlauf, aber weitere folgen: Schulzeugnisse dienen Arbeitgebern als ein wichtiges Kriterium der Auswahl von Lehrstellenbewerbern; mit dem Abitur wird die allgemeine Hochschulreife bescheinigt, also eine Prognose über die Studierfähigkeit gegeben, und der Notendurchschnitt wird mehr und mehr zum zentralen Kriterium der Studienzulassung. Aber sind gute Schulzeugnisse tatsächlich valide Prädiktoren für den Erfolg im folgenden Ausbildungsabschnitt? Schuler (1998) gibt einen Überblick über Studien, in denen die prognostische Validität von Schulnoten für den Erfolg in Ausbildung, Studium oder Beruf untersucht wurde. Dabei kann er für die Kriterien Erfolg in der Berufsausbildung und im Studium auf deutsche Studien zurückgreifen, während er Zusammenhänge zwischen Schulnoten und Berufserfolg fast nur in amerikanischen Studien untersucht findet.

Das Bild, das sich beim Vergleich verschiedener Einzelstudien ergibt, ist nicht konsistent. Die berichteten Korrelationen zwischen Abitur- und Studiennoten variieren ebenso wie die Zusammenhänge zwischen Schulnoten und Berufsausbildungserfolg stark (Schuler 1998: 370 f.). Zu ähnlichen Ergebnissen kommen Baron-Boldt et al. (1988, 1989) in Metaanalysen zu Zusammenhängen zwischen Schulzeugnissen und Ausbildungserfolg sowie Abiturnoten und Studienerfolg, während Trapmann et al. (2007) in einer Metaanalyse europäischer Studien seit den 1980er Jahren positive Zusammenhänge zwischen Schulabschlussnoten und Studiennoten ermitteln konnten.

## Lebensverlaufsparadigma und Wert-Erwartungstheorie

Huinink und Schröder (2008) erweitern das Lebensverlaufsparadigma (Mayer 1990, 2004) auf der Mikroebene individuellen Handelns um Elemente einer Wert-Erwartungstheorie, bei dem die jeweils zu treffenden Entscheidungen dem Ziel der Produktion und Sicherung individueller Wohlfahrt dienen. Dabei wird unterstellt, dass die Akteure an den verschiedenen Übergängen des Lebensverlaufes effizient zu handeln suchen, also Kosten und Nutzen der Handlungen zur Verwirklichung unterschiedlicher und teilweise konkurrierender Lebensziele in den verschiedenen, hochgradig interdependenten Dimensionen des Lebensverlaufes gegeneinander abwägen. Von besonderer Bedeutung für die Bestimmung des Referenznutzens, an dem Kosten und Nutzen alternativer Entscheidungen gemessen werden, sind neben dem direkten biografischen Nutzen und dem Anspruchsniveau die subjektive Gewichtung konkurrierender Ziele (Salienz) sowie die Erwartungen, die Ziele auch tatsächlich erreichen zu können.

Ähnliche Modelle werden in der Bildungsforschung zur Erklärung von Bildungsentscheidungen schon lange genutzt. Boudon (1974) weist darauf hin, dass Bildungsentscheidungen in einem Spannungsfeld zwischen schulischen Leistungen, spezifischen Gegebenheiten des Bildungssystems mit seinen Selektionsmechanismen und schichtspezifischen Aspirationen, Erwartungen und Bewertungen getroffen werden. Dabei wird zwischen primären Effekten der sozialen Herkunft, die zu einem mit der Herkunftsschicht kovariierenden schulischen Leistungsniveau führen, und einem sekundären Effekt, der sich aus den spezifischen Entscheidungskalkülen unterschiedlicher sozialer Schichten ergibt, unterschieden. Die sekundären Effekte führen dazu, dass bei gleichen Leistungen Bildungsentscheidungen mit der sozialen Herkunft variieren. Weiter entwickelt werden diese Überlegungen zu Bildungsentscheidungen im Rahmen der Wert-Erwartungstheorie u.a. von Erikson und Jonsson (1996), Breen und Goldthorpe (1997) sowie Esser (1999: 265ff.). In Essers Modell lässt sich der subjektiv erwartete Nutzen einer Entscheidung zwischen zwei Bildungswegen in einer Funktion des Bildungsertrags, der materiellen und immateriellen Bildungskosten, sowie des Wertes eines möglichen Statusverlustes darstellen. Darüber hinaus wird die Funktion durch die Erwartung, dass ein Statusverlust tatsächlich eintritt, und die Erwartung, dass der Bildungsertrag auch wirklich erzielt werden kann, beeinflusst. Eine Entscheidung zugunsten der anspruchsvolleren, weiterführenden Alternative wird umso wahrscheinlicher, je mehr die Bildungsmotivation die Risiken übersteigt. Die konkrete Entscheidung wird also nicht nur durch eine Bilanzierung des erwarteten Nutzens und der Kosten, sondern auch von der Erwartung, das Ziel erreichen zu können, beeinflusst. Je besser die eigenen Erfolgchancen eingeschätzt werden, umso geringer fallen die Kosten ins Gewicht.

Das gilt auch für die Entscheidung ein Studium aufzunehmen. Verschiedene Studien (Becker / Hecken 2008; Becker 2009; Schindler / Reimer 2010) zeigen, dass soziale Ungleichheiten beim Hochschulzugang weniger durch primäre als durch sekundäre Effekte der sozialen Herkunft beeinflusst werden, und dass ein zwischen den verschiedenen sozialen Strata variierendes Investitionsrisiko maßgeblich dazu beiträgt, Arbeiterkinder von einem Studium in andere Bildungswege umzulenken. Mangels anderer Maßstäbe dürften sich insbesondere

Jugendliche aus den unteren Schichten an den Einschätzungen ihrer Lehrer, ob sie ein Studium erfolgreich zu bewältigen können, orientieren. Solche Einschätzungen werden – auch wenn sie den Schülern gegenüber unausgesprochen bleiben – im Schulalltag auf vielfältige Weise kommuniziert (Brophy / Good 1974; Brattesani et al. 1984; Good / Brophy 2003). Die Einschätzungen der Lehrer – ob es sich nun um valide Urteile handelt oder nicht – werden so zum Bestandteil des Selbstkonzeptes der Schüler (Dalbert / Stöber 2008) und beeinflussen über die Erfolgserwartungen zukünftige Bildungsentscheidungen. Während Schulnoten und Abschlüsse die objektiven Übergangsmöglichkeiten der Schüler bestimmen, haben Einstellungen und Erwartungen der Lehrer, indem sie für die Schüler zum Element ihrer Definition der Situation werden, Einfluss auf die Entscheidungen der Schüler, ihre durch Noten und Abschlüsse eröffneten Möglichkeitsräume auch tatsächlich zu nutzen. Durch diesen Mechanismus kann das Lehrerurteil zu einer Self-Fulfilling Prophecy werden, die den weiteren Lebensverlauf beeinflusst.

### Wie kann die Logik einer möglichen Self-Fulfilling Prophecy durchbrochen werden?

Merton (1948: 195 f.) fasst seine Überlegungen zur Self-Fulfilling Prophecy folgendermaßen zusammen: “The Self-Fulfilling Prophecy is, in the beginning, a false definition of the situation evoking a new behavior which makes the originally false conception come true. The specious validity of the Self-Fulfilling Prophecy perpetuates a reign of error. For the prophet will cite the actual course of events as proof that he was right from the very beginning. [...] Such are the perversities of social logic.”

Die Frage, inwieweit Lehrerurteile den weiteren Ausbildungs- und Berufsverlauf und damit letztlich die individuellen Lebenschancen valide prognostizieren können, oder ob sie über die Definition der Situation der Schülerinnen und Schüler deren Entscheidungen beeinflussen und dadurch als Self-Fulfilling Prophecy den Lebensverlauf erst in eine bestimmte Richtung lenken, kann nicht durch einen einfachen Abgleich der Vorhersagen mit den Ereignissen im Lebensverlauf beantwortet werden. Ein parallel zu den Lehrerurteilen erhobenes konkurrierendes Kriterium für die Validität des Lehrerurteils, das zugleich ein guter Prädiktor für Erfolge in der Ausbildungs- und Berufslaufbahn ist, kann aber helfen, die Logik einer Self-Fulfilling Prophecy zu durchbrechen. Dieses Kriterium sollte möglichst objektiv gemessen worden werden, also nicht durch subjektive Eindrücke und Maßstäbe des Lehrers beeinflusst sein. Es sollte sich auch um ein Maß handeln, dass den meritokratischen Ansprüchen des Bildungswesens genügt, also um einen Leistungsindikator. Dabei ist ein Indikator, der über die durch unterschiedliche situative Einflüsse beeinflussten aktuellen Leistungen hinausweist und die generelle Leistungsfähigkeit indiziert, zu bevorzugen.

Auch wenn Schulnoten Leistungsindikatoren sind, so haben sie doch die angesprochenen Reliabilitäts- und Validitätsprobleme, so dass sie als konkurrierendes Kriterium ungeeignet erscheinen. Objektive Kompetenztests, so wie sie u.a. in PISA (z.B. OECD 2006) durchgeführt wurden, wären ideal, stehen aber nicht zur Verfügung. Allerdings messen solche Tests neben bereichsspezifischen Kompetenzen vor allem auch die Intelligenz als einen Indikator der allgemeinen kognitiven Leistungsfähigkeit (vgl. Rindermann 2006; Baumert et al. 2007; Prenzel et al. 2007). Enge Zusammenhänge zwischen Testergebnissen und allgemeiner Intelligenz werden auch für den in den USA verwendeten Scholastic Aptitude / Assessment Test (SAT) berichtet (Frey / Detterman 2004). Intelligenz ist also ein guter Indikator kognitiver Leistungsfähigkeit im Bildungswesen; sie prognostiziert darüber hinaus die Leistungsfähigkeit im Beruf sehr gut (Schmidt / Hunter 1998; Schmidt 2002). In der vorliegenden Untersuchung wird daher auf die Ergebnisse eines Intelligenztests als konkurrierendes Kriterium zu den Lehrerurteilen zurückgegriffen. Den „Prophezeiungen“ der Lehrer und den Ereignissen im Lebens-

verlauf der Schüler wird so einen externen Maßstab gegenübergestellt, mit dessen Hilfe untersucht werden kann, ob und wie weit Lehrerurteile zur Self-Fulfilling Prophecy werden können, indem sie Bildungsentscheidungen der Schüler beeinflussen.

