

2 Selbsterfüllende Prophezeiungen im schulischen Kontext – Konzeption und Forschungsstand

Der Umfang an Studien zu Erwartungseffekten im Schulkontext ist immens. Schon 1978, zehn Jahre nach der Veröffentlichung der Ergebnisse von *Pygmalion in the Classroom* (Rosenthal und Jacobson 1968), konnten Rosenthal und Rubin (1978) 345 Studien zählen, die sich mit der Thematik beschäftigen. Auch in den Folgejahren riss die Flut an Studien nicht ab. Das Phänomen wurde mit der Zeit jedoch immer differenzierter betrachtet. Nicht nur die direkten Auswirkungen von Lehrererwartungen auf Intelligenzmaße und auf schulische Fähigkeiten wurden untersucht, sondern auch vor-, nach- und zwischengelagerte Aspekte des Phänomens. Die zentralen Fragestellungen in der Literatur lauten: Wie zutreffend sind Lehrererwartungen? Haben zugeschriebene und unveränderliche Eigenschaften der Schülerinnen und Schüler einen Einfluss auf die Lehrererwartungen? Können sich lehrerseitige Über- und Unterschätzungen der Leistungen von Schülerinnen und Schülern positiv beziehungsweise negativ auf deren Kompetenzentwicklung auswirken? Auf welchen Wegen werden Lehrererwartungen an die Heranwachsenden vermittelt?

Gegenwärtig erscheint die Anzahl an Publikationen, die sich mit den Ursachen von Lehrererwartungen beschäftigen, ähnlich hoch zu sein wie die Anzahl an Publikationen zu den Auswirkungen und zu Vermittlungsprozessen. Jedoch liegt derzeit keine Arbeit vor, die den Entstehungsprozess selbsterfüllender Prophezeiungen sowohl theoretisch als auch empirisch von der Entstehung der Lehrererwartungen bis zu den Auswirkungen auf den Schulerfolg von Kindern und Jugendlichen nachzeichnet. Es bleibt stets bei der Beantwortung von Teilfragen. Auch Einschätzungen einer möglichen Bedeutung des Phänomens für die Herausbildung von Bildungsungleichheiten lassen sich auf der Grundlage der vorliegenden Literatur nur implizit treffen. Bereits 1981 zogen Berka und Westhoff eine Bilanz, die im Prinzip auch auf den gegenwärtigen Forschungsstand zutrifft:

„[Es] fehlt den bisher vorliegenden empirischen Arbeiten durchweg die Bemühung um theoretische Eingliederung in eine Taxonomie von Hypothesen. Damit verbunden ist fast immer eine nicht kontinuierliche Arbeit in diesem Forschungsgebiet. Beides zusammen führt zu einem Sammelsurium an empirischen Arbeiten, die wenig Aussagekraft haben.“ (a. a. O.: S. 20)

Das Anliegen dieses Kapitels besteht deshalb nicht nur darin, einen deutschsprachigen Überblick der empirischen Befunde zu Erwartungseffekten im schulischen Kontext zu präsentieren. Anhand der Systematisierung des Forschungsstandes sollen vielmehr Schlussfolgerungen gezogen werden, die sowohl eine soziologisch-theoretische Betrachtung als auch eine umfassende empirische Untersuchung des Entstehungsprozesses selbsterfüllender Prophezeiungen erlauben. Außerdem erfolgt die Darstellung der Literatur stets vor dem Hintergrund einer Bewertung der Relevanz von Befunden für die Entstehung ethnischer Bildungsungleichheiten. Letztendlich sollen Erwartungseffekte als eine soziale Tatsache aufgefasst werden, die genau wie andere, bereits bekannte Ursachen ethnischer Bildungsungleichheit anhand der Anwendung eines einheitlichen theoretischen Frameworks erklärt und auf dieser Grundlage empirisch untersucht werden kann.

Zunächst wird dazu im Folgenden kurz das Pygmalion-Experiment vorgestellt, mit dem der Grundstein der Untersuchung von Erwartungseffekten gelegt wurde (Kapitel 2.1). Im gleichen Zug werden solche Arbeiten rezitiert, die als direkte und indirekte Reaktion auf das Experiment zu verstehen sind und die einige der zentralsten Kritikpunkte an der Studie hervorbrachten. Daraufhin folgt ein Überblick über die wichtigsten nicht-experimentellen empirischen Feldstudien (Kapitel 2.2). Dort wird im Gegensatz zu einigen bereits vorliegenden Überblicksarbeiten (z. B. Spitz 1999) kein historischer Ansatz verfolgt, der die Befunde nach ihrer zeitlichen Entstehung ordnet. Vielmehr wird zunächst deutlich gemacht, dass hinter der Entstehung von selbsterfüllenden Prophezeiungen ein dreistufiger Entstehungsprozess steht. Für jeden der drei Teilschritte dieses Prozesses wird dann systematisch das bislang vorliegende Wissen zusammengetragen.

Entsprechend bilden die drei Schritte des Entstehungsprozesses von Erwartungseffekten die Basis für die Gliederung des Abschnitts 2.2. So nimmt Abschnitt 2.2.1 die Befunde zu den Lehrererwartungen in den Blick. Abschnitt 2.2.2 widmet sich daraufhin dem Wissen zu Vermittlungsprozessen. In Abschnitt 2.2.3 stehen dann die Befunde zu den Reaktionen der Schülerinnen und Schüler im Fokus. Abschließend werden in Abschnitt 2.2.4 Faktoren erörtert, die entsprechend des aktuellen Forschungsstandes bedeutsam für die Moderation von Erwartungseffekten sind, die also zu einer Verstärkung oder Abschwächung der im Rahmen des Entstehungsprozesses ablaufenden Mechanismen führen können.

In erster Linie muss bei allen Ausführungen auf die internationale Literatur verwiesen werden. Wenn vorhanden, werden Arbeiten aus Deutschland gesondert vorgestellt. Liegen diese nicht vor, so werden Überlegungen darüber angestellt, wie sich die internationalen Befunde auf den deutschen Kontext übertragen lassen. Außerdem wird der Stand der bisherigen Forschung stets vor dem Hintergrund der Frage beleuchtet, welche Rolle das Phänomen im Rahmen der Entstehung ethnischer Bildungsungleichheiten spielen könnte.

2.1 *Pygmalion in the Classroom* und die experimentelle Erforschung selbsterfüllender Prophezeiungen

Die Idee selbsterfüllender Prophezeiungen wurde ursprünglich anhand von anekdotischen Evidenzen illustriert (Merton 1948). Die erste *empirische* Untersuchung des Phänomens fand vor dem Hintergrund der Annahme statt, dass die Erwartungen von Untersuchungsleiterinnen und -leitern von Experimenten Einfluss auf die Untersuchungsobjekte und damit auf den Ausgang von Experimenten nehmen können (Rosenthal und Fode 1963).

Die Ergebnisse der Studie konnten einen solchen Zusammenhang tatsächlich belegen. Sie waren schließlich der Anlass, der Frage nachzugehen, ob derartige Erwartungseffekte auch in gesellschaftlichen Kontexten von Bedeutung sind. So entstand die Idee, dass Lehrkräfte über ihre Erwartungen unbewusst Einfluss auf die intellektuelle Entwicklung von Schülern nehmen könnten.

Zur Überprüfung von Erwartungseffekten im Schulkontext entwarf Rosenthal ein Forschungsdesign, das mittlerweile zu einiger Berühmtheit gelangt ist. Dabei handelt es sich um das unter dem Namen *Pygmalion in the Classroom* bekannte Experiment, das dem Phänomen der sich selbst erfüllenden Prophezeiung im Kontext der Schule den Namen Pygmalion-Effekt gab (Rosenthal und Jacobson 1968). Die Studie wurde 1964 an einer öffentlichen Grundschule in San Francisco in Zusammenarbeit mit Lenore Jacobson durchgeführt, die zu diesem Zeitpunkt Direktorin der Schule war. Vier Monate vor Beginn des Schuljahres absolvierten die Kinder einen nonverbalen Intelligenztest. Die Lehrkräfte erhielten allerdings die unzutreffende Information, der Test sei entwickelt worden, um die zu erwartende akademische Leistungssteigerung der Schülerinnen und Schüler erfassen zu können. Das Forscherteam teilte den Lehrkräften außerdem mit, dass bei einigen Kindern, die zuvor zufällig und ohne Wissen des Lehrpersonals ausgewählt wurden, mit ungewöhnlich hohen Intelligenzzuwächsen für das bevorstehende Schuljahr zu rechnen sei. In Bezug auf die restlichen Kinder wurden gegenüber den Lehrkräften keine Entwicklungsprognosen abgegeben. Nach sechs, zwölf und 24 Monaten erfolgten weitere Intelligenztests. Dort zeigte sich, dass die Experimentalgruppe tatsächlich höhere Werte in den Tests erzielen konnte als die Kontrollgruppe. Während die 65 Kinder, denen hohe Erwartungen entgegengebracht wurden, auf der verwendeten Skala einen mittleren Zuwachs von zwölf Punkten erfuhr, fiel der Zuwachs unter den 255 Kindern, denen keine Erwartungen entgegengebracht wurden, mit durchschnittlich acht Punkten signifikant geringer aus.

Rosenthal und Jacobson sahen sich durch diese Ergebnisse in ihrer Vermutung bestätigt, dass Erwartungen auch im Alltag und gesellschaftlich relevanten Situationen selbsterfüllend sein können. Entsprechende Aufmerksamkeit erhielt die Studie in der Forschung. Die Veröffentlichung des Buches *Pygmalion in the*

Classroom (Rosenthal und Jacobson 1968) zählte zeitweilig zu den am häufigsten zitierten Studien der Bildungsforschung. Im Rahmen dieser Zitationen wurden allerdings auch einige Kritikpunkte laut, von denen die zentralsten im Folgenden zusammengetragen werden sollen.

Unmittelbare Kritik

Kritisiert wurden zunächst die Intelligenzmessungen des Experiments. Thorndike (1968) betont, dass die Testwerte zum Zeitpunkt des ersten Tests einerseits so niedrig und andererseits so unterschiedlich waren, dass die Messungen vor allem für die sehr jungen Kinder relativ wertlos gewesen sein sollten (a. a. O.: 710).³ Dieses Argument greift auch Snow (1969) auf, der in erster Linie auf ungewöhnlich hohe Zuwächse auf der für die Messung der Intelligenz verwendeten Skala im Zeitverlauf verweist.⁴ Vor dem Hintergrund des üblicherweise als relativ stabil angesehenen Konstrukts der Intelligenz erscheint dieser Kritikpunkt als besonders schwerwiegend (Jussim und Harber 2005: S. 136 ff.).

Weitere Kritik wurde an der Tatsache geübt, dass Rosenthal und Jacobson gegenüber den Lehrkräften darauf verwiesen hatten, dass der angebliche Test zur Identifikation von Hochbegabten durch die National Science Foundation gefördert sei. Das Lehrpersonal hätte dadurch ein besonderes Verantwortungsgefühl gegenüber den vermeintlichen Hochbegabten spüren können. Da die Lehrkräfte die Intelligenztests zu Anfang des Schuljahres auch noch selbst an den Kindern durchgeführt hatten und sich unter Umständen mit den Anforderungen der Tests vertraut gemacht haben könnten, sei es möglich, dass die hohen Intelligenzzuwächse in der Experimentalgruppe die Folge einer lehrerseitigen Beeinflussung der Tests selbst war (Pellegrini und Hicks 1972).⁵

Anlass für weitere Zweifel gab der Befund des Pygmalion-Experiments, dass signifikante Unterschiede in der Intelligenzentwicklung zwischen Experimental-

3 Zum ersten Messzeitpunkt besuchten einige der Kinder noch den Kindergarten.

4 Auf die Wiederveröffentlichung von *Pygmalion in the Classroom* im Jahr 1992 reagierte Snow (1995) mit einer Re-Analyse der Originaldaten. Die Ergebnisse belegen, dass signifikante Unterschiede im Intelligenzzuwachs zwischen Experimental- und Kontrollgruppe verschwinden, wenn diejenigen Kinder aus den Daten entfernt werden, die in einer der beiden Messungen Werte außerhalb der Norm des Instruments aufwiesen.

5 Pellegrini und Hicks (1972) unterstrichen ihre Vermutungen mit einem eigenen Experiment. Das Design glich dem des Pygmalion-Experiments bis auf die Tatsache, dass die Autoren die Experimentalgruppe danach unterteilten, ob die Lehrkräfte mit dem Testmaterial vertraut gemacht wurden oder nicht. Es zeigte sich, dass die Experimentalgruppe nur dann signifikant höhere Intelligenzzuwächse gegenüber der Kontrollgruppe aufwies, wenn die Lehrkräfte mit dem Testmaterial vertraut gewesen waren.

und Kontrollgruppe in der erwarteten Richtung ausschließlich in den Klassenstufen eins und zwei gefunden werden konnten (Jussim und Harber 2005: S. 134; Spitz 1999: S. 204). Während sich in den Klassenstufen drei und vier kein Erwartungseffekt finden ließ, zeigte sich in den Klassenstufen fünf und sechs, dass die Experimentalgruppe hier sogar geringere Intelligenzzuwächse erzielte als die restlichen Kinder.⁶

In die Kritik gerieten die Schlussfolgerungen Rosenthals außerdem durch die Veröffentlichung der Ergebnisse einer Lehrerbefragung, die am Ende des Experiments durchgeführt wurde. Ein Großteil der 18 Lehrerinnen und Lehrer, die an dem Experiment teilnahmen, hatten keinerlei Erinnerung an die Namen der vermeintlichen Hochbegabten in ihrer Schulklasse. Von 72 Kindern aus der Experimentalgruppe wurden lediglich 18 namentlich richtig bestimmt (Spitz 1999: S. 206 f.). Den Ergebnissen wurde damit eine gewisse Zufälligkeit unterstellt. Ob diese Kritik zutrifft, oder ob der Befund der Tatsache geschuldet ist, dass Erwartungen unbewusst entstehen und kommuniziert werden, bleibt allerdings offen.

Replikationen und Kontroversen

Neben Re-Interpretationen und Re-Analysen des Pygmalion-Experimentes entstanden zahlreiche Replikationen der Studie. Einige empirische Untersuchungen konnten die Ergebnisse des Original-Experiments replizieren, andere nicht (vgl. Elashoff und Snow 1971; Spitz 1999: S. 212 f.).

