



TUM School of Education
Susanne Klatten-Stiftungslehrstuhl für Empirische Bildungsforschung

Leistungsstarke Jugendliche in Naturwissenschaften

Vertiefende Analysen zu PISA 2006 und PISA 2012

Stefanie Sieglinde Schmidtner

Vollständiger Abdruck der von der Fakultät *TUM School of Education* der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Philosophie (Dr. phil.)

genehmigten Dissertation.

Vorsitzende: Prof. Dr. Claudia Nerdel

Prüfer der Dissertation: 1. Prof. Dr. Manfred Prenzel
2. Prof. Dr. Christina Seidel
3. Prof. Dr. Johannes Bauer (schriftliche Prüfung)
Prof. Dr. Kristina Reiss (mündliche Prüfung)

Die Dissertation wurde am 30.05.2017 bei der Technischen Universität München eingereicht und durch die Fakultät TUM School of Education am 31.08.2017 angenommen.

Inhaltsverzeichnis

Danksagung	vi
Zusammenfassung	vii
1. Problemstellung und Ziel der Untersuchung	1
1.1. Einbettung der Arbeit und offene Fragen	3
1.2. Zielsetzung und Vorgehen	5
1.3. Gliederung der Arbeit	6
2. Leistungsstarke Schülerinnen und Schüler in Naturwissenschaften	7
2.1. Definition und Forschungsstand im Kontext internationaler Vergleichs- studien	7
2.1.1. Naturwissenschaftliche Kompetenz als zugrunde liegendes Leis- tungskonstrukt	8
2.1.2. Definition von Top und High Performern in Naturwissenschaften	11
2.2. Individuelle, familiäre und institutionelle Merkmale leistungsstarker Ju- gendlicher in Naturwissenschaft	15
2.2.1. Individuelle Merkmale	16
2.2.2. Familiäre Merkmale	21
2.2.3. Institutionelle Rahmenbedingungen	24
2.2.4. Zusammenfassung	24
3. Naturwissenschaftsbezogenes Interesse und Fähigkeitsselbstkonzept	27
3.1. Interesse an den Naturwissenschaften	29
3.1.1. Die Person-Gegenstand-Theorie als theoretischer Rahmen	29
3.1.2. Interessenentwicklung und Ausdifferenzierung im Jugendalter .	31
3.1.3. Interesse leistungsstarker Jugendlicher in Naturwissenschaften	33

3.2.	Fähigkeitsselbstkonzept in Naturwissenschaften	38
3.2.1.	Fähigkeitsselbstkonzept leistungsstarker Schülerinnen und Schüler in Naturwissenschaften	38
3.2.2.	Soziale und dimensionale Vergleiche bei der Genese des Selbstkonzepts	41
3.2.3.	Wechselseitige Beziehung mit Interesse	43
3.3.	Profile aus Leistung, Selbstkonzept und Interesse	48
3.4.	Motivationsunterstützende Lehr-Lernbedingungen	53
3.5.	Zusammenfassung	58
4.	Die Bedeutung des Unterrichts für die Kompetenz- und Interessenentwicklung	59
4.1.	Rahmenmodell zur Untersuchung des Naturwissenschaftsunterrichts .	59
4.1.1.	Forschungsparadigmen der Unterrichtsforschung	60
4.1.2.	Angebot-Nutzungsmodell des Naturwissenschaftsunterrichts . .	61
4.2.	Der Einfluss von Unterrichtsaktivitäten auf die naturwissenschaftliche Kompetenz und das Interesse leistungsstarker Jugendlicher in Naturwissenschaften	65
4.2.1.	Interaktives Lehren und Lernen	66
4.2.2.	Forschend-entdeckendes Lernen	73
4.2.3.	Anwendungsbezug im Naturwissenschaftsunterricht	81
4.2.4.	Zusammenfassung	92
4.3.	Einfluss individueller Schülervoraussetzungen auf die Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen	94
4.4.	Erfassung der Unterrichtswahrnehmung von Schülerinnen und Schülern	102
5.	Fragestellungen	107
5.1.	Charakterisierung von High Performern in Naturwissenschaft	108
5.1.1.	Individuelle Merkmale	108
5.1.2.	Familiäre und institutionelle Merkmale	111
5.2.	Unterrichtsaktivitäten im Zusammenhang mit naturwissenschaftlicher Kompetenz und Interesse von High Performern	114

5.3.	Selbstkonzept-Interesse-Typen unter den High Performern und ihre Unterrichts- wahrnehmung	118
5.3.1.	Selbstkonzept-Interesse-Typen unter den High Performern	118
5.3.2.	Wahrnehmung motivationsunterstützender Lehr-Lernbedingun- gen von High Performern unterschiedlicher Selbstkonzept-Inter- esse-Typen	120
6.	Methodisches Vorgehen	123
6.1.	Forschungskontext	123
6.2.	Stichprobe	126
6.3.	Instrumente	127
6.3.1.	Kognitive Merkmale	127
6.3.2.	Familiäre und institutionelle Merkmale	129
6.3.3.	Motivational-affektive Merkmale	130
6.3.4.	Motivationsunterstützende Lehr-Lernbedingungen	131
6.3.5.	Unterrichtaktivitäten	131
6.4.	Auswertungsverfahren	134
6.4.1.	Umgang mit Plausible Values	134
6.4.2.	Mehrebenenanalysen	135
6.4.3.	Latente Profilanalyse	142
7.	Ergebnisse	147
7.1.	Charakterisierung von High Performern in Naturwissenschaft	147
7.1.1.	Individuelle Merkmale	147
7.1.2.	Familiäre und institutionelle Merkmale	157
7.1.3.	Zusammenfassung	171
7.2.	Unterrichtsaktivitäten im Zusammenhang mit naturwissenschaftlicher Kompetenz und Interesse von High Performern	173
7.2.1.	Faktorenstruktur der Unterrichtsmerkmale auf Schüler- und Klas- senebene	174
7.2.2.	Messinvarianz der Skalen zwischen High und Nicht-High-Perfor- mern	182
7.2.3.	Unterrichtsaktivitäten als Prädiktoren naturwissenschaftlicher Kompetenz von High Performern	186

7.2.4.	Unterrichtsaktivitäten als Prädiktoren des Interesses von High Performern an Naturwissenschaft	201
7.2.5.	Zusammenfassung	207
7.3.	Selbstkonzept-Interesse-Typen unter den High Performern und ihre Unterrichtswahrnehmung	209
7.3.1.	Selbstkonzept-Interesse-Typen unter den High Performern . . .	211
7.3.2.	Individuelle, familiäre und institutionelle Merkmale der Selbstkonzept-Interesse-Typen	217
7.3.3.	Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen von High Performern unterschiedlicher Selbstkonzept-Interesse-Typen	225
7.3.4.	Zusammenfassung	243
8.	Diskussion	247
8.1.	Methodische Überlegungen	247
8.2.	Diskussion zentraler Befunde	251
8.2.1.	Individuelle, familiäre und institutionelle Merkmale leistungsstarker Jugendlicher in Naturwissenschaften	251
8.2.2.	Unterrichtsaktivitäten im Zusammenhang mit der naturwissenschaftlichen Kompetenz und dem Interesse von leistungsstarken Jugendlichen in Naturwissenschaft	260
8.2.3.	Selbstkonzept-Interesse-Typen und mit ihnen verbundene Unterschiede in der Unterrichtswahrnehmung	265
8.3.	Implikationen für die pädagogische und bildungspolitische Praxis	270
8.4.	Ausblick	276
	Literaturverzeichnis	283
	Abbildungsverzeichnis	335
	Tabellenverzeichnis	337

A. Anhang	343
A.1. Latent-manifeste Mehrebenenanalysen	343
A.1.1. Unterrichtsaktivitäten im Zusammenhang mit naturwissenschaftlicher Kompetenz	343
A.1.2. Unterrichtsaktivitäten im Zusammenhang mit Interesse an den Naturwissenschaften	346
A.2. Mehrebenenanalysen zum Zusammenhang von Selbstkonzept-Interesstypen mit Lehr-Lernbedingungen	349
A.3. Instrumente	352
A.3.1. Naturwissenschaftsbezogene motivational-affektive Merkmale .	352
A.3.2. Unterrichtsaktivitäten im Naturwissenschaftsunterricht	353
A.3.3. Motivationsunterstützende Lehr-Lernbedingungen im Naturwissenschaftsunterricht	355

Danksagung

Zum Gelingen dieser Dissertation haben zahlreiche Menschen beigetragen und ich möchte mich bei allen bedanken, die mich auf dem Weg begleitet und unterstützt haben. Mein besonderer Dank gilt meinem Erstbetreuer Prof. Dr. Manfred Prenzel für die äußerst kompetente Betreuung und die Möglichkeit, meinen Interessen in der Bildungsforschung nachzugehen. Mit beeindruckendem Wissen und differenzierter Denkweise hat er mir in meiner wissenschaftlichen und persönlichen Entwicklung ebenso geholfen wie durch seine eigene Begeisterung für die Forschungsthemen und dem Bestreben, immer einen Bezug zur pädagogischen Praxis herzustellen. Meiner Zweitbetreuerin, Frau Prof. Dr. Tina Seidel, möchte ich für die äußerst fachkundige und wertschätzende Begleitung bedanken. Von ihrer Fähigkeit, an jeder Punkt sofort mitdenken und gezielt die passenden Ratschläge geben zu können, habe ich sehr profitiert. Auch ihrem gesamten Lehrstuhl-Team möchte ich für die bereichernden Diskussionen danken. Darüber hinaus möchte ich Anja Schiepe-Tiska, meiner Mentorin, sehr herzlich danken. In zahlreichen Diskussionen hat sie mir wichtige inhaltliche und persönliche Impulse gegeben und mich darin unterstützt, meine Ideen zu strukturieren und präzisieren. Auch Inga Simm möchte ich für ihre fachliche und emotionale Unterstützung sowie ihren unglaublichen Humor sehr danken.

Mein Dank gilt auch dem Doktorandenprogramm REASON und dem Elitennetzwerk Bayern für das anregende Forschungsumfeld und die Unterstützung, die ich dort erfahren habe. Danken möchte ich auch Dr. Aikling Tan und dem NIE in Singapur, die mir in einem Forschungsaufenthalt interessante und lehrreiche Einblicke in den Naturwissenschaftsunterricht und die Bildungsforschung in Singapur ermöglichten. Darüber hinaus möchte ich dem gesamten PISA- und Lehrstuhl-Team für die schöne Zeit danken. Mit euch gemeinsam an Projekten zu arbeiten und dabei viele lustige Momente zu erleben, empfand ich als großes Privileg. Auch den Hilfskräften möchte ich für ihre Unterstützung danken. Ein großer Dank gilt meiner Familie für ihre Geduld, ihren Rückhalt und das Vertrauen in meine Fähigkeiten. Mein ganz besonderer Dank gilt Martin, der mich in allem, was mir wichtig ist, unglaublich unterstützt und bestärkt.

Zusammenfassung

Die Ergebnisse von PISA 2006 und 2012 zeigen, dass Deutschland einen geringeren Anteil leistungsstarker Jugendlicher in Naturwissenschaften erreicht als vergleichbare Länder und sich ein großer Teil dieser Jugendlicher nur gering für Naturwissenschaften interessiert. Ausgehend davon beschäftigt sich diese Arbeit mit den Fragen (a) welche individuellen, familiären und institutionellen Merkmale leistungsstarke Jugendliche in Deutschland auszeichnen und (b) welche Unterrichtsaktivitäten eine höhere naturwissenschaftliche Kompetenz dieser Jugendlichen vorhersagen. Darüber hinaus wurde untersucht, (c) welche Selbstkonzept-Interesse-Typen unter leistungsstarken Jugendlichen identifiziert werden können und welcher Zusammenhang zur Unterrichtswahrnehmung besteht. Zur Untersuchung dieser Fragen wurden Sekundäranalysen von Daten aus PISA 2006 und PISA 2012 durchgeführt. Wie der Vergleich von leistungsstarken und anderen Jugendlichen ergab, zeichnen sich leistungsstarke Jugendliche in Naturwissenschaften durch bestimmte individuelle, familiäre und institutionelle Merkmale aus, sind in vielen dieser Merkmalen selbst jedoch eine heterogene Gruppe. Mehrebenenanalysen zeigten weiterhin, dass das häufigere Vorkommen einzelner wahrgenommener Unterrichtsaktivitäten nicht vorhersagt, wie hoch die Kompetenz sowie das Interesse leistungsstarker Jugendlicher ist. Leistungsstarke Jugendliche profitierten möglicherweise aufgrund ihres höheren Vorwissens von verschiedenen Unterrichtsumgebungen. Mit einer latenten Profilanalyse konnten darüber hinaus unterschiedliche Selbstkonzept-Interesse-Typen unter den leistungsstarken Jugendlichen identifiziert werden. Unter den leistungsstarken Jungen befand sich ungefähr ein Zehntel, unter den leistungsstarken Mädchen ungefähr ein Fünftel an Jugendlichen, die im Vergleich zu ihrer Leistung ein sehr geringes Selbstkonzept und Interesse aufwiesen. Es bestand ein Zusammenhang zwischen den Selbstkonzept-Interesse-Typen und der Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen (Autonomieunterstützung, Kompetenzunterstützung und inhaltlichen Relevanz) im Naturwissenschaftsunterricht. Gleichzeitig erwiesen sich häufiger berichtete forschend-entdeckende Unterrichtselemente auf Ebene der Schulklassen als Prädiktor dieser motivationsunterstützenden Lehr-Lernbedingungen. Die vorliegende Arbeit zeigt damit wichtige Charakteristika leistungsstarker Jugendlicher auf. Sie identifiziert Problemfelder aber auch Möglichkeiten für die Identifikation und multidimensionale Förderung dieser Jugendlichen.

1. Problemstellung und Ziel der Untersuchung

„Wie Deutschland die besten Schüler vernachlässigt“ (Vitzthum, 04.12.13) titelte eine Tageszeitung Ende 2013 als Reaktion auf die Ergebnisse von PISA 2012. Gut ein Jahrzehnt nach PISA 2000 hat sich Deutschland vom „PISA-Schock“ erholt und gehört zur internationalen Leistungsspitze in Naturwissenschaften unter den OECD-Ländern (Schiepe-Tiska, Schöps, Rönnebeck, Köller & Prenzel, 2013). Tiefergehende Analysen zeigen jedoch, dass diese Verbesserung Deutschlands vor allem darauf zurückzuführen ist, dass Deutschland den Anteil leistungsschwacher Jugendlicher verringern konnte (Schiepe-Tiska et al., 2013). Der Anteil leistungsstarker Jugendlicher konnte dagegen nicht verbessert werden (Schiepe-Tiska et al., 2013). Im internationalen Vergleich weist Deutschland zwar nach wie vor einen signifikant höheren Anteil an leistungsstarken Schülerinnen und Schülern in Naturwissenschaften auf als die OECD-Länder im Mittel (Schiepe-Tiska et al., 2013). In Relation zu Ländern mit vergleichbarem oder leicht niedrigerem mittleren Kompetenzniveau besitzt Deutschland jedoch nur durchschnittliche Anteile und wird beispielsweise von Ländern wie Finnland, Australien und Neuseeland übertroffen (Schiepe-Tiska et al., 2013).

Angesichts der hohen wirtschaftlichen, gesellschaftlichen und individuellen Bedeutung eines hohen Anteils leistungsstarker Jugendlicher in Naturwissenschaften sind diese Befunde besorgniserregend. Der Gesamtbedarf an MINT-Nachwuchskräften der Wirtschaft kann unter Berücksichtigung der Expansionsdynamik und Veränderungen der Arbeitsmarktstruktur nach wie vor nicht gedeckt werden (IW, 2016; IAB, 2015). Vertreter der Wirtschaft fordern deshalb, leistungsstarke Schülerinnen und Schüler als Nachwuchskräfte in Naturwissenschaften besser zu fördern und betonen den resultierenden „Return on Investment“ (z. B. sdw, 2015b, 2015a).

Sowohl die Wahl eines naturwissenschaftlich-technischen Studiums oder einer

Berufsausbildung als auch das Engagement in lebenslangen Lernprozessen in Naturwissenschaften erfordert neben einer hohen Leistung auch eine positive subjektive Einschätzung der eigenen Fähigkeiten in Naturwissenschaften und die Bereitschaft, sich weiter mit naturwissenschaftlichen Themen auseinanderzusetzen (Prenzel & Schütte, 2007). Die Förderung leistungsstarker Schülerinnen und Schüler in Naturwissenschaften muss deshalb die persönliche Interessenentwicklung und -ausdifferenzierung von Schülerinnen und Schülern unterstützen und ihnen dabei helfen, eigene Fähigkeiten und Interessen erkennen und einschätzen zu können bzw. zumindest eine Interessiertheit im Sinne einer grundlegenden Aufgeschlossenheit gegenüber Naturwissenschaften zu entwickeln (vgl. Prenzel & Schütte, 2007; Schütte, Frenzel, Asseburg & Pekrun, 2007).

Auch hier zeigt PISA jedoch, dass ein beträchtlicher Teil der leistungsstarken Jugendlichen in Naturwissenschaften sich nur geringfügig für Naturwissenschaften interessiert (Prenzel & Schütte, 2007) und Deutschland den im internationalen Vergleich der OECD-Länder zweithöchsten Anteil an leistungsstarken Jugendlichen mit im Verhältnis zur Leistung unerwartet geringer zukunftsorientierter Motivation in Naturwissenschaften aufweist (OECD, 2009b). Deutschland verfügt hier über „kognitives Potential für die Naturwissenschaften ..., das motivational noch nicht richtig erschlossen ist“ (Prenzel & Schütte, 2007, S. 122).

Angesichts dieser Problematik rückt die Förderung leistungsstarker Jugendlicher auch in der Bildungspolitik zunehmend in den Vordergrund. Während seit längerem die Initiative zur Förderung leistungsschwacher Jugendlicher besteht (z. B. KMK, 2010), bekräftigt die Kultusministerkonferenz der Länder nun in einer *Förderstrategie für leistungsstarke Schülerinnen und Schüler* das Ziel, leistungsstarke Jugendliche gezielt zu identifizieren und sowohl in ihrer Leistung als auch in ihren motivational-affektiven Merkmalen wie Interesse zu fördern (KMK, 2015). Der ehemalige Präsident der Kultusministerkonferenz nennt in diesem Zusammenhang das Ziel eines Anteils von 25% auf den beiden obersten Kompetenzstufen in PISA (Vitzthum, 04.12.13). Dies wirft jedoch sowohl für Beteiligte der Bildungssteuerung als auch für Lehrkräfte und Nachwuchsgewinnung der Wirtschaft die Fragen auf, welche Merkmale Schülerinnen und Schüler mit hohem Potential in Naturwissenschaften aufweisen und welche Faktoren eine hohe naturwissenschaftliche Kompetenz sowie Interesse an Naturwissenschaften fördern.

1.1. Einbettung der Arbeit und offene Fragen

Forschungsansätze zu schulischen und außerschulischen Lernumwelten (Doll & Prenzel, 2004b) spannen einen Suchraum für besondere Merkmale auf, die leistungsstarke Jugendliche in Naturwissenschaften möglicherweise auszeichnen und das Erreichen hoher naturwissenschaftlicher Kompetenz unterstützen.

So könnte man annehmen, dass leistungsstarke Jugendliche in Naturwissenschaften möglicherweise besondere individuelle Lernvoraussetzungen wie eine hohe Intelligenz oder eine hohe mathematische Kompetenz aufweisen. Auch ihre familiären Lernumgebung zeichnet sich möglicherweise durch charakteristische Merkmale wie einem akademischen Elternhaus oder einem höheren sozioökonomischen Bildungshintergrund aus und begünstigt so die Chancen, hohe Leistungen in Naturwissenschaften zu erreichen. Schließlich stellt sich die Frage, in welchen institutionellen Kontexten leistungsstarke Jugendliche in Naturwissenschaften anzutreffen sind, d. h. wie sie sich auf Schularten verteilen und in welchen Klassen sie gegebenenfalls gehäuft vorkommen. Auch im Hinblick auf die Identifikation leistungsstarker Jugendlicher ist dabei von Bedeutung, inwieweit leistungsstarke Jugendliche in Naturwissenschaften an ihren Schulnoten zu erkennen sind.

Obwohl Studien bereits wichtige Erkenntnisse zu individuellen und familiären Einflussfaktoren leistungsstarker Jugendlicher lieferten (z. B. OECD, 2009b; Zimmer, Brunner, Lüdtke, Prenzel & Baumert, 2007), wurden leistungsstarke Jugendliche in Naturwissenschaften bisher nur in einzelnen Merkmalen charakterisiert. Eine Frage, die bisher nicht ausreichend beantwortet wurde, ist daher, wer die leistungsstarken Jugendlichen in Naturwissenschaften sind, d. h. durch welche Merkmale sie sich einerseits auszeichnen und in welchen Merkmalen sie andererseits möglicherweise selbst eine heterogene Gruppe sind.

Neben dem Wissen über besondere Merkmale liegt eine weitere Herausforderung darin, leistungsstarke Jugendliche in Naturwissenschaften sowohl in ihrer Leistung als auch ihrem Selbstkonzept und Interesse zu fördern. Dem Naturwissenschaftsunterricht kommt hier als regelmäßige Gelegenheit für Erfahrungen und Lernprozessen in Naturwissenschaften eine besondere Bedeutung zu. Sowohl die Bildungspolitik als auch Lehrpersonen sollten für eine angemessene Förderung leistungsstarker Jugendlicher in Naturwissenschaften wissen, wie Curriculum- bzw. Unterrichtsinhalte und Unterrichts-

methoden zu wählen sind, damit leistungsstarke Jugendliche in Naturwissenschaften in ihrer Leistung und ihrem Interesse gezielt unterstützt werden.

Die Lehr-Lernforschung der letzten Jahrzehnte hat einerseits domänenübergreifende Unterrichtsaktivitäten (z. B. interaktives und kooperatives Lernen) und andererseits eher naturwissenschaftsspezifische Unterrichtsaktivitäten (z. B. forschend-entdeckendes Lernen und Anwendung von Naturwissenschaften in lebensnahen Kontexten) als Merkmale eines effektiven Naturwissenschaftsunterrichts identifiziert (Bransford & Donovan, 2005; Furtak, Seidel, Iverson & Briggs, 2012; Kobarg et al., 2011; Seidel & Shavelson, 2007). Differenzielle Sichtweisen von Unterricht zeigen jedoch, dass die Nutzung und Wirkung des Unterrichts von individuellen Merkmalen der Schülerinnen und Schüler wie der Leistung abhängt (Helmke, 2007; Slavin, 1987). Es stellt sich daher die Frage, inwiefern diese als effektiv geltende Unterrichtsaktivitäten die Leistung und das Interesse von Jugendlichen mit hohem Potential in Naturwissenschaften fördern. Bisher wurden diese Unterrichtsaktivitäten nur vereinzelt im Kontext leistungsstarker Jugendlicher, jeweils einzeln und jeweils nur im Zusammenhang mit Leistung oder motivational-affektiven Merkmalen untersucht (z. B. White & Frederiksen, 1998). Hier bleibt unklar, welche Unterrichtsaktivitäten die Leistung und das Interesse leistungsstarker Jugendlicher unterstützen und inwieweit auch weniger leistungsstarke Jugendliche von diesen Unterrichtsaktivitäten profitieren.

Eine besondere Problematik im Zusammenhang von Unterricht und Interesse stellen die leistungsstarken Jugendlichen in Naturwissenschaften dar, die sich in Relation zu ihrer Leistung nur gering für Naturwissenschaften interessieren (vgl. Prenzel & Schütte, 2007). Vor allem im Hinblick auf zukünftige Bildungswege und Karriereentscheidungen gilt es als „kritisch, wenn sich zahlreiche Schülerinnen und Schüler mit einem hohen Leistungspotenzial nicht in einem ähnlichen Maße für die Gegenstandsbereiche begeistern können“ (vbw, 2015, S. 155). Zentrale Studien zum Naturwissenschaftsunterricht wie beispielsweise die IPN-Videostudie (Seidel et al., 2006) zeigen, dass weniger die Organisation von Unterricht als vielmehr Qualitätsmerkmale wie die motivationsunterstützende Wahrnehmung von Unterricht ausschlaggebend für die intrinsische Motivation in Naturwissenschaften sind. Unterricht wird dabei als Angebot gesehen, das maßgeblich durch die Persönlichkeit, das Professionswissen und das Verhalten der Lehrperson geprägt wird (z. B. Shulman, 1986, 2001) und dessen Nutzung von der individuellen motivationsunterstützenden Wahrnehmung des Unterrichtsangebots durch die Schülerinnen

und Schüler abhängt (vgl. Fend, 2001; Helmke, 2007; Seidel, 2003). Studien zu individuellen Schülervoraussetzungen offenbaren darüber hinaus, dass sich unterschiedliche Schülerprofile hinsichtlich Leistung, Selbstkonzept und Interesse identifizieren lassen und sich diese Profile in der motivationsunterstützenden Wahrnehmung des Naturwissenschaftsunterrichts unterscheiden, d. h. unterschiedliche Mikro-Lernumgebungen (innerhalb einer Klasse) existieren können (z. B. Huber, Häusler, Jurik & Seidel, 2015; Seidel, 2006). Ein unterrichtsbezogener Grund dafür, dass ein Teil der leistungsstarken Jugendlichen motivational-affektiv nicht erreicht wird, ist daher möglicherweise, dass sie ihren Unterricht weniger motivationsunterstützend wahrnehmen als leistungsstarke Mitschüler mit günstigerem motivational-affektivem Profil. Für leistungsstarke Jugendliche in Naturwissenschaften ist bisher jedoch unklar, inwieweit Gruppen unterschiedlicher motivational-affektiver Profile differenziert werden können, die ihren Naturwissenschaftsunterricht unterschiedlich wahrnehmen.

1.2. Zielsetzung und Vorgehen

Diese Arbeit greift die oben beschriebenen Problematiken zu Merkmalen leistungsstarker Schülerinnen und Schüler, der Förderung von Leistung und Interesse leistungsstarker Schülerinnen und Schüler durch Unterrichtsaktivitäten sowie der Unterrichtswahrnehmung uninteressierter leistungsstarker Jugendlicher auf. Ziel dieser Arbeit ist es erstens, leistungsstarke Jugendliche zu charakterisieren und herauszufinden, welche besonderen individuellen, familiären und institutionellen Merkmale sie auszeichnen. Zweitens soll untersucht werden, welche Unterrichtsaktivitäten im Naturwissenschaftsunterricht Prädiktoren für die naturwissenschaftliche Kompetenz und das Interesse leistungsstarker Jugendlicher in Naturwissenschaften sind und welche Zusammenhänge sie für die anderen Schülerinnen und Schüler aufweisen. Drittens soll aufgezeigt werden, inwieweit unter den leistungsstarken Jugendlichen Gruppen mit unterschiedlichen Profilen aus Selbstkonzept und Interesse identifiziert werden können und mit einer unterschiedlich motivationsunterstützenden Wahrnehmung des Unterrichts einhergehen. Damit verfolgt diese Arbeit die Absicht, bestehende Stärken und Defizite in der Förderung leistungsstarker Jugendlicher auf Basis einer internationalen Vergleichsstudie zu identifizieren und Hinweise für zukünftige Förderbemühungen zu geben.

1.3. Gliederung der Arbeit

Zur Beantwortung der aufgeworfenen Fragen wird in der Arbeit zunächst in einem theoretischen Teil der Erkenntnisstand bezogen auf die Fragestellungen aufgearbeitet. Das bedeutet, es wird erstens dargelegt, wie leistungsstarke Jugendliche in dieser Arbeit definiert werden, und mögliche individuelle, familiäre und institutionelle Bedingungsfaktoren für hohe Leistung in Naturwissenschaften diskutiert. Letztere bilden den Suchraum für mögliche besondere Merkmale leistungsstarker Jugendlicher im Vergleich zu den restlichen Schülerinnen und Schülern. Zweitens werden theoretische Grundlagen von naturwissenschaftsbezogenem Interesse und Selbstkonzept erläutert sowie der Zusammenhang von Selbstkonzept und Interesse diskutiert. Dabei wird ein erster Fokus auf Selbstkonzept und Interesse im Kontext leistungsstarker Jugendlicher bzw. im Zusammenhang mit Leistung gelegt. Im Anschluss daran wird der Einfluss motivationsunterstützender Lehr-Lernbedingungen theoretisch modelliert. Drittens werden der Forschungsstand zum Einfluss von Merkmalen des Naturwissenschaftsunterrichts auf Leistung und Interesse aufgearbeitet und dabei interaktives Lehren und Lernen sowie forschend-entdeckendes Lernen und Anwendungen als effektiv geltende Unterrichtsaktivitäten in Naturwissenschaften im Kontext leistungsstarker Schülerinnen und Schüler diskutiert. Abschließend werden als Teil des dritten Theoriekapitels mögliche Unterschiede der Unterrichtswahrnehmung in Abhängigkeit individueller Schülervoraussetzungen und die Relevanz der Erfassung von Unterrichtsmerkmalen aus Schülersicht beschrieben.

Die daraus gewonnenen differenzierten Fragestellungen und Hypothesen werden in Kapitel 5 dargelegt und in Kapitel 6 für die Untersuchung der Fragestellungen relevanten Methoden dargestellt. Zentrale Ergebnisse zur Beantwortung der Fragestellungen werden in Kapitel 7 berichtet und abschließend in der Diskussion (Kapitel 8) aus methodischer, theoretischer und anwendungsbezogener Perspektive diskutiert.

2. Leistungsstarke Schülerinnen und Schüler in Naturwissenschaften

2.1. Definition und Forschungsstand im Kontext internationaler Vergleichsstudien

Ausgangspunkt dieser Arbeit ist der Befund, dass Deutschland im Vergleich zu Ländern wie Finnland und Neuseeland einen geringeren Anteil an Top Performern in Naturwissenschaft (Jugendlichen auf den beiden obersten Kompetenzstufen) in PISA aufweist (Prenzel & Schütte, 2007; Schiepe-Tiska et al., 2013) und ein bildungspolitisches Interesse an der Erhöhung des Anteils an Top Performern besteht (KMK, 2015). Ausgehend davon wird in dieser Arbeit eine verbreiterte Gruppe leistungsstarker Jugendlicher in den Blick genommen, die sich bereits durch hohe Leistungen und ein hohes Potential in Naturwissenschaften auszeichnet, in der jedoch noch nicht alle Jugendlichen die oberen beiden Kompetenzstufen erreicht haben. Die so definierte Gruppe leistungsstarker Jugendlicher wird *High Performer* genannt. Sie wird als Zielgruppe für Fördermaßnahmen hin zu einem international vergleichbaren Anteil an Top Performern angesehen und ihre Definition erfolgt daher orientiert an internationalen Maßstäben.

Ziel dieser Arbeit ist es, (a) diese High Performer in Naturwissenschaften in Deutschland mit Hilfe des Large-Scale-Datensatzes von PISA 2006 (OECD, 2007) im Vergleich zu den übrigen Jugendlichen zu charakterisieren, (b) zu untersuchen, inwieweit diese Gruppe leistungsstarker Jugendlicher in ihren besonderen Merkmalen der Leistungsspitze der Top Performer bereits gleicht sowie (c) den Zusammenhang ihrer Unterrichtswahrnehmung mit ihrer naturwissenschaftlichen Kompetenz und ihrem Interesse zu untersuchen und somit Hinweise auf eine Förderung dieser Jugendlichen hin zu Top Performance und Interesse für Naturwissenschaften zu erlangen.

Im nachfolgenden Kapitel wird zunächst die naturwissenschaftliche Grundbildung als verwendetes Leistungsmaß für die Definition leistungsstarker Jugendlicher begründet. In einem zweiten Unterkapitel wird daraufhin auf Basis bisheriger Forschungsansätze und internationaler Ergebnisse zu leistungsstarken Jugendlichen dargestellt, welche Größenordnung für die verbreiterte Leistungsspitze der High Performer angenommen werden kann und welche Schwelle daraus für die Definition dieser Gruppe abgeleitet wird.

2.1.1. Naturwissenschaftliche Kompetenz als zugrunde liegendes Leistungskonstrukt

Mit dem Rahmenkonzept der naturwissenschaftlichen Kompetenz (für eine detaillierte Beschreibung vgl. Bybee & McCrae, 2009; OECD, 2006; Prenzel et al., 2007) entwickelte PISA ein Leistungsmaß, mit dem nicht nur internationale Benchmarks und Vergleiche, sondern auch unterschiedlich leistungsstarke Gruppen innerhalb eines Staats identifiziert werden können (vgl. Zimmer et al., 2007). Das Rahmenkonzept der naturwissenschaftlichen Kompetenz in PISA basiert auf dem Konzept der *Scientific Literacy* und hat gemeinsam mit anderen Large-Scale-Assessments wie TIMSS (z. B. Baumert et al., 2000) *Scientific Literacy* als Maß für internationale und standardisierte Vergleiche naturwissenschaftlicher Bildungsergebnisse etabliert. Das Konzept der *Scientific Literacy* wurde dabei durch PISA in zweifacher Hinsicht erweitert: Zum einen wurde der Aspekt der Grundbildung eingeführt und *Scientific Literacy* in Teilkompetenzen, Wissensbereiche und Kontexte ausdifferenziert und vertieft. Zum anderen wurden motivational-affektive Merkmale im Sinne multipler Bildungsziele als zusätzliche Dimension berücksichtigt.

Naturwissenschaftliche Kompetenz wird als Grundbildung verstanden, weil sie für die generelle Teilhabe an der Gesellschaft erforderlich ist. Sie basiert deshalb auf der Frage, was Gesellschaftsmitglieder in Situationen, in denen sie mit Naturwissenschaften und Technik konfrontiert werde, wissen und können sollten (OECD, 2006). Über das ursprüngliche Literacy-Konzept hinaus werden innerhalb der naturwissenschaftlichen Kompetenz Teilkompetenzen und Wissensbereiche als Dimensionen unterschieden sowie nach Kontexten differenziert. Als Teilkompetenzen naturwissenschaftlicher Grundbildung unterscheidet die Rahmenkonzeption in PISA zwischen (1) *naturwissenschaftliche Fragestellungen erkennen*, (2) *naturwissenschaftliche Phänomene beschreiben*,

erklären und vorhersagen und (3) naturwissenschaftliche Evidenz nutzen, um Entscheidungen zu treffen (Prenzel et al., 2007, S. 66). Zu den Wissenskomponenten, die zu den Bereichen *naturwissenschaftliches Wissen* und *Wissen über die Naturwissenschaften* zusammengefasst werden, zählt neben deklarativem und prozeduralem Wissen auch das Verständnis naturwissenschaftlicher Denk- und Arbeitsweisen (Prenzel et al., 2007, S. 66). Das Grundverständnis von Kompetenz ist, dass Wissen flexibel in alltagsnahen Situationen angewendet werden kann. In der Operationalisierung von naturwissenschaftlicher Kompetenz in PISA spielt die Anwendung in verschiedenen Situationen eine zentrale Rolle und wird durch die Einbettung der Testaufgaben in lebensweltnahe Kontexte mit verschiedenen Inhaltsbezügen und Reichweiten erreicht (OECD, 2006). Die Kontextualisierung unterstützt dabei das Ziel, aktives und anwendbares im Gegensatz zu lediglich reproduziertem passiven Wissen abzufragen und begünstigt die Differenzierung zwischen sehr kompetenten und weniger kompetenten Schülerinnen und Schülern (Prenzel et al., 2007).

Teil des Kompetenzkonstrukts in PISA sind zudem motivational-affektive Merkmale wie das Interesse und die Wertschätzung naturwissenschaftlicher Vorgehensweisen (Prenzel et al., 2007). Bereits Weinert definierte Kompetenz als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösung in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, 2001, S. 27) und schloss damit motivational-affektive Merkmale in die Definition von Kompetenz mit ein. Mit der Konzeption des Rahmenmodells naturwissenschaftlicher Kompetenz in PISA werden diese und weitere Überlegungen zur Wichtigkeit motivational-affektiver Merkmale aufgegriffen und implementiert (Prenzel, Rost, Senkbeil, Häußler & Klopp, 2001; Prenzel et al., 2007; J. Rost, Walter, Carstensen, Senkbeil & Prenzel, 2004; Schiepe-Tiska et al., 2013).

Weil PISA die Zielpopulation der 15-Jährigen repräsentativ abbildet und der Naturwissenschaftstest auch im oberen Leistungsbereich ausreichend differenziert, kann eine Teilgruppe der Zielpopulation auf Basis der Testkonzeption und des Studiendesigns von PISA aus methodischer Sicht sinnvoll identifiziert und charakterisiert werden (OECD, 2009b; Zimmer et al., 2007). Aus inhaltlicher Sicht zeichnet sie sich dadurch aus, dass sie anschlussfähige und anwendungsbezogene Fähigkeiten erfasst sowie eine

domänenspezifische Definition ermöglicht.

Sowohl eine auf naturwissenschaftliche Ausbildungswege als auch auf die gesellschaftliche Teilhabe konzentrierte Sichtweise erfordert von leistungsstarken Schülerinnen und Schülern anschlussfähiges Wissen als Grundlage für nachfolgende Bildungsprozesse sowie für lebenslanges Lernen in Naturwissenschaften und die Fähigkeit, dieses Wissen und Verständnis anwenden zu können (vgl. Gräber & Nentwig, 2002; Prenzel et al., 2007; Zimmer et al., 2007). Eine Definition leistungsstarker Jugendlicher in Naturwissenschaften, die auf dem Konzept der naturwissenschaftlichen Kompetenz basiert, entspricht diesem Anspruch. Sie stellt sicher, dass so definierte leistungsstarke Jugendliche ihr naturwissenschaftliches Wissen in verschiedenen Kontexten und lebensweltnahen Situationen anwenden können. Weil naturwissenschaftliche Grundbildung in dieser Form eine wichtige Rolle für Lern- und Berufschancen, lebenslanges Lernen und die gesellschaftliche Teilhabe spielt (Bybee, McCrae & Laurie, 2009; Prenzel, Rost et al., 2001; Prenzel et al., 2007; J. Rost et al., 2004; Schiepe-Tiska et al., 2013), zeichnen sich auf Basis der naturwissenschaftlichen Kompetenz identifizierte leistungsstarke Schülerinnen und Schüler durch ein hohes Maß anschlussfähiger Kompetenzen und dadurch ein hohes Potential für zukünftige Leistungs- und Lernprozesse in Naturwissenschaften aus (OECD, 2009b; Zimmer et al., 2007).

Eine solche Definition leistungsstarker Schülerinnen und Schüler unterscheidet sich aus verschiedenen Gründen sowohl in der Zielstellung als auch in der Vorgehensweise von Ansätzen, die sich auf ein generelles kognitives Leistungspotential in Form von Intelligenz beziehen, beispielsweise in der Hochbegabungsforschung (D. H. Rost, 2009a): (a) sie ist domänenspezifisch angelegt, (b) sie schreibt dem Lebensweltbezug und der Anwendbarkeit von Wissen einen hohen Stellenwert zu und (c) sie fokussiert auf das durch die aktuelle Leistung bzw. das Vorwissen vorhandene Potential zum Weiterlernen im Unterschied zu einem generellen kognitiven Potential. Darüber hinaus schließt sie für den weiteren Ausbildungsweg und die gesellschaftliche Teilhabe wichtige motivational-affektive Merkmale der Bereitschaft, sich mit naturwissenschaftlichen Themen auseinanderzusetzen und das Vertrauen in die eigene Leistung explizit mit ein (Prenzel et al., 2007). Damit berücksichtigt die in dieser Arbeit verwendete Definition leistungsstarker Jugendlicher in Naturwissenschaften Studien, die zeigen, dass nicht der Intelligenzquotient, sondern die Vorleistung am verlässlichsten Leistungsexzellenz vorhersagt und eine hohe Intelligenz umgekehrt beispielsweise fachliche Lerndefizite in

der Grundschule nur bedingt kompensieren kann (Krajewski & Schneider, 2006; Lohman, 2005; Stern, 1994, 1997).

2.1.2. Definition von Top und High Performern in Naturwissenschaften

Nachdem naturwissenschaftliche Kompetenz als theoretisch zugrunde liegendes Konstrukt der Definition der High Performer charakterisiert wurde, stellt sich die Frage, wie die Gruppenzugehörigkeit auf Basis dieses Konstruktes definiert wird. Grundsätzlich sind mehrere Arten der Schwellensetzung für leistungsstarke Jugendliche in Naturwissenschaften denkbar, darunter eine relative Definition auf der Kompetenzskala, aber auch eine relative Definition auf Basis der Verteilung der Population.

Eine Schwellensetzung, bei der der Schwellenwert relativ auf einer inhaltlichen Kompetenzskala festgelegt wird, erfolgt beispielsweise durch die Definition von Kompetenzstufen (Adams, 2009). Eine solche Definition ermöglicht, die Größe des Anteils leistungsstarker Schülerinnen und Schüler eines Staates zu ermitteln sowie Anteile leistungsstarker Schülerinnen und Schüler international zwischen Staaten zu vergleichen. Erfolgt sie anhand der Kompetenzstufen in PISA, so bietet sie aufgrund der in Kombination mit der Kompetenzstufenbildung stattfindenden Beschreibung der kognitiven Anforderungen jeder Kompetenzstufe die Möglichkeit, die hohe Leistungsfähigkeit auch inhaltlich klar und anschaulich zu beschreiben und Kompetenzniveaus inhaltlich voneinander abzugrenzen (Carstensen, Frey, Walter & Knoll, 2007). So sind Schülerinnen und Schüler der Kompetenzstufe V in der Lage „sowohl ihr konzeptuelles Wissen als auch ihr Wissen über Naturwissenschaften auf (...) [komplexe] Situationen an[zu]wenden“ und „besitzen ein gut entwickeltes Verständnis naturwissenschaftlicher Untersuchungen, können ihr Wissen verknüpfen und Situationen kritisch bewerten“ (Prenzel et al., 2007, S. 77). Eine Definition leistungsstarker Schülerinnen und Schüler anhand der Kompetenzstufen wird deshalb international im Rahmen von PISA für die Gruppe der Top Performer in Naturwissenschaften (Jugendliche auf Kompetenzstufe V und VI) genutzt (Buccheri, Gürber & Brühwiler, 2011; OECD, 2007, 2009b). Deutschland verfügte in PISA 2006, 2012 und 2015 mit etwas mehr 10% über Top-Performer-Anteile, die jeweils über dem Mittelwert der OECD-Länder lagen (OECD, 2013b, S. 237). Die Definition der Top Performer eignet sich dazu, international vergleichend und mit klarer

inhaltlicher Beschreibung des erreichten Kompetenzniveaus die Leistungsspitze eines Staates abzubilden. Die Top Performer werden in dieser Arbeit daher als Definition für eine international vergleichbare Leistungsspitze verwendet, mit der die Gruppe der High Performer verglichen wird und für deren Vergrößerung die Gruppe der High Performer als Zielgruppe betrachtet wird.

Dagegen bildet eine relative Definition auf Basis der Populationsverteilung, beispielsweise das beste Viertel eines Staates, statt eines bestimmten Kompetenzbereichs einen bestimmten Abschnitt einer Populationsverteilung ab. Eine solche Definition ist zwar abhängig vom jeweils betrachteten Land und im Kompetenzniveau international nicht vergleichbar, sie ermöglicht jedoch eine sozial vergleichende Perspektive innerhalb eines Landes. Gleichzeitig erlaubt eine solche Definition einen in Anteil bzw. Größe festgelegten Bereich einer Population zu betrachten. Die verbreiterte Gruppe leistungsstarker Jugendlicher, High Performer genannt, wird in dieser Arbeit deshalb relativ an der Population definiert. Dies ermöglicht die Definition einer Gruppe, deren Größe sich am anvisierten Anteil an leistungsstarken Jugendlichen eines Landes orientiert und somit die Definition einer Gruppe, die die bereits existierende Leistungsspitze in Kombination mit einer Zielgruppe für die Förderung hin zu Top Performance enthält.

Bisher wurden solche relativ definierten Gruppen leistungsstarker Jugendlicher in Vergleichsstudien mit unterschiedlicher Größenordnung definiert. In PISA werden regelmäßig die besten 25% einer Domäne als leistungsstarke Jugendliche (OECD, 2007, 2013b; Prenzel, 2007) und im Rahmen vertiefter Analysen von Zimmer et al. (2007) auch die in allen Domänen zu den besten 10% gehörenden Schülerinnen und Schüler als „hochkompetente“ Jugendliche (Zimmer et al., 2007, S. 198) betrachtet. Auch im Rahmen des IQB-Ländervergleichs werden die besten 25% in einer Betrachtung zum Zusammenhang von Leistung und Interesse als leistungsstarke Schülerinnen und Schüler mit besonderem Fokus betrachtet und in drei Leistungsgruppen unterteilt: (a) besten 5%, (b) besten 5-10% und (c) besten 10-25% (Jansen, Schroeders & Stanat, 2013, S. 359). Die genannten relativen Definitionen orientierten sich vor allem an Einteilungen, die zur Untersuchung einer Verteilung dienen (z. B. Quartilen).

Da in dieser Arbeit die relativ definierte Gruppe der High Performer als Zielgruppe für Förderbemühungen hin zu Top Performance verstanden wird, soll ihre Größenordnung anders als in den berichteten Definitionen an internationalen Benchmarks zum Anteil an Top Performern orientiert werden. Wie bereits erwähnt, verfügte Deutschland

in PISA 2006 und 2012 mit etwas mehr 10% über Top-Performer-Anteile, die jeweils über dem Mittelwert der OECD-Länder lagen (Prenzel et al., 2007, S. 86; Schiepe-Tiska et al., 2013, S. 202). Andere westlich geprägte Länder wie Finnland und Neuseeland wiesen jedoch deutlich höhere Werte von 13% bis 21% auf, asiatisch geprägte Länder wie Japan und Singapur erreichten sogar Werte von 15% bis 24% (Prenzel et al., 2007, S. 86; Schiepe-Tiska et al., 2013, S. 202). In der mittleren naturwissenschaftlichen Kompetenz erreicht Neuseeland zum Vergleich jedoch lediglich einen zehn Punkte geringeren mittleren Kompetenzwert als Deutschland (Schiepe-Tiska et al., 2013, S. 199). Die Werte von Finnland und Neuseeland zeigen, dass kulturell mit Deutschland vergleichbare Staaten Anteile von um die 20% erreichen. Dies kann als internationaler Maßstab für einen realistisch-optimistischen Top-Performer-Anteil in Naturwissenschaften angesehen werden, den sich Deutschland zum Ziel setzen kann. Orientiert an dieser Größenordnung werden High Performer in dieser Arbeit als die besten 20% der deutschen Schülerinnen und Schüler in Naturwissenschaften definiert und als mögliche Zielgruppe für Förderbemühungen gesehen. Der Begriff der „High Performer“ wurde in bisherigen Untersuchungen im Rahmen von PISA im Gegensatz zu Begriffen wie „Top Performer“, „Strong Performer“, „Middle Performer“ oder „hochkompetente Jugendliche“ noch nicht verwendet und wurde daher in Abgrenzung zu den bereits mit einer Definition belegten Begriffen hier gewählt. In Relation zur Größenordnung der Top Performer in Deutschland in PISA 2006 (11.8%) sind die als beste 20% definierte High Performer in Naturwissenschaften eine ungefähr verdoppelte Leistungsspitze (OECD, 2013b, S. 237). Im Vergleich dazu wird diese Zielgruppe im Rahmen von TIMSS etwas größer aufgefasst, indem Schülerinnen und Schüler auf Kompetenzstufe IV und damit ungefähr die besten 30% als Zielgruppe für die Erhöhung des Anteils an Schülerinnen und Schüler auf der höchsten Kompetenzstufe V bezeichnet werden (Wendt, Willems, Tarelli, Euen & Bos, 2013, S. 32).

Auch wenn das in dieser Arbeit verwendete Verständnis von High Performance sich von der Hochbegabungsforschung unterscheidet, soll ein Blick in dieses Feld geworfen werden, um die Anteile, die als „hoch“ bezeichnet werden, zu beurteilen. Auch in der Hochbegabungsforschung lassen sich neben häufig verwendeten enger definierten Leistungsspitzen (z. B. 3%), ähnliche Größenordnungen zur Definition in dieser Arbeit finden. In seinem *Triad/Revolving Door System* (Renzulli, 1976, 1984) bezeichnet Renzulli (1984) die besten 15-20% als Talentpool bzw. Zielgruppe für Differenzierungs- und Förderungsbemühungen. Darüber hinaus beschreibt Mönks (2006) 15-20% der Kinder

als sehr begabt/hochbegabt und (Ziegler, Finsterwald & Grassinger, 2005) gibt an, dass ungefähr 16% der Jugendlichen als mindestens *middly gifted* bezeichnet werden können. Auch politisch werden diese Größenordnungen aufgegriffen, indem der ehemalige Kultusminister Dorgerloh einen Top-Performer-Anteil von 25% als Zielgrößenordnung für Deutschland vorschlägt (Vitzthum, 04.12.13). Auch in der Bildungspolitik in den Niederlanden beispielsweise werden die besten 20% der Schülerinnen und Schüler im Rahmen des Förderprogramms *Sirius Programma (2009-14)* als Zielgruppe der Exzellenzförderung gesehen werden (OECD, 2014; Sirius Programme, 2008).

Zusammenfassend werden in dieser Arbeit leistungsstarke Schülerinnen und Schüler in (a) Top Performer und (b) High Performer unterteilt. Unter Top Performern werden Schülerinnen und Schüler verstanden, deren naturwissenschaftliche Kompetenz auf Kompetenzstufe V oder VI in Naturwissenschaften in PISA liegt. Sie zeichnen sich durch eine anschaulich beschreibbare hohe Kompetenz in Naturwissenschaften aus. High Performer als verbreiterte Spitzengruppe werden als die besten 20% in Naturwissenschaften definiert und beinhalten sowohl Top Performer als auch Schülerinnen und Schüler, die bereits hohe Leistungen aber noch keine Top Performance aufweisen. Letztere können als mögliche Zielgruppe für Förderbemühungen hin zu Top Performance gesehen werden.

2.2. Individuelle, familiäre und institutionelle Merkmale leistungsstarker Jugendlicher in Naturwissenschaft

Mit der Frage nach Merkmalen, die leistungsstarke Jugendliche in Naturwissenschaften auszeichnen, werden besonderer Merkmale angesprochen, in denen sich leistungsstarke Jugendliche im Vergleich von den übrigen Schülerinnen und Schülern abgrenzen. Damit in Verbindung steht auch die Frage danach, inwiefern der Talentpool in Deutschland ausgeschöpft wird und Schülerinnen und Schüler mit einem hohen Potential in Naturwissenschaften unabhängig von individuellen, familiären und institutionellen Merkmalen Top Performance erreichen oder umgekehrt bestimmte besondere Merkmale aufweisen.

Ansätze zu Bedingungen der Schulleistung (z. B. Schrader & Helmke, 2008) sowie schulischen und außerschulischen Lernumwelten (Doll & Prenzel, 2004b) und Metaanalysen zu Einflussfaktoren auf Lernen (M. C. Wang, Haertel & Walberg, 1993) geben einen Überblick über mögliche Bedingungsfaktoren für Lernentwicklungen. Die Forschungszugänge dieser Arbeiten werden von einer differenzialpsychologischen Perspektive geprägt und verfolgen regressionsanalytische bzw. vergleichende Vorgehensweisen. Im Unterschied zu prozessorientierten Ansätzen untersuchen sie keine kausalen Zusammenhänge, sondern konzentrieren sich auf die Erklärung der Varianz einer abhängigen Variable, indem sie Prädiktoren für unterschiedliche Ausprägungen dieser Variable identifizieren. Analog zum differenziellen Vorgehen werden besonderer Merkmale leistungsstarker Schülerinnen und Schüler in dieser Arbeit als Merkmale verstanden, die mit hoher naturwissenschaftlicher Kompetenz assoziiert sind und leistungsstarke Schülerinnen und Schüler von anderen Jugendlichen unterscheiden.

Als ein Vertreter dieser Forschungszugänge zu Prädiktoren von Lernergebnissen spannt unter anderem das Modell schulischer und außerschulischer Lernumwelten aus dem DFG-Schwerpunktprogramm „Bildungsqualität von Schule (BIQUA)“ (Doll & Prenzel, 2004a) einen Rahmen auf, der sich für eine systematische Charakterisierung leistungsstarker Jugendlicher in Naturwissenschaften eignet. Aufbauend auf der Annahme systemisch vernetzter Lernumwelten (Bronfenbrenner, 1981) wird darin unter anderem zwischen individuellen, familiären und institutionellen Bedingungsfeldern für Lernergebnisse unterschieden und verschiedene Merkmale zur Operationalisierung dieser Bereiche genannt. Insbesondere individuelle Merkmale wie kognitive Voraussetzungen und Geschlecht, familiäre Bedingungsfaktoren zur sozialen und sprachlichen

Herkunft sowie institutionelle Rahmenbedingungen wie die Schulart und Klassenzusammensetzung werden dort genannt und haben sich in verschiedenen Studien als prädiktiv für die Leistungsentwicklung erwiesen (z. B. Baumert, Trautwein & Artelt, 2003; Baumert, Stanat & Watermann, 2006b; Ehmke & Baumert, 2007; Prenzel et al., 2007). Diese individuellen, familiären und institutionellen Bedingungsfaktoren werden nachfolgend als mögliche besondere Merkmale von Top und High Performance näher betrachtet.

2.2.1. Individuelle Merkmale

Als individuelle Merkmale sind im Zusammenhang mit Top und High Performance Faktoren relevant, die dem Schüler bzw. der Schülerin zugeordnet werden und im Zusammenhang mit Leistung bzw. Lernprozessen in Naturwissenschaften stehen. In alltagspsychologischen Vorstellungen wird eine hohe Leistung in Naturwissenschaften häufig mit individuellen Merkmalen wie hohen kognitiven Fähigkeiten, einer gleichzeitig geringeren sprachlichen Fähigkeit oder dem Geschlecht assoziiert. Welche empirischen Belege oder widersprechende Evidenz sich für solche Vorstellungen finden lassen und welche Forschungslücken im Kontext leistungsstarker Schülerinnen und Schüler bestehen, wird in den folgenden Abschnitten dargestellt.

Intelligenz

Wie bereits im ersten Kapitel des theoretischen Teils diskutiert, wurde in der Definition von Top und High Performance bewusst ein domänenspezifisches kompetenzbezogenes Vorgehen im Unterschied zu einer Definition über allgemeine kognitive Fähigkeiten gewählt. Es stellt sich jedoch die Frage, welche Rolle die allgemeine kognitive Fähigkeit bzw. Intelligenz als domänenübergreifendes Konstrukt für eine hohe naturwissenschaftliche Kompetenz spielt. Da die Intelligenz ein vergleichsweise stabiles individuelles Merkmal ist (Asendorpf & Neyer, 2012), ist für die Identifikation und Förderung von Jugendlichen mit hohem Potential in Naturwissenschaften von hoher Bedeutung, ob und in welcher Form die Intelligenz eine relevante Voraussetzung für das Lernen in Naturwissenschaften ist.

Verschiedene Forschungsbereiche wie die Hochbegabungsforschung nehmen an, dass die Intelligenz Lernprozesse und Leistungsentwicklungen weitgehend vorhersagen kann (Heller, 2004). Durchschnittlich finden Studien eine Korrelation von $r = .50$

zwischen Intelligenz und Schulleistung (Schrader & Helmke, 2008, S. 291), wobei sich je nach Alter, verwendetem Intelligenzmaß und Art der Leistung beträchtliche Unterschiede ergeben (für eine genauere Übersicht siehe Gustafson & Undheim, 1996; D. H. Rost, 2009c). Die Intelligenz dient zwar häufig als Identifikationsinstrument für Jugendliche mit hohem Leistungspotential, ist dabei aber wegen der geringen Varianzaufklärung domänenspezifischer Leistungen und dadurch zu erwartender Fehlzuordnungen umstritten (vgl. Ziegler & Stoeger, 2012).

Forschung zum Expertiseansatz (vgl. Bloom & Sosniak, 1985; Ericsson, 1996) vertritt dagegen die Annahme, dass für hohe fachspezifische Leistungen lediglich eine Intelligenz oberhalb einer bestimmten Schwelle erreicht werden muss (vgl. W. Schneider, 1997; W. Schneider & Stumpf, 2007). Leistungsexzellenz wird oberhalb dieser Schwelle nicht durch höhere Intelligenz, sondern durch gezieltes und intensives Üben („deliberate practice“) erreicht. (vgl. W. Schneider, 1997; W. Schneider & Stumpf, 2007). Die in diesem Zusammenhang vielfach genannte 10-Jahres-Regel, nach der für Leistungsexzellenz ungefähr zehn Jahre intensives Üben notwendig sind, wurde in vielen Fachbereichen belegt (vgl. Ericsson, Krampe & Tesch-Römer, 1993). Damit vertritt die Expertiseforschung eine Schwellenannahme und betrachtet die kognitive Fähigkeit nicht als linearer Prädiktor hoher fachspezifischer Leistungen.

Verschiedene umfangreiche Längsschnittstudien, vor allem im Bereich der Mathematik, unterstützen diese Annahme, indem sie die Rolle des Vorwissens im Vergleich zur Intelligenz betrachten und zeigen, dass bereichsspezifisches Vorwissen der stärkste Prädiktor von Leistung bzw. Leistungsexzellenz ist (vgl. Krajewski & Schneider, 2006; Prenzel, Reiss & Hasselhorn, 2009; Schraw, 2006; Stern, 1994, 1997). Aus kognitionspsychologischer Sicht ist die hohe Bedeutung des Vorwissens darauf zurückzuführen, dass ein nachhaltiger Wissenserwerbsprozess auf der Verknüpfung von Lerninhalten mit bestehendem Wissen beruht und dadurch in seiner Geschwindigkeit und Effektivität entscheidend durch die Quantität und Qualität (Vernetztheit und Dichte) vorhandener Wissensnetzwerke beeinflusst wird (vgl. Renkl, 1996). Insgesamt legen Forschungsergebnisse nahe, dass der Wissenserwerbsprozess vor allem am Anfang durch eine hohe Intelligenz begünstigt wird, im weiteren Verlauf jedoch das Vorwissen zunehmend an Bedeutung gewinnt und auch eine hohe Intelligenz Rückstände im Vorwissen nicht kompensieren kann (Krajewski & Schneider, 2006; Stern, 1994, 1997).

Überträgt man diese Ergebnisse auf die in dieser Arbeit betrachteten leistungs-

starken 15-Jährigen und die Kompetenz im Bereich der Naturwissenschaften, so ist entsprechend des Expertiseansatzes anzunehmen, dass High und Top Performer zwar kognitive Fähigkeiten oberhalb einer gewissen Schwelle aufweisen, höhere kognitive Fähigkeiten oberhalb dieser Schwelle jedoch nicht zwingend hohe naturwissenschaftliche Leistungen mit sich bringen. Vielmehr ist davon auszugehen, dass hohe naturwissenschaftliche Kompetenz ab einer gewissen kognitiven Fähigkeit vor allem aus über zehn Jahren gezieltem und intensivem Üben resultiert. Inwieweit sich eine solche Schwellenannahme auch für High Performer in Naturwissenschaften im Kontext der Large-Scale-Assessments nachweisen lässt, wurde in bisheriger Forschung noch nicht ausreichend untersucht.

Kompetenz in Mathematik und Lesen

Wie oben erläutert, nehmen Vorkenntnisse Einfluss auf Lernprozesse, indem sie Verstehensprozesse erleichtern (z. B. Schrader & Helmke, 2008). Dies kann auch fächerübergreifend zum Tragen kommen, wenn beispielsweise mathematische Kenntnisse zum Verstehen und Lösen physikalischer Inhalte und Aufgabenstellungen nötig sind. Mathematik spielt in den Naturwissenschaften eine wichtige Rolle und dient insbesondere in Physik als Werkzeug, Sprache und Methode für Modellierungsprozesse (Uhlen, Karam, Pietrocola & Pospiech, 2012). Mathematische Fähigkeiten werden als Voraussetzung aber nicht hinreichende Bedingung für Kompetenzen bzw. Lernprozesse in Physik gesehen (Hudson & McIntire, 1977). Es ist daher anzunehmen, dass mathematische und naturwissenschaftliche Kompetenz identische Elemente enthalten und daher ein gewisser Überlappungsbereich zwischen beiden Kompetenzen besteht. Im Kontext leistungsstarker Jugendlicher in Naturwissenschaften ist daher insbesondere die Frage interessant, wie viele der High Performer in Naturwissenschaften auch zu den High Performern in Mathematik gehören.

Neben Fähigkeiten in Mathematik spielen aber auch das Textverstehen und andere sprachliche Fähigkeiten eine wichtige Rolle in Naturwissenschaften. Naturwissenschaftliches Wissen ist häufig in Textform gespeichert und wird über Texte kommuniziert. Im Naturwissenschaftsunterricht, im Studium und auch im beruflichen Leben werden Inhalte und Aufgaben zum Großteil in unterschiedlichen Textsorten präsentiert und eine schriftliche, ausformulierte Lösung wird verlangt. Cano, García, Berbén und Justicia (2014) konnten beispielsweise in einer querschnittlichen Studie mit ca. 600 Neuntkläss-

lern zeigen, dass Schülerinnen und Schüler mit einem höheren Leseverständnis tiefere Lernstrategien und höhere Schulleistungen in Naturwissenschaften aufweisen. Auch in Large Scale Assessments wie PISA oder TIMSS/IGLU erfordert das Lösen naturwissenschaftlicher Aufgaben neben dem Abruf und Transfer naturwissenschaftlicher Konzepte einerseits mathematische Fähigkeiten anzuwenden und andererseits in Texten gegebene naturwissenschaftliche Informationen zu verstehen (Bybee et al., 2009). Zwischen naturwissenschaftlicher Kompetenz und Lesekompetenz sind daher ebenso wie zwischen naturwissenschaftlicher und mathematischer Kompetenz gemeinsame Elemente und daher ein Überlappungsbereich anzunehmen. Auch hier stellt sich daher die Frage, welcher Anteil der High Performer in Naturwissenschaften ebenso zu den High Performern in Lesen gehört und damit sowohl naturwissenschaftlich als auch sprachlich leistungsstark ist.

In ersten Analysen zu diesen Überlappungsbereichen zeigt eine Untersuchung der OECD, dass im OECD-Durchschnitt 2.8% (Dtl.: 3.9%) der Schülerinnen und Schüler sowohl in Naturwissenschaften als auch Mathematik Top Performer (Schülerinnen und Schüler auf Kompetenzstufe V und VI) sind, 0.8% (Dtl.: 0.9%) sowohl in Naturwissenschaften als auch Lesen (OECD, 2009b) und 5.2% in allen drei Domänen zu den Top Performern gehören (OECD, 2009b, S. 81). Innerhalb der Top Performer in Naturwissenschaften sind damit 44.1% (Dtl.: 44.2%) auch in Mathematik und Lesen Top Performer (OECD, 2009b, S. 81).

Die Ergebnisse des Berichts der OECD zeigen, dass ein Großteil der Top Performer in Naturwissenschaften auch in Mathematik und Lesen zur Leistungsspitze gehört (OECD, 2009b). Dennoch bleibt offen, ob dieser Zusammenhang auch in einer breiteren Gruppe der High Performer anzutreffen ist. Unklar ist darüber hinaus, inwiefern Top und High Performer diese Bild nicht nur in standardisierten Kompetenztests, sondern auch in ihrer Schulleistung aufweisen.

Geschlecht

Auch das biologische Geschlecht stellt ein Merkmal dar, in dem sich leistungsstarke Schülerinnen und Schüler im Mittel von den restlichen Jugendlichen möglicherweise unterscheiden. Wie die anderen Merkmale wird das Geschlecht hier aus einer differenziellen Perspektive betrachtet, d. h. es stehen Unterschiede zwischen verschiedenen Leistungsgruppen im Mittelpunkt. Nicht im Fokus dagegen steht eine prozessuale Per-

spektive, die sich mit dem biologischen Geschlecht als Prädiktor und dahinter stehenden vermittelnden Prozessen und Mechanismen beschäftigen würde.

Geschlechterunterschiede in Naturwissenschaften werden seit Jahrzehnten intensiv in verschiedenen Kontexten untersucht und die Ergebnisse zeigen ein uneinheitliches Bild (vgl. Übersicht von Brotman & Moore, 2008; Scantlebury & Baker, 2007). In einigen Studien zeigen sich Unterschiede zugunsten der Jungen in physikalischen und naturwissenschaftlichen Fähigkeiten (Beller & Gafni, 1996; Dresel, Ziegler & Broome, 1998). Demgegenüber finden sich jedoch auch vermehrt Studien, die eine differenziertere empirische und theoretische Sicht auf Geschlechterdifferenzen vermitteln und häufig keine Geschlechterunterschiede nachweisen (Benbow, Lubinski & Hyde, 1997; Hyde, Fennema & Lamon, 1990), darunter auch Large-Scale-Assessments wie PISA. In PISA 2006 unterschieden sich Mädchen und Jungen international nicht in ihrer naturwissenschaftlichen Gesamtkompetenz, zeigten jedoch Unterschiede in den Teilkompetenzen (OECD, 2007). Jungen übertrafen Mädchen in der Teilkompetenz „Phänomene naturwissenschaftlich erklären“. Mädchen erreichten dagegen in der Teilkompetenz „Naturwissenschaftliche Fragestellungen identifizieren“ – auch unter Kontrolle der Lesekompetenz – signifikant höhere Werte als Jungen. Keine Unterschiede waren bei der Teilkompetenz „Naturwissenschaftliche Evidenz nutzen“ vorhanden (OECD, 2007). Unklar ist jedoch, wie hoch die Effektstärken und damit wie bedeutsam die Geschlechterunterschiede in diesen Teilkompetenzen waren.

Auch im Bereich der Forschung zu Hochbegabung und Leistungsexzellenz finden Studien Leistungsunterschiede in Naturwissenschaften zugunsten der Jungen (Colangelo, Assouline, Cole, Cutrona & Maxey, 1996; Lee & Burkam, 1996). Beller und Gafni (1996) berichten zudem über eine Zunahme der Differenz vom Kindesalter bis hin zur Adoleszenz, Beerman, Heller und Menacher (1992) über eine Zunahme der Unterschiede mit steigendem Kompetenzniveau. Demgegenüber stehen jedoch Studien, die ähnliche oder sogar bessere Leistungen von Mädchen in Naturwissenschaften finden (Freeman, 2004; Pajares, 1996; Stocking & Goldstein, 1992) und kaum Unterschiede in der mathematisch-naturwissenschaftlichen Begabung zwischen Mädchen und Jungen erkennen (Hedges & Nowell, 1995; Naglieri & Rojahn, 2001). Für die domänenübergreifend hochkompetenten 15-Jährigen in Deutschland (Zimmer et al., 2007) ergaben sich keine Geschlechterunterschiede aber sowohl international als auch für Deutschland wiesen die Jungen höhere Anteile an Top Performern in Naturwissenschaften auf als die Mädchen (OECD, 2009b;

Prenzel et al., 2007).

Während Ergebnisse standardisierter Tests keine oder Geschlechterunterschiede zugunsten der Jungen feststellten, zeigte sich für Schulleistungen ein entgegengesetztes Bild. Mädchen zeigten im Vergleich zu Jungen sowohl in der Grundschule als auch in der Sekundarstufe bessere oder ähnliche Noten in allen Fächern – Naturwissenschaften eingeschlossen (Hannover & Kessels, 2011; M. Neumann, Milek, Maaz & Gresch, 2010; Valtin, Wagner & Schwippert, 2005).

Eine bisher unbeantwortete Frage ist jedoch, ob die in der Betrachtung der Top Performer der OECD (OECD, 2009b) gefundenen Geschlechterunterschiede auch unter den High Performern auftreten und inwiefern sich Geschlechterunterschiede in der Spitzengruppe seit 2006 verändert haben. Darüber hinaus ist unklar, inwiefern Geschlechtsdifferenzen auch in Schulnoten leistungsstarker Schülerinnen und Schüler in Naturwissenschaften festzustellen sind und welcher Zusammenhang zu Kompetenzprofilen standardisierter Testleistungen besteht.

2.2.2. Familiäre Merkmale

Die Lebenslagen von 15-Jährigen unterscheiden sich unter anderem durch die Lernumwelten und damit verbundenen sozialen Ungleichheiten ihrer Familien. Die Soziologie unterscheidet zwischen vertikalen Ungleichheiten, die mit unterschiedlichen sozialen Klassen oder Schichten einer Gesellschaft verbunden sind (z. B. Bildungsstand, Beruf, Vermögen) und horizontalen Ungleichheiten, die Unterschiede innerhalb einer sozialen Schicht (z. B. durch kulturelle oder regionale Unterschiede) beschreiben (z. B. Büchner, 2003). Darüber hinaus wird zwischen Strukturmerkmalen wie dem sozioökonomischen Hintergrund oder dem Bildungsniveau der Eltern und Prozessmerkmalen wie den kulturellen und sozialen Lebensverhältnissen unterschieden (Watermann & Baumert, 2006). Beide Unterscheidungen berücksichtigen ausgehend von den sozialkapitaltheoretischen Ansätzen von Bourdieu (1998, 1983) und Coleman (1988) neben dem sozioökonomischen Status als relative Stellung der Eltern innerhalb der sozialen Hierarchie einer Gesellschaft auch das kulturelle und soziale Kapital der familiären Lernumgebung, d. h. vertikale Unterschiede. Auch im Bereich der familiären Merkmale wird hier ein differenzieller Ansatz, der sich mit Merkmalsunterschieden beschäftigt, im Gegensatz zu einem prozessualen Ansatz zu Wirkmechanismen und vermittelnden Prozessen (z. B. Lerngelegenheiten, Ressourcen oder elterliche Unterstützung) verfolgt.

Im Vergleich dieser Prädiktoren konnten internationale Vergleichsstudien zeigen, dass der sozioökonomische Status vor dem Migrationsstatus, dem Kulturbesitz und Besuch von Kindergärten die stärkste Vorhersagekraft für Leistung aufweist (z. B. Ehmke, Hohensee, Heidemeier & Prenzel, 2004). Neben primären Effekten konnten sekundäre Effekte wie eine höhere Wahrscheinlichkeit, ein Gymnasium zu besuchen, als sekundäre Wirkmechanismen identifiziert werden (Baumert & Maaz, 2006; Maaz, Baumert & Trautwein, 2011; Ehmke & Baumert, 2007; T. Schneider, 2004).

Da Merkmale wie Einkommen, gesellschaftliche Anerkennung oder Einfluss in Large-Scale-Assessments nur schwer erfasst werden können (z. B. aus Datenschutzgründen), wird der sozioökonomische Status üblicherweise über Informationen zur Berufstätigkeit operationalisiert und verschiedene Indizes verwendet (für eine Übersicht siehe Ehmke & Baumert, 2007; Müller & Ehmke, 2013; Watermann & Klieme, 2006), darunter vor allem der *International Socio-Economic Index of Occupational Status (ISEI)* (Ganzeboom, de Graaf, Treiman & de Leeuw, 1992; Ganzeboom & Treiman, 1996), die *Erikson-Goldthorpe-Portocarero-Klassifikation (EGP)* (Erikson, Goldthorpe & Portocarero, 1979; Erikson & Goldthorpe, 2002) und der *Index of Economic, Social and Cultural Status (ESCS)* (OECD, 2009b). Der ISEI ordnet Berufe einer Berufsrangskala zu und hat einen Wertebereich von 16 bis 90 Punkten. Die Rangskala bildet Merkmale von Berufen ab, die sich als Vermittler zwischen Bildungsabschluss und Einkommen erwiesen haben (Ganzeboom & Treiman, 1996). Der ISEI wird daher mit der funktionalen Bedeutung, den sozioökonomischen Status basierend auf dem Status des Berufs und in Form eines quantitativen eindimensionalen Wertes abzubilden, verwendet. Üblicherweise wird der höchste ISEI-Wert der beiden Elternteile (HISEI = Highest ISEI) als Maß für den sozioökonomischen Status der Familie verwendet. Der HISEI gilt als Standardindikator in vielen internationalen Vergleichen (Ehmke & Baumert, 2007). Während der HISEI bzw. ISEI den sozioökonomischen Status auf Basis des Berufs abbilden, schließt der ESCS zusätzlich Prozessmerkmale wie kulturelle und ökonomische Merkmale mit ein (vgl. Ehmke & Baumert, 2007). Er wurde induktiv und mit dem Ziel einer maximalen Vorhersagekraft konstruiert (vgl. Ehmke & Baumert, 2007).

Die EGP-Klassifikation stellt im Unterschied zum eindimensionalen kontinuierlichen Index des ISEI eine kategoriale Einteilung dar und verfolgt die Absicht, qualitative Unterschiede zwischen sozialen Schichten (z. B. gesellschaftliche Anerkennung, Zugang zu Bildung und Macht) sichtbar zu machen. Diese qualitativen Unterschiede werden

in der EGP-Klassifikation dadurch erreicht, dass die Berufe nach Tätigkeit, Stellung im Beruf, Weisungsbefugnis und erforderlichen Qualifikationen geordnet werden (Ehmke & Baumert, 2007). Der kategoriale Ansatz der EGP-Klassen ist theoretisch besser fundiert als der HISEI und ermöglicht es, anschaulich zu untersuchen, inwiefern Jugendliche aus unterschiedlichen Entwicklungsmilieus für naturwissenschaftliche Kompetenz stammen (vgl. Baumert & Maaz, 2006; Ehmke & Baumert, 2007).

Betrachtet man den Forschungsstand zu leistungsstarken Schülerinnen und Schülern, so zeigen PISA 2000 und PISA 2003, dass vielseitig hochkompetente Schülerinnen und Schüler im Mittel einen höheren sozioökonomischen Status als weniger kompetente Jugendlichen aufweisen, wobei sich mehr als die Hälfte der hochkompetenten Jugendlichen in Deutschland dem obersten Quartil des sozioökonomischen Status und nur ungefähr 5% dem untersten Quartil zuordnen lassen (Zimmer et al., 2007). Auch in PISA 2006 haben Top Performer in Naturwissenschaften international und in Deutschland im Mittel einen niedrigeren sozioökonomischen Status (OECD, 2009b). In TIMSS/IGLU zeigt sich bereits für leistungsstarke Schülertypen unter Viertklässlern, dass sich Sozialschicht und Bildungsabschluss der Eltern von denen weniger leistungsstarken Viertklässler unterscheidet (Wendt et al., 2013).

Obwohl PISA hier bereits Erkenntnisse zu Top Performern in Naturwissenschaften und fächerübergreifend leistungsstarken Jugendlichen liefern konnte, wurde der sozioökonomische Hintergrund leistungsstarker Jugendlicher in Naturwissenschaften noch nicht abschließend untersucht. Der Großteil der genannten Betrachtungen des sozioökonomischen Status leistungsstarker Schülerinnen und Schüler in Naturwissenschaften erfolgte bisher auf Basis des ESCS oder HISEI und ließen eine genauere Betrachtung der Kategorisierung der Elternberufe nach EGP-Klassen außer Acht (vgl. OECD, 2009b; Zimmer et al., 2007) bzw. untersuchte sie nur für jüngere Schülerinnen und Schüler und verfolgten einen domänenübergreifenden Fokus (vgl. Wendt et al., 2013). Gleichzeitig wurde bisher vor allem der mittlere Unterschied verschiedener Leistungsgruppen untersucht. Inwieweit leistungsstarke Schülerinnen und Schüler jedoch selbst eine heterogene Gruppe darstellen und welcher Anteil leistungsstarker Schülerinnen und Schüler in diesem Zusammenhang beispielsweise aus nicht-akademischen Bildungshäusern stammt, wurde bisher nicht als Fragestellung betrachtet. Hier bleiben noch einige Fragen offen, die diese Arbeit aufgreift.

2.2.3. Institutionelle Rahmenbedingungen

Die Schule stellt einen zentralen institutionellen Rahmen für Lernprozesse in Naturwissenschaften dar, indem sie allen Schülerinnen und Schülern die Gelegenheit für eine regelmäßige und systematische Auseinandersetzung mit naturwissenschaftlichen Inhalten bietet. Unter den verschiedenen Schularten hat vor allem das Gymnasium den Bildungsauftrag zur Förderung leistungsstarker Schülerinnen und Schüler (Heller, 2002b). Baumert und Köller (1998b) schlussfolgern aus Ergebnissen zu TIMSS, dass Jugendliche bei gleicher Ausgangsleistung auf Gymnasien eine höhere Kompetenz erreichen als auf Realschulen, Gesamtschulen oder Hauptschulen. Aufgrund der differenziellen Lernumgebungen und den Kompositionseffekten, die Schularten und Bildungsgänge in der Sekundarstufe bieten, kann die Schulart ein entscheidender Prädiktor für hohe Leistungen in Naturwissenschaften sein. Zimmer et al. (2007, S. 201) konnten zeigen, dass mehr als 90% der Jugendlichen, die in allen drei Domänen in PISA 2000 und 2003 zu den besten 10% in Deutschland gehören, ein Gymnasium besuchen und ungefähr ein Zehntel der Gymnasiasten zu dieser Gruppe der vielseitig Hochkompetenten gehört. Wie aus Selektionsprozessen in mehrgliedrigen Schulsystemen zu erwarten, zeigte die internationale Betrachtung der Top Performer in Naturwissenschaften, dass leistungsstarke Schülerinnen und Schüler Schulen mit höherer durchschnittlicher Leistung besuchen (OECD, 2009b). Darüber hinaus wiesen auch die Klassen, denen Top Performer angehören, im Mittel höhere Kompetenzen auf als die Klassen, denen Nicht-Top Performer angehören (OECD, 2009b). Aufgrund der internationalen Betrachtung und der Betrachtung 15-Jähriger konnten bildungssystemspezifische Rahmenbedingungen wie Schularten nicht berücksichtigt und keine vollständigen Schulklassen untersucht werden (OECD, 2009b). Eine Analyse der Zusammenhänge von Schulart, Klassenleistung und -komposition mit hoher Leistung in Naturwissenschaften auf Basis einer klassenbasierten Stichprobe könnte hier weitere Erkenntnisse liefern.

2.2.4. Zusammenfassung

Die Bildungsforschung hat zahlreiche individuelle, familiäre und institutionelle Faktoren identifiziert, die in einem prädiktiven Zusammenhang zur Leistungsentwicklung von Jugendlichen stehen. Ausgehend von bestehenden Forschungsergebnissen bezüglich individueller Merkmale ist anzunehmen, dass Fähigkeiten in Mathematik und Lesen

sowie kognitive Grundfähigkeiten und das Geschlecht Eigenschaften sind, in denen sich leistungsstarke Schülerinnen und Schüler in Naturwissenschaften von den anderen Schülerinnen und Schülern unterscheiden. Gleichzeitig ist davon auszugehen, dass innerhalb der leistungsstarken Schülerinnen und Schüler eine gewisse Varianz in diesen Merkmalen besteht und die kognitive Grundfähigkeiten weniger ein Prädiktor, sondern bis zu einem gewissen Niveau eine Schwelle für hohe naturwissenschaftliche Leistungen sind.