Der skizzierte Ansatz unterscheidet sich von Untersuchungen eines Effektes von Lehrerwartungen auf die Leistungen der Schülerinnen und Schüler, der unter der Bezeichnung „Pygmalioneffekt“ (Rosenthal / Jacobson 1968) kontrovers diskutiert wird (u.a. Wineburg 1987; Rosenthal 1994; Jussim / Harber 2005). Das Explanandum liegt in der vorliegenden Untersuchung nicht auf der Ebene der individuellen Schülerleistung, sondern auf der einer Untersuchung von Mechanismen des Statuserwerbs, wenn mögliche langfristige Effekte von Lehrerurteilen bzw. -erwartungen auf Entscheidungen im Ausbildungs- und Berufserfolg untersucht werden. Die Frage, ob und inwieweit Lehrerwartungen als Self-Fulfilling Prophecy auch die Schülerleistungen zu beeinflussen vermögen, kann dabei offen bleiben, denn die in Form von Noten dokumentierten Leistungen sind als objektive Gegebenheit und in ihrer subjektiven Relevanz Bestandteil der Situation und fließen in die Entscheidungen über den nächsten Bildungsabschnitt ein.

## Daten und Hypothesen

### Daten

Datenbasis der Untersuchung ist das „Kölner Gymnasiasten-Panel“ (KGP).<sup>1</sup> Im Rahmen der Primärerhebung (PE) 1969 wurden 3240 nordrhein-westfälische Gymnasiasten des 10. Schuljahres klassenweise u.a. über ihre schulischen Interessen, Leistungen und Pläne, ihre soziale Herkunft und ihre Einstellungen zu Elternhaus und Schule befragt. Dabei wurde auch ein Intelligenz-Struktur-Test (IST) mit vier Subskalen (Amthauer 1953) durchgeführt (Meulemann 1979: 194), dessen Ergebnisse weder den Lehrern noch den Schülern oder ihren Eltern mitgeteilt wurden. Parallel zur PE wurden die Eltern (n=2646) und Lehrer (n=1701) befragt, wobei die Lehrer, die in den befragten Klassen unterrichteten, in einem Zusatzbogen um Einschätzungen zu ihren Schülern<sup>2</sup> gebeten wurden. 1984 / 85 wurde eine erste Wiederbefragung (WB1, n=1987) durchgeführt, in der die damals rund 30-jährigen ehemaligen Gymnasiasten in standardisierter Form detailliert über ihren beruflichen und privaten Werdegang zwischen dem 15. und dem 30. Lebensjahr Auskunft gegeben haben. Eine zweite Wiederbefragung 1996 / 97 (WB2, n=1596) knüpft an die erste an und ergänzt die Lebensverläufe um den Zeitraum bis zum 43. Lebensjahr. Analysen zur Panelmortalität (Birkelbach 1998 b) zeigen, dass die Ausfallneigung wesentlich stärker durch Indikatoren für die Kooperationsbereitschaft und für das Interesse am Thema beeinflusst wird als durch biografische Erfolgskriterien. Dennoch ist an den Variablen zum Bildungsverlauf ein schwacher Erfolgsbias ablesbar, der die Varianz der Bildungsverläufe etwas einschränkt, aber gerade deshalb zu einer strengeren Überprüfung möglicher Effekte der Lehrerurteile auf die Erfolge im Ausbildungs- und Berufsverlauf führt.

- 1 Die PE (ZA-Studie Nr. 600, 639, 640) wurde am Kölner Forschungsinstitut für Soziologie unter Leitung von H.-J. Hummell, M. Klein, M. Wieken-Mayser und R. Ziegler durchgeführt. Projektleiter der WB1 (ZA-Studie Nr. 1441) waren H. Meulemann, H.-J. Hummell, M. Wieken-Mayser und R. Ziegler. WB2 (ZA-Studie Nr. 4228) wurde im CATI-Labor des Instituts für Angewandte Sozialforschung der Universität zu Köln durchgeführt und von der DFG finanziert (Birkelbach et al. 2000). Ergebnisse von WB1 und WB2 werden von Meulemann (1995), Birkelbach (1998 a) und Meulemann et al. (2001) berichtet. Eine durch die DFG finanzierte dritte Wiederbefragung wurde 2010 durchgeführt. Die hier vorgelegten Analysen basieren auf zuvor noch nicht ausgewerteten Zusatzbögen zur Lehrerbefragung, deren Aufbereitung und Auswertung von der DFG gefördert wurde. Besonderer Dank gilt M. Wieken-Mayser und H. Meulemann für ihre Unterstützung bei der Aufbereitung und Integration der Daten in das KGP.
- 2 Zur sprachlichen Vereinfachung wird im Folgenden nur von „Schülern“ statt von „Schülerinnen und Schülern“ gesprochen.

*Lehrerurteile und kognitive Leistungsfähigkeit*

Die Lehrer wurden 1969 in zwei aufeinander folgenden Fragen gebeten, die Schüler ihrer Klasse zu benennen, die sie (a) „unabhängig vom derzeitigen Leistungsstand für besonders geeignet halten, ein Studium zu ergreifen“, bzw. die, von denen sie glauben, „dass sie auf keinen Fall zum Studium geeignet“ seien.<sup>3</sup> Für die übrigen liegt kein explizites Lehrerurteil vor. Aus diesen Informationen wurde eine Variable mit drei Kategorien konstruiert: (1) „Ungeeignet“, (2) „Ohne Beurteilung“ und (3) „Geeignet“, deren Verteilung Tabelle 1 entnommen werden kann.<sup>4</sup>

*Tabelle 1: Lehrerurteil zur Studieneignung und kognitiven Leistungsfähigkeit (IST) (Spaltenprozente)\**

	Lehrerurteil			Gesamt
	– Ungeeignet	0 Kein Urteil	+ Geeignet	
IST: MW	107,8	109,6	114,7	110,8
SD	9,6	11,3	11,8	11,4
IST<Median	–/–	–/0	–/+	46,5
–	Valides negatives Urteil 58,7	IST unterdurch., kein Urteil 47,6	Überschätzt 35,6	
IST>=Median	+/-	+/0	+/+	53,5
+	Unterschätzt 41,3	IST überdurch., kein Urteil 52,4	Valides positives Urteil 64,4	
N	288	473	371	1132
	25,4	41,8	32,8	100

\* am Median (110) dichotomisiert

Stichprobe: Teilnehmer WB2 (1996 / 07)

MW: arithmetisches Mittel, SD: Standardabweichung

Die durch den IST (Amthauer 1953) erhobene kognitive Leistungsfähigkeit dient als ein konkurrierendes Kriterium, an dem die Prognose der Studieneignung gemessen wird. Tabelle 1 weist im oberen Bereich die Mittelwerte und Standardabweichungen des IST in den durch die Lehrerurteile gebildeten Gruppen aus.<sup>5</sup> Die durchschnittliche Intelligenz der von ihren Lehrern als für ein Studium geeignet beurteilten Schüler liegt wie erwartet signifikant ( $p < .0001$ ) über den Ergebnissen der als ungeeignet Beurteilten, während es den Lehrern schwer fällt, die die große Gruppe der durchschnittlich Leistungsfähigen zu beurteilen. Offenkundig besteht ein positiver Zusammenhang zwischen der kognitiven Leistungsfähigkeit der Schüler und den Beurteilungen durch die Lehrer.

- Es wird auf das Urteil der Klassenlehrer zurückgegriffen, da diese eher als Fachlehrer in der Lage sein dürften, ihre Schüler auch über die Leistungen in einzelnen Fächern hinaus zu beurteilen.
- Zusätzliche Analysen haben gezeigt, dass Schüler ohne Beurteilung hinsichtlich ihrer Noten zwischen den Gruppen mit positiver bzw. negativer Beurteilungen liegen. Offenbar haben die Lehrer vor allem dann eine Beurteilung abgegeben, wenn sie sich ihres Urteils aufgrund der dokumentierten Leistungen sicher waren. Als fehlend ausgeschlossen wurden die Lehrer, die keinen Schüler beurteilt haben.
- Der Mittelwert des IST liegt für die Gymnasiastenstichprobe um etwa eine Standardabweichung über der Eichstichprobe.

Durch Kreuztabellierung der Lehrerurteile mit dem am Median dichotomisierten IST werden sechs Gruppen gebildet.<sup>6</sup> Das Lehrerurteil und die kognitive Leistungsfähigkeit stimmt bei zwei Gruppen (–/– und +/+) überein. Gemessen am Kriterium der kognitiven Leistungsfähigkeit geben die Lehrer jeweils in der Mehrheit der Fälle, in denen sie sich eine Einschätzung zutrauen, ein valides Urteil über ihre Schüler ab. Komplementär dazu sind die beiden Gruppen, bei denen beide Kriterien nicht übereinstimmen. Rund 41,3% der als ungeeignet beurteilten Schülerinnen und Schüler sind überdurchschnittlich leistungsfähig, werden also von ihren Lehrern unterschätzt (+/–). Eine Überschätzung der unterdurchschnittlich Leistungsfähigen (–/+) ist in etwas geringerem Umfang (35,6%) zu beobachten. Trifft bei diesen Gruppen die Prognose der Lehrer ein und entwickelt sich der weitere Lebensverlauf entsprechend, dann ist dies ein Hinweis auf die Wirkung einer Self-Fulfilling Prophecy. Die beiden Gruppen, bei denen die Lehrer keine Beurteilung abgeben konnten oder wollten (–/0 und +/0), sollen in den Analysen als Kontrollgruppen zeigen, wie die Entscheidungen im Ausbildungsverlauf und der beruflichen Erfolg unbeeinflusst durch ein Lehrerurteil mit der kognitiven Leistungsfähigkeit zusammenhängen.