Der Überblick über derartige Studien macht deutlich, dass Erwartungseffekte wahrscheinlich weniger wie ursprünglich angenommen auf die Intelligenz, sondern eher auf die Entwicklung schulischer Leistungen und Kompetenzen wirken (Elashoff und Snow 1971). Entsprechend verzichtet Rosenthal, der sich im Verlauf der Jahrzehnte zahlreicher Kritikpunkte annahm (1975, 1987, 1994, 1995, 2003), seit einem Überblick über insgesamt 242 Studien aus den 1970er Jahren darauf, bei selbsterfüllenden Prophezeiungen in der Schule von Effekten auf die Intelligenz zu sprechen (Rosenthal 1973).

6 Als Reaktion auf diese Kritik wurden die Ergebnisse so interpretiert, dass Lehrererwartungen die Kompetenzentwicklung vor allem bei den jüngeren Kindern beeinflussen. In aktuelleren Studien finden sich Hinweise für eine Bestätigung dieser Vermutung (vgl. Jussim, Robustelli and Cain 2009).

Meta-analytische Reviews

Die immer weiter wachsende Zahl an empirischen Studien zu Erwartungseffekten machte schließlich auch übergreifende Auswertungen der Ergebnisse verschiedener experimenteller Studien anhand von Meta-Analysen möglich. Bei diesem methodischen Ansatz werden die Ergebnisse unabhängiger Studien mit der gleichen Fragestellung zusammenfassend statistisch ausgewertet. Das Ziel entsprechender Studien besteht darin, Aufschluss in Hinblick auf die Evidenz, Generalisierbarkeit und den Umfang von Erwartungseffekten zu liefern.

Die Ergebnisse von Meta-Analysen zeigen, dass der Einfluss von Lehrererwartungen auf die Intelligenz umso stärker ist, je kürzer der Kontakt zwischen Lehrkräften und Schülerinnen und Schülern vor der experimentellen Manipulation der Erwartungen war. Die stärksten Erwartungseffekte treten außerdem in den ersten und zweiten Klassen auf (Rosenthal 1987; Raudenbush 1984). Ein meta-analytischer Vergleich verschiedener *outcomes* macht deutlich, dass Erwartungseffekte auf die intellektuellen Fähigkeiten der Schulkinder eine geringere Stärke aufweisen als Erwartungseffekte auf die schulischen Leistungen (Smith 1980). Dieselbe Studie belegt außerdem, dass manipulierte Lehrererwartungen stärkere Auswirkungen auf Leseleistungen als auf Mathematikleistungen haben.

Die meta-analytischen Studien entkräften die Kritik an der Pygmalion-Studie zum Teil, indem sie zeigen, dass sich selbsterfüllende Prophezeiungen auch über viele Studien hinweg konsistent nachweisen lassen. Außerdem liefern sie Hinweise dafür, dass unterschiedliche Ergebnisse der verschiedenen Replikationen der Pygmalion-Studie nicht zwingend als inkonsistent einzustufen sind, sondern eher aus unterschiedlichen Versuchsanordnungen resultieren und somit lediglich bestimmte Muster in Bezug auf die Entstehungsbedingungen von selbsterfüllenden Prophezeiungen widerspiegeln. Dies betrifft zum Beispiel die Kritik, dass sich Erwartungseffekte im Pygmalion-Experiment nur für die untersten Klassenstufen belegen lassen. Dass sich dieses Muster auch in anderen Experimenten findet, deutet weniger auf ein mangelndes Untersuchungsdesign der Pygmalion-Studie als vielmehr auf die Annahme hin, dass sich die Meinungen und Einschätzungen der Lehrerschaft in Bezug auf ihre Schülerinnen und Schüler zu früheren Zeitpunkten der Schullaufbahnen der Kinder einfacher beeinflussen lassen als zu späteren Zeitpunkten, weil die Lehrkräfte dann noch kein gefestigtes Bild von den Kindern verankert haben.

Zwischenfazit

Was lässt sich dem Pygmalion-Experiment, seiner Kritik und den zahlreichen Replikationsversuchen entnehmen? Grundsätzlich scheinen von Lehrererwartungen selbsterfüllende Prophezeiungen ausgehen zu können. Allerdings gibt es Hinweise dafür, dass das Phänomen im schulischen Kontext weniger die Intelligenz als vielmehr die Entwicklung schulischer Fähigkeiten und Kompetenzen betrifft. Schließlich besteht Grund zur Annahme, dass Erwartungseffekte umso stärker auszufallen, je früher im Bildungsverlauf sie manipuliert werden.

2.2 Der Entstehungsprozess selbsterfüllender Prophezeiungen im schulischen Kontext und die Bedeutung von Feldstudien

Den meta-analytischen Befunden zufolge sollten keine Zweifel daran bestehen, dass verzerrte Lehrererwartungen in experimentellen Settings selbsterfüllende Prophezeiungen erzeugen können. Diese Befunde eignen sich zwar dazu, Einschätzungen über die Konsequenzen von experimentell verzerrten Lehrererwartungen zu treffen. Es können jedoch keine ökologisch validen Schlüsse über die Häufigkeit und das Ausmaß von natürlich auftretenden Erwartungseffekten gezogen werden. Insbesondere geben die bisher aufgeführten Untersuchungen keinerlei Aufschluss darüber, ob Lehrkräfte im realen Schulalltag tatsächlich unterschiedlich hohe Erwartungen an verschiedene Schülerinnen und Schüler oder (ethnische) Schülergruppen stellen. Auch bleibt offen, in welchem Ausmaß die Einschätzungen von Lehrkräften auf einer zutreffenden Einschätzung tatsächlicher Kompetenzen und Potentiale der Kinder gründen oder nicht.

Vor allem der letztgenannte Aspekt ist von nicht zu unterschätzender Bedeutung. Denn sollten Lehrererwartungen unter natürlichen Bedingungen üblicherweise den Fähigkeiten der Schülerinnen und Schüler entsprechen oder sich diesen anpassen, so würden die experimentellen Hinweise erheblich an Relevanz für die Bildungsforschung verlieren. Schließlich können der Definition von selbsterfüllenden Prophezeiungen gemäß nur solche Erwartungen die Kompetenzen verändern, die dem tatsächlichen Zustand nicht schon von Anfang an entsprechen: „[...] the self-fulfilling prophecy is [...] a false definition of the situation, evoking a new behavior which makes the original false conception come true“ (Merton 1948: S. 195). Auf den Anwendungsfall bezogen heißt dies, dass Lehrererwartungen nur dann Veränderungen von Schülerleistungen hervorrufen können, wenn sie für die Kinder eine Leistungsentwicklung vorhersagen, die unabhängig von den Erwartungen der Lehrkräfte anders ausfallen würde. Lehrererwartungen, die diesem Kriterium entsprechen, werden im Folgenden als *unzutreffend* bezeichnet.

Dass *zutreffende* Lehrererwartungen *nicht selbsterfüllend* sein können, sei an einem Beispiel verdeutlicht. Ein Lehrer erwartet zu Beginn des ersten Schuljahres von einer Erstklässlerin, dass sie im späteren Verlauf der Grundschulzeit sehr gute Leistungen erbringen wird. Gleichzeitig erzielt die Schülerin in der ersten Klasse sehr gute schulische Leistungen. Außerdem genießt sie umfangreiche Unterstützungsleistungen von Seiten ihrer Familie und verfolgt selbst stets das Ziel zu den Klassenbesten zu gehören. Die Erwartung des Lehrers ist in diesem Fall zutreffend, weil sich in ihr leistungsrelevante Eigenschaften und das Potential des Kindes widerspiegeln. Gleichzeitig sollte die Lehrererwartung die Leistungsentwicklung nicht beeinflussen. Denn eine Förderung und Unterstützung des Kindes im Unterricht entsprechend der beschriebenen Erwartung sollte das Ziel des Kindes, sehr gute Leistungen zu erzielen, unterstützen, aber keine Veränderung dieses Ziels und des darauf ausgerichteten Verhaltens des Kindes bewirken.

Würde der Lehrer dagegen von derselben Schülerin geringere schulische Erfolge erwarten, so wäre diese Erwartung in Anbetracht der tatsächlichen Fähigkeiten des Kindes negativ verzerrt. Eine Unterforderung des Kindes von Seiten des Lehrers im Unterricht entsprechend dieser Erwartung könnte in einer selbsterfüllenden Prophezeiung münden. Dies wäre der Fall, wenn die Schülerin aufgrund des Verhaltens ihres Lehrers weniger gefördert oder unterfordert wird und/oder mit der Zeit von ihrem Ziel absieht zu den Klassenbesten zu gehören und dadurch im Vergleich zu ihren ursprünglichen Potentialen und Absichten geringere Schulleistungen erzielt.

Dieses Beispiel macht deutlich, dass neben dem Vorliegen unzutreffender Erwartungen eine weitere Bedingung gegeben sein muss, damit es zu Erwartungseffekten im Schulkontext kommt: Lehrererwartungen müssen im Unterrichtskontext an die Kinder kommuniziert werden. Das bedeutet, dass sich unterschiedlich hohe beziehungsweise mehr oder weniger zutreffende Lehrererwartungen in voneinander abgrenzbaren und dauerhaft auftretenden Verhaltensweisen der Lehrkräfte gegenüber den Schülerinnen und Schülern niederschlagen müssen. Schließlich besteht der eigentliche Effekt selbsterfüllender Prophezeiungen in einer Wirkung auf das Lernen der Kinder. Für den Nachweis selbsterfüllender Prophezeiungen muss also belegt werden, dass Interaktionsmuster der Lehrkräfte im Unterricht, welche in der Folge unterschiedlich hoher und mehr oder weniger zutreffender Erwartungen an verschiedene Kinder differenziert ausfallen, entsprechende Unterschiede in Bezug auf den Kompetenzerwerb der einzelnen Schülerinnen und Schüler bewirken.

Auf der Grundlage dieser Überlegungen kann der Entstehungsprozess selbsterfüllender Prophezeiungen wie in Abbildung 1 dargestellt als eine Abfolge dreier Teilschritte verstanden werden. An erster Stelle stehen verzerrte Lehrererwartungen. Daraufhin erfolgt die Vermittlung der Erwartungen über das Verhalten der

Lehrkräfte. Schließlich treten leistungsbezogene Auswirkungen auf Seiten der Kinder auf.

Abbildung 1: Der Prozess der Entstehung einer selbsterfüllenden Prophezeiung in drei Schritten.



Die seit den 1970er Jahren zahlreich vorgelegten konzeptionellen Modelle der Entstehung von Erwartungseffekten gleichen diesem Modell, auch wenn Unterschiede in der Bezeichnung der zentralen Schritte und in Bezug auf die Komplexität, zum Beispiel bei der Berücksichtigung von Rückkopplungsprozessen zwischen den Schritten, bestehen (vgl. Brophy und Good 1970; Braun 1976; Ferguson 2003; Jussim 1986; Rosenthal 1981; Harris und Rosenthal 1985).

Da experimentelle Studien insbesondere über die ersten beiden Teilschritte des Entstehungsprozesses selbsterfüllender Prophezeiungen keine Auskunft geben, werden die zentralen Ergebnisse empirischer Feldstudien im Folgenden anhand der Abarbeitung dreier Teilfragen zusammengetragen. Im *ersten* Schritt geht es darum zu klären, ob Lehrererwartungen unter natürlichen Bedingungen verzerrt sind und welche Schülerinnen und Schüler beziehungsweise Schülergruppen davon betroffen sind (Abschnitt 2.2.1). Im *zweiten* Schritt rücken Kommunikationsprozesse in das Blickfeld (Abschnitt 2.2.2): Auf welchem Weg und über welches Verhalten vermitteln Lehrkräfte ihre Erwartungen? Im *dritten* Schritt werden schließlich die Wahrnehmung differentiellen Lehrerverhaltens und die Folgen für die Motivation und das Lernen der Schülerinnen und Schüler betrachtet (Abschnitt 2.2.3): Nehmen Kinder und Jugendliche das durch unterschiedlich hohe Erwartungen beeinflusste differentielle Verhalten der Lehrkräfte wahr? Führt dies zu einer Bestätigung der ursprünglichen, gegebenenfalls unzutreffenden Erwartungen? Über welche Mechanismen werden Auswirkungen auf das schulische Lernen vermittelt? Schließlich berichtet Abschnitt 2.2.4 von moderierenden Faktoren, die dazu führen können, dass Erwartungseffekte mit einer niedrigeren oder höheren Wahrscheinlichkeit auftreten.

2.2.1 Befunde zu Verzerrungen in Lehrererwartungen

Begriffsbestimmung

Trotz der umfangreichen Erforschung von Lehrererwartungen lassen sich kaum Definitionen des Begriffs finden. Am häufigsten wird in der Literatur der Terminus „teacher expectancies“ verwendet, häufig in synonyme Verwendung mit „teacher perceptions“, jedoch ohne weitere Erläuterung. Eine der wenigen vorliegenden Begriffsbestimmungen liefern Dusek und Joseph (1983). Ihnen zufolge umfassen Lehrererwartungen lehrerseitige Überzeugungen hinsichtlich der Schulleistungen, Fähigkeiten, Begabungen und des möglicherweise erreichbaren Schulabschlusses individueller Schülerinnen und Schüler.

Diese Definition enthält nützliche Anhaltspunkte für eine präzise Konzeption von Lehrererwartungen. An erster Stelle macht sie deutlich, dass es sich um Annahmen oder Überzeugungen hinsichtlich der Ausprägungen schulleistungsrelevanter Merkmale individueller Schülerinnen und Schüler handelt. Die Definition deutet außerdem darauf hin, dass Lehrererwartungen danach unterschieden werden können, auf welche Merkmale von Kindern und Jugendlichen sie sich beziehen. Des Weiteren legen Dusek und Joseph nahe, dass sich Lehrererwartungen sowohl auf gegenwärtige Merkmalsausprägungen als auch auf solche, die in der Zukunft vorliegen werden, beziehen können. Da die Autoren Lehrererwartungen als Überzeugungen im Sinne des englischsprachigen Begriffs „beliefs“ verstehen, ist schließlich davon auszugehen, dass Lehrkräfte im Rahmen der Entstehung ihrer Erwartungen auf Informationen zurückgreifen, die in der gegebenen Situation vorliegen und die als Signale zur Bestimmung der „wahren“ Ausprägungen der schulleistungsrelevanten Merkmale bei den Kindern dienen. Da diese Informationen begrenzt, ungenau und unzuverlässig sein können, sollten Lehrererwartungen stets ein gewisses Potential an Unsicherheit und Ungenauigkeit bergen.