Im familiären Bedingungsfeld konnte der sozioökonomische Hintergrund in zahlreichen Untersuchungen als wichtigster Prädiktor für Leistung identifiziert werden und kann in Form verschiedener Indizes überwiegend über die Berufszugehörigkeit erfasst werden. Analog zu den individuellen Merkmalen ist für leistungsstarke Jugendliche unklar, welche Heterogenität innerhalb der Gruppe existiert. Darüber hinaus konzentrieren sich bisherige Untersuchungen auf kontinuierliche Indizes und qualitative Unterschiede in der Schichtzugehörigkeit, wie durch die EGP-Klassen erfasst, wurden nicht betrachtet.

Als wesentliches Merkmal institutioneller Rahmenbedingungen wurde die Schulart und Klassenzusammensetzung für leistungsstarke Schülerinnen und Schüler bisher nur in Form des mittleren Anteils der Gymnasiasten (Zimmer et al., 2007) und der mittleren Schulleistung der Schulen von leistungsstarken Jugendlichen in Naturwissenschaften (OECD, 2009b) betrachtet. Vertiefte Analysen zur Verteilung auf die Schularten sowie der mittleren Klassenleistung und -heterogenität auf Grundlage eines klassenbasierten Datensatzes stehen hier noch aus und könnten bisherige Erkenntnisse zu institutionellen Prädiktoren für hohe Leistung in Naturwissenschaften ergänzen.

3. Naturwissenschaftsbezogenes Interesse und Fähigkeitsselbstkonzept

Interesse an den Naturwissenschaften bezeichnet die grundsätzliche Bereitschaft, sich zukünftig mit naturwissenschaftlichen Inhalten auseinanderzusetzen und ist neben leistungsbezogenen Zielen ein wichtiges Ziel von Bildung und Unterricht (Gottfried, Marcoulides, Gottfried & Oliver, 2009; Krapp & Prenzel, 2011; Kunter, 2005; Rakoczy, Klieme & Pauli, 2008). Im Kontext High Performance in Naturwissenschaften hat Interesse für Naturwissenschaften im Sinne (a) einer Ausdifferenzierung von Interesse hin zu Naturwissenschaften und (b) einer allgemeinen Interessiertheit eine hohe Bedeutung (vgl. Krapp & Prenzel, 1992, 2011; Prenzel, 2000; Prenzel & Schütte, 2007). Bildungsweg- und Berufswahlentscheidungen von Jugendlichen sind in hohem Maße von ihrem persönlichen Interesse für Naturwissenschaften abhängig (Patrick, Care & Ainley, 2010; Prenzel & Schütte, 2007; Taskinen, Schütte & Prenzel, 2013; Taskinen, Asseburg & Walter, 2008). Die Klärung und Ausdifferenzierung dieses Interesses ist Teil der Persönlichkeitsentwicklung und sollte durch Schule und Bildungseinrichtungen unterstützt werden. Ähnlich zu Interesse ist das Fähigkeitsselbstkonzept als Einschätzung der Schülerinnen und Schüler ihrer eigenen Leistung sowohl für Wahlentscheidungen als auch lebenslange Lernprozesse in Naturwissenschaften ein zentraler Prädiktor und steht in wechselseitigem Zusammenhang mit Leistung und Interesse (z. B. Guay, Larose & Boivin, 2004; Marsh & Craven, 2006; Marsh, Trautwein, Ludtke, Koller & Baumert, 2005; Marsh & Yeung, 1997a, 1997b; Parker et al., 2012).

Weil leistungsstarke Jugendliche in Naturwissenschaften aus leistungsbezogener Sicht bereits gute kognitive Voraussetzungen für naturwissenschaftsbezogene Bildungs- und Karrierewege mitbringen, hat die Entwicklung eines starken persönlichen Interesses

für Naturwissenschaften sowie eines optimistisch-realistischen Selbstkonzepts einen hohen Stellenwert und kann das „Zünglein an der Waage“ für Studien und Berufsausbildungsentscheidungen sein. Andererseits sind auch High und Nicht-High Performer, die keine naturwissenschaftliche Ausbildungsrichtung einschlagen, in ihrem weiteren Lebensweg gefordert, sich mit naturwissenschaftlichen Themen auseinanderzusetzen, beispielsweise im Umgang mit gesellschaftlich relevanten naturwissenschaftlichen Fragen und Konflikten, zu denen evidenzbasiert Meinungen und Entscheidungen gefunden werden müssen. Die damit verbundenen lebenslangen Lernprozesse setzen die Bereitschaft, sich mit naturwissenschaftlichen Themen auseinanderzusetzen, sowie Zutrauen in die eigenen Fähigkeiten voraus. Darüber hinaus legt eine solche in die Breite geführte „Interessiertheit“ (Prenzel & Schütte, 2007, S. 107) unter leistungsstarken Jugendlichen die Grundlage für die Spitzenförderung und einen großen Talentpool für naturwissenschaftliche Nachwuchskräfte (Krapp & Prenzel, 2011; Prenzel & Schütte, 2007; Prenzel et al., 2009).

Für eine zielorientierte Nachwuchsförderung ist daher neben der Leistungsförderung vor allem die Entwicklung von Interesse und Selbstkonzept ein wichtiges Bildungsziel und eine zentrale Frage bei der Charakterisierung der aktuellen Leistungsspitze in Naturwissenschaften ist: Inwiefern werden die leistungsstarken Schülerinnen und Schüler motivational-affektiv erreicht und welche Unterschiede sind innerhalb der leistungsstarken Jugendlichen in Naturwissenschaften zu finden? Zur Beantwortung dieser Frage wird nachfolgend der Begriff und die Entwicklung des Interesses eingebettet in die Person-Gegenstand-Theorie dargestellt sowie der Forschungsstand im Kontext leistungsstarker Jugendlicher in Naturwissenschaften erläutert. Als weiteres motivational-affektives Bildungsziel wird darüber hinaus ein realistisch-optimistisches Fähigkeitsselbstkonzept und sein Zusammenhang mit Leistung sowie Interesse in Naturwissenschaften diskutiert. Zudem werden Vergleichsprozesse, die im Zusammenhang von Leistung und Selbstkonzept sowie Leistung und Interesse stattfinden, berichtet und schließlich personenzentrierte Profilanalysen als aktuelle Forschungsansätze zur Berücksichtigung der Interaktion von Leistung, Selbstkonzept und Interesse beleuchtet.

3.1. Interesse an den Naturwissenschaften

Historisch gesehen war Interesse bereits unter frühen Arbeiten wie denen von Herbart (1806), Dewey (1913), Kerschensteiner (1922) oder Piaget (1974, org. Piaget, 1939) Teil psychologischer und pädagogisch-psychologischer Theorien (für eine detaillierte Übersicht vgl. Prenzel, 1988). Eine enge Verbindung und Einbettung in die pädagogische und unterrichtspsychologische Forschung erfolgte vor allem ab den 1970er Jahren (z. B. Todt, 1974; H. Schiefele, 1978). Unter den, ab diesem Zeitraum entwickelten, neueren Interessentheorien hat sich besonders die pädagogischen Interessentheorie (Prenzel, 1988, 1992; H. Schiefele, Krapp, Prenzel, Heiland & Kasten, 1983) bzw. Person-Gegenstand-Theorie von Interesse (Krapp & Prenzel, 1992; Krapp, 2002a, 2002b) etabliert. Zum Teil in Anlehnung an diese Konzeption von Interesse entstanden weitere Theorien bzw. Erweiterungen (z. B. Häussler, Hoffmann, Langeheine, Rost & Sievers, 1998; Todt & Schreiber, 1998) und zahlreiche empirische Untersuchungen (für eine Übersicht vgl. Krapp & Prenzel, 2011). Nachfolgend wird die Person-Gegenstand-Theorie als theoretisches Rahmenmodell für das Interesse von High Performern in Naturwissenschaften beschrieben und daraufhin der aktuelle Forschungsstand zum Interesse leistungsstarker Schülerinnen und Schüler und Geschlechterunterschieden dargestellt.

3.1.1. Die Person-Gegenstand-Theorie als theoretischer Rahmen

In der Person-Gegenstand-Theorie wird Interesse als die Beziehung zwischen einer Person und einem Objekt verstanden (Krapp, 2002a; Krapp & Prenzel, 1992; H. Schiefele et al., 1983). Mögliche Objekte sind dabei Inhalte, Gegenstände, Wissensgebiete oder Tätigkeiten. Aufgrund des Gegenstandsbezugs stellt Interesse ein domänenspezifisches bzw. inhaltspezifisches Konstrukt dar und kann durch verschiedene Inhaltsstrukturen beschrieben werden (vgl. Krapp & Prenzel, 2011). Interesse in Physik wird von Häussler (1987) sowie Häussler und Hoffmann (2000) beispielsweise durch eine Inhaltsstruktur mit den drei Dimensionen Thema, Kontext und Aktivität charakterisiert.

Die Beziehung zwischen Person und Objekt beinhaltet sowohl kognitive als auch affektive und wertbezogene Komponenten (Krapp, 2002a; Krapp & Prenzel, 2011; Prenzel, Krapp & Schiefele, 1986; U. Schiefele, Krapp, Wild & Winteler, 1993). Im Kontext von Lernprozessen ist insbesondere die epistemische Valenz als Teil der kognitiven Komponente zentral. Sie beschreibt die Bereitschaft dafür, mehr über einen Gegenstand

zu erfahren (Prenzel, 1988). Die affektive Komponente zeigt sich in positiven Emotionen wie Freude oder Neugier in der Beschäftigung bzw. gegenüber des Objekts, während sich die Wertkomponente durch die Bedeutung und Wichtigkeit, die dem Gegenstand zugeschrieben wird, auszeichnet. Interesse kann daher ein Auslöser positiver Emotionen wie Freude sein, gleichzeitig ist sie jedoch nur ein möglicher Grund für Freude und als Konstrukt theoretisch davon zu trennen (z. B. Krapp & Prenzel, 2011; U. Schiefele, 1991; Silvia, 2006). Betrachtet man die Valenzen im Kontext hoher Kompetenz in Naturwissenschaften, so sind alle drei für leistungsstarke Schülerinnen und Schüler anzustreben. Da weiterführende Lernprozesse aber vor allem durch die epistemische Valenz begünstigt werden, ist für leistungsstarke Jugendliche insbesondere die epistemische Valenz als Ziel motivational-affektiver Förderung hervorzuheben.

Die Person-Gegenstand-Theorie von Interesse zeichnet sich gemeinsam mit anderen Ansätzen in der pädagogisch-psychologischen Interessenforschung (z. B. Hidi & Renninger, 2006) dadurch aus, dass sie zwischen individuellem und situationalem Interesse unterscheidet, wobei die drei Valenzen beiden Interessenformen zugeschrieben werden (Krapp, 2002b; Krapp, Hidi & Renninger, 1992; Krapp & Prenzel, 1992; U. Schiefele & Wild, 2000). Als individuelles Interesse wird eine zeitlich stabile positive Disposition gegenüber einem Gegenstand verstanden. Situationales Interesse dagegen beschreibt ein situationsspezifisches Interesse, das aus einer Person-Umwelt-Interaktion entsteht und sich als momentaner motivationaler Zustand und aktuelles Engagement während einer Tätigkeit äußert (Krapp & Prenzel, 2011).

Interessenbasiertes Lernen erfolgt intrinsisch und steht in enger Verbindung zu intrinsischer Lernmotivation (Krapp & Prenzel, 2011; Ryan & Deci, 2000a). Eine interessierte Tätigkeit geht dabei über die Definition einer intrinsisch motivierten Tätigkeit im Sinne der Selbstbestimmungstheorie (Deci, 1998; Deci & Ryan, 1993) hinaus, indem sie eine wertbezogene Komponente und damit die Bedeutung und Sinnhaftigkeit des Gegenstandes explizit mit einschließt. In der Selbstbestimmungstheorie gilt die Bedeutung dagegen nicht als Bedingung für intrinsisches Lernen, sondern wird im Integrationsprozess, der wiederum von der Erfüllung motivationsunterstützender Bedingungen abhängt, generiert. Das heißt, sind die Bedürfnisse nach Eingebundenheit, Kompetenz und Autonomie erfüllt, kann eine Person einen Gegenstand in ihr Selbst integrieren und erlebt ihn automatisch als bedeutungsvoll (Deci & Ryan, 1993).

Interesse für Naturwissenschaften entsteht grundsätzlich aus unterschiedlichen

Interaktionen zwischen Personen und Interessensobjekten in sozialen und institutionellen Umgebungen (Hidi & Renninger, 2006; Krapp & Prenzel, 2011; Silvia, 2006; Prenzel, 1992). Diese Entwicklung von Interesse lässt sich durch verschiedene Ansätze theoretisch einbetten (Daniels, 2008; Krapp & Prenzel, 2011). Neben biographischen Ansätzen (z. B. Todt & Schreiber, 1998; Travers, 1978) wird der Entwicklungsprozess von Interesse durch eine Überführung von situationalem zu individuellem Interesse, die in Abhängigkeit der Person-Umwelt-Interaktion erfolgt, erklärt (Hidi & Harackiewicz, 2000; Krapp, 2002b). Krapp (2002b) unterscheidet auf diesem Entwicklungskontinuum drei verschiedene Formen von Interesse bzw. zwei Entwicklungsschritte. Von außen induziertes situationales Interesse, beispielsweise in Form von Neugierde, kann sich in einem ersten Schritt in stabilisiertes, d. h. über Lernphasen andauerndes situationales Interesse entwickeln. In einem weiteren Schritt kann aus dem andauernden situationalen Interesse schließlich eine relativ stabile Disposition und damit individuelles Interesse entstehen. Damit greift Krapp (2002b) die von Mitchell (1993) und Hidi und Harackiewicz (2000) vorgeschlagene Unterscheidung zwischen „catch“ und „hold“ Aspekten situationalen Interesses auf. Ähnlich dazu unterscheiden Hidi und Renninger (2006) zwischen vier Phasen der Entwicklung von Interesse und unterteilen individuelles Interesse in aufkommendes und stabilisiertes individuelles Interesse.

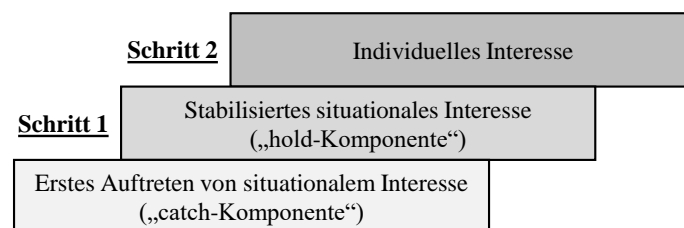


Abbildung 1. Entwicklungsschritte von Interesse (Krapp, 2002b, S. 399, übers. v. Verf.)

3.1.2. Interessenentwicklung und Ausdifferenzierung im Jugendalter

Die ontogenetische Entwicklung von Interesse war in den letzten Jahrzehnten insbesondere im Bereich mathematisch-naturwissenschaftlicher Themen und Schulfächer Gegenstand vieler Forschungsbemühungen (Krapp, 2002a). Zahlreiche Studien belegen, dass zwar am Ende der Grundschule noch hohes Interesse gegenüber Naturwissenschaften besteht (Krapp & Prenzel, 2011; Osborne, Simon & Collins, 2003), das Interesse für

Chemie, Physik und zum Teil auch Biologie jedoch während der Sekundarstufe abnimmt (Baumert & Köller, 1998a; Daniels, 2008; Gardner, 1998; Gottfried, Fleming & Gottfried, 2001; Krapp & Prenzel, 2011; Osborne et al., 2003). Dieser Abfall von Interesse ist bei Mädchen im Mittel stärker ausgeprägt als bei Jungen (Häussler, Hoffmann et al., 1998; Hoffmann, 2002).

Differenziertere Ergebnisse und zum Teil gegenläufige Entwicklungen ergeben sich bei Betrachtung von Studien, die Einflussfaktoren wie Fachinhalte bzw. Kontexte, Schulart und Geschlecht bzw. Interaktionen dieser Merkmale mit untersuchen (vgl. auch Krapp, 2002a; Krapp & Prenzel, 2011). So zeigt sich in Bezug auf Fachinhalte beispielsweise, dass sich das Interesse von Mädchen und Jungen in Naturwissenschaften in der Sekundarstufe je nach Kontext unterscheidet und die Geschlechter unterschiedliche Präferenzen zeigen (Hoffmann, 2002; Hoffmann, Krapp, Renninger & Baumert, 1998; J. Rost, Sievers, Häußler, Hoffmann & Langeheine, 1999). Werden physikalische Inhalte demnach in humanbiologischen oder gesellschaftlichen Kontexten präsentiert, zeigen Mädchen ein höheres Interesse als bei der Präsentation in technischen Kontexten. Jungen dagegen zeigen hohes Interesse für technische Kontexte, interessieren sich jedoch auch für soziale und humanbiologische Kontexte (Häussler, 1987; Häussler & Hoffmann, 2000; Hoffmann, Krapp et al., 1998). Ein Vergleich mit typischen Curricula in Physik zeigt, dass die verwendeten Kontexte bisher eher quantitative und technische Bezüge aufweisen und dadurch Jungen näher sind (Hoffmann, 2002). Ähnlich zeigt sich in Biologie, dass das Interesse für Zoologie und Pflanzen bei Mädchen sinkt, gleichzeitig das Interesse an Humanbiologie und Ökologie steigt (Todt, 1974).

Zur Erklärung des Interessenabfalls in Naturwissenschaften während der Sekundarstufe existieren verschiedene Ansätze. Daniels (2008) und Krapp und Prenzel (2011) geben einen Überblick und nennen unter anderem drei große Einflussfaktoren. Zum einen wird angenommen, dass die Passung der Lernumgebung zu den Bedürfnissen und Anlagen des Schülers, das Curriculum, Eigenschaften der Schule und andere Einflussfaktoren auf die Lernumgebung die Interessensentwicklung beeinflussen (z. B. Eccles, 1994; Hoffmann, Krapp et al., 1998; Rheinberg & Vollmeyer, 2000; Snow, Corno & Belmont, 1996; Turner et al., 1998). Instruktionale Merkmale des Unterrichts, insbesondere motivationsunterstützende Lehr-Lernbedingungen, beeinflussen daher die Entwicklung von Interesse und Lernmotivation (vgl. Prenzel, 1997) (vgl. auch Kapitel 3.4).

Zum anderen wird der Interessenabfall damit erklärt, dass Jugendliche aus ent-

wicklungspsychologischer Sicht mit für die Sekundarstufe typischen Entwicklungsaufgaben beschäftigt sind und immer weniger bereit sind, Ressourcen für akademische Lernprozesse aufzuwenden (Anderman & Maehr, 1994; Hofer, 2010; Wigfield et al., 1997). Soziale Beziehungen zu Gleichaltrigen werden wichtiger (z. B. Tajfel, 1982). Jugendliche weisen der Interaktion mit Gleichaltrigen und Freunden und den damit verbundenen Themen und Rollen deshalb möglicherweise mehr Bedeutung zu als der Interaktion mit naturwissenschaftlichen Themen als Interessensgegenstand.

Schließlich kann die Abnahme des Interesses an Naturwissenschaften auch durch die Ausdifferenzierung von Interessen erklärt werden (Daniels, 2008; Todt, Drewes & Heils, 1994; Travers, 1978). Im Jugendalter entwickeln Schülerinnen und Schüler zunehmend ihre eigene Identität und reflektieren Stärken und Schwächen sowie Vorlieben und Abneigungen. Interessen, die mit dem Selbstbild nicht vereinbar sind, verlieren im Vergleich zu zur Identität konformen Interessen an Wert und werden zunehmend abgelegt. Auch Geschlechtsrollenvorstellungen nehmen in der Adoleszenz zunehmend Einfluss auf die Identitätsbildung (z. B. Hannover, 1997, 1998) und können bei geschlechtsstereotypen Vorstellungen von Schulfächern zu niedrigerem Interesse (Hannover & Kessels, 2004; Kessels & Hannover, 2004). Nachdem Schülerinnen und Schüler am Ende der Grundschulzeit relativ breites und hohes Interesse aufweisen (Krapp & Prenzel, 2011; Todt & Schreiber, 1998) resultiert eine Ausdifferenzierung der Interessen auf Populationsebene betrachtet automatisch zu einer Reduzierung des Interesses für Naturwissenschaften mit zunehmendem Alter in der Sekundarstufe. Im Hinblick auf High Performer in Naturwissenschaften stellt sich daher die Frage, inwieweit sie ihre Stärke in Naturwissenschaften wahrnehmen und ein Selbstbild entwickeln, das eine hohe Fähigkeit und ein hohes Interesse in Naturwissenschaften miteinschließt.

3.1.3. Interesse leistungsstarker Jugendlicher in Naturwissenschaften

Ausgehend vom Zusammenhang von Leistung und Interesse (z. B. Krapp & Prenzel, 2011), kann von leistungsstarken Schülerinnen und Schülern ein höheres Interesse als von leistungsschwächeren Jugendlichen erwartet werden. Der Zusammenhang von Leistung und Interesse wurde in den letzten Jahrzehnten vielfach untersucht und in verschiedenen Übersichtsarbeiten zusammengefasst (Krapp et al., 1992; U. Schiefele, Krapp &

Winteler, 1992; U. Schiefele, 1998). Viele der Studien sind querschnittlich angelegt und zeigen einen mittleren Zusammenhang zwischen Interesse und Leistung von $r = .31$ (U. Schiefele et al., 1992, S. 198), wobei sich mit $r = .35$ (U. Schiefele et al., 1992, S. 200) und $r = .38$ (Hoffmann, Lehrke & Häussler, 1998, S. 113) für Physik wesentlich höhere Korrelationen zeigen als in Biologie ($r = .16$, U. Schiefele et al., 1992, S. 200). Gleichzeitig existieren jedoch auch Studien in Mathematik, die keinen Zusammenhang von Leistung und Interesse finden (z. B. Köller, Baumert & Schnabel, 2001). In vielen Studien beeinflussen darüber hinaus Geschlecht und Schwierigkeit des Faches diesen Zusammenhang (Baumert & Köller, 1998a; U. Schiefele et al., 1992). Im Hinblick auf Geschlechterunterschiede spielt jedoch auch die Art, wie Interesse gemessen wird, eine Rolle. Bei der dabei stattfindenden detaillierten Differenzierung nach unterschiedlichen Themen und Anwendungsbereichen findet sich bei Hoffmann, Lehrke und Häussler (1998) beispielsweise kein Unterschied zwischen Mädchen und Jungen (Hoffmann, Lehrke & Häussler, 1998).

Neben den zahlreichen querschnittlichen Studien existieren auch einige längsschnittliche Untersuchungen für Mathematik (Baumert, Schnabel & Lehrke, 1998; Köller et al., 2001; Köller, Trautwein, Lüdtke & Baumert, 2006; Lopez, Lent, Brown & Gore, 1997; Prenzel et al., 2006) und Physik (Hoffmann, Lehrke & Häussler, 1998). Sie belegen reziproke Effekte und zeigen, dass positive Lernerfahrungen die Entwicklung von Interesse unterstützen und Interesse wiederum Engagement in Lernprozessen und die Nutzung tieferer Lernstrategien fördert und damit Leistung positiv bedingt (Baumert et al., 1998; Köller et al., 2006). Baumert et al. (1998) und Köller et al. (2006) finden dabei einen stärkeren Effekt von Leistung auf Interesse als umgekehrt. Köller et al. (2001) konnten jedoch auch zeigen, dass das Ergebnis von der Art der Leistungsmessung abhängig ist. Während am Anfang des Schuljahres bestehendes Interesse spätere Schulnoten beeinflusst, zeigt es keinen Zusammenhang zur Leistung in einem standardisierten Test (Köller et al., 2001). Eine weitere Erklärung dafür, dass der Einfluss von Interesse auf Leistung schwächer als der umgekehrte Effekt ausgeprägt ist, ist, dass ein solcher Einfluss vor allem in schüleraktiven und selbstregulativen Unterrichtsformen wirksam wird (Köller et al., 2001). Diese sind im Naturwissenschaftsunterricht in Deutschland jedoch eher schwach ausgeprägt (vgl. Seidel et al., 2006).

Basierend auf dem reziproken Zusammenhang von Leistung und Interesse wird von leistungsstarken Schülerinnen und Schülern in Naturwissenschaften ein höheres Interesse in Naturwissenschaften vermutet. Viele der bisherigen Arbeiten, die sich mit der

Subgruppe der leistungsstarken Jugendlichen in Naturwissenschaften oder der vielseitig leistungsstarken Jugendlichen beschäftigen, konzentrieren sich auf die Beschreibung ihrer kognitiven Fähigkeiten (Wendt et al., 2013; Zimmer et al., 2007) und nur wenige schließen Motivation oder Interesse mit ein (z. B. OECD, 2009b; Prenzel & Schütte, 2007).

Wichtige Erkenntnisse für leistungsstarke Schülerinnen und Schüler in Naturwissenschaften liefern Large-Scale Assessments. Top Performer, d. h. Schülerinnen und Schüler auf den beiden obersten Kompetenzstufen in Naturwissenschaften, weisen international ein höheres Interesse auf als schlechtere Schülerinnen und Schüler (OECD, 2009b). Ähnlich sind Schülerinnen und Schüler im obersten Quartil in PISA 2006 (OECD, 2007) in Relation zur Gesamtpopulation interessierter und knapp 60% befinden sich in den beiden obersten Interessenquartilen (Prenzel & Schütte, 2007, S. 114).

In der Hochbegabungsforschung werden motivational-affektive Merkmale zwar in einzelnen Konzepten von Hochbegabung berücksichtigt (vgl. Ziegler, 2000) aber nur wenige Studien untersuchen allgemeine (z. B. Ziegler, Heller & Broome, 1996) oder naturwissenschaftsbezogene (z. B. Andersen & Chen, 2016; C. Anderson, Kim J. & Keller, 2014; Ziegler & Heller, 2000c) motivational-affektive Merkmale Hochbegabter bzw. Hochleistender im Vergleich zu anderen Schülern. Hochbegabte bzw. Hochleistende zeigen sich darin allgemein sowie in Physik und Chemie motivierter als weniger begabte.

Neben den Ergebnissen zum Interessenniveau leistungsstarker Schülerinnen und Schüler beschäftigen sich einige Studien damit, welcher Anteil leistungsstarker Jugendlicher ein vergleichsweise niedriges Interesse aufweist. Auch wenn Leistung und Interesse im Mittel (schwach) korrelieren, ist es möglich, dass leistungsstarke Jugendliche schwach oder nicht interessiert sind. Zwei im Rahmen von PISA 2006 durchgeführte Untersuchungen auf Basis kontextbasierter Interessensskalen (Prenzel & Schütte, 2007, 2008) liefern differenziertere Einblicke in die Interessensverteilung von Jugendlichen im oberen Quartil des PISA Tests in Naturwissenschaften. Sowohl in Deutschland (43%) als auch in Ländern wie Finnland (38%) gehören mehr als ein Drittel der leistungsstarken Jugendlichen den beiden unteren Interessenquartilen an und zeigen damit in Relation zur Gesamtpopulation ein vergleichsweise niedriges Interesse (Prenzel & Schütte, 2007, S. 115). Auf einem kriterienbasierten statt relativen Maßstab bedeutet das, dass die Schülerinnen und Schüler in alltäglichen Situationen, beispielsweise beim Nachrichten lesen, kein oder nur wenig Interesse dafür haben, etwas Neues über ein naturwissenschaftliches Thema zu erfahren (Prenzel & Schütte, 2007). Auch Ergebnisse zur zukunftsorientierten

Motivation, d. h. der Bereitschaft, sich auch zukünftig mit Naturwissenschaften zu beschäftigen, bestätigen diesen Befund (OECD, 2009b; Taskinen, 2010). Über ein Drittel der Top Performer, d. h. Schülerinnen und Schüler auf den beiden obersten Kompetenzstufen in Naturwissenschaften, möchte sich in Zukunft weniger gern mit Naturwissenschaften beschäftigen als etwas leistungsschwächere Jugendliche auf der Kompetenzstufe darunter (Kompetenzstufe IV) (OECD, 2009b). Damit weist Deutschland im internationalen Vergleich der Top Performer der OECD den zweithöchsten Anteil an in Relation zur Leistung wenig zukunftsmotivierten Top Performern auf (OECD, 2009b). Diese als unmotivierte Top Performer bezeichneten Jugendlichen unterscheiden sich von den motivierten Top Performern insbesondere signifikant im Interesse für Naturwissenschaften (OECD, 2009b).

Diese Ergebnisse beruhen einerseits auf einer kontextbasierten und auf die kognitiv-epistemologische Komponente fokussierten Operationalisierung von Interesse (Prenzel & Schütte, 2007) und andererseits auf zukunftsorientierter Motivation (OECD, 2009b). Offen bleibt daher, welcher Anteil leistungsstarker Schülerinnen und Schüler in Large-Scale-Assessments sich basierend auf einer Interessenskala, die ohne Kontextbasierung allgemein das Interesse für Naturwissenschaften abfragt, wenig für die Naturwissenschaften interessiert. So betonen Zimmer et al. (2007), dass neben kognitiven auch motivationale Merkmale hochleistender Jugendlicher näher betrachtet werden sollten und insbesondere Geschlechterunterschiede und der Vergleich von domänenspezifisch sowie vielseitig Hochkompetenten in den Blick genommen werden soll.

Im Rahmen der Betrachtung des Interesses von leistungsstarken Jugendlichen werden auch Merkmale wie Geschlecht, soziale Lage, Schulart und Unterricht diskutiert (Prenzel & Schütte, 2007). Ähnlich zum allgemeinen Zusammenhang von Interesse und Leistung können auch bei High Performern Geschlechterunterschiede vermutet werden. Sowohl im Vergleich themenübergreifend in Naturwissenschaften zukunftsmotivierter und unmotivierter Top Performer der OECD (OECD, 2009b) als auch in der Betrachtung der Quartile des kontextbezogenen Interesses der besten 25% in Naturwissenschaften in Deutschland finden sich jedoch keinen nennenswerten Geschlechterdifferenzen (Prenzel & Schütte, 2007).

Neben dem Geschlecht beeinflusst möglicherweise der sozioökonomische Hintergrund das Interesse, wenn Schülerinnen und Schüler aus höheren sozialen Lagen bessere oder geeignetere materielle Ressourcen oder aufgrund der Berufszugehörigkeit

und des Bildungsstands ihrer Eltern mehr Zugänge zu naturwissenschaftlicher Förderung besitzen. Umgekehrt ist jedoch auch denkbar, dass Schülerinnen und Schüler aus niedrigeren sozialen Lagen und ländlichen Umgebungen beispielsweise wesentlich häufiger Berührungspunkte mit der Natur und naturwissenschaftlichen Phänomenen haben. Gerade bei Jugendlichen aus niedrigeren sozialen Schichten wecken Naturwissenschaften und Technik im Vergleich zu sprachlastigeren Themen möglicherweise mehr Interesse (vgl. Prenzel & Schütte, 2007). In ihrer Studie zu Prädiktoren einer hohen zukunftsorientierten Motivation und der Wahl eines naturwissenschaftlichen Berufs fanden Kjærnsli und Lie (2011) überraschenderweise, dass ein naturwissenschaftlicher Beruf von einem oder beiden Elternteilen im Vergleich zu einem anderen Beruf nur sehr wenig Varianz der zukunftsorientierten Motivation aufklärte. Leistungsstarke Jugendliche unterschiedlicher Motivation bzw. unterschiedlichen Interesses unterschieden sich sowohl international in der Untersuchung der OECD (OECD, 2009b) als auch national in der Analyse von Prenzel und Schütte (2007) nicht in ihrem sozioökonomischen Hintergrund. Bisher wurde der sozioökonomische Hintergrund in den genannten Studien mit Hilfe des höchsten *International Socio-Economic Index of Occupational Status (ISEI)* (Ganzeboom et al., 1992; Ganzeboom & Treiman, 1996) operationalisiert. Ähnlich zur Analyse von Kjærnsli und Lie (2011) könnte eine kategorielle Einteilung der Berufe – beispielsweise in Form *Erikson-Goldthorpe-Portocarero-Klassifikation (EGP)* (Erikson et al., 1979; Erikson & Goldthorpe, 2002) – weitere Aufschlüsse geben.

Zusammengefasst scheinen leistungsstarke Jugendliche in Naturwissenschaften sich bisherigen Studien zufolge für Naturwissenschaften im Mittel in höherem Maße interessieren als leistungsschwächere Schüler, gleichzeitig existiert jedoch ein beträchtlicher Anteil an leistungsstarken Jugendlichen mit relativ geringem Interesse. Bisherige Untersuchungen finden dabei keine Geschlechterunterschiede und Zusammenhänge zu sozialer Herkunft. Bisher existieren jedoch nur einzelne Studien zum Interesse leistungsstarker Jugendlicher in Naturwissenschaften, die jeweils unterschiedliche Definitionen für High Performance und unterschiedliche motivational-affektive Konstrukte betrachten. Insgesamt ist deshalb noch relativ wenig über das naturwissenschaftliche Interesse von High Performern in Naturwissenschaften und Einflussfaktoren wie Geschlecht und soziale Herkunft bekannt. Insbesondere über den repräsentativen Anteil von High Performern, die sich in Relation zu ihrer Leistung wenig für Naturwissenschaften interessieren, existieren nur einzelne Ergebnisse.

3.2. Fähigkeitsselbstkonzept in Naturwissenschaften

Wenn Schülerinnen und Schüler überlegen, welchen Bildungsweg oder welchen Beruf sie wählen sollen, trägt ihre eigene Einschätzung über ihre Fähigkeiten entscheidend zu ihrer Wahlentscheidung bei (z. B. Wigfield & Eccles, 2000). Das Konstrukt des domänenspezifischen akademischen Selbstkonzepts beschreibt dieses Zutrauen in die eigene Fähigkeit (Marsh, Byrne & Shavelson, 1988) und kann als ein wichtiges Ziel von Unterricht und Bildung verstanden werden (Marsh & Hau, 2003; Prenzel & Schütte, 2007). Neben seinem Beitrag zur Kurs- und Studienwahl (Köller, Daniels, Schnabel & Baumert, 2000; Marsh & Yeung, 1997b; Taskinen et al., 2013), steht das Selbstkonzept auch in positivem Zusammenhang zu engagierten Lernprozessen (Helmke & van Aken, 1995; Trautwein, Lüdtke, Köller & Baumert, 2006) und akademischer Leistung (Marsh, Liem, Martin, Morin & Nagengast, 2011). Es ist darüber hinaus eng mit Interesse assoziiert und entwickelt sich biographisch gemeinsam mit Leistung und Interesse (Marsh et al., 2005). Im Zusammenhang mit High Performance in Naturwissenschaften hat das naturwissenschaftsbezogene Selbstkonzept eine hohe funktionelle Bedeutung in mehrfacher Hinsicht: Ein optimistisch-realistisches Selbstkonzept ist einerseits ein eigenständiges Ziel von Unterricht und andererseits Prädiktor von Interesse und anderen Lernzielen (vgl. auch Prenzel & Schütte, 2007; Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006).

Nachfolgend wird der Zusammenhang von Leistung und Selbstkonzept sowie der Forschungsstand zum Selbstkonzept leistungsstarker Schülerinnen und Schüler in Naturwissenschaften beschrieben. Als theoretisches Modell dient dabei das *Internal/External Frame of Reference Model* (Marsh, 1986). Die darin postulierten sozialen und dimensional Vergleichenprozesse in der Genese von Selbstkonzept aus Leistung werden in einem nachfolgenden Abschnitt beschrieben und die Rolle von Schulnoten erläutert. Im letzten Teil wird schließlich die wechselseitige Beziehung von Selbstkonzept und Interesse sowie die eingeführten Vergleichsprozesse auf Interesse übertragen.

3.2.1. Fähigkeitsselbstkonzept leistungsstarker Schülerinnen und Schüler in Naturwissenschaften

Der Zusammenhang von Leistung und Selbstkonzept wurde in den letzten Jahrzehnten vielfach untersucht. Sowohl Metaanalysen (z. B. Huang, 2011) als auch Längsschnittstudien (Marsh & O'Mara, 2008) konnten zeigen, dass beide Merkmale reziprok aufeinander

wirken. Studien aus verschiedenen Altersbereichen und Fächern finden demnach einen Einfluss des Selbstkonzepts auf die Leistungsentwicklung („*self-enhancement*“), der unter Kontrolle verschiedener Faktoren wie Vorleistung, Intelligenz und Interesse beobachtet wird (z. B. Köller et al., 2006; Marsh, 2005; Steinmayr & Spinath, 2009). Umgekehrt beeinflusst die Kompetenzentwicklung die Entwicklung des Selbstkonzepts („*skill development*“) (z. B. Köller et al., 2001; Marsh et al., 2011).

Während für Mathematik und Sprachen zahlreiche Studien zum Zusammenhang von Leistung und Selbstkonzept existieren, finden sich für Naturwissenschaften deutlich weniger Forschungsergebnisse (Jansen, Schroeders & Lüdtke, 2014). Dennoch kann auch in Naturwissenschaften ein Zusammenhang von sowohl domänenübergreifend erfasstem naturwissenschaftsbezogenem Selbstkonzept (J. Wang, Oliver & Staver, 2008) als auch Selbstkonzept in Biologie (Britner, 2008), Chemie (Nieswandt, 2007) und Physik (Möller, Streblov, Pohlmann & Köller, 2006) mit Leistung nachgewiesen werden. Die Zusammenhänge sind dabei stärker ausgeprägt, wenn zwischen den einzelnen naturwissenschaftlichen Fächern unterschieden wird und Leistung in Form von Schulnoten erfasst wird (Jansen et al., 2014). Im Vergleich zu anderen Domänen ist der Zusammenhang von Leistung und Selbstkonzept für Naturwissenschaften höher als für verbale Domänen und niedriger als für Mathematik (Marsh & Yeung, 1997a).

Ergebnisse über das Selbstkonzept leistungsstarker Schülerinnen und Schüler finden sich im Bereich der Large-Scale-Assessments sowie in der Hochbegabungsforschung. Dabei beschäftigen sich jedoch nur wenige Studien mit domänenspezifisch definierten leistungsstarken Jugendlichen in Naturwissenschaften. Eine dieser Ausnahmen bildet eine internationale Betrachtung der Jugendlichen auf den beiden obersten Kompetenzstufen in PISA 2006 (OECD, 2009b), in der das naturwissenschaftsbezogene Selbstkonzept von speziell in Naturwissenschaften leistungsstarken Schülerinnen und Schülern im OECD-Vergleich betrachtet wird. Dabei zeigen 80% (OECD, 2009b, S. 65) der Top Performer (Schülerinnen und Schüler auf Kompetenzstufe V und VI) im OECD-Durchschnitt ein hohes Selbstkonzept. Im Mittel übersteigt das Selbstkonzept der Top Performer zudem das Selbstkonzept einer nur wenig leistungsschwächeren Gruppe der *Strong Performer* (Schülerinnen und Schüler auf Kompetenzstufe IV). Weitere Studien thematisieren zwar das naturwissenschaftsbezogene Selbstkonzept leistungsstarker Jugendlicher in Naturwissenschaften in Large-Scale-Assessments, betrachten jedoch keinen Vergleich zu weniger leistungsstarken Schülerinnen und Schülern, sondern nutzen

das Selbstkonzept vor allem als Prädiktor für die Berufserwartung (z. B. Buccheri et al., 2011; Taskinen, 2010).

Auch die Hochbegabungsforschung liefert Ergebnisse zum Zusammenhang von Selbstkonzept und Hochbegabung bzw. hoher Leistung, die die genannten Erkenntnisse unterstützen. Demnach entwickeln begabte bzw. leistungsstarke Jugendliche mit höherer Wahrscheinlichkeit ein positives Selbstkonzept als weniger leistungsstarke Schülerinnen und Schüler (Hoge & Renzulli, 1993; Janos & Robinson, 1985; D. H. Rost, 2009b). Innerhalb leistungsstarker Schülerinnen und Schüler wird sowohl ein Zusammenhang von höherer Leistung mit höherem Selbstkonzept gefunden (Williams & Montgomery, 1995) als auch in der einzigen in diesem Zusammenhang auf Naturwissenschaften konzentrierte Studie kein Zusammenhang mit höherer Begabung gefunden (Ziegler & Heller, 2000a). Insgesamt gibt es keine Hinweise darauf, dass das Selbstkonzept bei begabten Schülerinnen und Schülern eine andere Struktur oder andere Entwicklungsverläufe aufweist als bei anderen Schülerinnen und Schülern (McCoach & Siegle, 2003; Meier, Vogl & Preckel, 2014; Plucker & Stocking, 2001; Reis & Park, 2001). Die meisten dieser Studien betrachten jedoch das allgemeine akademische Selbstkonzept und nutzen fachunspezifische Konzeptionen von Begabung bzw. eine Identifikation durch Intelligenzmaße (Dixon, 1998). Nur wenige Studien differenzieren zwischen mathematischem und verbalem Selbstkonzept (Plucker & Stocking, 2001) und bis auf wenige Ausnahmen (z. B. Ziegler & Heller, 2000a) werden naturwissenschaftsbezogene leistungsbezogene Kognitionen bisher nicht betrachtet.

Daneben lassen verschiedene Ergebnisse Geschlechterunterschiede im Zusammenhang von Leistung und Selbstkonzept bei Top und High Performern vermuten. Sowohl allgemeine (OECD, 2007; Schütte et al., 2007) als auch auf leistungsstarke Jugendliche in Naturwissenschaften fokussierende Studien (Buccheri et al., 2011; Reis & Park, 2001) finden für Mädchen niedrigere Selbstkonzeptwerte als für Jungen mit gleicher Leistung. So zeigen Studien von Reis und Park (2001) und Ziegler und Heller (2000b) für die Jungen unter den besten 10% der Schülerinnen und Schüler in Naturwissenschaften ein höheres allgemeines akademisches bzw. chemisches Selbstkonzept als für die Mädchen.

3.2.2. Soziale und dimensionale Vergleiche bei der Genese des Selbstkonzepts

Marsh (1986) beschreibt im *Internal/External Frame of Reference Model* für den Einfluss von Leistung auf Selbstkonzept zwei Arten von Vergleichsprozessen bzw. Bezugsrahmen: Soziale und dimensionale Bezugsrahmen. Schülerinnen und Schüler setzen dabei in sozialen Vergleichsprozessen ihre eigene Leistung in einem bestimmten Fach in Beziehung zur Leistung ihrer Bezugsgruppe, d. h. im Schulkontext zur Leistung ihrer Mitschüler in der Schulklasse. Dieser soziale Vergleich mit der Schulklasse als externen Bezugsrahmen bewirkt, dass Schülerinnen und Schüler sich bei gleicher Leistung in einer Klasse, die insgesamt sehr gute Leistungen aufweist, weniger kompetent fühlen als in einer Klasse mit niedrigerer mittlerer Leistung. Dieser Effekt wird auch *Big Fish Little Pond Effect (BFLPE)* (Marsh & Yeung, 1997a) genannt. Neben einem negativen Kontrasteffekt durch die Leistung der Bezugsgruppe kann auch ein positiver Assimilationseffekt durch die Zugehörigkeit zur Bezugsgruppe stattfinden (z. B. Marsh, Köller & Baumert, 2001). Köller et al. (2006) finden beispielsweise einen positiven Einfluss der Zugehörigkeit zu einem Leistungskurs in Mathematik auf das Selbstkonzept, während Trautwein, Lüdtke, Marsh et al. (2006) diesen Effekt unter Kontrolle der Noten nicht nachweisen können. Der BFLP-Effekt bildet aus statistischer Sicht den Nettoeffekt von Kontrast und Assimilation ab (Köller et al., 2006; Marsh, 1991). In einem zweiten Vergleichsprozess vergleichen Schülerinnen und Schüler ihre Leistung dimensional an einem internen Bezugsrahmen, indem sie die Leistung in einem Fach in Relation zu ihrer Leistung in anderen Fächern setzen. Daraus resultiert, dass eine höher Leistung in einem numerischen Kompetenzbereich, beispielsweise Mathematik, das Selbstkonzept in verbalen Domänen, beispielsweise Deutsch, negativ beeinflusst und umgekehrt (Möller et al., 2006). Einen detaillierteren Überblick über soziale und dimensionale Vergleichsprozesse inklusive einer Berücksichtigung temporaler Vergleiche geben Möller und Köller (2004) sowie Möller und Marsh (2013) gesondert zum Forschungsstand dimensionaler Vergleiche. Auch für leistungsstarke Schülerinnen und Schüler sind diese Vergleichsprozesse gültig (Plucker & Stocking, 2001).

Die Bedeutung von Schulnoten für die Vergleichsprozesse

Schülerinnen und Schüler erkennen in Vergleichsprozessen einerseits, in welchen Fächern sie Stärken und Schwächen haben und andererseits, ob sie in einem Fach fähiger sind als ihre Mitschüler. Diese Vergleiche können sich ganz grundsätzlich auf verschiedene Situationen und Handlungen wie Ideen, Beiträge und Erfolge in Gruppenarbeiten, Experimenten oder auch im Klassengespräch beziehen. Unter diesen Vergleichsgegenständen spielen Noten eine wichtige Rolle für Vergleichsprozesse im Zusammenhang von Leistung und Selbstkonzept (Trautwein, Lüdtke, Köller & Baumert, 2006). Noten sind in der Schule im Rahmen der Leistungsbeurteilung gebräuchlich und prägen, wie Schülerinnen und Schüler ihre Leistungen einschätzen und welche Wahlentscheidungen sie treffen (Prenzel et al., 2009). Viele Studien zeigen den wechselseitigen Zusammenhang von Noten und Selbstkonzept (Skaalvik & Skaalvik, 2002; Trautwein, Lüdtke, Marsh et al., 2006) und finden zwischen Noten und Selbstkonzept allgemein (Marsh et al., 2005) und in Naturwissenschaften (Jansen et al., 2014) einen höheren Zusammenhang als zwischen der Leistung in standardisierten Tests und Selbstkonzept. Weil Noten sehr regelmäßig und in Form von Ziffern gegeben werden, sind sie für Schülerinnen und Schüler eines der am leichtesten greifbaren und präsentesten Kommunikationsinstrumente von Leistungen und ermöglichen einen direkten „quantitativen“ Vergleich mit den Leistungen von Mitschülern (Trautwein, Lüdtke, Marsh et al., 2006). Gleichzeitig erfolgt die Benotung in der Schule nicht unabhängig vom Niveau der Klasse, sondern vielfach nach einem „grading-on-a-curve“ (Baumert et al., 2003; Ingenkamp, 1995; Ingenkamp & Lissmann, 2008), d. h. bester bzw. schlechtester Schüler bekommen tendenziell beste bzw. schlechteste mögliche Note und der Rest verteilt sich tendenziell normalverteilt dazwischen. Eine Leistungsgruppierung in unterschiedlichen Zweigen und Schularten kann deshalb zusätzlich zu unterschiedlichen Noten bei gleicher Leistung beitragen. Verschiedene Autoren vermuten deshalb, dass Vergleichsprozesse mediiert über Noten wirken und bestätigen in Studien, dass Noten allgemein und in Naturwissenschaften den Effekt von Leistung auf Selbstkonzept teilweise mediiieren (z. B. Trautwein, Lüdtke, Marsh et al., 2006).

3.2.3. Wechselseitige Beziehung mit Interesse

Ziel dieser Arbeit ist es, Selbstkonzept und Interesse als wichtige Ziele der Förderung von High Performern zu betrachten. Dazu wurde in den vorangegangenen Teilen zunächst der Zusammenhang von Interesse und Leistung bzw. High Performance in Naturwissenschaften und anschließend der Zusammenhang von Selbstkonzept und Leistung bzw. High Performance in Naturwissenschaften erläutert. Weite Teile der Forschung nehmen neben dem Zusammenhang mit Leistung einen reziproken Zusammenhang zwischen Selbstkonzept und Interesse und eine gemeinsame biographische Entwicklung an (für eine Übersicht vgl. Daniels, 2008; Marsh, 2005). Zusätzlich postulieren Baumert und Köller (1998a), dass das Selbstkonzept den Einfluss von Leistung auf Interesse mediert. Aus theoretischer Sicht ist dieser Zusammenhang zwischen Selbstkonzept und Interesse ebenso für leistungsstarke Jugendliche in Naturwissenschaften plausibel. Nachfolgend werden Forschungsergebnisse zum wechselseitigen Zusammenhang von Selbstkonzept und Interesse dargelegt sowie dimensionale und soziale Vergleiche auf die Interessenentwicklung übertragen.

Betrachtet man den Forschungsstand, zeigen zahlreiche querschnittliche und korrelative Studien einen engen Zusammenhang von Selbstkonzept und Interesse (z. B. U. Schiefele, 1996; U. Schiefele & Schreyer, 1992). Deutlich seltener sind jedoch längsschnittlich angelegte Studien, die reziproke Zusammenhänge zwischen Selbstkonzept und Interesse bzw. Leistung, Selbstkonzept und Interesse untersuchen und dabei zum Teil auch eine mögliche Mediationsrolle von Selbstkonzept analysieren (für einen Überblick vgl. Daniels, 2008; Marsh, 2005). Da aus der bisherigen Forschungslage nur wenig über den Zusammenhang von Selbstkonzept und Interesse speziell für leistungsstarke Schülerinnen und Schüler bekannt ist, werden nachfolgend Ergebnisse zum allgemeinen Zusammenhang berichtet. Da unter diesen Studien überdies nur wenige naturwissenschaftsbezogenes Selbstkonzept und Interesse betrachten, wird nachfolgend zusätzlich auf Erkenntnisse der Forschung zu mathematikbezogenem Selbstkonzept und Interesse zurückgegriffen.

Betrachtet man den Einfluss des Selbstkonzepts auf das Interesse in Naturwissenschaften, so geben einige Studien Belege für einen signifikanten Zusammenhang. Häussler, Hoffmann et al. (1998) und Hoffmann, Lehrke und Häussler (1998) weisen in Physik nach, dass das Selbstkonzept das Fachinteresse in Physik unterstützt und dabei in den jeweils getesteten Modellen stärkster und zum Teil einziger Prädiktor ist. Auch in

einer Studie zu Studenten beeinflusst der wahrgenommene Erfolg in Naturwissenschaften das Interesse an den Naturwissenschaften (Sjöberg, 1984). Mehr Längsschnittstudien existieren dagegen in Mathematik. In einer Analyse zweier längsschnittlicher Stichproben der BIJU-Studie (Köller, 1998) und TIMSS/III-Studie (Baumert et al., 1997) wurde ein deutlicher und für beide Geschlechter gleicher Einfluss des Selbstkonzeptes auf das Interesse nachgewiesen (Marsh et al., 2005). Eine weitere groß angelegte Längsschnittstudie von der ersten bis zur zwölften Jahrgangsstufe bestätigt das Ergebnis (Denissen, Zarrett & Eccles, 2007). Insgesamt finden Längsschnittstudien in Mathematik sowohl in der Mittelstufe (z. B. Köller et al., 2001) als auch der Oberstufe (z. B. Köller et al., 2006) einen deutlichen Einfluss des Selbstkonzeptes auf das Interesse.

Neben dem Einfluss von Selbstkonzept und Interesse gibt es auch Hinweise auf den Umgekehrten Effekt von Interesse auf Selbstkonzept, die wiederum im Bereich der mathematikbezogenen Forschung gefunden wurden. Während in den reziproken Analysen von Marsh et al. (2005) in beiden Stichproben das Selbstkonzept das Interesse erklärt, besteht umgekehrt nur in der TIMSS-Stichprobe ein geringer umgekehrter Effekt von Interesse nach Selbstkonzept. In der BIJU-Stichprobe wird der Zusammenhang nur unter Kontrolle anderer Merkmale knapp signifikant (Marsh et al., 2005). Auch in einer weiteren Analyse mit dem BIJU-Datensatz wird kein direkter Einfluss des Interesses auf das Selbstkonzept nachgewiesen (Köller et al., 2001). Das Interesse beeinflusst jedoch indirekt die Leistung, indem es die Kurswahl in der zehnten Jahrgangsstufe beeinflusst (Köller et al., 2001). Weitere Belege für den Einfluss von Interesse auf Selbstkonzept bzw. den reziproken Zusammenhang beider Merkmale liefern die Längsschnittstudie von Denissen et al. (2007) und die auf dem RIASEC-Modell basierende Studie von (Tracey, 2002). Obwohl Studien den Einfluss von Interesse auf Selbstkonzept belegen, sind die Ergebnisse weniger eindeutig als für die umgekehrte Einflussrichtung.

Eine mögliche Erklärung liegt im Alter der Schülerinnen und Schüler und damit einhergehenden Unterrichtsformen. Jacobs, Lanza, Osgood, Eccles und Wigfield (2002) weisen darauf hin, dass Längsschnittstudien mit breiteren Altersbereichen (Denissen et al., 2007; Marsh et al., 2005) Hinweise darauf liefern, dass der Zusammenhang zwischen Selbstkonzept und Interesse mit steigender Jahrgangsstufe stärker wird. Als mögliche Erklärung dafür wird argumentiert, dass der Einfluss von Interesse auf Selbstkonzept vor allem in schülerzentrierten Lernumgebungen, in denen intrinsische Motivation bzw. Interesses das Engagement und Lernen beeinflusst, wirksam wird und nachgewiesen

werden kann (vgl. auch Daniels, 2008; Köller et al., 2001). Solche Lernumgebungen werden vor allem der Oberstufe in Verbindung mit gewählten Leistungskursen zugeordnet (Köller et al., 2001).

Insgesamt wird zwischen Selbstkonzept und Interesse über Forschungsbereiche hinweg ein reziproker Zusammenhang angenommen und argumentiert, dass sowohl die Ergebnisse einer nur auf Leistung abzielenden Förderung als auch einer nur auf Selbstkonzept und Interesse fokussierenden Lernumgebung nur kurzfristig erreicht werden, eine langfristig erfolgreiche Strategie jedoch beide Merkmale fördert (Marsh, 2005). Während in Naturwissenschaften nur wenig wissenschaftliche Evidenz zum reziproken Zusammenhang von Selbstkonzept und Interesse existiert, geben Erkenntnisse aus dem Bereich Mathematik wichtige Anknüpfungspunkte. Einschränkend ist jedoch auch für Mathematik anzumerken, dass ein beträchtlicher Teil der existierenden zentralen Studien (Köller, Daniels et al., 2000; Köller et al., 2001, 2006; Marsh, 2005) auf einem identischen Datensatz aus der BIJU-Studie beruht.

Soziale und dimensionale Vergleiche bei der Genese von Interesse

Wie die vorherigen Abschnitte darlegen, stehen sowohl Leistung und Selbstkonzept als auch Leistung und Interesse sowie Selbstkonzept und Interesse in wechselseitigem Zusammenhang. Zusätzlich mediiert das Selbstkonzept den Einfluss von Leistung auf Interesse. Daraus folgt, dass Vergleichsprozesse, die im Zusammenhang Leistung-Selbstkonzept wirken, auch auf den Zusammenhang Leistung-Interesse übertragen werden können und möglicherweise einen Einfluss auf das Interesse von High Performern in Naturwissenschaften haben. Für die Betrachtung von Selbstkonzept und Interesse leistungsstarker Jugendlicher ist dies insofern relevant, als Vergleichsprozesse möglicherweise neben anderen Faktoren erklären können, warum Schülerinnen und Schüler trotz hoher Leistung vergleichsweise geringes Selbstkonzept und Interesse zeigen.

Weil Leistung sowohl Interesse als auch Selbstkonzept beeinflusst und Selbstkonzept den Zusammenhang Leistung-Interesse teilweise mediiert, postulieren Autoren, dass das I/E-Modell zur Genese des Selbstkonzepts auch auf Interesse übertragbar ist (Daniels, 2008). Im Rahmen der Selbstkonzeptforschung wurde zudem immer wieder kritisiert, dass sich die Forschung zu Referenzgruppeneffekten lediglich auf Selbstkonzept fokussiert und vorgeschlagen, auch andere Konstrukte in den Blick zu nehmen (vgl. Trautwein & Lüdtke, 2005). Verschiedene neuere Studien greifen dies auf und nehmen

dimensionale und soziale Vergleiche unter Berücksichtigung von Kontrast- und Assimilationseffekten bei der Entwicklung von Interesse für Mathematik in den Blick (Daniels, 2008; Köller, Schnabel & Baumert, 2000; Köller et al., 2006).

Blickt man zunächst auf die vermuteten sozialen Kontrasteffekte durch die Schul- bzw. Klassenleistung, so finden sich sowohl in der Oberstufe (Köller et al., 2006) als auch in der Mittelstufe (Trautwein, Lüdtke, Marsh et al., 2006) Belege dafür. Köller et al. (2006) zeigen für ihre Stichprobe, dass dieser BFLP-Effekt für Interesse vollständig vom Selbstkonzept mediiert wird. Darüber hinaus können positive Assimilationseffekte der Leistungskurszugehörigkeit in der Oberstufe (Köller et al., 2006), nicht jedoch der Schulart- bzw. Schulweizugehörigkeit in der neunten Klasse, auf das Mathematikinteresse nachgewiesen werden (Trautwein, Lüdtke, Köller & Baumert, 2006). Für die Förderung leistungsstarker Jugendlicher ist es in gewisser Weise besonders nachteilig, wenn Schulzweig bzw. Schulart keine Assimilationseffekte bewirken, denn in der Diskussion um die Leistungsgruppierung zur Förderung leistungsstarker Jugendlicher (vgl. Ireson & Hallam, 2009) wird häufig argumentiert, dass sie die Interessensentwicklung begünstigt (Plucker et al., 2004). Für die Interessenförderung leistungsstarker Jugendlicher bedeutet dies, dass eine Leistungsgruppierung entweder nicht interessenfördernd wirkt und das Argument ungültig ist oder, dass durch Leistungsgruppierung geschaffene Möglichkeiten zur Interessenförderung, z. B. auch durch soziale Peer-Effekte, nicht genutzt werden.

Während für soziale Vergleiche einige Untersuchungen existieren, beschäftigen sich mit dimensional Vergleichen zwischen Leistung und Interesse nur wenige Studien (vgl. Daniels, 2008). In ihnen zeigt sich ein negativer Einfluss der Leistung in anderen Leistungsbereichen auf die intrinsische Motivation (Skaalvik & Rankin, 1995) bzw. ein dimensionaler Effekt auf die Zufriedenheit mit der Leistung in Mathematik (Möller & Köller, 2001). Auch direkt für Interesse konnte ein dimensionaler Vergleich, der wiederum über das Selbstkonzept vermittelt wird, nachgewiesen werden (Köller, Daniels et al., 2000). Für eine Generalisierung fehlen hier jedoch noch weitere Studien.

Die meisten Studien betrachten dimensionale Vergleiche zwischen mathematischen und verbalen Domänen und nur wenige beziehen auch Naturwissenschaften mit ein. Zwischen Mathematik und Physik werden darin dimensionale Assimilationseffekte gefunden (Möller et al., 2006), zwischen Mathematik und Naturwissenschaften schwache dimensionale Kontrasteffekte (Chiu, 2012) und innerhalb der Naturwissenschaften keine

signifikanten (Netto-)Effekte (Jansen et al., 2014). Obwohl der bisherige Forschungsstand noch keine detaillierten Aussagen zulässt, kann angenommen werden, dass kontrastierende dimensionale Effekte zwischen mathematisch-naturwissenschaftlichen und verbalen Domänen stärker ausgeprägt sind als innerhalb naturwissenschaftlicher Domänen und davon abhängen, wie ähnlich Schülerinnen und Schüler Fächer wahrnehmen (Möller & Marsh, 2013).

Insgesamt kann vermutet werden, dass Schülerinnen und Schüler in einer leistungsstarken Klasse ein geringeres Interesse aufweisen als in einer leistungsschwachen Klasse und die Zugehörigkeit zu einem Leistungskurs positiven Einfluss auf das Interesse hat. Darüber hinaus interessiert sich ein Jugendlicher für Naturwissenschaften mit höherer Wahrscheinlichkeit, wenn es eines seiner besten Fächer ist, und interessiert sich umgekehrt weniger dafür, wenn er in anderen Fächern bessere Leistungen erreicht. Insgesamt liegen jedoch nur wenige Studien zum sozialen Vergleich vor und noch weniger zum dimensional Vergleich. Ähnlich zum Zusammenhang von Selbstkonzept und Interesse sind sie überwiegend im Bereich Mathematik angesiedelt. Für Naturwissenschaften und insbesondere leistungsstarke Jugendliche in Naturwissenschaften ist deshalb relativ unklar, in welchem Zusammenhang Leistung, Selbstkonzept und Interesse bei leistungsstarken Jugendlichen genau stehen und inwiefern soziale und dimensionale Vergleiche sich negativ auf das Selbstkonzept und Interesse in Naturwissenschaften auswirken und die positiven Voraussetzungen der hohen Leistung relativieren.

3.3. Profile aus Leistung, Selbstkonzept und Interesse

Leistung, Selbstkonzept und Interesse stehen in positivem wechselseitigem Zusammenhang und können sowohl als Ziele als auch als individuelle Einflussfaktoren für erfolgreiche Lernprozesse gesehen werden (Huang, 2011; Krapp & Prenzel, 2011; Seidel, 2006). Die vorangegangenen Kapitel zeigen, dass leistungsstarke Jugendliche in Naturwissenschaften demnach auch mit höherer Wahrscheinlichkeit ein hohes Selbstkonzept und Interesse in Naturwissenschaften aufweisen als leistungsschwächere Jugendliche (Kapitel 3). Gleichzeitig verdeutlichen Studien jedoch, dass ein beträchtlicher Teil der leistungsstarken Jugendlichen wenig leistungszuversichtlich und uninteressiert ist. Dies bedeutet, dass auch Kombinationen aus Selbstkonzept und Interesse unter leistungsstarken Jugendlichen existieren, die inkonsistent zum positiven linearen Zusammenhang von Leistung und Selbstkonzept bzw. Interesse sind – beispielsweise sich unterschätzende Jugendliche – und ein Teil der Jugendlichen motivational-affektiv indifferent bleibt. Für eine angemessene motivational-affektive Förderung leistungsstarker Jugendlicher ist es wichtig, zu wissen, wie viele und welche der leistungsstarken Jugendlichen in Naturwissenschaften motivational-affektiv indifferent sind und welche Wirkmechanismen dahinterstecken. Nachfolgend wird dazu zunächst dargestellt, wie solche inkonsistenten Kombinationen theoretisch erklärt werden können, welche methodischen Implikationen ihre Betrachtung hat und welche Forschungsergebnisse allgemein und für leistungsstarke Jugendliche bekannt sind.

Studien zu Interesse und vor allem Selbstkonzept beschäftigen sich üblicherweise mit linearen Zusammenhängen zwischen Leistung, Selbstkonzept und Interesse. Grundsätzlich ergeben die Befunde einen positiven wechselseitigen Zusammenhang zwischen Leistung, Selbstkonzept und Interesse (z. B. Köller et al., 2006). Darüber hinaus wurden jedoch auch Einflussfaktoren wie Geschlecht (Hoffmann, 2002), Leistung der Klasse (Big Fish Little Pond Effect, Marsh & Yeung, 1997b), Leistungen in anderen Fächern (dimensionale Vergleiche, Möller et al., 2006) oder auch dem sozioökonomischen Hintergrund (Prenzel & Schütte, 2007) gefunden, die sich negativ auf das Selbstkonzept und Interesse auswirken und inkonsistente Kombinationen erklären könnten. Insgesamt bedeutet dies, dass eine hohe Leistung zwar ein hohes Selbstkonzept und Interesse wahrscheinlicher macht, bestimmte Mechanismen wie Referenzrahmeneffekte oder Geschlechterdifferenzen aber möglicherweise auch verhindern, dass diese drei Merkmale parallel ausgeprägt

sind.

Die Analyse solcher unerwarteter Kombinationen aus Leistung und Selbstkonzept und Interesse erfordert jedoch, Merkmalskombinationen anstatt Stichprobenmittelwerte einzelner Variablen zu betrachten und Gruppen leistungsstarker Schülerinnen und Schüler mit ähnlichen Profilen zu identifizieren. Aus methodischer Sicht impliziert dies ein personenzentriertes anstatt des üblicherweise verwendeten variablenzentrierten Vorgehens. Die Vorteile der Analyse von Merkmalsprofilen bzw. personenzentrierten Vorgehen sind, dass durch die gleichzeitige Betrachtung mehrerer Merkmale ein holistischer Ansatz verfolgt werden kann (Bergman & Magnusson, 1997). Dies berücksichtigt zum einen die Bedeutung funktional organisierter Systeme und ermöglicht gleichzeitig, Merkmalskombinationen anschaulich und greifbar darzustellen und zu beschreiben (Lau & Roeser, 2008; Roeser, Eccles & Sameroff, 1998; Stemmler & Heine, 2015). Gleichzeitig kann eine Stichprobe mit unterschiedlichen Verteilungen in den Merkmalen „entmischt“ werden und schwache Korrelationen zwischen Merkmalen können tiefergehend analysiert werden.

Während in weiten Teilen der Selbstkonzeptforschung bisher die Erforschung der linearen variablenzentrierten Prozesse im Vordergrund stand und weniger beabsichtigt wurde, mögliche Merkmalskombinationen methodisch abzubilden, gewann die Untersuchung von Merkmalskombinationen in der Lern- und Unterrichtsforschung in den letzten beiden Jahrzehnten zunehmend an Bedeutung und es wurden vermehrt personenzentrierte Analysen verwendet (z. B. Lau & Roeser, 2008; Linnenbrink-Garcia, Pugh, Koskey, Kristin L. K. & Stewart, 2012a; Seidel, 2006; Wormington, Corpus & Anderson, 2012). Während personenzentrierte Analysen dabei für Merkmale wie Zielorientierungen (z. B. Köller, 1998), Lernmotivation (z. B. Vansteenkiste, Sierens, Soenens, Luyckx & Lens, 2009; Wormington et al., 2012) und Interesse (z. B. Häussler, Hoffmann et al., 1998) bereits häufiger verwendet wurden, existieren deutlich weniger Studien, die kognitive und motivational-affektive Merkmale in Naturwissenschaften kombinieren und das Geschlecht miteinbeziehen (z. B. Jurik, Gröschner & Seidel, 2013).

Eine Ausnahme für die Kombination kognitiver und motivational-affektiver Merkmale bildet die Studie von Seidel (2006, vgl. auch Seidel, 2003). Sie untersucht Profile aus Leistung, Selbstkonzept und Interesse in Physik und findet fünf verschiedene Typen, die sich durch unterschiedliche Merkmalsprofile auszeichnen, darunter sich über- und unterschätzende sowie trotz höherem Selbstkonzept wenig interessierte Schülerinnen und

Schüler. Auch Lau und Roeser (2008) und Linnenbrink-Garcia et al. (2012a) betrachten unterschiedliche Kombinationen aus kognitiven und motivational-affektiven Merkmalen allgemein bzw. aus Leistung, Selbstwirksamkeit bzw. Selbstkonzept und Interesse in Biologie. Obwohl nur auf Interesse fokussiert, liefern auch Häussler, Hoffmann et al. (1998) Hinweise auf Selbstkonzept-Interesse-Typen, indem sie das Selbstkonzept von Interessenprofilen in Physik betrachten. In der Studie werden drei Interessen-Typen identifiziert, die den Interessensbereichen (a) Physik und Technik, (b) Mensch und Natur sowie (c) Gesellschaft zugeordnet werden. Ein Vergleich des Selbstkonzepts der Typen in Physik zeigt, dass ein Zusammenhang zwischen Interesseprofil und Selbstkonzept besteht. Demnach sind Schülerinnen und Schüler des physikalisch-technisch interessierten Profils leistungszuversichtlicher als Jugendliche des gesellschaftlich interessierten Profils (Häussler, Hoffmann et al., 1998).

Zahlreiche Studien haben den Einfluss des Geschlechts auf individuelle Schüler-voraussetzungen wie Leistung, Selbstkonzept und Interesse in Naturwissenschaften gezeigt (z. B. Buccheri et al., 2011; Hannover, 1998; Hoffmann, 2002; OECD, 2007). Unter den personenzentrierten Studien berücksichtigen aber nur wenige das Geschlecht (z. B. Jurik et al., 2013; Jurik, Gröschner & Seidel, 2014; Linnenbrink-Garcia et al., 2012a) und belegen, dass es mit den Profilen im Zusammenhang mit verbalem Engagement (Jurik et al., 2013), Conceptual Change (Linnenbrink-Garcia et al., 2012a) sowie kognitiven Lernaktivitäten und intrinsischer Lernmotivation (Jurik et al., 2014) interagiert. Eine getrennte Ermittlung der Profile für beide Geschlechter führen von den genannten Untersuchungen jedoch nur Lau und Roeser (2008) durch. Sie finden unterschiedliche Profile für Mädchen und Jungen und lediglich zwei Profilverläufe, die in beiden Geschlechtergruppen hoch korrelieren. Allerdings untersuchen sie motivational-affektive Merkmale wie Selbstbewusstsein und Emotionen und nicht direkt Selbstkonzept und Interesse an den Naturwissenschaften. Inwiefern innerhalb der Mädchen und Jungen unterschiedliche Profile hinsichtlich Selbstkonzept und Interesse bzw. Leistung, Selbstkonzept und Interesse bestehen, kann aus dem momentanen Forschungsstand daher nicht abgeleitet werden.

Für leistungsstarke Schülerinnen und Schüler in Naturwissenschaften deuten variablenzentrierte Studien an, dass ein beträchtlicher Anteil ein in Relation zu ihrer Leistung geringes Interesse hat (z. B. Prenzel & Schütte, 2007), Kombinationen bzw. Profile aus Selbstkonzept und Interesse wurden bisher jedoch nicht betrachtet. Hinweise darauf,

dass innerhalb der leistungsstarken Schülerinnen und Schüler mehrere Profile existieren, können aus der personenzentrierten Analyse von Seidel (2006) abgeleitet werden. So wurden innerhalb der fünf Typen zwei leistungsstarke Profile mit hohem und niedrigem Selbstkonzept bzw. Interesse gefunden (s. Abbildung 2). Schülerinnen und Schüler des ersten Profils (*Gruppe 1*) zeigen neben einer hohen Leistung auch hohes Selbstkonzept und Interesse, Schülerinnen und Schüler des zweiten Profils (*Gruppe 2*) sind in Relation zur Leistung wenig leistungszuversichtlich und interessiert (Seidel, 2006). Wenngleich dies darauf hinweist, dass innerhalb der leistungsstarken Jugendlichen Profile mit hohem und niedrigem Interesse identifizierbar sind, bleibt unklar, wie viele und welche Profile innerhalb der Gruppe der leistungsstarken Jugendlichen existieren. Aus bisherigen Analysen können zudem keine Anhaltspunkte dafür gefunden werden, ob innerhalb leistungsstarker Mädchen und Jungen unterschiedliche Selbstkonzept-Interessenprofile oder wie bei Lau und Roeser (2008) parallele Merkmalskombinationen existieren.

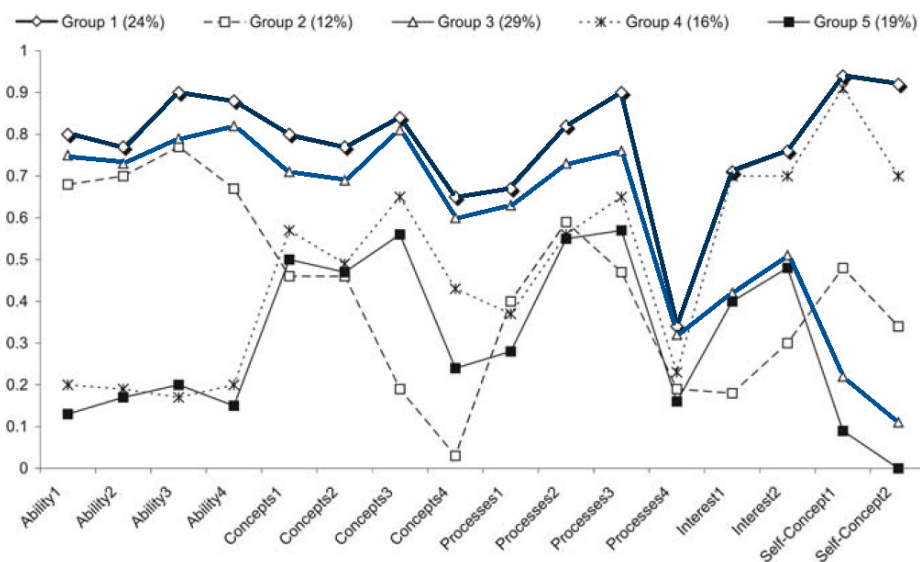


Abbildung 2. Profile aus Leistung, Selbstkonzept und Interesse von Seidel (2006, S. 260, leistungsstarke Profile blau markiert)

Darüber hinaus ist unklar, in welchen für Selbstkonzept und Interesse relevanten Eigenschaften sich solche Gruppen unterscheiden. Bisherige Studien, die Profile aus Leistung, Selbstkonzept und Interesse betrachten, untersuchen Unterrichtswahrnehmung, Schülerverhalten und Geschlecht (z. B. Jurik et al., 2013). Mögliche andere Einflüsse wie soziale und dimensionale Vergleiche wurden bisher nicht analysiert. Für Schülerinnen und Schüler allgemein und insbesondere leistungsstarke Jugendliche ist beispielsweise unklar, welchen Selbstkonzept-Interesse-Profilen leistungsstarke Jugendliche angehören,

3. Naturwissenschaftsbezogenes Interesse und Fähigkeitsselbstkonzept

die in anderen Fächern genauso gut bzw. besser sind als in Naturwissenschaften oder inwiefern umgekehrt ein relevanter Anteil an leistungsstarken, leistungszuversichtlichen aber uninteressierten Jugendlichen möglicherweise in anderen Fächern noch bessere Leistungen als in Naturwissenschaften zeigt.

3.4. Motivationsunterstützende Lehr-Lernbedingungen

Die vorausgehenden Abschnitte haben gezeigt, dass die Leistung in verschiedenen Fachbereichen sowie das Selbstkonzept für bestimmte Fragestellungen zu Interesse berücksichtigt werden sollten. Neben diesen individuellen Merkmalen spielen auch Unterrichtsmerkmale eine bedeutende Rolle für Interesse bzw. interessierte Lernmotivation und können auf High Performer übertragen werden. Zentrale Unterrichtsfaktoren für die Entwicklung von interessierter und intrinsischer Lernmotivation sind motivationsunterstützende Lehr-Lernbedingungen (z. B. Prenzel, 1997; Seidel & Prenzel, 2006b). Sie haben sowohl die Funktion, intrinsische und interessierte Lernmotivation zu erzeugen und aufrechtzuerhalten als auch die Entwicklung extrinsisch induzierter Lernmotivation hin zu intrinsischer Lernmotivation zu unterstützen, indem sie die Internalisierung externer Anforderungen anregen (Prenzel, 1993; Seidel, 2003). Prenzel (1997) formuliert sechs Bedingungen zur Unterstützung interessierter und intrinsischer Lernprozesse (vgl. auch Seidel, 2003; Prenzel, Kramer & Drechsel, 2001).

- Kompetenzunterstützung (z. B. Zutrauen in die Leistung signalisieren, informierendes Feedback, individuelle Bezugsnorm)
- Soziale Eingebundenheit (z. B. Akzeptanz und Ernstnehmen der Schüler durch die Lehrkraft, kollegiales und kooperatives Arbeiten, Empathie)
- Autonomieunterstützung (z. B. Wahl- bzw. Entscheidungsmöglichkeiten, Freiräume, Unterstützung des selbstständigen, entdeckenden Handelns und Lernens)
- Inhaltliche Relevanz (z. B. Anwendungsbezüge, Verknüpfung mit realen Situationen sowie Lebenswelt und Präferenzen der Schüler, Transparenz von Absichten und Zielen)
- Instruktionsqualität (z. B. klare Struktur, Verständlichkeit, Handlungsorientierung, Anstreben anspruchsvollerer Ziele wie Verstehen und Abstraktion)
- Interesse der Lehrperson (z. B. Engagement, Enthusiasmus, Offenbarung von Empfindungen)

Die ersten drei Bedingungen beruhen auf der anthropologischen Überlegung, dass das Verhalten von Menschen neben Trieben und physiologischen Bedürfnisse durch drei

Bedürfnisse nach Autonomie, Kompetenz und sozialer Eingebundenheit beeinflusst wird. Je mehr diese Grundbedürfnisse in einer Situation erfüllt werden, desto selbstbestimmter erlebt die Person die jeweilige Situation und desto intrinsischer motiviert handelt sie (Ryan & Deci, 2000a, 2000b, 2002).

Darüber hinaus nennt (Prenzel, 1997) drei weitere Bedingungen selbstbestimmten Lernens (Inhaltliche Relevanz, Instruktionsqualität und Interesse der Lehrperson), die im Einklang mit Erkenntnissen der Forschung zu Lernmotivation (Lewalter, Krapp, Schreyer & Wild, 1998; Rheinberg, 1998), Interesse (Hoffmann, Krapp et al., 1998; Krapp & Prenzel, 1992; Renninger, Hidi & Krapp, 1992) und konstruktivistischen Lehr-Lern-Modellen (J. S. Brown, Collins & Duguid, 1989; Collins, Brown & Newman, 1989) stehen (vgl. Prenzel, Kramer & Drechsel, 2001; Seidel, 2003). Sie berücksichtigen auch instruktionale und inhaltsbezogene Aspekte, indem sie die Wahrnehmung und Identifizierung mit Anforderungen sowie inhaltlichen Anreizen und Bedeutungen miteinschließen (Prenzel, Kramer & Drechsel, 2001; Seidel, 2003). Während die Selbstbestimmungstheorie annimmt, dass autonom erlebte Lernprozesse automatisch in das Selbst integriert und als bedeutsam erlebt werden (z. B. Deci & Ryan, 1993), unterscheidet Prenzel (1997) zwischen *Autonomie* im Sinne von Wahlmöglichkeiten und Entscheidungsfreiräumen und *inhaltlicher Relevanz* als Zieltransparenz und Bedeutung von Lerninhalten. Letztere beruht auf der Beziehung zwischen Anforderungen und Inhalten der Lernsituation und Vorlieben, Absichten und Interessen des Lerner.