### Hypothesen

Die Entscheidungen zum Abitur und Studieneintritt werden als Voraussetzungen des Studienerfolgs, d.h. des Ereignisses, auf das die Prognose des Lehrerurteils zielt, in der Gesamtgruppe der Befragten, für die aus der ersten Wiederbefragung entsprechende Lebensverlaufsdaten vorliegen, untersucht.<sup>7</sup> Wenn das Lehrerurteil im Sinne einer Self-Fulfilling Prophecy die Entscheidungen beeinflusst, dann sollte dies an diesen beiden Übergängen sichtbar werden. Anschließend werden Zusammenhänge des Lehrerurteils mit dem Studienerfolg bei den Befragten, die ein Studium aufgenommen haben, analysiert. Im Berufserfolg am Ende des Beobachtungszeitraums (i.d.R. im 43. Lebensjahr) kumulieren schließlich die bisherigen Entscheidungen und Erfolge im Ausbildungs- und Berufsverlauf. Untersuchungsgesamtheit sind daher alle in der zweiten Wiederbefragung 1997 Erwerbstätigen. Erhoben wird der Berufserfolg durch das auf Wegeners (1988) Magnitude-Prestigeskala gemessene Prestige des letzten Berufs im Beobachtungszeitraum. Im Folgenden werden vier Hypothesen zu den Zusammenhängen zwischen den durch Kreuztabellierung von Lehrerurteilen und kognitiver Leistungsfähigkeit gebildeten sechs Gruppen mit den genannten Kriterien des Lebensverlaufes begründet.

Die *Intelligenz-Hypothese* bezieht sich auf die Qualität einer Prognose, die sich auf die kognitive Leistungsfähigkeit stützt. Wenn Erfolg im Ausbildungs- und Berufsverlauf ehemaliger Gymnasiasten vor allem von der kognitiven Leistungsfähigkeit abhängt, dann ist zu erwarten, dass die Gruppen mit überdurchschnittlicher Intelligenz (+/+, +/0, +/–) jeweils häufiger das Abitur erreichen, ein Studium aufnehmen, es erfolgreich abschließen und langfristig beruflich erfolgreicher sind als die Vergleichsgruppen mit unterdurchschnittlicher Intelligenz (–/+, –/0, –/–).

6 Durch die Dichotomisierung des IST entsteht im mittleren Bereich eine Unschärfe. Während man sich bei den Extremwerten schnell einig werden dürfte, ob eine spezifische Beurteilung valide ist oder nicht, schwand die Sicherheit, je mehr man sich einer durchschnittlichen Intelligenz nähert. Es existiert kein „natürlicher“ Schwellenwert, anhand dessen absolut sicher zwischen validen Urteilen und Fehlurteilen unterschieden werden könnte. Allerdings erhält man durch die Kreuztabellierung Gruppen, die sich hinsichtlich der Wahrscheinlichkeit eines validen bzw. invaliden Urteils unterscheiden. Ein positives Urteil trifft z.B. bei Schülern mit überdurchschnittlicher Intelligenz mit einer höheren Wahrscheinlichkeit zu als bei Schülern mit unterdurchschnittlicher Intelligenz.

7 Da neben dem Abitur auch andere Wege in ein Studium möglich sind, werden beide Entscheidungen in der Gesamtgruppe analysiert. Tatsächlich haben weniger Befragte das Abitur an einer Schule erworben als ein Studium begonnen.



Die *Hypothese einer wechselseitigen Validierung* geht davon aus, dass sich kognitive Leistungsfähigkeit und Lehrerurteil bei Übereinstimmung wechselseitig validieren. Bei den Gruppen, bei denen kognitive Leistungsfähigkeit und Lehrerurteil übereinstimmen, kann daher von einer besonders sicheren Prognose des Erfolgs ausgegangen werden, so dass in der Gruppe  $+/+$  der Anteil der an den jeweiligen Übergängen Erfolgreichen am höchsten, in der Gruppe  $-/-$  dagegen am niedrigsten liegen dürfte.

Im Zentrum der Untersuchung steht die Überlegung, dass Lehrerurteile als Self-Fulfilling Prophecy die Entscheidungen im Vorfeld des Studiums beeinflussen können und damit auch Einfluss auf die Weichen für den weiteren Ausbildungs- und Berufsverlauf zu nehmen vermögen (*Self-Fulfilling Prophecy-Hypothese*). Beeinflusst das im Schulalltag direkt oder indirekt kommunizierte Lehrerurteil im Sinne einer Self-Fulfilling Prophecy die Entscheidung, das Abitur und ein Studium anzustreben, dann sollten die unterschätzten ehemaligen Gymnasiasten  $(+/-)$  bei gleicher Leistungsfähigkeit an diesen beiden Übergängen seltener erfolgreich sein als die Vergleichsgruppe mit positivem Lehrerurteil  $(+/+)$  und die Kontrollgruppe ohne Beurteilung  $(+/0)$ . Umgekehrt sollten im Falle einer Self-Fulfilling Prophecy die überschätzten Befragten  $(-/+)$  nur unwesentlich seltener erfolgreich sein als die Befragten der Gruppe mit einem durch den IST validierten positiven Lehrerurteil  $(+/+)$ . Sie sollten zudem häufiger erfolgreich sein als die Gruppen mit vergleichbarer Leistungsfähigkeit, aber negativem  $(-/-)$  oder ohne Lehrerurteil  $(-/0)$ .

Die Richtigkeit der Überlegungen zur Wirkung einer Self-Fulfilling Prophecy vorausgesetzt, sollten bei Kontrolle der Entscheidungen und Erfolge im bisherigen Ausbildungs- und Berufsverlauf keine direkten Zusammenhänge mehr zwischen dem Lehrerurteil und dem Studienerfolg sowie dem Berufsprestige am Ende des Beobachtungszeitraums bestehen. Die *Kontroll-Hypothese* erwartet daher: Sind die Entscheidungen erst einmal zugunsten eines Studiums gefallen, dann sollte der beschriebene Mechanismus einer Self-Fulfilling Prophecy durch die vorausgegangenen (Selbst-)Selektionsprozesse seine Bedeutung verloren haben und das Lehrerurteil bei Kontrolle dieser Entscheidungen keinen direkten Effekt mehr auf den Erfolg im Studium und Beruf haben. Umgekehrt wären bei Kontrolle der Entscheidungen empirische Zusammenhänge zwischen dem Lehrerurteil und dem Studienerfolg sowie dem Berufsprestige als Beleg der prognostischen Validität des Lehrerurteils zu werten.

### *Entscheidungen und Erfolge im Ausbildungs- und Berufsverlauf*

In den Analysen wird zunächst ein nach dem Lehrerurteil und der kognitiven Leistungsfähigkeit differenzierter Überblick über den Erfolg beim Abitur (a), den Studieneintritt (b) und -abschluss (c) sowie das Berufsprestige mit 43 Jahren (d) gegeben. Diese Ergebnisse werden anschließend in multivariaten Analysen überprüft. Beim Abitur und beim Studieneintritt werden die mit monatlicher Genauigkeit erhobenen zeitlichen Informationen der vorliegenden Ereignisdaten genutzt und Verfahren der Ereignisanalyse (Blossfeld et al. 2009) eingesetzt. Bei der multivariaten Analyse dieser Ereignisse werden semiparametrische (Cox-)Regressionen gerechnet, um private Verlaufsumstände des Lebensverlaufes als zeitabhängige Variablen zu kontrollieren. Es werden also streng genommen Ereignisse – Übergänge – betrachtet, aber diese setzen entsprechende Entscheidungen im Lebensverlauf voraus. Der Studienerfolg wird mit einer binären logistischen Regression analysiert, da hier die Studiengänge zu kontrolliert sind, aber durch deren unterschiedliche Dauer die Proportionalitätsannahme der Cox-Regression verletzt würde (Blossfeld et al. 2009: 233ff.). Schließlich werden für die Analyse des Prestiges im letzten Beruf Mittelwerte und die Ergebnisse einer OLS-Regression berichtet.

In den multivariaten Analysen des Abiturs, des Studieneintritts und -erfolgs werden zusätzliche Aspekte der Situation kontrolliert, die die Entscheidungen, über das Abitur ein Studium und ein entsprechend ehrgeiziges Berufsziel anzustreben, beeinflussen können. Der Noten-

durchschnitt über alle Fächer des ersten Halbjahres der Klasse 10 dokumentiert die in diesem Zeitraum erbrachten Leistungen und ist – anders als die Testintelligenz – ein gut sichtbarer Leistungsindikator für Schüler und Lehrer. Hinzu kommt die in der Primärerhebung erfragte subjektive Selbsteinschätzung der eigenen schulischen Leistungen im Vergleich zur Klasse. Beides sind Faktoren, die – wie das Lehrerurteil – die subjektive Erwartung, ein Studium leistungsmäßig erfolgreich bewältigen zu können, beeinflussen dürften. Weiterhin wird die soziale Herkunft kontrolliert, die durch das Berufsprestige (Treiman 1977) des Vaters operationalisiert wird. Mit der sozialen Herkunft korreliert einerseits die Erwartung, ein Studium zu bewältigen, andererseits kann die Furcht vor einem drohenden Statusverlust die Entscheidung beeinflussen. Der subjektive Wert, der dem Abitur und dem Studium zugemessen wird, hängt von den langfristigen Zielen ab, die man sich im Leben setzt. Die Aspirationen werden bei den Analysen zum Abitur durch Fragen an die 15-Jährigen, ob sie ein festes Lebensziel haben und ob sie dafür das Abitur brauchen, und bei der Entscheidung zum Studieneintritt durch die Frage, ob ein Studium angestrebt wird, operationalisiert. Als Indikator geschlechtsspezifischer Lebenspläne wird das Geschlecht einbezogen. Darüber hinaus werden jeweils biografische Verlaufsumstände mit zeitabhängigen Variablen berücksichtigt, die die persönliche Lebenssituation verändern und daher den jeweiligen Übergang negativ beeinflussen (Birkelbach 1998 a). Kontrolliert wird bei den Analysen zum Abitur die Tatsache, bereits aus dem Elternhaus ausgezogen zu sein, und beim Studieneintritt die Elternschaft. Zusätzlich wird hier die Tatsache eines Wehr- oder Zivildienstes als zeitabhängige Variable berücksichtigt, weil sie den Studieneintritt der Männer um mehr als ein Jahr verschiebt. Bei der Analyse des Studienerfolgs wird auch der Studiengang kontrolliert. Als ein Indikator möglicher konkurrierender privater Ziele wird die Geburt eines Kindes vor dem Ende des Studiums berücksichtigt, mögliche geschlechtsspezifische Effekte werden durch eine Interaktionsvariable der Elternschaft mit dem Geschlecht (Mutterschaft) erfasst.