Allerdings kann auch angenommen werden, dass es sich bei Lehrererwartungen um ein latentes Konstrukt im Sinne einer Disposition zur Beurteilung beziehungsweise Vorhersage des zukünftigen Bildungserfolgs einzelner Kinder und Jugendlicher handelt. Die von Dusek und Joseph genannten Schülermerkmale, auf die sich Lehrererwartungen beziehen können, sind diesem Verständnis nach nichts weiter als unterschiedliche Indikatoren, anhand derer sich das latente Merkmal der Lehrererwartung messen lässt (Brophy und Good 1970). Hier kommen schulische Leistungen, das Talent und die Begabungen sowie zukünftig erreichbare Noten und Schulabschlüsse, aber auch leistungsrelevante Merkmale wie Aufmerksamkeit, Motivation, Fleiß, Disziplin und die soziale Kompetenz in Frage. Auf der

Grundlage dieser Überlegungen lassen sich Lehrererwartungen wohl am treffendsten als latentes Wahrscheinlichkeitsurteil in Hinblick auf den globalen Schulerfolg einzelner Schülerinnen und Schüler definieren (a. a. O.: S. 33).

Definition von zutreffenden, verzerrten und systematisch verzerrten Lehrererwartungen

Leistungserwartungen von Lehrkräften bilden die Grundlage pädagogischen Handelns im Unterricht (Schrader und Helmke 2001). Die während des Lehramtsstudiums vermittelten diagnostischen Fertigkeiten und Fähigkeiten sollten es Lehrerinnen und Lehrern ermöglichen, ähnliche Überzeugungen darüber zu teilen, welche Schülermerkmale und -Verhaltensweisen als schulleistungsrelevant einzustufen sind und in Leistungserwartungen eingehen sollten. Außerdem sollten leistungsbezogene Urteile grundsätzlich mit dem Ziel einer *zutreffenden* Einschätzung gefällt werden. Der pädagogischen Literatur gemäß wäre dieses Ziel erreicht, wenn Schülermerkmale auf der Basis adäquater Einschätzungen individueller Lern- und Aufgabenanforderungen angemessen beurteilt werden würden (vgl. Artelt und Gräsel 2009).

Auf der anderen Seite gelten Maße von Lehrererwartungen in der Literatur meistens dann als *unzutreffend* oder *verzerrt*, wenn eine Diskrepanz zwischen testbasierten Leistungsmessungen und lehrerseitigen Einschätzungen vorliegt (Ready und Wright 2011: S. 338 f.). Erwartet beispielsweise eine Lehrkraft kurz nach Schulbeginn, dass eine Schülerin am Ende des Schuljahres Leistungen im Fach Mathematik erzielt, die im Klassendurchschnitt liegen, dann wäre diese Einschätzung entsprechend verzerrt, wenn die Schülerin in einem altersgerechten Mathematiktest (der ebenfalls zu Beginn des Schuljahres stattfindet) ein Ergebnis erzielt, das über dem Klassendurchschnitt liegt. Maße von Lehrererwartungen können folglich sowohl zutreffend (Akkuratheit) also auch positiv (Überschätzung) oder negativ (Unterschätzung) verzerrt (Inakkuratheit) sein.

Sollen Erwartungseffekte wie in der vorliegenden Arbeit mit ethnischen Bildungsungleichheiten in Verbindung gebracht werden, so reicht es allerdings nicht zu belegen, dass Lehrererwartungen verzerrt sind. Vielmehr ist dazu der Nachweis *systematischer Verzerrungen* nach der ethnischen Herkunft notwendig. Dies wäre der Fall, wenn Lehrkräfte die durchschnittlichen Fähigkeiten einer Schülergruppen mit einer bestimmten ethnischen Herkunft über- oder unterschätzen. Denn Erwartungen, die gegenüber einzelnen Schülerinnen oder Schülern irrtümlicherweise zu hoch oder zu niedrig ausfallen, können zwar zu individuellen Vor- oder Nachteilen führen, nicht jedoch Effekte auf Gruppenebene bewirken. Soll zum Beispiel belegt werden, dass selbsterfüllende Prophezeiungen schulleistungsbezogene Disparitäten von Kindern mit einem türkischen Zuwanderungshintergrund

im Vergleich zu einheimischen Kindern vergrößern, so muss zunächst die Frage beantwortet werden, ob die durchschnittlichen Leistungserwartungen an die türkischstämmige Schülergruppe negativ verzerrt sind, während die Erwartungen an die Gruppe ohne Zuwanderungshintergrund weniger negativ verzerrt, zutreffend oder positiv verzerrt sind.

Die folgende Zusammenfassung der Forschung zu Lehrererwartungen erfolgt deshalb mit Blick auf die Frage nach systematischen Verzerrungen nach der ethnischen Herkunft der Kinder. Rezitiert werden außerdem Befunde, die auf Effekte sozioökonomischer Herkunft der Kinder auf Lehrererwartungen hindeuten. Dies liegt darin begründet, dass Zugewanderte, die im Rahmen der Arbeitsmigration und des Familiennachzugs in der zweiten Hälfte des 20. Jahrhunderts nach Deutschland kamen, im Vergleich zur Mehrheitsbevölkerung überproportional häufig den unteren sozialen Schichten angehören und deshalb entsprechend häufig von möglicherweise auftretenden Verzerrungen der Lehrererwartungen nach der sozialen Herkunft betroffen sein könnten.

Empirische Befundlage zu systematischen Verzerrungen der Lehrererwartungen

Wie in den vorherigen Abschnitten angedeutet wurde, fallen Erwartungseffekte in den meisten empirischen Studien relativ klein aus. In einem aktuellen Überblick wird die durchschnittliche Effektstärke mit $r = 0,17$ ausgewiesen. Als Hauptursache für diesen Befund gilt, dass Maße von Lehrererwartungen die schulischen Fähigkeiten der Kinder häufig zutreffend vorhersagen (für einen Überblick siehe Brophy 1983; Jussim et al. 2009b; Ready und Wright 2011). Beispielsweise wurde belegt, dass zu Beginn des sechsten Schuljahres gemessene lehrerseitige Leistungserwartungen stark mit den Ergebnissen aus unabhängigen Leistungstests (die ebenfalls zu Beginn des Schuljahrs durchgeführt wurden) korrelieren (Jussim et al. 1996; Madon et al. 1998).

Kritisch angemerkt werden muss in diesem Zusammenhang allerdings, dass Lehrerurteile in den meisten Studien zu einem Zeitpunkt erhoben wurden, zu dem die Lehrkräfte ihre Schülerschaft bereits seit längerer Zeit kannten. Die in den betreffenden Studien gemessenen Kompetenzen könnten bereits durch vorausgegangene selbsterfüllende Prophezeiungen beeinflusst worden sein. Ebenso könnten die Lehrkräfte ihre Erwartungen an die beobachteten Schülermerkmale angepasst haben. In welche Richtung Anpassungen stattgefunden haben, lässt sich ausgehend von den verfügbaren Messungen nicht feststellen. Erst seit Kurzem liegen Belege dafür vor, dass Lehrererwartungen auch gegenüber Erstklässlerinnen und Erstklässlern, die nicht auf eine lang andauernde Interaktion mit ihren Lehrkräften zurückblicken können, zu einem bedeutsamen Anteil zutreffend sind (Hinnant et al. 2009; Sorhagen 2013).

Gleichzeitig zeigt eine aktuelle Studie, die sich zwar nicht mit Lehrererwartungen, dafür aber mit Leistungsbeurteilungen beschäftigt, dass diese ein nicht unbeachtliches Potential an Ungenauigkeit in sich bergen. Die Meta-Analyse weißt die mittlere Korrelation zwischen den Beurteilungen durch die Lehrkräfte und den tatsächlich erbrachten Schülerleistungen mit $r = .63$ aus (Südkamp et al. 2012). Ob der durch die tatsächlichen Schülerleistungen nicht erklärbare Anteil der Varianz in den Leistungsbeurteilungen eher unsystematisch zustande kommt oder ein systematischer Zusammenhang mit nicht veränderlichen Eigenschaften der Kinder und Jugendlichen wie der ethnischen Herkunft vorliegt, konnte jedoch nicht ermittelt werden.

Allerdings finden sich zahlreiche Belege dafür, dass die Leistungen *bestimmter Zuwanderungsgruppen* gemessen an objektiven Leistungen systematisch unterschätzt werden (Baron et al. 1985; Jussim und Harber 2005; Jussim et al. 1996; Madon et al. 1998; McKown und Weinstein 2008; Ready und Wright 2011; Tenenbaum und Ruck 2007). Auch Schülerinnen und Schüler aus sozial schwachen Familien werden im Vergleich zu Kindern aus besser gestellten Familien häufig unterschätzt (Alvidrez und Weinstein 1999; Baron et al. 1985; Dusek und Joseph 1983; Jussim et al. 1996; Madon et al. 1998). Diese beiden Herkunftsmerkmale (soziale und ethnische Herkunft) scheinen konfundiert auf die Entstehung von Lehrererwartungen zu wirken (Jussim et al. 1996: S. 313 ff.).

Auch experimentelle Studien liefern Hinweise für einen Zusammenhang zwischen askriptiven Schülermerkmalen und Lehrererwartungen. Diese arbeiten häufig mit Vignetten. Dem Lehrpersonal werden beispielsweise gleiche Leistungen (z. B. in Form eines immer gleichen Essays oder der Lösung einer Mathematikaufgabe) mit unterschiedlichen Schülervornamen präsentiert (Anderson-Clark et al. 2008; van Ewijk, Reyn J.G. 2010; Harari und McDavid 1973). Andere Studien verwenden Bilder von Schülerinnen und Schülern mit einer scheinbar unterschiedlichen Herkunft als Stimuli und kombinieren diese mit fiktiven Schulleistungen (Parks und Kennedy 2007). Beide Methoden zeigen, dass ein scheinbar bestehender Zuwanderungshintergrund durchschnittlich mit niedrigeren Leistungsbewertungen einhergeht. Aber auch von der sozialen Herkunft der Kinder scheinen deutliche Effekte auszugehen (Figlio 2005).

Andere Arbeiten belegen, dass Lehrerinnen und Lehrer höhere Erwartungen an Schulkinder richten, die die gleiche ethnische Zugehörigkeit wie sie selbst aufweisen (Beady und Hansell 1981; Ehrenberg et al. 1995; Dee 2005). Der gleiche Zusammenhang zeigt sich in Bezug auf den sozioökonomischen Status von Lehrkräften und Schulkindern (ebd.). Allerdings konnte neuerdings auch gezeigt werden, dass der Einfluss der Übereinstimmung zwischen Schüler- und Lehrerherkunft davon abhängt, welche ethnische Herkunft die Lehrkraft hat. So scheinen afro-amerikanische Schülerinnen und Schüler zwar von anglo-amerikanischen

Lehrkräften schlechter bewertet zu werden als ihre Mitschülerinnen und Mitschüler; gleichzeitig scheinen afro-amerikanische Schülerinnen und Schüler jedoch nicht zu profitieren, wenn sie von afro-amerikanischen Lehrkräften beurteilt werden (McGrady und Reynolds 2013).

Das Ausmaß an ethnischen Verzerrungen in den Lehrererwartungen scheint außerdem mit dem Umfang an alltäglichen Erfahrungen mit bestimmten Schülergruppen zu variieren. Zum Beispiel wurde gezeigt, dass Lehrerurteile in ethnisch homogenen Stadtteilen eher zutreffend ausfallen als in anderen Stadtteilen (Jussim et al. 1996: S. 328 ff.). Gleichzeitig liegen Belege dafür vor, dass Verzerrungen der Lehrererwartungen nach der ethnischen Herkunft häufiger in sozial schwächeren Kontexten auftreten (Ready und Wright 2011). Eingehende Untersuchung zu Kontexteffekten auf das Ausmaß an systematischen Verzerrungen in den Lehrererwartungen stehen aber noch aus. Dies betrifft insbesondere eine Untersuchung entsprechender Effekte auf der Grundlage theoretisch fundierter Annahmen.

Mehr bekannt ist über Unterschiede im Ausmaß an systematischen Verzerrungen zwischen verschiedenen ethnischen Herkunftsgruppen. Da der Großteil der Forschung aus den USA stammt, steht in den bislang aufgeführten Studien häufig der Vergleich zwischen anglo- und afro-amerikanischen Kindern im Fokus. Gelegentlich werden auch latein-amerikanische und asiatische Herkunftsgruppen einbezogen. Dabei sind zum Teil erhebliche Unterschiede je nach untersuchter Herkunftsgruppe sichtbar. Die Erwartungen an afro-amerikanische Schülerinnen und Schüler sind durchgängig stärker negativ verzerrt als die Erwartungen an anglo-amerikanische Jugendliche. Die latein-amerikanische Gruppe scheint im Vergleich zur Mehrheitsbevölkerung der USA ebenfalls unterschätzt zu werden, während negative Verzerrungen der Erwartungen an Kinder asiatischer Herkunft in geringerem Umfang auftreten (Ready und Wright 2011).

Lehrerurteile im deutschen Kontext

Für den deutschen Kontext liegen Hinweise dafür vor, dass Lehrkräfte die Fähigkeiten von Kindern mit einem Zuwanderungshintergrund unterschätzen. So zeigt sich, dass Lehrerurteile in Bezug auf die Lesefähigkeiten von Kindern der dritten Klassenstufe nicht nur mit unabhängigen Testergebnissen im Lesen zusammenhängen, sondern darüber hinaus signifikante Effekt des sozialen Status der Eltern bestehen (Stahl 2007). Jugendliche aus sozial schwächeren Familien erhalten schlechtere Bewertungen in Bezug auf die Lesefähigkeit und die sprachliche Ausdrucksfähigkeit. In Bezug auf die Einschätzung der allgemeinen Begabung lässt sich unter Kontrolle von Maßen der sozialen Herkunft außerdem ein Einfluss des Zuwanderungshintergrundes nachweisen. Da dieser anhand der Muttersprache

operationalisiert ist (deutsch vs. nicht-deutsch), bleibt unklar, ob Unterschiede in Bezug auf die Beurteilung verschiedener Herkunftsgruppen bestehen.

Auch für Jugendliche der zehnten Klassenstufe zeigt sich, dass Lehrkräfte die Wahrscheinlichkeit des Erreichens eines Abiturs bei Schülerinnen und Schülern aus sozioökonomisch besser gestellten Elternhäusern bei gleichen Leistungen eher überschätzen (Becker 2013). Der Zuwanderungshintergrund wurde in der Untersuchung jedoch nicht berücksichtigt.

Aus dem deutschen Kindergartenkontext ist zudem bekannt, dass sich die Erwartungen von Erzieherinnen und Erziehern nicht systematisch zwischen Kindern mit einem türkischen Zuwanderungshintergrund und einheimischen Kindern unterscheiden (Kratzmann 2013). Dies verdeutlichen Analysen von Daten, die aus der Studie "Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vor- und Grundschulalter" (BiKS) stammen. Die Erwartungen wurden über die Abfrage von Fähigkeitseinschätzungen und Prognosen über den Einschulungszeitpunkt und den Bildungsabschluss erhoben. Ob die Erwartungen von Lehrkräften in einer ähnlichen Weise unverzerrt sind, bleibt fraglich.