Aus empirischer Sicht konnten zahlreiche Studien zeigen, dass Schülerinnen und Schüler umso selbstbestimmter motiviert sind, umso mehr sie ihren Unterricht autonomieunterstützend, kompetenzunterstützend und sozial einbindend erleben (Lewalter et al., 1998; Wild, 2000). Ebenso bestätigt sich die Bedeutung der inhaltlichen Relevanz, Instruktionsqualität und Inhaltliches Interesse der Lehrperson für interessierte bzw. intrinsische Lernmotivation empirisch in Naturwissenschaften (z. B. Seidel, 2003). Je mehr Schülerinnen und Schüler in den sechs genannten Lehr-Lernbedingungen unterstützt werden, desto mehr selbstbestimmte bzw. intrinsische Formen der Lernmotivation erleben sie (Prenzel, Häussler, Rost & Senkbeil, 2002; Prenzel, Kramer & Drechsel, 2001; Seidel, 2003). Umgekehrt berichten Jugendliche umso mehr fremdbestimmte Motivation, je weniger motivationsunterstützende Bedingungen erfüllt werden (Prenzel, Kramer & Drechsel, 2001; Seidel, 2003). Auch die Entwicklung stabilen Interesses in einem Unterrichtsfach (Krapp, 2005) sowie komplexe Wissenserwerbsprozesse werden durch

motivationsunterstützende Lehr-Lernbedingungen begünstigt (U. Schiefele, 1996; Seidel, 2003). Das Kompetenzerleben bzw. die wahrgenommene Kompetenzunterstützung spielen außerdem eine wichtige Rolle für die Entwicklung eines positiven Selbstkonzepts (Jansen, 2015).

Da die Selbstbestimmungstheorie die psychologischen Grundbedürfnisse nach Autonomie, Kompetenz und sozialer Einbindung unabhängig von individuellen Merkmalen annimmt (z. B. Deci & Ryan, 2000) und die prozessbezogenen Begründungen der instruktionalen und inhaltsbezogenen Bedingungen von Instruktionsqualität, inhaltlicher Relevanz und Interesse der Lehrperson ebenfalls unabhängig von der Leistung gelten (vgl. Prenzel, 1997), können diese motivationsunterstützenden Lehr-Lernbedingungen auf die Entwicklung interessierter Lernmotivation von High Performern in Naturwissenschaften übertragen werden. Sie sind demnach Bedingungsfaktoren, die die interessierte Lernmotivation leistungsstarker Jugendlicher in Naturwissenschaften in Unterrichtsumgebungen einerseits induzieren und erhalten sowie andererseits zur Internalisierung externer Anforderungen führen können. Prenzel (1997) greift auch die umgekehrte Denkrichtung auf und beschreibt, wie Schülerinnen und Schüler demotiviert werden können, indem motivationsunterstützende Bedingungen nicht erfüllt werden.

Dieser Ansatz wurde in der IPN-Videostudie für Physik genutzt und gezeigt, dass nicht die Organisation des Unterrichts, sondern die Qualität der Lehr-Lernbedingungen ausschlaggebend für motivationale Merkmale ist (Seidel, 2003; Seidel & Prenzel, 2006a). So führten Klassengespräche mit hoher Beteiligung und gleichberechtigter Gesprächspartnerschaft der Schülerinnen und Schüler im Gegensatz zu eng geführten Klassengesprächen zu höher wahrgenommener Motivationsunterstützung und mehr intrinsischer Motivation (Seidel, Rimmele & Prenzel, 2003). Ebenso konnten Zielklarheit und Kohärenz sowie Lernorientierung die Lernmotivation der Schülerinnen und Schüler in Physik vorhersagen (Seidel, Rimmele & Prenzel, 2005; Seidel, Prenzel, Rimmele et al., 2007). Darüber hinaus zeigte sich, dass Beispiele bzw. Anwendungen zu mehr inhaltlicher Relevanz und mehr Beteiligungsmöglichkeiten zu höher wahrgenommener Autonomie führten und so die Motivation unterstützten (Dalehefte, 2001; Seidel, 2003). Übertragen auf den Kontext High Performer bietet dieser Ansatz mögliche instruktionale Erklärungen dafür, dass ein Teil der leistungsstarken Jugendlichen trotz hoher Leistungen ein niedriges Selbstkonzept und/oder Interesse für Naturwissenschaften zeigt, d. h. motivational-affektiv unerreicht bzw. indifferent bleibt.

Verschiedene Studien haben den Zusammenhang von motivationsunterstützenden Lehr-Lernbedingungen und Lernmotivation, interessierter Motivation oder verwandten motivational-affektiven Merkmalen auch bereits für leistungsstarke Jugendliche untersucht. Entscheidungsmöglichkeiten als autonomieunterstützende Gestaltungsmaßnahme unterstützen intrinsische Motivation leistungsstarker Jugendlicher (Garn & Jolly, 2014). Dies belegen auch Linnenbrink-Garcia, Patall und Messersmith (2013) in einer Studie, in der $N = 125$ talentierte Jugendliche (identifiziert über einen Mathematiktest) an einem dreiwöchigen Sommerprogramm in Naturwissenschaften teilnehmen. Die wahrgenommene Wahlmöglichkeit und Bedeutung für das reale Leben sagten das situationale Interesse vorher und beeinflussten – mediiert über das situationale Interesse – das individuelle Interesse der talentierten Jugendlichen an Naturwissenschaften. Darüber hinaus förderte die Wahlmöglichkeit – ebenso mediiert durch das situationale Interesse – die Selbstwirksamkeit der Jugendlichen (Linnenbrink-Garcia et al., 2013). Selbstwirksamkeit und Selbstkonzept bilden zwar zwei getrennte Konzepte ab, stehen jedoch in engem Zusammenhang zueinander (Jansen, Scherer & Schroeders, 2015). Ebenso zeigen leistungsstarke Schülerinnen und Schüler in der Grundschule einen Zusammenhang von Kompetenzerleben und Freude sowie Autonomieerleben und Neugierde (Miserandino, 1996).

Auch Forschung zu Schülerprofilen hinsichtlich Leistung, Selbstkonzept und Interesse deutet an, dass unter den leistungsstarken Jugendlichen leistungszuversichtliche und interessierte Schülerinnen und Schüler tendenziell mehr inhaltliche Relevanz, Autonomieunterstützung und Kompetenzunterstützung wahrnehmen als sich unterschätzende und uninteressierte leistungsstarke Schülerinnen und Schüler (Seidel, 2006) sowie eher intrinsisch motiviert sind (Jurik et al., 2014). Während in der genannten Studie paarweise Unterschiede zwischen Profilen statistisch nicht geprüft wurden, bestätigen weiterführende Analysen statistisch signifikant, dass mehr Kompetenzunterstützung, soziale Eingebundenheit und Autonomieunterstützung mit einer positiven Selbstkonzept- und Interessensentwicklung von sich unterschätzenden leistungsstarken Jugendlichen innerhalb eines Schuljahres einhergeht (Huber et al., 2015).

Wenngleich einige Studien die Bedeutung motivationsunterstützender Lehr-Lernbedingungen für leistungsstarke Jugendliche zeigen, beruht ein Großteil der Ergebnisse auf kleinen Stichprobenzahlen (für eine Ausnahme s. Linnenbrink-Garcia et al., 2013). Gleichzeitig definieren alle genannten Studien leistungsstarke Jugendliche auf Basis von

Mathematiktests (Linnenbrink-Garcia et al., 2013) oder allgemeinen Leistungstests bzw. Noten (Garn & Jolly, 2014; Miserandino, 1996). Es bleibt daher offen, inwiefern Interesse bzw. Selbstkonzept von domänenspezifisch und auf Basis eines standardisierten Tests definierten leistungsstarken Schülerinnen und Schüler in Naturwissenschaften im Zusammenhang mit motivationsunterstützenden Lehr-Lernbedingungen stehen und ob dieser Zusammenhang in einer repräsentativen Studie belegt werden kann. Personenzentrierte Studien liefern darüber hinaus wichtige Hinweise auf einen positiven Zusammenhang zwischen Selbstkonzept-Interesse-Typen leistungsstarker Jugendlicher und motivationsunterstützenden Lehr-Lernbedingungen, untersuchen leistungsstarke Jugendliche jedoch nur indirekt. Offen bleibt daher, inwiefern in einer explizit auf High Performer fokussierten personenzentrierten Analyse Gruppen von wenig leistungszuversichtlichen, uninteressierten High Performern („motivational-affektiv indifferente High Performer“) und leistungszuversichtliche und/oder interessierte High Performern gefunden werden und Unterschiede in der Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen als differenzielle Merkmale dieser Gruppen identifiziert werden können.

3.5. Zusammenfassung

Die vorherigen Kapitel hatten zum Ziel, Interesse und Selbstkonzept in Naturwissenschaften als Konzepte sowie den Forschungsstand zu leistungsstarken Jugendlichen in Naturwissenschaften zu präzisieren. Ebenso wurde die Frage nach dem Zusammenhang von Selbstkonzept und Interesse, auch unter Berücksichtigung des Einflusses von Leistung in Naturwissenschaften, diskutiert.

Auf Basis des aktuellen Forschungsstandes kann demzufolge ein reziproker Zusammenhang zwischen Leistung, Selbstkonzept und Interesse angenommen werden. Dabei treten sowohl im Einfluss von Leistung auf Selbstkonzept als auch im Einfluss von Leistung auf Interesse soziale und dimensionale Vergleichsprozesse auf. Mögliche Merkmalskombinationen aus Leistung, Selbstkonzept und Interesse werden darüber hinaus zunehmend in personenzentrierten Analysen untersucht und damit einer holistischen Sicht der Entwicklung von kognitiven und motivational-affektiven Merkmalen von Schülerinnen und Schülern Rechnung getragen. Trotz erster Ergebnisse bleiben Forschungslücken darüber, inwiefern sich Leistung-Selbstkonzept-Interesse-Profile zwischen den Geschlechtern unterscheiden.

Die Gruppe der leistungsstarken Jugendlichen in Naturwissenschaften wurde bisher generell nur in wenigen Studien hinsichtlich ihres Interesses und Selbstkonzeptes beschrieben. Diese Studien deuten an, dass ein beträchtlicher Teil dieser Jugendlichen wenig Interesse für Naturwissenschaften zeigt. Über das Selbstkonzept und Interesse einer breiter definierten Gruppe leistungsstarker Jugendlicher in Naturwissenschaften in Deutschland ist allerdings bisher wenig bekannt. Insbesondere darüber, welche Noten sie in erzielen und inwiefern ihr Leistungsprofil in Verbindung zum Selbstkonzept und Interesse in Naturwissenschaften steht, d. h. inwiefern sich beispielsweise Jugendliche, die nicht nur in Naturwissenschaften, sondern auch Sprachen exzellent sind, sich für Naturwissenschaften interessieren, ist unklar. Darüber hinaus deuten Ergebnisse personenzentrierter Studien an, dass auch unter leistungsstarken Jugendlichen in Naturwissenschaften unterschiedliche Typen aus Selbstkonzept und Interesse existieren. Bisher wurden Selbstkonzept-Interesse-Profile innerhalb der leistungsstarken Jugendlichen in Naturwissenschaften sowie Eigenschaften dieser Profile nicht untersucht. Insbesondere ist unklar, inwiefern leistungsstarke Mädchen und Jungen unterschiedliche Merkmalskombinationen hinsichtlich Selbstkonzept und Interesse aufweisen.

4. Die Bedeutung des Unterrichts für die Kompetenz- und Interessenentwicklung

Der Naturwissenschaftsunterricht in der Schule bietet für leistungsstarke und alle anderen Schülerinnen und Schüler einen Kontext, in dem sie sich systematisch und organisiert mit naturwissenschaftlichen Inhalten auseinandersetzen (Seidel & Prenzel, 2006b, 2006a; Seidel, Prenzel, Rimmele et al., 2007). Naheliegende Frage ist daher, welche Merkmale des Unterrichts naturwissenschaftliche Kompetenz und Interesse an den Naturwissenschaften bei leistungsstarken Schülerinnen und Schülern fördern und welche Auswirkungen ein solcher Unterricht auf die Mitschüler leistungsstarker Jugendlicher hat. Unter „fördern“ wird eine angemessene bzw. vergleichbare Förderung leistungsstarker Jugendlicher verstanden, in der das Ziel verfolgt wird, bei leistungsstarken ebenso wie bei Jugendlichen anderer Leistungsniveaus alle vorhandenen Potentiale zu entfalten und sie durch differenzierte und auf ihr Leistungsniveau angepasste Lernangebote in ihrer Entwicklung zu unterstützen.

4.1. Rahmenmodell zur Untersuchung des Naturwissenschaftsunterrichts

Betrachtet man die Unterrichtsforschung der letzten Jahrzehnte, so lassen sich verschiedene Sichtweisen auf Unterricht identifizieren, die auf den Kontext High Performance übertragen werden können. So lassen sich einerseits Ansätze finden, die sich auf Strukturen und Prozessen effektiven Unterrichts fokussieren und wichtige Hinweise auf Merkmale effektiven Unterrichts liefern. Andererseits existieren jedoch auch Ansätze, die

sich eher aus einer differentialpsychologischen Perspektive mit Unterricht beschäftigen, indem sie individuelle Merkmale in Interaktion mit Unterricht betrachten. Beide Sichtweisen in Kombination legen die Grundlage für die Betrachtung effektiven Unterrichts im Kontext High Performance in dieser Arbeit und werden nachfolgend kurz beschrieben und anschließend im Angebots-Nutzungs-Modell von Unterricht aufgelöst. Im Anschluss daran werden zentrale Unterrichtsaktivitäten des Naturwissenschaftsunterricht im Kontext High Performance beschrieben. In einem abschließenden Teil wird die differenzielle Sicht auf Unterrichtsprozesse erneut vertieft, indem dargestellt wird, dass gleicher Unterricht unterschiedlich wahrgenommen werden kann und Unterschiede in individuellen Schülervoraussetzungen zur Erklärung dieser Wahrnehmungsdifferenzen beitragen.

4.1.1. Forschungsparadigmen der Unterrichtsforschung

Lange Zeit war das Prozess-Produkt-Modell von Dunkin und Biddle (1974) das vorherrschende Modell der Unterrichtsforschung (Helmke, 2010). Das Modell berücksichtigt Produkt-, Prozess-, Voraussetzungs- und Kontextvariablen von Unterricht. Forschung auf Grundlage des Prozess-Produkt-Modells untersucht den Einfluss von Prozessmerkmalen des Unterrichts auf Lernergebnisse, meistens Leistung oder Leistungszuwachs (Helmke, 2010). Verschiedene Metaanalysen geben eine Übersicht über die im Rahmen der Prozess-Produkt-Forschung identifizierten Merkmale effektiven Unterrichts und ihre Effektstärken (Fraser & Walberg, H., Welch, W., & Hattie, J., 1987; Scheerens & Bosker, 2011; Seidel & Shavelson, 2007). Demnach erweisen sich die Lernzeit, die Zielorientierung und Strukturierung von Unterricht, interaktives und kooperatives Lehren und Lernen und kognitive Aktivierung und Herausforderung als einflussreich auf Lernen (Seidel & Shavelson, 2007).

Wenngleich sich ein Großteil der Forschung zu Unterrichtseffektivität auf domänenübergreifende Unterrichtsmerkmale fokussiert, wurden in den letzten Jahrzehnten auch zunehmend domänenspezifische Unterrichtsstrukturen und Unterrichtsprozesse untersucht und ihre Berücksichtigung gefordert (vgl. Helmke & Schrader, 2010; K. Neumann, Kauertz & Fischer, 2012; Seidel & Prenzel, 2006b; Seidel & Shavelson, 2007). Im Naturwissenschaftsunterricht legen deklarative, prozedurale und epistemische Lerninhalte und -Ziele, z. B. zur „Natur der Naturwissenschaft“ (*nature of science*), domänenspezifische Unterrichtsaktivitäten wie naturwissenschaftliche Untersuchungen (*scientific inquiry*), die Nutzung von Modellen und Repräsentationen und Anwendung von Naturwis-

senschaften auf Alltagssituationen nahe (Duit & Treagust, 1998; Harlen, 1999; Hofstein & Lunetta, 2004; Seidel & Prenzel, 2006b). Metaanalysen (z. B. Scheerens, Luyten, Steen & Luyten-Thouars, 2007; Seidel & Shavelson, 2007) zeigen, dass domänenspezifische bzw. nahe an Lerninhalte geknüpfte Merkmale höhere Effektstärken zeigen als eher distale Merkmale.

Der ursprüngliche Prozess-Produkt-Ansatz geht davon aus, dass solche Unterrichtsstrukturen und Prozessmerkmale unabhängig von individuellen Eigenschaften der Schülerinnen und Schüler wirken bzw. verfolgt vielmehr das Ziel, global effektive Merkmale von Unterricht aus einer systemischen Sichtweise zu identifizieren. In späteren Arbeiten auf Basis des Prozess-Produkt-Paradigmas wurde diese Annahme erweitert und individuelle Wahrnehmungs- und Verarbeitungsprozesse von Schülerinnen und Schüler sowie der dabei mögliche Einfluss individueller Schülervoraussetzungen in Prozess-Mediations-Produkt-Modellen berücksichtigt (z. B. Ditton, 2002a).

Eine solche Interaktion zwischen individuellen Schülervoraussetzungen und Unterrichtsmerkmalen ist der zentrale Aspekt einer zweiten, differenziellen Sichtweise auf Unterricht. Sie findet sich z. B. in der Forschung zu Aptitude Treatment Interaction (Brünken & Leutner, 2005). Grundlegende Annahme ist hier, dass individuelle Lernvoraussetzungen die Wirksamkeit einer Lernumgebung beeinflussen können bzw. Treatments in Abhängigkeit individueller Voraussetzungen unterschiedlich wirken und Lernumgebungen aus diesem Grund möglichst an individuelle Lernvoraussetzungen angepasst werden sollen (Hasebrook & Brünken, 2010; Snow, 1989). Der Ausgangspunkt solcher differenzieller Ansätze ist die Suche nach statistischen Interaktionen, häufig zwischen individuellen Merkmalen und Treatment.

4.1.2. Angebot-Nutzungsmodell des Naturwissenschaftsunterrichts

Für die Förderung von High Performern in Naturwissenschaften ist sowohl eine prozessbezogene als auch differenzialpsychologische Sichtweise auf Unterricht von Bedeutung. Aus ersterer können zentrale Merkmale effektiven Unterrichts identifiziert werden, aus zweiterer kann der Einfluss der hohen Leistung und Unterschiede zu den übrigen Schülerinnen und Schüler auf die Wirkungsweise der Unterrichtsmerkmale abgeleitet werden. Eine gleichzeitige Berücksichtigung beider Perspektiven ermöglichen Angebots-Nutzungs-Modelle von Unterricht (s. Abbildung 3) (Fend, 1980, 2001; Helmke & Weinert, 1997; Helmke, 2007). Sie beschreiben Unterrichtsaktivitäten als Angebot bzw. Lern-

4. Die Bedeutung des Unterrichts für die Kompetenz- und Interessenentwicklung

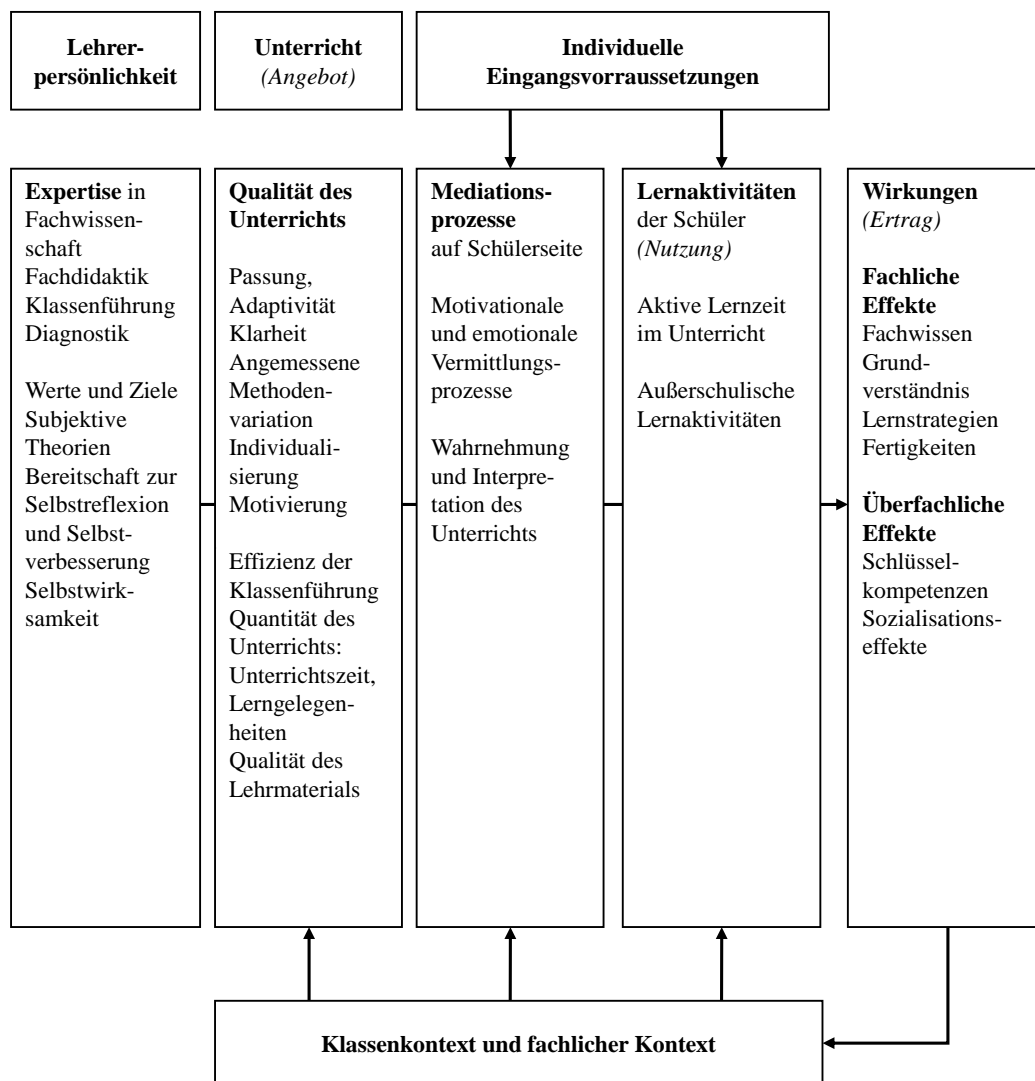


Abbildung 3. Angebots-Nutzungs-Modell der Unterrichtsforschung (Helmke, 2007, S. 42)

gelegenheit, dessen Wirkung auf Lernergebnisse durch die Nutzung der Schülerinnen und Schüler bestimmt wird. Der Einfluss des Unterrichtsangebotes auf die Unterrichtsziele (Ertrag) erfolgt dabei vermittelt über Wahrnehmungsprozesse des Unterrichts, motivationalen und volitionalen Merkmale sowie Lernaktivitäten (Nutzung). Ausgehend von konstruktivistischen Überlegungen in Bezug auf Lernen und die Wahrnehmung der Umwelt berücksichtigen Angebots-Nutzungs-Modelle darüber hinaus wechselseitige Interaktionen zwischen Merkmalen der Schülerinnen und Schüler und der Unterrichtsgestaltung, -Wahrnehmung und -Wirkung (Seidel, 2014).

In der deutschsprachigen Forschung zum Naturwissenschaftsunterricht werden Angebots-Nutzungs-Modelle bereits seit längerer Zeit in größer angelegten Studien mit ausreichender Stichprobenzahl genutzt (z. B. K. Neumann et al., 2012; Fischer, Neumann, Labudde & Viiri, 2014). International greifen Seidel und Prenzel (2006b) im Rahmen von PISA 2006 (OECD, 2007) das Angebots-Nutzungs-Modell erstmals für den Naturwissenschaftsunterricht auf und entwickeln in Anlehnung an PISA 2000 (Baumert, Stanat & Demmrich, 2001) sowie die IPN-Videostudie für den Physikunterricht (Seidel & Prenzel, 2004; Seidel, Prenzel, Rimmel et al., 2007) und aktuelle Lehr-Lern-Forschung (Fend, 2001; Fischer, Neumann et al., 2014; Harlen, 1999; Helmke, 2007; K. Neumann et al., 2012; Reusser & Pauli, 2003; Scheerens & Bosker, 2011; Seidel & Shavelson, 2007) ein Rahmenmodell zur Untersuchung des Naturwissenschaftsunterrichts (s. Abbildung 4).

Während ursprüngliche Angebot-Nutzungsmodelle sich vor allem durch die differenzielle Sichtweise und dadurch auszeichnen, dass sie durch den Einschluss zahlreicher Merkmale sehr umfassend sind (vgl. z. B. Fend, 2001; Helmke, 2007), verfolgt das Rahmenmodell von PISA 2006 die Absicht, den Naturwissenschaftsunterricht spezifisch und theoretisch kohärent zu beschreiben. Es bezieht Lehr-Lerntheorien mit ein und basiert daher auch auf prozessbezogenen Grundlagen. Als instruktionale Merkmale berücksichtigt das Rahmenmodell von PISA 2006 interaktives Lehren und Lernen, forschend-entdeckendes Lernen sowie naturwissenschaftliche Anwendungsbezüge als Merkmale effektiven Naturwissenschaftsunterrichts. Grundlegende Annahmen des Modells sind:

- Lehren und Lernen wird vom Hintergrund, den Bedürfnissen und individuellen Voraussetzungen der Schülerinnen und Schüler beeinflusst
- Unterricht wird durch die Orchestration von Unterrichtsaktivitäten charakterisiert
- Die Qualität der Unterrichtsaktivitäten bzw. Orchestration bestimmt die Möglich-

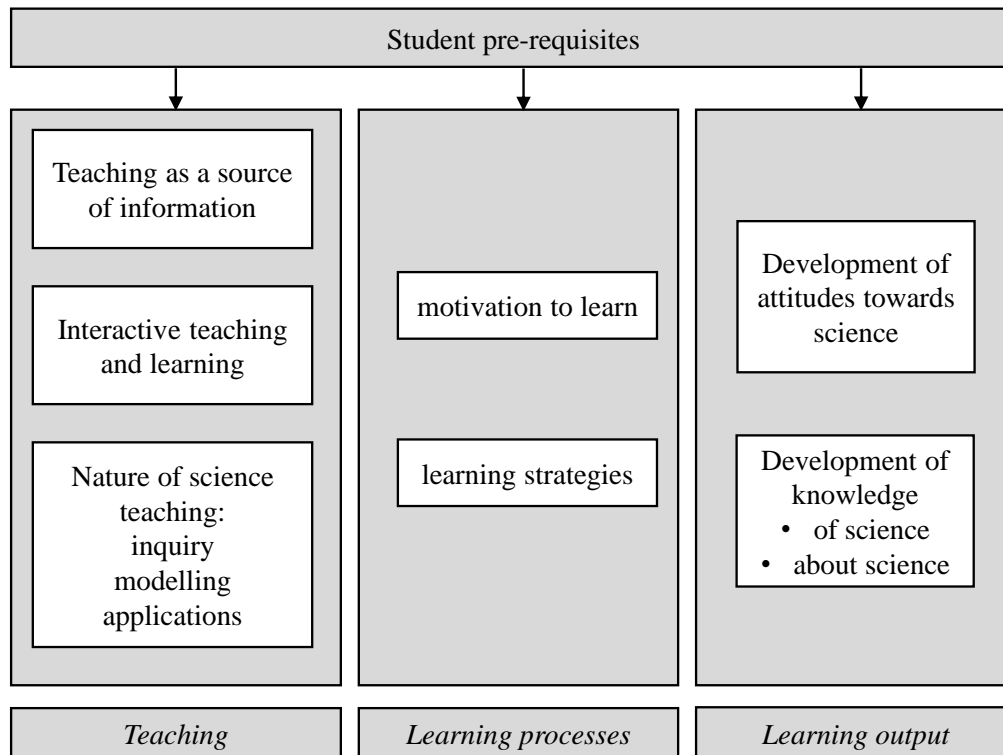


Abbildung 4. Rahmenmodell zur Untersuchung des Naturwissenschaftsunterrichts in PISA 2006 (Seidel & Prenzel, 2006b, S. 60)

keit zu aktivem Engagement der Schülerinnen und Schüler in Lernaktivitäten

- Die Nutzung des Unterrichtsangebots hängt von der individuellen Schülerwahrnehmung der unterstützenden Lernumgebungen ab, der Qualität der Schülermotivation und der Anwendung von Lernstrategien
- Lernergebnisse sind multipel und enthalten kognitive wie motivational-affektive Merkmale

4.2. Der Einfluss von Unterrichtsaktivitäten auf die naturwissenschaftliche Kompetenz und das Interesse leistungsstarker Jugendlicher in Naturwissenschaften

Das Rahmenmodell des Naturwissenschaftsunterrichts von Seidel und Prenzel (2006b) berücksichtigt sowohl die eingangs genannten zentralen domänenübergreifenden (Lernzeit, interaktives Lehren und Lernen) als auch naturwissenschaftsbezogenen Merkmale (forschend-entdeckendes Lernen, Anwendungsbezug) effektiven Naturwissenschaftsunterrichts. In den nachfolgenden Abschnitten wird das Rahmenmodell zur Untersuchung des Naturwissenschaftsunterrichts von Seidel und Prenzel (2006b) aufgegriffen, die Unterrichtsaktivitäten beschrieben und um eine differenzielle Sichtweise für hohe Leistung in Naturwissenschaften ergänzt. Das Kapitel gliedert sich in drei Abschnitte zu den drei Unterrichtsaktivitäten „interaktives Lehren und Lernen“, „forschend-entdeckender Unterricht“ und „Anwendungsbezüge“. Für jede Unterrichtsaktivität wird zunächst die Relevanz und der Forschungsstand zur Förderung leistungsstarker Jugendlicher in Naturwissenschaften getrennt für die beiden Lernergebnisse Leistung und Interesse dargelegt sowie Unterschiede zu Nicht-High-Performern diskutiert. Innerhalb dieser Abschnitte werden die nur zum Teil vorhandenen empirischen Ergebnisse zu leistungsstarken Jugendlichen durch lerntheoretische und motivationspsychologische Überlegungen ergänzt.

4.2.1. Interaktives Lehren und Lernen

Interaktives Lehren und Lernen beschreibt ein Merkmal von Unterricht, das durch Lehrer-Schüler- und Schüler-Schüler-Interaktionen sowie kollaboratives Arbeiten geprägt ist (Seidel & Prenzel, 2006b). In interaktiven Unterrichtsphasen wird Schülerinnen und Schülern ermöglicht, eigene Ideen zu äußern und zu erklären sowie selbstbestimmt und kooperativ zu lernen. Im Gegensatz dazu stehen transmissive Unterrichtsphasen, in denen Lerninhalte lehrerzentriert durch Lehrervorträge und Demonstrationsexperimente vermittelt werden. Interaktives Lehren und Lernen erlaubt durch die Interaktion von Lehrern und Jugendlichen bzw. Jugendlichen untereinander eine differenzielle Förderung, kooperativer Unterricht ermöglicht durch eine gezielte Einbindung von Schülerinnen und Schülern gegenseitiges und selbstgesteuertes Lernen. Metaanalysen und Studien zu kooperativem Lehren und Lernen belegen, dass interaktives Lehren und Lernen sowohl Lernprozesse (Scheerens & Bosker, 2011; Seidel & Shavelson, 2007) als auch motivational-affektive Merkmale wie Interesse (Lazarowitz, Baird, Bowlden & Hertz-Lazarowitz, 1996a; Seidel & Shavelson, 2007) fördert. Im Naturwissenschaftsunterricht wird interaktivem Lehren und Lernen deshalb eine hohe Bedeutung für Lernprozesse zugesprochen (Hofstein & Lunetta, 2004) und im Rahmen des Frameworks für Kompetenz in Naturwissenschaften in PISA 2006 (OECD, 2006) insbesondere mit der Teilkompetenz *naturwissenschaftliche Phänomene erklären* in Verbindung gebracht (Kobarg et al., 2011). Klassenunterricht mit Gelegenheiten für Diskussionen, der Raum für kognitive Aktivierung der Schülerinnen und Schüler lässt und Schülerinnen und Schüler aktiv einbindet, führt darüber hinaus zu höher wahrgenommenen Unterstützung motivationsrelevanter Bedürfnisse sowie zu höherem Interesse (Kobarg & Seidel, 2007; Seidel, 2003; Seidel & Prenzel, 2006a; Seidel et al., 2003, 2006).

Darüber hinaus zeigen Videostudien (Roth et al., 2006; Prenzel, 2001; Seidel et al., 2006), dass das Unterrichtsgespräch – eine von Lehrer-Schüler- und Schüler-Schüler-Interaktionen geprägte Instruktionsform, die unterschiedlich interaktiv gestaltet sein kann – im Naturwissenschafts- bzw. Physikunterricht in Deutschland sehr häufig vorkommt und einen Großteil der Unterrichtszeit einnimmt sowie vorwiegend lehrerzentriert gestaltet ist (Reyer, Trendel & Fischer, 2004; Roth et al., 2006; Seidel et al., 2003). Das im Unterrichtsgespräch in unterschiedlichen Ausprägungen stattfindende interaktive Lehren und Lernen prägt deshalb möglicherweise die Unterrichtspraxis von High Performern. Angesichts dessen stellt sich die Frage, inwiefern häufigeres interaktives

Lehren und Lernen im Zusammenhang mit höherer Leistung und höherem Interesse von High Performern steht und, ob Unterschiede zu Nicht-High Performern zu erwarten sind.

Forschungsstand zum Zusammenhang von Interaktivem Lehren und Lernen und naturwissenschaftlicher Kompetenz leistungsstarker Jugendlicher

Zwei Fragen leiten die Betrachtung des Zusammenhangs von interaktivem Lehren und Lernen und der naturwissenschaftlichen Kompetenz leistungsstarker Schülerinnen und Schüler: (a) Wie effektiv ist interaktives Lehren und Lernen für Lernprozesse leistungsstarker Jugendlicher? und (b) Unterscheidet sich der Zusammenhang zwischen leistungsstarken Jugendlichen und anderen Schülerinnen und Schülern? Für die grundsätzliche Wirksamkeit interaktiven Lernens für leistungsstarke Jugendliche liefern Überlegungen zum Lernprozess einen theoretischen Rahmen, während der differenzielle Aspekt der zweiten Frage theoretisch vor allem durch den Einfluss von Vorwissen und der Forschung zur Gruppenzusammensetzung beleuchtet werden kann. Nachfolgend werden sowohl die theoretischen Überlegungen zum Lernprozess als auch der Einfluss des Vorwissens und der Erkenntnisse zur Gruppenzusammensetzung beschrieben und im Anschluss der Forschungsstand zu leistungsstarken Jugendlichen dargelegt.

Seit geraumer Zeit beschreiben verschiedene Ansätze Lernen als einen konstruktiven und sozialen Prozess (Bransford, 2000; Donovan & Bransford, 2005; Piaget, 1974). Unter den verschiedenen Lernaktivitäten, mit denen Teilprozesse des Wissenserwerbs beschrieben werden, identifizierte Seidel (2003) Elaboration und Organisation von Wissensstrukturen als zentrale übergreifende Prozesse für den Aufbau und die Modifikation von Wissensstrukturen. Als grundlegende Aktivitäten in Wissenserwerbsprozessen können sie auch auf das Lernen leistungsstarker Schülerinnen und Schüler übertragen werden. Elaboration beschreibt die Verknüpfung neuer Informationen mit bestehendem Vorwissen, d. h. eine Anreicherung bzw. Erweiterung des Wissens, Organisation beschreibt die Organisation und Strukturierung von Informationen, d. h. eine Reduktion der Informationsmenge durch Selektion, Generalisation oder Ähnliches (vgl. Seidel, 2003). Interaktives Lehren und Lernen fördert den Elaborationsprozess durch verschiedene Komponenten. So unterstützt Konversation die kognitive Elaboration (O'Donnell, 2006) und führt zu tieferem Verständnis (N. M. Webb, 2009). Im Naturwissenschaftsunterricht ist verbale Interaktion darüber hinaus wichtig dafür, eine angemessene Fachsprache zu entwickeln und die Rolle des Diskurs für die naturwissenschaftliche Erkenntnisgewin-

nung zu erkennen (Bransford & Donovan, 2005). Darüber hinaus begünstigen Elemente des interaktiven Unterrichts wie Argumentation (z. B. Berland & McNeill, 2010; Osborne, 2012), kritischer Diskurs (z. B. Henderson, MacPherson, Osborne & Wild, 2015; Osborne, 2010), Reden (z. B. Haug & Ødegaard, 2014) sowie Schülerfragen (z. B. Andersson-Bakken & Klette, 2016; Chin & Osborne, 2008) Elaborationsprozesse im Naturwissenschaftsunterricht (vgl. Müller, Prenzel, Seidel, Schiepe-Tiska & Kjærnsli, 2016).

Vergleicht man diesen Zusammenhang zwischen leistungsstarken und weniger leistungsstarken Schülerinnen und Schülern, so deuten verschiedene theoretische Ansätze Unterschiede an. Bereits Piaget, Aebli und Seiler (1996) und vor allem Aebli (1980, 1981) in der didaktischen Weiterentwicklung von Piagets Annahmen zum Wissensaufbau haben die entscheidende Rolle des Vorwissens betont. Sie beschreiben Wissenserwerbsprozesse als Erweiterung und Anpassung kognitiver Schemata, die wesentlich durch die Akkommodation (Anpassung oder Neuerschaffung von Schemata) und Assimilation (Zuordnung zu bzw. Erweiterung von vorhandenen Schemata) von Informationen bestimmt werden. Das Vorwissen von Schülerinnen und Schülern entspricht in diesem Verständnis der Ontogenese von Wissensstrukturen den zu einem gewissen Zeitpunkt bereits vorhandenen kognitive Strukturen und beeinflusst daher die Wahrnehmung und das Verständnis von Ereignissen (vgl. auch Seidel, 2003). Tiefe Verarbeitung von Lerninhalten äußert sich dann auch darin, dass flexibel zwischen verschiedenen Repräsentationsformen gewechselt werden kann (Aebli, 2006).

Empirische Befunde bestätigen dies, indem sie zeigen, dass Schülerinnen und Schülern mit höherem Vorwissen Elaborationsprozesse beim Lernen von Texten, und damit vermutlich auch die Verarbeitung von Informationen aus gesprochener Sprache, leichter fallen als Schülerinnen und Schülern mit niedrigerem Vorwissen (Lehtinen, 1992; Seidel, 2003). Darüber hinaus verdeutlicht Forschung zu Lehrer-Schüler-Interaktion, dass sich Schülerinnen und Schüler mit höherer Leistung bzw. allgemein besseren Lernvoraussetzungen in interaktiven Unterrichtsumgebungen häufiger verbal im Unterricht engagieren und insbesondere häufiger Antworten geben (Jurik et al., 2013). Ein erhöhtes Vorwissen führt daher – vermittelt über erleichterte Elaborationsprozesse und Engagement – im interaktiven Unterricht möglicherweise zu höheren Lernzuwächsen leistungsstarker Schülerinnen und Schüler im Vergleich zu leistungsschwächeren Schülerinnen und Schülern.

Eng verbunden mit Überlegungen zur Wirksamkeit von kooperativem Lernen für

leistungsstarke Jugendliche und Unterschieden zu leistungsschwächeren Schülerinnen und Schülern, ist die Frage danach, ob leistungsstarke Schülerinnen und Schüler eher von homogenen oder heterogenen Klassen bzw. Lerngruppen profitieren. Aus theoretischer Sicht plädiert sowohl Forschung zu leistungsstarken Jugendlichen (z. B. Dar & Resh, 1986; Johnson, Johnson & Taylor, 1993) als auch allgemeine Forschung zur Wirksamkeit von interaktivem Lernen (z. B. N. M. Webb, 1992) für heterogene Gruppen. Wesentlicher Grund dafür ist die Annahme, dass heterogene Gruppen leistungsschwachen Schülerinnen und Schülern Hilfe durch leistungsstärkere Schülerinnen und Schüler bieten, während leistungsstarke Schülerinnen und Schülern durch die Erklärung von Konzepten und Diskussion unterschiedlicher inhaltlicher Sichtweisen in ihrer Elaboration und ihren kognitiven Lernprozessen unterstützt werden (Johnson & Johnson, 1989; Slavin, 1987, 1991). Im Naturwissenschaftsunterricht haben unterschiedliche und konträre Ansichten eine besondere Bedeutung, da Schülerinnen und Schüler aus dem Alltag häufig Fehlvorstellungen mitbringen und verbale Interaktionen durch Fragen und widersprechende Ansichten die Reflexion und Korrektur von Fehlvorstellungen anstoßen und befördern (Leman, Skipper, Watling & Rutland, 2016; Lonning, 1993). Manche Studien beobachten darüber hinaus, dass sich leistungsstärkere Schülerinnen und Schüler mehr in der Auflösung dieser konträrer Aussagen engagieren und mehr von kognitiven Konflikten profitieren als leistungsschwächere (vgl. Zohar & Aharon-Kravetsky, 2005). Dies wird unter anderem damit erklärt, dass die Entstehung und Lösung eines bedeutsamen kognitiver Konflikt eine gewisse Menge an Vorwissen und Schlussfolgerungsfähigkeit benötigt (Limón, 2001).

Reviews von Slavin (1990a, 1990b) fassen den Forschungsstand zu Leistungsgruppierung und kooperativem Lernen bei leistungsstarken Schülerinnen und Schülern zusammen. Es gilt als relativ gesichertes Ergebnis (Slavin, 1990a), dass Leistungsgruppierung alleine keine positive Auswirkung auf leistungsstarke Schülerinnen und Schüler hat (Slavin, 1987, 1990b; Kulik & Kulik, 1987). Sowohl leistungsstarke als auch durchschnittliche oder leistungsschwache Schülerinnen und Schüler profitieren in ihrer Leistungsentwicklung dagegen von kooperativen im Vergleich zu traditionellen Lernumgebungen (Slavin, 1990a, 1990b). Eine Metaanalyse, die getrennt für leistungsstarke und (hoch-)begabte Schülerinnen und Schüler kooperatives und traditionelles Lernen kontrastiert, findet ebenfalls mit einer mittleren Effektstärke von $\Delta = .26$ ($n = 302$), dass leistungsstarke Jugendliche in heterogenen kooperativen Gruppen besser lernen als bei individuellem

Lernen (Neber, Finsterwald & Urban, 2001, S. 206). Unter anderem konnten Stevens und Slavin (1995) explizit nachweisen, dass kooperatives Lernen in heterogenen Gruppen bei leistungsstarken Jugendlichen zu positiveren Leistungsentwicklungen führt als ein Enrichment-Programm, in dem leistungsstarke Jugendliche durch eine Anreicherung der Inhalte in homogenen kooperativen Lerngruppen gefördert werden (vgl. Neber et al., 2001). Auch in der Metaanalyse von Hattie (2010) zeigte sich, dass kooperatives Lernen traditionellem Lernen und individuellem Lernen überlegen ist und sowohl leistungsstarke als auch durchschnittliche und leistungsschwache Schülerinnen und Schüler vom Lernen in kleinen Gruppen profitieren.

Kooperative Lernumgebungen sind vor allem dann für leistungsstarke Jugendliche wirksam, wenn in ihnen „echtes“ kooperatives Lernen stattfindet, d. h. wenn die Ziele als Gruppenziele modelliert sind und gleichzeitig eine individuelle Verantwortlichkeit sichergestellt wird (Slavin, 1990a, 1990c). Dies wird in Lernumgebungen beispielsweise erreicht, indem die Beurteilung der Schülerinnen und Schüler am durchschnittlichen Leistungszuwachs der Gruppenmitglieder erfolgt. Leistungsstarke Jugendliche ermöglicht das Erklären der Inhalte für leistungsschwächere Schülerinnen und Schüler, tieferes Verständnis für die Inhalte zu entwickeln (N. Webb, 1985) und sie profitieren in motivational-affektiven Merkmalen wie ihrem Selbstbewusstsein oder ihrer Einstellung zu Fächern sowie ihrer Fähigkeit, in kooperativen Gruppen zu arbeiten (Slavin, 1990c, 2010).

Insgesamt ist aus theoretischer Sicht anzunehmen, dass leistungsstarke Schülerinnen und Schüler von kooperativen Lernumgebungen und auch von heterogenen Lerngruppen profitieren. Auch empirisch konnte dieser Zusammenhang weitgehend belegt werden (Slavin, 1990a, 1990b), wenngleich die Studien zur Wirksamkeit kooperativer Lerngruppen eine große Varianz in den Studiendesigns und behandelten Fragestellungen, z. B. in der Rolle homogener bzw. heterogener Gruppen, aufweisen und zum Teil nur geringe Stichprobenzahlen erreichen (Neber et al., 2001).

Forschungsstand zum Zusammenhang von interaktivem Lehren und Lernen und Interesse von High Performern

Ausgehend von Annahmen der pädagogischen Interessentheorie (Prenzel, 1988, 1992; H. Schiefele et al., 1983) bzw. Person-Gegenstand-Theorie von Interesse (Krapp, 2002a, 2002b) sowie der Selbstbestimmungstheorie (Deci & Ryan, 1985, 2000, 2002) werden

Interesse bzw. intrinsische Lernmotivation – unabhängig vom Leistungsniveau – durch die Erfüllung des Bedürfnisses nach Selbstbestimmung bzw. motivationsrelevanter Lehr-Lernbedingungen erfüllt (vgl. Kapitel 3.4). Interaktives Lehren und Lernen begünstigt je nach Gestaltung die Erfüllung dieser Lehr-Lernbedingungen und führt dadurch mediiert zu höherer intrinsischer Motivation (z. B. Hänze & Berger, 2007; Lazarowitz et al., 1996a). So zeigt sich, dass Unterricht mit qualitativ hochwertigen Diskussionen, d. h. interaktiver Unterricht, der autonomieunterstützend und inhaltlich relevant gestaltet ist, (Seidel, 2003) zu höherer intrinsischer Motivation führt wohingegen eng geführte Unterrichtsgespräche mit reproduktiven Fragen und ohne die Möglichkeit für Schülerinnen und Schüler, eigene Ideen zu äußern, intrinsische Motivation und Interesse verringern (Krapp & Prenzel, 2011; Seidel et al., 2003; Seidel, Prenzel, Rimmel et al., 2007). Auch in PISA 2006 zeigt sich die Bedeutung von motivationsunterstützend gestaltetem interaktivem Unterricht, indem Unterricht mit Fokus auf das Erklären eigener Ideen, Ziehen von Schlussfolgerungen aus Experimenten und Bezug zum Alltag im Vergleich zu anderen Unterrichtsmustern mit der höchsten Leistung und dem höchsten Interesse in Naturwissenschaften assoziiert ist (Kobarg et al., 2011). Aus motivationstheoretischer Sicht unterstützt interaktives Lehren und Lernen daher grundsätzlich sowohl das Interesse von High Performern als auch Nicht-High-Performern.

Empirisch wurde der Zusammenhang von interaktivem Lehren und Lernen und Interesse oder andere motivational-affektiven Merkmalen leistungsstarker Jugendlicher nur vereinzelt und mit unterschiedlichen motivational-affektiven Konstrukten, untersucht. In einer Betrachtung allgemein leistungsstarker Jugendlicher in der fünften Jahrgangsstufe finden Johnson et al. (1993), dass kooperatives Lernen im Vergleich zu individuellem Lernen zu höherer Leistung und positiverem akademischen Selbstkonzept führt. Ähnlich ergab eine experimentelle Studie mit fächerübergreifend leistungsstarken taiwanesischen Schülerinnen und Schüler (Hong, 2010), dass eine 20-wöchige kollaborative Intervention in Naturwissenschaften zu einer Verbesserung der Einstellung führt. Die Autoren berichten zudem, dass qualitative Interviews andeuten, dass sich das Interesse gegenüber Naturwissenschaften durch die Intervention verbessert.

Zweifel daran, dass interaktives Lehren und Lernen leistungsstarke Schülerinnen und Schüler ebenso wie leistungsschwächere Schülerinnen und Schüler motiviert, äußert die Hochbegabungsforschung in der Beschreibung motivationaler Effekte, die entweder leistungsstarke oder leistungsschwächere Schülerinnen und Schülern in heterogenen

Gruppen bevorteilen (vgl. Neber et al., 2001). Zum wird mit dem *free-rider effect* die Möglichkeit beschrieben, dass leistungsstarke Jugendliche in kooperativen Lernumgebungen die Hauptarbeit übernehmen während leistungsschwächere Schülerinnen und Schüler –freiwillig oder unfreiwillig – eine passive Rolle übernehmen (vgl. Neber et al., 2001; Robinson, 2003). In diesem Fall engagieren sich vor allem leistungsstarke Schülerinnen und Schüler, werden kognitiv aktiviert und profitieren von der Lerngelegenheit. Als umgekehrte Situation wird als *sucker-effect* bezeichnet, dass leistungsstarke Schülerinnen und Schülern in heterogenen Gruppen den Eindruck gewinnen können, sie werden als Lernunterstützung von schwächeren Schülerinnen und Schülern ausgenutzt und als Resultat geringere Motivation und geringeres Engagement zeigen (z. B. Neber et al., 2001; Robinson, 1990). Grundsätzlich sind diese Effekte jedoch durch eine entsprechende Gestaltung der Unterrichtsmethoden instruktional reduzierbar bzw. vermeidbar. Wie oben bereits erwähnt, kann dies beispielsweise durch die Beurteilung am durchschnittlichen Leistungszuwachs der Gruppe sichergestellt werden (Slavin, 1990c, 2010)

Insgesamt bleibt festzuhalten, dass interaktives Lehren und Lernen Potential zur Förderung von sowohl Leistung als auch Interesse von High Performern bietet. Vergleicht man den Zusammenhang von Leistung bzw. Interesse mit interaktivem Lehren und Lernen zwischen High Performern und Nicht-High Performern, so gibt es einerseits Hinweise darauf, dass High Performer in interaktiven Lernumgebungen aufgrund ihres höheren Vorwissens stärker als leistungsschwächere Schülerinnen und Schüler von interaktivem Lernen profitieren können. Trotz relativ umfangreicher Forschungslage unterscheiden sich die Studien zum Teil stark und enthalten überwiegend kleine Stichproben. Es ist darüber hinaus unklar, ob sich ein möglicher Zusammenhang zwischen High und Nicht-High Performern unterscheidet.

4.2.2. Forschend-entdeckendes Lernen

Forschend-entdeckendes Lernen im Naturwissenschaftsunterricht (engl. *inquiry-based learning*) beschreibt naturwissenschaftsbezogene schüleraktive Instruktionsformen wie Schülerexperimente und forschende Elemente (Duit & Treagust, 1998). Mit Hilfe des forschend-entdeckenden Lernens wird im Naturwissenschaftsunterricht das Ziel verfolgt, neben naturwissenschaftlichem Wissen auch Wissen über die Naturwissenschaften im Sinne naturwissenschaftlicher Denk- und Arbeitsweisen zu vermitteln, denn es ermöglicht Schülerinnen und Schülern naturwissenschaftliche Fragen zu identifizieren, Untersuchungen durchzuführen und Daten zu interpretieren (Hofstein & Lunetta, 2004; Seidel & Prenzel, 2006b).

Schülerexperimente und forschende Elemente fördern die Entwicklung naturwissenschaftlicher Kompetenz, führen zu tieferem Verständnis (Minner, Levy & Century, 2010; Scheerens, Seidel, Witziers, Hendriks & Doornekamp, 2005; Yoon, 2009) und unterstützen das Interesse an den Naturwissenschaften und weitere motivational-affektive Merkmale (George & Kaplan, 1998; Minner et al., 2010; Scheerens et al., 2005). In einer Metaanalyse zu Merkmalen effektiven Unterrichts belegen Seidel und Shavelson (2007), dass domänenspezifische Merkmale wie forschend-entdeckendes Lernen die höchste Effektstärke in Bezug auf Leistung sowie motivational-affektive Merkmale aufweisen. Im Durchschnitt erreicht forschend-entdeckender Unterricht in Metaanalysen eine Effektstärke von $d = 0.35$ (Hattie, 2010, S. 209). Eine Metaanalyse zum Naturwissenschaftsunterricht (Schroeder, Scott, Tolson, Huang & Lee, 2007) in den USA bestätigt dies, indem sie für forschend-entdeckendes Lernen den höchsten Einfluss auf die Leistung fand, gefolgt von Kontextstrategien, kollaborativem Lernen und Fragestrategien. Furtak et al. (2012, S. 315) verdeutlichen in einer aktuellen Metaanalyse zu (quasi-)experimentellen Studien eine mittlere Effektstärke von $d = .50$ für forschend-entdeckenden Unterricht im Zusammenhang mit Lernerfolg und zeigen, dass rein schülerzentrierte Lernumgebungen den Lernumgebungen, die auch lehrerzentrierte Elemente einschließen, unterlegen sind. Large-Scale-Assessments finden darüber hinaus, dass der Unterricht in leistungsstarken Ländern in Naturwissenschaften häufiges Ideen erklären und Schlüsse ziehen (Taylor, Stuhlsatz & Bybee, 2009) sowie praktisches Arbeiten (Lavonen & Laaksonen, 2009) enthält und sich Länder mit hoher zukunftsorientierter Motivation der Jugendlichen häufige Schülerexperimente berichten (Kjærnsli & Lie, 2011). Darüber hinaus konnte in PISA 2006 gezeigt werden, dass Naturwissenschaftsunterricht mit gezielt eingesetzten

Experimenten, in dem Schülerideen häufig eingebunden sowie regelmäßig Schlussfolgerungen aus Experimenten gezogen werden, mit der höchsten naturwissenschaftlichen Kompetenz sowie dem höchsten Interesse assoziiert ist (Kobarg et al., 2011).

Nach wie vor kritisieren jedoch Unterrichtsforscher Ansätze des forschend-entdeckenden Lernens (Kirschner, Sweller & Clark, 2006; Mayer, 2004) und Studien finden gemischte Ergebnisse (z. B. Klahr & Nigam, 2004; Lazarowitz, Baird, Bowlden & Hertz-Lazarowitz, 1996b). Dies kann einerseits dadurch erklärt werden, dass forschend-entdeckendes Lernen in der bisherigen Forschung nicht einheitlich definiert wurde und in verschiedenen Konzeptionen – sowohl als Wissensart als auch als Instruktionsart, Lernart und Curriculuminhalt – verwendet wurde (Furtak et al., 2012; Furtak & Kunter, 2012). Darüber hinaus wird aus bestehender Forschung deutlich, dass die Einbettung und Gestaltung forschend-entdeckender Lernelemente (Tesch & Duit, 2004) sowie die Orchestration mit anderen Unterrichtsaktivitäten (Kobarg et al., 2011) über ihre Wirksamkeit bestimmt. Folgende Prinzipien werden als Schlüsselfaktoren für eine wirksame Gestaltung von forschend-entdeckendem Unterricht genannt (Harlen, 1999; Hofstein & Lunetta, 2004; Furtak et al., 2012; Kobarg et al., 2011; Ostermeier, Prenzel & Duit, 2010; Seidel & Prenzel, 2006b; Seidel et al., 2006; Seidel, Prenzel, Rimmele et al., 2007; Tesch & Duit, 2004):

- Zielorientierung und Einbettung des Experiments in den Unterricht (gezielter und lernzielorientierter Einsatz, Transparenz und Kommunikation von Zielen, Orchestration mit anderen Unterrichtsaktivitäten)
- Lernbegleitung (z. B. Anregung von Denkprozessen, Strukturhilfen, Unterstützung, Rückmeldung)
- Berücksichtigung sowohl prozeduraler (z. B. Umgang mit Messgeräten) als auch kognitiver (z. B. tiefes Verständnis) und epistemologischer Ziele (z. B. wissenschaftliches Denken und Arbeiten)
- Möglichkeit für Schülerinnen und Schüler, eigene Fragestellungen zu generieren und eigene Experimente zu planen und zu interpretieren (im Gegensatz zu bloßer Durchführung vorgegebener Experimente nach Anleitung)
- Möglichkeit zu selbstreguliertem und autonomen Lernen sowie dem Erleben von Kompetenz

- Kontextbasierung und inhaltlich relevante Gestaltung (z. B. Bezug zur Lebenswelt der Schülerinnen und Schüler)

Forschung zu forschend-entdeckendem Lernen steht deshalb weniger vor der Frage nach dem Einsatz, sondern vielmehr vor der Frage nach Bedingungen und Kontexten, die zur Wirksamkeit von forschend-entdeckendem Unterricht für einzelne Lernergebnisse führen (Hmelo-Silver, Duncan & Chinn, 2007). Eine zentrale Fragestellung in diesem Zusammenhang ist, ob forschend-entdeckender Unterricht für unterschiedliche Populationen, beispielsweise für leistungsstarke und weniger leistungsstarke Schülerinnen und Schüler, unterschiedlich wirkt (Hmelo-Silver et al., 2007). Der dazu bisher bekannte Forschungsstand zum Zusammenhang von forschend-entdeckendem Lernen und Leistung sowie Interesse wird nachfolgend beschrieben und sowie mögliche Unterschiede zwischen High und Nicht-High-Performern diskutiert.

Forschungsstand zum Zusammenhang von forschend-entdeckendem Lernen und Leistung von High Performern

Der Zusammenhang von forschend-entdeckendem Lernen und Leistung in Naturwissenschaften wurde für High Performer sowie im Hinblick auf mögliche Unterschiede zu Nicht-High-Performern bisher nur eingeschränkt untersucht. Im internationalen Vergleich auf Basis von PISA 2006 berichten Top Performer in Naturwissenschaften, d. h. Schülerinnen und Schüler auf den beiden obersten Kompetenzstufen, zunächst weniger forschend-entdeckenden Unterricht als leistungsschwächere Schülerinnen und Schüler (OECD, 2009b).

Für den Zusammenhang von forschend-entdeckendem Unterricht mit Leistung in Naturwissenschaften innerhalb der Gruppe der leistungsstarken Schülerinnen und Schüler belegen Studien, dass forschend-entdeckender Unterricht die naturwissenschaftliche Kompetenz leistungsstarker Schülerinnen und Schüler fördert (D. Schwartz, Chase, Oppezzo & Chin, 2011; J.-R. Wang, Wang, Tai & Chen, 2010; White & Frederiksen, 1998). Auch in der Hochbegabtenförderung werden offene Unterrichtsformen und schülerzentrierte naturwissenschaftliche Untersuchungen an lebensnahen Problemstellungen gefordert (Heller, 2002c; Neber & Schommer-Aikins, 2002) und finden sich beispielsweise im *Triad/Revolving Door System* von Renzulli (1976, 1984) explizit als anvisierte Enrichment-Aktivität (*Type-III-Enrichment-Activity*).

Betrachtet man Eigenschaften und Gestaltungskriterien forschend-entdeckender Lernumgebungen prozessbezogen im Zusammenhang mit Erkenntnissen zu hohem Vorwissen, ergeben sich verschiedene theoretische Hinweise auf Unterschiede des Zusammenhangs von forschend-entdeckendem Unterricht und naturwissenschaftlicher Kompetenz zwischen leistungsstarken und leistungsschwächeren Jugendlichen. Nachfolgend werden drei Aspekte näher erläutert, die auf einen Vorteil leistungsstarker gegenüber leistungsschwachen Schülerinnen und Schülern beim forschend-entdeckenden Lernen hinweisen: (a) allgemeine Vorteile im Lernprozess durch ein hohes Vorwissen, (b) Vorteile im Umgang mit der inhaltlichen Offenheit forschend-entdeckenden Lernens und (c) Vorteile im Umgang mit der strukturellen Offenheit durch Selbstregulation.

Ein wichtiges Gestaltungskriterium erfolgreicher forschend-entdeckender Lernumgebungen ist die Möglichkeit zum schüleraktiven, konstruktiven Lernen (z. B. Euler, 2001). Erfolgreiches forschend-entdeckendes Lernen erfordert kognitive Aktivität und Beteiligung der Schülerinnen und Schüler (Minner et al., 2010; Taraban, Box, Myers, Pollard & Bowen, 2007). Dabei erklärt weniger das Ausmaß an Schülerzentrierung und praktischer Tätigkeit (*hands-on*), sondern vielmehr das Ausmaß an domänenspezifischen kognitiv aktivierenden und konzeptuell anregenden Lernaktivitäten (*minds-on*) die Unterstützung von Lernprozessen (Prenzel et al., 2009; Seidel & Prenzel, 2006a). Dies erfordert das Anknüpfen an und die Erweiterung und Neusortierung von bestehenden Wissensstrukturen. Ein hohes Vorwissen in Naturwissenschaften erleichtert Elaborationsprozesse (vgl. Kapitel 4.2.1) und begünstigt so den konstruktivistischen Lernprozess. Es ist deshalb davon auszugehen, dass leistungsstarken Jugendlichen aufgrund ihres Vorwissens die Wissensgenerierung im forschend-entdeckenden Lernprozess leichter fällt.

Neben dem konstruktiven schüleraktiven Lernen, gehört die Möglichkeit zum selbstbestimmten, kontextbezogenen und problembasierten Lernen zu den zentralen Gestaltungsmerkmalen forschend-entdeckenden Lernens (Engeln & Euler, 2005; Tesch & Duit, 2004). Forschend-entdeckendes Lernen bietet dadurch eine inhaltliche Offenheit, die den Schülerinnen und Schülern einerseits ermöglicht, eigene Ideen einzubringen und eigene Problemlösungen zu finden, und andererseits auch das Lerntempo selbst zu bestimmen. Gleichzeitig führen Kontext und problembasierte Gestaltung unter Umständen zu einer erhöhten Schwierigkeit, da große Mengen an Informationen dekodiert und verarbeitet sowie Problemlösestrategien angewandt werden müssen. Ein höheres Vorwissen

erleichtert die Verarbeitung größerer Informationsmengen, da es die Organisation von Wissen in größeren Bündeln, sogenannten *chunks* erlaubt (Chase & Simon, 1973; Gruber & Stamouli, 2015). Es ist daher wahrscheinlich, dass leistungsstarke Jugendliche Kontext und Informationsmenge in forschend-entdeckenden Lernumgebungen leichter dekodieren als leistungsschwächere Jugendliche. Klahr und Dunbar (1988) beschreiben forschend-entdeckendes Lernen als Suche im Hypothesen- sowie im Experimentraum. Gleichzeitig belegt Nehring, Nowak, zu Belzen und Tiemann (2015), dass die Fähigkeit zum naturwissenschaftlichen Argumentieren größtenteils durch Vorwissen zu naturwissenschaftlichen Konzepten sowie allgemeinen kognitiven Merkmalen vorhergesagt wird. Aufgrund ihres höheren Vorwissens und der höheren methodischen Fähigkeiten beim naturwissenschaftlichen Denken und Arbeiten, ist davon auszugehen, dass leistungsstarke Jugendliche sowohl über einen größeren Hypothesen- als auch Experimentraum verfügen und so sowohl inhaltlich als auch methodisch Vorteile bei der Durchführung naturwissenschaftlicher Untersuchungen haben. Im Hinblick auf die inhaltliche Offenheit und mögliche höhere Schwierigkeit ist es daher plausibel, dass leistungsstarken Jugendlichen Lernprozesse leichter fallen. Andererseits zeigt Forschung zu Vorwissen in Mathematik jedoch, dass auch leistungsschwächere Schülerinnen und Schüler von anspruchsvollen Lernaufgaben profitieren (Renkl & Stern, 1994; Renkl, 1996).

Schließlich zeichnet sich effektiv gestalteter forschend-entdeckender Unterricht neben einer inhaltlichen Offenheit auch durch eine strukturelle Offenheit aus, indem er von den Schülerinnen und Schülern selbstreguliertes und selbständiges Arbeiten fordert. Die Forschung der letzten Jahrzehnte hat gezeigt, dass effektives forschend-entdeckendes Lernen ein Gleichgewicht aus kognitiver Aktivierung, inhaltlicher Offenheit bzw. Selbstbestimmung und passender Lernbegleitung, Strukturierung und Unterstützung durch die Lehrkraft erfüllen muss (Alfieri, Brooks, Aldrich & Tenenbaum, 2011; Blanchard et al., 2010; Harlen, 1999; Lazonder & Harmsen, 2016; Schütte & Köller, 2015; Seidel, Prenzel, Rimmele et al., 2007). Im in Deutschland anzutreffenden forschend-entdeckenden Naturwissenschaftsunterricht wird jedoch häufig eine eingeschränkte Lernbegleitung und Strukturierung beobachtet (vgl. Seidel, Prenzel, Rimmele et al., 2007; Tesch & Duit, 2004; Tesch, 2005). Im Umgang mit der strukturellen Offenheit, insbesondere bei fehlender Strukturierung und Begleitung durch die Lehrkraft, haben leistungsstarke Jugendliche möglicherweise einen Vorteil, denn sie weisen oftmals (für eine Ausnahme siehe Neber & Schommer-Aikins, 2002) höhere Selbstregulationsstrategien

auf als leistungsschwache Schülerinnen und Schüler (Chan, 1996; Schunk & Swartz, 1993; Zimmerman & Martinez-Pons, 1990). Ähnlich argumentieren Vertreter des Aptitude-Treatment-Interaction-Modells, dass weniger leistungsstarke Schülerinnen und Schüler eher von stärker strukturierten Lernumgebungen profitieren, während offene Lernumgebungen sich für leistungsstarke Schülerinnen und Schüler eignen (Cronbach & Snow, 1977; Corno & Snow, 1986). Analog empfiehlt die Hochbegabungsforschung sowie Forschung zu Vorwissen, dass mit höherer Leistung ein höheres Maß an selbstregulativen, kontextbezogenen und forschend-entdeckenden Lernumgebungen einhergeht (Heller, 2002b; Renkl, 1996; Robinson, Shore & Enersen, 2006a; VanTassel-Baska, Bass, Ries, Poland & Avery, 1998). Insgesamt wird aus theoretischer Sicht vielfach angenommen, dass leistungsstarke Jugendliche stärker von forschend-entdeckenden Lernumgebungen profitieren als leistungsschwache Schülerinnen und Schüler.

Der empirische Vergleich des Zusammenhangs von forschend-entdeckendem Unterricht und Leistung zwischen leistungsstarken und weniger leistungsstarken Jugendlichen in Naturwissenschaft liefert gemischte Ergebnisse. D. Schwartz et al. (2011) verglichen Unterricht, in dem Schülerinnen und Schüler naturwissenschaftliche Formeln forschend-entdeckend herausfinden, indem sie Experimente zu kontrastierenden Fällen durchführen, mit Unterricht, in dem Experimente zu bereits bekannten Formeln durchgeführt werden. Sowohl leistungsstarke als auch leistungsschwache Schülerinnen und Schüler profitierten von der forschend-entdeckenden Unterrichtsumgebung stärker als von der Vergleichsumgebung und unterschieden sich nicht in der Höhe des Zusammenhangs.

Im Gegensatz dazu fanden White und Frederiksen (1998), dass die Leistung von sowohl leistungsstarken als auch leistungsschwachen Schülerinnen und Schüler von forschend-entdeckendem Unterricht gefördert wird, leistungsschwache Schülerinnen und Schüler jedoch in ihrem Leistungszuwachs stärker davon profitieren als andere Jugendliche. Ähnlich beobachteten J.-R. Wang et al. (2010) in einer enger begleiteten forschend-entdeckenden Lernumgebung, dass Schülerinnen und Schüler mit niedrigem naturwissenschaftlichem Vorwissen (untersten 27%) im Vergleich zu Schülerinnen und Schülern mit mittlerem und hohem Vorwissen (oberste 27%) stärker gefördert wurden. Deckeneffekte, die auf die Gestaltung und Leistungsmessung zurückzuführen sind, werden jedoch als mögliche Einschränkung aufgeführt (J.-R. Wang et al., 2010) und führten möglicherweise dazu, dass das Potential forschend-entdeckender Lernumgebungen für

leistungsstarke Schülerinnen und Schüler nicht ausgeschöpft wurde und dementsprechend nur eingeschränkte Effekte beobachtet werden konnten. Die genannten Studien unterscheiden sich zudem sowohl in der Gestaltung der forschend-entdeckenden Lernumgebungen als auch in den untersuchten Altersbereichen, der Operationalisierung von Vorwissen bzw. Vorleistung sowie den Lernergebnissen. Eindeutige Schlussfolgerungen zu möglichen Unterschieden zwischen leistungsstarken und leistungsschwachen Schülerinnen und Schülern können daher aus bestehender Forschung nicht abgeleitet werden.

Forschungsstand zum Zusammenhang von forschend-entdeckendem Lernen und Interesse von High Performern

Der Zusammenhang von forschend-entdeckendem Lernen und Interesse wurde für leistungsstarke Jugendliche in Naturwissenschaften bisher nicht im Detail untersucht. Studien zu forschend-entdeckendem Unterricht im Zusammenhang mit High Performance in Naturwissenschaft und Unterschieden zu Nicht-High-Performern beschäftigen sich überwiegend mit Leistung (D. Schwartz et al., 2011; White & Frederiksen, 1998). Indirekte Hinweise liefern jedoch Studien, die sich mit Vorwissen bzw. Leistung und Interesse bzw. anderen motivational-affektiven Merkmalen wie Selbstkonzept oder zukunftsbezogener Motivation, die im Zusammenhang mit Interesse stehen (Trautwein, Lüdtke, Marsh et al., 2006), beschäftigen.

Eine vertiefte Analyse leistungsstarker Jugendlicher in Naturwissenschaften in PISA 2006 (OECD, 2007) verdeutlichte, dass forschend-entdeckender Unterricht anders als erwartet nicht das Interesse leistungsstarker Jugendlicher vorhersagt (Prenzel & Schütte, 2007). Taskinen et al. (2013) untersuchten den Zusammenhang verschiedener Unterrichtsaktivitäten mit der zukunftsorientierten Motivation leistungsstarker Schülerinnen und Schüler in Naturwissenschaften im Rahmen von PISA 2006. Sie fanden für forschend-entdeckendes Lernen keinen Zusammenhang. Ähnlich beobachteten Stake und Mares (2001) in einem Enrichmentprogramm im Schülerlabor keine signifikante Veränderung beim Selbstkonzept und der Karrieremotivation leistungsstarker Jugendlicher.

Entgegengesetzte Hinweise liefern jedoch Studien, die zeigen, dass forschend-entdeckender Unterricht als motivationsunterstützend erlebt wird. Die Erfüllung motivationsunterstützender Lehr-Lernbedingungen im Sinne der Selbstbestimmungstheorie (Deci & Ryan, 1985, 2000; Ryan & Deci, 2002) und pädagogischen Interessentheorie

(Prenzel, 1992, 1997; H. Schiefele et al., 1983) unterstützt aus motivationstheoretischer Sicht die interessierte und intrinsische Lernmotivation Jugendlicher aller Leistungsniveaus. Forschend-entdeckender Unterricht fördert die Erfüllung dieser motivationsunterstützender Bedingungen: Zum Einen werden offene selbstbestimmte forschende Unterrichtsprozesse als autonom erlebt und Schülerinnen und Schüler berichten mehr Interesse und Motivation (Berg, Bergendahl, Lundberg & Tibell, 2003; Chin & Chia, 2006; Ornstein, 2006). Zum Anderen fördern relativ offene Schülerexperimente die Kooperation der Schülerinnen und Schüler untereinander, ohne negative Auswirkungen auf die Autonomie zu haben (Sadeh & Zion, 2012), und schaffen so die Bedingungen dafür, dass sich Schülerinnen und Schüler sozial eingebunden wahrnehmen (vgl. Deci & Ryan, 1993). Darüber hinaus haben Schülerinnen und Schüler in Schülerexperimenten aufgrund der schüleraktiven Gestaltung die Möglichkeit, sich als wirksam und kompetent zu erleben und werden durch Hands-On-Aktivitäten in der positiven Wahrnehmung ihres Selbstkonzepts und ihrer Selbstwirksamkeit unterstützt (Jansen et al., 2015).

Insgesamt existieren lediglich vereinzelte Studien zum Zusammenhang von forschend-entdeckendem Unterricht und Interesse leistungsstarker Jugendlicher, die häufig lediglich verwandte Merkmale untersuchen. Befunde zum positiven Zusammenhang von forschend-entdeckendem Unterricht und motivationsunterstützenden Lehr-Lernbedingungen legen jedoch nahe, dass forschend-entdeckender Unterricht das Interesse leistungsstarker Jugendlicher unterstützen kann.

4.2.3. Anwendungsbezug im Naturwissenschaftsunterricht

Neben interaktivem Lehren und Lernen und forschend-entdeckendem Unterricht zählen Anwendungsbezüge zu den Unterrichtsaktivitäten, die auf Basis konstruktivistischer Annahmen Lernprozesse unterstützen und für eine Förderung leistungsstarker Jugendlicher in Frage kommen (Bransford & Schwartz, 1999; Donovan & Bransford, 2005). Unter Anwendungsbezügen im Naturwissenschaftsunterricht versteht man die Anwendung naturwissenschaftlicher Konzepte sowie naturwissenschaftlicher Denk- und Arbeitsweisen auf lebensnahe Kontexte und Problemstellungen (Kobarg et al., 2011; Seidel & Prenzel, 2006b). In der Auseinandersetzung mit naturwissenschaftlich-technischen Anwendungen werden Naturwissenschaften als Modell und Repräsentation der Natur und Welt benutzt.

Während zum Zusammenhang von Anwendungsbezügen und Leistung gemischte Ergebnisse gefunden werden, kann ein positiver Zusammenhang mit Interesse relativ übereinstimmend belegt werden (Fey, Gräsel, Puhl & Parchmann, 2004; Kobarg et al., 2011; Murphy & Whitelegg Elizabeth, 2006; Seidel & Prenzel, 2006b). Ähnlich zeigt sich, dass das Interesse an Physik vom neunten zum zehnten Schuljahr stabil bleibt oder sich verbessert, wenn der Unterricht direkte Verbindungen zu lebensnahen Praxissituationen herstellt, die Schülerinnen und Schüler persönlich interessieren (Häussler, Hoffmann et al., 1998; Häussler & Hoffmann, 2000). Anwendungsbezüge können sowohl in traditionellem Unterricht als auch neueren Lehrplänen und Unterrichtsansätzen mit Fokus auf forschend-entdeckenden Unterricht eingebettet sein. Die Anwendung von Gesetzmäßigkeiten sowie naturwissenschaftlicher Phänomene gehört zu den Zielen von Schülerexperimenten (Kircher, Girwidz & Häußler, 2009). Auch Forschungszweige, die für eher transmissiven im Gegensatz zu forschend-entdeckendem Unterricht plädieren, betonen, dass Informationen und Lerninhalte eine Bedeutung für den Lernenden haben müssen, damit transmissiver Unterricht erfolgreich ist (Ausubel, Novak & Hanesian, 1981; Kircher et al., 2009). Naturwissenschaftsunterricht, in dem die Lehrperson durch Anwendungsbezüge Zielklarheit und inhaltliche Kohärenz herstellt, berichten Schülerinnen und Schüler mehr inhaltliche Relevanz und Kompetenzunterstützung (Seidel, Rimmel & Prenzel, 2005). Für Deutschland konnte die IPN-Videostudie jedoch aufzeigen, dass lebensnahe Anwendungsbezüge in Form außerschulischer Beispiele in traditionellen Unterrichtsaktivitäten wie dem Klassengespräch lediglich 9% der Unterrichtszeit ausmachen (Seidel et al., 2003, S. 145).

Relativ unklar ist bisher, inwiefern Anwendungsbezüge auch im Zusammenhang mit Kompetenz und Interesse von High Performern bzw. allgemein leistungsstarken Jugendlichen in Naturwissenschaften stehen und ob Unterschiede zu weniger leistungsstarken Jugendlichen bestehen. Nachfolgend wird jeweils für Leistung und Interesse getrennt der Forschungsstand im Kontext leistungsstarker Jugendlicher dargestellt sowie mit Hilfe prozessbezogener theoretischer Überlegungen ergänzt.

Forschungsstand zum Zusammenhang von Anwendungsbezug und Leistung von High Performern

Betrachtet man den Zusammenhang von Anwendungsbezügen und naturwissenschaftlicher Kompetenz leistungsstarker Schülerinnen und Schüler, so impliziert bereits die Konzeption naturwissenschaftlicher Kompetenz eine funktionale Rolle von Anwendungen: Die Anwendung von Wissen und Fähigkeiten auf alltagsnahe und relevante Kontexte ist ein wesentlicher Aspekt des Rahmenkonzepts naturwissenschaftlicher Kompetenz in PISA 2006 (z. B. Bybee & McCrae, 2011). Schülerinnen und Schüler sind in ihrem Alltag und in außerschulischen Situationen mit vielfältigen naturwissenschaftsbezogenen Anwendungsfeldern konfrontiert. Die Rahmenkonzeption naturwissenschaftlicher Kompetenz greift dies auf. Insbesondere High Performer in Naturwissenschaften sollten daher Anwendungsgelegenheiten von naturwissenschaftlichen Konzepten und Vorgehensweisen meistern und nutzen können. Naturwissenschaftsunterricht, der diese Fähigkeit, naturwissenschaftliches Wissen und Fähigkeiten in lebensnahen Kontexten anzuwenden, zum Ziel hat, muss Lerngelegenheiten dafür bieten und daher lebensnahe Anwendungsbezüge mit einschließen (vgl. Kobarg et al., 2011).

Ergebnisse zu Unterrichtsmethoden mit engem Bezug zu Anwendungen wie situiertem Lernen sowie grundlegende Erkenntnisse zum Wissenserwerbsprozess ermöglichen es, die Bedeutung von Anwendungen für die Leistungsentwicklung von High Performern zunächst aus theoretischer und prozessbezogener Sicht zu betrachten. Da Anwendungen in unterschiedlichen Unterrichtsphasen mit unterschiedlicher Funktion eingesetzt werden können, werden nachfolgend drei Aspekte aufgegriffen und in ihrer Bedeutung für den Wissenserwerb leistungsstarker Jugendlicher beschrieben: (a) Anwendungsbezüge als Teil der Erarbeitung von Inhalten (z. B. situiertes, kontextbezogenes und problembasiertes Lernen), (b) Anwendungsbezüge als Transfermöglichkeit bereits erarbeiteter Inhalte und (c) Anwendungsbezüge als Möglichkeit für Enrichmentprozesse

(z. B. Curriculumentwicklungen).

Anwendungsbezüge als Elemente zur Erarbeitung von Inhalten findet man in situierten, kontextbezogenen oder problembasierten Lernumgebungen. Ansätze zum situierten Lernen nehmen an, dass Wissen an Kontexte oder Situationen gebunden ist und sich aus der Interaktion von Person und Situation ergibt (Greeno, Moore & Smith, 1993; Renkl, 2015). Wissen, das in Wissenserwerbsprozessen ohne Kontext bzw. Anwendungsbezug erarbeitet wird, erweist sich häufig als „träges Wissen“ (Renkl, 2010; Whitehead, 2008), das deklarativ zwar memorisiert wurde, jedoch in Anwendungssituationen nicht effektiv genutzt werden kann. Aus diesen Überlegungen folgern Vertreter situierten Lernens, dass auch Kognition und Lernen situiert und kontextgebunden gestaltet sein sollte (vgl. Renkl, 2015). Diese Annahme kann auf den Wissenserwerb leistungsstarker Jugendlicher übertragen werden. Auch im Zusammenhang mit Hochbegabung stellen Wissenschaftler die Wichtigkeit aktiven und anwendbaren Wissens heraus und betonen, dass die bloße Wissensvermittlung nicht zu Expertise, sondern tragem Wissen führt (Gruber, 2007; Stark, Gruber, Renkl & Mandl, 1998).