Der durch Wegeners Magnitude-Prestigeskala (1988) erfasste Berufserfolg am Ende des Beobachtungszeitraumes wird maßgeblich durch die Entscheidung für das Abitur, für die Aufnahme eines bestimmten Studiums und dessen erfolgreichen Abschluss geprägt (Birkelbach et al. 1997), so das frühere Bildungserfolge zu kontrollieren sind. In der multivariaten Analyse sollte das Lehrerurteil wegen seines Einflusses auf frühere Erfolge in der Bildungslaufbahn gemäß der Kontrollhypothese keinen eigenständigen Effekt auf den Berufserfolg mehr haben. Dies sollte auch für die soziale Herkunft und die Aspirationen gelten, aber frühere Analysen (Birkelbach et al. 1997) haben gezeigt, dass auch bei Kontrolle der Leistungen im bisherigen Bildungsweg sowohl die soziale Herkunft als auch die Aspirationen der Schüler einen signifikanten Effekt auf das Berufsprestige mit 43 Jahren behalten. Zugleich können eigene Entscheidungen (z.B. zugunsten der Familie zumindest zeitweise aus dem Erwerbsleben aussteigen, Elternschaft), aber auch Veränderungen der Situation, wie Arbeitslosigkeit oder eine längere Krankheit, die Berufskarriere beeinträchtigen. In der Gleichung werden daher die Elternschaft und die Dauer der Lebensphasen als Hausfrau/-mann sowie erzwungene Erwerbsunterbrechungen kontrolliert.

## Ergebnisse

### *Bivariate Analysen*

Die Ergebnisse (Tabelle 2) zeigen, dass die Mehrheit der Stichprobe den mit dem Besuch des Gymnasiums eingeschlagenen Pfad über das Abitur und ein Studium in einen prestigeträchtigen Beruf weiterverfolgt hat. Knapp drei Viertel der ehemaligen Gymnasiasten haben das Abitur erreicht. Der Anteil, der ein Studium aufgenommen hat, liegt noch etwas höher, was darauf hinweist, dass eine kleine Gruppe mit der Fachhochschulreife oder auf Umwegen den Weg ins Studium gefunden hat. Die große Mehrheit (85%) derjenigen, die sich für ein Studium

entschieden haben, schließt dieses auch erfolgreich ab und übt mit 43 Jahren einen angesehenen Beruf (vgl. Wegener 1988: 239ff.) aus. Allerdings ergeben sich im Hinblick auf die Entscheidung zum Abitur und zum Studieneintritt deutliche Unterschiede zwischen den sechs Gruppen. Sind diese beiden Hürden erst einmal überwunden und betrachtet man den Studienerfolg in der Gruppe, die ein Studium aufgenommen hat, dann fallen die Differenzen, wie mit der Kontroll-Hypothese zu erwarten ist, wesentlich geringer aus. Die Erfolge im Bildungsweg spiegeln sich wiederum erwartungsgemäß deutlich in den Ergebnissen für das Berufsprestige wider.

Tabelle 2: Erfolge im Ausbildungs- und Berufsverlauf

	Voraussetzungen			Gegenstand der Prognose		Langfristige Folgen	
	(a) Abitur		(b) Studieneintritt	(c) Studienerfolg		(d) Berufsprestige (MPS)	
	N	Anteil(*)	Anteil(*)	N	Anteil(*)	N	MW (*)
+/+ Valides positives LU	239	95,8 (1)	92,0 (1)	220	88,2 (1)	213	111,9 (1)
-/+ Überschätzt	132	85,6 (2)	81,8 (2)	108	87,0 (2)	109	107,0 (2)
+/0 IST hoch, kein LU	248	78,2 (3)	78,6 (3)	195	86,7 (3)	216	96,4 (3)
-/0 IST niedrig, kein LU	225	69,8 (4)	71,5 (4)	161	81,4 (5)	198	95,7 (4)
+/- Unterschätzt	119	50,4 (5)	62,2 (5)	74	75,7 (6)	106	90,6 (6)
-/- Valides negatives LU	169	43,2 (6)	60,9 (6)	103	85,4 (4)	137	94,4 (5)
Gesamt**	1384	73,9	77,6	861	85,0	979	99,9

LU: Lehrerurteil.

\* Rangplatz, gemessen am Anteil der Erfolgreichen (a-c), bzw. am Berufsprestige (d).

\*\* Stichproben: (a/b) Ehemalige Gymnasiasten, die mindestens an WB1 teilgenommen haben; (c) Teilstichprobe aus (a/b) mit Studieneintritt; (d) Teilnehmer an WB2.

Betrachtet man zunächst nur die Zahlen für die beiden Gruppen, über die ihre Lehrer geurteilt haben, sie seien für ein Studium nicht geeignet, dann scheint die prognostische Validität des negativen Lehrerurteils gering zu sein. Immerhin nehmen auch aus diesen Gruppen mehr als 60 Prozent ein Studium auf und schließen es entgegen der Einschätzung ihrer Klassenlehrer in ihrer großen Mehrheit erfolgreich ab. Der hohe Anteil Erfolgreicher in der Gruppe mit unterdurchschnittlicher Intelligenz und negativem Lehrerurteil dürfte teilweise auf die oben diskutierte Unschärfe durch die Dichotomisierung zurückzuführen sein, aber vor allem aber auf die für diese Kohorte günstigen Gelegenheitsstrukturen durch den Ausbau der Hochschulen in den frühen siebziger Jahren des 20. Jahrhunderts (vgl. Müller / Haun 1994; Müller et al. 2009).

Gemäß der Intelligenz-Hypothese sollten sich die Gruppen mit überdurchschnittlicher kognitiver Leistungsfähigkeit häufiger als die Vergleichsgruppen für das Abitur entscheiden und auch häufiger ein Studium aufnehmen, also die drei ersten Rangplätze einnehmen. Empirisch aber ergibt sich für beide Übergänge eine Rangfolge, die sich an erster Stelle an den Beurteilungen durch die Lehrer und erst an zweiter Stelle an der Intelligenz ausrichtet. Erst wenn man das Lehrerurteil konstant hält, zeigt sich für die Gruppen mit überdurchschnittlicher Testintelligenz ein Vorteil. Am erfolgreichsten sind die beiden Gruppen mit positivem Lehrerurteil, gefolgt von den Kontrollgruppen ohne Lehrerurteil. Am Ende der Rangfolge findet man diejenigen, die ihren Lehrern nach als ungeeignet für ein Studium galten.

Die Überlegung, dass sich die gemessene kognitive Leistungsfähigkeit und das Lehrerurteil bei Übereinstimmung wechselseitig validieren und die Prognose insofern doppelt abgesichert ist, wird ganz überwiegend bestätigt. Die Panelteilnehmer, die im Alter von 15 Jahren im Test ihre überdurchschnittliche Leistungsfähigkeit belegt haben und über die ihre Lehrer positiv geurteilt haben (+/+), liegen durchgängig an der Spitze der Erfolgsrangliste. Sie bewältigen zu

fast 96 Prozent das Abitur, nehmen mit 92 Prozent am häufigsten ein Studium auf, schließen es dann auch am häufigsten mit einem Abschluss (88 Prozent) ab und üben mit 43 Jahren im Schnitt die prestigeträchtigsten Berufe aus. Umgekehrt bewältigen diejenigen, bei denen eine unterdurchschnittliche Leistungsfähigkeit mit einem negativen Lehrerurteil kombiniert ist (–/–), am seltensten die Schwellen des Abiturs und des Studieneintritts.

Die Hypothese einer Self-Fulfilling Prophecy der Wirkung des Lehrerurteils nimmt an, dass sich bei gleicher kognitiver Leistungsfähigkeit die von ihren Lehrern positiv beurteilten Panelteilnehmer eher für Abitur und Studium entscheiden als diejenigen ohne oder mit negativer Beurteilung. Tatsächlich liegt die als „überschätzt“ bezeichnete Gruppe (–/+) an beiden Übergängen und auch im weiteren Ausbildungs- und Berufsverlauf auf dem zweiten Rangplatz gleich hinter der Gruppe mit einem positiven validen Lehrerurteil. Umgekehrt sind die „Unterschätzten“ (+ /–) unter den Befragten mit überdurchschnittlichem Leistungsvermögen an den beiden Übergängen deutlich seltener erfolgreich als die Gruppen ohne explizites oder mit positivem Lehrerurteil. Dass das Lehrerurteil mehr als eine Prognose sein kann und als Self-Fulfilling Prophecy einen eigenständigen Effekt auf den Erfolg hat, wird zusätzlich durch die mittleren Ränge der Gruppen ohne Lehrerurteil unterstrichen.

### *Multivariate Analysen*

#### *Entscheidungen im Vorfeld: Abitur und Aufnahme eines Studiums*

Tabelle 3 gibt die Ergebnisse von Cox-Regressionen des Abiturs und des Studieneintritts auf die durch das Lehrerurteil und die kognitive Leistungsfähigkeit gebildeten Gruppen, die spezifischen Aspirationen, Schulleistungen, die soziale Herkunft sowie private Lebenspläne und biografische Verlaufsumstände wieder.