Eine experimentelle Studie aus dem deutschen Kontext belegt, dass Lehrkräfte inhaltlich identische Essays signifikant schlechter bewerten, wenn die Verfasserin oder der Verfasser vermeintlich türkischstämmig ist (Sprietsma 2013). Die in Folge der Bewertung des Essays abgegebenen Bildungsempfehlungen fallen ebenfalls deutlich zu Gunsten der Kinder mit deutschen Vornamen aus. Dabei nimmt das Ausmaß an wahrgenommenen Unterschieden zu, wenn die Lehrkräfte über mehr unterrichtliche Erfahrung im Umgang mit Kindern aus Zuwandererfamilien verfügen.

Weitere Anhaltspunkte liefern aktuelle Arbeiten zur allgemeinen Qualität von Lehrerurteilen. Aus den Befunden kann ein grundlegender Eindruck der Akkuratheit von Lehrerurteilen im deutschen Kontext gewonnen werden, schließlich umfasst die diagnostische Kompetenz die Fähigkeit, „Schülerleistungen und -merkmale [...] korrekt einzuschätzen“ (Lorenz 2011: S. 11). Der entscheidende Unterschied zum Konzept der Lehrererwartungen liegt demnach darin, dass sich die diagnostische Kompetenz auf die Einschätzungen gegenwärtiger Merkmalsausprägungen bezieht, während Lehrererwartungen die Einschätzungen dieser Ausprägungen in der Zukunft meinen.⁷ Sollten Lehrkräfte nicht in der Lage sein, die gegenwärtigen Leistungen eines Kindes adäquat einzuschätzen, so kann davon

7 Übergangsempfehlungen stellen eine Ausnahme dar, da sie eine prognostische Form der diagnostischen Kompetenz widerspiegeln.

ausgegangen werden, dass auch Einschätzungen, die sich auf die Zukunft beziehen, ungenau ausfallen (vgl. Schrader und Helmke 2001: S. 47 f.).⁸

Mehrere Arbeiten, die ebenfalls mit den BiKS-Daten arbeiten, kommen zu dem Schluss, dass Lehrkräfte die Leistungen und Interessen in den Fachbereichen Deutsch und Mathematik durchaus zutreffend einschätzen (Karing 2009, Karing et al. 2011a; Lorenz 2011). Mögliche bestehende Verzerrungen in Hinblick auf askriptive Schülermerkmale wurden ebenfalls analysiert: Grundschul Kinder, deren Eltern einen hohen sozioökonomischen Status genießen, scheinen in Bezug auf die Fähigkeitsbereiche Arithmetik, Wortschatz, Textverstehen und Rechtschreibung öfter überschätzt und seltener unterschätzt zu werden als Kinder aus sozial vergleichsweise benachteiligten Familien (Lorenz 2011: S. 224 f.). Für den frühen Sekundarschulbereich (fünfte Klassenstufe) lassen sich diese Zusammenhänge nicht bestätigen (Karing et al. 2011b). Ein Einfluss der ethnischen Herkunft der Kinder auf die Genauigkeit von Lehrerurteilen scheint weder für den Grundschul noch für den Sekundarschulbereich belegbar zu sein.

Mit den BiKS-Daten konnte kürzlich außerdem gezeigt werden, dass Lehrerurteile umso höher ausfallen, je ähnlicher sich die Persönlichkeitseigenschaften von Lehrkräften und Schülerinnen und Schülern sind (Rausch et al. 2015). Welche Rolle die ethnische Herkunft von Lehrkräften und Kindern im Zusammenhang mit Lehrerurteilen spielt, wurde allerdings nicht untersucht. Trotzdem legen die Befunde nahe, dass das in der internationalen Literatur berichtete Phänomen eines positiven Einflusses der Übereinstimmung der ethnischen Herkunft von Lehrkraft und Kindern auf die Leistungserwartungen auch im deutschen Kontext auftreten könnte.

Zwischenfazit

Lehrererwartungen sind systematisch verzerrt, wenn sie unabhängig von den tatsächlichen Fähigkeiten und Potentialen der Kinder zwischen bestimmten Schülergruppen variieren. Für selbsterfüllende Prophezeiungen im Kontext ethnischer Bildungsungleichheiten sind in erster Linie systematische Über- und Unterschätzungen von Schülerinnen und Schülern mit einer bestimmten ethnischen Herkunft von Bedeutung.

8 Diese konzeptionelle Überschneidung bezieht sich auf die Vergleichskomponente diagnostischer Kompetenz. Die Aspekte der Niveau- und der Differenzierungskomponente ließen sich sicherlich auch auf Lehrererwartungen übertragen. Allerdings interessieren im Rahmen der Untersuchung von Erwartungseffekten in erster Linie Einschätzungen der relativen Leistungsposition einzelner Schülerinnen und Schüler.

Der internationale Forschungsstand zu dieser Frage erweist sich als gemischt. Zum Teil finden sich Belege dafür, dass die ethnische Herkunft von Schülerinnen und Schülern einen systematischen Einfluss auf die Leistungserwartungen von Lehrkräften hat, in einigen Fällen aber auch nicht (für eine Zusammenschau siehe Jussim und Harber 2005; Jussim et al. 2009b; Schofield 2006). Gleichzeitig ist eine Bewertung dieser Befunde schwierig, da eine hohe Anzahl von Studien auf Messungen von Lehrererwartungen zurückgreift, denen bereits intensive Schüler-Lehrer-Interaktionen vorausgingen.

Auf die Forschung aus dem deutschen Kontext trifft diese Schlussfolgerung ebenfalls zu. Allerdings gründet sie hier auf einer weit weniger abgesicherten empirischen Befundlage. So konnte bislang nur eine einzige Feldstudie belegen, dass lehrerseitige Einschätzungen der Begabung von Schülerinnen und Schülern negativ verzerrt sind, wenn sie sich auf Kinder mit einem Zuwanderungshintergrund beziehen. Auf der Grundlage eines Experiments liegen außerdem Hinweise dafür vor, dass türkischstämmige Kinder unterschätzt werden. Welche schulischen Leistungen Lehrkräfte an deutschen Schulen von einheimischen Kindern und von Kindern mit einem Zuwanderungshintergrund erwarten, ist nicht geklärt. Zu beantworten bleibt außerdem die Frage, ob Verzerrungen der Lehrererwartungen zwischen unterschiedlichen Herkunftsgruppen variieren.

Die internationale Forschung macht außerdem deutlich, dass das Vorliegen einer Übereinstimmung der ethnischen Herkunft von Lehrkraft und Kind positive Auswirkungen auf die Höhe von Leistungserwartungen zu haben scheint. Sollte dieser Zusammenhang auch auf den deutschen Schulalltag zutreffen, so könnten sich daraus negative Verzerrungen der Erwartungen an Kinder aus Zuwandererfamilien ergeben, da ein Großteil der Lehrerschaft an deutschen Schulen selbst keinen Zuwanderungshintergrund aufweist.

Im Gegensatz zu dieser widersprüchlichen Befundlage zeichnet der Überblick über Studien zu Verzerrungen der Lehrererwartungen nach der sozialen Herkunft der Heranwachsenden ein weit klareres Bild. Verschiedene Untersuchungen, die sowohl aus dem internationalen als auch aus dem deutschen Kontext stammen, legen nahe, dass Lehrererwartungen systematisch höher ausfallen, je höher der sozioökonomische Status der Familien ist, aus denen die Kinder kommen. Im deutschen Kontext könnte dies vor allem für die Nachkommen türkischstämmiger Arbeitsmigranten weitreichende Folgen haben. Denn selbst wenn Verzerrungen, die direkt mit der ethnischen Herkunft verbunden sind, relativ klein ausfallen, so wären türkischstämmige Kinder aufgrund der durchschnittlich sozioökonomisch schwachen Stellung ihrer Eltern im Mittel wahrscheinlich häufiger von negativ verzerrten Lehrererwartungen betroffen als Kinder aus der Mehrheitsbevölkerung.

Ausblick

Die unterschiedlichen Befunde zu Verzerrungen in Lehrererwartungen drängen die Frage auf, wie diese inkonsistenten Befunde trotz der hohen Anzahl an Studien zustande kommen könnten. Hierfür kommen verschiedene mögliche Erklärungen in Betracht. Nur die wenigsten nicht-experimentellen Studien können auf Indikatoren zurückgreifen, die auf einer tatsächlichen Messung von Erwartungen im Sinne lehrerseitiger Prognosen basieren. Stattdessen werden meistens Beurteilungen einer oder mehrerer Kompetenzen oder des Talents – bezogen auf den Ist-Zustand – analysiert. Ob sich systematische Verzerrungen auch dann finden lassen, wenn sich die gemessenen Erwartungen explizit auf zukünftige Schülerleistungen beziehen, ist nicht bekannt.

Zukünftige Untersuchungen sollten daher explizit die Erwartungen der Lehrenden an individuelle Kinder erheben. Die entsprechenden Messungen könnten Items einschließen, die sich auf mehr oder weniger weit in der Zukunft liegende Fähigkeiten, Schulleistungen und Bildungserfolge (in Form von Bildungsübergängen und Bildungsabschlüssen) beziehen. Gleichzeitig müssten Messungen des Potentials der Kinder vorliegen, um bestimmen zu können, inwieweit Leistungserwartungen begründet sind. An dieser Stelle bieten sich Messungen der Leistungen an, die zum gleichen Zeitpunkt stattfinden wie die Messungen der Lehrererwartungen. Anhand eines solchen Vorgehens könnte abschließend geklärt werden, ob die Lehrererwartungen an verschiedene (ethnische) Schülergruppen bei vergleichbaren Ausgangslagen systematisch voneinander abweichen oder nicht.

Andererseits unterscheiden sich die in den verschiedenen Studien analysierten Daten in Hinblick auf den Zeitpunkt der Messung von Lehrererwartungen. Wie bereits erläutert könnte damit das Problem verbunden sein, dass bis zum Zeitpunkt der Erhebungen eine gegenseitige Anpassung von Lehrererwartungen und Schülerleistungen in einem schwächeren (frühe Messung im Verlauf der Interaktion zwischen Lehrkraft und Kind) oder stärkeren Ausmaß (spätere Messung im Verlauf der Interaktion zwischen Lehrkraft und Kind) stattgefunden haben könnte, wobei unklar bleibt, in welche Richtung eine mögliche Anpassung stattgefunden haben könnte. Für einen konsistenten Nachweis von (ethnischen) Verzerrungen in den Lehrererwartungen bietet es sich daher an, frühe Zeitpunkte der Bildungslaufbahnen von Kindern in den Blick zu nehmen. Erhebungen von Lehrererwartungen sollten aus dem gleichen Grund im Idealfall zu Beginn eines Schuljahres stattfinden.

Um statistisch abgesicherte Aussagen über die Gestalt von Lehrererwartungen treffen zu können, müssten schließlich genügend Schulkinder und Lehrkräfte in einer Stichprobe enthalten sein. Eine Schülerstichprobe müsste hinreichend groß sein, um zu klären, ob Heranwachsenden mit Zuwanderungshintergrund

niedrigere Erwartungen entgegengebracht werden als einheimischen Kindern und ob Unterschiede zwischen verschiedenen Herkunftsgruppen bestehen. Darüber hinaus müsste die Zahl der Lehrkräfte hinreichend groß sein, um klären zu können, ob das Ausmaß an systematischen Verzerrungen in den Erwartungen zwischen verschiedenen Lehrkräften variiert und welche individuellen und kontextuellen Bedingungen dafür verantwortlich sind.

2.2.2 Befunde zur Vermittlung von Lehrererwartungen über das Lehrerverhalten

Zu den Wegen, über die Lehrkräfte ihre Erwartungen an Schülerinnen und Schüler herantragen, existieren einige eigenständige Forschungsarbeiten. Für eine Veranschaulichung der Vermittlung von Lehrererwartungen eignet sich das „10-Arrow Model“ (Harris und Rosenthal 1985). Hierbei handelt es sich um eine etwas kleinteiligere Version des in Abschnitt 2.2 vorgestellten konzeptionellen Modells der Entstehung von selbsterfüllenden Prophezeiungen in der Schule. Die Autoren unterscheiden Mediatorvariablen, welche die Verhaltensweisen von Lehrkräften betreffen, über die Erwartungen kommuniziert werden, von Moderatorvariablen, welche das Ausmaß von Erwartungseffekten beeinflussen (a. a. O.: S. 364).

Innerhalb des Modells stehen die Moderatorvariablen zu Anfang des Entstehungsprozesses von selbsterfüllenden Prophezeiungen (Schritt 1). Hierzu gehören Merkmale der Lehrkräfte ebenso wie Eigenschaften der Kinder. Moderatorvariablen haben sowohl einen Einfluss auf das Ausmaß an Verzerrungen in den Lehrererwartungen als auch auf die Wirkung von Mediatorvariablen und das Ausmaß an Veränderungen in Bezug auf das Outcome bei den Kindern. Nachdem Lehrererwartungen entstanden sind (Schritt 2), werden sie über das Verhalten der Lehrkräfte vermittelt (Schritt 3). Die Folge sind Veränderungen auf Seiten der Kinder, welche die Motivation, das Selbstkonzept und die Interaktion mit der Lehrkraft betreffen (Schritt 4). Schließlich sollten die genannten Prozesse eine Veränderung von Lernprozessen herbeiführen (Schritt 5). Die eigentliche Vermittlung von Erwartungseffekten sehen Harris und Rosenthal in Schritt 3 des Modells. Dabei geht es um die Frage, welche Verhaltensweisen von Lehrkräften durch ihre zuvor gebildeten Erwartungen strukturiert werden und dadurch eine Kommunikation der Erwartungen herbeiführen.