Kontextbasiertes (King & Ritchie, 2012) und problembasiertes Lernen (Dochy, Segers, van den Bossche & Gijbels, 2003) greifen die Annahmen situierten Lernens auf und sind eng miteinander verbunden. Kontextbasierte Lernumgebungen basieren auf lebensweltnahen Anwendungsbezügen (King & Ritchie, 2012) und wurden im Rahmen von Kontextprojekten vor allem in Chemie (*Salters* (Ramsden, 1997), *Chemie im Kontext* (Parchmann et al., 2006)) aber auch Physik (*PLON* (Wierstra, 1984)) untersucht (für eine Übersicht vgl. King & Ritchie, 2012). Sie finden im Vergleich zu traditionellem Unterricht keine oder nur geringe Unterschiede hinsichtlich der Leistungsentwicklung (Fey et al., 2004; Ramsden, 1997; Wierstra, 1984), jedoch in Falle des *Salters*-Programms ein tieferes Konzeptverständnis (Barker & Millar, 2000). Ähnlich wie kontextbasierte Lernumgebungen benutzen problembasierte Lernumgebungen (für ein Review siehe Wilder, 2015) meist einen lebensweltnahen bzw. authentischen Kontext. Studien im Bereich der Hochbegabungsforschung zeigen für allgemein leistungsstarke Schülerinnen und Schüler, dass problembasiertes Lernen traditionellem Unterricht überlegen ist (Gallagher, Stepien & Rosenthal, 1992). Tyler-Wood, Mortenson, Putney und Cass (2000) entwickelten für die besten zehn Prozent der Schülerinnen und Schüler in Mathematik und Physik in der achten Jahrgangsstufe ein Curriculum für ein zweijähriges Programm, das unter anderem realitätsnahe Anwendungsbezüge in Laborumgebungen als zentralen Aspekt

enthält. Im Vergleich zur ähnlich starken Kontrollgruppe konnten leistungsstarke Schülerinnen und Schüler des zweijährigen Programms höhere Leistungen in Mathematik und Naturwissenschaften am Ende der zehnten Jahrgangsstufe entwickeln und wiesen bis zum Ende der High School höhere Leistungen auf. Das in Kooperation mit Lehrkräften und hohem Arbeitsaufwand entwickelte interdisziplinäre und integrierte Curriculum zeichnet diese Studie aus, einschränkend ist jedoch die geringe Stichprobengröße und damit verbundene begrenzte Verallgemeinerung zu nennen. Ähnliche Ergebnisse finden (Etkina, Matilsky & Lawrence, 2003) in einem Astrophysikprogramm, in dem sich Schülerinnen und Schüler mit authentischen Forschungskontexten beschäftigten.

Anstatt zur direkten Erarbeitung von Inhalten können Anwendungen im Unterricht als Transfermöglichkeit bereits bekannter Inhalte eingesetzt werden. Transfer bezeichnet die Anwendung von Wissen bzw. Fertigkeiten in neuen Situationen bzw. Kontexten (für eine Übersicht vgl. Bransford & Schwartz, 1999). Verschiedene Taxonomien von Lehrzielen (Bloom, 1956; Merrill, 1987) sowie Konzeptionen von Instruktionstheorien (z. B. Gagne, 1985; Klauer, 1985) schließen Anwendenden und Transfer als kognitives Lehrziel bzw. wichtigen Lehr-Lernschritt mit ein. Anwendungen ermöglichen die Gestaltung von Transferaufgaben, in denen bereits bekannte Inhalte auf lebensnahe Fragestellungen bzw. Kontexte übertragen werden oder die Relevanz von bekannten Inhalten für den Alltag verdeutlicht wird. Für intelligente Schülerinnen und Schüler zeigt sich ebenso wie für andere Schülerinnen und Schüler, dass sie Transferleistungen nicht automatisch erbringen, sondern vielmehr durch entsprechende Lernumgebungen im Erwerb von Transferfähigkeiten unterstützt werden müssen (Mähler & Stern, 2010). Im Hinblick auf die Förderung naturwissenschaftlicher Kompetenz leistungsstarker Schülerinnen und Schüler ermöglichen Anwendungen daher aus theoretischer Sicht Wissenserwerbs- sowie Transferprozesse.

Schließlich können Anwendungsbezüge in jeder Unterrichtsphasen als praktikable Enrichment- oder Differenzierungsmöglichkeit verwendet werden. Enrichment bezeichnet die Anreicherung und Vertiefung von Themen und Inhalte für leistungsstarke Schülerinnen und Schüler und gehört neben Akzeleration (beschleunigtes Durchlaufen des Unterrichtsstoffes) zu einer der prominenten Fördermaßnahmen der Hochbegabungsforschung (für eine detaillierte Übersicht siehe Reis & Renzulli, 2010). Anwendungen können als Enrichment eingesetzt werden, da sie zusätzlich zum vorgesehenen Unterrichtsstoff eine Verbreiterung sowie Vertiefung der Inhalte ermöglichen, bei der

leistungsstarke und leistungsschwächere Schülerinnen und Schüler trotzdem gleichzeitig beim gleichen inhaltlichen Stoffgebiet bleiben können. Curriculumentwicklungen für leistungsstarke Schülerinnen und Schüler greifen diese Überlegung auf und betonen die funktionale Bedeutung lebensnaher Kontexte und Anwendungsbezüge als Teil des Curriculums (z. B. VanTassel-Baska, 2003; VanTassel-Baska & Wood, 2010). Insgesamt ist davon auszugehen, dass Anwendungen leistungsstarke Jugendliche in Naturwissenschaften in ihrer Kompetenzentwicklung unterstützen, da sie im Sinne situierten bzw. problembasierten Lernens Lernprozesse fördern, Transfer ermöglichen sowie Enrichtmentmöglichkeiten bieten.

Explizit untersucht wurden Anwendungsbezüge im Zusammenhang mit Leistung von High Performern in Naturwissenschaften im Rahmen deskriptiver Analysen von PISA 2006 (OECD, 2007). Im internationalen Vergleich berichten Top Performer (Schülerinnen und Schüler auf Kompetenzstufe V und VI) in zwölf von 28 OECD-Staaten mehr Anwendungsbezüge im Naturwissenschaftsunterricht als Strong Performer (Schülerinnen und Schüler auf Kompetenzstufe IV) (OECD, 2009b, S. 58). Demgegenüber weisen Top Performer jedoch in den meisten der restlichen Länder, darunter Deutschland, zwar tendenziell höhere Anwendungsbezüge auf, unterscheiden sich jedoch nicht signifikant von den Strong Performern. High Performer, definiert als die besten 20% der Schülerinnen und Schüler, enthalten Schülerinnen und Schüler bis einschließlich eines oberen Teils auf Kompetenzstufe IV und weisen daher eine Schnittmenge zu den Strong Performern der OECD Analyse auf (vgl. Kapitel 2.1). Auch innerhalb der High Performer sind daher ähnliche Ergebnisse zu erwarten. Im Detail wurde ein Zusammenhang von Anwendungsbezügen und naturwissenschaftlicher Kompetenz innerhalb der Top und High Performer in PISA bisher jedoch empirisch nicht untersucht.

Die angestellten prozessbezogenen Überlegungen zum Zusammenhang von Anwendungen und naturwissenschaftlicher Kompetenz leistungsstarker Jugendlicher basieren auf allgemeinen Überlegungen zum Lernprozess und gelten daher grundsätzlich für Jugendliche aller Leistungsniveaus. Aus differenzieller Perspektive stellt sich jedoch die Frage, inwiefern sich der Zusammenhang in der Höhe für leistungsstarke und den Rest der Schülerinnen und Schüler unterscheidet. Aufgrund der geringen Erkenntnislage können hier keine direkten Schlüsse getroffen werden. Rückschlüsse auf differenzielle Effekte können jedoch getroffen werden, indem bekannte Einflüsse von Vorwissen auf Lernprozesse in Verbindung mit prozessbezogenen Aspekten von anwendungsbasier-

tem Lernen betrachtet werden. Dabei wird nachfolgend zwischen Unterschieden in der Informationsverarbeitung einerseits und in den verwendeten Problemlösestrategien andererseits unterschieden.

Betrachtet man die Informationsverarbeitung in anwendungsbezogenen Unterrichtsumgebungen als ersten Punkt, so liefert Forschung zu „chunking“, der Perspektive der fokussierten Informationsverarbeitung und der Prozeduralisierung von Wissen Ergebnisse zu möglichen Unterschieden. Anwendungsbezüge in situierten und problembezogenen Lernumgebungen aber auch in Transferaufgaben erhöhen durch den Einschluss eines Kontextes unter Umständen die von Jugendlichen zu dekodierende Informationsmenge oder -Komplexität. Die Experten-Novizen-Forschung belegt, dass ein höheres Vorwissen die Dekodierung großer Informationen erleichtert, indem sie die Verarbeitung in größeren Einheiten („chunks“) erlaubt und bereits bei der Informationswahrnehmung eine Verbindung von Wissen und Erfahrungen stattfindet (vgl. Gruber, Renkl & Schneider, 1994; Renkl, 1996).

Neben der Möglichkeit, größere Informationsmengen bei Anwendungskontexten zu verarbeiten, fällt leistungsstarken Jugendlichen möglicherweise die Konzentration auf die wesentlichen Informationen innerhalb dieser großen Informationsmenge leichter. Kognitionspsychologen argumentieren in der *Perspektive der fokussierten Informationsverarbeitung* auf Wissenserwerbsprozesse (Renkl & Atkinson, 2007; Renkl, 2015), dass Lernen nicht nur durch konstruktivistische kognitive Aktivitäten allein bedingt wird, sondern davon abhängt, dass Lernende bzw. Lernmaterialien einen Fokus auf zentrale Konzepte und Prinzipien legen. Diese Fokussierung wirkt sich beispielsweise bei der aktiven Verarbeitung von Texten, denen relevante Informationen unter interessanten aber unwichtigen Details entnommen werden müssen, auf Lernprozesse aus (Garner, Gillingham & White, 1989). Schmidt, de Volder, de Grave, Moust und Patel (1989) untersuchten in diesem Zusammenhang den Einfluss einer Diskussion zu Beginn einer Lernumgebung, mit deren Hilfe die Aufmerksamkeit der Schülerinnen und Schüler auf eine lebensnahe Fragestellung zu Blutzellen gelenkt wurde. Die Fokussierung erhöhte den Lernerfolg der Schülerinnen und Schüler und half insbesondere leistungsschwachen Schülerinnen und Schülern. Ausgehend vom höheren Vorwissen und dadurch bedingten leichteren themenbezogenen Informationsverarbeitung kann davon ausgegangen werden, dass leistungsstarken Schülerinnen und Schülern eine fokussierte Informationsverarbeitung von Anwendungsbezügen – sofern sie nicht automatisch durch die

Lernumgebung unterstützt wird – leichter fällt als leistungsschwächeren Schülerinnen und Schülern.

Schließlich betont die ACT*-Theorie des Wissenserwerbs („Adaptive Control of Thought Theory“; J. R. Anderson, 1982), dass die Anwendung von Wissen die Umwandlung von deklarativem in prozedurales Wissen, d. h. Wenn-dann-Regeln, erfordert. Experten besitzen mehr prozedurales Wissen und können ihr Wissen daher leichter anwenden und Kontextinformationen bereits in der Wahrnehmung mit Vorwissen verknüpfen (z. B. Gruber & Stamouli, 2015). Ähnlich spricht Gruber (2007) in seinem Handlungskompetenzmodell von Expertise von einer „Überlegenheit bezüglich Routinen“, die mit zunehmender Leistung in einem Bereich einhergeht. Leistungsstarke Jugendliche lernen daher möglicherweise mit einer höheren Lerngeschwindigkeit oder in höherem Ausmaß in anwendungsbezogenen Lernumgebungen als leistungsschwächer Jugendliche.

Neben der Informationsverarbeitung liefert die Experten-Novizen-Forschung Hinweise darauf, dass leistungsstarke Jugendliche bei anwendungsbezogenen Lernumgebungen möglicherweise Vorteile durch die Nutzung von Problemlösestrategien aufweisen. Problembasierte Lernumgebungen als ein Beispiel anwendungsorientierten Lernens können unterschiedlich komplexe Aufgabenstellungen und unterschiedlich gut definierte Anforderungen bereitstellen. Dörner (1979) spricht in diesem Zusammenhang von wohl bzw. schlecht definierten Problemen (vgl. Gruber & Stamouli, 2015). Experten-Novizen-Forschung zeigt, dass Experten aufgrund ihres höheren Vorwissens und bestehender Schemata in komplexen Aufgabenstellungen hochwertigere Problemrepräsentationen entwickeln als Novizen (Chi, Feltovich & Glaser, 1981). Sie nutzen zudem andere Problemlösestrategien, die vom gegebenen Istzustand ausgehend vorwärts denkend angelegt sind, während Novizen auf umgekehrte Strategien, die vom Sollzustand ausgehen, angewiesen sind (Larkin, McDermott, Simon & Simon, 1980). Insgesamt wird zunehmende Leistungsexpertise mit zunehmender Flexibilität hinsichtlich Problemrepräsentationen, Analyseebenen und Verarbeitungsstrategien in Verbindung gebracht (vgl. Modell der Handlungskompetenz von Gruber, 2007). Insgesamt sprechen die Befunde dafür, dass High Performer in Naturwissenschaft sowohl in der Informationsverarbeitung kontextbasierter Aufgabenstellungen als auch in der Nutzung von Lösungsstrategien bei problembasierten Lernumgebungen Nicht-High-Performern überlegen sind und einen höheren Lernerfolg bzw. eine höhere Lerngeschwindigkeit aufweisen.

Forschungsstand zum Zusammenhang von Anwendungsbezug und Interesse von High Performern

Prozessbezogene Überlegungen und Ergebnisse zum Zusammenhang von Anwendungen und Lernmotivation bzw. Interesse liefern eine theoretische Begründung eines positiven Zusammenhangs von Anwendungsbezügen und Interesse leistungsstarker Jugendlicher in Naturwissenschaften. Anwendungen und Beispiele generieren inhaltliche Relevanz (Dalehefte, 2001; Seidel, 2003) und unterstützen damit Bedingungen für die Entwicklung von Lernmotivation (z. B. Prenzel, 1997). Analog dazu wird in der Interessenforschung argumentiert, dass Unterrichtsmaterial mit hoher Bedeutsamkeit und Bezug zum Alltag situatives Interesse fördert (Hoffmann, 2002; Hulleman & Harackiewicz, 2009; Hulleman, Godes, Hendricks & Harackiewicz, 2010; Linnenbrink-Garcia et al., 2013). Unterschiedliche Kontextprojekte in Physik, Chemie oder Biologie in Deutschland (*ChiK*; Parchmann et al., 2006; *BiK*; Bayrhuber et al., 2007; *PhiK*; Mikelskis-Seifert & Duit, 2007), den USA (*Chemistry in Context*; A. T. Schwartz, 1997), Großbritannien (*Salters*; Bennett & Lubben, 2006; Ramsden, 1997), Israel (*Industrial Science*; Hofstein & Kesner, 2006) und den Niederlanden (*PLON*; Eijkelhof & Lijnse, 1988) konnten diesen Zusammenhang von lebensweltnahen Anwendungsbezügen und Interesse oder verwandten motivational-affektiven Merkmalen im Unterschied zum Zusammenhang von Anwendungsbezügen und Leistung nachweisen (für weitere Projekte und eine detaillierte Übersicht siehe Fensham, 2009; King & Ritchie, 2012).

Forschung zu Typologien individuellen Schülermerkmalen bestätigt diesen Zusammenhang indirekt auch für leistungsstarke Schülerinnen und Schüler. Im Rahmen der IPN Videostudie (Seidel & Prenzel, 2004; Seidel, Prenzel, Rimmele et al., 2007) wurden Profile aus Leistung, Selbstkonzept und Interesse in Physik unter den Schülerinnen und Schülern der neunten Jahrgangsstufe identifiziert und der Zusammenhang mit der Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen untersucht (Seidel, 2006). Zwei der Profile können als leistungsstarke Jugendliche mit hohem bzw. niedrigem Interesse angesehen werden. Die Unterrichtsanalysen zeigen, dass leistungsstarke Schülerinnen und Schüler mit interessiertem Profil ihren Unterricht inhaltlich relevanter wahrnehmen als leistungsstarke Schülerinnen und Schüler mit uninteressiertem Profil. Da Anwendungen zu höherer wahrgenommener inhaltlicher Relevanz führen, ist anzunehmen, dass leistungsstarke Jugendliche durch Anwendungen in ihrem Interesse unterstützt werden.

Darüber hinaus greifen Forschungsinitiativen, die Curricula für leistungsstarke Jugendliche entwickeln, die inhaltliche Relevanz als Bedingung für die motivationale Förderung leistungsstarker Jugendlicher auf und setzen sie in Form von Anwendungen um. Prominente Curriculumentwicklungen für hochleistende bzw. hochbegabte Schülerinnen und Schüler wie das *Intergrated Curriculum Model* (VanTassel-Baska & Wood, 2010) bzw. das darauf aufbauende *National Science Curriculum Project for High-Ability Learners* (VanTassel-Baska et al., 1998) betonen die Wichtigkeit lebensweltnaher Problemstellungen und inhaltlich relevanter Lerninhalte sowie der Multidisziplinarität von Unterrichtsinhalten für leistungsstarke Schülerinnen und Schüler (VanTassel-Baska et al., 1998; Robinson, Shore & Enersen, 2006b). Lehrkräfte bestätigen, dass die anwendungsbezogenen Curriculumthemen zu mehr Motivation, Interaktion und Engagement führen (VanTassel-Baska et al., 1998). Auch talentierte Jugendliche in Naturwissenschaften selbst äußern nach ihrem Naturwissenschaftsunterricht gefragt den Wunsch nach mehr problem-basiertem Lernen und einem höheren Anwendungsbezug zur realen Welt der Naturwissenschaften (Cross & Coleman, 1992).

Prenzel und Schütte (2007) haben den Zusammenhang von Anwendungen und Interesse von leistungsstarken Jugendlichen in PISA 2006 (OECD, 2007) explizit untersucht. Naturwissenschaftliche Anwendungen konnten im Unterschied zu Schülerexperimenten als einziges Unterrichtsmerkmal signifikant das Interesse von Schülerinnen und Schülern des obersten Leistungsquartils in Naturwissenschaften vorhersagen. Ähnlich beobachteten (Taskinen et al., 2008) einen Zusammenhang von Anwendungen im Naturwissenschaftsunterricht und der zukunftsorientierten Motivation leistungsstarker Schülerinnen und Schüler in PISA 2006.

Mögliche Unterschiede zwischen High und Nicht-High-Performern im Zusammenhang von Anwendungsbezügen und Interesse liefert wiederum eine differenzielle Betrachtung. Die Annahme, dass Anwendungen, die die inhaltliche Relevanz unterstützen, im Lernprozess das Interesse fördern können, gilt aus Sicht der pädagogischen Interestheorie (Prenzel, 1988, 1992; H. Schiefele et al., 1983) und Selbstbestimmungstheorie (Deci & Ryan, 1985, 2000, 2002) unabhängig von der Leistung. Für Nicht-High-Performer ist daher aus theoretischer Sicht zunächst ein ähnlicher Zusammenhang von Anwendungen und Interesse zu erwarten wie oben für High Performer begründet. Da Kontexte und Anwendungen aufgrund der hohen Informationsmenge unter Umständen die Komplexität erhöhen, sind Unterschiede vor allem dann zugunsten leistungsstarker

Schülerinnen und Schüler denkbar, wenn die Gestaltung der anwendungsbezogenen Lernumgebung nicht an die Lernvoraussetzungen leistungsschwächerer Schülerinnen und Schüler angepasst ist und diese überfordert (vgl. zur Passung von Anforderung und Fähigkeit: Csikszentmihalyi & Schiefele, 1993). Im Rahmen des Salters-Projekts zeigt sich der Kontext beispielsweise zwar im Mittel interessenförderlich, weniger leistungsstarke Schülerinnen und Schüler berichteten jedoch, dass sie die Fragestellung und die wenig vorhandene Möglichkeit, Routinen anzuwenden, schwierig fanden (King & Ritchie, 2012, orig. Barber, 2000).

Zwischen leistungsstarken und weniger leistungsstarken Jugendlichen sind auch dann Unterschiede denkbar, wenn Anwendungen nicht den Lebensweltbezug aller Schülerinnen und Schüler treffen, weil sie beispielsweise eher technisch ausgerichtet sind (vgl. auch die Interessentypen in Kapitel 3). Leistungsstarke Schülerinnen und Schüler beschäftigen sich möglicherweise auch außerhalb der Schule häufiger mit Naturwissenschaften und konnten aufgrund ihrer höheren Kompetenz mehr Verbindungen zwischen Alltag und Naturwissenschaften herstellen. Trifft diese Annahme zu, werden unterschiedliche Anwendungen leichter ihren Lebensweltbezug treffen können als den von weniger leistungsstarken Schülerinnen und Schülern. Sowohl die Unterschiede aufgrund der Komplexität als auch die Unterschiede aufgrund des Lebensweltbezugs sind jedoch von der Gestaltung abhängig und für eine anwendungsbezogene Lernumgebung nicht zwingend zutreffend.

Insgesamt ist aus prozessbezogener Sicht davon auszugehen, dass Anwendungsbezüge im Naturwissenschaftsunterricht sowohl für die Leistungs- als auch Interessenentwicklung von High Performern förderlich eingesetzt werden können. Abgesehen von einzelnen Untersuchungen zu Anwendungen im Zusammenhang mit Kompetenz (OECD, 2009b) bzw. Interesse (Prenzel & Schütte, 2007) leistungsstarker Jugendlicher, existieren jedoch bisher lediglich Nachweise im Bereich kontextbasiertem, situativem und problembasiertem Lernen sowie Transfer und Enrichment. Empirisch steht deshalb aus, den Zusammenhang von Anwendungsbezügen im alltäglichen Naturwissenschaftsunterricht und Kompetenz bzw. Interesse auf Basis einer großen Stichprobe zu belegen. Ähnlich zu interaktivem Lehren und Lernen und forschend-entdeckendem Unterricht kann dabei auch im Rahmen von Anwendungsbezügen vermutet werden, dass High Performer aufgrund ihrer Expertise in der Informationsverarbeitung und in Problemlösestrategien schwächeren Schülerinnen und Schülern überlegen sind und daher bei

4. Die Bedeutung des Unterrichts für die Kompetenz- und Interessenentwicklung

durch Anwendungen möglicherweise gegebenen anspruchsvollen Kontexten und Problemlöseaufgabenstellungen erfolgreicher lernen.

4.2.4. Zusammenfassung

Leitende Fragen dieses Kapitels waren: Welche Unterrichtsaktivitäten im Naturwissenschaftsunterricht haben aus prozessbezogener Sicht das Potential, High Performer in Naturwissenschaften hinsichtlich ihrer Kompetenz und/oder ihrem Interesse angemessen, d. h. differenziell, zu fördern? und was ist aus differenzieller Sicht bei diesen Unterrichtsaktivitäten über Unterschiede zwischen High Performern und Nicht-High-Performern bekannt bzw. kann aus anderen Forschungsbereichen abgeleitet werden?

Eingebettet in das Rahmenmodell zur Untersuchung naturwissenschaftlichen Unterrichts aus PISA 2006 (Seidel & Prenzel, 2006b), das als konzeptionelle Grundlage die Berücksichtigung von sowohl prozessbezogenen Aspekten der Unterrichtsaktivitäten als auch differenziellen Einflüssen ermöglicht, wurden interaktives Lehren und Lernen, forschend-entdeckender Unterricht sowie Anwendungen als mögliche Unterrichtsaktivitäten zur Förderung leistungsstarker Schülerinnen und Schüler beschrieben. Sowohl interaktives Lehren und Lernen als auch forschend-entdeckender Unterricht haben sich in verschiedenen Studien als effektiv zur Förderung von Leistung und Interesse erwiesen, auch bei leistungsstarken Jugendlichen. Untersuchungen zur Orchestration von Unterricht konnten zeigen, dass interaktives Lehren und Lernen im Sinne eines Unterrichtsgesprächs vor allem mit hoher Leistung assoziiert ist, forschend-entdeckendes Lernen im Sinne häufiger Experimente vor allem mit hohem Interesse. Mit der höchsten Kompetenz und dem höchsten Interesse geht allerdings Unterricht einher, der Schülerinnen und Schüler häufig interaktiv einbezieht und in dem gezielt Schülerexperimente eingesetzt werden (Kobarg et al., 2011). Neben der Orchestration hat auch die motivationsunterstützende Gestaltung der Unterrichtsaktivitäten einen Einfluss darauf, wie wirksam sie motivational-affektive Merkmale fördern und zu Lernaktivitäten anregen (Rakoczy et al., 2008; Prenzel, Kramer & Drechsel, 2002; Kunter, 2005; Seidel, 2003). Insbesondere die IPN-Videostudie konnte hier wichtige Zusammenhänge im Physikunterricht verdeutlichen und zeigen, dass vor allem bei Schülerexperimenten die Einbettung einen entscheidenden Einfluss hat (z. B. Seidel, Prenzel, Rimmele et al., 2007). Insgesamt haben daher alle genannten Unterrichtsaktivitäten das Potential, Kompetenz und Interesse auch von High Performern zu unterstützen. In welchem Ausmaß sie das in der Unterrichtsrealität tun, hängt jedoch von der Orchestration, Einbettung und qualitativen Gestaltung der Unterrichtsaktivitäten sowie den individuellen Schülervoraussetzungen ab.

Zur Beantwortung der zweiten Frage nach möglichen Unterschieden zwischen High und Nicht-High-Performern wurde einerseits auf Forschungsergebnisse direkt zu den Unterrichtsaktivitäten und andererseits auf Erkenntnisse aus Forschung zu Vorwissen, Expertise und Hochbegabung zurückgegriffen. In Bezug auf die Leistung sprechen Elaborations- und Organisationsprozesse und ein erhöhtes verbales Engagement dafür, dass High Performer in interaktiven Lehr-Lernprozessen schneller und dadurch ggf. mehr lernen als Nicht-High-Performer. Ähnlich lassen bei forschend-entdeckendem Lernen Annahmen der Aptitude-Treatment-Interaction (Corno & Snow, 1986; Snow, 1989) sowie prozessbezogene Überlegungen zur Informationsverarbeitung und inhaltlichen und strukturellen Offenheit vermuten, dass leistungsstarke Schülerinnen und Schülern den restlichen Jugendlichen in Lernprozessen überlegen sind. Aus empirischer Sicht existieren jedoch nur begrenzte Ergebnisse und deuten darüber hinaus einen umgekehrten Zusammenhang an (White & Frederiksen, 1998). Auch bei Anwendungsbezügen deuten Ergebnisse zur Informationsverarbeitung und zu Problemlösestrategien leistungsstarker Jugendlicher darauf hin, dass High Performer Vorteile im Lernprozess aufweisen.

Im Hinblick auf Interesse fördern aus motivationstheoretischer Sicht alle Unterrichtsmerkmale sowohl die Lernmotivation von High als auch Nicht-High-Performern. Unterschiede zugunsten der Nicht-High-Performer sind bei interaktivem Lehren und Lernen und forschend-entdeckendem Unterricht nur dann denkbar, wenn für High Performer demotivierende Prozesse in heterogenen Gruppen auftreten. Umgekehrt ist denkbar, dass Anwendungsbezüge High Performer stärker in ihrem Interesse unterstützen, wenn sie nur deren Lebenswelt treffen oder zu komplex gestaltet sind.

Grundsätzlich hängt der Zusammenhang von Unterrichtsaktivitäten und Kompetenz bzw. Interesse und, ob es Unterschiede zwischen High und Nicht-High-Performern gibt, auch von der qualitativen Gestaltung der Unterrichtsaktivitäten ab (Furtak & Kunter, 2012; Tesch & Duit, 2004; J.-R. Wang et al., 2010). Insbesondere ob eine innere Differenzierung umgesetzt wird und High Performer die Möglichkeit haben, Lerninhalte zu erweitern und zu beschleunigen, beeinflusst, ob der mit der Unterrichtsaktivität maximal mögliche Zusammenhang zu Kompetenz und Interesse erreicht wird oder High Performer umgekehrt nicht die Möglichkeit haben, mehr oder schneller zu lernen als Nicht-High-Performer.

4.3. Einfluss individueller Schülervoraussetzungen auf die Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen

Die vorangegangenen Kapitel zeigen, dass leistungsstarke Jugendliche in Naturwissenschaften mit höherer Wahrscheinlichkeit ein hohes Selbstkonzept und Interesse in Naturwissenschaften aufweisen als leistungsschwächere Jugendliche (Kapitel 3) und Unterrichtsaktivitäten wie forschend-entdeckender Unterricht das Interesse leistungsstarker Jugendlicher zusätzlich unterstützen (Kapitel 4). Gleichzeitig verdeutlichen Studien jedoch, dass ein beträchtlicher Teil der leistungsstarken Jugendlichen wenig leistungszuversichtlich und uninteressiert ist (vgl. Kapitel 3). Dies wird insbesondere auch an unterschiedlichen Profilen hinsichtlich Leistung, Selbstkonzept und Interesse (z. B. leistungsstarke aber wenig leistungszuversichtliche und uninteressierte Schüler) anschaulich, die bisherige Forschungsbemühungen identifizieren konnten (vgl. Kapitel 3.3). Die Suche nach Gründen und zukünftigen Fördermöglichkeiten wirft die Frage danach auf, wie diese leistungsstarken Jugendlichen ihren Unterricht erleben und, ob Unterrichtsaktivitäten wie forschend-entdeckender Unterricht einen Teil der High Performer motivational-affektiv möglicherweise nicht erreicht und ihr kognitives Potential damit nicht ausschöpft. Nachfolgend wird zur Beantwortung dieser Fragen zunächst dargestellt, dass motivationsunterstützende Lehr-Lernbedingungen eine vermittelnde Rolle in der Förderung interessierter Lernmotivation spielen. Daraufhin wird erläutert, wie individuelle Voraussetzungen wie Selbstkonzept und Interesse die Wahrnehmung dieser Lehr-Lernbedingungen beeinflussen können.

Die Unterrichtsforschung der letzten Jahrzehnte weist darauf hin, dass oberflächliche Strukturen der Unterrichtsgestaltung wie Schüler-/Lehrerzentrierung weniger bedeutsam sind als vielfach angenommen, sondern vor allem tiefergehende Strukturen und vermittelnde Prozesse entscheidend sind (Brophy, 1999; Ditton, 2002a; Kunter, 2005; Weinert & Helmke, 2010). Oser und Patry (1990) und Oser und Baeriswyl (2001) haben diesen Forschungsbereich durch die Unterscheidung von Sichtstrukturen (Unterrichtsorganisation) und Basismodellen des Lernprozesses (sequenzierte und verkettete Operationen bzw. Lernaktivitäten zu einem Lerninhalt) als Elemente der „Choreographien unterrichtlichen Lernens“ geprägt. Aufbauend darauf wird zwischen Sichtstrukturen

und Gelegenheitsstrukturen (Seidel, 2003) bzw. Tiefenstrukturen (Fischer, Neumann et al., 2014) im Naturwissenschaftsunterricht unterschieden. Motivationsunterstützende Lehr-Lernbedingungen wie Kompetenzunterstützung, Soziale Eingebundenheit, Autonomieunterstützung, inhaltliche Relevanz, Instruktionsqualität und Interesse der Lehrperson gehören zu diesen Tiefenstrukturen bzw. vermittelnden Prozessen (vgl. Rakoczy, 2008) (s. Kapitel 3.4). Ihre vermittelnde Rolle wird in dieser Arbeit exemplarisch am Beispiel forschend-entdeckenden Lernens dargestellt. Da forschend-entdeckendes Lernen durch sozial-konstruktive Lernprozesse charakterisiert wird und Schülerinnen und Schülern aktive und eigenständige Lernerfahrungen erlaubt, erfüllt es aus theoretischer Sicht vor allem Gestaltungsmerkmale für das Erleben von Autonomie, Kompetenz, sozialer Einbindung und inhaltlicher Relevanz (z. B. Rakoczy, 2008; Prenzel, 1997; Prenzel, Kramer & Drechsel, 2001)(vgl. auch Kapitel 3.4 und 4.2.2). Da die soziale Eingebundenheit im Rahmen von PISA im Hinblick auf das Verhalten der Lehrperson gegenüber dem Schüler bzw. der Schülerin operationalisiert wird und nicht im Hinblick auf die Eingebundenheit unter Gleichaltrigen (Frey & Asseburg, 2009), wird sie in dieser Arbeit nicht betrachtet.

Aus empirischer Sicht unterstützen verschiedene Forschungsbemühungen die Annahme, dass forschend-entdeckendes Lernen vermittelt über bzw. unter der Bedingung von wahrgenommener Autonomieunterstützung, Kompetenzunterstützung und inhaltlicher Relevanz motivationsfördernd wirkt. In Anlehnung an andere Autoren (z. B. Blumenfeld et al., 1991; Hardy, Jonen, Möller & Stern, 2006; Mayer, 2004) greifen Furtak und Kunter (2012) die Wichtigkeit der Autonomieunterstützung bei forschend-entdeckendem Lernen auf und unterscheiden analog zu Stefanou, Perencevich, DiCintio und Turner (2004) zwischen kognitiver und prozeduraler Autonomie. Kognitive Autonomie betont dabei, dass Schülerinnen und Schüler nicht nur Geräte auswählen und Experimente zusammenbauen dürfen (prozedurale Autonomie), sondern auch eigene Ideen und Lösungen entwickeln dürfen sowie informatives Feedback (kognitive Autonomie) bekommen. Darauf aufbauend argumentieren sie, dass viele gemischte Ergebnisse durch fehlende oder falsch interpretierte Autonomie erklärt und viele kritische Ansichten zu forschend-entdeckendem Lernen, die sich beispielsweise auf die geringe Lenkung, Lernbegleitung und Anleitung durch die Lehrperson beziehen (z. B. Kirschner et al., 2006; Mayer, 2004), durch die Umsetzung einer kognitiv autonomen Gestaltung aufgelöst werden können (Furtak & Kunter, 2012). Eine experimentelle Untersuchung ergibt jedoch – anders als erwartet, dass eine kognitiv autonom gestaltete forschend-entde-

ckende Lernumgebung nicht zu mehr wahrgenommenem Autonomieerleben und nicht zu mehr intrinsischer Motivation führt als eine lediglich prozedural autonom gestaltete forschend-entdeckende Lernumgebung (Furtak & Kunter, 2012). Die Autoren erklären dies damit, dass kognitiv aktivierende Lernumgebungen anstrengender sind (Blumenfeld et al., 1991) und die Stichprobe gleichzeitig im Umfeld traditioneller lehrerzentrierter Methoden sozialisiert wurde sowie aufgrund der kurzen Dauer der Intervention kein vertrauensvolles Verhältnis zur Lehrperson aufgebaut werden konnte (Furtak & Kunter, 2012). Andere Studien zeigen jedoch, dass kognitiv autonome Lernaktivitäten wie aktives Denken und Schlüsse ziehen aus Experimenten (Blanchard et al., 2010; Minner et al., 2010) und Argumentatieren (z. B. Berland & McNeill, 2010; Osborne, 2010) die Effektivität forschend-entdeckender Lernumgebungen beeinflusst. Bereits die „Mrs. Oublier“-Studie von Cohen (1990) betont für Mathematik, dass „hands-on“ alleine nicht ausreicht, sondern auch ein „minds-on“ ermöglichen muss, was auch Studien zu Naturwissenschaften, beispielsweise die IPN-Videostudie (Tesch & Duit, 2004) und PISA 2006 (Kobarg et al., 2011) bestätigen.

Neben der Autonomie unterstützt forschend-entdeckender Unterricht das Erleben von Kompetenz, indem er Schülerinnen und Schülern die Möglichkeit bietet, sich wirksam in Interaktion mit der Umwelt zu erleben und ihre Fähigkeiten erfolgreich anzuwenden (Deci & Ryan, 2000; Ryan & Deci, 2002). Forschend-entdeckende Lernumgebungen ermöglichen darüber hinaus Differenzierungsmaßnahmen und begünstigen das für das maximale Kompetenzerleben wichtige optimale Anforderungsniveau, bei dem individuelle Voraussetzungen zu den Anforderungen einer Aktivität passen (Kunter, 2005; Rheinberg, 2004; Rheinberg & Vollmeyer, 2000). Weiterführende Analysen zu PISA 2006 (OECD, 2007) zeigen darüber hinaus, dass forschend-entdeckender Unterricht die Selbstwirksamkeitserwartung von Schülerinnen und Schülern vorhersagt (Jansen et al., 2015) und daher das Erleben von Selbstvertrauen und Wirksamkeit unterstützt.

Schließlich spielt die inhaltliche Relevanz eine wichtige Rolle dafür, dass forschend-entdeckender Unterricht intrinsische Motivation fördert. Inhaltliche Relevanz beschreibt, dass die Bedeutung und Relevanz von Lerninhalten sowie die Ziele, die mit der Vermittlung der Lerninhalte im Hinblick auf die Lebenswelt der Schülerinnen und Schüler verbunden werden, transparent und klar dargestellt werden (Prenzel, 1997; Prenzel, Kramer & Drechsel, 2001). Ergebnisse zum Physikunterricht in Deutschland (Tesch & Duit, 2004) und Schülerlaboren (Brandt, Möller & Kohse-Höinghaus, 2008; Engeln, 2004;

Guderian & Priemer, 2008) deuten an, dass weniger die Häufigkeit oder Zeit von Schülerexperimenten, sondern die Einbettung bzw. Vor- und Nachbereitung die Effekte auf motivational-affektive Merkmale und Leistung beeinflusst.

Insgesamt wird als Grund für gemischte Ergebnisse verschiedener Unterrichtsaktivitäten im Zusammenhang mit Interesse bzw. Leistung daher vermutet, dass motivationsunterstützende Lehr-Lernbedingungen unzureichend umgesetzt werden und dadurch bedingt ein Mangel an selbstbestimmtem Lernen entsteht (vgl. Rakoczy, 2008). Eine inhaltliche Gestaltung der Unterrichtsaktivitäten hin zu konzeptuell anregendem Unterricht, der das Vorwissen und die Lebenswelt der Schülerinnen und Schüler berücksichtigt, wird dagegen mit positiver Motivations- und Leistungsentwicklung in Verbindung gebracht (Baumert et al., 2000; Kunter, 2005; Turner et al., 1998).

Wie motivationsunterstützend forschend-entdeckender Unterricht erlebt wird, hängt jedoch auch von den individuellen Voraussetzungen der Schülerinnen und Schüler ab (Snow et al., 1996). Individuelle Schülervoraussetzungen wie Leistung, Selbstkonzept und Interesse beeinflussen die Entwicklung von Schülerinnen und Schülern (Weinert & Helmke, 2010) und sind wichtige Prädiktoren von Bildungsprozessen (Corno & Snow, 1986; Shuell, 1996). Darüber hinaus nehmen sie unter Umständen Einfluss auf die Wahrnehmung des Unterrichts (Wittwer, 2008). Die Unterrichtsforschung berücksichtigt sie daher häufig als Kontrollvariablen neben Unterrichtsmerkmalen (Seidel, Prenzel, Wittwer & Schwindt, 2007). Da Lehr-Lernprozesse als Interaktion zwischen Lehrperson bzw. Merkmalen des Unterrichtsangebots einerseits und Lernenden bzw. individuellen Schülervoraussetzungen andererseits erfolgen (Helmke, 2007; Seidel, 2014), können individuelle Schülervoraussetzungen jedoch auch als Interaktionsfaktoren gesehen werden (Seidel, Jurik, Häusler & Stubben, 2016). So beeinflussen individuelle Schülervoraussetzungen Lernumgebungen und umgekehrt (z. B. Perry, Turner & Meyer, 2006; Shuell, 1996).

Die Interaktion von individuellen Schülervoraussetzungen und Unterricht kann sowohl direkt im Verhalten von Lehrperson oder Schüler deutlich werden als auch lediglich die Wahrnehmung des Unterrichts durch Schülerinnen und Schüler beeinflussen. Im Hinblick auf das Verhalten zeigen Untersuchungen zur Lehrer-Schüler-Interaktion beispielsweise, dass Lehrer Jugendlichen, die sich im Unterricht stärker engagieren, mehr Beachtung, Strukturierung und Autonomieunterstützung entgegenbringen (Skinner & Belmont, 1993). Ähnlich findet sich in einer Videostudie zum Mathematikunterricht, dass

Lehrer mit leistungsstarken Jugendlichen tendenziell häufiger interagieren (Lipowsky, Rakoczy, Pauli, Reusser & Klieme, 2007). Lehrer loben Jugendliche, an die sie positive Leistungserwartungen haben, zudem häufiger für richtige Unterrichtsbeiträge und kritisieren Schülerinnen und Schüler, mit denen sie negative Erwartungen verbinden, häufiger für falsche Beiträge (Brophy & Good, 1970). Schülerinnen und Schüler, die von der Lehrperson als leistungsstark eingeschätzt werden, werden darüber hinaus häufiger aufgerufen (Good, Cooper & Blakey, 1980; Good & Brophy, 1990).

Auf das Verhalten von Schülerinnen und Schüler bezogen zeigt sich, dass Jugendliche mit hoher Leistung, hohem Selbstkonzept und hohem Interesse sich im Physikunterricht häufiger, länger und intensiver beteiligen als Jugendliche, die sich in ihrer Leistung unterschätzen, und Jugendliche, die eine niedrige Leistung, niedriges Selbstkonzept und niedriges Interesse aufweisen (Jurik et al., 2013). Individuelle Schülervoraussetzungen wie Leistung, Selbstkonzept und Interesse sagen darüber hinaus Lernaktivitäten von Schülerinnen und Schülern vorher und moderieren den Einfluss von Unterricht auf Lernaktivitäten (Jurik et al., 2014). So werden insbesondere sich unterschätzende Schülerinnen und Schüler in Physik durch vermehrte Verständnisfragen zu häufigeren positiven Lernaktivitäten angeregt (Jurik et al., 2014).

Neben dem Verhalten führen individuelle Schülervoraussetzungen zu einer unterschiedlichen Wahrnehmung des Unterrichts. So argumentiert Hartinger (2005), dass Jugendliche mit einer höheren Kompetenzüberzeugung eher Entscheidungsmöglichkeiten wahrnehmen können und ihren Unterricht dadurch autonomieunterstützender empfinden. Ähnlich vermutet Rakoczy (2008), dass Schülerinnen und Schüler, die ihre Leistung höher einschätzen, im Unterricht häufiger Fragen beantworten und das Unterrichtsgeschehen beeinflussen können. Verschiedene Studien konnten die Moderation von Schülervoraussetzungen wie der wahrgenommenen Kompetenzunterstützung belegen und können in variablenzentrierte und personenzentrierte Studien eingeteilt werden. So zeigt sich in einer variablenzentrierten Betrachtung, dass Schülerinnen und Schüler mit einer höheren Einschätzung ihrer eigenen Kompetenz eine höhere Motivationsunterstützung berichten (Rakoczy, 2008). Hänze und Berger (2007) untersuchen einen möglichen Moderationseffekt von Selbstkonzept auf den Zusammenhang von kooperativem Unterricht und Kompetenz- bzw. Autonomieunterstützung sowie sozialer Eingebundenheit – ebenfalls in einem variablenzentrierten Vorgehen. Sie nehmen an, dass insbesondere Schülerinnen und Schüler mit niedrigem Selbstkonzept von kooperativem

im Gegensatz zu traditionellem Physikunterricht profitieren und eine stärkere Kompetenzunterstützung berichten. In einem experimentellen Design, in dem kooperatives Lehren und Lernen in Form eines Gruppenpuzzles direkter Instruktion gegenübergestellt wurde, bestätigten sie die Mediation des Zusammenhangs von kooperativem Unterricht und intrinsischer Motivation durch motivationsunterstützende Lehr-Lernbedingungen. Auch die Moderationshypothese können sie belegen und zeigen, dass Schülerinnen und Schüler mit positivem Selbstkonzept keine signifikanten Differenzen in der wahrgenommenen Kompetenzunterstützung zwischen beiden Instruktionmethoden (kooperativ und direkt) berichten, Schülerinnen und Schüler mit negativem Selbstkonzept jedoch in kooperativem Unterricht eine signifikant höhere Kompetenzunterstützung wahrnehmen als bei der direkten Instruktion.

Deutlich wird diese unterschiedliche Unterrichtswahrnehmung in Abhängigkeit von Selbstkonzept und Interesse sowie Leistung in personenzentrierten Analysen (z. B. Linnenbrink-Garcia, Pugh, Koskey, Kristin L. K. & Stewart, 2012b; Seidel, 2006). Individuelle Schülermerkmale wie Leistung, Selbstkonzept und Interesse stehen in Wechselwirkung zueinander und zeigen insbesondere als Merkmalskombination betrachtet einen Zusammenhang zur Wahrnehmung von Unterricht und Lernprozessen (Seidel, 2006; Trautwein et al., 2006). Personenzentrierte Analysen erlauben Gruppen innerhalb der Stichprobe, die sich durch Profile hinsichtlich Leistung, Selbstkonzept und Interesse in Naturwissenschaften auszeichnen, zu identifizieren (vgl. auch Kapitel 3.3) und finden zunehmend Anwendung in der Unterrichtsforschung (z. B. Hornstra, van der Veen, Peetsma & Volman, 2013; Linnenbrink-Garcia et al., 2012a; Seidel, 2006; Wormington et al., 2012).

In einer personenzentrierten Analyse zur IPN-Videostudie (z. B. Seidel, Prenzel, Rimmele et al., 2007) zeigte sich, dass Schülerinnen und Schüler mit aus motivationaler Sicht positiv ausgeprägten Leistung-Selbstkonzept-Interesse-Profilen (starkes und überschätzendes Profil, für eine Übersicht der Profile siehe Abbildung 2) ihren Unterricht in Physik innerhalb des Klassenverbandes motivationsunterstützender wahrnehmen als Schülerinnen und Schüler weniger positiv ausgeprägter Profile (Seidel, 2006). Weiterführende Analysen zum Profil der sich unterschätzenden Schülerinnen und Schüler ergeben zudem, dass differenzielle Entwicklungen dieser Subgruppe im Zusammenhang mit der Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen stehen: Sich unterschätzende Schülerinnen und Schüler mit einer positiven Selbstkonzeptentwicklung

innerhalb eines Schuljahres nehmen im Unterschied zu sich unterschätzenden Jugendlichen mit einer negativen bzw. keiner Selbstkonzeptentwicklung mehr Autonomieunterstützung, Kompetenzunterstützung und soziale Eingebundenheit wahr, sind stärker intrinsisch motiviert und verwenden tiefere Lernstrategien (Huber et al., 2015). Auch in einer neueren Studie, die an die IPN-Videostudie anknüpft und sie auf den Deutsch- und Mathematikunterricht erweitert, bestätigt sich, dass Interesse im Mathematikunterricht und Selbstkonzept im Deutschunterricht die Wahrnehmung der Unterrichtsqualität positiv vorhersagen (Jurik, Häusler, Stubben & Seidel, 2015).

Seidel (2006) prägt im Hinblick auf diesen Einfluss individueller Schülervoraussetzungen auf die Unterrichtswahrnehmung den Begriff der *Mikro-Lernumwelten*, die möglicherweise als eine Art Zwischenebene zwischen Klassen- und individueller Ebene unterschieden werden müssen (vgl. auch Shavelson & Seidel, 2006). Sie beschreiben, dass innerhalb einer Klassen verschiedene Mikro-Umgebungen existieren können, die dadurch entstehen, dass Schülerinnen und Schüler mit ähnlichen Voraussetzungen das Unterrichtsangebot ähnlich wahrnehmen und sich unterschiedliche Schülerprofile in der Unterrichtswahrnehmung systematisch unterscheiden (Seidel et al., 2016). Damit bilden die Mikro-Lernumwelten eine Art Zwischenebene zwischen Klassenebene und individueller Ebene, deren Berücksichtigung differenzierte Zusammenhänge offenbart (Seidel et al., 2016).

Wie bereits in Kapitel 3 erwähnt, ist das Geschlecht ein weiterer wesentlicher Prädiktor von Interesse an den Naturwissenschaften (z. B. Hannover, 1998; Hannover & Kessels, 2002, 2004; Hoffmann, 2002). Nur wenige der personenzentrierten Analysen berücksichtigen jedoch diesen Einfluss (für eine Ausnahme siehe Jurik et al., 2013; Lau & Roeser, 2008, vgl. auch Kapitel 3.3). Mädchen und Jungen verteilen sich einerseits unterschiedlich auf Profile unterschiedlicher Schülervoraussetzungen in Naturwissenschaften, indem Jungen häufiger zu den leistungsstarken, leistungszuversichtlichen und interessierten bzw. sich überschätzenden Profilen gehören und Mädchen häufiger den uninteressierten, sich unterschätzenden sowie „schwachen“ Profilen mit niedriger Leistung, Selbstkonzept und niedrigem Interesse angehören, wobei Naturwissenschaften und Mathematik sich ähneln, in Deutsch einige Unterschiede zu beobachten sind (Jurik et al., 2013, 2015). Andererseits beeinflusst das Geschlecht auch das Verhalten sowie die Unterrichtswahrnehmung und interagiert in diesem Zusammenhang zum Teil mit den Profilen hinsichtlich Leistung, Selbstkonzept und Interesse. Mädchen nehmen die

Unterrichtsqualität im Rahmen der Interaction Studie insgesamt positiver wahr als Jungen (Jurik et al., 2015). Wie Forschung zum Mathematikunterricht (Rakoczy, 2008) zeigt, verschwinden Geschlechterunterschiede in der Wahrnehmung der Unterrichtsqualität möglicherweise, wenn die Kompetenzeinschätzung als Einflussvariable berücksichtigt wird. Blickt man auf das Verhalten, so beteiligen sich Mädchen insgesamt verbal weniger am Physikunterricht als Jungen (Jurik et al., 2013). Differenzierte Analysen ergeben jedoch, dass dieser Zusammenhang von der Profiltugehörigkeit moderiert wird, indem sich innerhalb des „starken“ Profils mit positiven Schülervoraussetzungen Mädchen häufiger am Unterricht beteiligen als Jungen (Jurik et al., 2013). Da das starke Profil einen Teil der leistungsstarken Jugendlichen in Physik verkörpert, bedeutet dies, dass innerhalb der High Performer in Naturwissenschaft das Geschlecht möglicherweise den Zusammenhang unterschiedlicher Profile hinsichtlich Selbstkonzept und Interesse und der Unterrichtseteiligung – bzw. als vermittelnden Faktor die Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen – moderiert.

Ausgehend von Hinweisen auf leistungsstarke Schülerinnen und Schüler aus den personenzentrierten Analysen kann insgesamt davon ausgegangen werden, dass sich High Performer in Naturwissenschaften mit unterschiedlichen Selbstkonzept-Interesse-Profilen ebenso wie Schülerinnen und Schüler allgemein in ihrer Unterrichtswahrnehmung unterscheiden und sich unterschiedlich in ihrer Kompetenz, Autonomie und inhaltlichen Relevanz unterstützt fühlen. Unterschiedliche Selbstkonzept-Interesse-Profile beeinflussen darüber hinaus möglicherweise den Zusammenhang einzelner Unterrichtsaktivitäten mit der wahrgenommenen Motivationsunterstützung. Beispielsweise könnten sich analog der Ergebnisse von Hänze und Berger (2007) wenig leistungszuversichtliche High Performer bei häufigerem forschend-entdeckendem Unterricht (im Vergleich zu wenig forschend-entdeckendem Unterricht) stärker kompetenzunterstützt fühlen als bereits leistungszuversichtliche High Performer. Dies würde Hinweise darauf liefern, dass insbesondere wenig leistungszuversichtliche High Performer von forschend-entdeckendem Unterricht profitieren. Bisher ist jedoch weitgehend unklar, inwiefern High Performer unterschiedlicher Profile allgemein bzw. innerhalb des Klassenverbands ihren Unterricht unterschiedlich motivationsunterstützend wahrnehmen, d. h. unterschiedliche Mikro-Lernumgebungen für High Performer existieren, und ob sich der Zusammenhang von Unterrichtsaktivitäten und motivationsunterstützenden Lehr-Lernbedingungen zwischen den Profilen unterscheidet.

4.4. Erfassung der Unterrichtswahrnehmung von Schülerinnen und Schülern

In den vorangegangenen Kapiteln wurde beschrieben, welche Unterrichtsaktivitäten im Naturwissenschaftsunterricht eine funktionale Rolle bei der Förderung von Kompetenz und Interesse von High Performern einnehmen, mögliche Unterschiede zu Nicht-High-Performern beschrieben und Forschungslücken in diesem Bereich aufgezeigt. Darüber hinaus wurde erläutert, dass diese Unterrichtsaktivitäten als Angebot für Lernprozesse gesehen werden müssen und die Nutzung dieser Lerngelegenheiten von der individuellen Wahrnehmung der Schülerinnen und Schüler abhängt. Auch hier wurden offene Punkte aufgezeigt, die in dieser Arbeit behandelt werden sollen. Dies führt jedoch zu der Frage, aus welcher Perspektive Unterricht für die Untersuchung dieser Forschungslücken erfasst, eingeschätzt und untersucht werden soll. Verschiedene Methoden wie die Befragung von Schülerinnen und Schülern oder Lehrpersonen sowie die Videobeobachtung von Unterricht dienen als Verfahren zur Rekonstruktion von Unterricht (z. B. Prenzel & Lankes, 2013). Für die in dieser Arbeit behandelten Fragestellungen wurden Schülereinschätzungen zur Unterrichtserfassung genutzt. Nachfolgend wird beschrieben, inwieweit sie methodisch und für unterschiedliche Unterrichtsmerkmale geeignet sind.

Wie in Kapitel 4.1 näher dargestellt, wird die Wirkung von Unterricht maßgeblich durch die Wahrnehmung des Unterrichtsangebots durch die Schülerinnen und Schüler bestimmt. Schülereinschätzungen werden in der Unterrichtsforschung daher sehr häufig verwendet (Fraser, 1998; Levy, Wubbels, den Brok & Brekelmans, 2003; Wagner et al., 2016). Objektives Angebot und subjektive Wahrnehmung durch die Schülerinnen und Schüler stimmen jedoch nicht zwingend überein (vgl. Clausen, 2002; Kunter & Baumert, 2006). Verschiedene Studien haben Schülereinschätzungen deshalb aus unterschiedlichen methodischen Perspektiven untersucht und sich beispielsweise mit der Übereinstimmung zwischen Lehrpersonen, Schülern und geschulten Beobachtern (z. B. Clausen, 2002; Kunter & Baumert, 2006; Wagner et al., 2016), der zeitlichen Kohärenz (z. B. Wagner et al., 2016), dem Vorhersagewert für Lernergebnisse (z. B. Kunter & Baumert, 2006; Wagner et al., 2016), der Konstruktvalidität (z. B. Clausen, 2002; Wagner, Göllner, Helmke, Trautwein & Lüdtke, 2013) und der Reliabilität und Interaktion der Unterrichtseinschätzung mit individuellen Schülermerkmalen (z. B. Wittwer, 2008) beschäftigt.

Dabei zeigt sich, dass Lehrpersonen und Schülerinnen und Schüler nur mittel bis gering in ihrer Unterrichtseinschätzung übereinstimmen (Kunter & Baumert, 2006). Dies trifft auch bei verhaltens- und ereignisnahen Skalen zu und die Ergebnisse legen nahe, dass sich Lehrer und Schüler bei der Einschätzung auf unterschiedliche Aspekte konzentrieren und beide Gruppen für unterschiedliche Perspektiven des Unterrichts valide Einschätzungen liefern (Kunter & Baumert, 2006).

Untersuchungen der Konstruktvalidität ergeben darüber hinaus, dass Schülerinnen und Schüler verlässlich verschiedene Dimensionen von Unterricht unterscheiden und einschätzen können (Wagner et al., 2013). Eine Generalisierung über Klassen und Fächer konnte für einzelne Merkmale (Klassenmanagement und Struktur) nachgewiesen werden, für andere nicht (Verständlichkeit, Motivation und Einbindung der Schülerinnen und Schüler). Andere Studien zeigen jedoch auch, dass sowohl Lehrer als auch Schülerinnen und Schüler nur gering mit der Unterrichtseinschätzung geschulter Beobachter übereinstimmen (Clausen, 2002; de Jong & Westerhof, 2001).

Sowohl Lehrer als auch Schüler schätzen ihren Unterricht in einem Zeitraum von drei Monaten jedoch in mittlerem bis hohem Ausmaß zeitlich konsistent ein (Wagner et al., 2016). Differenzierte Analysen ergeben, dass einzelne Messzeitpunkte bereits 92 % der Varianz zwischen Schulklassen aufklären und in vielen Fällen daher Schülereinschätzungen des Unterrichts zu einzelnen Messzeitpunkten für eine allgemeine Unterrichtsbeschreibung ausreichen (Wagner et al., 2016, S. 716).

Im Hinblick auf die Vorhersagekraft in Bezug auf Leistung und Selbstkonzept erwiesen sich vor allem die zeitlich konsistenten Schülereinschätzungen als prädiktiv – wenngleich sie deshalb nicht als objektiver angesehen werden können –, während Einschätzungen zu einzelnen Messzeitpunkten nur teilweise Leistung und Selbstkonzept vorhersagen (Wagner et al., 2016). Auch in den Untersuchungen im Rahmen von COACTIV zeigt sich eine begrenzte Vorhersagekraft der Unterrichtseinschätzungen für Leistung (Kunter & Baumert, 2006). Insgesamt kommen verschiedene methodische Studien zu dem Schluss, dass die Wahrnehmung der Schülerinnen und Schüler grundsätzlich eine prädiktive Unterrichtseinschätzung liefern kann, der gewinnbringende und methodisch angemessene Nutzen dieser Schülereinschätzungen jedoch von der Fragestellung und den Unterrichtsmerkmalen abhängt (Kunter & Baumert, 2006; Prenzel & Lankes, 2013; Wagner et al., 2016).

Die methodische Eignung von Schülereinschätzungen des Unterrichts hängt auch

entscheidend von der inhaltlichen Fragestellung und den verwendeten Unterrichtsmerkmalen ab (Clausen, 2002; Prenzel & Lankes, 2013). Eine grundsätzliche Frage, die sich dabei stellt ist: „Will man beschreiben, wie der Unterricht ist – oder wie er wahrgenommen beziehungsweise erlebt wird?“ (Prenzel & Lankes, 2013, S. 95). Innerhalb der Unterrichtsmerkmale kann dafür zwischen Sicht- und Tiefenstrukturen unterschieden werden (Oser & Baeriswyl, 2001). Sichtstrukturen beschreiben eher formale, organisatorische und methodische Elemente des Unterrichts und werden zur Erfassung von Unterrichtsaktivitäten und -Skripts verwendet (Hugener et al., 2009; Seidel, 2003; Seidel et al., 2006). Tiefenstrukturen sind weniger leicht objektiv beobachtbare qualitative Merkmale des Unterrichts, die häufig zur Erklärung und Analyse der Wirkung des Unterrichts dienen (Clausen, 2002; Ditton, 2000). Motivationsunterstützende Lehr-Lernbedingungen wie Kompetenzunterstützung, Autonomieunterstützung oder inhaltliche Relevanz gehören beispielsweise zu den Tiefenstrukturen (z. B. Seidel, 2003).

Obwohl Sichtstrukturen beobachtbar sind, stimmen Schülerinnen und Schüler in ihrer Einschätzung dieser Merkmale nicht automatisch überein (z. B. Wittwer, 2008). Sie werden daher immer häufiger und zuverlässig durch Videobeobachtungen erfasst (Fischer, Labudde, Neumann & Viiri, 2014; Petko, Waldis, Pauli & Reusser, 2003; Seidel, Prenzel & Kobarg, 2005). Resultat dieser objektiveren Erfassung von Sichtstrukturen in Form von Videos ist jedoch zum Teil, dass geringere bzw. keine Zusammenhänge mehr mit Bildungsergebnissen wie Leistung beobachtet werden (z. B. Hugener et al., 2009).

Die Reliabilität und Verlässlichkeit von Schülereinschätzungen von Sichtstrukturen kann jedoch auch optimiert werden, beispielsweise durch (1) das Design der Items und Skalen und (2) die Nutzung aggregierter Daten. Schülereinschätzungen von Sichtstrukturen von Unterricht sind vor allem dann verlässlich, wenn sie mit verhaltens- und ereignisnahen Skalen erfasst werden (Clausen, 2002). Die Erfassung der Unterrichtsaktivitäten im Naturwissenschaftsunterricht in PISA 2006 setzt diese Bedingung beispielsweise um, indem sie nach beobachtbaren Ereignissen in einem klaren Bezugsrahmen fragt (z. B. Itemstamm: „Wie oft kommen die folgenden Aktivitäten in deinem Physik-, Chemie-, Biologie- und Naturwissenschaftsunterricht vor?“, Item: „Die Schülerinnen und Schüler sollen Schlüsse aus einem Experiment ziehen, das sie durchgeführt haben.“) und mit einem konkreten häufigkeitsbasiertem Antwortformat („in allen Stunden“, „in den meisten Stunden“, „in manchen Stunden“, „nie oder fast nie“) kombiniert (Frey & Asseburg, 2009). Die so erfassten Unterrichtsmerkmale sind daher unabhängiger

von subjektiven Vergleichen (Kobarg et al., 2011). Schülereinschätzungen, die sich auf qualitative Meinungen und Eindrücke beziehen, sind im Unterschied dazu nicht zur reliablen Erfassung von Merkmalen zur Beschreibung des Unterrichts und vor allem nicht den Sichtstrukturen geeignet (Clausen, 2002).

Neben dem Design der Erfassungsinstrumente kann die Reliabilität von Schülereinschätzungen zu Sichtstrukturen verbessert werden, indem auf innerhalb von Schulklassen aggregierte Daten zurückgegriffen und damit die genestete Struktur von Schülereinschätzungen des Unterrichts berücksichtigt wird (Lüdtke, Robitzsch, Trautwein & Kunter, 2009). Auch hier besteht die Einzeleinschätzung eines Schülers sowohl aus Unterrichtsangebot als auch weiteren Einflüssen wie persönlichen Merkmalen, durch die Aggregation ergibt sich jedoch ein zuverlässiger Indikator für Unterrichtsprozesse (Ditton, 2002b; Kunter & Baumert, 2006; Marsh et al., 2009). Im Rahmen von PISA 2006 konnte Wittwer (2008) zeigen, dass Schülerinnen und Schüler einer Schulklasse die häufigkeitsbasiert und ereignisnah erfassten Unterrichtsaktivitäten insgesamt konsistent und reliabel einschätzen und aggregierte Klassenwerte sinnvoll verwendet werden können, wobei ein Teil der Konsistenz auf das ähnliche Leistungsniveau der Schülerinnen und Schüler durch die Aufteilung in Schularten zurückgeht. Insbesondere für häufigkeitsbasiert und ereignisnah erfassbare Sichtstrukturen wie die Anzahl von Schülerexperimenten oder die Häufigkeit von Klassendiskussionen eignen sich daher aggregierte Schülereinschätzungen. Im Hinblick auf dabei nutzbare Unterrichtsergebnisse weisen aktuelle Befunde darauf hin, dass Leistungsmerkmale häufig eine höhere Varianz auf Klassenebene aufweisen und daher eher durch aggregierte Unterrichtsmerkmale erklärt werden können als motivational-affektive Merkmale (vgl. Kunter, 2005; Rakoczy, 2008; Willems, 2011).

Tiefenstrukturen werden im Unterschied zu Sichtstrukturen aufgrund ihrer subjektiven Natur und ihrer Abhängigkeit von individuellen Wahrnehmungen üblicherweise in Form von Schülereinschätzungen erfasst (z. B. Kunter, 2005; Rakoczy et al., 2008; Seidel, Prenzel & Kobarg, 2005). Durch Schülereinschätzungen erfasste Tiefenstrukturen, darunter motivationsunterstützende Lehr-Lernbedingungen, haben sich in zahlreichen Studien als valide und reliabel erwiesen und sind Sichtstrukturen insbesondere in der Vorhersage motivational-affektiver Merkmale überlegen (z. B. Kramer, 2002; Kunter, 2005; Rakoczy et al., 2008; Seidel, 2003; Seidel, Prenzel, Rimmelé et al., 2007).

Insgesamt sind Schülereinschätzungen für Tiefenstrukturen wie motivationsun-

terstützende Lehr-Lernbedingungen, die sich auf die individuelle Wahrnehmung des Unterrichtsangebots unter Einfluss von Schülervoraussetzungen konzentrieren, valide und reliabel. Auch für verhaltensnahe Sichtstrukturen wie Unterrichtsaktivitäten kann gefolgert werden, dass aggregierte Schülereinschätzungen für die Abbildung dieser Unterrichtsmerkmale auf Klassenebene verlässlich und aussagekräftig eingesetzt werden können. In Anlehnung an Prenzel und Lankes (2013, S. 97) kann insgesamt argumentiert werden, dass die Perspektive der Schülerinnen und Schüler zur Einschätzung von Unterricht alleine nicht ausreicht und eine objektive Rekonstruktion von Unterricht den Einsatz von Videostudien erfordert, es im Umkehrschluss jedoch auch keinen Sinn macht, auf die „Informationsquelle „Schülerinnen und Schüler“ zu verzichten“. Für eine gewinnbringende Verwendung sei stattdessen entscheidend, dass Antwortperspektiven und damit verbundene eingeschränkte bzw. fokussierte Sichtweisen bewusst berücksichtigt werden und die jeweiligen Messinstrumente überlegt und in Passung zur Fragestellung eingesetzt werden.

5. Fragestellungen

Ziel dieser Arbeit ist es erstens, die Gruppe der leistungsstarken Jugendlichen in Naturwissenschaften zu charakterisieren. Dafür werden besondere Merkmale, in denen sie sich von den übrigen Schülerinnen und Schülern (*Nicht-High-Performer*) unterscheiden, identifiziert und es wird untersucht, wie heterogen die Gruppe der leistungsstarken Jugendlichen selbst in diesen Merkmalen ist. Zweitens sollen Hinweise für eine Förderung von Leistung, Selbstkonzept und Interesse leistungsstarker Jugendlicher im Naturwissenschaftsunterricht ermittelt werden. Ein besonderer Fokus vertiefter Analysen liegt dabei auf der Betrachtung motivational-affektiv indifferenter leistungsstarker Schülerinnen und Schüler und möglichen Zusammenhängen ihrer Merkmale mit der Motivationsunterstützung im Unterricht. Als methodischer Zugang wurde dabei die Analyse der internationalen Vergleichsstudien PISA 2006 (OECD, 2007) und PISA 2012 (OECD, 2013a) gewählt. Diese Arbeit beschäftigt sich dabei mit den drei übergreifenden Fragestellungen:

1. Durch welche individuellen, familiären und institutionellen Merkmale zeichnen sich leistungsstarke Jugendliche in Naturwissenschaften aus?
2. Inwieweit kann die Häufigkeit einzelner Unterrichtsaktivitäten die naturwissenschaftliche Kompetenz und das Interesse von High Performern vorhersagen und bestehen Unterschiede zu den übrigen Schülerinnen und Schülern?
3. Lassen sich innerhalb der leistungsstarken Jugendlichen unterschiedliche Selbstkonzept-Interesse-Typen finden und unterscheiden sich High Performer unterschiedlicher Typen darin, wie motivationsunterstützend sie ihren Unterricht wahrnehmen?

5.1. Charakterisierung von High Performern in Naturwissenschaft

Gegenstand der ersten übergreifenden Fragestellung ist die Charakterisierung der leistungsstarken Jugendlichen in Naturwissenschaften. Sie erfolgt anhand individueller, familiärer und institutioneller Merkmale. Leitende Fragen sind dabei (a) Welche besonderen Merkmale zeichnen High Performer im Vergleich zu Nicht-High-Performern aus? und (b) Wie heterogen ist die Gruppe der High Performer selbst in diesen Merkmalen? Über verschiedene Faktoren hinweg wird darüber hinaus die Frage, inwieweit die High Performer aus PISA 2006 mit den High Performern aus PISA 2012 in ihren Merkmalen vergleichbar sind, behandelt.

5.1.1. Individuelle Merkmale

In der Literatur findet man verschiedene individuelle Merkmale, die Prädiktoren von naturwissenschaftlicher Kompetenz sein bzw. Leistungsvarianz erklären können (vgl. Kapitel 2.2). Dazu gehören die Intelligenz (z. B. Schrader & Helmke, 2008) oder die Leistung in anderen Fachbereichen wie Mathematik und Lesen (z. B. Cano et al., 2014; Uhden et al., 2012).

Leistungen in Mathematik und Lesen

Informationen zur Leistung von naturwissenschaftlichen High Performern in anderen Fachbereichen und Überlappungen der High Performer in Naturwissenschaften, Mathematik und Lesen können einerseits Hinweise für die Identifikation von High Performern geben und zeigen, ob mit High Performance in Naturwissenschaften möglicherweise auch hohe Leistungen in anderen Fachbereichen einhergehen. Andererseits kann bei Schülerinnen und Schülern, die in mehreren Domänen hohe Leistungen erreichen, angenommen werden, dass ihr Interessenprofil nicht allein aufgrund der Leistung naturwissenschaftlich geprägt ist, sondern Naturwissenschaften vielmehr mit anderen Fächern konkurrieren. Erkenntnisse über die Anzahl und die Eigenschaften dieser Gruppe könnten daher in motivational-affektiven Förderbemühungen aufgenommen werden.

Aus bisherigen Forschungsprojekten ist bekannt, dass naturwissenschaftliche Kompetenz und mathematische Kompetenz bzw. Lesekompetenz korrelieren (OECD, 2007)

und Schülerinnen und Schüler auf den beiden obersten Kompetenzstufen in Naturwissenschaften in PISA überwiegend auch in Mathematik oder Lesen die beiden obersten Kompetenzstufen erreichen (OECD, 2009b). Für die Gruppe der besten 20% in Naturwissenschaften ist bisher aber unbekannt, ob sie auch in Mathematik und Lesen höhere Leistungen als die übrigen Schülerinnen und Schüler erreichen und wie viele dieser Jugendlichen auch in Mathematik und Lesen zu den leistungsstärksten Schülerinnen und Schülern gehören. Offen ist zudem, welche Schulleistungen High Performer in Naturwissenschaften, Mathematik und Deutsch im Vergleich zu ihrer Leistung in PISA erreichen und, welche Leistungsprofile sich bei den Schulnoten finden lassen. Folgende Forschungsfragen werden daher untersucht:

1. Wie leistungsstark sind die High Performer in Naturwissenschaften in anderen Fachbereichen und bestehen Unterschiede zu Nicht-High-Performern?
2. Welcher Anteil der High Performer gehört auch in Mathematik oder Lesen bzw. allen drei Domänen zu den High Performern und welcher Anteil hat ein naturwissenschaftliches Leistungsprofil?
3. Welche Schulleistungen erreichen High Performer in Naturwissenschaften und welche Notenprofile lassen sich unter High Performern in Naturwissenschaften, Mathematik und Deutsch finden?

Hypothesen:

High Performer in Naturwissenschaften erreichen im Mittel auch in Mathematik und Lesen höhere Kompetenzwerte als Nicht-High-Performer. Ähnlich zu den Befunden für Top Performer (OECD, 2009b) gehört ein großer Teil der High Performer auch in einer oder zwei anderen Domänen zu den High Performern. Die Überlappungen sind insgesamt jedoch geringer ausgeprägt als bei den Top Performern.

High Performer erreichen gute Schulleistungen und im Mittel bessere Schulnoten als Nicht-High-Performer. Ausgehend von der Annahme, dass Naturwissenschaftsaufgaben in der Schule vergleichsweise häufiger mathematisch geprägt sind und seltener Textinformationen enthalten als in standardisierten Kompetenztests, wird erwartet, dass innerhalb der High Performer häufig ein mathematisch-naturwissenschaftliches Profil und seltener ein naturwissenschaftlich-sprachliches oder sprachliches Notenprofil vorkommt.

Kognitive Fähigkeiten

Inwiefern eine hohe kognitive Grundfähigkeit eine Voraussetzung für hohe naturwissenschaftliche Kompetenz ist, hat sowohl für die Identifikation als auch dafür, wie relevant Gelegenheiten für Erfahrung und Übung in Naturwissenschaften für High Performance sind, Bedeutung.

Die Expertiseforschung stimmt für spezialisierte Leistungen darin überein, dass zwar ein gewisses Maß an Intelligenz für Lernprozesse hin zu hohen Leistungen notwendig ist, höhere kognitive Fähigkeiten jedoch nicht zwingend zu Leistungsexpertise führen (z. B. Krajewski & Schneider, 2006). Es existieren jedoch nur wenige domänen-spezifische Ergebnisse zu Naturwissenschaften und es ist unklar, ob High Performer eine höhere Intelligenz aufweisen als Nicht-High-Performer bzw., ob auch Schülerinnen und Schüler mit niedrigerer Intelligenz eine hohe naturwissenschaftliche Kompetenz erreichen. Folgende Fragestellungen werden daher behandelt:

4. Unterscheiden sich High Performer in ihren kognitiven Grundfähigkeiten von Nicht-High-Performern?
5. Welche Spannweite der kognitiven Grundfähigkeiten lässt sich innerhalb der High und Nicht-High-Performer finden und geht eine überdurchschnittliche kognitive Grundfähigkeit durchweg mit High Performance einher?

Hypothesen:

High Performer zeigen im Durchschnitt eine höhere kognitive Fähigkeit als Nicht-High-Performer. Innerhalb der High Performer zeigt sich dennoch eine gewisse Varianz in der kognitiven Fähigkeit. High Performer erreichen demnach zwar einen gewissen Schwellenwert kognitiver Grundfähigkeiten, hohe kognitive Fähigkeiten sind jedoch nicht zwingend mit High Performance in Naturwissenschaften verbunden.

Geschlecht

Mit der Betrachtung des Geschlechts in Verbindung mit High Performance in Naturwissenschaften können Hinweise darauf gefunden werden, ob Mädchen und Jungen gleich wahrscheinlich hohe Leistungen in Naturwissenschaften erreichen oder Unterschiede

beobachtet werden, die möglicherweise auf das soziale Geschlecht bzw. damit verbundene Wirkmechanismen auf Einstellung, Sichtweisen und Verhalten von Mädchen (z. B. Hannover & Kessels, 2004) zurückgehen.

Neben den erwähnten allgemeinen Wirkmechanismen des sozialen Geschlechts, die auftreten können, wurde in PISA gezeigt, dass Mädchen und Jungen sich generell nicht in ihrer naturwissenschaftlichen Kompetenz unterscheiden aber Unterschiede in den Teilkompetenzen bestehen und Jungen häufiger die beiden obersten Kompetenzstufen (Top Performer) erreichen als Mädchen (OECD, 2009b). Unklar ist, ob diese Unterschiede auch unter den High Performern gefunden werden können und High Performer in Geschlechterunterschieden mit der Leistungsspitze oder der ganzen Population vergleichbar sind.

6. Erreichen unter Mädchen und Jungen unterschiedliche Anteile High Performance in Naturwissenschaften?

Hypothesen:

Unter den Jungen erreicht ein höherer Anteil als unter den Mädchen High Performance in Naturwissenschaften. High Performer in Naturwissenschaften ähneln damit in Geschlechterunterschieden den Top Performern in Naturwissenschaften (OECD, 2009b).

5.1.2. Familiäre und institutionelle Merkmale

Neben individuellen Eigenschaften prägen familiäre Determinanten wie der sozioökonomische Hintergrund und institutionelle Rahmenbedingungen wie die Schulart und die Klassenzusammensetzung die Leistungsentwicklung (vgl. Kapitel 2.2.2).

Sozioökonomischer Hintergrund

Erkenntnisse dazu, ob sich High Performer in ihrem sozioökonomischen Hintergrund von Nicht-High-Performern unterscheiden, zeigen, ob die familiäre Lernumwelt und damit verbundene Wirkmechanismen durch Ressourcen, Einstellungen oder Übertrittsentscheidungen (z. B. Baumert et al., 2006b) möglicherweise Einfluss darauf nehmen, welche Chancen Jugendliche haben, hohe naturwissenschaftliche Kompetenz zu erreichen. Umgekehrt können sie jedoch auch Stärken des Bildungssystems aufzeigen, indem

sie möglicherweise sichtbar machen, dass Jugendliche unterschiedlicher sozialer Hintergründe High Performance erlangen.

Large Scale Assessments konnten bereits zeigen, dass Top Performer in Naturwissenschaften einen höheren sozioökonomischen Hintergrund aufweisen als leistungsschwächere Schülerinnen und Schüler (z. B. OECD, 2009b). Diese Ergebnisse wurden bisher jedoch noch nicht auf High Performer erweitert. Darüber hinaus beschränken sich bisherige Ergebnisse zu Top Performern in Naturwissenschaften auf eindimensionale Indizes zum sozioökonomischen Hintergrund auf Basis verschiedener Einflussfaktoren wie Einkommen oder Bildungsabschluss und ermöglichen keine qualitative Einteilung der Berufe der Eltern. Bisher ist daher unbekannt, wie sich leistungsstarke Schülerinnen und Schüler in Naturwissenschaften auf unterschiedliche soziale Klassen verteilen und inwiefern beispielsweise Schülerinnen und Schüler aus bildungsfernen Schichten zu den Top und High Performern gehören. Als Fragen werden daher untersucht:

7. Weisen High Performer einen höheren sozialen Hintergrund auf als Nicht-High-Performer?
8. Wie heterogen sind High Performer in ihrem sozialen Hintergrund und in welchem Ausmaß erreichen Schülerinnen und Schüler aus niedrigen sozialen Lagen High Performance in Naturwissenschaften?

Hypothesen:

High Performer kommen auf Basis des HISEI im Mittel aus höheren sozialen Lagen als Nicht-High-Performer. Bei der Betrachtung unterschiedlicher sozialer Schichten auf Basis der EGP-Klassen sind nicht-akademische Elternhäuser unter den High Performern zwar seltener vertreten als unter den Nicht-High-Performern. Dennoch ist auch innerhalb der High Performer eine gewisse Varianz zu beobachten und auch Schülerinnen und Schüler aus nicht-akademischen Bildungshäusern erreichen High Performance in Naturwissenschaften.

Schulart und Klassenkomposition

Unter den institutionellen Bedingungsfaktoren stehen vor allem die Schulartzugehörigkeit und die Klassenzusammensetzung in Zusammenhang mit der naturwissenschaftlichen Kompetenz (vgl. Kapitel 2.2.2). Im deutschen Bildungssystem hat vor allem das

Gymnasium den Bildungsauftrag, leistungsstarke Jugendliche zu fördern (vgl. Heller, 2002a). Der institutionelle Kontext leistungsstarker Schülerinnen und Schüler auf Ebene der Schulklasse wird darüber hinaus dadurch charakterisiert, wie sich High Performer auf Klassen verteilen und welche Klassenkomposition in Schulklassen mit unterschiedlich vielen High Performern zu finden ist.