Tabelle 3: Entscheidungen im Vorfeld: Abitur und Studienaufnahme (Cox-Regressionen)

	(a) Abitur		(b) Studienaufnahme		
	min / max	unstandard. exp(b)	standard. exp(b*SD)	unstandard. exp(b)	standard. exp(b*SD)
<b>Lehrerurteil und kognitive Leistungsfähigkeit</b>					
+/+ Valides positives LU (Referenz)					
-/- Überschätzt	0/1	0.98	1.01 <sup>-1</sup>	0.78	1.08 <sup>-1</sup>
+/0 IST hoch, kein LU	0/1	0.85	1.07 <sup>-1</sup>	0.85	1.07 <sup>-1</sup>
-/0 IST niedrig, kein LU	0/1	0.68	1.17 <sup>-1</sup>	0.69	1.16 <sup>-1</sup>
+/- Unterschätzt	0/1	0.50	1.24 <sup>-1</sup>	0.66	1.14 <sup>-1</sup>
-/- Valides negatives LU	0/1	0.46	1.32 <sup>-1</sup>	0.82	1.07 <sup>-1</sup>
<b>Leistung in der Schule</b>					
Notendurchschnitt Klasse 10	2.3/7.0	0.71	1.26	0.76	1.20
Selbstschätzung der Schulleistungen	-2/2	1.20	1.17	1.10	1.09
<b>Aspirationen</b>					
(a) Kein Abitur angestrebt (Referenz)					
(b) Kein Studium angestrebt (Referenz)					
(a) Abitur angestr., kein fest. Lebensziel	0/1	3.42	1.65	3.14	1.63
(b) Studium: Weiß noch nicht	0/1	3.70	1.75	3.27	1.54
(a) Eventuell Abitur für Lebensziel	0/1	4.18	2.03	3.43	1.83
(b) Studium: Eventuell	0/1				
(a) Abitur notwendig für Lebensziel	0/1				
(b) Studium: Sicher	0/1				
<b>Soziale Herkunft</b>					
Berufsprestige des Vaters	19/82	1.01	1.08	1.01	1.09
<b>Private Lebenspläne u. biografische Verlaufsumstände</b>					
Geschlecht: Frau	0/1	0.99	1.01 <sup>-1</sup>	Schichtungsvariable	
Eig. Wohnung vor Abitur (z)	0/1	0.27	1.85 <sup>-1</sup>		
Mit Kind (z)	0/1			.41	1.45 <sup>-1</sup>
In Wehr- / Zivildienst (z)	0/1			.003	1.19 <sup>-1</sup>
N =	1097 (zensiert: 27.3%)		1097 (zensiert: 23.7%)		
-2 LogLikelihood =	10138.87		9125.27		
Likelihood-Ratio Chi² =	425.7*** (13 FG)		673.5*** (13 FG)		

LU: Lehrerurteil, SD: Standardabweichung, (z): zeitabhängige Variablen. Notendurchschnitt des ersten Halbjahres der Klasse 10 im Klassenverband z-transformiert. Negative standardisierte Koeffizienten werden als Kehrwert dargestellt. \* p<05, \*\* p<01, \*\*\* p<001

## (a) Abitur

Auch bei Kontrolle der genannten Variablen wird für das Abitur, ausgehend von dem validen positiven Lehrerurteil (+/+), exakt die schon aus Tabelle 2 bekannte Rangfolge der Gruppen durch die Stärke der Effekte reproduziert. Auch hier stützen die Ergebnisse also die Hypothese der wechselseitigen Validierung von Intelligenz und Lehrerurteil, während die Intelligenz-Hypothese eines Primats der kognitiven Leistungsfähigkeit nur eingeschränkt aufrecht erhalten werden kann. Wichtiger allerdings ist, dass die Daten die Hypothese des Lehrerurteils als einer Self-Fulfilling Prophecy stützen, denn verglichen mit dem validen positiven Lehrerurteil (+/+ ) ergeben sich keine signifikanten Unterschiede zu den Überschätzten (-/+) und umgekehrt unterscheiden sich die Koeffizienten für die Unterschätzten (+/-) kaum von denen der Gruppe mit einem validen negativen Lehrerurteil (-/-). Die Gruppen ohne explizites Lehrerurteil liegen wiederum in der Mitte.

Die schulischen Leistungen haben neben dem Lehrerurteil und der durch den IST erhobenen kognitiven Leistungsfähigkeit einen eigenständigen Einfluss auf die Wahrscheinlichkeit, das Abitur zu erreichen. Leistungen in der Vergangenheit sind Voraussetzungen des Erfolgs und beeinflussen über die subjektiven Erwartungen, ein Studium erfolgreich abschließen zu können, die Entscheidungen. Allerdings sind die Effekte der Leistung schwächer als die der Aspirationen, was die Bedeutung individueller Zielsetzungen unterstreicht. Bei der Interpretation des Effektes der Aspirationen ist zu beachten, dass diese positiv mit der sozialen Herkunft korrelieren, so dass in diesem Effekt auch das Stuserhaltungsmotiv sichtbar wird. Die soziale Herkunft hat dagegen nur einen schwachen positiven Effekt, der neben dem Stuserhaltungsmotiv von Kinder aus höheren Schichten auch darauf zurückzuführen sein dürfte, dass die Möglichkeit materieller Unterstützung durch die Eltern die Erwartung, nach dem Abitur ein Studium erfolgreich absolvieren zu können, beeinflusst. Die Bedeutung der Lebensplanung für die Entscheidung zum Abitur wird auch an dem starken negativen Effekt einer eigenen Wohnung vor einem grundsätzlich noch möglichen Abitur sichtbar. Wer sich für die Selbstständigkeit einer eigenen Wohnung entschieden hat, der hat sein Leben bereits in andere Bahnen gelenkt und entscheidet sich nur in den seltensten Fällen noch für das Abitur. Das Geschlecht bleibt ohne signifikanten Effekt.

## (b) Aufnahme eines Studiums

Bei der zweiten Entscheidung, die dem Studienerfolg voraus geht, der Studienaufnahme, sind im multivariaten Modell die Effekte für die sechs durch Lehrerurteil und kognitive Leistungsfähigkeit gebildeten Gruppen nicht mehr ganz so eindeutig wie im Falle des Abiturs. Verglichen mit dem durch überdurchschnittliche Intelligenz validierten positiven Lehrerurteil (+/+) senkt zwar die Zugehörigkeit zu jeder anderen Gruppe die Wahrscheinlichkeit der Studienaufnahme. Für prognostische Validität der kognitiven Leistungsfähigkeit sprechen immerhin die Effekte der Gruppen ohne Lehrerurteil: Überdurchschnittlich intelligente Befragte (+/0) nehmen häufiger ein Studium auf als Befragte mit einem unterhalb des Durchschnittes liegenden IST (-/0). Während der Effekt für erstere (+/0) sich nicht signifikant von dem der Gruppe mit validem Lehrerurteil (+/+) unterscheidet, ist dies aber für die zweite Gruppe (-/0) der Fall. Eine darüber hinaus gehende Ordnung ist auf den ersten Blick allerdings nicht erkennbar, so dass die Intelligenz-Hypothese und die Hypothese einer wechselseitigen Validierung nur eingeschränkt aufrechterhalten werden können.

Aber die in Tabelle 3 dargestellten Ergebnisse stützen die Hypothese einer Self-Fulfilling Prophecy, denn diese erwartet, dass die von ihren Lehrern überschätzten Befragten (-/+) ähnlich erfolgreich sind wie die Befragten mit validem positiven Urteil (+/+), während unterschätzte Befragte (+/-) deutlich seltener erfolgreich sein sollten. Genau dies zeigen die Daten.

Eine Einschränkung muss gemacht werden: Die Gruppe mit einem validen negativen Urteil (–/–) entscheidet sich wider Erwarten nicht signifikant seltener als die Gruppe mit einem validen positiven Urteil (+/+) für ein Studium. Dies könnte – wie bereits angesprochen – mit den günstigen Gelegenheitsstrukturen der Bildungsexpansion zusammenhängen. Wichtiger aber scheint aus der Lebensverlaufsperspektive aufeinander folgender Entscheidungen, dass die Effekte für die Gruppen mit negativem Lehrerurteil gegenüber dem Modell für das Abitur insgesamt schwächer sind. Dies weist darauf hin, dass ein negatives Lehrerurteil als subjektive Entscheidungsgrundlage im weiteren Verlauf an Bedeutung verliert, weil neue Erfahrungen – wie das bestandene Abitur – die Erwartung eines erfolgreichen Studiums positiv beeinflussen. Dagegen entsprechen die Effekte für die Leistungen in der Schule, die Aspirationen und die soziale Herkunft in Höhe und Tendenz etwa denjenigen, die auch in dem Modell für das Abitur ermittelt wurden. Die Effekte der privaten Verlaufsumstände entsprechen den Erwartungen.

### *Erfolg im Studium*

Ist die Entscheidung für ein Studium erst einmal gefallen, dann sollte, wie in der Kontroll-Hypothese begründet, das Lehrerurteil keinen signifikanten Einfluss mehr auf den Erfolg im Studium haben, da es vermittelt über die subjektiven Erfolgserwartungen die dem Studium vorgelagerten Entscheidungen beeinflusst und die Gruppe, die letztlich ein Studium aufnimmt, dadurch bereits selektiert ist. Wie der unterhalb der Tabelle 4 ausgewiesene niedrige Wert für McFadden's Pseudo- $R^2$  ausweist, ist die Erklärungskraft des Modells insgesamt gering.<sup>8</sup> Die Lehrerurteile haben erwartungsgemäß keinen signifikanten Effekt mehr auf den Studienerfolg. Das gilt auch für die Aspirationen, für die ein vergleichsweise starker Einfluss auf die beiden Entscheidungen im Vorfeld nachgewiesen werden konnte, die aber, nachdem die Weichen für das Studium gestellt worden sind, dessen Abschluss nicht mehr beeinflussen. Verglichen mit den klassischen Professionen Medizin und Recht bestehen auch zwischen den Studiengängen und dem Studienerfolg keine signifikanten Zusammenhänge. Allenfalls das Vorzeichen der Koeffizienten lässt vermuten, dass die in der Gruppe der Sprach-, Kultur- und Orientierungswissenschaften zusammengefassten Studiengänge seltener den Abschluss erreichen. Auch für das Geschlecht und die private Lebenssituation ist kein signifikanter Effekt zu berichten, obwohl auch hier die Vorzeichen den Erwartungen entsprechen.