Mit dem „10-Arrow Model“ kann nicht erklärt werden wie die eigentliche Vermittlung von Erwartungseffekten funktioniert. Allerdings haben Harris und Rosenthal (1985) auf der konzeptionellen Grundlage dieses Modells versucht zu explizieren, welche Verhaltensweisen der Lehrkräfte bei der Vermittlung eine zentrale Rolle spielen. Dafür werteten sie die Ergebnisse einer Meta-Analyse aus,

in die insgesamt 135 Studien zu Erwartungseffekten und den Folgen für Interaktionsprozesse eingingen. Im Rahmen der Auswertungen konnten 31 Verhaltensweisen von Lehrkräften als bedeutsame Mediatoren identifiziert werden. Anhand von Faktoren-Analysen konnten schließlich vier Verhaltenskategorien gebildet werden, die zu Erwartungseffekten führen: (1) das sozio-emotionale Klima zwischen Lehrenden und Kindern (z. B. unterstützende Reaktionen, Lächeln, Aufmunterung, physische Nähe, positive oder negative Haltung gegenüber den Schülerinnen und Schülern), (2) das Ausmaß an negativem oder positivem Feedback (z. B. Lob und Kritik, positive Reaktionen auf Beiträge der Kinder beziehungsweise Ignorieren ihrer Beiträge), (3) der Input (z. B. die Präsentation von mehr oder weniger anspruchsvollen und/oder zusätzlichen Aufgaben, Aufmerksamkeit, Zeitaufwendung) und (4) der geforderte Output (Möglichkeiten zur Beteiligung, z. B. Aufrufen, eingeräumte Zeit für die Beantwortung).

Weiterführende empirische Untersuchungen deuten darauf hin, dass für die Vermittlung von Lehrererwartungen vor allem dem sozio-emotionalen Klima und dem Input eine wichtige Bedeutung zukommen (Babad 1993; Rosenthal 1989). Außerdem liegen Studien vor, in denen nonverbale Verhaltensweisen von Lehrkräften im Fokus stehen (Babad 1990; Babad et al. 1991; Tal und Babad 1989). Dabei handelt es sich um Arbeiten, in denen Videos des Unterrichtsgeschehens ausgewertet wurden. Bei der Auswahl der zu analysierenden Videos wurde versucht, kontextuelle Faktoren zu minimieren. Das heißt, dass sich die Auswertungen auf kleine Ausschnitte des Filmmaterials beziehen. Außerdem waren die Schülerinnen und Schüler in einem Teil der Videos nicht zu sehen. In anderen Filmausschnitten fehlte der Ton oder es war nur ein kleiner Gesichtsausschnitt der Lehrkraft sichtbar. Anhand dieses Vorgehens wurde deutlich, dass sich die Kommunikation von Erwartungen häufig auf subtilen Wegen vollzieht und dass dabei vor allem nonverbale Verhaltensweisen bedeutsam sind. Beispielsweise scheinen Lehrerinnen und Lehrer dieselbe Aufgabe in einem unterschiedlichen Tonfall zu stellen, je nachdem, welche Leistungen sie von einem Kind erwarten.

Interessanterweise wurden Unterschiede im Lehrerverhalten im Rahmen der genannten Untersuchungen nicht nur von den Forscherinnen und Forschern wahrgenommen. Vor allem die an den Studien teilnehmenden Schülerinnen und Schüler identifizierten anhand der Videos korrekt, wann eine Kommunikation gegenüber Kindern stattfand, von denen viel erwartet wird und wann sich die Lehrkraft an vermeintlich leistungsschwache Kinder wendete (Babad 2005). Die Bewertungen der Videos durch das Forscherteam und die Kinder und Jugendlichen waren außerdem mit den Bewertungen außenstehender Beobachter konsistent, die nicht der in der Klasse gesprochenen Sprache mächtig waren und die aus anderen kulturellen Kontexten stammten (vgl. auch Archer et al. 1993). Diese Ergebnisse unterstreichen, dass sich Lehrererwartungen nicht nur in relativ offensichtlichen,

sondern auch in sehr subtilen und sowohl für die Lehrkräfte als auch für die Kinder bewusst kaum wahrnehmbaren Interaktionsmustern zu manifestieren scheinen.

Schließlich liegen Arbeiten vor, die zeigen, dass verschiedene Lehrkräfte ihre Erwartungen in einem unterschiedlichen Ausmaß in Ungleichbehandlungen der Kinder übersetzen und es dadurch zu interindividuellen Unterschieden im Auftreten von Erwartungseffekten zwischen verschiedenen Lehrkräften kommt (Weinstein 2002: S. 51). Zu diesem Ergebnis kommen Auswertungen der schülerseitigen Wahrnehmung unterschiedlicher Lehrerverhaltensweisen im Unterricht (Brattesani et al. 1984; Kuklinski und Weinstein 2001; McKown und Weinstein 2008; Weinstein und Middlestadt 1979). Erwartungseffekte traten mit einer höheren Wahrscheinlichkeit in solchen Klassen auf, in denen unter der Schülerschaft mehr Ungleichbehandlungen berichtet wurden.⁹

Entsprechende interindividuelle Unterschiede zwischen den Lehrkräften in Bezug auf die Kommunikation verzerrter Erwartungen scheinen mit Persönlichkeitseigenschaften der Lehrkräfte zusammenzuhängen. In den betreffenden Studien zeigt sich empirisch, dass zum einen ein autoritärer kognitiver Stil bei der Lehrkraft Ungleichbehandlungen wahrscheinlicher zu machen scheint. Zum anderen zeigen manche Lehrpersonen einen starken und ausdrucksvollen Stil der Kommunikation und sind deshalb anfälliger für die Kommunikation ihrer Erwartungen (Weinstein 2002: S. 55).

Zwischenfazit

Im schulischen Alltag scheinen Lehrkräfte ihre Erwartungen über unterschiedliche Kanäle an die Schülerinnen und Schüler zu kommunizieren. Dabei lassen sich verbale und nonverbale Wege unterscheiden. Über diese beiden Wege scheinen Lehrkräfte in einem unterschiedlichen Maße Rückmeldungen, Input und Output zu vergeben und ein unterschiedliches sozio-emotionales Klima herzustellen. Gleichzeitig scheinen Unterschiede in Bezug auf das Ausmaß an Ungleichbehandlungen zwischen verschiedenen Lehrkräften zu bestehen.

9 Die Messung des Lehrerverhaltens erfolgte in den entsprechenden Studien anhand der Befragung der Kinder zu ihrer Sicht auf das Unterrichtsgeschehen. Anhand des *Teacher Treatment Inventory* sollten sie einschätzen, ob eine fiktive Mitschülerin, die sehr gute Leistungen erzielt, von ihrer Klassenlehrkraft in Bezug auf bestimmte Verhaltensweisen anders behandelt werden würde als eine fiktive Schülerin mit eher schlechten Leistungen. Auf diese Weise konnten Schulklassen identifiziert werden, in denen die Kinder durchschnittlich ein hohes Ausmaß an Ungleichbehandlungen erlebten und Klassen, in denen die Mehrzahl der Kinder berichtete, dass die fiktiven Schülerinnen gleich behandelt werden.

Aus dem deutschen Kontext liegen zu den Vermittlungsprozessen keine Studien vor. Die an deutschen Schulen ablaufenden Prozesse der Schüler-Lehrer-Interaktion sollten sich aber nicht grundlegend von den in der internationalen Forschung berichteten Prozessen unterscheiden.

Sollten Lehrererwartungen systematisch negativ verzerrt sein, wenn sie an Kinder mit einem bestimmten Zuwanderungshintergrund gestellt werden, so müssten verbale und nonverbale Verhaltensweisen der Lehrkräfte, die zu den Dimensionen Input, Output, Feedback und Herstellung eines sozio-emotionalen Klimas zählen, systematisch unterschiedlich ausfallen, je nachdem, ob Interaktionen mit Kindern mit oder ohne Zuwanderungshintergrund betrachtet werden. Bislang liegt keine Forschung vor, die belegen kann, dass sich Lehrerinnen und Lehrer – beispielsweise bezogen auf den US-amerikanischen Kontext – afroamerikanischen Kindern gegenüber anders verhalten als gegenüber anglo-amerikanischen Kindern. Für den deutschen Kontext bleibt zu klären, ob unterrichtliche Interaktionsmuster mit dem Zuwanderungshintergrund der Kinder variieren.

In Bezug auf die in diesem Abschnitt behandelte Fragestellung der Vermittlung von Lehrererwartungen muss außerdem betont werden, dass sie in den meisten Fällen eigenständig, das heißt losgelöst von vor- und nachgelagerten Prozessen der Entstehung selbsterfüllender Prophezeiungen abgehandelt wird. Auch wenn bekannt ist, welche Verhaltensweisen von Lehrkräften im Unterricht durch die Höhe von Leistungserwartungen strukturiert sind, so kann kaum eine Studie Aussagen dazu treffen, in welchem Maße die erhobenen Erwartungen verzerrt waren. Gleichzeitig fehlt es den zitierten Studien an Möglichkeiten, Aussagen über die Folgen für die Kompetenzentwicklung der Heranwachsenden zu treffen. So wurde nirgends geprüft, ob die gemessenen lehrerseitigen Ungleichbehandlungen tatsächlich Differenzen in den schulischen Leistungen der Kinder herstellen (oder verstärken), die den Leistungserwartungen der Lehrkräfte entsprechen.

Ausblick

Neben dem Nachweis unzutreffender Lehrererwartungen ist die Messung des Unterrichtsverhaltens der Lehrkräfte von zentraler Bedeutung für den Nachweis selbsterfüllender Lehrererwartungen. Denn Studien, die statistisch zeigen, dass Lehrererwartungen über die Ausgangsleistungen der Schülerinnen und Schüler hinaus einen Einfluss auf die später erzielten Leistungen haben, lassen streng genommen die Frage danach offen, ob dieser Effekt tatsächlich durch die Kommunikation der unterschiedlich hohen und in unterschiedlichem Ausmaß verzerrten Lehrererwartungen zustande kommt.

Um diese Frage zu beantworten, müssten Datenerhebungen idealerweise nicht nur einen Eindruck über die Vorgänge während des Unterrichts und in Bezug

auf die Lehrer-Schüler-Interaktion ermöglichen. Hierfür würden sich beispielsweise Videoaufnahmen des Unterrichtsgeschehens eignen. Vielmehr müsste das so gemessene Verhalten einzelner Lehrkräfte gegenüber individuellen Schülerinnen oder Schülern mit den im Ausblick von Abschnitt 2.2.1 vorgeschlagenen Messungen der Lehrererwartungen in Verbindung gebracht werden. Über wiederholte Messungen der Leistungen und Fähigkeiten zu späteren Zeitpunkten könnten außerdem mögliche Auswirkungen differentieller Verhaltensmustern der Lehrkräfte auf die Kompetenzentwicklung unterschiedlicher Schüler(gruppen) ermittelt werden.

Können Vermittlungsprozesse nicht gemessen werden, so eröffnen sich für den statistischen Zusammenhang zwischen verzerrten Lehrererwartungen und Schülerleistungen Erklärungen, die von dem Szenario einer selbsterfüllenden Prophezeiung abweichen. So ergibt sich insbesondere der Einwand, dass ein derartiger Zusammenhang durch unbeobachtete leistungsrelevante Merkmale zustande kommt, die sowohl die Höhe der Leistungserwartungen als auch die tatsächliche Kompetenzentwicklung beeinflussen. Dieser Aspekt wird in Kapitel 5 im Rahmen der Diskussion der empirischen Analysen erneut aufgegriffen und ausführlich diskutiert (siehe Abschnitt 5.4).

Abschließend soll darauf hingewiesen werden, dass zwar weitestgehend geklärt ist, in welchen Verhaltensdimensionen sich Lehrererwartungen ausdrücken. Allerdings müssen lehrerseitige Ungleichbehandlungen verschiedener Schülerinnen und Schüler nicht zwangsläufig als Benachteiligungen beziehungsweise Bevorteilungen gedeutet werden, welche in einem differentiellen beziehungsweise ungerechten Kompetenzerwerb resultieren. Vielmehr könnte die ungleiche Behandlung einzelner Kinder einerseits auch als Ausdruck proaktiver Versuche interpretiert werden, sich im Unterricht den individuellen Bedürfnissen der Kinder anzupassen. Andererseits könnten Ungleichbehandlungen reaktive Versuche der Lehrkräfte widerspiegeln, auf die verschiedenen Schülerverhaltensweisen zu reagieren. Weiterführende Forschungsarbeiten müssten also auch die Frage beantworten, welche Formen der Ungleichbehandlungen überhaupt Benachteiligungen für individuelle Schülerinnen und Schüler in dem Sinne mit sich bringen, dass Lernergebnisse entstehen, die bei einem anderen Verhalten der Lehrkräfte anders ausgefallen wären.

2.2.3 Befunde zu den Auswirkungen von Erwartungen und Verhaltensweisen der Lehrkräfte auf die Wahrnehmung, Motivation und Kompetenzentwicklung von Schülerinnen und Schülern

Aus Sicht des in Abschnitt 2.2.2 erwähnten „10-Arrow Models“ der Entstehung selbsterfüllender Prophezeiungen haben unterschiedliche Verhaltensweisen der

Lehrkräfte im Rahmen des Unterrichtskontextes Effekte auf die Motivation, die Wahrnehmung und das Verhalten der Kinder (Harris und Rosenthal 1985). Die Autoren betrachten diese Effekte als separaten Teilschritt der Entstehung selbsterfüllender Prophezeiungen getrennt. Als Resultat ergeben sich erst in einem vierten Schritt differentielle Lernentwicklungen auf Seiten der Kinder.

Allerdings konnten die bislang vorgelegten Forschungsarbeiten nicht klären, ob Schülerinnen und Schüler unterschiedliche Lehrerverhaltensweisen tatsächlich zwingend wahrnehmen müssen, damit sich Effekte auf das Lernen ergeben. Gleiches trifft auf die kausale Bedeutung motivationaler Variablen bei den Kindern für die Entstehung von Erwartungseffekten zu. So lässt sich anhand des aktuellen Forschungsstandes nicht bestimmen, ob sich unterschiedliche Interaktionsmuster von Lehrkräften (infolge ihrer unterschiedlich hohen Erwartungen) zwingend vermittelt über die Motivation der Kinder auf das Lernen auswirken. Aufgrund dieser Unklarheiten werden die Wahrnehmung und die Motivation der Schülerinnen und Schüler im Rahmen dieser Arbeit, wie in Abbildung 1 (Abschnitt 2.2) dargestellt, gemeinsam mit den konkreten Auswirkungen des Lehrerverhaltens auf das Lernen der Kinder als dritter und letzter Schritt des Entstehungsprozesses selbsterfüllender Prophezeiungen verstanden.

Schülerwahrnehmung

Zunächst soll es um nicht abgeschlossene Diskussion darüber gehen, ob unterschiedliche Lehrerverhaltensweisen eher auf direkten, eher auf indirekten oder sowohl auf direkten als auch auf indirekten Wegen auf die Schulleistungen von Kindern wirken. Diese Diskussion wurde ursprünglich von Brophy und Good (1974) angestoßen, die zwei Arten der Vermittlung zwischen Lehrerverhalten und schülerseitigem Lernen unterscheiden.