Im Rahmen von Large-Scale-Assessments wurde bereits gezeigt, dass vielseitig hochkompetente Jugendliche (besten 10% in Naturwissenschaften, Mathematik und Lesen) fast ausschließlich ein Gymnasium besuchen (Zimmer et al., 2007) und Klassen, denen Top Performer in Naturwissenschaften angehören, eine höhere Leistung aufweisen (OECD, 2009b). Auch hier ist jedoch offen, wie sich High Performer als breitere Leistungsspitze auf Schularten aufteilen. Darüber hinaus wurde bisher lediglich die mittlere Klassenleistung von Top Performern betrachtet und nicht untersucht, welche Anteile leistungsstarker Jugendlicher Klassen aufweisen und, mit welcher Klassenkomposition ein höherer High-Performer-Anteil einhergeht. Zur Analyse dieser institutionellen Merkmale werden daher folgende Forschungsfragen untersucht:

9. Wie verteilen sich High Performer und Nicht-High-Performer auf Schularten?
10. Wie verteilen sich High Performer auf Schulklassen?
 - 10.1. Kommen sie als gehäufte oder verteilte Gruppe in Klassen vor?
 - 10.2. Kommen High Performer eher in leistungsstarken oder leistungsschwachen und eher in homogenen oder heterogenen Klassen vor?

Hypothesen:

Ein Großteil der High Performer besucht ein Gymnasium, während ein Großteil der Nicht-High-Performer nicht-gymnasiale Schularten besucht. Auf Klassenebene verteilen sich High Performer am Gymnasium relativ gleichmäßig auf Schulklassen, kommen in einigen jedoch gehäuft bzw. in kleiner Zahl vor. Schulklassen mit einem hohen Anteil an High Performern gehören tendenziell zu den leistungsstarken Klassen und sind vergleichsweise heterogen.

5.2. Unterrichtsaktivitäten im Zusammenhang mit naturwissenschaftlicher Kompetenz und Interesse von High Performern

Die Lehr-Lernforschung der letzten Jahrzehnte zeigt, dass Unterrichtsaktivitäten wie interaktives Lehren und Lernen, forschend-entdeckender Unterricht sowie Anwendungsbezüge die Entwicklung naturwissenschaftlicher Kompetenz und naturwissenschaftsbezogenem Interesse unterstützen (vgl. Kapitel 4). Interaktives Lehren und Lernen wird aufgrund der Unterstützung von Elaborationsprozessen (O'Donnell, 2006), tieferem Verständnis (N. M. Webb, 2009), Entwicklung von Fachsprache (Bransford & Donovan, 2005), Argumentation (Osborne, 2012) und kritischem Diskurs (Henderson et al., 2015) aus theoretischer Sicht als förderlich für High Performer angenommen. Forschungsbemühungen bestätigen auch empirisch, dass leistungsstarke Jugendliche von kooperativem Lernen in heterogenen Gruppen profitieren (Slavin, 1990a, 1990b). Insgesamt liegen jedoch nur wenige Studien mit zumeist geringen Stichprobenzahlen und nur begrenzt vergleichbaren Operationalisierungen von interaktivem Lehren und Lernen vor (Neber et al., 2001).

Betrachtet man Unterschiede zwischen High- und Nicht-High-Performern im Zusammenhang von interaktivem Lehren und Lernen und naturwissenschaftlicher Kompetenz, so spricht ein höheres Vorwissen (z. B. Piaget et al., 1996) sowie in Verbindung damit ein höheres verbales Engagement (Jurik et al., 2013) für Vorteile zugunsten von High Performern. Bisher wurde in Studien jedoch die Frage, ob High Performer von häufigerem interaktivem Naturwissenschaftsunterricht mehr profitieren als Nicht-High-Performer, nicht beantwortet.

Auch für forschend-entdeckenden Unterricht gibt die bestehende Forschungslage bereits Hinweise auf einen positiven Zusammenhang mit der naturwissenschaftlichen Kompetenz von High Performern (z. B. White & Frederiksen, 1998). Aus theoretischer Sicht sprechen Wissenserwerbs- und Informationsverarbeitungsprozesse und strukturelle sowie inhaltliche Offenheit von forschend-entdeckenden Lernumgebungen für eine Überlegenheit leistungsstarker Schülerinnen und Schüler in der Kompetenzentwicklung bei Schülerexperimenten (vgl. Kapitel 4.2.2), in empirischen Untersuchungen finden sich jedoch entweder keine Unterschiede (D. Schwartz et al., 2011) oder Unterschiede

zugunsten leistungsschwächerer Jugendlicher (J.-R. Wang et al., 2010; White & Frederiksen, 1998). Insgesamt liegt nur eine begrenzte Anzahl an Studien vor, die sich in untersuchten Altersbereichen, Definition leistungsstarker Schülerinnen und Schüler und der Gestaltung forschend-entdeckender Lernumgebungen unterscheidet. Eindeutige Aussagen können daher bisher nicht getroffen werden.

Für Anwendungsbezüge deuten lediglich weiterführende Analysen im Rahmen von PISA 2006 einen Zusammenhang mit naturwissenschaftlicher Kompetenz von High Performern an, beschränken sich jedoch auf Mittelwertsunterschiede und nutzen eine andere Definition leistungsstarker Schülerinnen und Schüler (OECD, 2009b). Ergänzt werden diese Ergebnisse indirekt durch Forschung zu situiertem Lernen, problembasierem Lernen und Enrichment bzw. Curriculumentwicklungen im Kontext leistungsstarker Schülerinnen und Schüler (vgl. Kapitel 4.2.3). Aufgrund der derzeitigen Forschungslage bleibt jedoch unklar, wie häufigere Anwendungsbezüge und naturwissenschaftliche Kompetenz von High Performern zusammenhängen. Für Vorteile von High Performern gegenüber Nicht-High-Performern beim Lernen aus Anwendungsbezügen sprechen aus theoretischer Sicht erneut bessere Wissenserwerbs- und Informationsverarbeitungsprozesse sowie Problemlösestrategien, empirisch wurde bisher jedoch nicht untersucht, inwiefern häufigere Anwendungsbezüge im Naturwissenschaftsunterricht in Deutschland bei High und Nicht-High-Performern mit unterschiedlichen Leistungen einhergehen.

Zum Zusammenhang von interaktivem Lehren und Lernen und Interesse liefern lediglich Studien zu Selbstkonzept und Einstellungen Hinweise auf positive Zusammenhänge, untersuchen jedoch nicht explizit Interesse (vgl. Kapitel 4.2.1). Dem positiven Zusammenhang widersprechend führen Hochbegabungsstudien motivationsmindernde Effekte für leistungsstarke Jugendliche in heterogenen Gruppen auf. Wenngleich die empirische Evidenz beschränkt und uneindeutig ist, ist aus motivationstheoretischer Sicht davon auszugehen, dass interaktives Lehren und Lernen das Interesse von High wie Nicht-High-Performern unterstützt.

Da sich Forschung zu forschend-entdeckendem Lernen im Kontext leistungsstarker Schülerinnen und Schüler vorwiegend mit Leistung beschäftigt, wurde bisher nicht immer ein Zusammenhang mit Interesse nachgewiesen (vgl. Kapitel 4.2.2). Studien zu verwandten Bildungsergebnissen und Vorwissen konnten keinen Zusammenhang nachweisen (Stake & Mares, 2001; Taskinen et al., 2013). Für einen Zusammenhang sprechen jedoch Ergebnisse zur motivationsunterstützenden Lehr-Lernbedingungen im

Zusammenhang mit Interesse. Detaillierte Analysen zum Zusammenhang mit Interesse leistungsstarker im Vergleich zu leistungsschwächeren Schülerinnen und Schüler im alltäglichen Naturwissenschaftsunterricht standen jedoch bisher nicht im Fokus der Forschung zu forschend-entdeckendem Lernen.

Im Rahmen von PISA 2006 konnte bereits nachgewiesen werden, dass häufigere Anwendungsbezüge im Zusammenhang mit dem Interesse des obersten Leistungsquartils in Naturwissenschaften stehen (Prenzel & Schütte, 2007). Dies wird von der Annahme, dass Anwendungen die inhaltliche Relevanz des Naturwissenschaftsunterrichts erhöhen, bestätigt (vgl. Kapitel 4.2.3). Unterschiede zwischen High und Nicht-High-Performern sind nur dann denkbar, wenn Anwendungsbezüge zu anspruchsvoll oder zu weit weg von der Lebenswelt leistungsschwächerer Schülerinnen und Schüler gestaltet sind.

Insgesamt bleiben daher folgende Fragen zum Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz bzw. Interesse von High Performern offen:

11. Welche Rolle spielt die Klassenkomposition und der Naturwissenschaftsunterricht für die naturwissenschaftliche Kompetenz von High Performern?
 - 11.1. Lassen sich für die naturwissenschaftliche Kompetenz von High Performern Kompositionseffekte durch die kognitive Grundfähigkeit, den sozioökonomischen Hintergrund, die Anzahl der High Performer oder die Kompetenzstreuung der Klasse finden, d. h. erreichen High Performer in Klassen mit höherer kognitiver Grundfähigkeit, höherem sozioökonomischen Hintergrund, mehr High Performern oder kleinerer Leistungsstreuung höhere Kompetenzwerte?
 - 11.2. Sind gefundene Kompositionseffekte in gleicher Weise auch für Nicht-High-Performer zu beobachten?
 - 11.3. Inwieweit kann die von den Schülerinnen und Schüler wahrgenommene Häufigkeit von interaktivem Lehren und Lernen, forschend-entdeckendem Unterricht und Anwendungen vorhersagen, welche naturwissenschaftliche Kompetenz High Performer erreichen?
 - 11.4. Welche Unterschiede lassen sich im Vergleich zu Nicht-High-Performern finden?
12. Welche Rolle spielt der Naturwissenschaftsunterricht für das Interesse von High Performern an den Naturwissenschaften?

- 12.1. Inwieweit kann die von den Schülerinnen und Schülern wahrgenommene Häufigkeit von interaktivem Lehren und Lernen, forschend-entdeckendem Unterricht und Anwendungen vorhersagen, wie sehr sich High Performer für Naturwissenschaften interessieren?
- 12.2. Welche Unterschiede lassen sich im Vergleich zu Nicht-High-Performern finden?

Hypothesen:

Weder die Zusammensetzung einer Gymnasialklasse hinsichtlich der kognitiven Grundfähigkeit noch des sozioökonomischen Hintergrunds bewirken Kompositionseffekte im Zusammenhang mit der naturwissenschaftlichen Kompetenz von High und Nicht-High-Performern. Da High Performer von heterogenen Lerngruppen profitieren, dies jedoch von der Gestaltung der Lernumgebung abhängig ist (vgl. Kapitel 4.2.1), werden für die naturwissenschaftliche Kompetenz der High Performer keine oder positive Kompositionseffekte durch die Leistungsstreuung erwartet. Da in Gymnasialklassen mit zahlreichen High Performern möglicherweise eher eine Ausrichtung des Unterrichts an leistungsstarken Schülerinnen und Schülern, beispielsweise durch Akzeleration des Inhaltes, stattfindet, sind Kompositionseffekte durch die Anzahl der High Performer einer Klasse denkbar. Für die Unterrichtsaktivitäten wird erwartet, dass sowohl häufigeres interaktives Lehren und Lernen als auch häufigerer forschend-entdeckender Unterricht eine höhere naturwissenschaftlichen Kompetenz der High Performer in Naturwissenschaften vorhersagen. Beide Unterrichtsaktivitäten sind in ihrem Vorkommen gleichzeitig mit Interesse der High Performer an Naturwissenschaften assoziiert. Während häufigere Anwendungsbezüge mit einem höheren Interesse von High Performern zusammenhängen, ist der Zusammenhang zu naturwissenschaftlicher Kompetenz tendenziell positiv aber weniger eindeutig.

Der Vergleich mit Nicht-High-Performern ergibt, dass der Zusammenhang von häufigerem interaktivem Lehren und Lernen und naturwissenschaftlicher Kompetenz für High Performer stärker ausgeprägt ist. Zu Unterschieden im Zusammenhang zwischen forschend-entdeckendem Unterricht und naturwissenschaftlicher Kompetenz können keine eindeutigen Annahmen getroffen werden und Vorteile von High- wie Nicht-High-Performern sind denkbar. Bestehen Zusammenhänge von Anwendungsbezügen und naturwissenschaftlicher Kompetenz, so sind sie

tendenziell für leistungsstarke Jugendliche höher ausgeprägt.

Der Zusammenhang von häufigerem interaktivem Lehren und Lernen und Interesse an Naturwissenschaften wird für High Performer in einer höheren Ausprägung als für Nicht-High-Performer erwartet. Sowohl bei forschend-entdeckendem Lernen als auch Anwendungsbezügen wird angenommen, dass ein häufigeres Vorkommen bei beiden Leistungsgruppen ähnlich stark mit Interesse assoziiert ist.

5.3. Selbstkonzept-Interesse-Typen unter den High Performern und ihre Unterrichtswahrnehmung

5.3.1. Selbstkonzept-Interesse-Typen unter den High Performern

Immer häufiger werden in der Lehr-Lernforschung in den letzten beiden Jahrzehnten Merkmalskombinationen aus Leistung, Selbstkonzept und Interesse untersucht, da sie eine mehrdimensionale und anschauliche Betrachtung von Schülervoraussetzungen ermöglichen (vgl. Kapitel 3.3). Die dabei verwendeten personenzentrierten Analysen ermöglichen unter anderem, motivational indifferente Schülerinnen und Schüler, d. h. Schülerinnen und Schüler mit hoher Leistung aber niedrigem Selbstkonzept und/oder Interesse zu identifizieren. Bestehende Forschungsergebnisse deuten an, dass auch unter leistungsstarken Schülerinnen und Schülern unterschiedliche Selbstkonzept-Interesse-Profile, darunter motivational indifferente Gruppen, existieren (Seidel, 2006) und das Geschlecht im Zusammenhang mit Unterricht mit diesen Profilen interagiert (Jurik et al., 2013). Dies wurde bisher jedoch nur mit Analysen der Gesamtstichprobe untersucht. Personenzentrierte Analysen, die identifizieren, welche Selbstkonzept-Interesse-Typen innerhalb der leistungsstarken Jugendlichen existieren, wurden bisher nicht durchgeführt. Auch zu Einflüssen des Geschlechts auf die Häufigkeit und den Verlauf von Selbstkonzept-Interesse-Profilen von High Performern gibt es keine Evidenz. Theoretische Überlegungen und empirische Ergebnisse zu variablenzentrierten Zusammenhängen verschiedener Merkmale mit Selbstkonzept und Interesse lassen vermuten, dass Einflussvariablen wie der soziale Hintergrund (vgl. Kapitel 2.2.2) und Noten- bzw. Kompetenzprofile in Naturwissenschaften, Mathe und Deutsch (vgl. Kapitel 3.2) Gründe für motivational unterreichte Profilverläufe sein können. Darüber hinaus zeigen Vergleiche der in Studien

identifizierten Typen, dass sie sich unterschiedlich häufig auf Schulklassen verteilen (Seidel, 2006) und deuten damit an, dass auch der Unterricht oder die Klassenkomposition mögliche Einflussfaktoren sind. Bisher wurde jedoch nicht untersucht, inwiefern sich Selbstkonzept-Interesse-Typen unter den High Performern in diesen Merkmalen differenziell unterscheiden. Folgende Forschungsfragen werden daher untersucht:

13. Welche Selbstkonzept-Interesse-Typen lassen sich innerhalb der High Performer identifizieren?
14. Unterscheiden sich die Profilverläufe und Anteile der Selbstkonzept-Interesse-Typen für Mädchen und Jungen?
15. Zeichnen sich High Performer unterschiedlicher Selbstkonzept-Interesse-Typen durch besondere Merkmale aus?
 - 15.1. Unterscheiden sich High Performer unterschiedlicher Selbstkonzept-Interesse-Typen in ihrem sozioökonomischen Hintergrund?
 - 15.2. Unterscheiden sich High Performer unterschiedlicher Selbstkonzept-Interesse-Typen in ihrem Noten- bzw. Kompetenzprofil?
 - 15.3. Wie homogen bzw. heterogen verteilen sich High Performer unterschiedlicher Selbstkonzept-Interesse-Typen auf Schulklassen?

Hypothesen:

Unter den High Performern in Naturwissenschaften lassen sich unterschiedliche Selbstkonzept-Interesse-Typen identifizieren. Darunter sind ein Typ leistungszuversichtlicher und interessierter sowie ein Typ wenig leistungszuversichtlicher und uninteressierter High Performer (motivational indifferente High Performer) zu finden. Keine eindeutige Aussage kann darüber getroffen werden, ob sich die Selbstkonzept-Interesse-Typen weiblicher und männlicher High Performer im Verlauf unterscheiden. Es wird jedoch angenommen, dass weibliche High Performer signifikant häufiger Selbstkonzept-Interesse-Typen mit niedrigem Selbstkonzept und niedrigem Interesse angehören.

Für den Vergleich der Selbstkonzept-Interesse-Typen wird erwartet, dass High Performer eines Typs mit hohem Interesse unabhängig vom Selbstkonzept einen höheren sozioökonomischen Hintergrund aufweisen als High Performer eines

Typs mit niedrigem Interesse. Darüber hinaus wird angenommen, dass Jugendliche der Typen mit höherem Selbstkonzept und niedrigem bzw. hohem Interesse bessere Noten und höhere Kompetenzwerte in Naturwissenschaften aufweisen als Typen mit niedrigem Selbstkonzept und hohem bzw. niedrigem Interesse. Darüber hinaus ist davon auszugehen, dass High Performer eines Selbstkonzept-Interesse-Typs mit niedrigem Selbstkonzept, d. h. sich unterschätzende High Performer, häufiger in Klassen mit hoher Leistung vorkommen und/oder häufiger ein weniger naturwissenschaftsbezogenes, sondern eher ausgeglichenes oder verbales Kompetenz- und Notenprofil aufweisen als High Performer eines Selbstkonzept-Interesse-Typs mit hohem Selbstkonzept. Ähnlich zu bestehender Forschung wird angenommen, dass Schulklassen sich im Vorkommen der Typen unterscheiden und ein Teil beispielsweise einen hohen Anteil eines Typs aufweist, ein anderer Teil eine hohe Anzahl verschiedener Typen.

5.3.2. Wahrnehmung motivationsunterstützender Lehr- Lernbedingungen von High Performern unterschiedlicher Selbstkonzept-Interesse-Typen

Neben außerschulischen Merkmalen stellen auch Qualitätsmerkmale des Unterrichts mögliche Gründe dafür da, dass High Performer motivational nicht erreicht werden und Selbstkonzept-Interesse-Typen in Relation zur Leistung inkonsistente Ausprägungen von Selbstkonzept und Interesse aufweisen. Die Wahrnehmung motivationsunterstützender Lehr-Lernbedingungen wie Kompetenzunterstützung, Autonomieunterstützung und inhaltliche Relevanz beeinflusst als vermittelnder Prozess, welchen Zusammenhang Unterrichtsaktivitäten wie forschend-entdeckender Unterricht mit Lernmotivation und Interesse zeigen (vgl. Kapitel 3.4 und 4.3). Dies unterstützend konnte bereits gezeigt werden, dass Schülerinnen und Schüler unterschiedlicher Leistungs-Selbstkonzept-Interesse-Typen sich in ihrer motivationsunterstützenden Wahrnehmung des Unterrichts unterscheiden, d. h. unterschiedliche Mikro-Lernumgebungen existieren (Seidel, 2006) und eine positive Entwicklung von Jugendlichen eines sich unterschätzenden Typs mit höher berichteten motivationsunterstützenden Bedingungen einhergeht (Huber et al., 2015). Die genannten Analysen beruhen jedoch jeweils auf Typen der Gesamtstichprobe. Unklar bleibt deshalb, inwiefern sich die Wahrnehmung motivationsunterstützender Be-

dingungen von motivational indifferenten High Performern von der Wahrnehmung von High Performern leistungszuversichtlicher und/oder interessierter Typen unterscheidet. Erkenntnisse zu solchen Unterschieden werden in Form folgender Forschungsfragen untersucht:

16. Unterscheiden sich High Performer unterschiedlicher Selbstkonzept-Interesse-Typen darin, wie kompetenzunterstützend, autonomieunterstützend und inhaltlich relevant sie ihren Naturwissenschaftsunterricht wahrnehmen? Inwieweit lassen sich dabei Mikro-Lernumgebungen in Schulklassen identifizieren?
17. Sind die Unterschiede zwischen High Performern unterschiedlicher Selbstkonzept-Interesse-Typen innerhalb Klassen (Mikro-Lernumgebungen) von der Häufigkeit forschend-entdeckender Unterrichtsaktivitäten abhängig, d. h. wird häufigerer forschend-entdeckender Unterricht von High Performern unterschiedlicher Typen unterschiedlich stärker motivationsunterstützend erlebt?

Hypothesen:

High Performer unterschiedlicher Selbstkonzept-Interesse-Typen unterscheiden sich in ihrer im Naturwissenschaftsunterricht wahrgenommenen Kompetenzunterstützung, Autonomieunterstützung und inhaltlichen Relevanz. High Performer eines leistungszuversichtlichen Selbstkonzept-Interesse-Typs nehmen (unabhängig von ihrem Interesse) mehr Kompetenzunterstützung wahr als High Performer eines weniger leistungszuversichtlichen Typs. High Performer mit hohem Selbstkonzept und Interesse unterscheiden sich von High Performern mit hohem Selbstkonzept und niedrigem Interesse durch mehr wahrgenommene inhaltliche Relevanz und Autonomieunterstützung. High Performer mit niedrigem Selbstkonzept und hohem Interesse unterscheiden sich von High Performern mit niedrigem Selbstkonzept und niedrigem Interesse in ihrer wahrgenommenen inhaltlichen Relevanz.

Unterricht mit mehr forschend-entdeckendem Lernen wird von High Performern leistungszuversichtlicher und interessierter Selbstkonzept-Interesse-Typen mehr autonomieunterstützend, kompetenzunterstützend und inhaltlich relevant wahrgenommen. Für (a) wenig leistungszuversichtliche Selbstkonzept-Interesse-Typen ist der Zusammenhang zwischen forschend-entdeckendem Lernen und Kompetenzunterstützung, (b) für wenig interessierte Selbstkonzept-Interesse-Typen

der Zusammenhang zwischen forschend-entdeckendem Lernen und Autonomieunterstützung bzw. inhaltlicher Relevanz sowie (c) für sowohl wenig leistungszuversichtliche als auch wenig interessierte Selbstkonzept-Interesse-Typen der Zusammenhang zwischen forschend-entdeckendem Lernen und allen dreien motivationsunterstützenden Lehr-Lernbedingungen weniger stark ausgeprägt als bei leistungszuversichtlichen und interessierten Selbstkonzept-Interesse-Typen.

6. Methodisches Vorgehen

6.1. Forschungskontext

Methodischer Zugang dieser Arbeit ist eine Sekundäranalyse von Daten aus PISA 2006 und PISA 2012. Internationale Vergleichsstudien wie PISA ermöglichen internationales Benchmarking und Monitoring sowie die Abbildung von Trends (Drechsel, Prenzel & Seidel, 2015; Sälzer, 2015; Sälzer, Schiepe-Tiska & Prenzel, 2017; Seidel & Prenzel, 2008). Sie generieren damit vor allem beschreibendes Wissen zu Bedingungen und Ergebnissen von Bildung bzw. Unterricht und deren Zusammenhängen (Bromme, 2014; Sälzer, 2015) und ermöglichen, auf „gravierende Problemlagen und Handlungsbedarf“ (Bromme, 2014, S. 40) hinzuweisen. In dieser Funktion ermöglichen Large-Scale Assessments wie PISA jedoch auch die gezielte Betrachtung von Subgruppen wie den leistungsstarken Schülerinnen und Schülern (Zimmer et al., 2007) und damit verbunden die Generierung von beschreibendem Wissen zu Bedingungen, Ergebnissen, Zusammenhängen und Problemlagen dieser Subgruppen. Während zu leistungsschwächeren Schülerinnen und Schülern bzw. Jugendlichen mit besonderem Förderbedarf bereits vermehrt Untersuchungen außerhalb der Primäranalysen bzw. üblichen Berichtlegung von PISA durchgeführt wurden (z. B. Gebhardt, Sälzer, Mang, Müller & Prenzel, 2015; Müller, Prenzel, Sälzer, Mang & Gebhardt, 2017), existieren zu leistungsstarken Schülerinnen und Schülern nur vereinzelte Untersuchungen. Sie wurden in den vorangegangenen Kapiteln bereits gemeinsam mit Studien, die nicht aus dem Large-Scale-Bereich stammen, zur Darstellung des Forschungsstandes inhaltlich behandelt. In diesem Abschnitt werden sie abschließend noch einmal für eine gezielte Darstellung des Forschungskontexts dieser Arbeit innerhalb der Forschung zu PISA aufgegriffen. Tabelle 1 zeigt eine Übersicht über diese bisher im Rahmen von PISA erfolgten Betrachtungen leistungsstarker Jugendlicher und die Daten, Subgruppen und Analysen, die dort verwendet wurden.

Prenzel und Schütte (2007) und der Top-Performer-Bericht der OECD (OECD,

Tabelle 1.
Untersuchungen zu leistungsstarken Jugendlichen in PISA

Studie	Daten	Subgruppe	Beschreibung
Prenzel und Schütte (2007)	PISA 2006 (15-Jährige), Deutschland	Beste 25% in Naturwissenschaften	Untersuchung der Interessenverteilung sowie des Zusammenhangs von Experimentieren und Anwendungsbezug im Unterricht mit Interesse leistungsstarker Jugendlicher
Zimmer et al. (2007)	PISA 2000 und 2003 (15-Jährige), Deutschland	Vielseitig hochkompetente Jugendliche (in Mathe, Naturw. und Lesen unter den besten 10%)	Beschreibung der Anteile, Geschlechterunterschiede, Kompetenzwerte in den Domänen, Verteilung auf Klassenstufen und Schulformen, Anteile in den Bundesländern sowie Bildungsabschlüsse der Eltern für 2000 und 2003
OECD (2009b)	PISA 2006 (15-Jährige), international	Top Performer (KS V+VI), Strong Performer (KS IV), Middle Performer (KS III), Low Performer (<=KS I+II)	Kompetenzen in Mathe und Lesen, Anteile in den OECD-Staaten sowie Vergleich von sozioökonomischem Hintergrund, Interesse, zukunftsorientierte Motivation, Unterrichtsaktivitäten und anderer Merkmale jeweils im Vergleich von Top und Strong Performern

Anmerkungen. KS = Kompetenzstufen.

2009b) beschäftigten sich mit unterschiedlichen Definitionen mit leistungsstarken Jugendlichen in Naturwissenschaften. Die OECD fokussierte sich in ihrem Bericht auf einen deskriptiven internationalen Vergleich einzelner Merkmale zwischen Top, Strong, Middle und Low Performern. Sowohl individuelle Merkmale wie Interesse als auch familiäre Merkmale wie der sozioökonomische Hintergrund und Unterrichtsmerkmale wie Unterrichtsaktivitäten wurden in einem Vergleich der Leistungsgruppen betrachtet. Obwohl hier eine breite Auswahl an Variablen betrachtet wurde, bleiben einige Forschungslücken: Klassenmerkmale werden beispielsweise nicht betrachtet und die Analysen begrenzen sich in weiten Teilen auf einen ersten Analyseschritt, in dem Mittelwerte verglichen wurden. Vertiefte Analysen, in denen hypothesengeleitet Merkmale kombiniert oder im Zusammenhang betrachtet werden, wurden nur selten eingeschlossen (vgl. Kapitel 2.2). Der Fokus lag darüber hinaus auf einem Vergleich zwischen Leistungsgruppen, d. h. leistungsstarken mit leistungsschwächeren Schülerinnen und Schüler. Analysen innerhalb der Gruppe leistungsstarker Jugendlicher, beispielsweise zum Zusammenhang von Kompetenz oder Interesse leistungsstarker Schülerinnen und Schüler mit Unterrichtsmerkmalen, wurden nicht betrachtet.

Prenzel und Schütte (2007) dagegen unternahmen eine vertiefte Untersuchung zum Interesse innerhalb leistungsstarker Jugendlicher und untersuchten unter anderem, wie Schülerexperimente und Anwendungsbezüge im Naturwissenschaftsunterricht mit dem Interesse der besten 25% zusammenhängen. Diese Studie zeigte, dass ein großer Anteil leistungsstarker Jugendlicher ein vergleichsweise niedriges Interesse aufwies. Sie wirft damit Fragen für vertiefte Analysen zu leistungsstarken aber uninteressierten Schülerinnen und Schülern, beispielsweise zu Zusammenhängen anderer Unterrichtsaktivitäten, der Rolle des Selbstkonzepts und dazu, wie motivationsunterstützend leistungsstarke Jugendliche ihren Unterricht wahrnehmen, auf. Beide Analysen verwendeten den Datensatz der 15-Jährigen in PISA. Als Erweiterungsstichprobe (*Oversample*) wurde in Deutschland sowohl in 2006 als auch 2012 eine klassenbasierte Stichprobe, die pro Schule zwei vollständige neunte Klassen enthält, gezogen. Diese klassenbasierten Datensätze bieten die Gelegenheit, bei der Untersuchung von Unterricht die genestete Klassenstruktur zu berücksichtigen und damit Unterrichtseffekte auch auf Klassenebene zu untersuchen. In dieser Arbeit werden daher die klassenbasierte Zusatzstichprobe von PISA 2006 und PISA 2012 verwendet.

Diese Arbeit setzt an aktuellen allgemeinen und PISA-bezogenen Forschungsergebnisse zu leistungsstarken Jugendlichen in Naturwissenschaften an und greift dort aufgeworfene Fragen sowie nicht ausgeschöpftes Potential von Large-Scale-Assessments für die Untersuchung leistungsstarker Jugendlicher in Naturwissenschaften auf. Da Naturwissenschaften in PISA 2006 als Hauptdomäne und damit in Kombination mit naturwissenschaftsbezogenen motivational-affektiven und Unterrichtsmerkmalen erhoben wurde, dient PISA 2006 als hauptsächliche Datengrundlage und insbesondere zur vertieften Untersuchung von Selbstkonzept und Interesse sowie Unterrichtsmerkmalen im Kontext leistungsstarker Jugendlicher. PISA 2012 wird dafür genutzt, charakteristische Merkmale leistungsstarker Jugendlicher auf Basis einer aktuellen Stichprobe zu analysieren und die Vergleichbarkeit der Gruppe der leistungsstarken Jugendlichen in 2006 und 2012 zu untersuchen.¹

¹Parallel zur Verfassung dieser Arbeit erfolgt die Berichtlegung von PISA 2015, das ebenfalls Naturwissenschaften als Hauptdomäne untersucht. Zu PISA 2015 wurde jedoch das Rahmenkonzept der naturwissenschaftlichen Kompetenz geändert, auf computerbasiertes Testen umgestellt sowie das Rahmenmodell zur Untersuchung motivationaler und unterrichtsbezogener Variablen verändert. Daher wurde in dieser Arbeit auf PISA 2012 als aktuelle und mit 2006 eher vergleichbare Stichprobe zurückgegriffen.

6.2. Stichprobe

Für diese Arbeit wurden die klassenbasierte Erweiterungsstichprobe von PISA 2006 und PISA 2012 verwendet. Für diese Stichprobe wurden in jeder für den internationalen Vergleich gezogenen Schule zwei vollständige neunte Klassen zufällig ausgewählt und damit die Stichprobe der 15-Jährigen erweitert (Carstensen et al., 2007; Frey & Asseburg, 2009; Heine, Sälzer, Borchert, Sibberns & Mang, 2013). Die damit zusätzlich ausgewählten Schülerinnen und Schüler nahmen ebenfalls an den zweistündigen Kompetenztests teil und beantworteten anschließend einen Fragebogen, unter anderem zu motivational-affektiven Merkmalen und Unterrichtswahrnehmung. Förderschulen und Berufsschulen wurden für die vorliegende Arbeit aus der Stichprobe ausgeschlossen. Insgesamt enthalten die klassenbasierten Stichproben damit 395 (2006) bzw. 412 (2012) neunte Klassen aus 203 (2006) bzw. 211 (2012) Schulen und insgesamt 9577 (2006) bzw. 9845 (2012) Schülerinnen und Schüler. Die mittlere Klassengröße betrug 24.25 (2006, $SD = 4.93$) bzw. 23.90 (2012, $SD = 4.25$). Die beiden Stichproben setzen sich zu etwa gleich großen Teilen aus Mädchen und Jungen zusammen (2006: 50% weiblich, 51% männlich; 2012: 50% weiblich, 50% männlich) und die Schülerinnen und Schüler sind im Mittel 15.67 (2006, $SD_{Alter} = 0.63$, $Alter_{min} = 13.08$, $Alter_{max} = 19.92$) und 15.47 (2012, $SD_{Alter} = 0.50$, $Alter_{min} = 13.08$, $Alter_{max} = 16.33$) Jahre alt. Die Stichprobenezusammensetzung im Hinblick auf die Schularten in den beiden Durchgängen ist jeweils repräsentativ für die jeweils aktuelle Verteilung der Schularten in Deutschland. Tabelle 2 zeigt die Verteilung der Schularten in der Stichprobe 2006 und 2012. Beim Vergleich der Verteilungen ist erkennbar, dass 2012 deutlich weniger Schülerinnen und Schüler eine Hauptschule besuchten und die anderen Schularten im Vergleich zu 2006 höhere Anteile an der Population hatten.

Tabelle 2.
Anteil der Schularten an Stichproben in PISA 2006 und PISA 2012

Jahr	HS (%)	IGS (%)	MBG (%)	RS (%)	GY (%)
2006	18.40	9.30	8.50	29.00	34.80
2012	9.90	10.40	14.20	26.70	38.90

6.3. Instrumente

6.3.1. Kognitive Merkmale

Naturwissenschaftliche Kompetenz

Die naturwissenschaftliche Kompetenz wurde in PISA 2006 und PISA 2012 mithilfe eines standardisierten Tests erfasst. Wie in Kapitel 2.1 beschrieben, enthält das Rahmenmodell naturwissenschaftlicher Kompetenz die drei Teilkompetenzen „Phänomene naturwissenschaftlich erklären“, „Naturwissenschaftliche Fragestellungen erkennen“ und „Naturwissenschaftliche Evidenz interpretieren“ sowie zwei verschiedene Wissensbereiche und fünf verschiedene Kontexte (OECD, 2006). Alle im Test verwendeten Items waren jeweils einer Teilkompetenz, einem Wissensbereich und einem Kontext zugeordnet. Mehrere Items wurden jeweils in Units zusammengefasst und die Units in einem Multi-Matrix-Design auf Testhefte verteilt (für eine detaillierte Beschreibung vgl. Carstensen et al., 2007). Anhand der Skalierung nach dem Rasch-Modell (Rasch, 1980) bzw. dem *Partial-Credit-Modell* (Masters, 1982) werden die Personenfähigkeiten unter Kontrolle der Aufgabenschwierigkeiten ermittelt und mit Hilfe eines Hintergrundmodells für jede Schülerin bzw. jeden Schüler fünf Plausible Values (PV) als Maß für die naturwissenschaftliche Kompetenz geschätzt.

Mathematische Kompetenz und Lesekompetenz

Mathematik und Lesen wurden in 2006 als Nebendomäne untersucht, Mathematik in 2012 als Hauptdomäne. Wie die naturwissenschaftliche Kompetenz wurden sie mit Testitems bzw. Units standardisiert geprüft. Für die Nebendomänen wurde jeweils eine geringe Itemanzahl verwendet als für die Hauptdomäne. Auch in Mathematik und Lesen wurde mithilfe des Bookletdesigns, der IRT-Skalierung und des Hintergrundmodells jeder Schülerin bzw. jedem Schüler fünf Plausible Values zugeordnet (vgl. Carstensen et al., 2007; Heine et al., 2013). Tabelle 3 enthält eine Übersicht über die Mittelwerte inklusive Standardfehler sowie die Streuung der Kompetenzen in der Stichprobe.

Kognitive Grundfähigkeit

Die kognitive Grundfähigkeit der Schülerinnen und Schüler wurde in PISA 2006 lediglich in Deutschland mithilfe der Subskala „Figurale Analogien“ des Kognitiven Fähigkeitstests

Tabelle 3.

Kennwerte der Kompetenzen

Skala	Jahr	N	M	(SE)	SD
Naturwissenschaftliche Kompetenz (PV)	2006	9577	513.65	(3.38)	89.00
	2012	9845	517.09	(2.43)	85.89
Mathematische Kompetenz (PV)	2006	9577	500.97	(3.20)	85.34
	2012	9845	508.56	(2.31)	83.80
Lesekompetenz (PV)	2006	9577	497.55	(3.54)	92.83
	2012	9845	503.19	(2.42)	82.37

(KFT) von Heller und Perleth (2000) am zweiten Testtag erfasst. Die Werte wurden IRT-skaliert und in Form eines WLE-Schätzers (*Weighted Likelihood Estimate*, Warm, 1989) angegeben, da sie ohne Berücksichtigung des Hintergrundmodells geschätzt wurden und nur auf den Itemschwierigkeiten und der Anzahl zugestimmter Items beruhen (Carstensen et al., 2007; Frey & Asseburg, 2009). Die Stichprobenwerte für PISA 2006 sind in Tabelle 4 zusammengefasst.

Schulnoten

Neben den Kompetenzen und der kognitiven Grundfähigkeit wurden die Halbjahresnoten in den einzelnen naturwissenschaftlichen Fächern Physik, Chemie und Biologie bzw. im Fach Naturwissenschaften sowie in Mathematik und Deutsch durch den Schulkoordinator mitgeteilt. Die Noten befinden sich auf einer Skala von 1 als beste Note bis 6 als schlechteste Note. Die Mittelwerte und Standardabweichungen in der Stichprobe für die kognitive Grundfähigkeit und die Noten sind in Tabelle 4 dargestellt. Da die naturwissenschaftlichen Noten nur in 2006 erfasst wurden, sind die Werte nur für die Stichprobe in 2006 dargestellt.

Tabelle 4.

Kennwerte der kognitiven Grundfähigkeit sowie der Schulnoten 2006 (1: bester Wert, 6: schlechtester Wert)

Skala	N	M	(SE)	SD	Min	Max
Kognitive Grundfähigkeit (WLE)	9228	-0.09	(0.05)	1.49	-4.89	3.34
Note Deutsch	9472	3.12	(0.02)	0.85	1.00	6.00
Note Mathe	9487	3.24	(0.02)	1.02	1.00	6.00
Note Biologie	7626	3.03	(0.03)	0.95	1.00	6.00
Note Chemie	7517	3.07	(0.03)	0.98	1.00	6.00
Note Physik	7957	3.13	(0.03)	0.98	1.00	6.00
Note Naturwissenschaft	853	2.87	(0.10)	0.98	1.00	6.00

6.3.2. Familiäre und institutionelle Merkmale

Sozioökonomischer Hintergrund

Die soziale Herkunft der Schülerinnen und Schüler wurde in Form der sozioökonomischen Stellung ihrer Familie erfasst. Dazu wurden die Berufsangaben der Eltern in PISA 2006 in die ISCO-88-Kodierung (International Labour Office, 1990), in PISA 2012 in die ISCO-08-Kodierung (International Labour Office, 2012) eingeordnet und daraus drei Indikatoren gebildet: Der HISEI als erster Index stellt den höchsten Wert des *International Socio-Economic Index of Occupational Status (ISEI)* der Familie dar (Ganzeboom et al., 1992). Er ordnet Berufe einer sozialen Hierarchie zu und bildet den sozioökonomischen Status anhand einer Berufsrankkala ab (für eine detaillierte Beschreibung vgl. Ehmke & Baumert, 2007; Watermann & Baumert, 2006). Tabelle 17 gibt eine Übersicht über die Stichprobenwerte.

Tabelle 5.
Kennwerte des „Highest International Socio-Economic Index of Occupational Status (HISEI)“

Jahr	N	M	(SE)	SD	Min	Max
2006	8927	49.06	(0.42)	16.35	16.00	90.00
2012	9998	60.80	(0.44)	24.71	10.00	99.00

Schließlich erlaubt die Zuordnung der Berufe zu den EGP-Klassen (Erikson et al., 1979) eine qualitative Einteilung und Klassifikation der Berufe, mit deren Hilfe eine Art soziale Schichten abgebildet werden können. Die Klassifikation beruht auf einem Kategoriensystem, das Berufe nach Tätigkeit, Stellung und Weisungsbefugnis abstuft. Damit berücksichtigen die EGP-Klassen mögliche Unterschiede in Bildungsressourcen und Bildungsbeteiligung, die auf Einkommen, Bildungsstand sowie gesellschaftliche bzw. berufliche Stellung der Eltern zurückzuführen sind. Üblicherweise, wenngleich umstritten, wird jeweils der Beruf des Vaters zugrunde gelegt und fehlende Werte durch die EGP-Klasse der Mutter aufgefüllt (Müller & Ehmke, 2013). Welche Verteilung die Stichprobenangaben auf die EGP-Klassen aufweisen, zeigt Tabelle 6.

Tabelle 6.
Verteilung der EGP-Klassen

Jahr	N	EGP-Klassen											
		1		2		3		4		5		6	
		%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
2006	8522	21.77	(0.79)	14.87	(0.44)	2.97	(0.19)	10.63	(0.38)	22.22	(0.71)	16.29	(0.55)
2012	8029	20.36	(0.81)	15.08	(0.38)	3.70	(0.20)	9.29	(0.32)	17.91	(0.55)	14.67	(0.51)

Anmerkung. 1: Obere Dienstklasse, 2: Untere Dienstklasse, 3: Routinedienstleistungen Handel und Verwaltung, 4: Selbstständige, 5: Facharbeiter und Arbeiter mit Leitungsfunktion, 6: Un- und angelernte Arbeiter, Landarbeiter (vgl. Ehmke & Baumert, 2007)

6.3.3. Motivational-affektive Merkmale

Naturwissenschaftsbezogenes Fähigkeitsselbstkonzept

Das naturwissenschaftsbezogene Selbstkonzept wurde mit sechs Items (z. B. „Es fällt mir leicht, neue Ideen im naturwissenschaftlichen Unterricht zu verstehen.“, für eine Übersicht aller Items siehe Tabelle A.5 im Anhang) erfasst (Frey & Asseburg, 2009). Schülerinnen und Schüler gaben eine über die naturwissenschaftlichen Fächer hinweg generalisierte Einschätzung ihrer eigenen Fähigkeit ab. Dies erfolgte auf einer vierstufigen Likert-Skala (1 = „stimme ganz zu“, 2 = „stimme eher zu“, 3 = „stimme eher nicht zu“, 4 = „stimme gar nicht zu“) mit guter Reliabilität ($\alpha = .90$, vgl. Tabelle 7).

Interesse an den Naturwissenschaften

Interesse an den Naturwissenschaften wurde in PISA 2006 mit Hilfe der Skala „Freude & Interesse“ gemessen. Die fünf Items erfassten sowohl das Vergnügen an der Beschäftigung mit naturwissenschaftlichen Themen als auch die Bereitschaft, Neues in diesem Bereich zu lernen (z. B. „Ich beschäftige mich gerne mit naturwissenschaftlichen Problemen.“, für eine Übersicht aller Items siehe Tabelle A.6 im Anhang). Auch hier wurden Schülerinnen und Schüler nach einer über alle naturwissenschaftlichen Fächer hinweg generalisierten Einschätzung gefragt und beantworteten sie auf einer vierstufigen Likert-Skala (1 = „stimme ganz zu“, 2 = „stimme eher zu“, 3 = „stimme eher nicht zu“, 4 = „stimme gar nicht zu“). Die Reliabilität der Skala war mit $\alpha = .92$ gut (vgl. Tabelle 7).

Tabelle 7.

Kennwerte der Skalen zu naturwissenschaftsbezogenem Selbstkonzept und Interesse

	N	Zahl der Items	Cronbach's α	M	(SE)	SD	Min	Max
Fähigkeitsselbstkonzept	9216	6	.90	2.75	(0.01)	0.66	1.00	4.00
Interesse	9311	5	.92	2.50	(0.01)	0.79	1.00	4.00

6.3.4. Motivationsunterstützende Lehr-Lernbedingungen

Motivationsunterstützende Lehr-Lernbedingungen wurden als nationale Ergänzungsvariablen an einem zweiten Testtag erfasst (Frey & Asseburg, 2009). Jede Schülerin bzw. jeder Schüler beantwortete diese Fragen für ein selbstgewähltes naturwissenschaftliches Fach (Biologie, Chemie oder Physik). Die Skalen zur Erfassung motivationsunterstützender Lehr-Lernbedingungen wurden im Rahmen der IPN-Videostudie entwickelt und erprobt (Dalehefte, 2001; Seidel, Prenzel & Kobarg, 2005). Alle drei Skalen sowie das Item zur Auswahl eines naturwissenschaftlichen Faches sind in den Tabellen A.11, A.12, A.13 und A.14 dargestellt. Tabelle 8 gibt eine Übersicht über die Kennwerte der Skalen und zeigt für alle drei Merkmale eine zufriedenstellende Reliabilität.

Tabelle 8.

Kennwerte der Skalen zu motivationsunterstützenden Lehr-Lernbedingungen

	N	Itemanzahl	Cronbachs α	M	(SE)	SD	Min	Max
Kompetenzunterstützung	9149	3	.76	2.73	(0.01)	0.74	1.00	4.00
Autonomieunterstützung	9149	3	.80	2.85	(0.01)	0.75	1.00	4.00
Inhaltliche Relevanz	9163	3	.80	2.53	(0.01)	0.73	1.00	4.00

6.3.5. Unterrichtaktivitäten

Fünf domänenübergreifende und naturwissenschaftsbezogene Unterrichtsaktivitäten, die auf Seidel und Prenzel (2006b) zurückgehen, wurden erfasst. Die Items wurden handlungsbezogen und ereignisnah konstruiert (Kobarg et al., 2011) und fragen, wie oft genannte Aktivitäten im Physik-, Chemie-, Biologie- und Naturwissenschaftsunterricht vorkommen. Vier Antwortoptionen standen jeweils zur Auswahl (1 = „in allen Stunden“; 2 = „in den meisten Stunden“; 3 = „in manchen Stunden“; 4 = „nie oder fast nie“). Eine

Übersicht über die Kennwerte der Skalen zeigt Tabelle 9.

Tabelle 9.

Kennwerte der Skalen zu Unterrichtsaktivitäten im Naturwissenschaftsunterricht

	N	Anzahl Items	Cron- bachs α	M	(SE)	SD	Min	Max
Interaktives Lehren und Lernen	9214	4	.76	2.53	(0.01)	0.69	1	4
Experimentieren	9221	4	.71	2.28	(0.02)	0.62	1	4
Schülerinnen und Schüler forschen	9232	3	.77	1.74	(0.02)	0.69	1	4
Anwendungen	9185	4	.72	2.34	(0.01)	0.64	1	4

Interaktives Lehren und Lernen

Interaktives Lehren und Lernen erfasst, wie häufig Schülerinnen und Schüler die Gelegenheit hatten, sich interaktiv am Unterricht zu beteiligen, und wurde mit vier Items (z. B. „Schülerinnen und Schüler bekommen Gelegenheit, ihre Ideen zu erklären.“, für eine Übersicht aller Items siehe Tabelle A.7 im Anhang) operationalisiert und mit einer vierstufigen Likertskala abgefragt (1 = „in allen Stunden“; 2 = „in den meisten Stunden“; 3 = „in manchen Stunden“; 4 = „nie oder fast nie“). Die Skala weist eine zufriedenstellende interne Konsistenz auf ($\alpha = .76$).

Forschend-entdeckender Unterricht

Der Anteil forschend-entdeckenden Unterrichts wurde mit zwei getrennten Skalen „Experimentieren“ (vier Items, z. B. „Die Schülerinnen und Schüler führen praktische Experimente im Labor durch.“, für eine Übersicht aller Items siehe Tabelle A.8 im Anhang) und „Schüler und Schülerinnen forschen“ (drei Items, z. B. „Die Schülerinnen und Schüler dürfen ihre eigenen Experimente entwickeln.“, für eine Übersicht aller Items siehe Tabelle A.9 im Anhang) erfasst (vierstufige Likert-Skala: 1 = „in allen Stunden“; 2 = „in den meisten Stunden“; 3 = „in manchen Stunden“; 4 = „nie oder fast nie“). Mit der Skala „Experimentieren“ wurde abgefragt, wie häufig Schülerinnen und Schüler im Naturwissenschaftsunterricht die Möglichkeit hatten, Experimente eigenständig durchzuführen – oft „Hands-On“-Aktivitäten genannt (Seidel, Prenzel, Wittwer & Schwindt, 2007). „Experimentieren“ misst damit den eher instrumentellen, „handwerklichen“ Teil des naturwissenschaftlichen Forschungsprozesses, der sich auf das Durchführen und Interpretieren von Experimenten konzentriert. Die Skala „Schülerinnen und Schüler

forschen“ dagegen erfasst, ob Schülerinnen und Schüler bei Experimenten auch aus ihrem eigenen Wissenshintergrund Fragestellungen entwickeln und dazu passende Experimente planen dürfen (Seidel & Prenzel, 2006b). Damit untersucht sie, inwieweit der naturwissenschaftliche Erkenntnisprozesse im Naturwissenschaftsunterricht nicht nur auf das Durchführen und Interpretieren von Experimenten beschränkt ist, sondern auch einen forschenden Aspekt miteinschließen, in dem auch Fragestellungen und Untersuchungszugänge Teil forschend-entdeckenden Lernens sind (Seidel, Prenzel, Wittwer & Schwindt, 2007). Mit $\alpha = .71$ und $\alpha = .77$ waren die Reliabilitäten beider Skalen zufriedenstellend (vgl. Tabelle 9).

Naturwissenschaftliche Anwendungen

Naturwissenschaftliche Anwendungen wurden mit vier Items untersucht (z. B. „Der Lehrer/die Lehrerin verwendet den naturwissenschaftlichen Unterricht, um den Schülerinnen und Schülern die Welt außerhalb der Schule verständlich zu machen.“, für eine Übersicht aller Items siehe Tabelle A.10 im Anhang). Darin werden die Schülerinnen und Schüler gefragt, wie häufig sie naturwissenschaftliche Konzepte oder Prinzipien auf Alltagsphänomene anwenden und Modellvorstellungen von Naturwissenschaften dazu verwenden, ihre Umwelt zu verstehen. Die Skala wies eine zufriedenstellende Reliabilität auf ($\alpha = .72$).

6.4. Auswertungsverfahren

6.4.1. Umgang mit Plausible Values

In den PISA-Studien werden die Ergebnisse der standardisierten Leistungstexts IRT-skaliert und für jede Schülerin bzw. jeden Schüler in jeder Domäne fünf Kompetenzwerte – sogenannte *Plausible Values (PVs)* – auf einer A-posteriori-Verteilung der latenten Kompetenz gezogen (für eine detaillierte Beschreibung vgl. OECD, 2009a). Damit wird für jede Testperson statt eines einzigen Punktschätzers eine Leistungsverteilung geschätzt und sichergestellt, dass trotz der diskontinuierlichen Messung (Testitems) einer kontinuierlichen latenten Variable (naturwissenschaftliche Kompetenz) Testwerte mit einer möglichst kleinen Verzerrung ermittelt werden.

Damit die Standardfehler in weiterführenden Rechnungen nicht unterschätzt werden, müssen sämtliche Analysen für jeden PV zunächst getrennt berechnet und das Ergebnis erst am Ende gemittelt werden (OECD, 2009a). Wenn θ_i das Einzelergebnis einer Analyse mit PV i als Variable darstellt (z. B. der Regressionskoeffizient von PV i auf Geschlecht), so berechnet sich das Analyseergebnis der Population θ (z. B. der Regressionskoeffizient der naturwissenschaftlichen Kompetenz auf Geschlecht) wie in Formel 1 dargestellt als Mittelwert der Ergebnisse der getrennten PV-Analysen (M : Anzahl der PVs).

$$\theta = \frac{1}{M} \sum_{i=1}^M \theta_i \quad (1)$$

Eine Mittelung der PVs und weiterführende Analyse mit einem gemittelten PV-Wert pro Schülerin bzw. Schüler würde zu verzerrten Ergebnissen bei allen varianzbasierten Koeffizienten führen. Der Standardfehler dieses Werts ergibt sich aus der Imputations- und Stichprobenvarianz (vgl. OECD, 2009a). Die Imputationsvarianz B_M bezeichnet die Messvarianz, die auf die Messungenauigkeit des Tests zurückzuführen ist. Sie hängt von der Varianz der Werte der einzelnen PV-Analysen ab und wird wie folgt berechnet:

$$B_M = \frac{1}{M-1} \sum_{i=1}^M (\theta_i - \theta)^2 \quad (2)$$

Die Stichprobenvarianz U beschreibt dagegen die allgemeine statistische Unsicherheit, die durch die Ziehung einer Stichprobe im Unterschied zur Analyse der Gesamtpopulation entsteht. Die finale Stichprobenvarianz ergibt sich wie der Mittelwert aus

der Mittelung der Stichprobenvarianzen der für jeden PV geschätzten Mittelwerte.

Die Gesamtvarianz des Analyseergebnisses θ wird dann wie in Gleichung 3 dargestellt aus Imputations- und Stichprobenvarianz berechnet (nach Rubin, 1987).

$$V = U + \left(1 + \frac{1}{M}\right)B_M \quad (3)$$

Überträgt man diese Überlegungen auf die für Werte $\hat{\mu}_i$, die in den Analysen mit den einzelnen PVs für die Statistik θ_i geschätzt werden, so werden Gesamtmittelwert $\hat{\mu}$, Imputationsvarianz $\sigma_{(test)}^2$, Stichprobenvarianz $\sigma_{\hat{\mu}}^2$ und Gesamtfehlervarianz σ_{error}^2 folgendermaßen berechnet:

$$\hat{\mu} = \frac{1}{5} \sum_{i=1}^5 \hat{\mu}_i = \frac{1}{5}(\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 + \hat{\mu}_4 + \hat{\mu}_5) \quad (4)$$

$$\sigma_{(test)}^2 = \frac{1}{4} \sum_{i=1}^5 (\hat{\mu}_i - \hat{\mu})^2 \quad (5)$$

$$\sigma_{\hat{\mu}}^2 = \frac{1}{5}(\sigma_{\hat{\mu}_1}^2 + \sigma_{\hat{\mu}_2}^2 + \sigma_{\hat{\mu}_3}^2 + \sigma_{\hat{\mu}_4}^2 + \sigma_{\hat{\mu}_5}^2) \quad (6)$$

$$\sigma_{error}^2 = \sigma_{\hat{\mu}}^2 + (1.2\sigma_{(test)}^2) \quad (7)$$

Der Standardfehler ist dann die Wurzel der Fehlervarianz. Für das Vorgehen in dieser Arbeit bedeutet dies, dass die Gruppe der Top bzw. High Performer im Folgenden zunächst jeweils getrennt für jeden Plausible Value betrachtet wird, d. h. es werden fünf Top- bzw. High-Performer-Datensätze gebildet. Sämtliche Analysen werden fünf Mal – für jeden PV einmal – gerechnet und die resultierenden Ergebnisse erst im letzten Schritt gemittelt (vgl. Formel 4) und aus den dazugehörigen fünf Standardfehlern der Gesamtfehler des Wertes nach Formel 7 berechnet.

6.4.2. Mehrebenenanalysen

Da für die klassenbasierte Stichprobe von PISA 2006 vollständige Schulklassen gezogen wurden, weist sie eine genestete Struktur auf. Im Unterschied zu einer Stichprobe, in der Schülerinnen und Schüler zufällig gezogen wurden, erfüllen die Daten auf Schülerebene daher nicht das Kriterium der statistischen Unabhängigkeit (zwei Schülerinnen und Schüler innerhalb einer Klasse sind sich ähnlicher als zwei Schülerinnen und Schüler

unterschiedlicher Klassen). Bei inferenzstatistischen Verfahren besteht damit die Gefahr, Standardfehler zu unterschätzen (Raudenbush & Bryk, 2010). Abhängig von der theoretischen Absicht bestehen zwei Möglichkeiten, die genestete Struktur zu berücksichtigen: Besteht lediglich Interesse an Zusammenhängen auf der Schülerebene, können die Standardfehler der Parameter angepasst werden (B. O. Muthen & Satorra, 1995). Alternativ ermöglichen Mehrebenenanalysen, nicht nur die Standardfehler zu korrigieren, sondern auch die Varianz der Zusammenhänge in Individual- und Klassenebene aufzuteilen und so sowohl Individual- als auch Kontexteffekte zu betrachten (Hox, 2010; Raudenbush & Bryk, 2010). In dieser Arbeit wird der zweite Ansatz verfolgt und in Mehrebenenanalysen mithilfe der Software Mplus 7.4 (Muthén & Muthén, 1998-2015) ermittelt, inwieweit Zusammenhänge zwischen Unterricht und Kompetenz sowie Interesse an Naturwissenschaften auf Individualebene oder Klassenebene zurückzuführen sind.

Dafür ermöglicht die Mehrebenenanalyse zunächst, in Form des Intraclass-Korrelations-Koeffizienten (ICC) zu quantifizieren, welcher Anteil der Gesamtvarianz durch Unterschiede auf Level-2 erklärt wird (Hox, 2010).

$$ICC_1 = \frac{\tau_{00}}{\tau_{00} + \sigma^2} \quad (8)$$

Dabei bezeichnen τ_{00} die Varianz der Residuen μ_{0j} auf Level 2 und σ^2 die Varianz der Residuen r_{ij} auf Level 1. Zusätzlich wird in aktuelleren Mehrebenenanalysen (z. B. Marsh et al., 2009, 2012) ein zweiter Intraclass-Korrelations-Koeffizient – häufig ICC_2 oder *design effect* genannt – als Maß der Reliabilität des aggregierten Gruppenmittelwerts von Level-1-Angaben vorgeschlagen (Bliese, 2000; Raudenbush & Bryk, 2010):

$$ICC_2 = \frac{\tau_{00}}{\tau_{00} + \frac{1}{n}\sigma^2} = \frac{n \cdot ICC_1}{1 + (n - 1) \cdot ICC_1} \quad (9)$$

Die Größe n bezeichnet dabei die Größe der Level-2-Gruppe und wird üblicherweise als durchschnittliche Gruppengröße eingefügt (Marsh et al., 2009). Der ICC_2 berücksichtigt, dass die Level-2-Werte durch eine begrenzte Anzahl an Level-1-Einheiten aus der jeweiligen Level-2-Einheit entstehen und berechnet ähnlich der Reliabilität eines Faktors die Verlässlichkeit des Gruppenmittelwerts aus diesen Einzelangaben unter Einbezug des Stichprobenfehlers. In den Analysen dieser Arbeit werden sowohl ICC_1 als auch ICC_2 berechnet. Bei der Interpretation des ICC_2 muss dabei bedacht werden, dass in

der vorliegenden Stichprobe komplette Klassen erhoben wurden und daher bereits die maximal mögliche Gruppengröße n ausgeschöpft wurde.

Kompositions- und Kontexteffekte

Grundsätzlich wird bei Mehrebenenanalysen zwischen Kompositions-/Kontext- und Klimateffekten unterschieden (vgl. Marsh et al., 2012). Kompositions- oder Kontexteffekte beschreiben, welchen Effekt eine Variable auf Level 2 über ihren Effekt auf Level 1 hinaus hat. In dieser Modellierung ist die individuelle Person Referenz des betrachteten Merkmals, d. h. die Variable ist personenspezifisch und die Personen auf Individualebene sind nicht austauschbar. Beispiele solcher Variablen sind die Leistung, der sozioökonomische Hintergrund oder das Geschlecht. Im Gegensatz dazu beschreiben Klimateffekte Zusammenhänge von Variablen, die sich aus inhaltlicher Sicht auf Level 2 befinden und lediglich als Aggregat von Schülerangaben gemessen werden. Referenz ist hier – möglichst auch in der Formulierung entsprechender Items – die Klassenebene. Im Vordergrund steht die geteilte Wahrnehmung des Level-2-Konstrukts durch Level-1-Individuen. Einzelindividuen auf Level 1 werden als austauschbar betrachtet und Unterschiede auf Level 1 repräsentieren Unsicherheiten in der Erfassung des Level-2-Merkmals. Die Residualvarianz auf Level 1 sollte daher idealerweise sehr klein sein und nicht bzw. minimal mit Level-1-Kriterien korrelieren (Marsh et al., 2012). Level-1-Residuen von Klimavariablen können jedoch in einzelnen Fällen als bedeutende Zusammenhänge statt als Messfehler interpretiert werden. Dies muss jedoch theoriegeleitet und in dem Bewusstsein, dass sie keine üblichen Level-1-Variablen darstellen, geschehen (Marsh et al., 2012). Ein Beispiel für Klimavariablen sind Unterrichtsmerkmale wie die in dieser Arbeit verwendeten Unterrichtsaktivitäten, die nach der von der Schulklasse geteilten Wahrnehmung der Häufigkeit einzelner Aktivitäten fragen.

Aus methodischer Sicht wird bei Klimavariablen der Effekt eines Level-2-Prädiktors auf ein Level-2-Kriterium untersucht. Bei Kontextvariablen wird dagegen der Level-2-Effekt nach Kontrolle des Level-1-Effekts, d. h. die Differenz der Effekte auf beiden Ebenen, betrachtet. Der ICC_1 für Klimavariablen ist häufig kleiner .10 und selten höher als .30 (Marsh et al., 2012, S. 115). Für Kontextvariablen ist er häufig wesentlich höher. Der ICC_2 hängt für beide Merkmale vom ICC_1 und der Anzahl der Schülerinnen und Schüler pro Klasse ab: Je höher der ICC_1 und je höher die Anzahl der Schülerinnen und Schüler, desto höher der ICC_2 .

In der vorliegenden Arbeit werden den Grundlagen zu Kontext- und Klimaefekten folgend die Unterrichtsaktivitäten (interaktives Lehren und Lernen, forschend-entdeckendes Lernen und Anwendungen) als Klimavariablen behandelt. Wie von Marsh et al. (2012) empfohlen, wird zunächst ihre Faktorstruktur in der Stichprobe auf Level 1 und Level 2 untersucht sowie mit Hilfe von ICC_1 und ICC_2 geprüft, ob eine Mehrebenenstruktur auch empirisch zu finden ist. Da die hier verwendete Stichprobe mehr als 50 Klassen enthält, entspricht sie der empfohlenen Stichprobengröße auf Level 2 und eine Betrachtung von Klimaeffekten wird als möglich angenommen (Marsh et al., 2012).

Latent-manifeste Modellierung der Variablen

Zwei verschiedene Arten statistischer Unsicherheiten spielen bei Mehrebenenanalysen eine Rolle und können mithilfe einer doppelt-latenten Modellierung berücksichtigt werden: (1) Die latente Aggregation von Level-1-Daten auf Level 2 (Stichprobenfehler) und (2) die latente Modellierung von Faktoren innerhalb der beiden Ebenen (Messfehler). Die vier Kombinationsmöglichkeiten der beiden latenten/manifesten Modellierungen werden in der Nomenklatur von Marsh et al. (2012) als doppelt-manifest, latent-manifest, manifest-latent und doppelt-latent bezeichnet (vgl. Tabelle 10).

Tabelle 10.
Vier Modelle zur Berücksichtigung von Mess- und Stichprobenfehlern

	Stichprobenfehler		
		Nein	Ja
Messfehler	Nein	Doppelt-manifest	Manifest-latent
	Ja	Latent-manifest	Doppelt-latent

Eine latente Aggregation der Schülerdaten auf Klassenebene (manifest/latent-latent) berücksichtigt den Stichprobenfehler der Aggregation der Angaben einzelner aus einer Klasse ausgewählter Schülerinnen und Schüler. Da in der vorliegenden Stichprobe nicht einzelne Schülerinnen und Schüler einer Klasse ausgewählt, sondern vollständige Klassen erfasst wurden, ist die Berücksichtigung dieses Stichprobenfehlers hier jedoch weniger notwendig als bei zufällig aus einer Klasse gezogenen Schülerinnen und Schüler. Darüber hinaus würde eine latente Aggregation der Schülerdaten in einem Multigruppenansatz, in dem für High Performer und Nicht-High-Performer getrennt gerechnet wird, dazu führen, dass auf Klassenebene jeweils nur die Angaben der High bzw. die der

Nicht-High-Performer berücksichtigt werden. Eine manifeste Aggregation ermöglicht stattdessen, auch bei der getrennten Analyse der beiden Leistungsgruppen jeweils die Unterrichtsmerkmale auf Klassenebene auf die Angaben der ganzen Klasse zu basieren. Die Schülerangaben werden in der vorliegenden Arbeit daher manifest auf Level 2 aggregiert.

Eine latente Modellierung der Faktoren (latent-manifest/latent) berücksichtigt den Messfehler bei der Skalenbildung. Sie erfordert jedoch mindestens drei Indikatoren, eine hohe Stichprobenanzahl und aufgrund der dabei ablaufenden numerischen Integrationen unter Umständen eine hohe Rechenkapazität (Marsh et al., 2009). Sie zu umgehen und manifest zu modellieren ist dann gerechtfertigt, wenn kleine Stichprobenzahlen auf Level 1 oder 2 zu instabilen latenten Schätzungen und nicht konvergierenden Modellen führen (Lüdtke et al., 2009; Marsh et al., 2009). In der vorliegenden Arbeit enthalten alle Skalen zu Unterrichtsaktivitäten und abhängigen Variablen genügend Indikatoren und die Stichprobenanzahl erwies sich als ausreichend für stabile und konvergente Schätzungen linearer Zusammenhänge. In den Analysen zum Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz bzw. Interesse wurde daher eine latente Skalenbildung, d. h. Faktoren, verwendet. Insgesamt liegt damit eine latent-manifeste Modellierung der Unterrichtsaktivitäten vor (Marsh et al., 2009).

In Anlehnung an Forschungsarbeiten mit vergleichbaren Mehrebenenanalysen (Lüdtke et al., 2009; Marsh & Rowe, 1996; Morin, Marsh, Nagengast & Scalas, 2014) wurden für eine bessere Vergleichbarkeit der Koeffizienten und zur Vermeidung unnötiger Multikollinearitäten alle Variablen zunächst auf Level 1 z-standardisiert ($M = 0$, $SD = 1$) und dann auf Level 2 aggregiert. Werte auf Level 2 sind damit in Einheiten der Standardabweichung auf Level 1 zu interpretieren. Die Koeffizienten auf Level 1 können darüber hinaus als standardisierte Regressionsgewichte interpretiert werden (Hox, 2010). Als Software wurde *Mplus* 7.4 (Muthén & Muthén, 1998-2015) verwendet.

Standardisierung von Koeffizienten

In der bei Mehrebenenanalysen in *Mplus* implementierten automatischen Standardisierung werden Koeffizienten für beide Analyseebenen getrennt standardisiert, da dies für eine getrennte Interpretation von Ergebnissen verschiedener Ebenen angemessen ist (Marsh et al., 2012). Die Standardisierung von Koeffizienten zwischen den beiden Ebenen erfordert jedoch eine Standardisierung an der Gesamtvarianz einer Variable,

d. h. der Summe aus Level-1- und Level-2-Varianz (Marsh et al., 2009). In Anlehnung an Marsh et al. (2009, 2012) wurden daher alle Level-2-Koeffizienten mit der Gesamtvarianz standardisiert.

Verwendete Mehrebenenmodelle

Bei einer Mehrebenenanalyse werden verschiedene Modelle, die sich jeweils darin unterscheiden, wie viele Koeffizienten sie innerhalb und zwischen der Level festsetzen bzw. variieren lassen, unterschieden (vgl. Luke, 2010):

- *Random-Intercept-Models*: Nur der Achsenabschnitt wird als variable Größe modelliert
 - *Nullmodell (Intercept-Only-Model, One-Way-Random-Effects-ANOVA)*: Mehrebenenmodell ohne Prädiktoren; Eignet sich zur Ermittlung der ICCs bzw. allgemein der Varianzen von abhängigen Variablen auf beiden Ebenen
 - *Random-Intercept-Model (One-Way-Random-Effects-ANCOVA)*: Prädiktoren nur auf Level 1
 - *Random-Intercept-Regression (Means-as-Outcomes-Regression)*: Prädiktoren nur auf Level 2, d. h. es werden nur Level-2-Unterschiede erklärt
 - *Random-Intercept-ANCOVA (Means-as-Outcomes-ANCOVA)*: Kovariaten auf Level 1 und Level 2; Steigungen über Level-2-Einheiten hinweg konstant
- *Random-Slope-Models*: Eine oder mehrere Level-1-Steigungen sowie der Achsenabschnitt werden als variable Größen modelliert
 - *Random-Coefficient-Regression*: Prädiktoren nur auf Level 1; Unterschiede in Steigung und Achsenabschnitt zwischen Level-2-Einheiten zugelassen
 - *Intercept-and-Slopes-as-Outcomes-Model (Full-Random-Coefficient-Model)*: Prädiktoren auf Level 1 und Level 2; Unterschiede in Steigung und Achsenabschnitt zwischen Level-2-Einheiten zugelassen

In der vorliegenden Arbeit wird der Zusammenhang von Unterrichtsmerkmalen und Outcomes auf Level 2 über Klassen hinweg als vergleichbar betrachtet (Steigung Level 2 konstant). Darüber hinaus wird angenommen, dass sowohl die abhängige Variable auf Level 2 als auch auf Level 1 bei Unterrichtsmerkmal auf Level 2 oder Level 1 gleich

Null zwischen Klassen bzw. Schülerinnen und Schülern variiert (Intercept Level 2 und Level 1 variabel). Insgesamt wird daher ein *Random-Intercept-ANCOVA-Modell* gerechnet. Mit folgenden Gleichungen lässt sich das Modell beschreiben (z. B. Hox, 2010; Luke, 2010):

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + r_{ij} \quad (10)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} w_j + u_{0j} \quad (11)$$

$$\Rightarrow y_{ij} = \gamma_{00} + \gamma_{01} w_j + u_{0j} + \beta_1 x_{ij} + r_{ij} \quad (12)$$

Wie von Hox (2010) empfohlen, wird das Modell schrittweise aufgebaut, indem zunächst ein Nullmodell zur Ermittlung des ICC spezifiziert wird und dann die Prädiktoren hinzugenommen werden, wobei zunächst die fixierten Variablen und danach die variabel modellierten Variablen hinzugefügt werden. Da die Unterrichtsaktivitäten aus theoretischer Sicht möglicherweise konfundiert sind und forschend-entdeckender Unterricht beispielsweise auch interaktiv ist und Anwendungen enthält, werden sie nacheinander in folgender Reihenfolge hinzugenommen: (1) Interaktives Lehren und Lernen, (2) Experimente und Forschend-entdeckende Elemente und (3) Anwendungen.

In späteren Analysen zur Interaktion von Profilen und Unterricht im Zusammenhang mit Interesse werden zusätzlich *Cross-Level-Interaction-Modelle* als eine Art *Intercepts-and-Slopes-as-Outcomes-Modell* gerechnet. In diesen Modellen wird zugelassen, dass ein Zusammenhang auf Level 1 zwischen Level-2-Einheiten, hier Schulklassen, variiert.

Multigruppenvergleich zwischen High und Nicht-High-Performern

Da in dieser Arbeit Unterschiede im Zusammenhang von Unterrichtsmerkmalen mit Kompetenz und Interesse zwischen High und Nicht-High-Performern untersucht werden, ist neben der Mehrebenenstruktur ein Multigruppen-Analysedesign notwendig. Asparouhov und Muthén (2012) unterscheiden zwischen Multigruppen-Modellen mit Gruppierungsvariable auf Level 1 (*within level group variable*) und Level 2 (*between level group variable*). Die Gruppierung nach High und Nicht-High-Performern stellt eine Gruppierungsvariable auf Level 1 dar. Eine Gruppierung auf Level 1 hat zur Konsequenz, dass die Variablen der beiden Gruppen abhängig voneinander sind und die Gesamt-Log-Likelihood daher nicht als einfache Summe der Log-Likelihoods der beiden Gruppen

berechnet werden können (Asparouhov & Muthén, 2012). Sieben verschiedene Modelle sind zur Durchführung eines solchen Multigruppenvergleichs möglich und erfordern unterschiedlich viele Dimensionen numerischer Integration, d. h. sehr unterschiedliche Rechenkapazitäten (Asparouhov & Muthén, 2012). Da die Unterrichtseinschätzungen auf Klassenebene im vorliegenden Modell die Einschätzung der kompletten Klasse (anstatt jeweils nur die Angaben der High- bzw. Nicht-High-Performer) enthalten sollen, eignet sich nur ein Modell mit manifest gebildeten Mittelwerten auf Level-2. Dies entspricht Modell H3 von Asparouhov und Muthén (2012), d. h. einem Zwei-Ebenen-Zwei-Gruppen-Modell mit Level-1-Gruppierungsvariable und unabhängigen Effekten.

Inwieweit Unterschiede zwischen High und Nicht-High-Performern im Multigruppenmodell signifikant sind, wurde mit Log-Likelihood-Differenzen-Tests zwischen Modellen mit zwischen High und Nicht-High-Performern gleichgesetzten Parametern und Modellen mit freien Parametern geprüft (Asparouhov & Muthén, 2012).

6.4.3. Latente Profilanalyse

Zur Identifikation verschiedener Selbstkonzept-Interesse-Typen unter den High Performern in Naturwissenschaften wird eine Latente Profilanalyse (*latent class analysis, LCA*, Lazarsfeld, 1950) verwendet. Sie wird zum Einen verwendet, weil sie als personenzentrierte Analyse aus inhaltlicher Sicht Vorteile zur Beantwortung der Forschungsfragen zu wenig leistungszuversichtlichen und uninteressierten High Performern bietet. Zum Anderen hat sie innerhalb der personenzentrierten Analysen einen methodischen Mehrwert, indem sie eine probabilistische statt einer deterministischen Typenzuordnung ermöglicht (für eine detaillierte Darstellung s. Vermunt & Magidson, 2002).

Mehrwert personenzentrierter Analysen

Die Latente Profilanalyse (LPA) ist eine personenzentrierte Analyse zur Identifikation verschiedener Merkmalsprofile bzw. Typen. Wie Köller (1998) im Detail darlegt, waren typologische Ansätze in der differenziellen Psychologie lange Zeit umstritten und dimensionale Ansätze, die Traits statt Typen benutzen, wurden bevorzugt. Demnach wurde an typologischen im Vergleich zu dimensional Ansätzen kritisiert, dass sie zu stark vereinfachen, indem sie in Kategorien statt kontinuierlichen Skalen denken und explorativ statt hypothesengeleitet vorgehen (z. B. Amelang & Bartussek, 1990). Andere Autoren (z. B.

Meehl, 1992) sahen jedoch Typen als Ursache der kontinuierlich wirkenden Ausprägung von Merkmalen und die Kritik richtete sich vor allem gegen typologische Verfahren zur Untersuchung polarer Typen (extreme Ausprägungen von Indikatoren, Typ 1. Art nach Herrmann, 1976) und Typen, die lediglich begrifflicher statt empirischer Art gebildet wurden (Typ 3. Art nach Herrmann, 1976). Typologische Ansätze zur Identifikation von Personenklassen mit bestimmter Merkmalsstruktur (Taxa nach Meehl, 1992, Typ 2. Art nach Herrmann, 1976) standen weniger in der Kritik (vgl. Köller, 1998). Während sie in der Forschung zur Zielorientierung bereits seit mehreren Jahrzehnten angewandt werden (z. B. Ames & Archer, 1988; Köller, 1998; Meece & Holt, 1993), werden personenzentrierte Analysen in der Lehr-Lernforschung erst in den letzten zehn Jahren zunehmend verwendet und mittlerweile gehäuft eingesetzt (z. B. Seidel, 2006; Lau & Roeser, 2008; Wormington et al., 2012). Personenzentrierte Analysen wie die Latente Profilanalyse ermöglichen, mehrere Variablen in ihrer Kombination und Organisation in einer Person zu betrachten (Lau & Roeser, 2008).

Die Latente Profilanalyse als probabilistisches personenzentriertes Verfahren

Die Latente Klassenanalyse (LCA, Lazarsfeld, 1950) bzw. Latente Profilanalyse (LPA, Lazarsfeld & Henry, 1968) schätzen die Anzahl homogener Klassen in einer heterogenen Stichprobe in Bezug auf Muster an beobachtbaren Antworten (Vermunt & Magidson, 2002). Während sich die Latente Klassenanalyse nur für kategoriale Variablen eignet, existiert mit der Latenten Profilanalyse auch ein entsprechendes Verfahren für kontinuierliche Variablen (z. B. Pastor, Barron, Miller, B. J. & Davis, 2007). Latente Klassenanalyse und Latente Profilanalyse sind Mischverteilungsmodelle, bei der die mischende Variable ebenso wie beim Rasch-Modell als latente, jedoch diskrete Variable angenommen wird (z. B. J. Rost, 2004). Dabei ist das Ziel, die Daten zu entmischen, indem eine latente Variable Unterschiede erklärt und beobachtbare Zusammenhänge verschwinden, sobald die latente Variable identisch ist (lokale stochastische Unabhängigkeit).

Gegenüber deterministischen Clusterverfahren haben LCA und LPA den Vorteil, dass unterschiedliche Skalenarten und -verteilungen möglich sind, d. h. die Nichtvergleichbarkeit bei der Gewichtung von Variablen wie bei deterministischen Verfahren unproblematisch ist (Bacher, Pöge & Wenzig, 2010b). Dies ist insbesondere für die Betrachtung der Subgruppe der High Performer von Bedeutung, da hier die Bedingung der Normalverteilung von Selbstkonzept und Interesse aufgrund des Zusammenhangs mit

Leistung nicht eingehalten werden könnte.

Darüber hinaus modellieren LCA und LPA aufgrund ihrer probabilistischen Natur Messfehler mit und werden durch irrelevante Variablen weniger verzerrt (Bacher et al., 2010b; J. Rost, 2004). Dies wird erreicht, indem Personen bei einer LCA bzw. LPA einer Klasse nicht manifest zugeordnet werden, sondern für jede Person für jede Klasse eine Zuordnungswahrscheinlichkeit bestimmt wird. LCA und LPA bieten zudem formal begründete Maßzahlen für die Klassenanzahl und errechnen – anders als deterministische Verfahren bei überlappenden Clustern – erwartungstreue Schätzer für Klassenzentren (Bacher et al., 2010b). Bei Clusteranalysen müssen darüber hinaus „relativ willkürlich“ (J. Rost, 2004, S. 156) die Distanzmaße zwischen Personen und Clustern sowie der Algorithmus zur Clusterung festgelegt werden, was bei der LPA ebenfalls nicht notwendig ist.