8 Zur Diskussion und Bewertung dieser Maßzahl vgl. Andreß et al. (1997: 288 f.) oder Gautschi (2010: 227 f.).

Tabelle 4: Studienerfolg (Logistische Regression)

	min / max	unstandardisiert exp(b)	standardisiert exp(b*SD)
<b>Interzept</b>		13.32 *	
<b>Lehrerurteil und kogn. Leistungsfähigkeit</b>			
+/+ Valides positives LU (Referenz)			
-/+ Überschätzt	0/1	0.92	1.03 <sup>-1</sup>
+/0 IST hoch, kein LU	0/1	1.15	1.06
-/0 IST niedrig, kein LU	0/1	0.68	1.16 <sup>-1</sup>
+/- Unterschätzt	0/1	0.55	1.18 <sup>-1</sup>
-/- Valides negatives LU	0/1	1.23	1.07
<b>Leistung in der Schule und Studiengang</b>			
Notendurchschnitt Klasse 10	2.3/7.0	0.63 *	1.37 <sup>-1</sup>
Klass. Prof.: Medizin, Jura (Referenz)			
Fachhochschule / Lehramt Grundschule	0/1	1.53	1.21
Lehramt Sek. II	0/1	1.58	1.23
Sprach-, Kultur- u. Orientierungswissensch.	0/1	0.74	1.10 <sup>-1</sup>
Wirtschaft, Technik, Naturwissenschaften	0/1	0.98	1.01 <sup>-1</sup>
<b>Aspirationen</b>			
Kein Studium angestrebt (Referenz)			
Studium: Weiß noch nicht	0/1	1.01	1.01
Studium: Eventuell	0/1	1.14	1.05
Studium: Sicher	0/1	1.49	1.22
<b>Soziale Herkunft</b>			
Berufsprestige des Vaters	19/82	1.02 **	1.38
<b>Private Lebenspläne u. Situation</b>			
Geschlecht: Frau	0/1	1.36	1.16
Mit Kind	0/1	0.94	1.03 <sup>-1</sup>
Interaktion Frau*Kind	0/1	0.62	1.12 <sup>-1</sup>
N=841			
-2 LogLikelihood=663,54			
Likelihood-Ratio $\chi^2=43,5^{***}$ (17 FG)			
McFadden Pseudo-R <sup>2</sup> =,06			

LU: Lehrerurteil, SD: Standardabweichung

Notendurchschnitt des ersten Halbjahres der Klasse 10 im Klassenverband z-transformiert.

Negative standardisierte Koeffizienten werden als Kehrwert dargestellt.

\* p<.05, \*\* p<.01, \*\*\* p<.001

Signifikante Effekte haben nur zwei der kontrollierten Variablen: Erstens trägt die Durchschnittsnote der Klasse 10 unabhängig vom Lehrerurteil und der Testintelligenz zur Vorhersage des Studienerfolgs bei. Auch wenn einzelne Noten möglicherweise problematisch sind, scheinen sie doch zum Notendurchschnitt aggregiert über die gemessenen Leistungen hinaus ein guter Indikator individueller Leistungsbereitschaft und -fähigkeit zu sein. Zweitens lässt sich beim Studienerfolg ein signifikanter Effekt der sozialen Herkunft beobachten, der angesichts der kontrollierten Variablen auf im Studium verwertbares kulturelles und soziales Kapital, aber auch auf die ökonomische Absicherung durch die Herkunftsfamilie zurückzuführen sein dürfte. Darüber hinaus werden ähnliche Entscheidungskalküle wie beim Abitur und dem Studieneintritt eine Rolle spielen, denn zumindest die Entscheidung das Studium abzubrechen (vgl. Meuleman 1995: 186) dürfte auch durch das Stuserhaltungsmotiv beeinflusst werden.

### Berufserfolge mit 43 Jahren

In der bivariaten Analyse wurden deutliche Zusammenhänge zwischen dem Lehrerurteil und dem durch Wegeners Magnitude-Prestigeskala (1988) erhobenen Berufserfolg mit 43 Jahren



nachgewiesen: Das Berufsprestige liegt in den Gruppen mit einem positiven Lehrerurteil um bis zu 21 Punkte höher als in den Vergleichsgruppen ohne oder mit negativen Urteil (vgl. Tabelle 2). Wie bereits für den Studienerfolg gezeigt werden konnte, kann auch hier mit der Kontroll-Hypothese erwartet werden, dass das Lehrerurteil im multivariaten Modell, bei Kontrolle der Entscheidungen und Erfolge im Bildungsverlauf, keinen direkten Einfluss mehr auf den Berufserfolg hat. Die sehr niedrigen, nicht signifikanten Koeffizienten für die durch Lehrerurteil und kognitive Leistungsfähigkeit gebildeten Gruppen in Tabelle 5 stützen die Kontroll-Hypothese.

Tabelle 5: Berufserfolg mit 43 Jahren: OLS-Regression auf das Berufsprestige (MPS)

klein	min / max	Regressionskoeffizienten	
		unstandardisiert	standardisiert
<b>Interzept</b>		9.21	***
<b>Lehrerurteil und kognitive Leistungsfähigkeit</b>			
+/+ Valides positives LU (Referenz)			
-/+ Überschätzt	0/1	.28	.03
+/0 IST hoch, kein LU	0/1	-.28	-.03
-/0 IST niedrig, kein LU	0/1	.21	.02
+/- Unterschätzt	0/1	.31	.03
-/- Valides negatives LU	0/1	.56	.05
<b>Aspirationen</b>			
Kein Studium angestrebt (Referenz)			
Studium: Weiß noch nicht	0/1	.32	.04
Studium: Eventuell	0/1	.12	.01
Studium: Sicher	0/1	.72	*
<b>Leistungen im Bildungsverlauf</b>			
Notendurchschnitt Klasse 10	2.3/7.0	-.61	***
Kein Studium (Referenz)			
Studium nicht abgeschlossen	0/1	-.01	-.00
Abgeschlossenes Studium	0/1	2.69	***
Promotion	0/1	6.31	***
<b>Soziale Herkunft</b>			
Berufsprestige des Vaters	19/82	.03	***
<b>Erwerbsunterbrechungen in Monaten</b>			
Hausfrau / -mann	0/269	-.01	-.05
Sonstige erzwungene Unterbrechungen	0/133	-.00	-.02
<b>Private Lebenspläne u. Situation</b>			
Geschlecht: Frau	0/1	-.65	-.09
Mit Kind	0/1	.10	.01
Interaktion Frau*Kind	0/1	.71	.09
N=973			
R <sup>2</sup> =.37			

Wegeners (1988) Magnitude-Prestigeskala (MPS), zur besseren Darstellung der Koeffizienten durch 10 dividiert.

LU: Lehrerurteil

Notendurchschnitt des ersten Halbjahres der Klasse 10 im Klassenverband z-transformiert.

Sonstige erzwungene Erwerbsunterbrechungen: Arbeitslosigkeit, Krankheit.

\* p<.05, \*\* p<.01, \*\*\* p<.001

Hohe Aspirationen, deren besondere Bedeutung für die Entscheidungen auf dem Weg ins Studium (vgl. Tabelle 3) nachgewiesen wurde, die dann aber keinen Einfluss mehr auf den Studienerfolg hatten, haben nun wieder einen (geringeren) Effekt auf den Berufserfolg mit 43 Jahren. Dies dürfte darauf zurückzuführen sein, dass sich die Aspirationen zwar zunächst auf ein Studium richten, aber das Studium selber Mittel für einen angesehenen Beruf ist. Die mit Abstand stärksten Effekte haben erwartungsgemäß die Erfolge in der Bildungsbiografie, auf

die sowohl das Lehrerurteil als auch die Aspirationen gerichtet waren, also das abgeschlossene Studium und darauf aufbauend die Promotion. Der signifikante Effekt des Notendurchschnitts der 10. Klasse, der selbst bei Kontrolle des Studienerfolgs noch einen eigenständigen Effekt auf das Berufsprestige mit 43 Jahren hat, bestätigt die oben formulierte Vermutung, dass – ungeachtet der Kritik an Einzelnoten und trotz des zeitlichen Bezugs – der Notendurchschnitt als ein Indikator für Leistungsfähigkeit und -motivation interpretiert werden kann. Zugleich belegt der eigenständige Effekt des Notendurchschnitts, dass sich die Lehrerurteile nicht nur an meritokratischen Kriterien orientieren, sondern auch andere Merkmale der Schüler berücksichtigen. Dazu gehört u. a. auch deren soziale Herkunft, die das Lehrerurteil in stärkerem Maße als den Notendurchschnitt beeinflusst, weil Lehrer neben meritokratischen Kriterien – bewusst oder unbewusst – in ihren Erfolgsprognosen die mit der sozialen Herkunft variierende Ausstattung mit kulturellen, sozialen und ökonomischen Ressourcen berücksichtigen (Becker / Birkelbach 2011). Indem die Lehrer in ihren Prognosen neben Leistungsindikatoren die Bedeutung dieser Ressourcen für den Erfolg antizipieren, tragen sie zur sozialen Selektivität des Bildungswesens bei. Die durch das väterliche Berufsprestige gemessene soziale Herkunft hat auch bei Kontrolle von Studiengängen, Aspirationen, Leistungen und Bildungserfolgen noch einen signifikanten Einfluss auf das Berufsprestige. Einen hier möglicherweise wirksamen Mechanismus beschreibt Blaneck (1994), wenn er zeigt, dass insbesondere in den prestigeträchtigen klassischen Professionen Jura und Medizin häufig Prozesse der Berufs- und Statusvererbung zu beobachten sind. Vor allem aber dürfte das im Elternhaus erworbene kulturelle und soziale Kapital nicht nur im Bildungsverlauf und beim Einstieg in einen prestigeträchtigen Beruf, sondern auch in der weiteren Berufskarriere hilfreich sein (vgl. Meulemann 1995: 240ff.; Hartmann 2001). Erwerbsunterbrechungen, sowie Geschlecht und Elternschaft (auch als Interaktion mit dem Geschlecht) haben im Modell keinen eigenständigen Effekt.

## **Diskussion und Resümee: Lehrerurteile im Bildungs- und Berufsverlauf**

Betrachtet man die kognitive Leistungsfähigkeit von Schülern als ein Kriterium für die Validität von Lehrerurteilen mit prognostischem Anspruch, dann wird deutlich, dass bei einer knappen Mehrheit beides übereinstimmt und das Lehrerurteil durch die kognitive Leistungsfähigkeit validiert wird. Die vergleichsweise große Minderheit, bei der beides auseinanderfällt, aber zeigt zugleich, dass solche Prognosen teilweise auf unsicheren Füßen stehen: Auf der einen Seite werden überdurchschnittlich leistungsfähige Schülerinnen und Schüler von ihren Lehrern unterschätzt und andererseits unterdurchschnittlich leistungsfähige Schülerinnen und Schüler überschätzt. Im weiteren Verlauf aber tritt unabhängig von der kognitiven Leistungsfähigkeit auch für diese Gruppen häufig das ein, was das Lehrerurteil vorhergesagt hat. Eine solche Entwicklung entspricht genau Mertons (1948) Definition einer Self-Fulfilling Prophecy, nach der eine ursprünglich falsche Definition der Situation das Handeln derart beeinflusst, dass die auf der falschen Situationsanalyse beruhenden Erwartungen tatsächlich eintreffen. Die vorgestellten Analysen haben gezeigt, dass Lehrerurteile mit einem prognostischen Anspruch sich im Lebensverlauf zu einer Self-Fulfilling Prophecy im Sinne Mertons entwickeln können, die eine eigene Macht auf die eingeschlagene Richtung des Bildungs- und Berufsverlaufs und das Niveau des Erreichten ausüben.