Die *direkte* Vermittlung würde funktionieren, indem bestimmte Lehrerverhaltensweisen, wie das Ausmaß des zur Verfügung gestellten Outputs, in einem unterschiedlichen Ausmaß Lernopportunitäten bereitstellen. Beispielsweise könnte das häufige Aufrufen eines Schülers während des Unterrichts dazu führen, dass dieser lernt sich angemessen auszudrücken, indem er viele Möglichkeiten zum Antworten beziehungsweise Reden und Formulieren erhält.

Die *indirekte* Vermittlung würde funktionieren, indem die Kinder Ungleichbehandlungen von Seiten der Lehrerschaft wahrnehmen und dadurch in ihrer Motivation zu Lernen beeinflusst werden. Beispielsweise könnte die ungleiche Vergabe von Rückmeldungen der Lehrkräfte an verschiedene Schulkinder während des Unterrichts dazu führen, dass eine Schülerin sich selbst einer Gruppe von Kindern zuordnet, die wenig oder nur negative Rückmeldungen bekommt. In der Folge könnte die Schülerin ein niedriges Fähigkeitsselfkonzept entwickeln und

die Motivation am Lernen verlieren. Dieser Weg der Vermittlung würde voraussetzen, dass den Kindern Ungleichbehandlungen in einem gewissen Maße bewusst sind und dass dieses Bewusstsein die beobachteten Leistungsunterschiede hervorruft.

Zunächst wurde angenommen, dass direkten und indirekten Vermittlungswegen eine relativ gleichwertige Bedeutung bei der Entstehung von Erwartungseffekten zukommt (Brophy und Good 1974; Brattesani et al. 1984). Neuere Studien kommen allerdings zu dem Schluss, dass der indirekte Weg über die schülerseitige Wahrnehmung unterschiedlichen Lehrerverhaltens eine wichtige Rolle bei der Vermittlung von Lehrererwartungen spielt. Dort konnte gezeigt werden, dass sowohl verzerrte Lehrererwartungen als auch Effekte dieser Erwartungen auf den Kompetenzerwerb in solchen Schulklassen häufiger vorkommen, in denen die Schülerinnen und Schüler stärkere Ungleichbehandlungen zwischen unterschiedlichen Kindern berichten (Kuklinski und Weinstein 2001; McKown und Weinstein 2008). Diese Befunde deuten darauf hin, dass Schülerinnen und Schüler unterschiedlich hohe Erwartungen und Verhaltensweisen ihrer Lehrkräfte korrekt wahrnehmen. Ob diese Wahrnehmung tatsächlich eine Voraussetzung für die Entstehung von Erwartungseffekten ist oder ob sie lediglich mit den Ungleichbehandlungen der Lehrkräfte einhergeht, können die betreffenden Studien allerdings nicht beantworten (Jussim et al. 2009b).

Schülermotivation

Gleichfalls ungeklärt ist die Frage, ob motivationale Variablen kausal zwischen dem Lehrerverhalten und den Schülerkompetenzen stehen oder ob Veränderungen in diesen Variablen lediglich die Folge unterschiedlicher Lernentwicklungen sind, die aus den direkten Wegen der Vermittlung von Lehrererwartungen resultieren.

Empirisch belegt wurde, dass Erwartungseffekte in den unteren Klassenstufen weniger über die Motivation vermittelt sind als in den oberen Klassenstufen (Kuklinski und Weinstein 2001). Entsprechend wurde geschlussfolgert, dass direkte Wege der Vermittlung bei den jüngeren Kindern eine stärkere Rolle spielen, während mit zunehmendem Alter den indirekten Wegen eine größere Bedeutung zukommt. Allerdings ist diese Schlussfolgerung kritisch zu betrachten. Denn motivationale Merkmale lassen sich bei den jüngsten Schülerinnen und Schülern nur schwer messen. Dafür spricht schon allein die Tatsache, dass sich in der betreffenden Studie in der jüngsten betrachteten Altersgruppe für die motivationalen Variablen kaum Varianz finden ließ. Dass die Motivation der Kinder im Rahmen der Entstehung von selbsterfüllenden Prophezeiungen kausal bedeutsam ist, kann die genannte Studie nicht zeigen.

Ausführliche konzeptionelle Überlegungen zu dieser Thematik finden sich bei Jussim et al. (2009b: S. 367 ff.). Die Autoren verweisen auf das psychologische Framework TARGET (Blumenfeld 1992), welches beschreibt, welche fünf lehrerseitigen Verhaltensdimensionen das schulische Lernen in erster Linie beeinflussen. Entsprechend der pädagogisch-psychologischen Literatur wirken alle Verhaltensdimensionen direkt auf die Schülermotivation. Dazu gehören die Gewährleistung von Schülerautonomie, Rekognitionen (Ausmaß an Feedback), Gruppenarbeiten, individuelle Evaluationen und Zeit zum Lernen. Diese Dimensionen lassen sich durchaus mit den von Harris und Rosenthal benannten Verhaltensdimensionen in Einklang bringen, die sie für die Vermittlung von Lehrererwartungen verantwortlich machen (Jussim et al. 2009b: S. 372). Implizit kann daher darauf geschlossen werden, dass Auswirkungen von Lehrererwartungen auf Schülerkompetenzen stets über die Motivation der Kinder vermittelt werden. Empirische Belege für diese Annahme stehen allerdings noch aus.

Effekte auf die Kompetenzentwicklung von Schülerinnen und Schülern – Internationaler Forschungsstand

Im Folgenden werden die zentralen Befunde in Bezug auf das eigentliche Resultat von Erwartungseffekten, nämlich die differentielle Kompetenzentwicklung von Schülerinnen und Schülern, zusammengefasst. Da der Nachweis derartiger Effekte eine Herausforderung für die zugrundeliegenden Messmodelle darstellt, werden auch die in der Literatur gewählten Methoden Beachtung finden.

Einer der wichtigsten Datensätze, welcher die Möglichkeit bietet, Erwartungseffekten im schulischen Kontext mit quantitativen Methoden auf den Grund zu gehen, ging aus der *Michigan Study of Adolescent Life Transitions* (MSALT) hervor. Das aus den USA stammende Forschungsprojekt verfolgte unter der Federführung von Jacqueline Eccles das Ziel, die Bildungs- und Berufskarrieren von mehr als 2000 Personen nachzuzeichnen. Die Erhebungen begannen 1983, als die teilnehmenden Jugendlichen die fünfte Klassenstufe besuchten. Die letzte Welle stammt aus dem Jahr 2000. Inhaltlich stand die Abbildung der zeitlichen Entwicklung von Selbstwahrnehmung, Überzeugungen und Werten und deren Auswirkungen auf die bildungs- und berufsbezogene Lebensplanung sowie auf die Umsetzung dieser Pläne im Fokus. Von den Erwartungen und Verhaltensweisen der Lehrkräfte im schulischen Alltag wurde angenommen, dass sie einen bedeutenden Beitrag zur Entstehung differentieller Bildungskarrieren von Jungen und Mädchen beziehungsweise Männern und Frauen leisten. Aus diesem Grund wurde das Lehrpersonal in den ersten beiden Erhebungswellen unter anderem nach der Einschätzung des mathematischen Talents, des Fleißes und nach der für das folgende

Schuljahr erwarteten absoluten und relativen Leistungsentwicklung einzelner Schulkinder gefragt.

Auf der Datengrundlage von MSALT entstand eine Reihe von Arbeiten, die Erwartungseffekte unter nicht-experimentellen Bedingungen untersuchen. Der Nachweis eines Einflusses von Lehrererwartungen auf die Schülerleistungen stellte sich dabei als statistische und methodische Herausforderung dar. Als geeigneter Weg erwies sich ein Vorgehen, bei dem die zu Anfang des Schuljahres gemessenen Lehrererwartungen mit den schulischen Leistungen im folgenden Schuljahr in Beziehung gesetzt wurden. Um mögliche Drittvariablen auszuschließen, die sowohl auf die Erwartungen als auch auf die Leistungen wirken und zu einer Überschätzung des Ausmaßes an Erwartungseffekten führen können, wurden die zum Zeitpunkt der Lehrerbefragung gemessenen Ausgangsleistungen berücksichtigt, teilweise ergänzt durch psychologische und demographische Variablen bei den Schülerinnen und Schülern.

Dieses Vorgehen verfolgten als erste Jussim (1989) sowie Jussim und Eccles (1992). Sie weisen den Einfluss von Lehrererwartungen auf die schulische Kompetenzentwicklung mit einer Effektstärke von $d = .13$ aus. Der Wert ist deutlich geringer als die in den meta-analytischen Arbeiten angegebene durchschnittliche Effektgröße experimenteller Studien (siehe Abschnitt 2.2). Als zentrale Ursache hierfür stellt sich die Tatsache dar, dass Lehrererwartungen die tatsächlichen Fähigkeiten und Kompetenzen der Kinder unter natürlichen Bedingungen häufig zutreffend abbilden (siehe Abschnitt 2.2.1).

Ein weiterer Datensatz, der wie MSALT sowohl geeignete Indikatoren für Lehrererwartungen als auch für die Leistungsentwicklung von Kindern beinhaltet, ist die kanadische „Carnegie Human Resources Data Bank“ (Williams 1976). Lehrererwartungen wurden hier über die Frage gemessen, ob individuelle Schülerinnen und Schüler das letzte High-School-Jahr vor dem Übergang in die Universität erfolgreich absolvieren werden oder nicht (West und Anderson 1976). Ähnlich wie in den Analysen der MSALT-Daten verdeutlichen die empirischen Analysen, dass Lehrererwartungen in erster Linie deshalb einen statistischen Zusammenhang mit Schülerleistungen aufweisen, weil ein Großteil der geäußerten Einschätzungen gemessen an den tatsächlichen Leistungen der Kinder zutreffend ausfiel. Entsprechend ließen sich keine Hinweise für selbsterfüllende Prophezeiungen finden.

Nicht manipulierte Lehrererwartungen waren auch Gegenstand einer Studie von Brattesani, Weinstein und Marshall (1984), die zeigen konnten, dass hohe Erwartungen der Lehrkräfte in Bezug auf das Bewältigen von Schulaufgaben und in Bezug auf die später im Schuljahr erzielbaren Leseleistungen zu höheren Lernzuwächsen im Bereich Lesen innerhalb eines Schuljahres führen. Lehrererwartungen erklären in dieser Studie unter Kontrolle der Ausgangsleistungen zwischen zwei

und sieben Prozent der Varianz im Leistungszuwachs bei Kindern der vierten, fünften und sechsten Klassenstufe.

Auch Alvidrez und Weinstein (1999) nutzten Sekundärdaten, um die langfristigen Folgen selbsterfüllender Prophezeiungen zu untersuchen. Mit dem Datensatz „Jack and Jeanne Block's 20-year Longitudinal Investigation of Ego and Cognitive Development“ können sie belegen, dass die Einschätzungen von Vorschullehrkräften gegenüber Vierjährigen in Bezug auf deren Intelligenz, soziale und kognitive Persönlichkeitseigenschaften sowie die zukünftig erzielbaren Leistungen in der High School teilweise nach der sozialen Herkunft verzerrt waren. In einem moderaten Ausmaß scheinen sich diese Verzerrungen sieben Jahre später selbst zu erfüllen. Bei diesem Ergebnis kann jedoch nicht ohne Zweifel von Erwartungseffekten die Rede sein. Denn erstens wurden die Kinder wahrscheinlich nicht bis zur siebten Klasse von ihren Vorschullehrkräften unterrichtet, während über die Erwartungen der Lehrkräfte aus höheren Klassenstufen gleichzeitig keine Informationen vorliegen. Zweitens ließen sich signifikante Einflüsse der Lehrererwartungen lediglich auf die Noten, nicht jedoch auf die Ergebnisse unabhängiger Leistungstests feststellen. Ob die Lehrererwartungen tatsächlich kausal auf die Schülerleistungen gewirkt hatten oder ob die Notenvergabe durch die Lehrkräfte unabhängig von tatsächlich bestehenden Kompetenzen nach der sozialen Herkunft der Kinder verzerrt war, lässt sich nicht bestimmen.

Zwei aktuelle Studien haben mögliche Erwartungseffekte bei sehr jungen Schülerinnen und Schülern mit den Daten der „NICHD Study of Early Child Care“ analysiert und verglichen dabei die Kompetenzbereiche Mathematik und Lesen. In den Studien wurden verzerrte Lehrererwartungen aus der ersten Klassenstufe mit den Schulleistungen von Schülerinnen und Schülern im Alter von 15 Jahren in Beziehung gesetzt. Zur Bestimmung von Verzerrungen regressierten die Autoren die Lehrererwartungen in einem ersten Schritt auf leistungsrelevante Merkmale der Kinder. Die bei diesem Vorgehen entstehenden Residuen bilden denjenigen Teil der Lehrererwartungen ab, der sich nicht durch die in der ersten Klassenstufe gemessenen Kompetenzen erklären lässt. Die Residuen können somit als Verzerrungen von Lehrererwartungen verstanden werden. In einem zweiten Schritt untersuchten die Autoren, ob sich signifikante Effekte der Residuen auf spätere Lese- und Mathematikkompetenzen (Hinnant et al. 2009) und auf vokabuläres Wissen und verbales logisches Denken (Sorhagen 2013) nachweisen lassen. Die empirischen Analysen liefern Hinweise für Erwartungseffekte in Bezug auf vier genannte Kompetenzbereiche.

Kürzlich hat sich außerdem eine Studie der Frage angenommen, ob Erwartungseffekte zur Vergrößerung leistungsbezogener Ungleichheiten im Rahmen des Kindergartenbesuchs beitragen (Ready und Chu 2015). Die zugrundeliegen-

den Daten stammen aus der Kindergartenkohorte der Early Childhood Longitudinal Study (ECLS-K). Die Analysen von knapp 14.000 Kindern verdeutlichen, dass die Erwartungen von Erzieherinnen und Erziehern systematisch nach der sozialen Herkunft der Kinder verzerrt zu sein scheinen. So zeigt sich, dass die Leseleistungen von Kindern aus privilegierten Familien im Vergleich zu tatsächlichen Fähigkeiten höher eingeschätzt werden, während die Leistungen von Kindern aus sozial schwächeren Familien unter der Kontrolle leistungsrelevanter Merkmale geringer eingeschätzt werden. Gleichzeitig belegen die Autoren, dass Überschätzungen der Leseleistungen mit einem stärkeren Zuwachs von Lesekompetenzen im Zeitverlauf einhergehen. Dieser Effekt besteht auch unter Kontrolle der sozialen und der ethnischen Herkunft, des Geschlechts, des Alters und zahlreicher anderer Hintergrundvariablen. Die Studie stellt im Vergleich zu allen bisher zitierten Arbeiten insofern eine Ausnahme dar, als dass explizit nach der potentiellen Bedeutung von Erwartungseffekten für die Entstehung schulleistungsbezogener Disparitäten nach der sozialen Herkunft gefragt wird. Allerdings bleibt unklar, inwieweit sich die für den Kindergartenkontext belegten Zusammenhänge auf den schulischen Kontext übertragen lassen.