Nachteile von LCA und LPA sind, dass sie eine große Stichprobe benötigen und sie verzerrt schätzen, wenn die Annahmen der lokalen stochastischen Unabhängigkeit nicht erfüllt sind (Bacher et al., 2010b; Kaufmann & Pape, 1984). Auch die hohe Parameterzahl des latenten Klassenmodells kann einen Nachteil haben: Das Modell passt auf jeden Datensatz, wenn nur genügend viele Klassen angenommen werden. Dies widerspricht jedoch dem Ansatz, hypothesentestend und theoriegeleitet vorzugehen und liefert wenig theoretischen Erklärungswert. Obwohl die LCA und LPA prinzipiell explorative Methoden sind, werden sie vor allem dann gewinnbringend nutzbar, wenn sie theoriegeleitet und hypothesenprüfend eingesetzt werden (J. Rost, 2004).

Latente Klassenanalyse und Latente Profilanalyse gehen wie oben erwähnt mit unterschiedlich vielen Möglichkeiten für das Skalenniveau von Variablen einher. Während in der Latenten Klassenanalyse nur kategorielle Merkmale, häufig Einzelitems, verwendet werden können, erlaubt eine latente Profilanalyse auch kontinuierliche Merkmale, z. B. Summenskalen. Vorteil der Verwendung aller Einzelitems einer Skala ist grundsätzlich, dass keine Informationen verloren gehen (*full information method*), da keine Datenaggregation erfolgt (J. Rost, 2004). Da sich die Zahl der möglichen Muster mit der Anzahl der Items exponentiell erhöht, erfordert eine LCA mit mehreren Items andererseits hohe Stichprobenzahlen (Bacher et al., 2010b; J. Rost, 2004). Ein weiterer Nachteil ist, dass Unreliabilitäten von Einzelitems nicht ausgeglichen werden können und in die Analyse eingehen. Bei Verwendung von Summenskalen wiederum haben Unreliabilitäten von Items eine geringere Auswirkung, Informationen einzelner Items gehen jedoch verlo-

ren. Die Bildung von Item-Bundles stellt eine mögliche Lösung dar, mit deren Hilfe die Aussagekraft der ganzen Skala behalten und gleichzeitig die Nutzung von Einzelitems vermieden werden kann (Seidel, 2006; Thissen, Steinberg & Mooney, 1989). In den LPA-Analysen in dieser Arbeit wurden daher analog dem Vorgehen von Seidel (2006) alle Items zu Selbstkonzept und Interesse dichotomisiert und zwei Bundles mit jeweils zwei Items (Selbstkonzept) bzw. zwei Bundles mit jeweils zwei und drei Items (Interesse) gebildet. Eine Dichotomisierung ist zwar grundsätzlich mit einem Informationsverlust verbunden (J. Rost, 2004), die bei vierstufigen Likert-Skalen vor allem abgefragte Information von „Ablehnung vs. Zustimmung“ bleibt jedoch auch bei einer Dichotomisierung erhalten. Darüber hinaus kann eine Zusammenlegung von Kategorien sinnvoll sein, wenn einzelne Antwortkategorien zu selten gewählt werden (J. Rost, 2004). Da Selbstkonzept und Interesse in Zusammenhang mit Leistung stehen, wird für High Performer eine schiefe Antwortverteilung und eine geringe Wahl der unteren Antwortkategorien erwartet. Die Dichotomisierung der Items zur Erfassung von Selbstkonzept und Interesse ist in dieser Arbeit daher auch für den Umgang der zu erwartenden nicht-normalen Verteilung der Items zu Interesse und Selbstkonzept für High und Nicht-High-Performer sinnvoll.

Anzahl der latenten Klassen

Mit den Itembundles wurde in dieser Arbeit eine Latente Profilanalyse gerechnet. Wie viele latente Klassen am besten zur Stichprobe passen, wurde ermittelt, indem Modelle mit aufsteigender Klassenanzahl berechnet und in ihrem Modellfit verglichen wurden. Auch ein 1-Klassenmodell wurde berechnet, um zu sehen, ob überhaupt eine Klassenstruktur vorliegt (Bacher et al., 2010b). Als Maß für den Vergleich der Modellpassungen wurden das Informationsmaß von Akaike (AIC) (Akaike, 1973, 1974; Kaufmann & Pape, 1984) und das Bayesian Informationskriterium (BIC) (Schwarz, 1978) verwendet. Der AIC berücksichtigt zwar, dass bei einer größeren Klassenzahl automatisch bessere Modellanpassungen erzielt werden, tendiert jedoch zu einer Überschätzung der Modellanpassung und Klassenzahl (Bacher et al., 2010b; von Davier, 1997). Auch beim BIC wird die Klassenzahl mit eingerechnet und er zeigt sich für quantitativ-kontinuierliche Variablen als am besten geeignetes Informationsmaß (Bacher et al., 2010b; Fonseca & Cardoso, 2007). Im Vergleich zum AIC gewichtet der BIC die Parameterzahl stärker, wodurch insbesondere bei vielen Parametern bzw. komplexen Modellen das Einfachheitskriterium besser berücksichtigt wird (J. Rost, 2004). In vielen Studien wird daher der BIC als wesentliches

Entscheidungskriterium herangezogen (z. B. Huber et al., 2015; Lau & Roeser, 2008; Seidel, 2006). In Anlehnung an Bacher et al. (2010b) werden zusätzlich die reaktive Verbesserung gegenüber dem Nullmodell der 1-Klassenlösung pv_{0k} sowie die relative Verbesserung gegenüber dem vorausgehenden Modell pv_k berechnet (vgl. Formeln 13 und 14).

$$pv_{0k} = 1 - \frac{|LL_k|}{|LL_1|} \quad (13)$$

$$pv_k = 1 - \frac{|LL_k|}{|LL_{k-1}|} \quad (14)$$

An den Werten der prozentualen Verbesserung wird sichtbar, in welchen Modellübergängen große Verbesserungen der Passung stattfinden und ab wann ungefähr eine Sättigung beginnt. In der vorliegenden LPA wird zunächst anhand von BIC und AIC sowie den prozentualen Verbesserungen die am besten passende Klassenanzahl ermittelt und durch den Likelihood-Quotiententest abgesichert (vgl. Bacher et al., 2010b). Darüber hinaus wurden Zuordnungswahrscheinlichkeiten, Klassengrößen und -profile inhaltlich überprüft, da eine auf statistischer Kriterien basierte Entscheidung nur dann sinnvoll ist, wenn eine Klassenlösung auch theoretisch gerechtfertigt werden kann (z. B. Bacher et al., 2010b).

Zur Beschreibung und Interpretation der Klassenlösung werden wie üblich die Profilverläufe beschrieben (z. B. Seidel, 2006; Lau & Roeser, 2008; Wormington et al., 2012) und darüber hinaus paarweise Unterschiede in den Variablen zwischen Klassen auf Signifikanz geprüft.

7. Ergebnisse

7.1. Charakterisierung von High Performern in Naturwissenschaft

Im diesem Kapitel werden High Performer anhand individueller, familiärer und institutioneller Merkmale charakterisiert. Dafür wurden einerseits besondere Merkmale im Vergleich zu Nicht-High-Performern identifiziert und andererseits untersucht, wie heterogen die Gruppe der High Performer in diesen Merkmalen ist. Das nachfolgende Kapitel gliedert sich in die Betrachtung individueller sowie familiärer und institutioneller Merkmale und enthält innerhalb dieser Abschnitte zunächst jeweils Vergleiche von High und Nicht-High-Performern sowie Untersuchungen zur Heterogenität innerhalb der High Performer. Es werden jeweils sowohl 2006 als auch 2012 betrachtet und untersucht, ob beide High-Performer-Jahrgänge miteinander vergleichbar sind. In einem letzten Unterkapitel wird schließlich die Verbindung zu bisherigen Forschungsergebnissen zu Top Performern (Jugendliche auf den beiden obersten Kompetenzstufen) in PISA (z. B. OECD, 2009b) hergestellt, indem untersucht wird, inwieweit sich die High Performer, die noch nicht zu den Top Performern gehören, in ausgewählten individuellen, familiären und institutionellen Merkmalen von den Top Performern unterscheiden.

7.1.1. Individuelle Merkmale

Gemäß der in dieser Arbeit verwendeten Definition von High Performern enthielten die Stichproben aus 2006 und 2012 je einen Anteil von 20% an High Performern in Naturwissenschaften. Sie erreichten eine mittlere naturwissenschaftliche Kompetenz von $M_{HP,06} = 628.04$ ($SD_{HP,06} = 44.68$, $SE_{HP,06} = 72.76$) bzw. $M_{HP,12} = 623.16$ ($SD_{HP,12} = 45.66$, $SE_{HP,12} = 6.92$). High Performer erreichten damit durchschnittlich eine Kompetenz im oberen Bereich der vierten Kompetenzstufe, während Nicht-High-Performer

eine mittlere Kompetenz von $M_{\text{NHP},06} = 485.03$ ($SD_{\text{NHP},06} = 72.76$, $SE_{\text{NHP},06} = 3.20$) bzw. $M_{\text{NHP},12} = 485.70$ ($SD_{\text{NHP},12} = 74.16$, $SE_{\text{NHP},12} = 2.84$), d. h. im Mittel Kompetenzstufe III aufwiesen. Wie aufgrund der Definition zu erwarten, charakterisieren sich High Performer nicht nur durch eine höhere Kompetenz, sondern auch durch eine wesentlich kleinere Standardabweichung als Nicht-High-Performer. Die mittleren Kompetenzwerte von High und Nicht-High-Performern waren jeweils zwischen 2006 und 2012 vergleichbar und unterschieden sich unter Berücksichtigung des *Linking Error* nicht statistisch bedeutsam.

In den einzelnen naturwissenschaftlichen Teilkompetenzen erreichten High Performer im Mittel ähnlich hohe Werte (vgl. Tabelle 11). In den Teilkompetenzen *Naturwissenschaftliche Evidenz nutzen* ($M_{\text{diff}} = 23.68$, $SE_{\text{diff}} = 3.20$, $p < .001$) und *Naturwissenschaftliche Phänomene erklären* ($M_{\text{diff}} = -29.17$, $SE_{\text{diff}} = 3.22$, $p < .001$) waren sie jedoch signifikant besser als in der Teilkompetenz *Naturwissenschaftliche Fragestellungen erkennen*. Nicht-High-Performer erreichten dagegen in allen Teilkompetenzen vergleichbare Werte. Sie wiesen höhere Standardabweichungen auf als High Performer, was in Anbetracht der Tatsache, dass sie 80% der Stichprobe erhalten, zu erwarten war.

Tabelle 11.

Naturwissenschaftliche Teilkompetenzen von High und Nicht-High-Performern in PISA 2006

Kategorie	Teilkompetenzen								
	Phänom. erklären			Fragest. erkennen			Evidenz nutzen		
	<i>M</i>	<i>SD</i>	<i>(SE)</i>	<i>M</i>	<i>SD</i>	<i>(SE)</i>	<i>M</i>	<i>SD</i>	<i>(SE)</i>
HP	631.49	53.53	(2.04)	607.81	49.82	(2.47)	636.98	53.82	(2.06)
NHP	486.96	78.06	(2.84)	484.57	71.23	(2.75)	483.77	84.06	(3.45)

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

Insgesamt bestätigten die Ergebnisse, dass die besten 20% der Schülerinnen und Schüler in Deutschland hohe Leistungen in Naturwissenschaften und im Mittel Kompetenzstufe IV erreichen. Sie zeigten auch, dass diese Schülerinnen und Schüler in allen Teilkompetenzen hohe Werte erzielten und keine „Schwächen“ in Naturwissenschaften hatten, sie jedoch in den Teilkompetenzen *Naturwissenschaftliche Phänomene erklären* und *Naturwissenschaftliche Evidenz nutzen* ihre Stärken zeigten. Hier unterschieden sie sich von der größeren Gruppe der Nicht-High-Performer, die sich nicht durch relative Stärken oder Schwächen in den Teilkompetenzen auszeichnete.

Kompetenz in Mathematik und Lesen

Wie Tabelle 12 zeigt, waren High Performer in Naturwissenschaften nicht nur in den Naturwissenschaften leistungsstark, sondern übertrafen Nicht-High-Performer im Mittel auch in ihrer mathematischen Kompetenz und Lesekompetenz (Mathe: $M_{\text{diff},06} = 127.59$, $SE_{\text{diff},06} = 4.64$, $p < .001$; $M_{\text{diff},12} = 130.38$, $SE_{\text{diff},12} = 5.14$, $p < .001$; Lesen: $M_{\text{diff},06} = 123.08$, $SE_{\text{diff},06} = 5.72$, $p < .001$; $M_{\text{diff},12} = 114.93$, $SE_{\text{diff},12} = 6.19$, $p < .001$). Durchschnittlich erreichten High Performer damit auch in Mathematik und Lesen jeweils ungefähr Kompetenzstufe IV, während Nicht-High-Performer im Durchschnitt Kompetenzstufe II erzielten. Zwischen 2006 und 2012 haben sich die Durchschnittswerte der Kompetenzen in Mathematik und Lesen von High und Nicht-High-Performern nicht signifikant verändert und beide High-Performer-Jahrgänge sind in ihrer Mathematik- und Lesekompetenz vergleichbar.

Tabelle 12.

Vergleich der mathematischen Kompetenz und Lesekompetenz von High und Nicht-High-Performern in 2006 und 2012

Jahr	Kategorie	Mathem. Kompetenz			Lesekompetenz		
		<i>M</i>	<i>SD</i>	<i>(SE)</i>	<i>M</i>	<i>SD</i>	<i>(SE)</i>
2006	HP	603.02	53.14	3.32	596.00	54.55	3.94
	NHP	475.43	71.68	2.91	472.92	83.63	3.46
2012	HP	609.04	51.32	3.96	591.21	49.67	4.76
	NHP	478.66	72.54	2.30	476.28	76.42	2.64

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

Neben der mittleren Kompetenz ist jedoch vor allem von Relevanz, welche Varianz innerhalb der High Performer existiert und, inwieweit High Performer in Naturwissenschaften eher in Naturwissenschaften begabte oder universal über Domänen hinweg leistungsstarke Jugendliche sind. Dazu wurde untersucht, welcher Anteil der High Performer in Naturwissenschaften auch in Mathematik oder Lesen zu den besten 20% gehörte. Abbildung 5 veranschaulicht diese Überschneidungsbereiche der High Performer in Naturwissenschaften, Mathematik und Lesen. Die Analysen ergaben, dass in beiden PISA-Durchgängen lediglich rund 13% der High Performer in Naturwissenschaften nur in Naturwissenschaften zu den besten 20% gehörten und eine Art „Inselbegabung“ in Naturwissenschaften besaßen. Jeweils ungefähr 70% gehörten gleichzeitig zu den besten 20% in Mathematik bzw. Lesen. Der Überlappungsanteil von Naturwissenschaften und

Mathematik war dabei höher als der Überlappungsanteil von Naturwissenschaften und Lesen. Insgesamt waren über die Hälfte der High Performer in Naturwissenschaften in allen drei Domänen leistungsstark und damit vielseitig hochkompetent statt inselbegabt in Naturwissenschaften.

Insgesamt bestätigten die Ergebnisse zu den Kompetenzen die Hypothese, dass High Performer auch in Mathematik und Lesen höhere Kompetenzen aufweisen als Nicht-High-Performer. Analog zu bestehenden Ergebnissen zu Top Performern (OECD, 2009b), konnte die Annahme, dass ein hoher Anteil leistungsstarker Schülerinnen und Schüler in Naturwissenschaften auch in Mathematik und Lesen exzellent ist, auch für die High Performer bestätigt werden. In diesen Eigenschaften waren die High Performer in PISA 2006 und 2012 dabei vergleichbar.

Schulnoten

Für die Beurteilung und Kommunikation von Schülerleistungen sind jedoch die Schulnoten und nicht das Ergebnis eines standardisierten Tests das entscheidende quantitative Maß. Noten prägen, wie erfolgreich sich Schülerinnen und Schüler in Fächern erleben und in welchen Bereichen sie ihre Stärken oder Schwächen sehen. In einem zweiten Schritt wurden daher neben den Kompetenzen die Schulnoten in den Naturwissenschaften, Mathematik und Lesen betrachtet.

Da die naturwissenschaftsbezogenen Noten nur in PISA 2006 verfügbar sind, beschränken sich die Analysen auf diesen Durchgang. Für eine bessere Übersicht wurden die Noten der Schülerinnen und Schüler in Naturwissenschaften, Mathematik und Deutsch hier in drei Kategorien eingeteilt, in denen zwischen (a) guten bzw. sehr guten („+“: Note 1 oder 2), (b) mittleren („o“: Note 3) und (c) ausreichend bis ungenügenden Noten („-“: Noten 4, 5 oder 6) unterschieden wurde. Für die Naturwissenschaften wurde dabei der gemittelte Wert der Noten in den einzelnen naturwissenschaftlichen Fächern (Physik, Chemie und Biologie) und, wenn vorhanden, im Fach Naturwissenschaften verwendet. Damit die Notenwerte vergleichbar sind, wurden in diese Analyse lediglich die Gymnasiasten eingeschlossen, zu denen 75% der High Performer und 21% der Nicht-High-Performer 2006 gehörten ($N = 3322$, $N_{HP} = 1571$, $N_{NHP} = 1751$).

High Performer erreichten zu $M_{HP,+N} = 42.77\%$ ($M_{NHP,+N} = 17.75\%$) eine gute oder sehr gute Note in Naturwissenschaften. Insgesamt $M_{HP,+N+M} = 26.10\%$ der High Performer erreichten sowohl in Naturwissenschaften als auch in Mathematik ei-

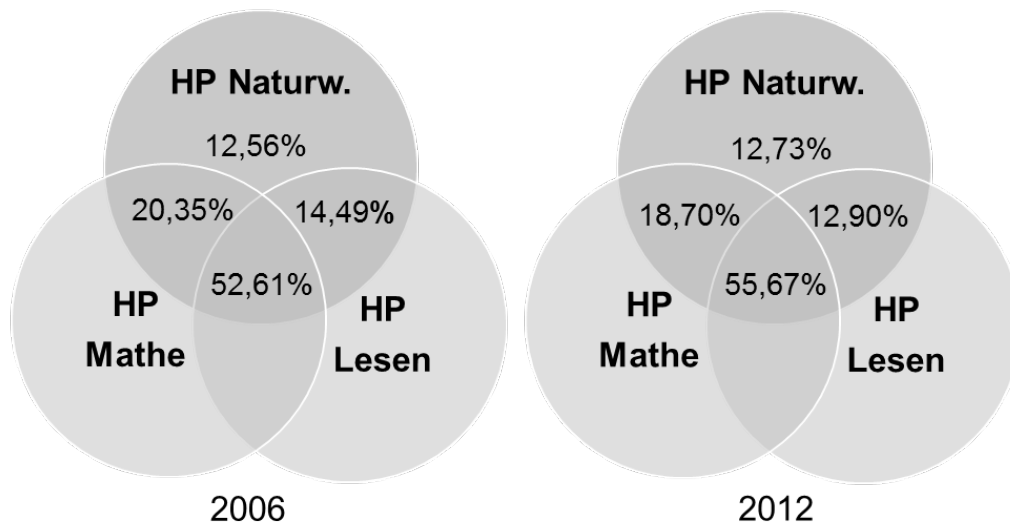


Abbildung 5. Überschneidung der High Performer in Naturwissenschaften, Mathematik und Lesen in PISA 2006 und 2012

ne sehr gute oder gute Note. Andererseits erhielten jedoch auch $M_{HP,-N} = 14.03\%$ ($M_{NHP,-N} = 36.77\%$) der High Performer eine Durchschnittsnote in den naturwissenschaftlichen Fächern von 3.5 oder schlechter. Die Leistungen dieser High Performer in Naturwissenschaft werden in der Schule als ausreichend, mangelhaft oder ungenügend eingestuft.

Neben der allgemeinen Schulleistung in Naturwissenschaften stellt sich die Frage, welches Notenprofil High Performer in Naturwissenschaften, Mathematik und Deutsch aufweisen. Dies zeigt einerseits, inwiefern sich die oben diskutierten Kompetenzprofile bzw. Überschneidungen von High Performern in der Schulleistung wiederfinden lassen. Andererseits ist anhand von Notenprofilen sichtbar, inwieweit High Performer hinsichtlich der Schulnoten eine relative Stärke in Naturwissenschaften bzw. ein naturwissenschaftliches Profil aufweisen oder beispielsweise in allen Fächern gute Noten erreichen. Tabelle 13 zeigt die prozentualen Anteile verschiedener Notenkombinationen in Naturwissenschaften, Mathematik und Deutsch für High und Nicht-High-Performer in Naturwissenschaften. Die Notenkombinationen sind danach geordnet, in welchem Fach bzw. Fächern sie ein Profil andeuten.

Rund ein Viertel der High Performer hatte in Naturwissenschaften ($M_{N,HP} = 12.51\%$) bzw. Naturwissenschaften und Mathematik ($M_{NM,HP} = 12.34\%$) bessere Noten als in Deutsch bzw. Mathematik und Deutsch und damit ein naturwissenschaftliches bzw. mathematisch-naturwissenschaftliches Notenprofil. Knapp 14% der High Performer hatten umgekehrt jedoch ein Notenprofil mit stärkster Note in Deutsch und

Tabelle 13.
Notenprofil von High und Nicht-High-Performern 2006

Profil	Notenkategorien			HP		NHP			
	Naturw.	Mathem.	Deutsch	M (%)	(SE)	M (%)	M (%)	(SE)	M (%)
Naturw.	o	-	-	4.08	(0.67)	12.51	5.24	(0.67)	10.47
	+	-	-	0.60	(0.25)		0.37	(0.18)	
	+	-	o	1.48	(0.37)		1.11	(0.29)	
	+	o	-	1.30	(0.38)		0.77	(0.29)	
	+	o	o	5.04	(0.74)		2.98	(0.52)	
Mathe	-	o	-	1.36	(0.40)	8.39	4.13	(0.51)	10.50
	-	+	-	0.46	(0.23)		0.56	(0.23)	
	-	+	o	0.31	(0.17)		0.43	(0.16)	
	o	+	-	1.16	(0.28)		1.57	(0.37)	
Deutsch	o	+	o	5.11	(0.66)	13.58	3.81	(0.53)	21.42
	-	-	o	4.40	(0.60)		10.42	(1.23)	
	-	-	+	0.73	(0.26)		2.53	(0.46)	
	-	o	+	0.48	(0.25)		0.97	(0.29)	
	o	-	+	2.71	(0.44)		2.95	(0.59)	
Naturw. u. Mathe	o	o	-	4.97	(0.86)	12.34	5.35	(0.71)	8.75
	+	+	-	1.24	(0.30)		0.59	(0.20)	
	+	+	o	6.13	(0.76)		2.82	(0.50)	
Naturw. u. Deutsch	o	-	o	6.26	(0.90)	14.51	8.50	(0.84)	11.94
	+	-	+	1.26	(0.31)		0.45	(0.18)	
	+	o	+	6.99	(0.77)		2.99	(0.70)	
Mathem. u. Deutsch	-	+	+	0.23	(0.13)	6.44	0.29	(0.15)	8.05
	-	o	o	1.93	(0.49)		4.32	(0.60)	
	o	+	+	4.29	(0.64)		3.44	(0.65)	
ausgeglichen	-	-	-	4.13	(0.63)	32.22	13.12	(1.35)	28.87
	o	o	o	9.37	(0.98)		10.08	(0.82)	
	+	+	+	18.73	(1.51)		5.68	(0.85)	

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer. -: Noten 4, 5 oder 6, o: Note 3, +: Note 1 oder 2.

14.51% hatten sowohl in Naturwissenschaften als auch Deutsch ihre besten Noten. Während bei den Kompetenzen gut 30% der High Performer ein rein mathematisch-naturwissenschaftliches Profil besaßen, gehörte damit bei den Noten nur ein Viertel der High Performer einem rein mathematisch-naturwissenschaftlichen Notenprofil an.

Rund ein Drittel der High Performer hatten in allen drei Fächern ähnlich gute Noten. Während innerhalb dieser High Performer gut 5% in allen Fächern nur ausreichend bis ungenügende Noten erreichten, sind 18.73% der High Performer Schülerinnen und Schüler, die sowohl in Naturwissenschaften als auch Mathematik und Deutsch Note Eins oder Zwei erzielten. Dieses knappe Fünftel erreichte zwar gute oder sehr gute Noten in Naturwissenschaften, hatte jedoch kein rein mathematisch-naturwissenschaftliches Profil, sondern auch in Deutsch gute Schulleistungen.

Auch unter den Nicht-High-Performern traten alle Notenprofile auf, wobei die Profile mit guten und sehr guten Noten wie erwartet tendenziell seltener und die Profile mit schlechteren Noten häufiger zu finden waren als bei den High Performern. Im Unterschied zu den High Performern in Naturwissenschaften besaßen Nicht-High-Performer zudem häufiger ein Notenprofil mit bester Note in Deutsch oder Mathematik und seltener in Naturwissenschaften oder Mathematik und Naturwissenschaften. Darüber hinaus erreichten nur knapp 5% der Nicht-High-Performer in allen drei Fachbereichen gute oder sehr gute Noten.

Insgesamt wurde die Hypothese, dass High Performer neben einer höheren Kompetenz auch bessere Schulnoten in Naturwissenschaften erreichen als Nicht-High-Performer bestätigt. Anders als erwartet erreicht jedoch nur weniger als die Hälfte der High Performer eine gute oder sehr gute Note in Naturwissenschaften und ein Teil der High Performer erreicht nur Note 4 oder schlechter. Die Hypothese, dass die Mehrheit der High Performer auch gute Schulnoten erzielt, muss daher verworfen werden.

Wie angenommen zeigte sich in den Ergebnissen zu den Schulnoten ähnlich zu den Kompetenzergebnissen, dass High Performer auch in Mathematik und Deutsch bessere Noten erzielten als Nicht-High-Performer. In Bezug auf naturwissenschaftliche bzw. mathematisch-naturwissenschaftliche Profile wurde in den Hypothesen angenommen, dass bei den Noten häufiger ein mathematisch-naturwissenschaftliches und seltener ein sprachlich-naturwissenschaftliches Profil auftrat als bei den Kompetenzen. Die Ergebnisse widerlegten diese Hypothese und zeigten umgekehrt, dass High Performer bei den Noten seltener ein naturwissenschaftliches oder mathematisch-naturwissenschaftliches

Profil und häufiger ein sprachliches oder naturwissenschaftlich-sprachliches Profil aufwiesen als bei den Kompetenzen. Lediglich bei der Überschneidung in allen drei Fächern wurde deutlich, dass High Performer bei den Noten seltener in allen drei Fächern gute oder sehr gute Noten erreichten als sie in allen drei Kompetenzen zu den High Performern gehörten.

High Performer zeigen also anders als erwartet hinsichtlich der Kompetenzen eher ein mathematisch-naturwissenschaftliches bzw. naturwissenschaftliches Profil und weniger Überlappungen zwischen den Fächern als hinsichtlich der Schulnoten. Sie sind damit nur schwer an ihrem Notenprofil zu erkennen, weil sie sich bezüglich der Noten auf ganz unterschiedliche Fächerprofile verteilen.

Kognitive Grundfähigkeit

Die Subskala „Figurale Analogien“ des KFT-12 (Heller & Perleth, 2000) wurde lediglich in PISA 2006 als Maß für die kognitiven Grundfähigkeit eingesetzt. Als Subskala bildet sie eine Facette der kognitiven Grundfähigkeit ab. High Performer wiesen im Mittel eine signifikant höhere kognitive Grundfähigkeit ($M_{\text{diff}} = 1.48$, $SE_{\text{diff}} = 0.05$, $p < .001$) auf als Nicht-High-Performer (vgl. Tabelle 14).

Tabelle 14.
Kognitive Grundfähigkeit (Subskala Figurale Analogien) im Vergleich zwischen High und Nicht-High-Performern in 2006

Kategorie	Kognitive Grundfähigkeit		
	<i>M</i>	<i>(SE)</i>	<i>SD</i>
HP	1.09	(0.03)	1.13
NHP	-0.39	(0.04)	1.42

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

Auch bei der kognitiven Grundfähigkeit ist neben dem Mittelwertsunterschied insbesondere von Bedeutung, welche Varianz innerhalb der Gruppe der High Performer existiert und, ob auch Schülerinnen und Schüler mit unterdurchschnittlicher kognitiver Grundfähigkeit High Performance erreichen. Zur Beantwortung dieser Frage wurden die Neuntklässler anhand der Standardabweichung in drei Kategorien eingeteilt: unterdurchschnittliche ($KFT < M - 1SD$), durchschnittliche ($M - 1SD \leq KFT \leq M + 1SD$)

und überdurchschnittliche ($KFT < M + 1SD$) kognitive Grundfähigkeit. Tabelle 15 zeigt die Verteilung dieser Kategorien unter den High und Nicht-High-Performern.

Wie aus der Gesamtverteilung zu erwarten, ließen sich sowohl ein Großteil der High als auch der Nicht-High-Performer der mittleren Kategorie zuordnen. Während rund 20% der Nicht-High-Performer eine in Relation zur Gesamtpopulation unterdurchschnittliche kognitive Grundfähigkeit aufwiesen, gehörten nur gut 1% der High Performer dieser Kategorie an. Umgekehrt hatten unter den High Performern fast 40% eine im Vergleich zur Gesamtpopulation überdurchschnittliche kognitive Grundfähigkeit und damit mehr als viermal so viele wie unter den Nicht-High-Performern. Es kann daher davon ausgegangen werden, dass High Performance in Naturwissenschaften mit einem gewissen Niveau kognitiver Grundfähigkeit einhergeht. Dennoch zeigten die Ergebnisse auch, dass eine überdurchschnittliche kognitive Grundfähigkeit nicht für das Erlangen von High Performance notwendig war, sondern auch Schülerinnen und Schüler mit unterdurchschnittlicher und durchschnittlicher kognitiver Grundfähigkeit zu den High Performern gehörten.

Tabelle 15.

Verteilung der kognitiven Grundfähigkeit unter High und Nicht-High-Performern

Kategorie	Kognitive Grundfähigkeit							
	unterdurchschn.		durchschn.		überdurchschn.		keine Angabe	
	<i>M (%)</i>	<i>(SE)</i>	<i>M (%)</i>	<i>(SE)</i>	<i>M (%)</i>	<i>(SE)</i>	<i>M (%)</i>	<i>(SE)</i>
HP	1.44	(0.31)	59.39	(1.18)	37.26	(1.30)	1.91	0.52
NHP	20.18	(1.11)	67.79	(1.04)	8.27	(0.46)	3.76	0.43

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

Die Ergebnisse bestätigen die Hypothese in Kapitel 5, wonach High Performer im Mittel höhere kognitive Grundfähigkeiten aufweisen als Nicht-High-Performer. Analog zu Befunden der Expertiseforschung (z. B. W. Schneider & Stumpf, 2007) konnte eine durchschnittliche kognitive Fähigkeit als eine Art Schwelle für High Performance in Naturwissenschaften identifiziert werden. Eine darüber hinaus höhere kognitive Fähigkeit war jedoch nicht hinreichend für High Performance. Vielmehr verdeutlichen die Ergebnisse, dass nicht alle Schülerinnen und Schüler mit überdurchschnittlichen kognitiven Fähigkeiten High Performance erreichten und umgekehrt nicht alle High Performer überdurchschnittliche kognitive Fähigkeiten aufwiesen. Dies bestätigt die zweite Hypothese, wonach High Performer in den kognitiven Grundfähigkeiten, obwohl sie Nicht-High-

Performer im Mittel übertreffen, eine gewisse Varianz zeigen und eine hohe kognitive Grundfähigkeit nicht hinreichend für High Performance ist. Insgesamt unterstützen die Ergebnisse damit die Annahme einer Schwelle im Sinne der Expertiseforschung.

Geschlechterunterschiede

Betrachtet man den Anteil von High und Nicht-High-Performern getrennt für Jungen und Mädchen (vgl. Tabelle 16), so wiesen Jungen sowohl in 2006 als auch 2012 einen signifikant höheren Anteil ($M_{\text{diff},06} = 5.1, SE_{\text{diff},06} = 1.68, p < .01; M_{\text{diff},12} = 3.2, SE_{\text{diff},12} = 1.39, p < .05$) an High Performern auf als Mädchen. Zwischen 2006 und 2012 fand dabei keine statistisch bedeutsame Veränderung statt. Umgekehrt betrachtet sind 42.2% (2006) bzw. 45.2% (2012) der High Performer weiblich und jeweils über die Hälfte männlich.

Tabelle 16.
Anteil der High Performer getrennt für Jungen und Mädchen

Jahr	Kategorie	Mädchen		Jungen	
		M (%)	(SE)	M (%)	(SE)
2006	HP	17.4	(1.00)	22.5	(1.36)
	NHP	82.6	(1.00)	77.5	(1.36)
2012	HP	18.5	(0.96)	21.7	(1.01)
	NHP	81.5	(0.96)	78.3	(1.01)

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

Dies bestätigt insgesamt die Hypothese, dass Jungen häufiger zu den High Performern in Naturwissenschaften gehören als Mädchen. Im Vergleich dazu wurden jedoch sowohl in 2006 als auch 2012 keine Geschlechterunterschiede in der mittleren naturwissenschaftlichen Kompetenz zwischen Mädchen und Jungen gefunden (Prenzel et al., 2007; Schiepe-Tiska et al., 2013). Die Ergebnisse deuten daher insgesamt an, dass Mädchen allgemein Jungen in der naturwissenschaftlichen Kompetenz nicht unterlegen sind, jedoch seltener die Leistungsspitze erreichen.

7.1.2. Familiäre und institutionelle Merkmale

Sozioökonomischer Hintergrund

Das familiäre Umfeld kann durch verschiedene Faktoren die Leistungsentwicklung von Kindern beeinflussen (z. B. Ehmke & Baumert, 2007). Ein zentraler und gut messbarer Einflussfaktor des familiären Umfelds ist der sozioökonomische Hintergrund der Familie. Er wird häufig durch den Index „ISEI“ quantifiziert und üblicherweise der höchste Wert (HISEI) der beiden Elternteile als Maß für den sozioökonomischen Hintergrund verwendet.

Vergleicht man High und Nicht-High-Performer 2006 und 2012 (Tabelle 17), so unterschieden sich High Performer von Nicht-High-Performern in beiden Erhebungen durch einen höheren sozioökonomischen Hintergrund ($M_{\text{diff},06} = 10.81$, $SE_{\text{diff},06} = 0.71$, $p < .001$; $M_{\text{diff},12} = 1.68$, $SE_{\text{diff},12} = 0.77$, $p < .05$). Der Unterschied zwischen beiden Leistungsgruppen ist in 2012 jedoch deutlich geringer als in 2006. Gleichzeitig sind in beiden Gruppen die Werte des HISEI von 2006 zu 2012 im Mittel angestiegen. Auch in der Gesamtstichprobe der 15-Jährigen war eine Erhöhung des HISEI von ca. 2 Punkten und eine Erhöhung der Standardabweichung zu beobachten (Müller & Ehmke, 2013). In 2012 wurde jedoch erstmals die ISCO-08 (International Labour Office, 2012) im Unterschied zur ISCO-88-Kodierung (International Labour Office, 1990) in PISA 2006 verwendet und eine Rekodierung der Werte in 2012 zeigte, dass die Differenz zwischen beiden Erhebungsrunden auch auf den Wechsel in der Kodierung zurückzuführen war (Müller & Ehmke, 2013). Auch die Veränderungen des HISEI zwischen 2006 und 2012 innerhalb der High und Nicht-High-Performer gehen damit zumindest zum Teil auf die Rekodierung zurück und können nur mit Einschränkungen interpretiert werden.

Der HISEI gibt einen kontinuierlichen quantitativen Index für den sozioökonomischen Status an. Eine Möglichkeit, Qualitätsunterschiede für die Sozialisation von Schülerinnen und Schülern in kategorialen Abschnitten sichtbar zu machen, bietet die EGP-Klassifikation (vgl. Kapitel 2.2). Tabelle 18 zeigt, wie sich die Familien von High Performern und Nicht-High-Performern auf die EGP-Klassen verteilen.

Zu den beiden Dienstklassen, d. h. den beiden höchsten EGP-Klassen, werden Berufe wie freie akademische Berufe, Beamte im mittleren, gehobenen oder höheren Dienst, Hochschul- und Gymnasiallehrer und technische Angestellte mit nicht-manueller Tätigkeit gezählt. PISA 2012 zeigte, dass Familien, die den beiden Dienstklassen ange-

Tabelle 17.
Vergleich des sozioökonomischen Status von High und Nicht-High-Performern anhand des HISEI

Jahr	Kategorie	HISEI			Differenz HP-NHP	
		<i>M</i>	<i>(SE)</i>	<i>SD</i>	<i>M</i>	<i>(SE)</i>
2006	HP	57.62	(0.61)	15.81	-10.81***	(0.71)
	NHP	46.81	(0.42)	15.74		
2012	HP	62.14	(0.64)	19.44	-1.68*	(0.77)
	NHP	60.46	(0.51)	25.86		

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

* $p < .05$, *** $p < .001$.

hören, überdurchschnittlich hohe sozio-ökonomische Ressourcen aufweisen, höhere Bildungsabschlüsse erreichen und mehr lernrelevante Besitztümer angeben, d. h. bessere kulturelle und familiäre Anregungsmilieus bieten (Müller & Ehmke, 2013). Kinder der beiden Dienstklassen besuchten darüber hinaus häufiger länger als ein Jahr einen Kindergarten bzw. nutzten andere frühkindliche Betreuungen und hatten eine höhere Gymnasialbeteiligung (Müller & Ehmke, 2013). Kinder der EGP-Klassen der Facharbeiter (V, VI) und un- und angelernten Arbeiter (VII) besuchten zwar mittlerweile häufiger ein Gymnasium als vor gut einem Jahrzehnt, insgesamt wurde jedoch auch in PISA 2012 nach wie vor ein deutlicher Unterschied in der Gymnasialbeteiligung der EGP-Klassen beobachtet (Müller & Ehmke, 2013). Umgekehrt wiesen Facharbeiter (V, VI) und un- und angelernte Arbeiter (VII) unterdurchschnittliche ökonomische, kulturelle und häusliche Ressourcen bzw. Anregungsniveaus auf und hatten üblicherweise einen nicht-akademischen Bildungshintergrund.

In beiden PISA Durchgängen stammen ungefähr die Hälfte der High Performer ($M_{HP,06} = 47.94\%$, $M_{HP,12} = 59.85\%$) aus den beiden Dienstklassen und damit günstigen sozialen Lagen. Unter den Nicht-High-Performern gehören lediglich $M_{NHP,06} = 27.15\%$ bzw. $M_{NHP,12} = 31.94\%$ den beiden obersten EGP-Klassen an. Sowohl bei den High Performern als auch bei den Nicht-High-Performern hat sich der Anteil der Schülerinnen und Schüler aus der oberen Dienstklasse von 2006 auf 2012 signifikant erhöht ($M_{HP} = 22.83\%$, $SE_{HP} = 1.89$, $p < .001$; $M_{NHP} = 9.90\%$, $SE_{NHP} = 0.83$, $p < .001$), der aus der unteren Dienstklasse reduziert ($M_{HP} = -10.92\%$, $SE_{HP} = 1.66$, $p < .001$; $M_{NHP} = -5.11\%$, $SE_{NHP} = 0.81$, $p < .001$). Auch hier ist davon auszugehen, dass Veränderungen zum Teil auf die neue ISCO-Kodierung zurückgehen (Müller & Ehmke, 2013).

Tabelle 18.

Verteilung der EGP-Klassen unter High und Nicht-High-Performern in 2006 und 2012

EGP-Klasse	2006				2012			
	HP		NHP		HP		NHP	
	M (%)	(SE)	M (%)	(SE)	M (%)	(SE)	M (%)	(SE)
Obere Dienstklasse (I)	14.35	(1.13)	7.12	(0.31)	37.18	(1.52)	17.02	(0.77)
Untere Dienstklasse (II)	33.59	(1.28)	20.03	(0.63)	22.67	(1.05)	14.92	(0.50)
Routinedienstleistungen Handel und Verwaltung (III)	27.12	(1.05)	25.42	(0.65)	4.52	(0.49)	5.30	(0.26)
Selbstständige (IV)	5.46	(0.52)	6.79	(0.37)	7.94	(0.65)	10.34	(0.42)
Facharbeiter und Arbeiter mit Leitungsfunktion (V, VI)	2.86	(0.50)	5.62	(0.26)	13.62	(0.82)	19.19	(0.64)
Un- und angelernte Arbeiter, Landarbeiter (VII)	9.47	(0.82)	19.86	(0.64)	8.10	(0.74)	16.77	(0.63)
k. A.	7.16	(0.72)	15.16	(0.67)	5.96	(1.18)	16.46	(1.04)

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

Aus den beiden niedrigsten EGP-Klassen und damit nicht-akademischen Elternhäusern stammen 2006 lediglich $M_{HP,06} = 12.33\%$ der High Performer, während $M_{NHP,06} = 25.48\%$ der Nicht-High-Performer diesen Klassen angehörten. Im Vergleich dazu stammten 2012 deutlich mehr High Performer aus den EGP-Klassen der Facharbeiter (V, VI) und un- und angelernten Arbeiter (VII) ($M_{HP,12} = 25.48\%$), allerdings ebenfalls ein höherer Anteil der Nicht-High-Performer ($M_{NHP,12} = 35.96\%$). In beiden Leistungsgruppen hat sich der Anteil der Facharbeiter (V, VI) erhöht ($M_{HP} = 10.76\%$, $SE_{HP} = 0.96$, $p < .01$; $M_{NHP} = 13.57\%$, $SE_{NHP} = 0.69$, $p < .001$). Der Anteil un- und angelernter Arbeiter (VII) hat sich bei den High Performern nicht verändert, bei den Nicht-High-Performern reduziert ($M_{HP} = -1.36\%$, $SE_{HP} = 1.10$, n.s.; $M_{NHP} = -3.09\%$, $SE_{NHP} = 0.89$, $p < .001$). Auch hier ist einschränkend zu beachten, dass 2012 eine neue Berufskodierung verwendet wurde.

Insgesamt bekräftigten die Ergebnisse die Hypothese, dass High Performer tendenziell einen höheren sozioökonomischen Hintergrund aufweisen als Nicht-High-Performer. Der Trend deutete jedoch an, dass die Differenz im sozioökonomischen Hintergrund abnahm und Schülerinnen und Schüler mit unterschiedlichen sozioökonomischen Hintergründen zunehmend ähnlich wahrscheinlich High Performance erreichen. Einschränkend muss für den Trend jedoch beachtet werden, dass die Kodierung der Berufe geändert wurde.

In Bezug auf die sozialen Schichten kamen 2012 insgesamt über die Hälfte der High Performer – damit deutlich mehr als 2006 – aus privilegierten sozialen Lagen. Ungefähr ein Fünftel der High Performer – ebenfalls ein deutlich höherer Anteil als 2006 – kamen aus niedrigeren sozialen Schichten und nicht-akademischen Elternhäusern. Damit hat sich die Zusammensetzung der High Performer derart verändert, dass sowohl hohe als auch niedrigere soziale Schichten mit einem höheren Anteil vertreten sind und sich somit die Varianz der sozialen Schichten innerhalb der High Performer verstärkt hat. Diese Verstärkung der hohen und niedrigen sozialen Lagen ist jedoch auch innerhalb der Nicht-High-Performer zu beobachten, wenngleich für die Dienstklassen in einem geringeren Ausmaß. Dies bestätigt die Hypothese, dass High Performer zu einem hohen Anteil Schülerinnen und Schüler aus hohen sozialen Schichten sind, und dennoch eine gewisse Varianz innerhalb der Gruppe der High Performer existiert. Die Ergebnisse zeigen, dass auch Schülerinnen und Schüler aus niedrigeren sozialen Lagen einen steigenden, wenngleich sehr geringen Anteil der High Performer ausmachen. In der Förderung der Jugendlichen aus diesen unteren Schichten hin zu hohen Leistungen in Naturwissenschaften liegt möglicherweise ein hohes Potential zur Erhöhung des High-Performer-Anteils. Auch bei den Ergebnissen zu den EGP-Klassen ist einschränkend zu beachten, dass Unterschiede teilweise auf eine geänderte Kodierung zurückgehen.

Verteilung auf Schularten

Die Schulart, die Schülerinnen und Schüler besuchen, kann einen bedeutenden Einfluss auf ihre Leistungsentwicklung haben (z. B. Baumert & Köller, 1998b). Für die Förderung leistungsstarker Schülerinnen und Schüler stellt sich daher die Frage, welchen Schularten diese Gruppe angehört und welcher Anteil dieser Schülerinnen und Schüler bereits am Gymnasium zu finden ist bzw. sich umgekehrt an anderen Schularten befindet.

Abbildung 6 zeigt für PISA 2006 und PISA 2012 die Schulartverteilung der High und Nicht-High-Performer. In beiden PISA Durchgängen besuchten ungefähr drei Viertel der High Performer ein Gymnasium und $M_{HP,06} = 17\%$ bzw. $M_{HP,12} = 11\%$ eine Realschule. Nur sehr geringe Anteile von jeweils unter 5% der High Performer besuchten in beiden Jahrgängen eine Hauptschule, Schulen mit mehreren Bildungsgängen oder eine integrierte Gesamtschule. In der Summe befanden sich damit dennoch ungefähr ein Viertel der High Performer in nichtgymnasialen Schularten und hatten keinen Zugang zu Lehrplänen und anderen inhaltlichen, strukturellen und pädagogischen Möglichkei-

ten, wie sie das Gymnasium zur Förderung leistungsstarker Schülerinnen und Schüler normalerweise bietet. Abgesehen vom Anteil der Realschüler, der um 5.9% ($SE = 2.05$, $p < .01$) gesunken ist, hat sich die Verteilung der High Performer auf Schularten von 2006 zu 2012 nicht verändert.

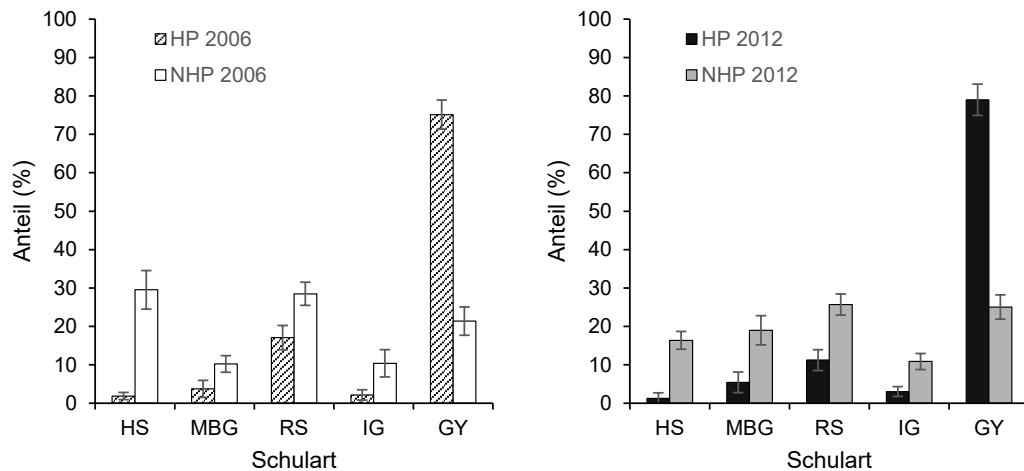


Abbildung 6. Verteilung der High und Nicht-High-Performer auf Schularten 2006 und 2012

Unter den Nicht-High-Performern besuchte wie erwartet ein wesentlich kleinerer Teil von ungefähr einem Viertel der Schülerinnen und Schüler ein Gymnasium. Vergleichsweise häufiger besuchten Nicht-High-Performer Realschulen und in 2006 Hauptschulen (vgl. Abbildung 6). Seit 2006 hat sich dabei der Anteil von Hauptschülern unter den Nicht-High-Performern signifikant verringert ($M_{diff} = -13.2\%$, $SE_{diff} = 2.76$, $p < .001$), der Anteil von Schülerinnen und Schülern aus Schulen mit mehreren Bildungsgängen signifikant erhöht ($M_{diff} = 8.8\%$, $SE_{diff} = 2.19$, $p < .001$). Da insgesamt in der Schullandschaft in Deutschland 2012 weniger Hauptschulen existierten als 2006, ist dieser Rückgang konform zur Entwicklung in der gesamten Stichprobe.

Bei der Gymnasialbeteiligung gab es sowohl unter den High als auch den Nicht-High-Performern keine signifikanten Veränderungen zu beobachten ($M_{diff,HP} = 3.9\%$, $SE = 2.78\%$, n.s.; $M_{diff,NHP} = 3.7\%$, $SE = 2.47\%$, n.s.). Die Steigerungen in beiden Gruppen von ca. 4% entsprechen der Erhöhung der Gymnasialklassen in der Stichprobe (vgl. Kapitel 6.2). In der nationalen Berichtlegung über alle Schülerinnen und Schüler hinweg zeigte sich dagegen ein Anstieg der Bildungsbeteiligung bei gleichbleibendem Anteil an Top Performern (Jugendliche auf Kompetenzstufe V und VI) in Naturwissenschaften (Sälzer, Reiss, Schiepe-Tiska, Prenzel & Heinze, 2013; Schiepe-Tiska et al., 2013). Diese Betrachtung ging jedoch auf die Stichprobe der 15-Jährigen zurück und ist mit der

hier verwendeten Stichprobe der Neuntklässler nur eingeschränkt vergleichbar, da sie auch 15-Jährige aus siebten, achten und zehnten Jahrgangsstufen enthält. Dennoch ist es auffällig, dass die Gymnasialbeteiligung unter den High Performern von 2006 auf 2012 sich nicht erhöht hat und 2012 nach wie vor ein Viertel der High Performer aus nicht-gymnasialen Schularten stammte.

Verteilung auf Schulklassen

Neben der Verteilung der High Performer auf Schularten, stellt sich die Frage, wie sich High Performer auf Schulklassen verteilen. Bisher ist unklar, ob sie als gehäufte oder verteilte Gruppe in homogenen oder heterogenen Klassen vorkommen. Im Folgenden wird daher gezeigt, wie viele High Performer im Mittel in einer Klasse vorkamen, wie viele Klassen mit besonders hohem bzw. niedrigem High-Performer-Anteil existieren und welche mittlere Leistung und Streuung die Klassen aufwiesen. Die Analysen wurden auf PISA 2006 begrenzt, da hier Naturwissenschaften Hauptdomäne waren und deshalb ein Großteil der Schülerinnen und Schüler über Klassen hinweg Naturwissenschaftstesthefte erhielt, wohingegen in 2012 zwar genügend Testhefte für eine Populationsbetrachtung vorhanden waren, eine klassenbasierte Betrachtung jedoch weniger verlässlich wäre.

Tabelle 19 zeigt den durchschnittlichen prozentualen High-Performer-Anteil von Schulklassen getrennt nach Schularten. Eine Gymnasialklasse bestand demnach im Mittel knapp zur Hälfte aus High Performern. In der Realschule gehörten gut ein Fünftel der Schülerinnen und Schüler einer Klasse zu den High Performern und in den restlichen Schularten jeweils ein Anteil unter zehn Prozent. Dabei betrug die Standardabweichung der High-Performer-Anteile unter Gymnasialklassen fast 20%. An den Realschulen und Schulen mit mehreren Bildungsgängen waren die Standardabweichungen im Vergleich zum jeweiligen Mittelwert noch höher. Am Gymnasium und vor allem auch den Realschulen und Schulen mit mehreren Bildungsgängen scheinen sich Schulklassen sehr stark in ihren High-Performer-Anteilen zu unterscheiden.

Für eine detaillierte Analyse, ob High Performer als gehäufte oder verteilte Gruppe in Schulklassen vorkommen, wurden die Schulklassen nach ihren High-Performer-Anteilen kategorisiert. Dabei wurde zwischen Klassen, in denen höchstens ein Viertel, weniger als die Hälfte, mehr als die Hälfte und mindestens drei Viertel der Schülerinnen und Schüler zu den High Performern gehörten, unterschieden. Es wurde untersucht, zu welchem Anteil Schulklassen einer Schulart sowie Schulklassen insgesamt den vier

Tabelle 19.
Durchschnittlicher High-Performer-Anteil einer Schulklasse getrennt nach Schularten

Schulart	<i>M (%)</i>	<i>(SE)</i>	<i>SD</i>
HS	1.49	(0.40)	3.23
MBG	7.97	(2.29)	9.54
RS	13.53	(1.50)	10.35
IG	4.72	(1.38)	5.40
GY	47.17	(3.98)	17.65

Anmerkung. HS: Hauptschule, MBG: Schule mit mehreren Bildungsgängen, RS: Realschule, IG: Integrierte Gesamtschule, GY: Gymnasium

unterschiedlichen Kategorien der High-Performer-Anteile angehörten (s. Tabelle 20).

Tabelle 20.
Prozentuales Vorkommen von Schulklassen mit unterschiedlichen High-Performer-Anteilen getrennt nach Schularten

Schulart	High-Performer-Anteil Klasse							
	< 25 %		25-50 %		50-75 %		> 75 %	
	<i>M (%)</i>	<i>(SE)</i>	<i>M (%)</i>	<i>(SE)</i>	<i>M (%)</i>	<i>(SE)</i>	<i>M (%)</i>	<i>(SE)</i>
HS	100.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
MBG	93.18	(9.72)	11.36	(8.96)	0.00	(0.00)	0.00	(0.00)
RS	85.94	(4.71)	13.65	(6.01)	0.51	(2.98)	0.00	(0.00)
IG	100.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
GY	10.00	(10.29)	42.92	(7.97)	40.62	(6.29)	6.46	(2.93)
Gesamt	64.39	(13.23)	19.15	(8.15)	14.22	(4.65)	2.24	(1.26)

Anmerkung. HS: Hauptschule, MBG: Schule mit mehreren Bildungsgängen, RS: Realschule, IG: Integrierte Gesamtschule, GY: Gymnasium

Alle untersuchten Hauptschulklassen und Klassen an integrierten Gesamtschulen sind Schulklassen, in denen höchstens ein Viertel der Schülerinnen und Schüler zu den High Performern gehörten. Auch an den Schulen mit mehreren Bildungsgängen und den Realschulen enthielten nur gut 10% der Klassen einen Anteil an High Performern zwischen 25% und 50%. Die vorherige Betrachtung des durchschnittlichen High-Performer-Anteils einer Schulklasse hatte gezeigt, dass an allen nichtgymnasialen Schularten im Mittel weniger als ein Viertel der Schülerinnen und Schüler einer Klasse High Performer sind. Die Betrachtung der unterschiedlichen High-Performer-Anteile von Klassen bestätigt

nun, dass dies nicht nur im Mittel zutraf und nur wenige Klassen von diesem mittleren Wert abwichen. An nichtgymnasialen Schularten waren daher auch bei einer genaueren Betrachtung keine einzelnen Klassen zu beobachten, in denen sich High Performer häuften. Lediglich an der Realschule existierte ein Anteil von einem halben Prozent an Klassen, die mindestens zur Hälfte aus High Performern bestehen und als eine Art „High-Performer-Klassen“ bezeichnet werden können. Insgesamt gehörten High Performer an nichtgymnasialen Schularten daher nicht gehäuft speziellen Klassenverbänden an, sondern sind zu einem kleinen Anteil breit verteilt in Klassen zu finden.

An den Gymnasien hatte ungefähr die Hälfte der Klassen einen Anteil von 50 Prozent oder größer und die andere Hälfte einen Anteil von 50 Prozent oder kleiner. Mit gut 80% enthielt die überwiegende Mehrheit der Gymnasialklassen einen High-Performer-Anteil zwischen einem Viertel und Dreiviertel. Nur in 6.46% der Gymnasialklassen waren mehr als Dreiviertel der Schülerinnen und Schüler High Performer in Naturwissenschaften, in 10% der Gymnasialklassen waren umgekehrt weniger als ein Viertel High Performer und Dreiviertel Nicht-High-Performer. Auch hier gruppieren sich die Klassen um den in der vorherigen Analyse gefundenen mittleren Klassenanteil und es konnte nicht beobachtet werden, dass manche Klassen sehr wenige High Performer und andere hohe Anteile enthielten. Nur ein kleiner Teil der Gymnasialklassen war eine Art High-Performer-Klasse mit sehr großen Anteilen an leistungsstarken Schülerinnen und Schülern in Naturwissenschaften. Die High-Performer-Anteile ergeben stattdessen das Bild einer leicht nach links verschobenen Normalverteilung. Obwohl sich die High Performer damit relativ gleichmäßig verteilen, ist hervorzuheben, dass knapp die Hälfte der Gymnasialklassen mehr High als Nicht-High-Performer enthielten und nur wenige Gymnasialklassen eine geringe Zahl an High Performern aufwiesen. In 90% aller Gymnasialklassen war mindestens ein Viertel der Schülerinnen und Schüler in Naturwissenschaften leistungsstark. Gymnasialklassen bieten damit nicht in einzelnen Fällen, sondern in der Breite Klassenverbände, in denen der vorhandene High-Performer-Anteil Fördermaßnahmen für diese Schülergruppe rechtfertigt.

Im Zusammenhang mit der oben untersuchten Frage, ob High Performer gehäuft oder verteilt in Schulklassen vorkommen, ist auch von Interesse, welches Leistungsniveau und welche Leistungsheterogenität Klassen aufweisen, in denen High Performer gehäuft vorkommen. Da nicht-gymnasiale Schularten in den unterschiedlichen Bundesländern sehr unterschiedlich häufig vorkommen und sich in ihrer konzeptionellen

Gestaltung unterscheiden, wurden diese beiden Aspekte nur für Gymnasialklassen untersucht. Gymnasien sind nicht nur die Schulart, die aus Sicht ihres Bildungsauftrags für die Förderung leistungsstarker Schülerinnen und Schüler am wichtigsten ist, sondern auch die einzige Schulart, die flächendeckend und vergleichbar in allen Bundesländern vorkommt.

Abbildung 7 zeigt, welche mittlere Leistung Gymnasialklassen mit unterschiedlichen High-Performer-Anteilen aufwiesen. Die mittlere Leistung wurde dahingehend kategorisiert, ob sie im Vergleich zu allen Klassen aller Schularten unterdurchschnittlich, durchschnittlich oder überdurchschnittlich ist. Sämtliche Klassen mit einem High-Performer-Anteil von unter 25% hatten eine durchschnittliche Leistung. Auch unter den Schulklassen mit 25 – 50% High Performer hatte die Mehrheit ($M = 62.88\%$, $SE = 9.99$) eine durchschnittliche sowie der Rest ($M = 37.12\%$, $SE = 9.99$) eine überdurchschnittliche Leistung. Gymnasialklassen, in denen mehr als die Hälfte bzw. mehr als Dreiviertel High Performer waren, hatten zum Großteil ($M_{50-75\%} = 94.27\%$, $SE_{50-75\%} = 4.82\%$) bzw. ausnahmslos ($M_{>75\%} = 100.00\%$, $SE_{>75\%} = 0.00\%$) eine überdurchschnittliche Klassenleistung.

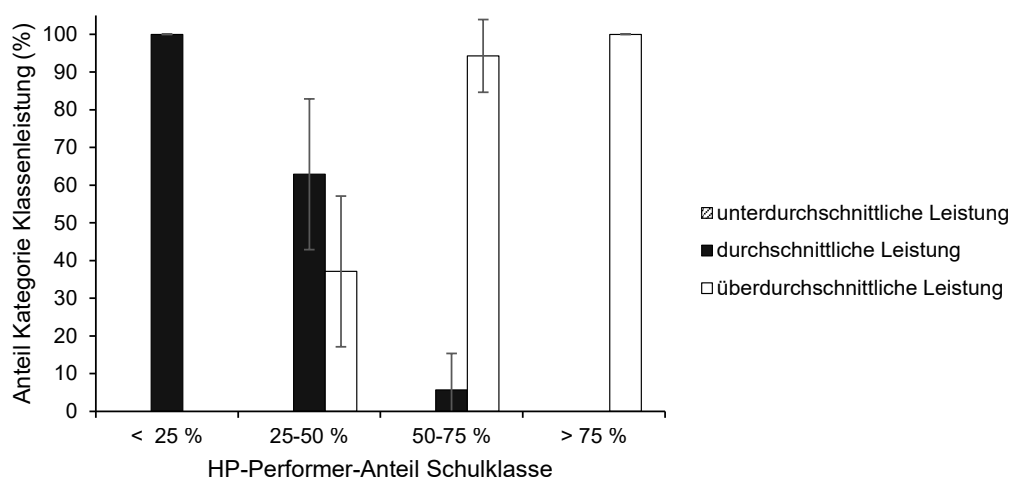


Abbildung 7. Leistungsniveau von Gymnasialklassen mit unterschiedlichen High-Performer-Anteilen (unterdurchschnittliche Leistung: $M_{\text{Klasse}} < M - 1SD$, durchschnittliche Leistung: $M_{\text{Klasse}} \in [M - 1SD, M + 1SD]$, überdurchschnittliche Leistung: $M_{\text{Klasse}} > M + 1SD$)

Insgesamt hatten alle Gymnasialklassen eine mindestens durchschnittliche Leistung. Wie als Nebenprodukt des High-Performer-Anteils zu erwarten, gehörte der Großteil der Klassen mit einem High-Performer-Anteil von über der Hälfte der Jugendlichen zu den überdurchschnittlich leistungsstarken Schulklassen.

Neben der Leistung der Gesamtklasse wurden auch die Kompetenzmittelwerte der High Performer und Nicht-High-Performer in Klassen mit unterschiedlichen High-Performer-Anteilen untersucht (vgl. Tabelle 21). Es zeigte sich, dass in Klassen mit höherem Anteil an High Performern sowohl High als auch Nicht-High-Performer höhere Leistungen erzielten. Signifikante Unterschiede ergaben sich bei beiden Leistungsgruppen im Vergleich (a) der Klassen mit Anteil unter einem Viertel und den Klassen mit Anteil bis 50% und (b) der Klassen mit Anteil bis 50% und den Klassen mit Anteil bis 75% (vgl. Tabelle 22). Klassen mit Anteil bis 75% und Klassen mit Anteil über 75% unterschieden sich dagegen nicht in der mittleren Leistung ihrer High und Nicht-High-Performer (c). Bis auf die mittleren Kompetenzwerte von High Performern in Klassen mit Anteil kleiner 25% und 25-50% sowie von Nicht-High-Performern in Klassen mit Anteil kleiner 25% unterschieden sich alle Werte von den über alle Klassen hinweg berechneten Klassenmittelwerten von High und Nicht-High-Performern. High und Nicht-High-Performer erzielten insgesamt also im Mittel in Klassen mit höherem High-Performer-Anteil höhere Kompetenzwerte. Das bedeutet, Klassen mit mehr High Performern erreichten nicht nur insgesamt eine höhere naturwissenschaftliche Kompetenz, sondern auch in den beiden Leistungsgruppen. Die Werte beziehen sich jedoch jeweils auf die mittlere Kompetenz der Leistungsgruppen gemittelt über Klassen einer Kategorie. Aussagen über die Breite des Wertebereichs und die damit verbundene Heterogenität der Klassen können daraus nicht getroffen werden.

Tabelle 21.
Mittlere Klassenleistung von High und Nicht-High-Performern getrennt nach Kategorien zur Beschreibung der High-Performer-Anteile

HP-Anteil	HP		NHP	
	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>
< 25 %	623.66	(2.88)	471.59	(10.73)
25-50 %	634.27	(2.20)	534.12	(3.18)
50-75 %	642.29	(2.40)	548.16	(2.39)
> 75 %	649.15	(8.13)	554.48	(8.77)
alle Klassen	630.62	(2.55)	496.30	(16.92)

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

Wie heterogen die Gymnasialklassen mit unterschiedlichen High-Performer-Anteilen im Detail sind, d. h. wie hoch die Leistungsstreuung in diesen Klassen ist, kann

Tabelle 22.
Vergleich der mittleren Kompetenzwerte von High und Nicht-High-Performern zwischen den Kategorien zur Beschreibung der High-Performer-Anteile

Differenz	HP		NHP	
	M_{diff}	(SE)	M_{diff}	(SE)
(< 25%) - (25-50%)	-10.61**	(3.63)	-62.53***	(11.19)
(25-50%) - (50-75%)	-8.02 *	(3.25)	-14.04**	(3.98)
(50-75%) - (>75%)	-6.86	(8.48)	-6.32	(9.09)

Anmerkung. HP: High Performer, NHP: Nicht-High-Performer.

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

mithilfe der Standardabweichung der Klassenleistung quantifiziert werden. Abbildung 8 veranschaulicht die Standardabweichung der Gymnasialklassen getrennt nach High-Performer-Anteilen. Wie bei der mittleren Klassenleistung wurde zwischen in Relation zur Gesamtpopulation unterdurchschnittlichen, durchschnittlichen und überdurchschnittlichen Standardabweichungen unterschieden und damit untersucht, ob die Klassen unter-, über- oder durchschnittlich heterogen waren.

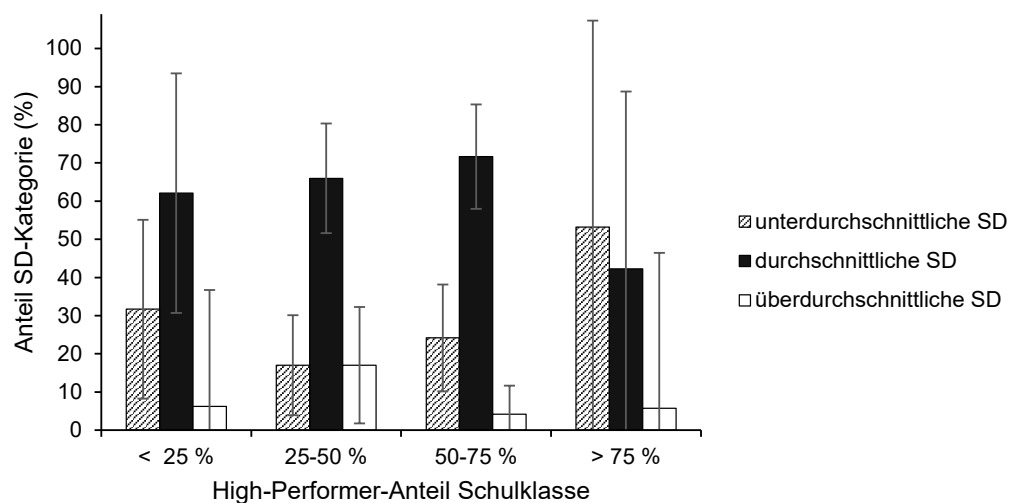


Abbildung 8. Leistungsheterogenität von Gymnasialklassen mit unterschiedlichen High-Performer-Anteilen (unterdurchschnittliche SD: $SD_{\text{Klasse}} < M - 1SD$, durchschnittliche SD: $SD_{\text{Klasse}} \in [M - 1SD, M + 1SD]$, überdurchschnittliche SD: $SD_{\text{Klasse}} > M + 1SD$)

Die Analysen zeigten, dass Gymnasialklassen mit einem hohen Anteil von High Performern eine unterdurchschnittliche ($M = 53.17\%$, $SE = 27.06\%$) oder durchschnittliche ($M = 42.22\%$, $SE = 23.25\%$) Standardabweichung aufwiesen, wobei bei beiden Anteilen jeweils der Standardfehler hoch war. Nur ein kleiner Teil der Klassen, in denen

mehr als die Hälfte High Performer waren, hatte eine überdurchschnittliche Standardabweichung ($M_{50-75\%} = 4.19\%$, $SE_{50-75\%} = 3.73\%$; $M_{>75\%} = 5.76\%$, $SE_{>75\%} = 20.34\%$). Das bedeutet, Klassen mit einem hohen Anteil an High Performern waren zwar selten besonders heterogen, umgekehrt aber nicht zwingend homogen. Vielmehr gehörte ein Großteil der Klassen mit vielen oder sehr vielen High Performern zu den durchschnittlich heterogenen Klassen. Klassen mit vielen oder sehr vielen High Performern waren damit ähnlich häufig durchschnittlich heterogen wie Klassen mit einem High-Performer-Anteil von unter 50%. Insgesamt unterschieden sich Klassen mit vielen und wenigen High Performern relativ wenig in ihrer Leistungsheterogenität, lediglich Klassen mit mehr als Dreiviertel High Performer waren tendenziell häufiger homogen, wobei auch hier der große Standardfehler keine signifikante Aussage zulässt.

Insgesamt unterstützen die Ergebnisse die Hypothese, dass ein Großteil der High Performer ein Gymnasium besucht und bestätigen frühere Ergebnisse zur Schularverteilung vielseitig hochkompetenter Jugendlicher (Zimmer et al., 2007). Am Gymnasium verteilen sich High Performer gemäß der angenommenen Hypothese relativ gleichmäßig auf die Klassen und es existieren wenige Klassen mit besonders vielen oder besonders wenigen High Performern. Wenngleich nicht unerwartet, zeigen die Ergebnisse anschaulich, dass ein beträchtlicher Teil der Gymnasialklassen mehr High als Nicht-High-Performer enthält und damit Anlass und angemessene Umgebung für Spitzenförderung bieten. Der Blick auf die mittleren Leistungen von Gymnasialklassen mit unterschiedlichen High-Performer-Anteilen bestätigte, dass Klassen mit geringen Anteilen eine durchschnittliche Leistung und Klassen mit hohen Anteilen eine überdurchschnittliche Leistung aufweisen. Gleichzeitig zeigte sich in den Analysen der Leistungsheterogenität von Klassen, dass Klassen mit vielen High Performern selten sehr heterogen oder sehr homogen sind, sondern vielmehr zu den durchschnittlich heterogenen Klassen zählen. Dies bestätigt die angenommene Hypothese, dass sich Klassen mit vielen High Performern nicht durch besondere Homogenität auszeichnen, sondern durchschnittlich homogen sind.

Vergleich von Top und High Performern

Neben dem Unterschied von High und Nicht-High-Performern sowie dem Vergleich von PISA 2006 und PISA 2012 stellt sich die Frage, inwieweit sich High Performer, die noch keine Top Performance erreicht haben, mit den Top Performern (Schülerinnen und Schüler auf Kompetenzstufe V und VI) in ihren Eigenschaften decken. Auftretende

Ähnlichkeiten und Unterschiede wurden anhand der Leistungen in Naturwissenschaften, Mathematik und Lesen sowie der kognitiven Grundfähigkeit, der Geschlechterverteilung und dem sozioökonomischen Hintergrund ermittelt.

Innerhalb der Gesamtpopulation waren 10.01% (2006) bzw. 9.99% (2012) der Population Top Performer, d. h. ungefähr die Hälfte der High Performer befanden sich 2006 und 2012 auf Kompetenzstufe V und VI. Zwischen 2006 und 2012 bestand dabei kein signifikanter Unterschied. Im Vergleich zur Stichprobe der 15-Jährigen (Anteil Top Performer: 11.8% (2006), 12.2% (2012)) enthielten die klassenbasierten Stichproben einen etwas niedrigeren Anteil an Top Performern. Dies liegt möglicherweise daran, dass die Stichprobe der 15-Jährigen auch Schülerinnen und Schüler aus höheren Jahrgangsstufen enthält.

Top Performer wiesen sowohl in 2006 als auch 2012 jeweils eine signifikant höhere mathematische Kompetenz und Lesekompetenz auf als die übrigen High Performer (vgl. Tabelle 23). Sie erreichten in Mathematik im Mittel Kompetenzstufe V, in Lesen Kompetenzstufe IV. Ähnlich wie die High Performer erzielten die Top Performer also auch in Mathematik und in Lesen im Mittel bessere Leistungen als die übrigen Schülerinnen und Schüler und insbesondere auch als die High Performer.

Tabelle 23.
Mathematische Kompetenz und Lesekompetenz von Top Performern sowie Differenzen zu den Werten der High Performer

		TP		Differenz TP-HP	
		<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>
Mathematik	2006	624.19	(4.89)	42.44***	(4.25)
	2012	630.87	(5.62)	43.42***	(4.33)
Lesen	2006	613.14	(6.18)	34.36***	(5.70)
	2012	607.81	(6.68)	33.04***	(4.95)

Anmerkung. TP: Top Performer, HP: High Performer.

*** $p < .001$.

Betrachtet man die Überschneidungen nur für die Leistungsspitze der Top Performer, so zeigte sich ein zu den High Performern ähnliches Bild. Im Unterschied zu den High Performern fanden sich in der Leistungsspitze jedoch weniger Überlappungen zwischen Naturwissenschaften und Lesen ($M_{HP,06} = 67.10\%$, $M_{TP,06} = 47.63\%$; $M_{HP,12} = 68.57\%$, $M_{TP,12} = 47.63\%$) und dementsprechend auch weniger Top Performer, die in allen drei Domänen zur Leistungsspitze gehörten. Ein Vergleich der beiden Jahrgänge zeigt, dass

mit 17.80% in 2006 zwar vergleichsweise mehr Top als High Performer existierten, die lediglich in Naturwissenschaften Spitzenleistungen aufwiesen. In 2012 waren jedoch die Anteile der nur in Naturwissenschaften zu den Top Performern gehörenden Jugendlichen (13.10%) vergleichbar zum entsprechenden Anteil unter High Performern (12.73%). Top und High Performer wiesen in beiden PISA-Durchgängen vergleichbare Kompetenzprofile auf. Die einzige Veränderung im Vergleich von 2006 und 2012 war, dass 2006 unter den Top Performern mehr „Inselbegabte“ in Naturwissenschaften waren als unter den High Performern und 2012 unter den Top Performern ähnlich wenige Jugendliche zu den „Inselbegabten“ gehörten wie unter den High Performern.

Ähnlich zur Kompetenz in Mathematik und Lesen erreichten Top Performer in der kognitiven Grundfähigkeit signifikant höhere Werte als High Performer ($M_{\text{diff},06} = 0.36$, $SE_{\text{diff},06} = 0.07$, $p < .001$). Auch unter den Top Performern gehörten jedoch 56.24% (HP: 60.8%) zu den Schülerinnen und Schülern mit unterdurchschnittlicher oder durchschnittlicher kognitiver Grundfähigkeit.

Betrachtet man innerhalb der Geschlechter die Anteile der High Performer, die nicht zu den Top Performern gehörten, so unterschieden sich die Anteile nicht zwischen Mädchen und Jungen ($M_{\text{Mae}} = 9.3\%$, $M_{\text{Ju}} = 10.7\%$). Signifikante Unterschiede existieren jedoch unter den Top Performern. Während unter den Jungen in 2006 und 2012 jeweils knapp 12% die oberen beiden Kompetenzstufen erreichten ($M_{\text{TP},06} = 11.8\%$, $SE_{\text{TP},06} = 0.95$; $M_{\text{TP},12} = 11.5\%$, $SE_{\text{TP},12} = 0.68$), gehörten unter den Mädchen mit rund 8% ein geringerer Anteil zu den Top Performern in Naturwissenschaften ($M_{\text{TP},06} = 8.2\%$, $SE_{\text{TP},06} = 0.57$; $M_{\text{TP},12} = 8.4\%$, $SE_{\text{TP},12} = 0.76$). Auch hier haben sich die Anteile zwischen 2006 und 2012 nicht statistisch bedeutsam verändert.

Beim familiären Hintergrund zeigten Top Performer ebenso wie die breitere Leistungsspitze der High Performer höhere Werte für den sozioökonomischen Hintergrund und unterschieden sich auch zusätzlich von den High Performern, die nicht zu den Top Performern gehörten. Die Verringerung der Unterschiede im sozioökonomischen Status von 2006 zu 2012 fand sich auch innerhalb der High Performer in einem Vergleich von Top Performern und den übrigen High Performern. Während Top Performer in 2006 noch einen signifikant höheren sozioökonomischen Hintergrund aufwiesen als die übrigen High Performer ($M_{\text{diff},06} = 2.66$, $SE_{\text{diff},06} = 0.90$, $p < .01$), unterschieden sich die beiden Gruppen innerhalb der High Performer 2012 nicht mehr in ihrem HISEI ($M_{\text{diff},12} = 0.65$, $SE_{\text{diff},12} = 1.61$, n.s.).

7.1.3. Zusammenfassung

Zusammenfassend verdeutlichen die Ergebnisse, dass sich die Gruppe der High Performer im Vergleich zu Nicht-High Performern im Mittel durch besondere Eigenschaften wie eine relative Stärke in zwei Teilkompetenzen, besseren Schulleistungen in Naturwissenschaften, höheren Kompetenzen in Mathematik und Lesen sowie besseren Schulnoten in Mathematik und Deutsch, höheren kognitiven Grundfähigkeiten und einem höheren sozioökonomischen Hintergrund auszeichneten sowie häufiger ein Gymnasium besuchten. Gleichzeitig zeigten die Analysen, dass innerhalb der High Performer in vielen Merkmalen eine gewisse Spannweite existierte und sie keine homogene Gruppe waren. So wurden sowohl High Performer mit einem naturwissenschaftsbezogenen Leistungsprofil als auch domänenübergreifend leistungsstarke Schülerinnen und Schüler gefunden. Hohe kognitive Grundfähigkeiten waren darüber hinaus nicht hinreichend oder notwendig für hohe Leistungen in Naturwissenschaften.

In Bezug auf die Noten legte der Vergleich von Kompetenzen und Noten dar, dass sowohl zahlreiche High Performer in Naturwissenschaft gute oder sehr gute Noten erzielten, über ein Fünftel jedoch Note ausreichend oder schlechter erhielt. Auch in Bezug auf Notenprofile zeigte sich, dass sich einerseits Kompetenz- und Notenprofile nicht deckten und andererseits ganz unterschiedliche Notenprofile unter den High Performern vorkamen. Besonders hervorzuheben ist, dass ein großer Teil in Naturwissenschaften, Mathematik und Deutsch gleich gute Noten erzielte.

Auch in Bezug auf den familiären Hintergrund wurde ersichtlich, dass die Unterschiede im sozioökonomischen Hintergrund zwischen High und Nicht-High-Performern abnahmen und möglicherweise weiteres Potential in der Förderung von Jugendlichen aus nicht-akademischen Bildungshäusern hin zu High Performance steckt. Darüber hinaus stammten High Performer in Bezug auf die mittlere Leistung und Heterogenität von Klassen nicht aus besonderen institutionellen Kontexten. Sie verteilten sich am Gymnasium relativ gleichmäßig auf Schulklassen und ein Großteil der Klassen mit vielen High Performern war durchschnittlich heterogen.