Die dabei wirksamen Mechanismen lassen sich mithilfe einer mikrosoziologischen Fundierung der Lebensverlaufstheorie durch die Wert-Erwartungstheorie verstehen. Die Analysen belegen die besondere Bedeutung der subjektiven Erfolgserwartungen für die Entscheidungen über den weiteren Bildungsverlauf, die wichtige Weichen im Lebensverlauf für den langfristigen beruflichen Erfolg stellen. Bei ihren Entscheidungen für das Abitur und ein Studium berücksichtigen die Schüler nicht nur Kosten und mögliche langfristigen Erträge, sondern auch die subjektiven Wahrscheinlichkeit, diese erreichen zu können. Bei der Einschätzung der Erfolgswahrscheinlichkeit orientieren sich die Schüler mangels anderer Informationsquellen

auch an den mehr oder weniger validen Beurteilungen ihrer Leistungsfähigkeit durch ihre Lehrer. Die wurden hier zwar in einem Fragebogen erhoben, aber sie werden in ähnlicher Form auch im schulischen Alltag kommuniziert. Wie in der Hypothese zur Self-Fulfilling Prophecy erwartet, entscheiden sich die durch ihre Lehrer unterschätzten Schüler seltener für Abitur und Studium als Schüler mit vergleichbarer kognitiver Leistungsfähigkeit, aber positivem oder keinem Lehrerurteil. Ein umgekehrter Effekt lässt sich bei den überschätzten Schülern beobachten. Wo es eine kleine Abweichung gibt – beim Studienzugang der Gruppe mit validem negativem Lehrerurteil –, bieten die Gelegenheitsstrukturen der Bildungsexpansion als Periodeneffekt eine ergänzende Erklärung.

Sind die Entscheidungen im Vorfeld aber gefallen und wurde ein Studium aufgenommen, dann besteht bei Kontrolle von Leistungen, Aspirationen, sozialer Herkunft und privater Lebenssituation kein signifikanter Zusammenhang mehr zwischen dem Lehrerurteil und dem Erfolg im Studium, der die prognostische Validität des Lehrerurteils stützen könnte. Die fehlenden Effekte auch der Gruppen mit negativem Lehrerurteil stützen also die Kontrollhypothese und bestätigen noch einmal die Annahme, dass die ursprünglichen Lehrerurteile insbesondere für die Unterschätzten nicht zutreffend waren. Wie in der bivariaten Analyse gezeigt wurde, beeinflussen die Lehrerurteile langfristig den Berufserfolg aufgrund der Pfadabhängigkeiten des Lebensverlaufes. Retrospektiv sind es vor allem die richtigen Weichenstellungen, die den Berufserfolg am Ende des Beobachtungszeitraums prägen.

Inwieweit die Ergebnisse zur Validität von Lehrerurteilen auf den heutigen Schulalltag übertragen werden können, kann man sich mit Blick auf das Alter der Daten fragen. Aber es spricht nichts dagegen, dass die beschriebenen Mechanismen noch genauso wirken. Wenn Lehrer derartige Beurteilungen mit prognostischem Anspruch aussprechen, dann berücksichtigen sie dabei oft nicht nur die Leistungen der Schüler, sondern auch leistungsfremde Merkmale. Auch wenn die Beurteilungen unverbindlich sind, fließen sie in die Bildungsentscheidungen ein, indem die Schüler bzw. deren Eltern ihre Erfolgserwartungen an den Lehrerurteilen orientieren. Dadurch können prognostische Lehrerurteile die Bildungsverläufe insbesondere an kritischen Übergängen, wie dem Wechsel von der Primar- in die Sekundarstufe oder der Entscheidung, ein Studium aufzunehmen, in eine bestimmte Richtung mit langfristigen Folgen für den Lebensverlauf umlenken, auch wenn meritokratische Kriterien eine andere Richtung nahelegen würden.

## Literaturverzeichnis

- Amthauer, R. (1953): Intelligenz-Struktur-Test, 2. erw. Auflage, Göttingen.
- Andreß, H.-J. / J.A. Hagenaars / S. Kühnel (1997): Analyse von Tabellen und kategoriale Daten. Log-lineare Modelle, latente Klassenanalyse, logistische Regression und GSK-Ansatz, Berlin u.a.
- Arnold, K.-H. / W. Bos / P. Kichert / T.C. Stubbe (2007): Schullaufbahnpräferenzen am Ende der vierten Klassenstufe, in: W. Bos / S. Hornberg / K.-H. Arnold / G. Faust / L. Fried, / E.-M. Lankes / K. Schwip-pert / R. Valtin (Hrsg.), IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich, Münster u.a., S. 271-297.
- Baron-Boldt, J. / H. Schuler / U. Funke (1988): Prädiktive Validität von Schulabschlussnoten: Eine Meta-analyse, in: Zeitschrift für Pädagogische Psychologie 2, S. 79-90.
- Baron-Boldt, J. / H. Schuler / U. Funke (1989): Prognostische Validität von Schulnoten. Eine Metaanalyse der Prognose des Studien und Ausbildungserfolgs., in: R.S. Jäger / R. Horn / K. Ingenkamp (Hrsg.), Tests und Trends. 7. Jahrbuch der Pädagogischen Diagnostik, Weinheim / Basel, S. 11-39.
- Baumert, J. / M. Brunner / O. Lüdtke / U. Trautwein (2007): Was messen internationale Schulleistungsstudien? – Resultate kumulativer Wissenserwerbsprozesse. Eine Antwort auf Heiner Rindermann, in: Psychologische Rundschau 58, S. 118-128.

- Becker, D. / K. Birkelbach (2010): Intelligenz und Schulleistung als Kontextmerkmale: Big Fish – Little Pond- oder Reflected-Glory-Effekt? Eine Mehrebenen-Analyse von Lehrerurteilen, in: T. Beckers / K. Birkelbach / J. Hagenah / U. Rosar (Hrsg.), *Komparative empirische Sozialforschung*, Wiesbaden, S. 113-141.
- Becker, D. / K. Birkelbach (2011): Teachers' Evaluations and the Definition of the Situation in the Classroom (CGS Working Paper Series Vol. 2, No. 4), Cologne.
- Becker, R. (2009): Wie können „bildungsferne“ Gruppen für ein Hochschulstudium gewonnen werden?, in: *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61, S. 563-593.
- Becker, R. / A.E. Hecken (2008): Warum werden Arbeiterkinder vom Studium an Universitäten abgelenkt? Eine empirische Überprüfung der „Ablenkungsthese“ von Müller und Pollak (2007) und ihrer Erweiterung durch Hillmert und Jacob (2003), in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60, S. 3-29.
- Birkelbach, K. (1998 a): Berufserfolg und Familiengründung. Lebensläufe zwischen institutionellen Bedingungen und individueller Konstruktion, Opladen.
- Birkelbach, K. (1998 b): Befragungsthema und Panelmortalität: Ausfälle in einer Lebenslaufferhebung, in: *ZA-Informationen* 42, S. 128-147.
- Birkelbach, K. / J.O. Hellwig / W. Hemsing / H. Meulemann (1997): Success and Satisfaction in the Occupational Career from Youth to Mid-Life in a Cohort of German Gymnasium Students, 1969, 1984, and 1997. Paper prepared for the Working Group 5.9 „Constraints and Opportunities in Life Histories“ at the World Congress of Sociology of the International Institute of Sociology (IIS), July 7-11, 1997 in Cologne, Germany.
- Birkelbach, K. / J.O. Hellwig / W. Hemsing / H. Meulemann (Hrsg.) (2000): Lebenserfolg und Erfolgsdeutung im frühen Erwachsenenalter. Eine Wiederbefragung ehemaliger Gymnasiasten im 43. Lebensjahr. Teil 1: Arbeitsbericht. Projektbericht zur Vorlage bei der DFG (Az. 577 / 7-1). Köln.
- Blaneck, S. (1994): *Der Berufseinstieg in Professionen*, Berlin.
- Blossfeld, H.-P. / K. Golsch / G. Rohwer (2009): *Event History Analysis with Stata*. Hove, East Sussex.
- Bos, W. / M. Pietsch (2004): *Erste Ergebnisse aus KESS 4*. Kurzbericht, Hamburg.
- Bos, W. / A. Voss / E.-M. Lankes / K. Schwippert / J. Thiel / R. Valtin (2004): Schullaufbahneempfehlungen von Lehrkräften für Kinder am Ende der vierten Jahrgangsstufe, in: W. Bos / E.-M. Lankes / M. Prenzel / K. Schwippert / R. Valtin / G. Walther (Hrsg.), *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*, Münster u.a., S. 191-228.
- Boudon, R. (1974): *Education, opportunity, and social inequality: Changing prospects in Western Society*, New York.
- Brattesani, K. A. / R.S. Weinstein, / H.H. Marshall, (1984): Student Perceptions of Differential Teacher Treatment as Moderators of Teacher Expectation Effects, in: *Journal of Educational Psychology* 76, S. 236-247.
- Breen, R. / Goldthorpe, J.H. (1997): Explaining Educational Differentials. Towards a Formal Rational Action Theory. *Rationality and Society* 9, S. 275-305.
- Brophy, J.E. / T.L. Good, (1974): *Teacher-Student Relationships. Causes and Consequences*, New York u.a.
- Dalbert, C. / J. Stöber (2008): Forschung zur Schülerpersönlichkeit, in: W. Helsper / J. Böhme (Hrsg.), *Handbuch der Schulforschung*, 2. durchgesehene und erweiterte Auflage, Wiesbaden, S. 905-925.
- Ditton, H. / J. Krüsken (2006): Der Übergang von der Grundschule in die Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft* 9, S. 348-372.
- Erikson, R. / J.O. Jonsson (1996): Explaining class inequality in education: The Swedish test case, in: Erikson, R. / J.O. Jonsson (Hrsg.), *Can education be equalized? The Swedish case in comparative perspective*, Oxford, S. 1-63.
- Esser, H. (1993): *Soziologie. Allgemeine Grundlagen*, Frankfurt / Main – New York.