Studien aus Europa und Deutschland

Der Großteil an Forschung zu selbsterfüllenden Prophezeiungen durch Lehrererwartungen stammt aus den USA. Auch aus dem europäischen Ausland lassen sich vereinzelt Studien finden (Jungbluth 1993, 1999; Morency et al. 2003; Navas et al. 1993; Tal und Babad 1989; Trouilloud et al. 2002, 2006; van den Bergh, Linda et al. 2010). Im deutschen Kontext werden Erwartungseffekte ebenfalls angesprochen (z. B. Dumke 1977; Ulich 2001; Schofield 2006; Hirschauer und Kullmann 2010). Einschlägige empirische Arbeiten liegen allerdings nur sehr vereinzelt und erst seit Kurzem vor.

Zunächst wurde sich der Thematik in einem Sammelband angenommen (Ludwig und Ludwig 2007). Die Beiträge beschäftigen sich mit den im deutschen Bildungssystem bestehenden geschlechtsspezifischen Bildungsungleichheiten. Im Mittelpunkt der Erklärung von Disparitäten stehen intra- und interindividuelle geschlechtsstereotype Rollenerwartungen. Die Herausgeber stellen in einem einführenden Kapitel als konzeptionelle Rahmung der weiteren Beiträge ein Modell der Entstehung selbsterfüllender Prophezeiungen vor. Leider orientieren sich die darauffolgenden Beiträge nicht explizit an diesem Modell, sondern arbeiten zum Teil recht unsystematisch verschiedene Teilfragen ab. Trotzdem liefern zwei Beiträge interessante Erkenntnisse zu Lehrererwartungen und ihren möglichen Auswirkungen.

So präsentieren Rustemeyer und Fischer (2007) die Befunde einer Studie, in der Sechstklässlerinnen und Sechstklässler zu ihrer Wahrnehmung der Lehrkräfte befragt wurden. Die Autoren berühren die Thematik der Erwartungseffekte insofern, als dass sie der Frage nachgehen, ob das schülerseitig wahrgenommene Lehrkraftverhalten und die schülerseitig wahrgenommenen Einstellungen der Lehrkräfte mit Geschlechterunterschieden in Mathematik zusammenhängen. Es wird also nicht nur danach gefragt, ob selbsterfüllende Prophezeiungen im Feld auftreten, sondern darüber hinaus auch versucht zu klären, ob das Phänomen eine Rolle bei der Entstehung von Bildungsungleichheiten, hier nach Geschlecht, spielt. Allerdings können die Lehrererwartungen nur über die Schülerperzeption erfasst werden. Verhalten sich Lehrkräfte aus Sicht der Schülerinnen und Schüler unterschiedlich gegenüber Jungen und Mädchen, so sprechen die Autoren von „Erwartungseffekten“. Die Auswertung der Daten ergibt, dass solche Lehrkräfte, die Jungen – aus der Sicht der Jugendlichen – mehr Beachtung schenken und „deren Sympathie und Antisympathie gegenüber Schülerinnen und Schülern im Unterricht offenkundig wird“ (a. a. O.: S. 97), in ihrem Verhalten Erwartungseffekte erkennen lassen und durch ihr Verhalten die Selbsteinschätzung von Mädchen und deren Interesse an Mathematik negativ beeinflussen. Die Befunde werden weiterhin so gedeutet, dass Erwartungseffekte Mitverursacher von geschlechtsspezifischen Bildungsungleichheiten sind.

Mit dieser Schlussfolgerung setzen sich die Autoren jedoch einiger Kritik aus. Es wird zwar gezeigt, dass Schülerinnen und Schüler bestimmte Verhaltensmuster der Lehrenden so wahrnehmen, dass diese davon in den psychologischen Dimensionen Selbstkonzept, Interesse, implizite Persönlichkeitstheorie und Lernzielorientierung betroffen sind. Welche Erwartungen die Lehrkräfte tatsächlich an die Jungen und Mädchen stellten, geht aus der Studie aber nicht hervor. Offen bleibt genauso, ob Verhaltensmuster wie die Aufgabenorientierung der Lehrenden unabhängig von der Schülerwahrnehmung tatsächlich zwischen Jungen und Mädchen variierten und ob dieses Verhalten von den vorausgegangenen Erwartungen beeinflusst war. Damit lassen die Daten einige alternative Erklärungen für die unterschiedliche Selbstwahrnehmung der Jugendlichen zu. Dazu zählt unter anderem, dass das differentielle Lehrerverhalten nur in der Wahrnehmung der Schülerinnen und Schüler existierte und sich intrapersonell selbst bestätigte. Dadurch bleibt letztendlich auch offen, welche Bedeutung selbsterfüllenden Prophezeiungen bei der Entstehung geschlechtsspezifischer Bildungsungleichheiten zukommt.

Der Beitrag von Winheller (2007) aus demselben Sammelband knüpft direkt an den soeben erläuterten Beitrag an. Wieder geht es um Geschlechterunterschiede in Bezug auf das Interesse und das Selbstkonzept im Fach Chemie. Den Ergebnissen einer Interventionsstudie in der siebten Klassenstufe entnimmt der Autor, dass autonomie- und kompetenzunterstützendes Lehrerverhalten im Unterricht zu einer

Angleichung von Interesse und Selbstkonzept zwischen Jungen und Mädchen führt. Machen sich die Lehrkräfte von geschlechtsspezifischen Stereotypen und Erwartungen frei, so verringern sich die Unterschiede in der Selbstwahrnehmung zwischen Mädchen und Jungen. Die Erwartungen des Lehrpersonals sowie das Verhalten der Lehrkräfte sind jedoch auch hier nicht Untersuchungsgegenstand der Studie. Der Beitrag kann damit nur sehr implizite Hinweise dahingehend liefern, dass stereotype Rollenerwartungen von Lehrkräften einen Beitrag im Rahmen der Entstehung geschlechtsspezifischer Bildungsungleichheiten leisten.

Eine andere Studie wird zwar nicht unter dem Label „Lehrererwartungen“ oder „selbsterfüllende Prophezeiung“ geführt, bietet aber aufschlussreiche Erkenntnisse in Bezug auf Erwartungseffekte im deutschen Schulsystem (Stahl 2007). Untersucht werden hier mögliche Zusammenhänge zwischen Lehrereinschätzungen der Lesefähigkeiten und tatsächlichen Leseleistungen und dem Leseinteresse. Es zeigt sich, dass die Leseleistungen am Ende der vierten Klasse nicht nur signifikant mit den bereits bestehenden Vorkenntnissen (Leseleistungen gemessen am Ende der dritten Klassen) zusammenhängen, sondern dass die Lehrereinschätzungen der Leseleistungen (gemessen am Ende der dritten Klasse) ebenfalls in signifikanter Beziehung zu den später erbrachten Leseleistungen stehen (a. a. O.: S. 181 f.). Die Arbeit liefert Hinweise dafür, dass Erwartungseffekte im deutschen Schulsystem wahrscheinlich sind. Welche Bedeutung dem Phänomen bei der Entstehung von Leistungsdisparitäten relativ zu anderen Einflüssen wie zum Beispiel der sozialen Herkunft zukommt, bleibt allerdings offen. Entsprechende Hintergrundvariablen gingen nicht in die Strukturgleichungsmodelle ein. Da aber in derselben Studie gezeigt werden konnte, dass die gemessenen Lehrurteile zum Teil systematisch nach der sozialen und ethnischen Herkunft der Schülerinnen und Schüler verzerrt waren (siehe Abschnitt 2.2.1), wäre es durchaus denkbar, dass selbsterfüllende Prophezeiungen im deutschen Kontext bestehende Nachteile von Kindern aus niedrigeren sozialen Schichten und mit Zuwanderungshintergrund verstärken könnten. Allerdings bedarf diese implizit getroffene Vermutung einer empirischen Überprüfung mit geeigneten Methoden. Insbesondere muss gezeigt werden, dass Verzerrungen von Lehrererwartungen nach der sozialen Herkunft unabhängig von der Wirkung der sozialen Herkunft auf den Schulerfolg selbst Effekte auf das Lernen haben.

Eine weitere Studie aus Deutschland kann selbsterfüllende Prophezeiungen für den deutschen Kontext mit Hilfe des oben erläuterten Vorgehens nachweisen, bei dem die Residuen aus Regressionen der Lehrererwartungen analysiert werden (Becker 2013). Im Fokus stehen hier lehrerseitige Einschätzungen der Wahrscheinlichkeit, dass Schülerinnen und Schüler der zehnten Klassenstufe später das Abitur erreichen werden. Dieses Maß für Lehrererwartungen hat den Ergebnissen zufolge einen signifikanten Einfluss auf die tatsächliche Wahrscheinlichkeit des

Erreichens eines Abiturs. Ein wichtiger Vorteil dieser Untersuchung liegt darin, dass in den statistischen Modellen solche Faktoren Berücksichtigung fanden, die entsprechend bewährter theoretischer Modelle zur Erklärung von Bildungsübergängen bedeutsam für den Übergang in ein Gymnasium sind. Als Nachteil ist allerdings zu betrachten, dass die der Studie zugrunde liegenden Daten des Kölner Gymnasiasten-Panel aus dem Jahr 1969 (Zeitpunkt der ersten Erhebungswelle, in der die untersuchten Schülerinnen und Schüler die zehnte Klassenstufe besuchten) stammen. Die Ergebnisse lassen sich deshalb nur bedingt auf aktuelle Kohorten von Schülerinnen und Schülern beziehungsweise Lehrkräften übertragen.

Im deutschen Kontext wurde kürzlich außerdem eine Untersuchung von Erwartungseffekten durchgeführt, die es sich zum Ziel gesetzt hat, Effekte auf die Leistungen und die Noten von individuellen Kindern und von Gruppen von Schülerinnen und Schülern in der fünften Klassenstufe zu untersuchen (Friedrich et al. 2015). Die Autoren diskutieren und analysieren in diesem Zusammenhang auch, ob selbsterfüllende Prophezeiungen vermittelt über das Selbstkonzept auf die Leistungsentwicklung wirken. Den Ergebnissen zufolge scheinen Lehrererwartungen mit der individuellen Leistungsentwicklung einzelner Kinder in Zusammenhang zu stehen. Dieser Zusammenhang scheint zum Teil über das Selbstkonzept vermittelt zu sein. Diese Ergebnisse müssen allerdings kritisch betrachtet werden. Denn im Rahmen von statistischen Modellen zur Schätzung der Leistungsentwicklung in Abhängig von Lehrererwartungen (die unabhängig davon in die Modelle eingingen, ob diese im Vergleich zu den tatsächlichen Leistungen der Kinder zutreffend waren oder nicht) versäumen es die Autoren, einige entscheidende Merkmale der Kinder zu berücksichtigen. Davon betroffen ist in erster Linie der sozio-ökonomische Hintergrund. Aber auch eine Unterscheidung von Kindern mit und ohne Zuwanderungshintergrund wurde nicht getroffen. Geht man davon aus, dass Leistungserwartungen von Lehrkräften (teilweise auch bei gleichen tatsächlichen Leistungen) mit der sozialen oder ethnischen Herkunft (oder beidem) variieren, so muss der gefundene Effekt von Lehrererwartungen auf die Leistungsentwicklung in Mathematik nicht zwangsläufig als selbsterfüllende Prophezeiung gedeutet werden. Genauso gut könnte er den Einfluss der sozialen beziehungsweise der ethnischen Herkunft (oder auch den Einfluss aller anderen nicht gemessenen leistungsrelevanten Merkmale, die mit den Lehrererwartungen zusammenhängen) auf die Mathematikleistungen abbilden. Da sich für Verzerrungen lehrerseitiger Einschätzungen nach der sozialen und nach der ethnischen Herkunft zahlreiche Belege finden (siehe Abschnitt 2.2.1) und weil der Einfluss der sozialen und ethnischen Herkunft auf mathematische Kompetenzen gleichzeitig hinlänglich bekannt ist, sollte man die Bedeutung dieses Kritikpunkts nicht unterschätzen. Weiterhin ist anzumerken, dass sich die aufgeworfenen Fragestellungen genau wie die ge-

troffenen Vorhersagen in Bezug auf die Richtung der untersuchten Effekten ausschließlich aus dem Forschungsstand und in keiner Weise aus theoretischen Überlegungen ableiten.

Zwischenfazit

Insgesamt liegen zahlreiche Arbeiten vor, die Erwartungseffekte im Feld untersuchen und die berücksichtigen, dass Lehrererwartungen nicht nur unterschiedlich hoch, sondern auch ein einem unterschiedlichem Ausmaß verzerrt beziehungsweise zutreffend sein können. Fasst man die Befunde zusammen, so sind dabei drei wesentliche Aspekte hervorzuheben. *Erstens* lassen sich unter natürlichen Bedingungen Einflüsse von Lehrererwartungen auf die Entwicklung schulischer Leistungen von Schülerinnen und Schülern belegen. *Zweitens* sind diese Effekte üblicherweise kleiner als in experimentellen Studien. *Drittens* liegen für den deutschen Schulkontext bislang nur vereinzelt Hinweise für die Existenz von selbsterfüllenden Prophezeiungen vor. Gleichzeitig gründet keine der betreffenden Untersuchungen auf belastbaren Datenquellen und Analysen.

Die zahlreichen Untersuchungen dieses Teilschritts der Entstehung selbsterfüllender Prophezeiungen konnten bislang allerdings nicht klären, welche Mechanismen hinter dem Zusammenhang zwischen lehrerseitigen Ungleichbehandlungen im Unterricht und einer differentiellen Leistungsentwicklung auf Seiten der Schülerinnen und Schülern stehen:

“[...] little or no research has clearly demonstrated how differential teacher treatment of high and low expectancy students actually affects those students in such a manner as to ultimately produce changes in achievement consistent with a self-fulfilling prophecy” (Jussim et al. 2009b: S. 376).

Diesbezüglich müssen zukünftige Forschungsarbeiten insbesondere zwei Fragen beantworten: Hat die schülerseitige Wahrnehmung von Unterschieden im Verhalten der Lehrkräfte eine kausale Bedeutung für die Entstehung von Erwartungseffekten? Wird der Zusammenhang zwischen dem Lehrerverhalten und der Leistungsentwicklung stets, nur zum Teil oder gar nicht über motivationale Merkmale der Kinder vermittelt?