Der Vergleich der High Performer in PISA 2006 und PISA 2012 verdeutlichte, dass die leistungsstarken Schülerinnen und Schüler in beiden PISA-Durchgängen in ihren Eigenschaften überwiegend vergleichbar sind. So zeigten sich ähnliche Werte und Profile in den Kompetenzen und ähnliche Geschlechterunterschiede. Einzig im sozioökonomischen Hintergrund unterschieden sich zwar High Performer in 2006 und 2012 signifikant von

Nicht-High-Performern, der Unterschied war jedoch 2012 signifikant geringer. Auch in den Anteilen auf den EGP-Klassen sowie der Schulartenverteilung gab es Veränderungen bei den High Performern von 2006 auf 2012. Diese Veränderungen sind jedoch in der gleichen Art auch bei den Nicht-High-Performern zu beobachten, so dass beide Gruppen leistungsstarker Schülerinnen und Schüler sich in ihren diesbezüglich besonderen Eigenschaften gegenüber den restlichen Schülerinnen und Schülern decken. Insgesamt sind die High Performer in 2006 und 2012 daher in ihren Eigenschaften vergleichbar und Ergebnisse von vertieften Analysen der High Performer in 2006 können mit hoher Wahrscheinlichkeit auf die High Performer in 2012 übertragen werden.

Der Vergleich von Top Performern als Teilgruppe der High Performer mit den restlichen High Performern ergab, dass sich beide Gruppen in den besonderen Merkmalen, die sie im Vergleich zu Nicht-High-Performern auszeichnen, deckten. In einem Teil dieser Eigenschaften setzten sich Top Performer jedoch nochmal von den restlichen High Performern ab. Sie erreichten demnach höhere Leistungen in Mathematik und Lesen, waren seltener gleichzeitig Top Performer in Naturwissenschaften und Lesen und wiesen 2006 einen höheren sozioökonomischen Hintergrund auf. Auch bei den kognitiven Grundfähigkeiten erzielten Top Performer zwar höhere Werte als die restlichen High Performer, sie waren darin jedoch ähnlich heterogen. Unterschiede zwischen Top Performern und den übrigen High Performern zeigten sich nur bei den Geschlechterunterschieden. Sie wurden lediglich für Top Performer gefunden.

7.2. Unterrichtsaktivitäten im Zusammenhang mit naturwissenschaftlicher Kompetenz und Interesse von High Performern

Nachdem High Performer im ersten Ergebnisteil hinsichtlich ihrer besonderen Merkmale charakterisiert wurden, beschäftigt sich der nachfolgende Ergebnisteil mit der Frage, inwieweit die Unterrichtsaktivitäten „interaktives Lehren und Lernen“, „Experimentieren“ und „Schülerinnen und Schüler forschen“ die naturwissenschaftliche Kompetenz und das Interesse von High Performern vorhersagen sowie welche Unterschiede zu Nicht-High-Performern bestehen.

In der Analyse wird dafür ein Mehrebenenansatz verfolgt, in dem gleichzeitig ein Multigruppenvergleich modelliert wird. Damit erweitert die Arbeit bisherige Analysen zu diesen Unterrichtsaktivitäten in PISA um ein mehrebenenanalytisches Vorgehen. In der Verwendung eines Unterrichtsklimaansatzes bei der mehrebenenanalytischen Modellierung der Unterrichtsaktivitäten greift diese Arbeit darüber hinaus neuere methodische Ansätze auf und setzt sich von anderen Untersuchungen zum Naturwissenschaftsunterricht ab. Eine Besonderheit im hier verwendeten Ansatz liegt in der latent-manifesten Modellierung der Merkmale in der Mehrebenenanalyse sowie der Kombination mit einem Zweigruppenansatz. Durch die latente Bildung von Faktoren auf Schüler- und Klassenebene werden Messfehler berücksichtigt und durch die manifeste Aggregation der Klassenwerte gleichzeitig sichergestellt, dass die Klassenwerte der Unterrichtsskalen trotz Multigruppendedesign die Angaben von sowohl High als auch Nicht-High-Performer enthalten. Durch die klassenbasierte Stichprobe und die Verwendung von Klassengewichten wurde dabei die mit dem Datensatz größtmögliche Repräsentativität der Ergebnisse angestrebt. Der methodische Ansatz wurde damit möglichst auf die inhaltliche Fragestellung abgestimmt. Indem dieser Ansatz messfehlerbedingte Varianz gleichzeitig in hohem Maße kontrolliert, schließt er Artefakte aufgrund einer subjektiven Unterrichtswahrnehmung von Jugendlichen aus bzw. macht sie auf Schülerebene sichtbar. Im Vergleich zu bisherigen Analysen mit den hier verwendeten Unterrichtsaktivitäten (z. B. Kobarg et al., 2011) ist dieser Ansatz daher eher varianzeinschränkend und relativ konservativ.

Die Mehrebenenanalysen erfordern zunächst, die Faktorenstruktur der Unterrichtsmerkmale auf Klassenebene zu analysieren und festzustellen, ob die Faktoren-

struktur auf Klassenebene der Struktur auf Schülerebene entspricht. Für den Multi-gruppenvergleich ist es darüber hinaus notwendig, die gefundene Faktorenstruktur auf Messinvarianz zwischen High und Nicht-High-Performern zu prüfen. In den nachfolgenden Abschnitten wird daher in einem eigenen Kapitel zunächst die Faktorenstruktur der Unterrichtsmerkmale auf Schüler- und Klassenebene untersucht und die Messinvarianz zwischen beiden Schülergruppen geprüft. Nach diesem methodischen Einschub werden schließlich die Zusammenhänge der Unterrichtsmerkmale mit naturwissenschaftlicher Kompetenz bzw. Interesse jeweils in einem eigenen Abschnitt dargestellt.

7.2.1. Faktorenstruktur der Unterrichtsmerkmale auf Schüler- und Klassenebene

Die Faktorenstruktur und Reliabilität der Unterrichtsskalen auf Individualebene wurde bereits im Rahmen von PISA für den Datensatz der 15-Jährigen untersucht (Frey & Asseburg, 2009). Offen ist jedoch, ob sich eine ähnlich differenzierte Faktorenstruktur auch für den Klassendatensatz und insbesondere auch auf Klassenebene finden lässt oder, ob die Items der Unterrichtsaktivitäten auf Klassenebene weniger Faktoren bilden als auf Schülerebene (vgl. Kunter, 2005). In diesem Abschnitt wird daher die Faktorenstruktur der Unterrichtsaktivitäten auf Klassenebene behandelt. Bevor die Faktorenstruktur aller Unterrichtsaktivitäten untersucht wird, werden zunächst die zwei Unterrichtsmerkmale „Experimentieren“ (Exper) und „Schülerinnen und Schüler forschen“ (Forsch) aufgrund ihrer konzeptuellen Ähnlichkeit gesondert untersucht.

Explorative Mehrebenenfaktorenanalyse der Unterrichtsskalen „Experimentieren“ und „Schülerinnen und Schüler forschen“

Die Skalen „Experimentieren“ und „Schülerinnen und Schüler forschen“ bilden forschend-entdeckendes Lernen ab (vgl. Frey & Asseburg, 2009). Beide Skalen fragen die Häufigkeit von Tätigkeiten im Rahmen von Schülerexperimenten mit unterschiedlichem Fokus ab: Die erste Skala beschäftigt sich mit der Durchführung, Auswertung und Interpretation von Experimenten, die zweite Skala erfasst, inwiefern Schülerinnen und Schüler darüber hinausgehend auch eigene Fragestellungen entwickeln sowie ihre experimentelle Untersuchung planen dürfen. Da sich beide mit Schülerexperimenten in einer unterschiedlich autonomen Form beschäftigen, bilden sie empirisch möglicherweise generell bzw. auf

Level 2 einen Faktor ab. Denkbar wäre auch, dass sie mehr als zwei Faktoren entsprechend den unterschiedlichen Facetten von Experimentieren abbilden. Wie viele Faktoren und mit welcher Itemzuordnung diese Items abbilden, wurde daher zunächst mit einer explorativen Mehrebenenfaktorenanalyse überprüft.

Die explorative Faktorenanalyse ergab zwei Hauptkomponenten mit Eigenwert größer Eins auf Level 1 und eine Hauptkomponente mit Eigenwert größer 1 auf Level 2 (vgl. Abbildung 9). Ausgehend vom Kaiser-Guttman-Kriterium kann damit von zwei Faktoren auf Level 1 und einem Faktor auf Level 2 ausgegangen werden. Da der zweite Eigenwert auf Level 2 annähernd Eins erreicht und hier der Knick im Eigenwertverlauf liegt, können auf Level 2 möglicherweise auch zwei Faktoren angenommen werden. Dies wurde überprüft, indem die Fitparameter der Modelle mit einem oder zwei Faktoren auf Schülerebene bzw. einem oder zwei Faktoren auf Klassenebene verglichen wurden (Tabelle 24).

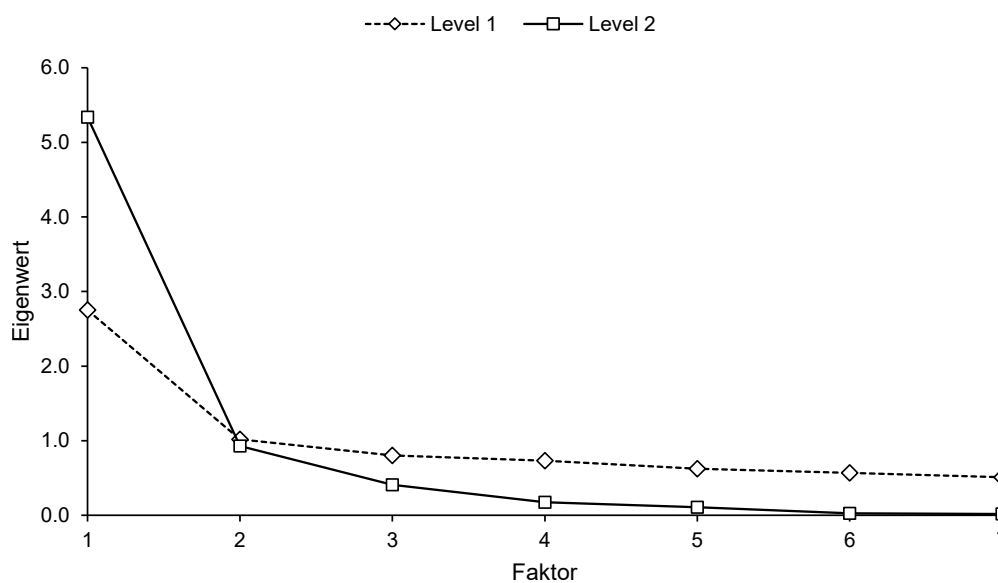


Abbildung 9. Verlauf der Eigenwerte in der explorativen Mehrebenenfaktorenanalyse

Als Indikatoren der Modellpassung werden zum einen der χ^2 -Test, sowie das Maß des Root Mean Square Error of Approximation (RMSEA) und des Comparative Fit Index (CFI) sowie das Aiken Information Criterion (AIC) und Bayesian Information Criterion (BIC) berichtet. Während der χ^2 -Test stark von der Stichprobengröße abhängt, sind RMSEA sowie CFI robuster gegenüber der Stichprobe und bevorzugen im Falle des RMSEA sparsamere Modelle (Schermelleh-Engel, Moosbrugger & Müller, 2003). AIC (Aiken Information Criterion) und BIC (Bayesian Information Criterion) stellen zwei

relative Fit-Indikatoren dar, die sich für relative Vergleiche nicht-genesteter Modelle eignen und ebenfalls die Sparsamkeit eines Modells gewichten (Schermelleh-Engel et al., 2003). Geringere Werte des AIC bzw. BIC deuten eine bessere Modellpassung an. Als Daumenregel werden Modelle mit $CFI \geq .97$ (Schermelleh-Engel et al., 2003) bzw. auch $CFI \geq .95$ (Byrne, 2012) als gut und Modelle mit $CFI \geq .95$ (Schermelleh-Engel et al., 2003) bzw. auch $CFI \geq .90$ (Byrne, 2012) als akzeptabel angesehen. Während beim CFI höhere Werte eine bessere Passung angeben, sprechen beim RMSEA niedrigere Werte für eine bessere Passung und ein $RMSEA < .05$ wird als guter, ein $RMSEA < .08$ als akzeptabler Fit interpretiert (Byrne, 2012; Schermelleh-Engel et al., 2003). Bei den Informationskriterien AIC und BIC spricht jeweils ein geringerer Wert für eine bessere Passung.

Tabelle 24 gibt eine Übersicht über die Indikatoren für die Passung der vier Modelle. Modell II und IV zeigen akzeptable bzw. gute Fit-Werte mit $RMSEA < .08$ bzw. $RMSEA < .05$ sowie $CFI \geq .95$ bzw. $CFI \geq .98$, während Modell I und II unakzeptable Fit-Werte aufweisen. Da Modell II und IV genestet sind, kann mithilfe eines Chi-Quadrat-Tests überprüft werden, welches Modell besser passt (z. B. Christ & Schlüter, 2012). Dies ergibt, dass Modell IV mit zwei Faktoren auf Level 1 und zwei Faktoren auf Level 2 am besten auf die Daten passt ($\Delta\chi^2 = 120.81, \Delta df = 6, p < .001$).

Tabelle 24.

Modellkennwerte der explorativen Faktorenanalyse von „Schülerexperimente“ und „Schülerinnen und Schüler forschen“

Nr.	Faktoren		χ^2	df	RMSEA	SRMR		CFI	TLI	AIC	BIC
	L1	L2				L1	L2				
I	1	1	540.76***	28	0.075	0.050	0.109	0.886	0.829	45267	45481
II	2	1	226.20***	22	0.053	0.022	0.107	0.955	0.913	44965	45215
III	1	2	388.48***	22	0.071	0.050	0.032	0.919	0.845	45127	45377
IV	2	2	105.40***	16	0.041	0.022	0.033	0.980	0.948	44856	45142

Anmerkung. L1 = Level 1 (Schülerebene); L2 = Level 2 (Klassenebene); df = Freiheitsgrade; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual (L1: Level 1, L2: Level 2); CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Aiken Information Criterion; BIC = Bayes Information Criterion.

*** $p < .001$.

Die Faktorladungen in Modell IV mit zwei Faktoren auf beiden Ebenen (Tabelle 25) entsprachen im Wesentlichen der theoretisch angenommenen Faktorenstruktur, die auch in der Stichprobe der 15-Jährigen nachgewiesen wurde (Frey & Asseburg, 2009). Der erste Faktor beschreibt „Schülerexperimente“, der zweite „Schülerinnen und Schüler

forschen“. Lediglich Item ST34Q03 und ST34Q06 zeigten keine eindeutig Zuordnung.

Tabelle 25.

Faktorladungen explorative Mehrebenen-Faktorenanalyse von „Schülerexperimenten“ und „Schülerinnen und Schüler forschen“

Item	Level 1		Level 2	
	Faktor 1	Faktor 2	Faktor 1	Faktor 2
ST34Q02	0.558*	0.062	1.093*	-0.006
ST34Q03	0.361*	0.190*	0.376*	0.604*
ST34Q06	0.527*	-0.067	0.439*	0.414*
ST34Q14	0.564*	-0.001	0.797*	0.093
ST34Q08	0.006	0.613*	0.090	0.881*
ST34Q11	-0.031	0.730*	-0.030	0.999*
ST34Q16	0.035	0.643*	0.000	0.990*

Anmerkung. Höchste Ladungen jeweils fettgedruckt.

Item ST34Q03 („festlegen, wie eine Fragestellung aus dem Naturwissenschaftsunterricht im Schullabor untersucht werden kann“) lädt auf beiden Ebenen auf beide Faktoren ähnlich hoch. Auch die Modifikationsindices zeigten an, dass Item ST34Q03 nicht eindeutig zugeordnet werden kann. Ursprünglich gehört es zum ersten Faktor „Schülerexperimente“, fragt im Wortlaut anders als die anderen Items jedoch nicht nur das Durchführen und Auswerten von Experimenten, sondern auch das Planen von Experimenten ab. Darüber hinaus enthält es mit „Schülerlabor“ einen starken Bezug zu einer eher forschenden Unterrichtssituation im Naturwissenschaftsunterricht und ist davon abhängig, ob die Schule ein Schülerlabor besitzt. Die Abhängigkeit davon, ob ein Schülerlabor vorhanden ist oder nicht, ist möglicherweise auch ein Grund dafür, warum dieses Item gerade auf Klassenebene auf „Schülerinnen und Schüler forschen“ lädt. Da es aus inhaltlicher und empirischer Sicht nicht eindeutig zugeordnet werden kann und die Qualität der Zwei-Faktorenlösung mindert, wurde Item ST34Q03 aus den weiteren Analysen ausgeschlossen.

Eine konfirmatorische Faktorenanalyse, in der ein Modell ohne Item ST34Q03 (Modell IV_a , $AIC = 36982.44$, $BIC = 37139.70$), ein Modell mit Item ST34Q03 bei Faktor eins (Modell IV_b , $AIC = 43109.48$, $BIC = 43290.68$) und ein Modell mit Item ST34Q03 bei Faktor zwei (Modell IV_c , $AIC = 43177.86$, $BIC = 43359.06$) verglichen wurde, bestätigte durch die relativen Indizes AIC und BIC, dass ein Modell ohne Item ST34Q03 (Modell IV_a) besser auf die Daten passt. In den CFAs wurden sehr geringe

negative Residualvarianzen auf Level 2 (*Heywood Cases*) gleich Null gesetzt, so dass eine negativ definite Residualmatrix verhindert wurde (z. B. Geiser, 2011; Hox, 2010).¹

Das zweite auffällige Item ST34Q06 lädt nur auf Level 2 auf beide Faktoren. Es erfasst, wie häufig Schülerinnen und Schüler Schlüsse aus einem Experiment ziehen. Da Schlüsse ziehen nicht nur wichtiges Element von Schülerexperimenten, sondern auch von forschendem Lernen ist, ist die Ladung auf Faktor zwei plausibel. Aus inhaltlicher Sicht bildet jedoch „Experimentieren“ die Durchführung von Schülerexperimenten (nach Vorgabe) inklusive des Schlussfolgerungsprozesses ab und dieses Item dient dazu, den Aspekt des Schlussfolgerns bei Schülexperimenten abzufragen. Es wird deshalb aus inhaltlichen Gründen beibehalten und der Skala „Experimentieren“ zugeordnet.

Konfirmatorische Mehrebenenanalyse aller Unterrichtsskalen

Nachdem die Faktorstruktur der zwei Unterrichtsskalen „Experimentieren“ und „Schülerinnen und Schüler forschen“ explorativ ermittelt wurde, wird in einer konfirmatorischen Mehrebenenfaktorenanalyse die Struktur aller Unterrichtsskalen auf beiden Ebenen betrachtet. Da frühere Analysen die Faktorstruktur und die Itemzuordnungen zu den Faktoren bestätigen (Frey & Asseburg, 2009), wurde für die restlichen Unterrichtsskalen keine explorative Mehrebenenfaktorenanalyse berichtet.²

Dem Vorgehen von Kunter (2005) folgend wird mit drei Modellen überprüft, ob die Faktorenstruktur auf Schülerebene auch auf Klassenebene zu finden ist: Modell I enthält einen Faktor auf beiden Ebenen, Modell II vier Faktoren auf Level eins (analog Frey & Asseburg, 2009) und einen Faktor auf Level zwei und Modell III besitzt schließlich vier Faktoren auf beiden Ebenen. Bei allen drei Modellen wurden die negativen Residualvarianzen nahe Null auf Level 2 analog zu oben gleich Null gesetzt. Die Ergebnisse sind in Tabelle 26 dargestellt.

Modell III mit den aus theoretischer Sicht sinnvollen vier Faktoren auf beiden Ebenen erreicht einen akzeptablen RMSEA-Wert von knapp über .05 aber einen CFI-Wert unter .90. Im Vergleich zu den anderen Modellen zeigt Modell III darüber hinaus die niedrigsten BIC- und AIC-Werte und damit die beste Passung. Die genesteten Modelle II und III wurden in robusten Chi-Quadrat-Tests nach Satorra (2000) und Satorra und

¹Da die Modelle nicht genestet sind, war nur ein Vergleich anhand relativer Modellindizes möglich.

²Zur Kontrolle wurde auch für diese Unterrichtsskalen eine explorative Faktorenanalyse gerechnet, die die auf Schülerebene in PISA 2006 gefundene Faktorenstruktur (Frey & Asseburg, 2009) für die Klassenebene bestätigt.

Tabelle 26.

Modellkennwerte der Mehrebenen-CFAs der Unterrichtsmerkmale

Nr.	Faktoren		χ^2	df	RMSEA	SRMR		CFI	TLI	AIC	BIC
	L1	L2				L1	L2				
I	1	1	5529.18***	168	0.103	0.070	0.412	0.538	0.500	91158	91495
II	4	1	2849.03***	162	0.074	0.048	0.429	0.769	0.740	89444	89817
III	4	4	1754.97***	156	0.058	0.047	0.124	0.862	0.839	88941	89349

Anmerkung. L1 = Level 1 (Schülerebene); L2 = Level 2 (Klassenebene); df = Freiheitsgrade; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual (L1: Level 1, L2: Level 2); CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Aiken Information Criterion; BIC = Bayes Information Criterion.

*** $p < .001$.

Bentler (2010) verglichen. Der Chi-Quadrat-Test wurde mit Loglikelihood-Werten unter Verwendung der Korrekturfaktoren für die Verwendung von MLR-Schätzern berechnet (vgl. auch B. Muthen & Muthen, 15.03.17). Der robuste Chi-Quadrat-Test ergibt, dass das weniger sparsame Modell III besser passt als Modell II ($\Delta\chi^2(6) = 41.82, p < .001$).

Invarianz der Unterrichtsmerkmale auf Klassen- und Schülerebene

Nachdem im vorherigen Abschnitt gezeigt wurde, dass auf beiden Analyseebenen für die Unterrichtsskalen die gleiche Faktorenstruktur angenommen werden kann, stellt sich darauf aufbauend die Frage, ob neben der Struktur auch die Faktorladungen auf beiden Ebenen als invariant angenommen werden können. Dies betrifft sowohl die Unterrichtsaktivitäten als unabhängige Variablen in den späteren Zusammenhangsanalysen als auch die Skala zu Interesse als spätere abhängige Variable. Da die naturwissenschaftliche Kompetenz jeweils nicht mit einer Skala aus mehreren Indikatoren, sondern in jeder der Analysen jeweils einer der fünf PVs für naturwissenschaftliche Kompetenz als Indikator verwendet wird, ist bei der naturwissenschaftlichen Kompetenz die Invarianz über die Ebenen per Definition gegeben.

Gleiche Faktorladungen können nicht automatisch angenommen werden (Lüdtke et al., 2009) und sind auch keine notwendige Bedingung zur Untersuchung von Klimaefekten aber sie haben verschiedene Vorteile (für eine Übersicht vgl. Morin et al., 2014). So sorgen invariante Faktorladungen einerseits dafür, dass auf beiden Ebenen die gleiche Metrik verwendet wird (Lüdtke, Marsh, Robitzsch & Trautwein, 2011), geben andererseits darüber Aufschluss, ob die Konstrukte auf beiden Ebenen einander entsprechen (Lüdtke et al., 2011; Marsh et al., 2012) und reduzieren die Komplexität des Modells. Wie

Lüdtke et al. (2011) in einer Simulationsstudie belegen konnten, führen invariante Faktorladungen zum Teil sogar zu genaueren Schätzungen, wenn ein Modell mit invarianten Ladungen schlechter auf die Daten passt als ein freies (vgl. Morin et al., 2014).

Um zu ermitteln, ob in dieser Arbeit invariante Faktorladungen für die Unterrichtsskalen angenommen werden können, wurden Modelle mit unterschiedlich vielen Invarianzen der Unterrichtsmerkmale und Interesse gerechnet und diese genesteten Modelle in robusten Chi-Quadrat-Tests verglichen. Tabelle 27 zeigt die Modellkennwerte für vier verschiedene Modelle, in denen keine Faktorladungen (Modell I), die Faktorladungen von Interesse (Modell II), die Faktorladungen aller Unterrichtsmerkmale (Modell III) und alle Faktorladungen (Modell IV) invariant gesetzt sind. Zur Vermeidung von Heywood Cases, wurden die Residualvarianzen von Interesse auf Level 2, die alle Werte um die Null aufwiesen, gemäß dem üblichen Vorgehen gleich Null gesetzt (Hox, 2010). Die Residualvarianzen der Unterrichtsskalen wurden mit Modellbedingungen auf Werte größer Null beschränkt, jedoch nicht fixiert (analog Vorgehen bei Marsh et al., 2012).

Tabelle 27.

Modelle zur Testung der Invarianz der Faktorladungen von Interesse und Unterrichtsmerkmalen zwischen Schüler- und Klassenebene

Nr.	Invarianz	χ^2	df	RMSEA	SRMR		CFI	TLI	AIC	BIC
					L1	L2				
I	keine	1739.94***	289	0.039	0.037	0.089	0.937	0.926	147247	148004
II	Unt.	1775.24***	299	0.039	0.038	0.092	0.936	0.927	147282	147978
III	Int.	1736.03***	293	0.038	0.037	0.090	0.938	0.927	147242	147976
IV	Unt. & Int.	1772.03***	303	0.038	0.038	0.092	0.936	0.928	147278	147950

Anmerkung. Unt. = Unterricht; Int. = Interesse; df = Freiheitsgrade; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual (L1: Level 1, L2: Level 2); CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Aiken Information Criterion; BIC = Bayes Information Criterion.

*** p<.001.

Chi-Quadrat-Tests ergeben, dass die Faktorladungen von Interesse über beide Ebenen hinweg invariant gesetzt werden können ($\Delta\chi^2(4) = 0.57$, n.s.), ein Modell mit variablen Faktorladungen für die Unterrichtsfaktoren jedoch besser auf die Daten passt als ein invariantes ($\Delta\chi^2(10) = 41.53$, p<.001). Analog dazu zeigt ein Modell mit invarianten Faktorladungen sowohl für Interesse als auch für Unterricht eine schlechtere Passung auf die Stichprobe ($\Delta\chi^2(10) = 44.66$, p<.001). Da der Chi-Quadrat-Test die Sparsamkeit von Modellen nicht berücksichtigt und auch das Modell IV mit invarianten Faktorladungen hinsichtlich RMSEA einen guten (Byrne, 2012; Kline, 2011) und hinsichtlich CFI und

TLI einen akzeptablen Modellfit (Schermelleh-Engel et al., 2003) aufweist und sich die Indikatoren im Vergleich zu den anderen Modellen kaum verändert haben, ist Modell IV mit invarianten Faktorladungen für Interesse und Unterricht dennoch als passend auf die Datenstruktur anzusehen (vgl. Morin et al., 2014). Für das invariante Modell spricht darüber hinaus die inhaltliche Sichtweise von Klimaeffekten von Unterricht, in der davon ausgegangen wird, dass die Unterrichtsmerkmale auf beiden Ebenen das gleiche Konstrukt abbilden (vgl. Marsh et al., 2012). Insgesamt wird daher die Annahme der Invarianz von Interesse und Unterrichtsaktivitäten über die Ebenen als vertretbar gesehen und im weiteren verwendet.

Tabelle 28 zeigt die standardisierten Faktorladungen der Indikatoren bei angenommener Invarianz zwischen beiden Ebenen. Zusätzlich sind in der Tabelle die Varianzen auf Schüler- und Klassenebene sowie die Intraklassenkoeffizienten angegeben. Die Varianzen sind jeweils in unstandardisierter Form angegeben, da in Mplus beide Analyseebenen getrennt standardisiert werden und die standardisierten Varianzen auf beiden Ebenen daher Eins betragen.

Tabelle 28.

Faktorladungen und Intraklassenkorrelation von Unterrichtsmerkmalen und Interesse

	Interesse		Inter		Exper		Forsch		Anwend	
	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)
<i>Schülerebene (Level 1)</i>										
Indikator 1	0.81	(0.01)	0.52	(0.02)	0.58	(0.02)	0.60	(0.02)	0.50	(0.02)
Indikator 2	0.79	(0.01)	0.61	(0.02)	0.47	(0.03)	0.68	(0.02)	0.60	(0.02)
Indikator 3	0.78	(0.01)	0.65	(0.02)	0.59	(0.02)	0.72	(0.02)	0.69	(0.01)
Indikator 4	0.86	(0.01)	0.73	(0.02)					0.65	(0.02)
Indikator 5	0.88	(0.01)								
Indikator 6										
<i>Klassenebene (Level 2)</i>										
Indikator 1	1.00	(0.00)	0.74	(0.04)	0.91	(0.04)	0.81	(0.04)	0.70	(0.05)
Indikator 2	1.00	(0.00)	0.84	(0.03)	0.78	(0.03)	0.91	(0.02)	0.81	(0.03)
Indikator 3	1.00	(0.00)	0.85	(0.03)	0.88	(0.03)	0.95	(0.02)	0.83	(0.03)
Indikator 4	1.00	(0.00)	0.93	(0.02)					0.81	(0.03)
Indikator 5	1.00	(0.00)								
Varianz Level 1	0.63	(0.03)	0.23	(0.02)	0.26	(0.02)	0.30	(0.03)	0.23	(0.02)
Varianz Level 2	0.05	(0.01)	0.07	(0.01)	0.17	(0.03)	0.10	(0.02)	0.04	(0.01)
ICC	0.07		0.23		0.39		0.25		0.14	

Anmerkung. Inter = Interaktives Lehren und Lernen; Exper = Schülerexperimente; Forsch = Schülerinnen und Schüler forschen; ICC = Intraklassenkorrelation. Alle Faktorladungen sind in standardisierten Werten angegeben und signifikant ($p < .05$); Die Varianzen und der ICC sind unstandardisierte Werte.

Alle Indikatoren weisen signifikante Ladungen im Bereich 0.78 – 0.88 bei Interesse bzw. 0.47 – 0.95 auf. Da bei Interesse aufgrund die Residualvarianzen auf Level 2,

die negativ aber nahe Null waren, gemäß der Empfehlung bei Mehrebenenanalysen (z. B. Christ & Schlüter, 2012) gleich Null gesetzt wurden, beträgt ihre Ladung auf Level 2 Eins.

Die Intraklassenkorrelation der Merkmale dient als Maß für den Anteil der Varianz, den die Merkmale auf Klassenebene besitzen und wird aus der Varianz auf Schüler- und Klassenebene berechnet (vgl. Kapitel 6.4.2). Bei Interesse liegen 7% der Gesamtvarianz auf Klassenebene und bei den Unterrichtsmerkmalen sogar 14% (Anwendungen) bis hin zu 39% (Experimente). Bereits ab einem ICC von 0.05 bzw. 0.01 wird die Berücksichtigung der Mehrebenenstruktur als sinnvoll angesehen (Geiser, 2011; Hox, 2010). Sowohl bei Interesse als auch allen Unterrichtsmerkmalen ist eine Mehrebenenanalyse daher angemessen.

7.2.2. Messinvarianz der Skalen zwischen High und Nicht-High-Performern

Unterschiede in Unterrichtszusammenhängen zwischen High und Nicht-High-Performern können mit Multigruppenmodellen untersucht werden. Voraussetzung für Multigruppenvergleiche bei Strukturgleichungsmodellen ist jedoch, dass die Messmodelle in den beiden Subgruppen der Stichprobe vergleichbar sind, d. h. Messinvarianz gilt. Man unterscheidet verschiedene Stufen der Messinvarianz (vgl. T. A. Brown, 2015; Christ & Schlüter, 2012): (a) konfigurale (Faktorstruktur gleich), (b) metrische (Faktorladungen gleich), (c) skalare (Intercepts der Indikatoren gleich) und (d) strikte faktorielle Invarianz (Residuen der Indikatoren gleich). Diese verschiedenen Formen der Messinvarianz stehen in hierarchischer Beziehung zueinander, wobei die konfigurale Invarianz die schwächste und die strikte faktorielle Invarianz die stärkste Form der Invarianz darstellen. In einem *Step-Up-* bzw. *Step-Down-Ansatz* genesteter Modelle wird daher mit χ^2 -Tests üblicherweise geprüft, ob eine stärkere Messinvarianz den Modellfit statistisch bedeutsam verschlechtert. Wie Simulationsstudien zeigen, reicht häufig bereits das Vorliegen partieller Messinvarianz für einen Teil der latenten Variablen dafür aus, Beziehungen und Mittelwerte zwischen Gruppen zu vergleichen (Byrne, Shavelson & Muthén, 1989; Christ & Schlüter, 2012; Steinmetz, 2013).

Ein Vergleich von Beziehungen von latenten Variablen zwischen zwei Subgruppen, wie er in dieser Arbeit beabsichtigt wird, setzt mindestens metrische Invarianz voraus, d. h. eine analoge Faktorenstruktur und vergleichbare Faktorladungen (T. A. Brown,

2015). Ein Vergleich der Mittelwerte von latenten Variablen zwischen beiden Gruppen würde zusätzlich skalare Messinvarianz voraussetzen.

Die Messinvarianz zwischen High und Nicht-High-Performern wurde mit drei Modellen mit verschiedenen Invarianzstufen überprüft (vgl. Tabelle 29). Da die Modellierung der Messinvarianz-Modelle zwischen High und Nicht-High-Performern auf zwei Ebenen und mit unterschiedlichen Faktorladungen der Unterrichtsmerkmale auf den beiden Ebenen insgesamt eine höhere Anzahl an freien Parametern als Klassen auf Ebene 2 vorhanden sind zufolge hätte, kann sie nicht berechnet werden. Darüber hinaus wird die Berücksichtigung der Mehrebenenstruktur bei Invarianzanalysen zwar als sinnvoll angesehen (Y. Kim & Muthén, 2009), es ist aktuell jedoch noch unklar, wie dabei vorgegangen werden kann (für verschiedene Ansätze vgl. E. S. Kim, Yoon, Wen, Luo & Kwok, 2015; Ryu, 2014). Die Messinvarianz zwischen High und Nicht-High-Performern wird hier daher auf die Schülerebene begrenzt. Da in den späteren Mehrebenenanalysen die aggregierten Werte der unabhängigen Variablen auf Klassenebene manifest gebildet werden, d. h. per Definition die Indikatoren bei beiden Gruppen auf Level 2 identisch sind, sollte die Beschränkung auf die Schülerebene hier jedoch ohnehin nur geringfügige Abweichungen mit sich bringen.

Die Modelle wurden jeweils nur für einen Plausible Value (PV1) gerechnet, da die Verwendung von fünf PVs vor allem bei der korrekten Ermittlung von Mittelwerten bzw. Zusammenhängen und den zugehörigen Standardfehlern notwendig ist. Hier wurde jedoch nur das Messmodell und seine Passung innerhalb der Rechnung mit einem PV betrachtet und es kann davon ausgegangen werden, dass die Modellpassung für alle PVs aufgrund ihrer hohen Korrelation vergleichbar ist. Als latente Variablen wurden alle Unterrichtsaktivitäten sowie Interesse in das Messmodell aufgenommen (vgl. Vorgehen in vorherigem Abschnitt). Da für jeden Plausible Value ein High-Performer-Datensatz verwendet wird und darin die naturwissenschaftliche Kompetenz nur mit einem Plausible Value gemessen wird, wird sie manifest gebildet und ist per Definition invariant.

Tabelle 29 gibt eine Übersicht über die Modellindikatoren der unterschiedlich invarianten Modelle. Der Chi-Quadrat-Test für robuste Schätzer (Satorra, 2000; Satorra & Bentler, 2010) ergibt, dass das metrische Modell im Vergleich zum konfigurationalen Modell einen schlechteren Modellfit zeigt ($\Delta\chi(19) = 53.62, p < .001$). Auch das skalare Modell muss aufgrund fehlender metrischer Invarianz und auch auf Basis eines Chi-Quadrat-Tests im Vergleich zur metrischen Invarianz ($\Delta\chi(19) = 105.23, p < .001$) verworfen

Tabelle 29.

Messinvarianz der Unterrichtsmerkmale und Interesse zwischen High und Nicht-High-Performern

Modell	χ^2	df	RMSEA	SRMR	CFI	TLI	AIC	BIC
Konfigural	2181.29***	520	0.044	0.035	0.950	0.942	194164	195262
Metrisch	2236.57***	539	0.044	0.036	0.948	0.943	194180	195162
Skalar	2340.33***	558	0.044	0.037	0.946	0.942	194249	195115

Anmerkung. df = Freiheitsgrade; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; AIC = Aiken Information Criterion; BIC = Bayes Information Criterion.

*** $p < .001$.

werden. Der Chi-Quadrat-Test berücksichtigt jedoch nicht die Sparsamkeit der Modelle. Darüber hinaus zeigen die Modellkennwerte auch für das metrische und skalare Modell eine gute ($RMSEA < .05$, $CFI \geq .95$) bzw. akzeptable ($TLI \geq .90$) Modellpassung, die sich insbesondere kaum zwischen den Modellen unterscheidet. Auch die Informationskriterien sprechen nicht klar für das konfigurale Modell, sondern im Falle des BIC für das skalare.

Modifikationsindices deuteten darüber hinaus an, dass partielle Invarianz für die Unterrichtsmerkmale gilt. In einer zweiten Invarianztestung wurden daher Modelle nur für Unterricht getestet (vgl. Tabelle 30). Sie ergeben, dass metrische Invarianz ($\Delta\chi(10) = 7.71$, ns) für die Unterrichtsskalen gilt, sich das skalar invariante Modell in der Passung jedoch signifikant vom metrischen Modell unterscheidet ($\Delta\chi(10) = 62.28$, $p < .001$). Auch für die Unterrichtsskalen allein ergeben sich jedoch gute und unveränderte Modellkennwerte des skalar invarianten Modells. Insgesamt ist daher sowohl die Annahme metrischer als auch skalarer Invarianz vertretbar.

Tabelle 30.

Partielle Messinvarianz für Unterrichtsskalen

Modell	χ^2	df	RMSEA	SRMR	CFI	TLI
Konfigural	2181.29***	520	0.04	0.04	0.95	0.94
Metrisch	2187.51***	530	0.04	0.04	0.95	0.94
Skalar	2248.24***	540	0.04	0.04	0.95	0.94

Anmerkung. df = Freiheitsgrade; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index.

*** $p < .001$.

Ziel dieses Kapitels war es, die Invarianz der Faktorstruktur und der Faktor-

ladungen aller verwendeter Instrumente über Schüler- und Klassenebene hinweg zu untersuchen sowie mindestens metrische Messinvarianz zwischen High und Nicht-High-Performern nachzuweisen. Dabei wurden zunächst die Unterrichtsskalen zu forschend-entdeckendem Lernen näher in den Blick genommen und ihre Trennbarkeit nachgewiesen. Darüber hinaus wurde gezeigt, dass die Skalen zu den Unterrichtsaktivitäten auf Klassenebene ebenso viele Faktoren abbilden wie auf Individualebene und die Faktorladungen aller Unterrichtsaktivitäten auf beiden Analyseebenen gleichgesetzt werden können. Obwohl zwischen High und Nicht-High-Performern ein Modell mit metrischer bzw. skalarer Messinvarianz im Vergleich zu konfiguraler Invarianz keinen besseren Modellfit aufweist, ist es auf Basis der Modellkennwerte sowie der partiellen Invarianz der Unterrichtsskalen vertretbar, skalare Invarianz zwischen High und Nicht-High-Performern anzunehmen. In den nachfolgenden Kapiteln werden daher Modelle mit invarianten Faktorladungen zwischen Schüler- und Klassenebene sowie zwischen High und Nicht-High-Performern verwendet.

7.2.3. Unterrichtsaktivitäten als Prädiktoren naturwissenschaftlicher Kompetenz von High Performern

Im Rahmen der Fragestellung zum Zusammenhang von Unterricht und Bildungsergebnissen beschäftigt sich diese Arbeit zunächst mit der naturwissenschaftlichen Kompetenz von High Performern und der Frage, inwieweit die Zusammensetzung der Klasse und die wahrgenommene Häufigkeit der Unterrichtsaktivitäten in einem Zusammenhang mit der naturwissenschaftlichen Kompetenz von High Performern stehen und ob Unterschiede zu Nicht-High-Performern existieren. Zur Beantwortung dieser Fragestellungen wurde im Rahmen der Arbeit ermittelt, welche Varianz die naturwissenschaftliche Kompetenz von High und Nicht-High-Performern auf Klassen- und Schülerebene aufweist und ob Mehrebenenanalysen ausgehend davon sinnvoll sind. Darauf aufbauend wurden mögliche Kompositionseffekte der Klassenzusammensetzung für die naturwissenschaftliche Kompetenz von High Performern analysiert. Anschließend wurden die Zusammenhänge von naturwissenschaftlicher Kompetenz und Unterrichtsmerkmalen in Mehrebenen-Strukturgleichungsmodellen untersucht. Schließlich wurden die Zusammensetzung der Klasse und Unterrichtsaktivitäten in einem gemeinsamen Modell auf ihre prädiktive Wirkung auf naturwissenschaftliche Kompetenz hin untersucht. Unterschiede in den Zusammenhängen zwischen High und Nicht-High-Performern wurden mit Modelltests ermittelt.

Aufbauend auf den Ergebnissen des vorigen Abschnitts wurde dabei angenommen, dass die Faktorladungen der Unterrichtsitems auf Schüler- und Klassenebene übereinstimmen und zwischen High und Nicht-High-Performern skalare Invarianz gilt. Zusätzlich wurden für diese Analysen Klassen- und Schüलगewichte in die Rechnung miteinbezogen, um so möglichst repräsentative Mittelwerte und Koeffizienten für die Population zu erreichen. Nachfolgend werden die Ergebnisse der einzelnen Analysen jeweils in einem eigenen Unterkapitel beschrieben.

Varianz auf Klassenebene

Mehrebenenanalysen sind aus methodischer Sicht vor allem dann sinnvoll, wenn ein bedeutender Anteil der Varianz eines Merkmals auf Klassenebene zu finden ist (Hox, 2010). Der Intraclass-Korrelations-Koeffizient dient als Maßzahl für die Varianzaufteilung auf Klassen- und Schülerebene (vgl. Kapitel 6.4.2).

Die Intraclass-Korrelations-Koeffizienten der naturwissenschaftlichen Kompetenz betragen $ICC_{1,HP} = .05$ für High Performer und $ICC_{1,NHP} = .16$ für Nicht-High-Performer. Grundsätzlich liegt eine Mehrebenenstruktur bzw. Varianz auf der Klassenebene vor, sobald der ICC_1 signifikant von Null abweicht (Hox, 2010), als Daumenregel wird häufig ab Werten von .05 bzw. zum Teil .01 davon ausgegangen (Geiser, 2011), dass die Modellierung einer Mehrebenenstruktur sinnvoll ist, was in diesem Falle für beide Leistungsgruppen zutrifft. Als weiteres Maß für die Reabilität der Daten auf Klassenebene wird der ICC_2 – auch *Design Effect* genannt – berechnet (vgl. Kapitel 6.4.2). Da die Klassenwerte manifest aus den Schülerangaben aggregiert wurden und bei beiden Leistungsgruppen auf den Angaben der gesamten Klasse beruhen, wird bei der Berechnung des ICC_2 als mittlere Klassengröße die mittlere Größe der gesamten Klasse ($n = 25.42$) und nicht der Teilgruppe der High bzw. Nicht-High-Performer verwendet. Die Werte des ICC_2 liegen sowohl für High Performer ($ICC_{2,HP} = 0.94$) als auch Nicht-High-Performer ($ICC_{2,NHP} = 0.97$) nicht über dem Wert zwei, der als Schwelle dafür angesehen wird, dass die Mehrebenenstruktur berücksichtigt werden muss. Dennoch wird aufgrund der Werte des ICC_1 angenommen, dass die Berücksichtigung der Mehrebenenstruktur bei der Analyse des Zusammenhangs von Unterricht und naturwissenschaftlicher Kompetenz angemessen ist.

Kompositionseffekte

Bei der Leistungsentwicklung in Schulklassen sind verschiedene Kompositionseffekte denkbar. Die Zusammensetzung einer Klasse bezüglich des sozioökonomischen Hintergrunds und der kognitiven Fähigkeit beeinflusst möglicherweise die naturwissenschaftliche Kompetenz, die High und Nicht-High-Performer erreichen. Wie in Kapitel 2.2 gezeigt wurde, enthalten Klassen unterschiedlich hohe Anteile an High Performern und unterschiedliche Kompetenzstreuungen. Die Ergebnisse zeigen, dass Klassen mit höheren High-Performer-Anteilen nicht zwingend homogener sind als Klassen mit weniger hohen Anteilen. Auch wenn damit kein Zusammenhang zwischen der Anzahl der High Performer und der Streuung einer Klasse beobachtet werden kann, hat die Streuung der Klasse möglicherweise einen Einfluss darauf, welches Kompetenzniveau High Performer und Nicht-High-Performer in der Klasse erreichen. Auch Kompositionseffekte der Kompetenzstreuung der Klasse werden neben Effekten der Zusammensetzung der Klasse hinsichtlich sozioökonomischem Hintergrund und kognitiver Grundfähigkeit des-

halb untersucht. Das Geschlecht wird dabei als Kontrollvariable auf der Schülerebene miteinbezogen.

Methodisch umgesetzt wird die Analyse der Kompositionseffekte, indem die Merkmale sowohl auf Schüler- als auch Klassenebene als Prädiktoren einbezogen werden. Auf Schülerebene werden sie innerhalb der Klasse zentriert (*groupmean*). Die Kompositionseffekte berechnen sich dann aus der Differenz von Level-2- und Level-1-Koeffizienten (z. B. Marsh et al., 2012) und wurden im Modell mit geschätzt. Tabelle 31 gibt eine Übersicht über die Ergebnisse. Sowohl bei High als auch bei Nicht-High-Performern ist eine höhere kognitive Grundfähigkeit im Mittel mit einer höheren naturwissenschaftlichen Kompetenz assoziiert. Unter den High Performern erreichen Mädchen eine geringere naturwissenschaftliche Kompetenz während unter den Nicht-High-Performern keine Unterschiede bestehen. Umgekehrt wirkt der sozioökonomische Hintergrund nur unter den Nicht-High-Performern als signifikanter Prädiktor naturwissenschaftlicher Kompetenz. Unter High Performern geht ein höherer sozioökonomischer Status dagegen nicht mit einer höheren naturwissenschaftlichen Kompetenz einher.

Weder die kognitive Grundfähigkeit, noch der sozioökonomische Status zeigen jedoch signifikante Kompositionseffekte auf die Leistung von High und Nicht-High-Performern. Dies bedeutet, dass die Zusammensetzung der Klasse hinsichtlich kognitiver Leistung oder sozioökonomischem Status keinen Einfluss auf die Kompetenz beider Leistungsgruppen hat. High Performer zeigen also in im Mittel intelligenteren Klassen oder Klassen mit durchschnittlich günstigeren Elternhäusern keine besseren Kompetenzniveaus, d. h. erreichen auch nicht häufiger Top Performance als in anderen Klassen. Umgekehrt scheint es auch für die Kompetenz von Nicht-High-Performern, keine Rolle zu spielen, ob sie sich in einer Klasse mit im Mittel höherer oder niedriger kognitiver Leistung oder höherem oder niedrigerem sozioökonomischen Status befinden.

Die Anzahl der High Performer einer Klasse und die Leistungsstreuung einer Klasse sind zwei Merkmale auf Klassenebene, die ebenfalls die Kompetenz beider Leistungsgruppen beeinflussen können. Da sie direkt auf Klassenebene gemessen werden und per Definition keine Varianz auf Individualebene haben, wird ihr Einfluss als Klimageffekt auf die Leistung auf Klassenebene modelliert, d. h. als Effekt auf die mittlere Kompetenz der High Performer einer Klasse sowie die mittlere Kompetenz der Nicht-High-Performer einer Klasse. Die Analysen ergaben, dass die Anzahl der High Performer in einer Klasse in einem positiven Zusammenhang zur naturwissenschaftlichen

Kompetenz von sowohl High als auch Nicht-High-Performern hat. Dies bedeutet, dass High Performer in Klassen mit mehr High Performern insgesamt bessere Leistungen erzielen als High Performer in Klassen mit weniger High Performern. Dasselbe gilt für die restlichen Schülerinnen und Schüler: Sie erzielen in Klassen mit vielen High Performern bessere Ergebnisse als in Klassen mit weniger High Performern. Dies impliziert, dass eine höhere Anzahl leistungsstarker Schülerinnen und Schüler in Naturwissenschaften sowohl für leistungsstarke als auch die restlichen Schülerinnen und Schüler einer Klasse eine positivere Lernumgebung mit sich bringt.

Tabelle 31.

Kompositionseffekte im Zusammenhang mit naturwissenschaftlicher Kompetenz von High und Nicht-High-Performern

	HP		NHP		Differenz HP-NHP	
	β	(SE)	β	(SE)	<i>b</i>	(SE)
<i>Schülerebene (Level 1)</i>						
Geschlecht	-0.08*	(0.03)	-0.04	(0.03)	-0.04	(0.05)
KFT	0.17***	(0.03)	0.24***	(0.03)	-0.06	(0.03)
HISEI	0.04	(0.04)	0.10***	(0.03)	-0.04	(0.02)
<i>Klassenebene (Level 2)</i>						
KFT	0.06	(0.04)	0.06	(0.04)	-0.01	(0.10)
HISEI	0.01	(0.03)	0.03	(0.04)	-0.03	(0.09)
Anzahl HP	0.12**	(0.04)	0.17***	(0.05)	-0.01	(0.01)
SD Kompetenz	0.18***	(0.03)	-0.19***	(0.03)	1.45***	(0.15)
<i>Kompositionseffekte</i>						
KFT	0.01	(0.08)	-0.04	(0.09)	0.04	(0.11)
HISEI	0.01	(0.06)	-0.01	(0.09)	0.02	(0.09)
R^2 Level 1	0.04***	(0.01)	0.07***	(0.02)		
R^2 Level 2	0.94***	(0.19)	0.86***	(0.19)		

Anmerkung. Standardisierte Koeffizienten für HP und NHP, Differenzwert nicht standardisiert. HP = High Performer; NHP = Nicht-High-Performer; KFT = Kognitiver Grundfähigkeitstest; HISEI = Highest Socio Economic Index; Anzahl HP = Anzahl der High Performer einer Klasse; SD Kompetenz = Standardabweichung der naturw. Kompetenz einer Klasse.

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

Ein wichtiger weiterer Aspekt für Aussagen zur Gruppierung ist der Einfluss der Leistungsstreuung einer Klasse. Dem Ergebnis zufolge erreichen die High Performer aus Klassen mit höherer Kompetenzstreuung, d. h. zunehmend heterogene Klassen, im Durchschnitt eine höhere naturwissenschaftliche Kompetenz als High Performer aus Klassen mit geringerer Streuung, d. h. eher homogenen Klassen. Hervorzuheben ist, dass dieser Zusammenhang auf Klassenebene stattfindet und die Heterogenität der Klasse demnach in einem positiven Zusammenhang zum mittleren Leistungsniveau der gesam-

ten High-Performer-Gruppe einer Klasse steht. Für die Gruppe der Nicht-High-Performer einer Klasse zeigt sich ein umgekehrter Zusammenhang: Je höher die Streuung einer Klasse, desto niedrigere Werte in der naturwissenschaftlichen Kompetenz erreicht die Gruppe der Nicht-High-Performer in dieser Klasse. Im Unterschied zu den High Performern, bei denen Leistungsheterogenität der Klasse mit höherer naturwissenschaftlicher Kompetenz einhergeht, ist Leistungsheterogenität bei Nicht-High-Performern mit niedrigeren Leistungen assoziiert. Beide Leistungsgruppen unterscheiden sich in diesem Zusammenhang signifikant.

Insgesamt wird durch die Ergebnisse deutlich, dass weder die kognitive Grundfähigkeit noch der soziökonomische Hintergrund zu Kompositionseffekten führte. Die bestätigt die angenommene Hypothese (vgl. Kapitel 5.2). Die Ergebnisse zeigen darüber hinaus, dass in Klassen mit mehr leistungsstarken Jugendlichen über Leistungsgruppen hinweg höhere Kompetenzwerte erreicht werden. Dies bedeutet, dass Klassen mit mehr High Performern aufgrund ihrer Zusammensetzung möglicherweise Lernprozesse unterstützen. Auch hier konnte daher der angenommene Zusammenhang, dass ein höherer High-Performer-Anteil einen positiven Zusammenhang zur naturwissenschaftlichen Kompetenz aufweist, belegt werden. Offen bleibt bei diesen Ergebnissen, ob eine höhere Anzahl von High Performern zu besseren Lernbedingungen bzw. höherer Leistung führt oder umgekehrt sich diese Klassen durch bessere Lernbedingungen auszeichnen und dadurch generell bei allen Schülerinnen und Schülern höhere Leistungen erzielen. Letzteres würde ebenfalls in einer höheren Anzahl leistungsstarker Jugendlicher und gleichzeitig höheren mittleren Leistungen innerhalb verschiedener Leistungsgruppen resultieren.

Im Hinblick auf die Leistungsstreuung wurde angenommen, dass kein Zusammenhang oder alternativ ein positiver Zusammenhang existiert. Auch hier wurde der Annahme entsprechend gefunden, dass eine höhere Leistungsstreuung in einer Klasse mit höheren Leistungen der High Performer dieser Klasse einhergeht. Nicht-High-Performer am Gymnasium scheinen zwar möglicherweise von einem höheren Anteil leistungsstarker Jugendlicher aber nicht von einer höheren Heterogenität der Klasse zu profitieren. Da beide Prädiktoren (Anteil High Performer und Streuung) gleichzeitig berücksichtigt wurden, bedeutet das Ergebnis, dass Nicht-High-Performer unter Kontrolle der High-Performer-Anzahl einer Klasse in heterogenen Klassen geringere Leistungen zeigen, High Performer umgekehrt höhere. Eine naheliegende Erklärung für das Ergebnis

ist daher auch, dass (unter Kontrolle der Anzahl der High Performer) das Vorhandensein einer höheren Heterogenität in einer Klasse bedeutet, dass Nicht-High-Performer geringere Kompetenzwerte und High Performer höhere Kompetenzwerte erreichen als in weniger heterogenen Klassen. Eine mögliche Erklärung könnte daher sein, dass eine höhere Heterogenität einer Schulklasse in der vorliegenden Analyse mit einer höheren Anzahl bzw. schlechteren Leistung von Nicht-High-Performern einhergeht und daher mit einer geringeren Leistung von Nicht-High-Performern assoziiert ist.

Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz von High Performern

Zur Beantwortung der Forschungsfrage, inwieweit Unterrichtsaktivitäten im Zusammenhang mit der naturwissenschaftlichen Kompetenz von High Performern stehen und, ob sich Unterschiede zu Nicht-High-Performern ergeben, wurden ebenfalls Mehrebenenanalysen gerechnet. Wie in Kapitel 6.4.2 dargestellt, wurden die Unterrichtsitems auf Klassenebene manifest aggregiert, so dass die Klassenwerte jeweils die Angaben der vollständigen Klasse, d. h. beider Leistungsgruppen innerhalb der Klassen, enthalten. Die Bildung von Faktoren auf Klassenebene erfolgte latent, um so Messfehler zu berücksichtigen.

Da sich die Formulierungen der Unterrichtsitems auf die Unterrichtssituation der gesamten Klasse beziehen und die Häufigkeiten bestimmter Unterrichtsaktivitäten mit Bezug zum Klassenunterricht abgefragt werden, werden die Zusammenhänge als Klimateffekte betrachtet (vgl. Marsh et al., 2012). Dies bedeutet, dass der Zusammenhang von Unterricht und abhängiger Variable auf Klassenebene im Vordergrund steht und der Zusammenhang auf individueller Ebene einen Zusammenhang zwischen individueller Abweichung der Unterrichtseinschätzung innerhalb einer Klasse und individueller Abweichung der Leistung innerhalb der Klasse darstellt. Üblicherweise sind Unterrichtseinschätzungen auf individueller Ebene daher eine Art Messfehler der Klassenvariable Unterricht und ein Zusammenhang mit der abhängigen Variable nicht gewünscht (vgl. Marsh et al., 2012; Lüdtke et al., 2009). Da bei den vorliegenden Unterrichtsaktivitäten jedoch auch inhaltlich bedeutsame Abweichungen auf Schülerebene auftreten können, z. B. wenn in differenzierenden Lernumgebungen ein Teil der Schülerinnen und Schüler häufiger Experimente durchführen dürfen als ihre Mitschüler oder Schülerinnen und Schüler innerhalb einer Schulklasse unterschiedliche Schulzweige besuchen, wird

7. Ergebnisse

nachfolgend auch der Zusammenhang auf individueller Ebene betrachtet.

Zum Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz von High und Nicht-High-Performern wurde ein Mehrebenenmodell gebildet, in dem neben den Unterrichtsmerkmalen Geschlecht, sozioökonomischer Hintergrund und kognitive Grundfähigkeit als Kontrollvariablen auf Schülerebene einbezogen wurden. Abbildung 10 zeigt das vollständige Modell.

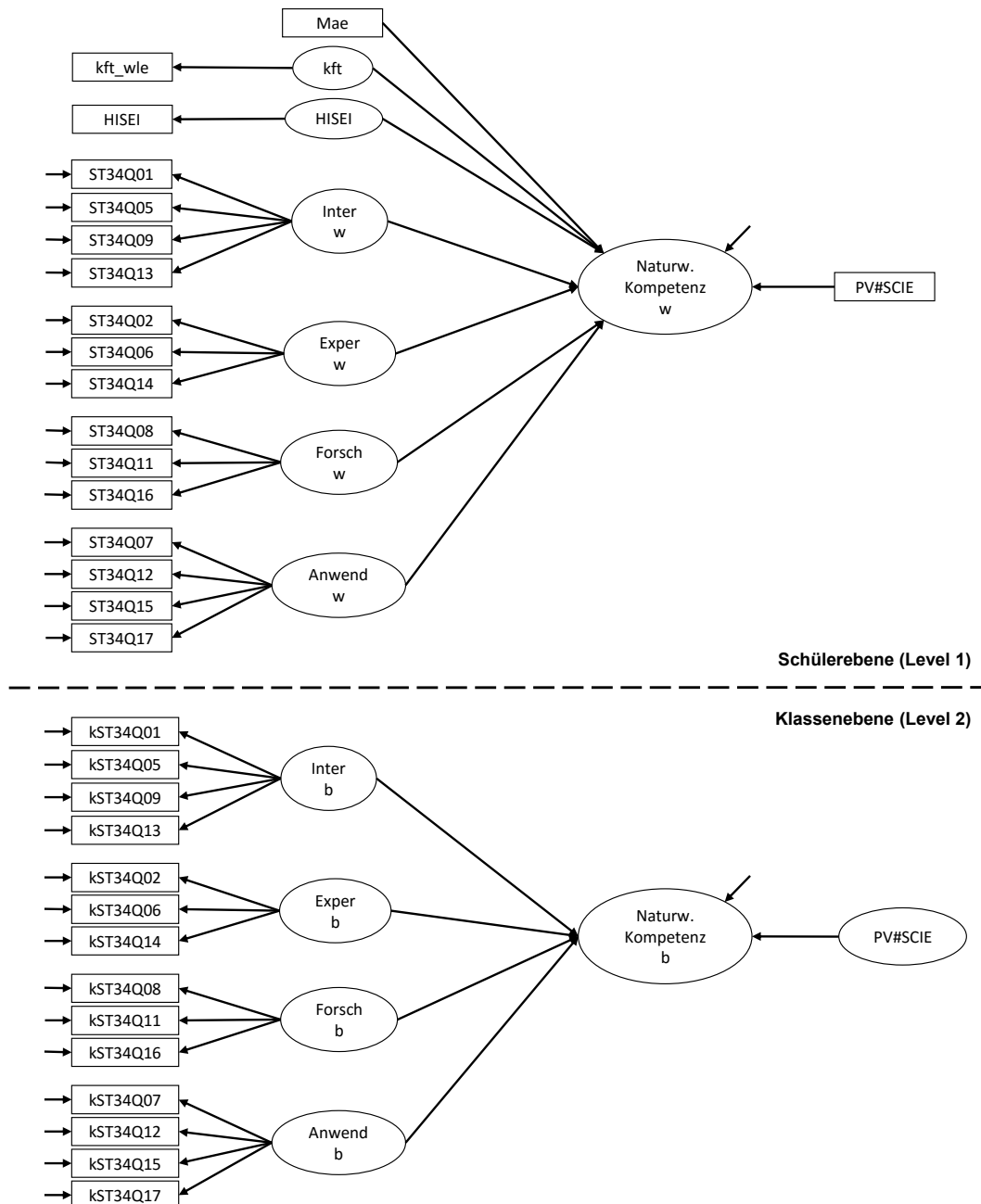


Abbildung 10. Latent-manifestes Modell für naturwissenschaftliche Kompetenz (Modell wurde in einem Multigruppenvergleich für High und Nicht-High-Performer gerechnet)

Das Messmodell für das Mehrebenen-Strukturgleichungsmodell zeigte je nach

betrachtetem Fitindikator einen nur eingeschränkt akzeptablen ($CFI = .88$, $TLI = .87$) bis sehr guten ($RMSEA = .04$, $SRMR_{Level1} = .05$, $SRMR_{Level2} = .08$) Fit. Die einzelnen Faktorladungen der Indikatoren sowie die Kovarianzen, Varianzen und Korrelationen der Faktoren sind in Tabellen A.1 und A.2 im Anhang zu finden.

Aufbauend auf dem Messmodell wurden vier verschiedene Strukturgleichungsmodelle gerechnet, in denen die einzelnen Unterrichtsaktivitäten sukzessive als Prädiktoren aufgenommen wurden. In allen Modellen wurde das gleiche Messmodell verwendet, d. h. bereits alle Faktoren modelliert, sowie zur besseren Vergleichbarkeit der Koeffizienten standardisierte Items verwendet. In jedem Modell wurden zusätzlich die Unterschiede der Koeffizienten zwischen High und Nicht-High-Performern geschätzt. Signifikant von Null verschiedene Differenzen belegen einen signifikanten Unterschied zwischen beiden Leistungsgruppen.

In Modell I wurden lediglich die Kontrollvariablen kognitive Grundfähigkeit, sozioökonomischer Status und Geschlecht als Prädiktoren modelliert (vgl. Tabellen 32 und 33). Da auch forschend-entdeckende Unterrichtsaktivitäten hohe Anteile an interaktivem Lehren und Lernen enthalten können, wurde in Modell II neben den Kontrollvariablen interaktives Lehren und Lernen als einzige Unterrichtsaktivität aufgenommen. In Modell III wurden daraufhin „Experimentieren“ und „Schülerinnen und Schüler forschen“ als Operationalisierung forschend-entdeckenden Lernens als Prädiktoren berücksichtigt. Aufgrund des hierarchischen Aufbaus untersucht Modell III, welche zusätzlichen Zusammenhänge durch forschend-entdeckendes Lernen unter Kontrolle interaktiven Lehren und Lernens zu beobachten sind und betont epistemologische und experimentelle Aspekte von forschend-entdeckendem Lernen. In Modell IV werden schließlich auch Anwendungsbezüge als Prädiktoren naturwissenschaftlicher Kompetenz aufgenommen. Modell IV zeigt wiederum, welche Zusammenhänge Anwendungen zusätzlich zu interaktiven und forschend-entdeckenden Unterrichtsaktivitäten zeigen. Die Ergebnisse beziehen sich jeweils auf Häufigkeiten im Vorkommen der Unterrichtsaktivitäten aus Schülersicht.

Tabelle 32 gibt eine Übersicht über die vier betrachteten Modelle für High Performer, Tabelle 33 zeigt die Ergebnisse für Nicht-High-Performer und in Tabelle 34 sind schließlich die Unterschiede beider Leistungsgruppen dargestellt. Modell I für die Kontrollvariablen zeigt, dass weibliche High Performer eine signifikant niedrigere naturwissenschaftliche Kompetenz erreichten als männliche High-Performer. Ebenso wurde

die kognitive Grundfähigkeit (Subskala „Figurale Analogien“) als signifikanter Prädiktor für High Performer identifiziert. Je höher die kognitive Grundfähigkeit, desto höher war die naturwissenschaftliche Kompetenz von High Performern auf Schülerebene. Im Gegensatz dazu zeigte der sozioökonomische Hintergrund für High Performer keinen Zusammenhang zur naturwissenschaftlichen Kompetenz. Dies entspricht den Ergebnissen zu den Kompositionseffekten.

Tabelle 32.

Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz von High Performern

	Modell I		Modell II		Modell III		Modell IV	
	β	(SE)	β	(SE)	β	(SE)	β	(SE)
<i>Schülerebene (Level 1)</i>								
Mädchen	-0.09**	(0.03)	-0.09**	(0.03)	-0.10**	(0.03)	-0.09**	(0.03)
KFT	0.18***	(0.03)	0.18***	(0.03)	0.18***	(0.03)	0.18***	(0.03)
HISEI	0.04	(0.04)	0.04	(0.04)	0.04	(0.04)	0.03	(0.04)
Intakt			-0.04	(0.04)	0.04	(0.05)	0.01	(0.06)
Exper					-0.03	(0.05)	-0.04	(0.06)
Forsch					-0.08	(0.07)	-0.12	(0.06)
Anwend							0.09	(0.06)
<i>Klassenebene (Level 2)</i>								
Intakt			-0.01	(0.04)	-0.04	(0.05)	-0.06	(0.05)
Exper					0.00	(0.05)	-0.02	(0.05)
Forsch					0.04	(0.06)	0.04	(0.06)
Anwend							0.05	(0.04)
R ² Level 1	0.04***	(0.01)	0.05***	(0.01)	0.05***	(0.01)	0.06***	(0.01)
R ² Level 2			0.02	(0.07)	0.07	(0.11)	0.12	(0.13)

Anmerkung. Alle Koeffizienten sind standardisierte Werte. KFT = Kognitiver Grundfähigkeitstest, HISEI = Highest Socio Economic Index, Intakt = Interaktives Lehren und Lernen, Exper = Experimentieren, Forsch = Schülerinnen und Schüler forschen, Anwend = Anwendungen.

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

Wie Modell I zeigt, war interaktives Lehren und Lernen auf Klassenebene kein signifikanter Prädiktor der naturwissenschaftlichen Kompetenz von High Performern. Auch unter Berücksichtigung der anderen Unterrichtsmerkmale in Modell III und Modell IV ergab sich kein Zusammenhang. Dies bedeutet, High Performer erreichten unabhängig von der Häufigkeit interaktiver Lerngelegenheiten in ihrer Klasse in der hier gemessenen Form ihre hohe naturwissenschaftliche Kompetenz. Auch innerhalb der Klassen auf Schülerebene sagte die Häufigkeit der Anwendungen die naturwissenschaftliche Kompetenz von High Performern nicht vorher.

Ähnlich zu interaktivem Lehren und Lernen macht es für High Performer einer Klasse keinen Unterschied, ob in der Klasse häufiger Gelegenheit für Schülerexperimente

oder forschendes Lernen geboten werden (Modell III). Modell III verdeutlicht zudem, dass auch innerhalb der Klasse häufigere forschend-entdeckende Tätigkeiten nicht mit höherer naturwissenschaftlicher Kompetenz im Vergleich zu anderen High Performern der Klasse einhergeht.

Modell IV zeigt, dass auch die Häufigkeit von Anwendungsbezügen keinen Zusammenhang mit der naturwissenschaftlichen Kompetenz von High Performern hatten. Auch hier konnten auf Schülerebene keine zusätzlichen Zusammenhänge beobachtet werden.

Für Nicht-High-Performer existierte in Modell I im Unterschied zu den High Performern kein Geschlechterunterschied, in Modell III und IV wurde jedoch auch dort ein Geschlechterunterschied zugunsten der Jungen signifikant. Ebenso wie bei den High Performern ergab sich die kognitive Grundfähigkeit auch bei den Nicht-High-Performern als positiver Prädiktor. Für Nicht-High-Performer zeigte der sozioökonomische Hintergrund im Unterschied zu den High Performern einen signifikant positiven und über die Modelle hinweg konsistenten Zusammenhang mit der naturwissenschaftlichen Kompetenz. Das bedeutet, Nicht-High-Performer mit höherem sozioökonomischen Status erreichen ein höheres Kompetenzniveau in Naturwissenschaften. Auch hier deckt sich das Ergebnis mit den Befunden zu den Kompositionseffekten.

Ähnlich zu den High Performern erwies sich interaktives Lehren und Lernen in Modell II auch bei den Nicht-High-Performern nicht als signifikanter Prädiktor naturwissenschaftlicher Kompetenz. Auch in Kombination mit den anderen Unterrichtsmerkmalen in den nachfolgenden Modellen konnte die Häufigkeit interaktiven Lehren und Lernens nicht die Leistung von Nicht-High-Performern auf Klassen- und Schülerebene vorhersagen. Dieses Ergebnis zeigte sich wie bei den High Performern auch auf Schülerebene, das heißt für Zusammenhänge innerhalb der Klassen.

Analog zu High Performern erreichten auch Nicht-High-Performer in Klassen, in denen häufiger forschend-entdeckender Unterricht berichtet wurde, keine höhere naturwissenschaftliche Kompetenz. Ein Zusammenhang zeigte sich jedoch auf Schülerebene: Nicht-High-Performer, die in ihrer Klasse häufiger Gelegenheiten für Schülerexperimente berichteten als andere Nicht-High-Performer, erreichten innerhalb ihrer Klasse höhere Leistungen. Umgekehrt zeichneten sich Nicht-High-Performer, die häufigere forschende Aktivitäten angaben, durch eine niedrigere naturwissenschaftliche Kompetenz aus.

Diese Zusammenhänge blieben auch in Modell IV für Nicht-High-Performer be-

Tabelle 33.

Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz von Nicht-High-Performern

	Modell I		Modell II		Modell III		Modell IV	
	β	(SE)	β	(SE)	β	(SE)	β	(SE)
<i>Schülerebene (Level 1)</i>								
Mädchen	-0.05	(0.03)	-0.05	(0.03)	-0.07*	(0.03)	-0.07*	(0.03)
KFT	0.26***	(0.03)	0.26***	(0.03)	0.24***	(0.04)	0.24***	(0.04)
HISEI	0.11***	(0.03)	0.11***	(0.03)	0.11***	(0.03)	0.11***	(0.03)
Intakt			-0.04	(0.04)	0.02	(0.06)	-0.01	(0.06)
Exper					0.20*	(0.10)	0.19*	(0.10)
Forsch					-0.26**	(0.09)	-0.32**	(0.10)
Anwend							0.10	(0.07)
<i>Klassenebene (Level 2)</i>								
Intakt			-0.08	(0.04)	0.01	(0.06)	-0.01	(0.07)
Exper					0.06	(0.05)	0.04	(0.05)
Forsch					-0.16	(0.10)	-0.17	(0.10)
Anwend							0.06	(0.05)
R ² Level 1	0.08***	(0.02)	0.09***	(0.02)	0.12***	(0.02)	0.13***	(0.03)
R ² Level 2			0.10	(0.08)	0.26	(0.18)	0.30	(0.15)

Anmerkung. Alle Koeffizienten sind standardisierte Werte. KFT = Kognitiver Grundfähigkeitstest, HISEI = Highest Socio Economic Index, Intakt = Interaktives Lehren und Lernen, Exper = Experimentieren, Forsch = Schülerinnen und Schüler forschen, Anwend = Anwendungen.

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

stehen. Anwendungsbezüge, die in Modell IV als unabhängige Variablen aufgenommen wurden, sagten jedoch sowohl auf Klassen als auch auf Schülerebene nicht die naturwissenschaftliche Kompetenz von Nicht-High-Performern vorher.

Betrachtet man die Unterschiede zwischen High und Nicht-High-Performern, was in dieser Arbeit durch den Multigruppenansatz in den Mehrebenenanalysen geschah, ergeben sich einige Unterschiede. Tabelle 34 zeigt die Koeffizienten von High und Nicht-High-Performern für Modell IV sowie deren Differenzen. Auf Schülerebene unterschieden sich die Koeffizienten von Geschlecht nicht signifikant zwischen High und Nicht-High-Performern und in beiden Gruppen gab es einen vergleichbaren Zusammenhang zugunsten der Jungen. Im Vergleich dazu ergaben sich in der Analyse der Gesamtstichprobe der 15-Jährigen in PISA 2006 keine signifikanten Geschlechterunterschiede in der naturwissenschaftlichen Kompetenz (Prenzel et al., 2007). Sowohl bei High als auch Nicht-High-Performern erwies sich die kognitive Grundfähigkeit als positiver Prädiktor von naturwissenschaftlicher Kompetenz. Zwischen High und Nicht-High-Performern waren dabei keine Unterschiede im Zusammenhang von kognitiver Grundfähigkeit und naturwissenschaftlicher Kompetenz zu beobachten. Bei beiden Leistungsgruppen trat der

Zusammenhang in vergleichbarer Höhe über alle Modelle hinweg auf. Anders als beim Geschlecht und der kognitiven Grundfähigkeit unterschieden sich beide Leistungsgruppen im Zusammenhang von sozioökonomischen Hintergrund und naturwissenschaftlicher Kompetenz statistisch bedeutsam. Während Nicht-High-Performer mit höherem sozioökonomischen Status ein höheres Kompetenzniveau in Naturwissenschaften erreichten, spielte der sozioökonomische Status bei High Performern für die naturwissenschaftliche Kompetenz keine Rolle.

Tabelle 34.
Differenz der Koeffizienten aus Modell IV zwischen High und Nicht-High-Performern

	HP		NHP		Diff. HP-NHP	
	β	(SE)	β	(SE)	b	(SE)
<i>Schülerebene (Level 1)</i>						
Mädchen	-0.09**	(0.03)	-0.07*	(0.03)	-0.02	(0.05)
KFT	0.18***	(0.03)	0.24***	(0.04)	-0.05	(0.03)
HISEI	0.03	(0.04)	0.11***	(0.03)	-0.05	(0.02)
Intakt	0.01	(0.06)	-0.01	(0.06)	0.02	(0.09)
Exper	-0.04	(0.06)	0.19*	(0.10)	-0.27*	(0.13)
Forsch	-0.12	(0.06)	-0.32**	(0.10)	0.20	(0.13)
Anwend	0.09	(0.06)	0.10	(0.07)	-0.02	(0.10)
<i>Klassenebene (Level 2)</i>						
Intakt	-0.06	(0.05)	-0.01	(0.07)	-0.05	(0.09)
Exper	-0.02	(0.05)	0.04	(0.05)	-0.06	(0.08)
Forsch	0.04	(0.06)	-0.17	(0.10)	0.21	(0.12)
Anwend	0.05	(0.04)	0.06	(0.05)	-0.01	(0.06)

Anmerkung. Standardisierte Koeffizienten für HP und NHP, Differenzwert nicht standardisiert. KFT = Kognitiver Grundfähigkeitstest, HISEI = Highest Socio Economic Index, Intakt = Interaktives Lehren und Lernen, Exper = Experimentieren, Forsch = Schülerinnen und Schüler forschen, Anwend = Anwendungen.

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

Im Hinblick auf die Unterrichtsmerkmale konnten auf Klassenebene keine Unterschiede in den Koeffizienten der Unterrichtsmerkmale zwischen beiden Leistungsgruppen identifiziert werden. Innerhalb der Klassen unterschied sich die Unterrichtsaktivität „Experimentieren“ in ihrer Vorhersagekraft für die naturwissenschaftliche Kompetenz von High und Nicht-High-Performern. Sie sagte die naturwissenschaftliche Kompetenz von Nicht-High-Performern innerhalb Klassen positiv vorher, während sie unter High Performern nicht als signifikanter Prädiktor identifiziert werden konnte. Obwohl in Modell III und IV bei High Performern kein Zusammenhang und bei Nicht-High-Performern ein negativer Zusammenhang auf Schülerebene für Forschen beobachtet wurde, war der Unterschied des Zusammenhangs von Forschen und naturwissenschaftlicher Kompetenz

zwischen High und Nicht-High-Performern nicht statistisch bedeutsam.

Insgesamt wurden damit auf Schülerebene lediglich für Nicht-High-Performer signifikante Prädiktoren unter den Unterrichtsmerkmalen identifiziert, auf Klassenebene für keine der Leistungsgruppen. Obwohl im Strukturgleichungsmodell keine der Unterrichtsaktivitäten naturwissenschaftliche Kompetenz vorhersagte, zeigen die Korrelationen im zugehörigen CFA-Modell (Tabellen A.1 und A.2), dass sowohl bei High als auch bei Nicht-High-Performern zumindest ein – wenngleich negativer – Zusammenhang zwischen interaktivem Lehren und Lernen sowie forschendem Lernen und naturwissenschaftlicher Kompetenz bestand. Der Zusammenhang war für beide Leistungsgruppen gleich hoch ausgeprägt. Damit stehen diese Unterrichtsaktivitäten zwar in paarweisem Zusammenhang mit naturwissenschaftlicher Kompetenz, sagen sie jedoch nicht vorher bzw. erklären keine Varianz.³

In den Hypothesen wurde angenommen, dass häufigeres interaktives Lehren und Lernen und häufigerer forschend-entdeckender Unterricht mit einer höheren naturwissenschaftlichen Kompetenz und einem höheren Interesse von High Performern einhergeht. Darüber hinaus wurde vermutet, dass der Zusammenhang von interaktivem Lehren und Lernen und naturwissenschaftlicher Kompetenz sowie ein eventuell vorhandener Zusammenhang von Anwendungen und naturwissenschaftlicher Kompetenz für High Performer höher ausgeprägt sind als für Nicht-High-Performer.

Die Ergebnisse widersprechen den angenommenen Hypothesen. Insgesamt wurden keine Zusammenhänge von interaktivem Lehren und Lernen, forschend-entdeckendem Unterricht und Anwendungen mit der naturwissenschaftlichen Kompetenz von High Performern auf Klassenebene beobachtet. Ebenso fanden sich für die restlichen Schülerinnen und Schüler keine signifikanten Zusammenhänge auf Klassenebene. Zusammenhänge fanden sich dagegen für Nicht-High-Performer und „Experimentieren“ und „Schülerinnen und Schüler forschen“ auf Schülerebene. Dabei bleibt jedoch offen, ob manche Nicht-High-Performer tatsächlich mehr Gelegenheiten für Schülerexperimente oder Forschen erhielten und das zu einer besseren Leistung führte oder ob leistungsstärkere Nicht-High-Performer ihren Unterricht in dieser Hinsicht positiver bzw. negativer wahrnahmen als leistungsschwächere Mitschüler (vgl. Wittwer, 2008).

³Alle Zusammenhänge wurden auch in einem separaten Modell für jede Unterrichtsaktivität untersucht. Auch in diesen Modellen, in denen jede Unterrichtsaktivität jeweils der einzige Unterrichtsprädiktor war, gab es keine Zusammenhänge auf Klassenebene und ähnliche Zusammenhänge auf Schülerebene.