- Esser, H. (1999): Soziologie. Spezielle Grundlagen. Band 1: Situationslogik und Handeln, Frankfurt / Main – New York.
- Esser, H. (2006): Eines für alle(s)? Das Weber-Paradigma, das Konzept des moderaten methodologischen Holismus und das Modell der soziologischen Erklärung, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie 58, S. 352-363.
- Frey, M.C. / D.K. Detterman (2004): Scholastic Assessment or g? The Relationship Between the Scholastic Assessment Test and General Cognitive Ability, in: Psychological Science 15, S. 373-378.
- Gautschi, T. (2010): Maximum-Likelihood Schätztheorie, in: C. Wolf / H. Best (Hrsg.), Handbuch der sozialwissenschaftlichen Datenanalyse, Wiesbaden, S. 205-235.
- Good, T. L. / J.E. Brophy, (2003): Looking in Classrooms, New York u.a.
- Hadjar, A. / R. Becker (2009): Erwartete und unerwartete Folgen der Bildungsexpansion in Deutschland, in: R. Becker (Hrsg.), Lehrbuch der Bildungssoziologie, Wiesbaden, S. 195-213.
- Hartmann, M. (2001): Klassenspezifischer Habitus und / oder exklusive Bildungstitel als soziales Selektionskriterium? – Die Besetzung von Spitzenpositionen in der Wirtschaft, in: B. Kraus (Hrsg.), An der Spitze. Von Eliten und herrschenden Klassen, Konstanz, S. 157-215.
- Hauser, R.M. (2002): Meritocracy, Cognitive Ability, and the Sources of Occupational Success. Paper prepared for the annual meeting of the American Sociological Association, Chicago, 2002 (CDE Working Paper 98-07), CDE Working Paper.
- Heinz, W.R. (1996): Status Passages as Micro-Macro Linkages in Life Course Research, in: A. Weymann / W.R. Heinz (Hrsg.), Society and Biography. Interrelationships between Social Structure, Institutions and the Life Course, Weinheim, S. 51-65.
- Hillmert, S. (2004): Soziale Ungleichheit im Bildungsverlauf: zum Verhältnis von Bildungsinstitutionen und Entscheidungen, in: R. Becker, / W. Lauterbach (Hrsg.), Bildung als Privileg? Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit, Wiesbaden, S. 72-98.
- Huinink, J. / T. Schröder (2008): Skizzen zu einer Theorie des Lebenslaufs, in: A. Diekmann / K. Eichner / P. Schmidt / T. Voss (Hrsg.), Rational Choice: Theoretische Analysen und empirische Resultate. Festschrift für Karl-Dieter Opp zum 70. Geburtstag, Wiesbaden, S. 291-308.
- Ingenkamp, K. (Hrsg.) (1971): Die Fragwürdigkeit der Zensurengebung, Weinheim – Basel.
- Ingenkamp, K. (1989): Diagnostik in der Schule. Beiträge zu Schlüsselfragen der Schülerbeurteilung, Weinheim.
- Jussim, L. / K.D. Harber (2005): Teacher Expectations and Self-Fulfilling Prophecies: Knowns and Unknowns, Resolved and Unresolved Controversies, in: Personality & Social Psychology Review 9, S. 131-155.
- Kristen, C. (2006): Ethnische Diskriminierung in der Grundschule? Die Vergabe von Noten und Bildungsempfehlungen, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie 58, S. 79-97.
- Lehmann, R.H. / R. Peek / R. Gänsfuß (1997): LAU 5 – Aspekte der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern, die im Schuljahr 1996/97 eine fünfte Klasse an Hamburger Schulen besuchten. Bericht über die Erhebung im September 1996, Hamburg.
- Mayer, K.U. (1990): Lebensverläufe und sozialer Wandel. Anmerkungen zu einem Forschungsprogramm, in: Ders. (Hrsg.), Lebensverläufe und sozialer Wandel, in: Sonderheft 31 / 1990 der KZfSS, Opladen, S. 7-21.
- Mayer, K.U. (2004): Whose Lives? How History, Societies and Institutions Define and Shape Life Courses, in: Research in Human Development 1, S. 161-187.
- Mayer, K.U. / W. Müller (1989): Lebensverläufe im Wohlfahrtsstaat, in: A. Weymann (Hrsg.), Handlungsspielräume. Untersuchungen zur Individualisierung und Institutionalisierung von Lebensläufen in der Moderne, Stuttgart, S. 41-60.
- Merton, R.K. (1948): The Self-Fulfilling Prophecy, in: The Antioch Review 8, S. 193-210.

- Meulemann, H. (1979): Soziale Herkunft und Schullaufbahn. Arbeitsbuch zur sozialwissenschaftlichen Methodenlehre, Frankfurt – New York.
- Meulemann, H. (1995): Die Geschichte einer Jugend. Lebenserfolg und Erfolgsdeutung ehemaliger Gymnasiasten zwischen dem 15. und 30. Lebensjahr, Opladen.
- Meulemann, H. / K. Birkelbach / J.O. Hellwig, (Hrsg.) (2001): Ankunft im Erwachsenenleben. Lebenserfolg und Erfolgsdeutung in einer Kohorte ehemaliger Gymnasiasten zwischen 16 und 43, Opladen.
- Müller, W. / D. Haun (1994): Bildungsgleichheit im sozialen Wandel. Kölner Zeitschrift für Soziologie und Sozialpsychologie 46, S. 1-42.
- Müller, W. / D. Pollack / D. Reimer / S. Schindler (2009): Hochschulbildung und soziale Ungleichheit, in: Becker, R. (Hrsg.), Lehrbuch der Bildungssoziologie, Wiesbaden, S. 281-319.
- OECD (2006): Assessing Scientific, Reading and Mathematical Literacy. A Framework for PISA 2006, Paris.
- Pietsch, M. / T.C. Stubbe (2007): Inequality in the Transition from Primary to Secondary School: school choices and educational disparities in Germany, in: European Educational Research Journal 6, S. 424-445.
- Prenzel, M. / O. Walter / A. Frey (2007): PISA misst Kompetenzen. Eine Replik auf Rindermann (2006): Was messen internationale Schulleistungsstudien?, in: Psychologische Rundschau 58, S. 128-136.
- Rindermann, H. (2006): Was messen internationale Schulleistungsstudien? Schulleistungen, Schülerfähigkeiten, kognitive Fähigkeiten, Wissen oder allgemeine Intelligenz?, in: Psychologische Rundschau 57, S. 69-86.
- Rosenthal, R. (1994): Interpersonal Expectancy Effects: A 30-Year Perspective, in: Current Directions in Psychology Science 3, S. 176-179.
- Rosenthal, R. / L. Jacobson (1968): Pygmalion in the Classroom: Teacher expectation and pupils' intellectual development, New York.
- Schelsky, H. (1957): Schule und Erziehung in der industriellen Gesellschaft, Würzburg.
- Schindler, S. / D. Reimer (2010): Primäre und sekundäre Effekte der sozialen Herkunft beim Übergang in die Hochschulbildung, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie 62, S. 623-653.
- Schmidt, F.L. (2002): The Role of General Cognitive Ability and Job Performance: Why There Cannot Be a Debate, in: Human Performance 15, S. 187-210.
- Schmidt, F.L. / J.E. Hunter, (1998): The Validity and Utility of Selection Methods in Personnel Psychology. Practical and Theoretical Implications of 85 Years of Research Findings, in: Psychological Bulletin 124, S. 262-274.
- Schrader, F.-W. / A. Helmke (2001): Alltägliche Leistungsbeurteilung durch Lehrer, in: F.E. Weinert (Hrsg.), Leistungsmessungen in Schulen, Weinheim – Basel, S. 45-58.
- Schuler, H. (1998): Noten und Studien- und Berufserfolg, in: D.H. Rost (Hrsg.), Handwörterbuch Pädagogische Psychologie, 2. Auflage, Weinheim, S. 370-374.
- Solga, H. (2005): Meritokratie – die moderne Legitimation ungleicher Bildungschancen, in: P.A. Berger, / H. Kahlert, (Hrsg.), Institutionalisierte Ungleichheiten: Wie das Bildungswesen Chancen blockiert, Weinheim, S. 19-38.
- Tent, L. (2001): Zensuren, in: D. H. Rost (Hrsg.), Handwörterbuch Pädagogische Psychologie, 2. Auflage, Weinheim, S. 580-584
- Trapmann, S. / B. Hell / S. Weigand / H. Schuler (2007): Die Validität von Schulnoten zur Vorhersage des Studienerfolgs – eine Metaanalyse, in: Zeitschrift für Pädagogische Psychologie 21, S. 11-27.
- Trautwein, U. / O. Lüdtke / M. Becker / M. Neumann / G. Nagy (2008): Die Sekundarstufe I im Spiegel der empirischen Bildungsforschung: Schulleistungsentwicklung, Kompetenzniveaus und die Aussagekraft von Schulnoten, in: E. Schlemmer / H. Gerstberger (Hrsg.), Ausbildungsfähigkeit im Spannungsfeld zwischen Wissenschaft, Politik und Praxis, Wiesbaden, S. 91-107.
- Treiman, D. J. (1977): Occupational Prestige in Comparative Perspective, New York.

Wegener, B. (1988): Kritik des Prestiges, Opladen.

Wineburg, S. S. (1987): The Self-Fulfillment of the Self-Fulfilling Prophecy, in: Educational Researcher 16, S. 28-37.

Ziegenspeck, J. W. (1999): Handbuch Zensur und Zeugnis in der Schule. Historischer Rückblick, allgemeine Problematik, empirische Befunde und bildungspolitische Implikationen, Bad Heilbrunn.

PD Dr. Klaus Birkelbach  
Universität zu Köln  
Forschungsinstitut für Soziologie  
Greinstr. 2  
50939 Köln  
birkelbach@wiso.uni-koeln.de