Die Sichtung der Literatur offenbart außerdem, dass es durchaus Arbeiten gibt, die darauf schließen lassen, dass selbsterfüllende Prophezeiungen zur Entstehung oder Verstärkung ethnischer Ungleichheiten beitragen könnten. So zeigen Analysen der MSALT-Daten, dass Lehrkräfte im US-amerikanischen Raum zum Teil geringere Erwartungen an afro-amerikanische Kinder stellen, obwohl diese in objektiven Leistungsmessungen die gleichen Ergebnisse erzielen wie andere Kin-

der. Kombiniert man diese Befunde mit den Belegen für Effekte verzerrter Lehrerwartungen auf die Kompetenzentwicklung der Schülerinnen und Schüler, so lässt sich zumindest implizit auf eine Verstärkung ethnischer Disparitäten durch Erwartungseffekte schließen. Explizit, das heißt sowohl in Bezug auf die Fragestellung als auch auf die theoretische Argumentation und die empirischen Analysen, widmet sich der Frage nach der Bedeutung selbsterfüllender Prophezeiungen für die Entstehung von Bildungsungleichheiten nur eine einzige Studie (Ready und Chu 2015). Diese Arbeit lässt eine Verstärkung von Disparitäten, die mit der sozialen und mit der ethnischen Herkunft der Kinder verbunden sind, vermuten.

Häufig stehen in Forschungsarbeiten zur Thematik allerdings entweder der Zusammenhang zwischen Schülermerkmalen und Lehrerwartungen *oder* das Ausmaß an Erwartungseffekten im Fokus, selten aber beides. Auch Aussagen über die relative Bedeutung selbsterfüllender Prophezeiungen im Vergleich zu anderen Ursachen für die Entstehung von ethnischen Bildungsungleichheiten lassen sich auf der Grundlage des aktuellen Forschungsstandes nicht treffen.

Ob sich die Befunde aus der internationalen Literatur auf den deutschen Schulkontext übertragen lassen, muss noch geklärt werden. Hierzulande konnten Erwartungseffekte auf Bildungsentscheidungen sowie Hinweise auf einen Zusammenhang mit sozioökonomisch bedingten Disparitäten beim Übergang vom Primar- in das Sekundarschulsystem belegt werden (Becker 2013). Nur zwei Arbeiten beschäftigen sich mit Effekten auf die Kompetenzen von Schülerinnen und Schülern (Friedrich et al. 2015; Stahl 2007). Ein Zusammenhang mit der Frage nach der Entstehung von Bildungsungleichheiten wird hier allerdings nicht versucht herzustellen. Außerdem weisen die Studien messmethodische Mängel auf, welche die Befunde leicht angreifbar machen.

Ausblick

Alle bisherigen Überlegungen lassen darauf schließen, dass Erwartungseffekte empirisch nur im Längsschnitt untersucht werden können. Denn es kann angenommen werden, dass die mit der Zeit wiederholt stattfindenden Interaktionen zwischen Lehrkräften und Kindern und die damit verbundene wiederkehrende Kommunikation unterschiedlich hoher Erwartungen Auswirkungen auf das Lernen der Kinder zeitigen.

Eine Identifizierung der hinter der Entstehung von selbsterfüllenden Prophezeiungen liegenden Prozesse wäre darüber hinaus nur möglich, wenn die Verhaltensweisen der Lehrkräfte gegenüber individuellen Kindern (wie im Ausblick von Abschnitt 2.2.2 beschrieben) gemessen werden und wenn dieses Verhalten gegenüber einzelnen Kindern statistisch sowohl mit den vorausgegangenen Lehrerer-

wartungen also auch mit den Ausgangsleistungen und später erbrachten Leistungen der Kinder sowie ihrer Motivation und Wahrnehmung des Unterrichts in Verbindung gebracht werden könnten. Neben der Erfassung der leistungsbezogenen Ausgangslagen der Schülerinnen und Schüler sowie ihrer Kompetenzentwicklung über die Zeit erscheint es deshalb auch erforderlich, ihre Wahrnehmungen in Bezug auf das Verhalten der Lehrkräfte sowie Veränderungen in motivationalen Variablen der Kinder zu erfassen.

Für die Messung der Schülerwahrnehmung in Bezug auf das unterrichtliche Verhalten ihrer Lehrkräfte würde sich allerdings nicht das in der Literatur häufig eingesetzte *Teacher Treatment Inventory* (Weinstein et al. 1987) eignen. Dabei sollen Kinder einschätzen, ob eine fiktive Schülerin, die sehr gute Leistungen erzielt, von ihrer Klassenlehrkraft in Bezug auf bestimmte Verhaltensweisen anders behandelt wird als eine fiktive Schülerin mit eher schlechten Leistungen.¹⁰ Vielmehr müssten Aussagen der Kinder über das wahrgenommene Lehrerverhalten (idealerweise in Bezug auf die Verhaltensdimensionen Input, Output, Feedback und sozioemotionales Klima) *ihnen selbst gegenüber* erfasst werden. Nur auf diesem Weg könnte über Befragungen der Kinder gemessen werden, ob Effekte von Lehrererwartungen nur bei denjenigen Kindern auftreten, die eine Andersbehandlung im Vergleich zu anderen Kindern berichten.

Über das hier und in den Ausblicken der vorausgegangenen Abschnitte beschriebene Vorgehen bei der Datenerhebung ließe sich schließlich feststellen, ob Lehrererwartungen systematisch nach bestimmten zugeschriebenen Merkmalen wie der ethnischen Herkunft der Kinder verzerrt sind. Gleichzeitig könnte überprüft werden, ob verschiedene Lehrkräfte ihre Erwartungen unterschiedlich stark in ihr Verhalten übersetzen, welche Verhaltensweisen von unterschiedlich hohen Erwartungen betroffen sind und ob die Kinder diese Unterschiede adäquat wahrnehmen. Schließlich könnte geklärt werden, ob differentielle Lernentwicklungen tatsächlich kausal mit unterschiedlich hohen und in einem unterschiedlichen Ausmaß verzerrten Lehrererwartungen zusammenhängen. Über die Messung der Schülermotivation könnte zudem getestet werden, ob sich Ungleichbehandlungen der Lehrkräfte eher auf direkten oder eher auf indirekten Wegen auf das Lernen der Kinder auswirken.

10 Über diese Schüleraussagen werden in den betreffenden Studien Schulklassen identifiziert, in denen die Kinder durchschnittlich ein hohes bzw. ein niedriges Ausmaß an Ungleichbehandlungen durch die Lehrkräfte berichten.

2.2.4 Befunde zur Moderation von Erwartungseffekten

Den Überblick über den Forschungsstand abschließend soll im Folgenden kurz zusammengefasst werden, unter welchen Bedingungen Erwartungseffekte in einem stärkeren oder schwächeren Ausmaß auftreten. Im Gegensatz zu den Mediationsprozessen lassen sich aus dem Wissen um die Moderation von Erwartungseffekten mögliche Folgen für den Schulerfolg von Kindern mit Zuwanderungshintergrund ableiten.

Positive und negative Erwartungen scheinen stärkere Effekte auf das Lernen von Kindern aus sozial benachteiligten Familien (Hinnant et al. 2009); Madon et al. 1997; Sorhagen 2013) und von Kindern, die ethnischen Minderheiten angehören, zu haben (Jussim et al. 1996, 2009b). Weiterhin sind die Effekte von negativ verzerrten Erwartungen in einigen Studien höher als die Effekte positiv verzerrter Erwartungen (Madon et al. 1997). Insgesamt lassen sich die stärksten Erwartungseffekte für Kinder aus sozioökonomisch benachteiligten Familien, für Kinder mit Zuwanderungshintergrund und für Kinder in Schulkontexten mit einer niedrigen Leistungszusammensetzung nachweisen (Jussim et al. 1996; Madon et al. 1997). Dabei kann das Vorliegen mehrerer dieser Merkmale auch zu doppelten Nachteilen führen, beispielsweise bei Kindern, die sowohl einen Zuwanderungshintergrund als auch eine niedrige soziale Herkunft vorweisen. Häufig wird daraus geschlossen, dass selbsterfüllende Prophezeiungen am häufigsten bei den am stärksten stigmatisierten Schülergruppen auftreten (Jussim et al. 2009b: S. 360). Eine theoretische Begründung für das Zustandekommen derartiger Zusammenhänge liegt bislang allerdings nicht vor.

Weiterhin wurde wiederholt belegt, dass Erwartungseffekte vor allem in den jüngsten Klassenstufen auftreten (Kuklinski und Weinstein 2001; Raudenbush 1984; Rosenthal und Jacobson 1968). Wie bereits an anderer Stelle dieser Arbeit diskutiert ist damit wahrscheinlich ein methodisches Problem verbunden: Je später Lehrererwartungen im Schuljahr gemessen werden, desto stärker könnten sie sich bereits an beobachtbare Schülerleistungen angepasst haben. Umgekehrt könnten Erwartungseffekte die Kompetenzen der Kinder zu späten Messzeitpunkten bereits in Richtung der ursprünglichen Lehrererwartungen verändert haben. Schließlich finden sich Belege dafür, dass Lehrererwartungen mit einer höheren Wahrscheinlichkeit in solchen Klassen einen Einfluss auf das Lernen haben, in denen unter der Schülerschaft ein höheres Maß an Ungleichbehandlungen berichtet wird (Brattesani et al. 1984; Kuklinski und Weinstein 2001; McKown und Weinstein 2008). Diese Befunde unterstützen die Vermutung, dass Lehrkräfte ihre Erwartungen unterschiedlich stark in ihr eigenes Verhalten übersetzen (siehe Abschnitt 2.2.2). Hinweise darauf, dass Kinder durchaus zutreffend einschätzen, in welchem

Maße ihre Lehrerinnen und Lehrer Mitschülerinnen und Mitschüler unterschiedlich behandeln, finden sich in einer Studie von Weinstein (2002: S. 114ff.). Dort wurden die quantitativ erhobenen Wahrnehmungen der Kinder mit qualitativen Lehrerinterviews und Beobachtungen des Unterrichtsgeschehens verglichen, wobei sich große Übereinstimmungen finden ließen.

Zwischenfazit

Bestimmte Charakteristika von Schülerinnen und Schülern scheinen verstärkend auf Erwartungseffekte wirken zu können. Als Moderatoren, die für die Fragestellung dieser Arbeit relevant sind, kommen in erster Linie die soziale und die ethnische Herkunft der Kinder in Betracht. Sollten diese Merkmale auch im deutschen Kontext einen Einfluss auf die Stärke von Erwartungseffekten haben, so dürften selbsterfüllende Prophezeiungen umso mehr zu einer Vergrößerung ethnischer Bildungsungleichheiten beitragen, wenn Kindern aus niedrigen sozialen Schichten und aus Zuwandererfamilien systematisch verzerrte Lehrererwartungen entgegengebracht werden. Dies sollte vor allem für die Nachfahren klassischer Arbeitsmigranten gelten, die häufig aus sozioökonomisch benachteiligten Familien stammen. Bevor einer solchen Vermutung empirisch auf den Grund gegangen wird, sollte allerdings geklärt werden, welche theoretischen Argumente für ihre Gültigkeit sprechen könnten.

Ausblick

Ausblickend soll dargelegt werden, dass sich die in der Literatur berichteten Moderatoren jeweils auf unterschiedliche Aspekte innerhalb des dreiteiligen Entstehungsprozesses von Erwartungseffekten beziehen. Beispielsweise wäre es möglich, dass variierende Effekte in Abhängigkeit der Leistungen der Kinder den dritten Teilschritt betreffen und damit zusammenhängen, dass Lernprozesse bei leistungsschwächeren Kinder aus bestimmten Gründen stärker durch das Handeln von Lehrkräften beeinflussbar sind als Lernprozesse bei Kindern mit höheren Ausgangskompetenzen. In Bezug auf das Ergebnis, dass verzerrte Erwartungen stärkere Effekte auf das Lernen von Kindern aus sozial benachteiligten Familien und von Kindern, die ethnischen Minderheiten angehören haben, ließen sich dagegen Unterschiede in Bezug auf das Ausmaß an Verzerrungen zwischen unterschiedlichen Schülergruppen als Ursache diskutieren. Gleichzeitig könnte diesbezüglich vermutet werden, dass diese Moderation ebenfalls auf differentielle Prozesse in Bezug auf den dritten Teilschritt zurückgeht. So wird in der Literatur beispiels-

weise diskutiert, dass stigmatisierte Kinder häufiger mit einem geringen Selbstkonzept ausgestattet und deswegen in ihrem Lernen anfälliger für eine Beeinflussung durch das Verhalten von Lehrkräften sind als Kinder aus einheimischen Familien (Jussim et al. 2009b: S. 355ff.). Zukünftige Forschungsarbeiten sollten derartige konzeptionelle Überlegungen unbedingt einbeziehen, um die Wirkung möglicher Moderatorvariablen einerseits theoretisch begründbar zu machen und andererseits präziser empirisch untersuchen zu können.

Dass Erwartungseffekte in den oberen Klassenstufen in einem geringeren Umfang empirisch nachweisbar sind als in den unteren Klassenstufen, lässt wiederum vermuten, dass sich die Erwartungen der Lehrkräfte und die Kompetenzen der Schülerinnen und Schüler im Verlauf der Schüler-Lehrer-Interaktion aneinander anpassen. Um dieser Vermutung auf den Grund gehen zu können, sollten zukünftige Forschungsdesigns die Lehrkräfte, ihre Erwartungen und Verhaltensweisen sowie die Kompetenzen der Kinder von Beginn des Kontaktes zwischen Lehrkräften und Kindern an und über einen längeren Zeitraum hinweg verfolgen. Das heißt, dass nicht nur die Leistungen der Kinder, sondern auch die Erwartungen der Lehrerschaft zu verschiedenen Zeitpunkt erhoben werden sollten. Nur so könnte geklärt werden, ob sich Lehrererwartungen und Schülerleistungen tatsächlich aneinander anpassen und in welche Richtung Anpassungen stattfinden. Vor diesem Hintergrund bietet es sich an, empirische Untersuchungen im Grundschulbereich durchzuführen und die Lehrkräfte und Kinder über mehrere Klassenstufen hinweg zu begleiten. Über dieses Vorgehen könnte schließlich auch überprüft werden, welche langfristigen Auswirkungen Erwartungseffekte möglicherweise haben.



<http://www.springer.com/978-3-658-19880-0>

Selbsterfüllende Prophezeiungen in der Schule
Leistungserwartungen von Lehrkräften und
Kompetenzen von Kindern mit
Zuwanderungshintergrund
Lorenz, G.
2018, XII, 252 S. 1 Abb., Softcover
ISBN: 978-3-658-19880-0