Kompositionseffekte und Unterrichtsaktivitäten im Zusammenhang mit naturwissenschaftlicher Kompetenz

In den beiden vorherigen Abschnitten wurde gezeigt, dass zwar Kompositionseffekte durch die Anzahl der High Performer und die Streuung der Klasse bestand, jedoch kein Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz auf Klassenebene beobachtet werden konnte. Lediglich Zusammenhänge auf Schülerebene wurden beobachtet. Nachfolgend werden beide Prädiktormodelle in einem Modell kombiniert und untersucht, ob Kompositionseffekte unter Kontrolle der Unterrichtsaktivitäten bzw. umgekehrt Zusammenhänge für die Unterrichtsaktivitäten unter Kontrolle der Kompositionseffekte auftraten. Tabelle 35 zeigt die Koeffizienten für High und Nicht-High-Performer sowie die Differenzen der Koeffizienten zwischen beiden Leistungsgruppen.

Tabelle 35.

Kompositionseffekte und Zusammenhänge von Unterrichtsaktivitäten mit der naturwissenschaftlichen Kompetenz von High und Nicht-High-Performern

	HP		NHP		Differenz HP-NHP	
	β	(SE)	β	(SE)	<i>b</i>	(SE)
<i>Schülerebene (Level 1)</i>						
Mädchen	-0.08**	(0.03)	-0.06*	(0.03)	-0.02	(0.05)
KFT	0.16***	(0.03)	0.22***	(0.04)	-0.04	(0.03)
HISEI	0.03	(0.04)	0.10***	(0.03)	-0.04	(0.02)
Intakt	0.01	(0.06)	-0.01	(0.06)	0.02	(0.09)
Exper	-0.05	(0.06)	0.20*	(0.10)	-0.27 *	(0.13)
Forsch	-0.12	(0.06)	-0.32**	(0.10)	0.21	(0.13)
Anwend	0.10	(0.07)	0.10	(0.07)	-0.02	(0.10)
<i>Klassenebene (Level 2)</i>						
Anteil HP	0.63***	0.15	0.60***	0.08	0.01	(0.00)
Streuung Klasse	0.81***	0.15	-0.56***	0.14	-1.36 ***	(0.16)
Intakt	-0.03	0.05	0.04	0.05	-0.07	(0.07)
Exper	-0.02	0.05	0.02	0.04	-0.04	(0.05)
Forsch	-0.02	0.05	-0.12	0.08	0.10	(0.09)
Anwend	0.04	0.04	-0.02	0.05	0.05	(0.05)
R ² Level 1	0.05***	(0.01)	0.12***	(0.03)		
R ² Level 2	0.95***	(0.21)	0.95***	(0.11)		

Anmerkung. Standardisierte Koeffizienten für HP und NHP, Differenzwert nicht standardisiert. KFT = Kognitiver Grundfähigkeitstest, HISEI = Highest Socio Economic Index, Intakt = Interaktives Lehren und Lernen, Exper = Schülerexperimente, Forsch = Schülerinnen und Schüler forschen, Anwend = Anwendungen.

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

In einem Modell, das sowohl Unterricht als auch den Anteil der High Performer einer Klasse sowie die Streuung einer Klasse berücksichtigt, zeigte sich, dass signifikante Kompositionseffekte bestehen, jedoch ähnlich wenige Unterrichtszusammenhänge.

Sowohl bei High als auch bei Nicht-High-Performern sagte der Anteil der High Performer einer Klasse die naturwissenschaftliche Kompetenz der High bzw. Nicht-High-Performer positiv vorher, d. h. in Klassen mit mehr High Performern erreichten sowohl die High als auch die Nicht-High-Performer eine höhere Leistung. Die Koeffizienten unterscheiden sich nicht zwischen beiden Leistungsgruppen. Die Streuung der Klasse zeigte dagegen bei beiden Leistungsgruppen unterschiedlich gerichtete Effekte. In Klassen mit einer höheren Leistungsstreuung, d. h. heterogenen Klassen, erreichten High Performer im Vergleich zu homogeneren Klassen eine höhere naturwissenschaftliche Kompetenz. Nicht-High-Performer dagegen erzielten in homogeneren Klassen höhere Leistungen. Bezüglich der Effekte des Anteils der High Performer und der Streuung der Klasse zeigten sich in den Analysen mit Unterrichtsaktivitäten demnach die gleichen Ergebnisse wie in den Analysen zu Kompositionseffekten. Die Koeffizienten erreichen in der kombinierten Analyse jedoch höhere Werte. Auffällig ist, dass das Modell mit Kompositionseffekten und Unterrichtsaktivitäten als Prädiktoren bei beiden Leistungsgruppen 95% der Varianz auf Klassenebene aufklärt, während in den Modellen, die nur Unterrichtsaktivitäten enthielten, lediglich 12% und 30% der Varianz aufgeklärt wurden (vgl. Tabellen 32 und 33).

Bei den Unterrichtsaktivitäten konnten die Häufigkeit von Gelegenheiten für Schülerexperimente in Relation zu den Mitschülern naturwissenschaftliche Kompetenz von Nicht-High-Performern positiv vorhersagen, die Häufigkeit von Gelegenheiten für forschende Aktivitäten in Relation zu den Mitschülern konnte die naturwissenschaftliche Kompetenz von Nicht-High-Performern negativ vorhersagen. Für High Performer konnte kein Zusammenhang nachgewiesen werden. Damit unterschieden sich auch die Koeffizienten der Unterrichtsaktivitäten im Modell mit Anteil der High Performer und Streuung der Klasse als Prädiktoren nicht vom ursprünglichen Modell mit lediglich den Unterrichtsaktivitäten.

7.2.4. Unterrichtsaktivitäten als Prädiktoren des Interesses von High Performern an Naturwissenschaft

Analog zur naturwissenschaftlichen Kompetenz wurde der Zusammenhang häufiger berichteter Unterrichtsaktivitäten mit dem Interesse von High Performern an Naturwissenschaften untersucht und Unterschieden zwischen High und Nicht-High-Performern betrachtet. Auch hier wurden die Unterrichtsmerkmale nacheinander als Prädiktoren eingefügt und Unterschiede zwischen den Leistungsgruppen durch die Schätzung der Koeffizientendifferenzen ermittelt. Abbildung 11 zeigt das Pfaddiagramm des latent-manifesten Modells, dass in einem Multigruppenvergleich berechnet wurde.

Varianz auf Klassenebene

Analog zur Analyse der naturwissenschaftlichen Kompetenz wurde auch für das naturwissenschaftliche Interesse untersucht, welcher Anteil der Varianz auf Klassenebene liegt. Die Intraclass-Korrelations-Koeffizienten von Interesse betragen für High Performer $ICC_{1,HP} = .06$ und für Nicht-High-Performer $ICC_{1,NHP} = .06$. Eine Berücksichtigung der Mehrebenenstruktur ist daher sinnvoll (Geiser, 2011). Das zweite Maß, der Design Effect, erreicht bei High Performern den Wert $ICC_{2,HP} = .92$ und bei Nicht-High-Performern $ICC_{2,NHP} = .95$. Beide Werte liegen damit unter der Schwelle zwei, ab der eine Mehrebenenanalyse als notwendig angesehen wird. Insgesamt lässt sich schlussfolgern, dass Interesse Varianz auf Klassenebene aufweist und eine Mehrebenenanalyse sinnvoll, jedoch nicht zwingend notwendig ist. Da hier jedoch insbesondere auch Unterrichtsmerkmale auf Klassenebene im Vordergrund stehen, wird die Mehrebenenstruktur im Folgenden berücksichtigt.

Zusammenhang von Unterrichtsaktivitäten und Interesse an Naturwissenschaften von High Performern

In der Mehrebenenanalyse wurden die kognitive Grundfähigkeit, das Geschlecht, der sozioökonomische Hintergrund und das Selbstkonzept in Naturwissenschaften als Kontrollvariablen berücksichtigt. Das Messmodell erreicht einen guten ($CFI = .93$, $TLI = .93$) bzw. sehr guten Fit ($RMSEA = .04$, $SRMR_{Level1} = .04$, $SRMR_{Level2} = .10$). Die Kovarianzen, Varianzen und Korrelationen der Faktoren finden sich in Tabellen A.3 und A.3 im Anhang.

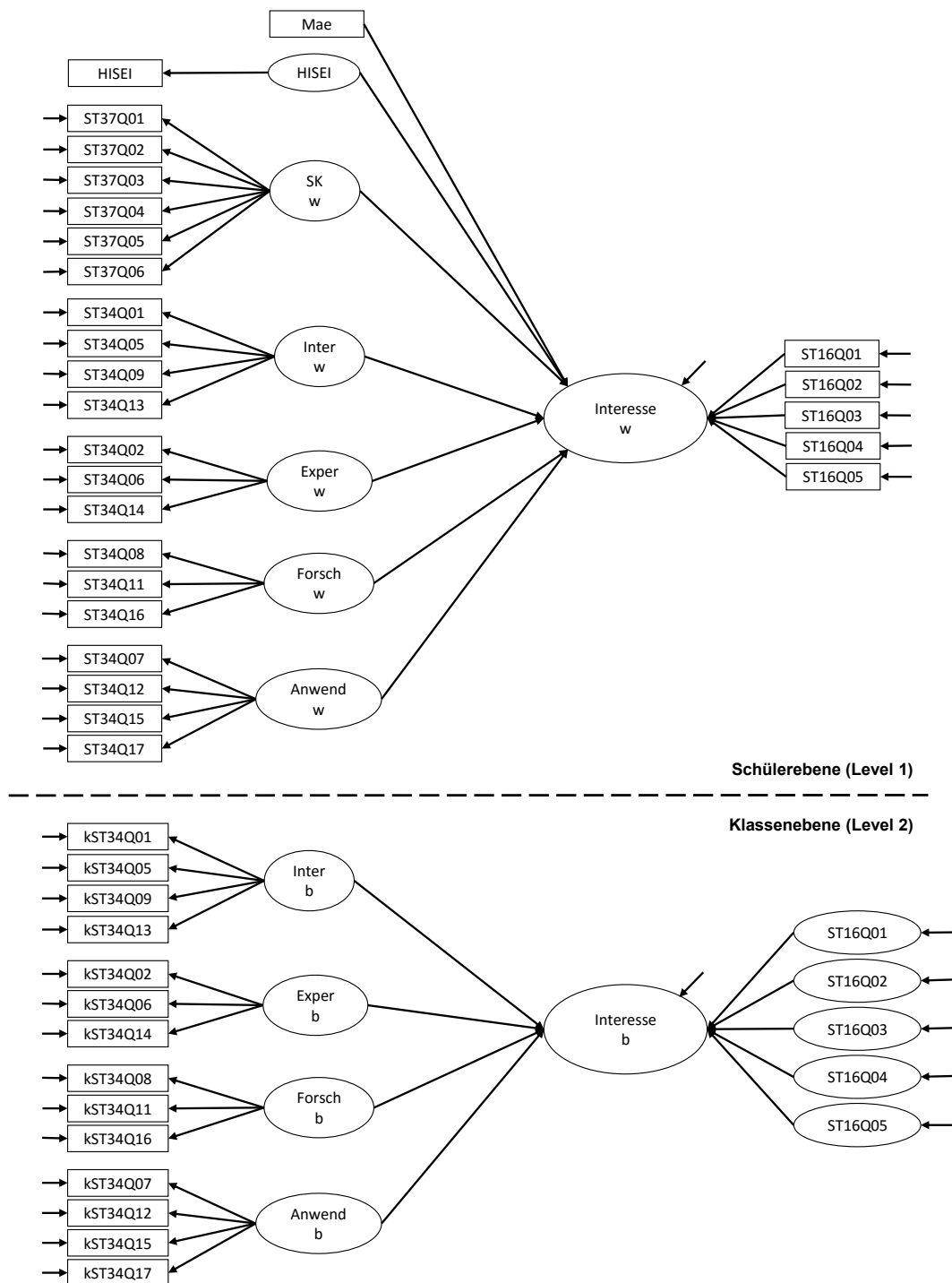


Abbildung 11. Latent-manifestes Modell für Interesse an den Naturwissenschaften (Modell wurde in einem Multigruppenvergleich für High und Nicht-High-Performer gerechnet)

Tabelle 36 gibt einen Überblick über die Ergebnisse für die High Performer. Wie die Ergebnisse für Modell I zeigen, sagte nur das Selbstkonzept von High Performern ihr Interesse an Naturwissenschaften signifikant positiv vorher. Dieser Zusammenhang konnte auch in den Modellen II-IV beobachtet werden. Im Unterschied dazu hatte das Geschlecht keine prädiktive Wirkung auf das Interesse von High Performern

Tabelle 36.

Unterrichtsaktivitäten im Zusammenhang mit dem Interesse von High Performern an Naturwissenschaften

	Modell I		Modell II		Modell III		Modell IV	
	β	(SE)	β	(SE)	β	(SE)	β	(SE)
<i>Schülerebene (Level 1)</i>								
Mädchen	-0.05	(0.03)	-0.05	(0.03)	-0.05	(0.03)	-0.05	(0.03)
HISEI	-0.01	(0.03)	-0.01	(0.03)	-0.01	(0.03)	-0.01	(0.03)
SK	0.61***	(0.03)	0.61***	(0.03)	0.62***	(0.03)	0.60***	(0.03)
Intakt			-0.02	(0.03)	-0.03	(0.04)	-0.07	(0.05)
Exper					0.08	(0.05)	0.05	(0.05)
Forsch					-0.06	(0.06)	-0.12	(0.06)
Anwend							0.17***	(0.04)
<i>Klassenebene (Level 2)</i>								
Intakt			0.02	(0.06)	0.01	(0.08)	-0.02	(0.05)
Exper					0.04	(0.11)	0.02	(0.05)
Forsch					-0.02	(0.12)	-0.02	(0.06)
Anwend							0.08	(0.04)
R ² Level 1	0.37***	(0.03)	0.38***	(0.03)	0.38***	(0.03)	0.40***	(0.03)
R ² Level 2			0.02	(0.04)	0.05	(0.10)	0.18	(0.18)

Anmerkung. Alle Koeffizienten sind standardisierte Werte. HISEI = Highest Socio Economic Index, SK = Selbstkonzept in Naturwissenschaften, Intakt = Interaktives Lehren und Lernen, Exper = Schülerexperimente, Forsch = Schülerinnen und Schüler forschen, Anwend = Anwendungen.

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

Interaktives Lehren und Lernen zeigte als erstes berücksichtigtes Unterrichtsmerkmal keinen signifikanten Zusammenhang mit dem Interesse von High Performern, weder als alleiniger Prädiktor (Modell II), noch in Kombination mit forschend-entdeckendem Unterricht oder Anwendungsbezügen (Modell III und IV). Auch auf Schülerebene, d. h. innerhalb der Klassen, war häufigeres interaktives Lehren und Lernen nicht mit einem höheren Interesse assoziiert.

Wie Modell III verdeutlicht, konnte auch die Häufigkeit des forschend-entdeckenden Unterrichts einer Schulklasse nicht vorhersagen, wie hoch das Interesse der High Performer dieser Schulklasse war. Auch innerhalb der Klassen war kein Zusammenhang zwischen häufigeren Gelegenheiten für forschend-entdeckendes Lernen und Interesse von High Performern zu beobachten.

Für Anwendungen (Modell IV) zeigte sich auf Klassenebene ein ähnliches Bild. Auf Schülerebene konnten Anwendungen jedoch als signifikanter Prädiktor für Interesse von High Performern identifiziert werden. High Performer, die im Vergleich zu ihren leistungsstarken Mitschülern berichten, dass in ihrem Naturwissenschaftsunterricht häufiger Anwendungsbezüge vorkommen, sind an Naturwissenschaften interessierter als ihre leistungsstarken Mitschüler.

Tabelle 37 gibt eine Übersicht über die Ergebnisse für Nicht-High-Performer. Analog zum Ergebnis der High Performer sagte das Selbstkonzept auch bei den Nicht-High-Performern das Interesse an Naturwissenschaften positiv vorher. Im Unterschied zu den High Performern war das Geschlecht bei den Nicht-High-Performern in Modell I und II zusätzlich positiver Prädiktor, d. h. Mädchen hatten ein höheres Interesse. Der Zusammenhang verschwand jedoch mit der Aufnahme der Häufigkeit forschend-entdeckenden Unterrichts als unabhängige Variable.

Tabelle 37.

Unterrichtsaktivitäten im Zusammenhang mit dem Interesse von Nicht-High-Performern an Naturwissenschaften

	Modell I		Modell II		Modell III		Modell IV	
	β	(SE)	β	(SE)	β	(SE)	β	(SE)
<i>Schülerebene (Level 1)</i>								
Mädchen	0.05*	(0.03)	0.05*	(0.03)	0.04	(0.03)	0.04	(0.03)
HISEI	0.01	(0.03)	0.00	(0.03)	0.01	(0.03)	0.01	(0.03)
SK	0.54***	(0.03)	0.54***	(0.03)	0.55***	(0.03)	0.53***	(0.03)
Intakt			0.01	(0.03)	0.06	(0.05)	0.00	(0.05)
Exper					0.12	(0.08)	0.11	(0.07)
Forsch					-0.18**	(0.09)	-0.29**	(0.09)
Anwend							0.21***	(0.06)
<i>Klassenebene (Level 2)</i>								
Intakt			0.03	(0.06)	0.03	(0.09)	-0.03	(0.05)
Exper					-0.03	(0.09)	-0.06	(0.04)
Forsch					0.02	(0.12)	0.00	(0.06)
Anwend							0.14***	(0.04)
R^2 Level 1	0.29***	(0.03)	0.30***	(0.03)	0.31***	(0.03)	0.33***	(0.03)
R^2 Level 2			0.04	(0.09)	0.07	(0.12)	0.56	(0.31)

Anmerkung. Alle Koeffizienten sind standardisierte Werte. HISEI = Highest Socio Economic Index, SK = Selbstkonzept in Naturwissenschaften, Intakt = Interaktives Lehren und Lernen, Exper = Schülerexperimente, Forsch = Schülerinnen und Schüler forschen, Anwend = Anwendungen.

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

Ebenso wie für High Performer zeigt Modell II für Nicht-High-Performer, dass interaktives Lehren und Lernen weder auf Klassen- noch auf Schülerebene vorhersagte, welche naturwissenschaftliche Kompetenz Nicht-High-Performer erreichten. In Modell

III wurde dagegen deutlich, dass die Häufigkeit von „Schülerinnen und Schüler forschen“ bei Nicht-High-Performern im Unterschied zu High Performern auf Schülerebene prädiktiv ist. Während die Häufigkeit forschend-entdeckenden Unterrichts allgemein auf Klassenebene keinen Zusammenhang zeigte, geben Nicht-High-Performer, die innerhalb ihrer Klasse häufiger Gelegenheiten für forschende Aktivitäten berichten als ihre Mitschüler, ein niedrigeres Interesse an als die anderen Nicht-High-Performer der Klasse. Dieser Zusammenhang zeigte sich auch in Modell IV unter Kontrolle der Häufigkeit von Anwendungen.

Wie Modell IV zeigt, sagte auch die Häufigkeit von Anwendungen auf Schülerebene das Interesse von Nicht-High-Performern vorher, im Unterschied zu forschenden Aktivitäten jedoch positiv. Berichteten Nicht-High-Performer häufigere Anwendungsbezüge als ihre Mitschüler, so waren sie interessierter als Nicht-High-Performer in ihrer Klasse, die weniger Anwendungen berichteten. Wie häufig Anwendungen in einer Schulklasse vorkamen, spielte jedoch statistisch keine Rolle für das Interesse der Nicht-High-Performer einer Schulklasse.

Betrachtet man die statistischen Unterschiede in den Koeffizienten zwischen der Gruppe der High Performer und der Gruppe der Nicht-High-Performer (vgl. Tabelle 38), so zeigte sich lediglich ein Unterschied im Koeffizient des Geschlechts. Obwohl das Geschlecht in Modell IV bei beiden Leistungsgruppen im Unterschied zu einigen anderen Modellen kein signifikanter Prädiktor von Interesse war, unterschied sich der Zusammenhang, indem er bei High Performern tendenziell ein höheres Interesse für Jungen, bei Nicht-High-Performern für Mädchen andeutete. Bei sämtlichen Unterrichtsmerkmalen wurde kein Unterschied im Zusammenhang mit Interesse zwischen High und Nicht-High-Performern nachgewiesen, weder auf Klassen- noch auf Schülerebene. Obwohl forschende Aktivitäten lediglich für Nicht-High-Performer als negativer Prädiktor für Interesse identifiziert werden konnten, unterschieden sich die Koeffizienten von forschenden Aktivitäten nicht zwischen beiden Leistungsgruppen.

Ähnlich zur naturwissenschaftlichen Kompetenz ließen sich auch bei Interesse zwar keine Prädiktoren im Strukturgleichungsmodell finden, aber Zusammenhänge im CFA-Modell (Tabellen A.3 und A.4). Anwendungsbezüge im Naturwissenschaftsunterricht einer Schulklasse hingen signifikant mit dem Interesse von Nicht-High-Performern dieser Klasse zusammen. Für High Performer wurde dagegen auch im CFA-Modell kein

Tabelle 38.
 Vergleich der Koeffizienten von Modell IV für Interesse zwischen High und Nicht-High-Performern

	HP		NHP		Differenz HP-NHP	
	β	(SE)	β	(SE)	<i>b</i>	(SE)
<i>Schülerebene (Level 1)</i>						
Mädchen	-0.05	(0.03)	0.04	(0.03)	-0.15*	(0.07)
HISEI	-0.01	(0.03)	0.01	(0.03)	-0.02	(0.04)
SK	0.60***	(0.03)	0.53***	(0.03)	0.04	(0.05)
Intakt	-0.07	(0.05)	0.00	(0.05)	-0.12	(0.11)
Exper	0.05	(0.05)	0.11	(0.07)	-0.09	(0.15)
Forsch	-0.12	(0.06)	-0.29**	(0.09)	0.20	(0.17)
Anwend	0.17***	(0.04)	0.21***	(0.06)	-0.06	(0.12)
<i>Klassenebene (Level 2)</i>						
Intakt	-0.02	(0.05)	-0.03	(0.05)	0.01	(0.07)
Exper	0.02	(0.05)	-0.06	(0.04)	0.08	(0.07)
Forsch	-0.02	(0.06)	0.00	(0.06)	-0.02	(0.09)
Anwend	0.08	(0.04)	-0.03	(0.04)	-0.06	(0.06)

Anmerkung. Standardisierte Koeffizienten für HP und NHP, Differenzwert nicht standardisiert. HISEI = Highest Socio Economic Index, SK = Selbstkonzept in Naturwissenschaften, Intakt = Interaktives Lehren und Lernen, Exper = Schülerexperimente, Forsch = Schülerinnen und Schüler forschen, Anwend = Anwendungen.

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

Zusammenhang auf Klassenebene gefunden.⁴

Für den Zusammenhang von Unterrichtsaktivitäten und Interesse an den Naturwissenschaften wurde in den Hypothesen angenommen, dass sowohl häufigeres interaktives Lehren und Lernen als auch forschend-entdeckender Unterricht und insbesondere Anwendungen mit einem höheren Interesse von High Performern einhergehen (vgl. Kapitel 5.2). Unterschiede wurden nur für interaktives Lehren und Lernen dahingehend wahrgenommen, dass der Zusammenhang für High Performer möglicherweise höher ausgeprägt ist als für Nicht-High-Performer. Insgesamt konnten für High Performer – ähnlich wie bei naturwissenschaftlicher Kompetenz – jedoch keine Unterrichtsaktivitäten als Prädiktoren für Interesse an Naturwissenschaften auf Klassenebene identifiziert werden. Das Selbstkonzept sowie im Vergleich zum Klassendurchschnitt häufiger berichtete Anwendungen auf Schülerebene standen als einzige Merkmale in einem positiven Zusammenhang zum Interesse in Naturwissenschaften. Bei Nicht-High-Performern wurden dagegen häufigere Anwendungen als prädiktives Unterrichtsmerkmal auf Klassenebene identifiziert. Analog zu den High Performern zeigten das Selbstkonzept und häufiger berichtete Anwendungen auf Individualebene einen Zusammenhang zu Interesse. Zusätzlich waren innerhalb der Klasse häufiger wahrgenommene forschende Aktivitäten bei Nicht-High-Performern negativ mit Interesse assoziiert.

7.2.5. Zusammenfassung

Zusammenfassend konnten sowohl für naturwissenschaftliche Kompetenz als auch für Interesse an Naturwissenschaften keine der betrachteten Unterrichtsaktivitäten als nennenswerte Prädiktoren auf Klassenebene, d. h. keine Unterrichtsklimaeffekte, identifiziert werden. Einzig Anwendungen zeigten für Nicht-High-Performer und Interesse einen positiven Zusammenhang. Unter den individuellen Merkmalen waren sowohl für High als auch Nicht-High-Performer das Geschlecht negativ und die kognitive Grundfähigkeit positiv mit der naturwissenschaftlichen Kompetenz assoziiert. Im Hinblick auf Zusammenhänge von Unterricht auf der Schülerebene, d. h. Zusammenhängen von innerhalb einer Klasse häufiger berichteten Unterrichtsaktivitäten mit innerhalb einer Klasse

⁴Alle Zusammenhänge wurden auch in einem separaten Modell für jede Unterrichtsaktivität untersucht. Dabei wurde in einem eigenen Modell für Anwendungsbezüge ein signifikanter Effekt auf Klassenebene für High Performer gefunden. In den restlichen Modellen, in denen jede Unterrichtsaktivität jeweils der einzige Unterrichtsprädiktor war, gab es keine Zusammenhänge auf Klassenebene und ähnliche Zusammenhänge auf Schülerebene.

höher bzw. niedriger ausgeprägten naturwissenschaftlichen Kompetenz bzw. Interesse, wurde deutlich, dass für Nicht-High-Performer Schülerexperimente und forschende Aktivitäten einen Zusammenhang zu naturwissenschaftlicher Kompetenz aufwiesen. In Bezug auf das Interesse sagten bei High Performern wie Nicht-High-Performern die individuell häufiger berichteten Anwendungen sowie das Selbstkonzept Interesse positiv vorher. Auch hier wurde für Nicht-High-Performer ein zusätzlicher negativer Zusammenhang mit forschendem Lernen gefunden. Betrachtet man die Ergebnisse für Interesse in Kombination mit den Ergebnissen zu naturwissenschaftlicher Kompetenz, so wird deutlich, dass kein Unterrichtsmerkmal innerhalb einer der beiden Leistungsgruppen gleichzeitig mit einer höheren naturwissenschaftlichen Kompetenz und einem höheren Interesse verbunden war.

Insgesamt scheint das Kompetenzniveau von High Performern vollkommen unabhängig von sowohl der auf Klassenebene aggregierten als auch individuell berichteten Häufigkeit unterschiedlicher Unterrichtsaktivitäten sein und auch das Interesse wird höchstens von innerhalb der Klasse häufiger berichteten Anwendungen unterstützt. Alle Merkmale, die bei High Performern als Prädiktoren identifiziert wurden, erwiesen sich auch bei Nicht-High-Performern als Prädiktoren. Dies impliziert, dass Nicht-High-Performer von für High Performer identifizierten Unterrichtsprädiktoren möglicherweise ebenfalls profitieren. In Bezug auf die Zusammenhänge auf Schülerebene ist einschränkend darauf hinzuweisen, dass unklar ist, ob die Schülerinnen und Schüler die Unterrichtsaktivitäten tatsächlich häufiger erleben oder aufgrund positiverer individueller Merkmale lediglich den Unterricht positiver einschätzen.

7.3. Selbstkonzept-Interesse-Typen unter den High Performern und ihre Unterrichtswahrnehmung

Bisherige Studien haben verdeutlicht, dass ein Teil der leistungsstarken Jugendlichen in Naturwissenschaften ein im Vergleich zu ihrer Leistung niedriges Interesse bekundet (z. B. Prenzel & Schütte, 2007). Die Unterrichtsforschung belegte in den letzten Jahrzehnten zudem, dass weniger die Organisation des Unterrichts bzw. Sichtstrukturen, sondern vielmehr die von Schülerinnen und Schülern wahrgenommene Tiefenstrukturen, beispielsweise die motivationsunterstützenden Lehr-Lernbedingungen, die intrinsische Motivation von Schülerinnen und Schülern beeinflussen (Seidel et al., 2006). In Übereinstimmung mit diesen Befunden haben die Ergebnisse im vorherigen Kapitel gezeigt, dass von den Schülerinnen und Schülern wahrgenommene Unterrichtsaktivitäten – eine Art Sichtstrukturen im Naturwissenschaftsunterricht – kaum Unterschiede in der naturwissenschaftlichen Kompetenz und dem Interesse von High Performern aufklären konnten, obwohl sie aus theoretischer Sicht als wirksam angenommen werden (vgl. Kapitel 4). Forschungsansätze zu Profilen individueller Schülervoraussetzungen in Verbindung mit Mikro-Lernumgebungen (*Micro-teaching-learning environments*; z. B. Seidel, 2006; Jurik et al., 2014; Huber et al., 2015) zeigen zudem, dass Jugendliche mit unterschiedlichen Profilen ihren Unterricht unterschiedlich motivationsunterstützend wahrnehmen bzw. unterschiedliche Lernaktivitäten berichten. Unterschiedlich leistungszuversichtliche und interessierte High Performer unterscheiden sich daher möglicherweise darin, wie motivationsunterstützend die Unterrichtsaktivitäten gestaltet sind bzw. sie diese wahrnehmen. Im folgenden Ergebnisteil wird daher der Fragen nachgegangen, welche Profile aus Selbstkonzept und Interesse innerhalb der High Performer existieren, welche Eigenschaften sie aufweisen und wie motivationsunterstützend sie ihren Unterricht wahrnehmen. Es wird außerdem untersucht, ob High Performer unterschiedlicher Profile aus Selbstkonzept und Interesse häufigeren forschend-entdeckenden Unterricht unterschiedlich motivationsunterstützend wahrnehmen.

Aus inhaltlicher Sicht orientiert sich die Arbeit damit an bisheriger Forschung zu Merkmalstypen und Mikro-Lernumgebungen, betrachtet die Merkmalskombinationen jedoch weniger als Voraussetzungen für Interaktions- und Lernprozesse, sondern eher als Veranschaulichung der motivational-affektiven Bildungsergebnisse, die unter leistungsstarken Neuntklässlern erreicht werden. Die Betrachtung der motivational-

affektiven Profile dient dabei der Vertiefung der Forschungsergebnisse zu wenig interessierten leistungsstarken Jugendlichen in Naturwissenschaften und den Ergebnissen in Kapitel 7.2. Indem die motivational-affektiven Typen in ihren Merkmalen charakterisiert und motivationsunterstützende Lehr-Lernbedingungen miteinbezogen werden, werden zwei verschiedene Einflussbereiche auf Hinweise für mögliche Erklärungen untersucht. Mit motivationsunterstützenden Lehr-Lernbedingungen werden damit im Anschluss an die Untersuchung von Unterrichtsaktivitäten in Kapitel 7.2 Merkmale der Unterrichtsqualität in den Blick genommen und mit der Häufigkeit forschend-entdeckendem Unterrichts als strukturelles Merkmal verknüpft. Damit wird die Rolle des Naturwissenschaftsunterrichts für motivational-affektive Merkmale von High Performern ausgehend von Ergebnissen aus Kapitel 7.2 auf Basis theoretischer Überlegungen und eines querschnittlichen Untersuchungsdesigns vertieft.

Für dieses vertiefte Bild in Zusammenhang mit Unterricht werden verschiedene methodische Zugänge verwendet, mit denen den Forschungsfragen entsprechende Befunde identifiziert und gleichzeitig Messfehler und statistische Unsicherheiten möglichst berücksichtigt werden können. Die Arbeit schließt dabei an Forschungsansätze zu Merkmalstypen und Mikro-Lernumgebungen an (vgl. Huber et al., 2015; Jurik et al., 2014; Seidel, 2006). In Erweiterung bzw. im Unterschied zu bisheriger Forschung werden in dieser Arbeit verschiedene Ansätze kombiniert und (a) ein Multigruppenansatz in der latenten Profilanalyse verwendet, um Unterschiede zwischen Mädchen und Jungen zu untersuchen, (b) in den Mehrebenenanalysen ein Unterrichtsklimaansatz zur korrekten Modellierung von Unterrichtsaktivitäten verwendet sowie (c) in den Interaktionsanalysen ausgehend von aktuellen Ansätzen (Preacher, Zhang & Zyphur, 2016) eine *cross-level*-Interaktion ohne Vermischung von Varianzanteilen beider Analyseebenen berechnet. Indem die Unterrichtsaktivitäten auf Klassenebene aus den Werten von sowohl High als auch Nicht-High-Performern aggregiert wurden, wurden Unreabilitäten der Schülerangaben soweit wie methodisch möglich verhindert. Schließlich wurde durch die klassenbasierte Stichprobe in Kombination mit der Verwendung von Klassengewichten der Versuch unternommen, möglichst repräsentative Ergebnisse zu erhalten. Da in dieser Arbeit der Begriff „Profile“ bereits im Rahmen der Kompetenzprofile in Kapitel 7.1 verwendet wurde, wird nachfolgend zur begrifflichen Klarheit von „motivational-affektiven Typen“ gesprochen.

7.3.1. Selbstkonzept-Interesse-Typen unter den High Performern

Da sich Mädchen und Jungen in ihrem Selbstkonzept und Interesse für Naturwissenschaften – anders als in der Leistung – deutlich unterscheiden (z. B. Schütte et al., 2007), wurden die Selbstkonzept-Interesse-Typen für weibliche und männliche High Performer getrennt ermittelt und untersucht, ob sich die Typen unter Mädchen und Jungen in Verlauf und Anteil unterscheiden. Als methodischer Zugang wurde die Latente Profilanalyse (LPA, Lazarsfeld & Henry, 1968) gewählt. Als probabilistisches Clusterverfahren zur Entmischung einer Stichprobe bietet sie verschiedene Vorteile im Vergleich zu anderen personenzentrierten Vorgehensweisen, unter anderem die Verwendung von Zugehörigkeitswahrscheinlichkeiten anstelle deterministischer Zuordnungen zu Typen (vgl. Kapitel 6.4.3).

Die Selbstkonzept- und Interesse-Items wurden für die latente Profilanalyse analog dem Vorgehen von Seidel (2006) dichotomisiert und entsprechend ihrer Reihenfolge zu jeweils zwei Bundles mit jeweils zwei bzw. drei Items zusammengefasst (vgl. Kapitel 6.4.3). Die getrennte LPA für Mädchen und Jungen wurde durch ein *Multigroup Mixture Model* modelliert, in dem neben der unbekanntem latenten Klassenvariable c eine bekannte Klassenvariable g , die das Geschlecht darstellt, spezifiziert wurde (L. K. Muthen & Muthén, 2012). Abbildung 12 veranschaulicht das verwendete Modell.

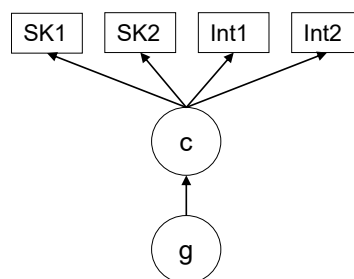


Abbildung 12. Modell der latenten Profilanalyse getrennt für Mädchen und Jungen (SK1: Selbstkonzept Bundle 1, SK2: Selbstkonzept Bundle 2; Int1: Interesse Bundle 1; Int2: Interesse Bundle2; c: latente Klasse; g: Geschlecht)

Das Modell wurde mit verschiedener Anzahl latenter Klassen berechnet und die Modellindizes verglichen (vgl. Tabelle 39). Die relativen Fitindizes AIC und BIC wiesen ein Modell mit fünf latenten Klassen als am besten aus. Die Werte der relativen Verbesserung gegenüber dem Nullmodell bzw. Modell mit einer latenten Klasse weniger zeigten jedoch, dass ab dem Modell mit vier latenten Klassen die relative Verbesserung abnimmt. Zudem erniedrigten sich bei fünf latenten Klassen die Anteile mehrerer latenter Klassen auf drei

Prozent und weniger. Die Anzahl latenter Klassen sollte jedoch so gewählt werden, dass die relativen Anteile bei mindestens 5% liegen (Bacher, Pöge & Wenzig, 2010a). Das Vier-Klassen-Modell war aus Sicht der relativen Verbesserung und der Schüleranteile einem Fünf-Klassen-Modell vorzuziehen. Da die Typen im Vier-Klassen-Modell darüber hinaus auch aus theoretischer Sicht plausibler waren als die Typen im Fünf-Klassen-Modell, wurde das Vier-Klassen-Modell gewählt.

Tabelle 39.

Modellparameter der Latenten Profilanalysen für Mädchen und Jungen mit unterschiedlicher Anzahl latenter Klassen

N Klassen	N Param.	Loglikelihood	$PV_{0K}(\%)$	$PV_K(\%)$	AIC	BIC	ssaBIC
1	13	-3428.80	-	-	6883.6	6953.1	6911.8
2	23	-2180.48	36.41	36.41	4407.0	4529.9	4456.8
3	33	-1753.53	48.86	19.58	3573.1	3749.4	3644.6
4	43	-1271.58	62.91	27.48	2629.2	2858.6	2722.0
5	53	-1081.69	68.45	14.93	2269.4	2552.3	2383.9

Anmerkung. $PV_{0K}(\%)$: Verbesserung gegenüber Nullmodell; $PV_K(\%)$: Verbesserung gegenüber Modell K-1. AIC = Aiken Information Criterion; BIC = Bayesian Information Criterion; ssaBIC = sample-size-adjusted BIC.

Abbildung 13 zeigt die Verläufe der Typen in der Vier-Klassen-Lösung, Tabelle 40 gibt einen Überblick über die Mittelwerte der Merkmale innerhalb der Typen. Sowohl für weibliche als auch männliche High Performer ergaben sich vier verschiedene latente Klassen bzw. Typen mit ähnlichen Verläufen. Aufgrund der Dichotomisierung der Items bzw. der ursprünglichen (rekodierten) Antwortkategorien (1=“stimme gar nicht zu“, 2=“stimme eher nicht zu“, 3=“stimme eher zu“, 4=“stimme ganz zu“) stehen Werte unter 0.5 für eine fehlende Zustimmung, Werte über 0.5 für eine Zustimmung zu den Items. Die mittleren Zugehörigkeitswahrscheinlichkeiten zu den latenten Klassen lagen in einem Wertebereich von 0.92 bis 0.97 und damit in einem sehr guten Bereich (Bacher et al., 2010a).

Typ 1 zeichnete sich bei Jungen und Mädchen durch ein niedriges Selbstkonzept und gleichzeitig niedriges Interesse aus. Schülerinnen und Schüler dieses Typs stimmten sowohl den Selbstkonzept- als auch den Interessenitems im Mittel nicht zu. Sie trauten sich beispielsweise in der Mehrheit nicht zu „anspruchsvollen Stoff im naturwissenschaftlichen Unterricht leicht (zu) lernen“ oder stimmten der Aussage „Ich bin interessiert, Neues in den Naturwissenschaften zu lernen.“ nicht zu. Diese High Performer können als *motivational-affektiv indifferente High Performer* bezeichnet werden, da sie Naturwissenschaften gegenüber motivational-affektiv unbeteiligt sind.

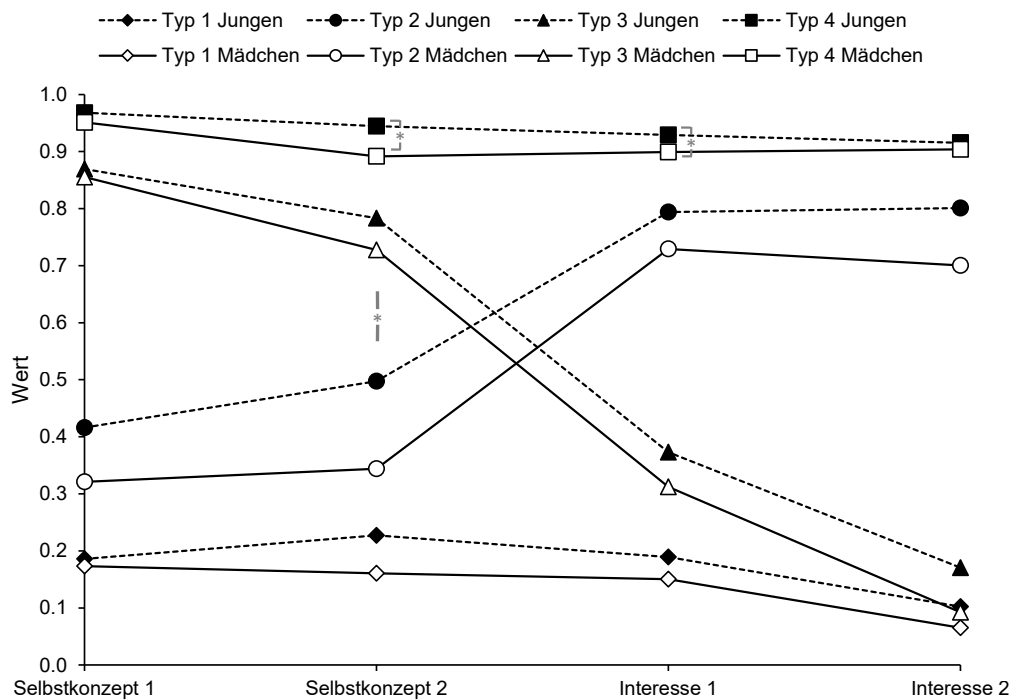


Abbildung 13. Selbstkonzept-Interesse-Typen der High Performer getrennt nach Mädchen und Jungen

Tabelle 40.

Mittelwerte von Selbstkonzept und Interesse in den latenten Klassen sowie Geschlechterdifferenzen

		Typ 1		Typ 2		Typ 3		Typ 4	
		<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>
Jungen	Selbstkonzept 1	0.19	(0.04)	0.42	(0.05)	0.87	(0.02)	0.97	(0.01)
	Selbstkonzept 2	0.23	(0.07)	0.50	(0.04)	0.78	(0.03)	0.94	(0.01)
	Interesse 1	0.19	(0.04)	0.79	(0.03)	0.37	(0.04)	0.93	(0.01)
	Interesse 2	0.10	(0.03)	0.80	(0.05)	0.17	(0.04)	0.92	(0.01)
Mädchen	Selbstkonzept 1	0.17	(0.03)	0.32	(0.05)	0.86	(0.02)	0.95	(0.01)
	Selbstkonzept 2	0.16	(0.03)	0.34	(0.03)	0.73	(0.04)	0.89	(0.02)
	Interesse 1	0.15	(0.04)	0.73	(0.04)	0.31	(0.04)	0.90	(0.01)
	Interesse 2	0.07	(0.02)	0.70	(0.05)	0.09	(0.02)	0.90	(0.01)
Differenz Jungen- Mädchen	Selbstkonzept 1	0.01	(0.05)	0.10	(0.07)	0.01	(0.03)	0.02	(0.01)
	Selbstkonzept 2	0.07	(0.09)	0.15	(0.06)	0.06	(0.04)	0.05	(0.02)
	Interesse 1	0.04	(0.06)	0.06	(0.05)	0.06	(0.05)	0.03	(0.01)
	Interesse 2	0.04	(0.04)	0.10	(0.07)	0.08	(0.04)	0.01	(0.02)

Anmerkung. Signifikante Differenzen ($p < .05$) fettgedruckt.

Typ 2 der Jungen und Mädchen enthielt leistungsstarke Jugendliche in Naturwissenschaften, die ein geringes Selbstkonzept aufwiesen aber sich für Naturwissenschaften interessieren. In Relation zu ihrer Leistung im standardisierten Kompetenztest unterschätzten diese Schülerinnen und Schüler ihre Leistungen in Naturwissenschaften und waren nur wenig leistungszuversichtlich, wenn sie nach ihren Fähigkeiten, Lernprozesse und Leistungssituationen in Naturwissenschaften zu meistern, gefragt wurden. Diese Schülerinnen und Schüler interessierten sich jedoch an Naturwissenschaften und berichteten mehrheitlich, dass sie Spaß an der Beschäftigung mit naturwissenschaftlichen Inhalten haben und gerne Neues in Naturwissenschaften lernen. Sie wurden daher *interessierte High Performer* genannt.

Sowohl unter den Mädchen als auch unter den Jungen konnte ein High-Performer-Typ identifiziert werden, der umgekehrt ein hohes Selbstkonzept aber niedriges Interesse berichtete (Typ 3). Leistungsstarke Jugendliche dieses Typs schätzten ihre Fähigkeiten in Naturwissenschaften zwar positiv ein, zeigten aber wenig bzw. kein Interesse an Naturwissenschaften. Aufgrund dieser Eigenschaft wurden diese beiden Profile *leistungszuversichtliche High Performer* genannt.

Schließlich fand sich in Typ 4 der weiblichen und männlichen High Performer Schülerinnen und Schüler, die sowohl ein hohes Selbstkonzept als auch ein hohes Interesse für Naturwissenschaften aufwiesen. Mit Mittelwerten der Itembundles nahe 1 stimmten diese Jugendlichen den Selbstkonzept- und Interessenitems überwiegend zu und berichteten sowohl ein hohes Zutrauen in ihre Fähigkeiten in Naturwissenschaften als auch Freude bei der Beschäftigung mit naturwissenschaftlichen Inhalten sowie die Bereitschaft, neue Inhalte in Naturwissenschaften zu lernen. Sie werden im Unterschied zu den motivational-affektiv indifferenten High Performern als *motivational-affektiv beteiligte High Performer* bezeichnet.

Zum Vergleich der Typen der latenten Profilanalyse für Mädchen und Jungen wurden (a) die Merkmalsverläufe der Typen und (b) die Verteilung der Typen innerhalb beider Geschlechter betrachtet. Bereits die Beschreibung der Typen zeigt eine große Ähnlichkeit der Typverläufe zwischen Mädchen und Jungen. Betrachtet man die Verläufe der Typen im Detail und vergleicht die mittleren Selbstkonzept- und Interessenwerte der Typen zwischen Mädchen und Jungen, bestätigt sich dieser Eindruck. Wie in Tabelle 40 bzw. mit Sternchen in Abbildung 13 dargestellt, unterschieden sich jeweils die Werte des zweiten Selbstkonzept-Bundles in den Typen der interessierten High Performer

sowie die Werte des zweiten Selbstkonzept- und ersten Interessen-Bundles unter den motivational-affektiv beteiligten High Performern. Jungen des Typs der interessierten High Performer sind demnach etwas leistungszuversichtlicher als Mädchen. Ebenso sind Jungen des motivational-affektiv beteiligten Typs leistungszuversichtlicher als Mädchen des entsprechenden Typs und berichten darüber hinaus etwas mehr Interesse. Die Unterschiede ergaben sich jedoch jeweils nur bei einem Bundle jedes Merkmals und konnten nicht durchgehend beobachtet werden. Insbesondere bei Typ 4 sind die Unterschiede nur minimal und die Signifikanz höchstwahrscheinlich auf die hohe Stichprobenzahl zurückzuführen. Bei Typ 2 sind die Unterschiede bei beiden Selbstkonzept-Bundles höher und möglicherweise aufgrund der Itemzuordnung beim zweiten Bundle signifikant. Die Items wurden zwar relativ zu ihrer Reihenfolge (erstes, drittes etc.) zu Bundles zusammengefasst und unterscheiden sich grundsätzlich nicht in der inhaltlichen Ausrichtung. Das zweite Bundle beinhaltet im Unterschied zum ersten jedoch auch eine Frage zur Leistung bei Prüfungsfragen und zur Einschätzung, ob der Stoff „leicht“ bzw. „einfach“ ist, das erste Bundle enthält Fragen, die häufiger „lernen“ in der Wortwahl verwenden. Während Mädchen des interessierten Profils beide Bundle im Mittel gleich beantworteten, scheinen Jungen zum zweiten Bundle stärker zuzustimmen. Insgesamt decken sich jedoch die Merkmalsverläufe zwischen Jungen und Mädchen weitgehend und es konnten nur vereinzelte Unterschiede beobachtet werden. Das bedeutet, dass innerhalb der weiblichen und männlichen High Performer die gleichen Typen aus Selbstkonzept und Interesse existierten und Selbstkonzept und Interesse bei den Geschlechtern nicht in unterschiedlichen Kombinationen auftraten.

Ein anderes Bild ergab sich bei der Verteilung der Typen. Tabelle 41 gibt einen Überblick über die relativen Häufigkeiten der Typen innerhalb der Mädchen und Jungen. Unter den Jungen gehörten mehr als die Hälfte der High Performer zu den motivational-affektiv beteiligten High Performern mit hohem Selbstkonzept und Interesse. Knapp ein Fünftel der männlichen High Performer war der Gruppe der leistungszuversichtlichen High Performer zuzuordnen und jeweils knapp 10% den interessierten aber wenig leistungszuversichtlichen und den motivational-affektiv indifferenten High Performern. Gut 80% der männlichen High Performer gehörten demnach Typen mit einem hohem Selbstkonzept an, gut 70% Typen mit hohem Interesse.

Bei den Mädchen konnte dagegen nicht wie bei den Jungen eine graduelle Abstufung der Häufigkeiten vom höchsten Wert des vierten Typs zum niedrigsten Wert des

Tabelle 41.

Häufigkeiten und Anteile der Typen innerhalb leistungsstarker Mädchen und Jungen

Typ	Jungen			Mädchen			Differenz	
	N	%	(SE)	N	%	(SE)	%	(SE)
1: Motiv.-aff. indifferente HP	64	7.52	(1.28)	120	17.66	(1.98)	-10.14 ***	(2.36)
2: Interessierte HP	82	9.60	(1.62)	110	15.58	(2.00)	-5.97 *	(2.57)
3: Leistungszuvers. HP	151	18.35	(2.04)	141	19.23	(1.83)	-0.87	(2.74)
4: Motiv.-aff. beteiligte HP	534	64.32	(2.20)	343	47.30	(2.37)	17.02 ***	(3.24)
Nicht zugeordnet	2	0.21	(0.15)	1	0.24	(0.25)		
Summe	832	100.00		716	100.00			

Anmerkung. HP = High Performer; N = Anzahl in Stichprobe; % = Anteil unter Berücksichtigung der Stichprobengewichte.

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

ersten Typs beobachtet werden. Zwar wies bei den Mädchen der vierte Selbstkonzept-Interesse-Typ ebenfalls mit knapp 50% den höchsten Wert auf, für alle drei anderen Klassen wurden jedoch ähnlich hohe Werte im Bereich von 15% bis 20% ermittelt. Während also knapp die Hälfte der Mädchen unter den High Performern motivational-affektiv beteiligt war, wiesen jeweils knapp ein Fünftel der weiblichen High Performer ein geringes Selbstkonzept oder ein geringes Interesse oder beides auf. Fasst man die Klassen jeweils zusammen, wiesen die Mädchen 65% der weiblichen High Performer Typen mit hohem Selbstkonzept auf, gut 60% ein hohes Interesse. Ein Vergleich der Typenanteile zwischen Mädchen und Jungen ergab, dass ein signifikant höherer Anteil der Jungen dem motivational-affektiv beteiligten Typ angehörte und ein höherer Anteil der Mädchen dem interessierten oder motivational-affektiv indifferentem Typ, d. h. häufiger einem Typ mit niedrigem Selbstkonzept (und Interesse), zuzuordnen war. Ähnlich häufig kommen sowohl unter weiblichen als auch männlichen High Performern leistungszuversichtliche aber uninteressierte Jugendliche vor. Weibliche und männliche High Performer unterscheiden sich daher zwar nicht in den Verläufen ihrer Selbstkonzept-Interesse-Typen, jedoch im Vorkommen der Typen. Unterschiede waren vor allem in den Häufigkeiten der beiden extremen Profile der motivational-affektiv beteiligten bzw. indifferenten Jugendlichen zu beobachten. Jungen waren häufiger motivational-affektiv beteiligt, Mädchen gehörten häufiger zu den motivational-affektiv indifferenten High Performern.

Insgesamt konnten daher vier Selbstkonzept-Interesse-Typen identifiziert und die Hypothese bekräftigt werden, dass unter den High Performern unterschiedliche Typen und darunter motivational-affektiv indifferente High Performer gefunden werden

können. Im Hinblick auf Unterschiede zwischen den Geschlechtern, konnte zwar bestätigt werden, dass weibliche High Performer häufiger einem Typ mit niedrigem Selbstkonzept und Interesse angehören. Gleichzeitig zeigte sich jedoch, dass sich Mädchen und Jungen in den Merkmalsverläufen der Typen nicht unterscheiden. Da innerhalb beider Geschlechter die gleichen Typen identifiziert wurden, wurden die Typen der Mädchen und Jungen in den weiteren Analysen jeweils zusammengefasst, so dass im Folgenden insgesamt vier Typen in den weiterführenden Analysen betrachtet wurden.

7.3.2. Individuelle, familiäre und institutionelle Merkmale der Selbstkonzept-Interesse-Typen

Neben der Identifikation von Selbstkonzept-Interesse-Typen soll in dieser Arbeit auch der Frage nachgegangen werden, welche Merkmale High Performer der unterschiedlichen Typen aufweisen und welche Unterschiede zwischen den Typen bestehen. Dafür wurden sowohl die Leistung in Form von Kompetenzen und Schulnoten als individuelle Merkmale als auch der sozioökonomische Hintergrund als familiäres Merkmal und die Verteilung auf Schulklassen als Veranschaulichung dafür, wo die Typen im institutionellen Kontext zu finden sind, untersucht.

Kompetenzen und Noten in unterschiedlichen Fachbereichen

High Performer sämtlicher Typen weisen per Definition eine hohe naturwissenschaftliche Kompetenz auf. In Mathematik und Lesen können jedoch unterschiedliche Kombinationen auftreten. In Kapitel 7.1.1 wurde bereits untersucht, welcher Anteil der High Performer in Naturwissenschaften auch in Mathematik und Lesen zu den 20 besten Prozent gehört. Auch im Zusammenhang mit den Selbstkonzept-Interesse-Typen ist von Bedeutung, wie sich Überschneidungen mit hoher Leistung in Mathematik und Lesen in den Typen verteilen. Ausgehend von der Theorie dimensionaler Vergleiche (z. B. Möller et al., 2006) bzw. allgemein dem Zusammenhang von Leistung und Selbstkonzept bzw. Interesse ist beispielsweise von Bedeutung, ob motivational-affektiv beteiligte High Performer häufiger High Performer mit einem naturwissenschaftlichen Leistungsprofil sind und in welchen Selbstkonzept-Interesse-Typen High Performer, die in allen Fächern hohe Leistungen zeigen, zu finden sind (vgl. Kapitel 3.2). Ersteres würde darauf hinweisen, dass dimensionale Vergleiche möglicherweise eine Rolle dafür spielen, dass ein Teil

der High Performer motivational-affektiv beteiligt wird und ein Teil indifferent gegenüber Naturwissenschaften bleibt. Zweiteres veranschaulicht, in welchen Typen High Performer landen, bei denen aufgrund der vergleichbaren Noten in mehreren Fächern dimensionale Vergleiche mit Noten nicht funktionieren.

Zur Beantwortung dieser Frage wurde die Verteilung der in Kapitel 7.1.1 gebildeten Kompetenzprofile innerhalb der Selbstkonzept-Interesse-Typen untersucht (vgl. 42). In allen Selbstkonzept-Interesse-Typen sind rund 50-60% der Jugendlichen sowohl in Naturwissenschaften als auch Mathematik und Lesen High Performer. Auch innerhalb aller High Performer wurde in Kapitel 7.1.1 ein vergleichbarer Anteil von 53% berichtet. Unter Beachtung der Standardfehler treten alle anderen High-Performance-Überschneidungen relativ gleichmäßig mit Mittelwerten zwischen 8 und 22% auf. Auch hier sind keine auffälligen Unterschiede zur Gesamtstichprobe der High Performer zu beobachten.

Tabelle 42.

Überschneidung der High Performance in unterschiedlichen Domänen innerhalb der Profile

Typ	Beschreibung	High Performance							
		N		N & M		N & L		N & M & L	
		M (%)	(SE)	M (%)	(SE)	M (%)	(SE)	M (%)	(SE)
1	Motivational-affektiv indifferente HP	13.05	(6.06)	16.12	(3.54)	21.06	(5.37)	49.77	(5.76)
2	Interessierte HP	15.06	(6.22)	16.34	(3.67)	15.32	(3.51)	53.28	(6.60)
3	Leistungszuversichtliche HP	9.91	(3.93)	20.43	(4.36)	13.12	(3.09)	56.54	(6.61)
4	Motivational-affektiv beteiligte HP	8.02	(1.47)	21.90	(2.28)	9.96	(1.71)	60.12	(3.29)

Anmerkung. N = Naturwissenschaften; M = Mathematik; L = Lesen.

Neben den Kompetenzprofilen wurde auch betrachtet, wie häufig die in Kapitel 7.1.1 untersuchten Notenprofile innerhalb der Selbstkonzept-Interesse-Typen vorkommen (vgl. Tabelle 43). In einer ersten Analyse wurden das Niveau der mittleren Note in Naturwissenschaften verglichen. Dies verdeutlichte, dass High Performer des motivational-affektiv indifferenten Profils die schlechtesten Noten aufwiesen und sich das Notenniveau über die Profile der interessierten High Performer und der leistungszuversichtlichen High Performer kontinuierlich verbesserte. Die Jugendlichen im Profil der motivational-affektiv beteiligten High Performer wiesen das höchste Notenniveau auf. Während unter den motivational-affektiv indifferenten Jugendlichen knapp ein Drittel

nur Noten im Bereich ausreichend und schlechter in Naturwissenschaften erreichen, beträgt dieser Anteil bei den interessierten High Performern ein Viertel und erreicht bei den motivational-affektiv beteiligten High Performern nur noch rund 8%. Dies könnte beispielsweise auch dadurch zustande kommen, dass sich motivierte High Performer stärker am Unterricht beteiligen und daher bei gleichen Leistungen positiver bewertet werden. Umgekehrt erreicht unter den motivational-affektiv beteiligten High Performern der größte Anteil Noten im Bereich gut und sehr gut. Während dort über die Hälfte der Jugendlichen ein solches hohes Notenniveau aufwiesen, gehörten nur rund ein Fünftel der motivational-affektiv indifferenten Schülerinnen und Schüler zu dieser Gruppe. Diese Ergebnisse zeigen in Übereinstimmung mit dem wechselseitigen Zusammenhang von Leistung und Selbstkonzept bzw. Interesse, dass High Performer mit positiveren Kombinationen von Selbstkonzept und Interesse höhere Schulleistungen erreichen. Sie verdeutlichen andererseits jedoch auch, dass selbst manche High Performer mit guten oder sehr guten Naturwissenschaftsnoten motivational-affektiv indifferent bleiben und einen Anteil von rund einem Fünftel am entsprechenden Selbstkonzept-Interesse-Typ ausmachen.

Neben dem Notenniveau in Naturwissenschaften wurde auch untersucht, welches fachliche Profil High Performer der unterschiedlichen Selbstkonzept-Interesse-Typen in den Naturwissenschaften, Mathematik und Deutsch aufwiesen. Ein Vergleich der Selbstkonzept-Interesse-Typen ergab, dass High Performer mit einem ausgeglichenen Notenprofil, d. h. mit in Naturwissenschaften, Mathematik und Deutsch ähnlich hohen Noten, innerhalb aller Selbstkonzept-Interesse-Typen den größten Anteil hatten. Lediglich bei den interessierten High Performern war das sprachliche Profil mit Mittel etwas häufiger vertreten, aufgrund der Standardfehler ist der Unterschied jedoch nicht signifikant. Bei den beiden Typen mit niedrigem Selbstkonzept (Profil 1 und 2) war ein sprachliches Notenprofil mit bester Note in Deutsch häufiger zu beobachten als bei den beiden Typen mit hohem Selbstkonzept (Profil 3 und 4). Umgekehrt wurde bei den Typen mit hohem Selbstkonzept häufiger ein sprachlich-naturwissenschaftliches Profil gefunden, die Unterschiede sind jedoch nicht signifikant. Auch ein mathematisch-naturwissenschaftliches Profil trat bei den Selbstkonzept-Interesse-Typen mit höherem Selbstkonzept häufiger auf. Eine Zusammenfassung der Anteile mit jeweils naturwissenschaftlichem, naturwissenschaftlich-sprachlichem und naturwissenschaftlich-mathematischem Notenprofil verdeutlichte, dass innerhalb der motivational-affektiv

Tabelle 43.
Verteilung der Notenniveaus und Notenprofile innerhalb der Selbstkonzept-Interesse-Typen der High Performer

	Selbstkonzept-Interesse-Typ							
	1		2		3		4	
	Motiv.-aff. indifferente HP		Interessierte HP		Leistungszuvers. HP		Motiv.-aff. beteiligte HP	
	<i>M</i> (%)	<i>(SE)</i>	<i>M</i> (%)	<i>(SE)</i>	<i>M</i> (%)	<i>(SE)</i>	<i>M</i> (%)	<i>(SE)</i>
<i>Notenniveau Naturw.</i>								
Ausreichend	32.56	(4.56)	24.66	(4.19)	13.33	(2.44)	8.03	(1.20)
Befriedigend	48.49	(4.49)	52.50	(3.65)	48.94	(3.55)	38.30	(2.06)
Gut bzw. sehr gut	18.95	(2.96)	22.84	(3.22)	37.74	(3.44)	53.66	(2.33)
<i>Summe</i>	100.00		100.00		100.00		100.00	
<i>Notenprofil</i>								
Naturw.	6.78	(1.77)	12.61	(3.82)	9.12	(2.51)	14.92	(1.68)
Mathem.	11.13	(2.46)	9.39	(2.97)	10.58	(2.73)	7.02	(1.17)
Sprachl.	24.55	(3.75)	25.48	(3.98)	10.67	(1.82)	9.28	(1.12)
Mathem.-Naturw.	4.61	(1.64)	6.33	(2.41)	12.26	(2.20)	15.33	(1.54)
Sprachl.-Naturw.	11.93	(3.36)	11.53	(3.37)	18.52	(2.46)	14.23	(1.55)
Sprachl.-Mathem.	11.69	(3.38)	10.26	(2.87)	8.42	(2.14)	4.12	(0.68)
Ausgeglichen	29.32	(5.23)	24.40	(3.76)	30.43	(3.62)	35.10	(2.14)
<i>Summe</i>	100.00		100.00		100.00		100.00	
Ausgeglichen: Überall gute Note	9.08	(2.14)	8.16	(2.38)	14.43	(2.73)	24.89	(1.89)

Anmerkung. HP = High Performer; Ausreichend: Note > 3.5; Befriedigend: 2.5 ≤ Note < 3.5; Gut bzw. sehr gut: Note < 2.5.

beteiligten High Performer der höchste Anteil an Jugendlichen mit solchen Profilen auftrat und der Anteil bis hin zum Selbstkonzept-Interesse-Type der motivational-affektiv indifferenten High Performer konstant abnahm. Umgekehrt wiesen die motivational-affektiv indifferenten High Performer den höchsten Anteil an Notenprofilen mit sprachlicher Komponente auf. Auch unter den interessierten High Performern existierte ein ähnlich hoher Wert, während die motivational-affektiv beteiligten High Performer den niedrigsten Anteil an sprachlichen Profilen enthielten. Mathematische Profile traten dagegen in allen Profilen ähnlich häufig auf. Insgesamt waren naturwissenschaftliche Notenprofile häufiger in Selbstkonzept-Interesse-Typen mit höherem Selbstkonzept und/oder Interesse zu finden, sprachliche Notenprofile häufiger in Selbstkonzept-Interesse-Typen mit geringerem Selbstkonzept und/oder Interesse und mathematische Profile überall relativ gleich verteilt. High Performer mit ausgeglichenen Notenprofilen kamen in jedem Selbstkonzept-Interesse-Type ähnlich häufig vor. Die Ergebnisse entsprechen damit weitestgehend den Erwartungen, die auf Basis der Theorie der dimensional Vergleichs (Möller & Marsh, 2013) getroffen wurden (vgl. Kapitel 5).

Unter der Annahme dimensionaler Vergleiche ist jedoch vor allem die Gruppe der High Performer mit in allen Fächern guten Noten – eine Teilgruppe der High Performer mit ausgeglichenem Notenprofil – beachtenswert, denn auf ihre Notenkonstellation sind dimensionale Vergleiche nicht anwendbar bzw. wirkungslos. Mit knapp einem Viertel erreichte das Profil der motivational-affektiv beteiligten High Performer den höchsten Anteil an High Performern mit überall guten Noten. Ein niedrigerer Anteil von um die 15% wurde innerhalb der leistungszuversichtlichen High Performern gefunden. Die beiden anderen Profile mit niedrigem Selbstkonzept zeigten jeweils Anteile von rund einem Zehntel. Ausgehend von der Theorie dimensionaler Vergleiche könnte man erwarten, dass sich die Interessen dieser Jugendlichen nur zum Teil zugunsten der Naturwissenschaften ausdifferenzieren und sie sich stärker für eines der anderen Fächer interessieren. Die Ergebnisse zeigten jedoch, dass die beiden Typen mit hohem Selbstkonzept und vor allem der Typ der High Performer mit hohem Selbstkonzept und Interesse die höchsten relativen Anteile dieser High Performer mit guten Noten in allen Fächern erreichten.

Insgesamt verdeutlichten die Ergebnisse, dass High Performer aus unterschiedlichen Selbstkonzept-Interesse-Typen in ihren Kompetenzüberschneidungen mit Mathematik und Lesen weitgehend vergleichbar waren und keine auffälligen Unterschiede zur Gesamtstichprobe der High Performer aufwiesen. Darüber hinaus wurde verdeutlicht,

dass High Performer aus Selbstkonzept-Interesse-Typen mit höherem Selbstkonzept und/oder Interesse tendenziell ein höheres Notenniveau in Naturwissenschaften erreichten, jedoch auch ein beträchtlicher Teil der motivational-affektiv indifferenten High Performer gute bzw. sehr gute Noten aufwies. Damit konnte für die Noten, nicht jedoch für die Kompetenzen die Hypothese bestätigt werden, dass die Selbstkonzept-Interesse-Typen mit höherem Selbstkonzept und/oder Interesse höhere Leistungen aufweisen. Gute Noten scheinen jedoch nicht hinreichend für ein hohes Selbstkonzept und Interesse zu sein.

Sozioökonomischer Hintergrund

Als familiärer Bedingungsfaktor wurde der mittlere sozioökonomische Hintergrund für die Selbstkonzept-Interesse-Typen betrachtet. Tabelle 44 zeigt den mittleren HISEI-Wert der Selbstkonzept-Interesse-Typen unter den High Performern. Die Werte der Selbstkonzept-Interesse-Typen waren vergleichbar hoch und unterschieden sich nicht statistisch signifikant. Das bedeutet, dass High Performer in Selbstkonzept-Interesse-Typen mit höherem Interesse – anders als in den Hypothesen erwartet – keinen höheren sozioökonomischen Hintergrund aufwiesen als High Performer in Selbstkonzept-Interesse-Typen mit niedrigerem Interesse. Da der HISEI den sozioökonomischen Status eindimensional auf Grundlage des Berufs quantifiziert, könnte mit mehrdimensional konstruierten Indizes wie dem ESCS, der z. B. auch kulturelle Besitztümer erfasst, möglicherweise Zusammenhänge beobachtet werden.

Tabelle 44.

Mittlere HISEI-Werte der unterschiedlichen Selbstkonzept-Interesse-Typen der High Performer

Profil	Beschreibung	HISEI		
		<i>M</i>	<i>(SE)</i>	<i>SD</i>
1	Motivational-affektiv indifferente HP	57.03	(1.26)	14.92
2	Interessierte HP	59.05	(1.20)	16.19
3	Leistungszuversichtliche HP	59.94	(1.04)	14.87
4	Motivational-affektiv beteiligte HP	60.41	(0.86)	15.49

Anmerkung. HP = High Performer.

Verteilung auf Schulklassen

Neben individuellen und familiären Merkmalen stellt sich die Frage, wie sich die Profile auf Schulklassen verteilen und, ob einzelne Schulklassen gegebenenfalls institutionelle Kontexte bieten, in denen günstige Selbstkonzept-Interesse-Typen gehäuft vorkommen. Analog dem Vorgehen von Seidel (2006) wurde betrachtet, welche Schulklassen hinsichtlich der Typen besonders heterogen sind, indem sie mindestens drei von vier Typen enthalten. Auch wie viele Schulklassen in den einzelnen Profilen mindestens einen Anteil von 30% unter ihren High Performern erreichten, wurde in Kombination damit untersucht. Darüber hinaus wurde die Varianz zwischen den Schulklassen mit Hilfe des Intraklassen-Korrelations-Koeffizienten quantifiziert.

Abbildung 14 veranschaulicht die relativen Anteile der vier Selbstkonzept-Interesse-Typen unter den High Performern in den einzelnen Schulklassen. Insgesamt enthielten 103 von 131 Schulklassen, d. h. 78.63%, mindestens drei verschiedene Selbstkonzept-Interesse-Typen. Das bedeutet, dass die High Performer im Großteil der Schulklassen, obwohl sie unterschiedliche Anteile an High Performern enthalten, unterschiedliche motivational-affektive Typen aufweisen und nicht einzelne Klassen für gehäufte Ausprägungen von einem oder zwei Profilen stehen. Umgekehrt scheint vielmehr eine Varianz innerhalb der Klassen zu existieren.

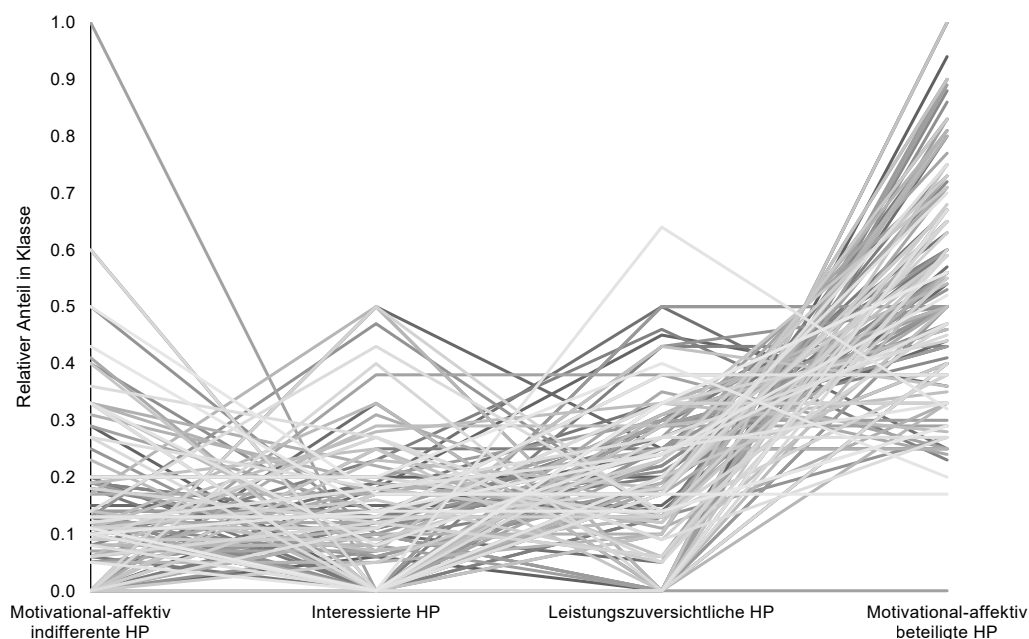


Abbildung 14. Anteile der Selbstkonzept-Interesse-Typen in den Schulklassen

Für eine genaue Erfassung der Varianz zwischen und innerhalb der Schulklassen

wurde der Intraklassen-Korrelations-Koeffizient (ICC) berechnet. Der ICC der Klassenzugehörigkeitsvariable zu den latenten Klassen, denen die Selbstkonzept-Interesse-Typen entsprechen, erreichte über alle PVs hinweg gemittelt einen Wert von $ICC_1 = .061$, d. h. 6% der Varianz der Zugehörigkeit zu den Selbstkonzept-Interesse-Typen liegt zwischen den Schulklassen. Die ICCs der Zugehörigkeitswahrscheinlichkeiten der einzelnen latenten Klassen betrug zwischen $ICC_1 = .023$ und $ICC_1 = .040$, d. h. auch hier lag Varianz zwischen den Schulklassen. Insgesamt bedeutet dies, dass der Großteil der Varianz innerhalb der Schulklassen liegt, ein kleinerer Teil im einstelligen Bereich jedoch auch zwischen Schulklassen.

Neben der Anzahl an Schulklassen mit mindestens drei Profilen und der Varianzaufteilung zwischen und innerhalb Schulklassen wurde analog Seidel (2006) untersucht, wie viele Schulklassen einen Anteil von 30% eines Typs unter ihren High Performern enthielten (vgl. Tabelle 45). In fast 90% aller Schulklassen gehörten mindestens 30% der High Performer dem motivational-affektiv beteiligten Typ an, d. h. sie hatten ein hohes Selbstkonzept und Interesse. In rund einem Fünftel erreichte auch bzw. stattdessen der leistungszuversichtliche Typ mindestens einen Anteil von 30%. Gut ein Zehntel der Gymnasialklassen zeichnete sich durch einen Anteil von mindestens 30% an motivational-affektiv indifferenten Jugendlichen unter den High Performern aus. Dem Anteil von 90% an Klassen mit mindestens rund einem Drittel motivierten High Performern stehen daher nur ein wesentlich geringerer Anteil an Klassen mit mindestens rund einem Drittel wenig motivierten High Performern gegenüber. Insgesamt gelingt es offensichtlich in fast allen Schulklassen, mindestens 30% der High Performer motivational-affektiv zu beteiligen.

Tabelle 45.

Anzahl und Anteil der Klassen mit jeweiligem Anteil eines Selbstkonzept-Interesse-Typs von mindestens 30%

Anteil	Motivational-affektiv indifferente HP	Interessierte HP	Leistungszuversichtliche HP	Motivational-affektiv beteiligte HP
N	18	9	24	114
%	13.74	6.87	18.32	87.02

Anmerkung. HP = High Performer.

7.3.3. Wahrnehmung motivationsunterstützender Lehr- Lernbedingungen von High Performern unterschiedlicher Selbstkonzept-Interesse-Typen

Nachdem die Selbstkonzept-Interesse-Typen identifiziert und hinsichtlich besonderer Merkmale untersucht wurden, beschäftigt sich das folgende Unterkapitel mit der Frage, wie High Performer dieser unterschiedlichen motivational-affektiven Typen Unterricht wahrnehmen. Als motivationsunterstützende Lehr-Lernbedingungen werden die Autonomieunterstützung, die inhaltliche Relevanz sowie die Kompetenzunterstützung in den Blick genommen. Als Hypothesen wurde für Selbstkonzept-Interesse-Typen mit unterschiedlich hohem Interesse vermutet, dass

- High Performer mit hohem Interesse bei gleichzeitig hohem Selbstkonzept sich von High Performern mit niedrigem Interesse bei gleichzeitig hohem Selbstkonzept in der wahrgenommenen Autonomieunterstützung und inhaltlichen Relevanz unterscheiden,
- High Performer aus Typen mit unterschiedlichem Interesse bei niedrigem Selbstkonzept sich lediglich in der wahrgenommenen inhaltlichen Relevanz unterscheiden.

In Bezug auf Unterschiede zwischen Typen mit unterschiedlich hohem Selbstkonzept wurde dagegen angenommen, dass

- High Performer mit leistungszuversichtlichem Typ sich unabhängig vom Interessenniveau in ihrer wahrgenommenen Kompetenzunterstützung von High Performern weniger leistungszuversichtlicher Typen unterscheiden.

Schließlich wurde in Verbindung mit forschend-entdeckendem Unterricht angenommen, dass

- High Performer unterschiedlicher Typen bei Unterricht mit häufigerem forschend-entdeckendem Lernen unterschiedlich positivere Lehr-Lernbedingungen berichten, d. h. eine Interaktion zwischen forschend-entdeckendem Lernen und Profilen im Zusammenhang mit den wahrgenommenen Lehr-Lernbedingungen besteht.

Diese Fragestellungen wurden in Mehrebenen-Pfadanalysen untersucht, indem sowohl für Autonomieunterstützung als auch Kompetenzunterstützung und inhaltliche

Relevanz in Kombination mit den Unterrichtsaktivitäten „Experimente“ und „Schülerinnen und Schüler forschen“ jeweils drei Modelle spezifiziert wurden: In Modell I wurden nur die Selbstkonzept-Interesse-Typen aufgenommen und die Unterschiede ihrer Unterrichtswahrnehmung untersucht. Dabei wurde auch betrachtet, inwieweit die Zusammenhänge auf Schüler- bzw. Klassenebene zu finden sind. In Modell II wurde analysiert, inwieweit die Unterrichtsaktivitäten die jeweilige Lehr-Lernbedingung vorhersagten und in Modell III wurde schließlich die Interaktion von Typen und jeweiliger Unterrichtsaktivität als Prädiktor für die betrachtete Lehr-Lernbedingung aufgenommen. Abbildung 15 zeigt Modell III als Pfadmodell. Die Interaktionen zwischen Unterrichtsaktivität und Selbstkonzept-Interesse-Typen sind Interaktionen von Variablen der Klassen- mit Variablen der Schülerebene (*cross-level interactions*). Durch die Verwendung eines Mehrebenen-Ansatzes wurde auch hier die genestete Struktur der Stichprobe berücksichtigt und darüber hinaus untersucht, ob sich High Performer unterschiedlicher Selbstkonzept-Interesse-Typen über Schulklassen hinweg oder umgekehrt vor allem innerhalb Schulklassen, d. h. in Mikro-Lernumgebungen innerhalb Schulklassen, in ihrer Unterrichtswahrnehmung unterscheiden. Im Unterschied zu den Analysen in Kapitel 7.2 wurden hier alle Merkmale doppelt-manifest modelliert, d. h. mit manifest gebildeten Skalenwerten statt Faktoren und manifest gebildeten Werten auf Klassenebene. Eine latent-manifeste Modellierung hätte Schwierigkeiten in der Rechenkapazität verursacht, da dafür numerische Integrationen mit mehreren Integrationspunkten unter Verwendung eines Monte-Carlo-Ansatzes notwendig gewesen wären.

Die Selbstkonzept-Interesse-Typen wurden in Form ihrer Klassenzugehörigkeitswahrscheinlichkeiten als Prädiktoren aufgenommen. Dieses regressionsanalytische Vorgehen mit Klassenzugehörigkeitswahrscheinlichkeiten als Prädiktoren entspricht einer ANOVA und erlaubt dabei die Berücksichtigung der Mehrebenenstruktur und die spätere Aufnahme von Interaktionen zwischen Klassen- und Schülerebene. Die Verwendung der Klassenzugehörigkeitswahrscheinlichkeiten anstelle dichotomer Klassenzugehörigkeitsvariablen ermöglicht eine genauere Schätzung, da sie die Unsicherheiten der Klassenzugehörigkeit mit berücksichtigt (Bacher et al., 2010a; Pastor et al., 2007). Als Referenzkategorie wurde der Typ der motivational-affektiv indifferenten High Performer (Typ 1), d. h. der Typ mit der niedrigsten Ausprägung von Selbstkonzept und Interesse gewählt. Zusätzlich wurden in den Modellen jeweils die Differenzen der Regressionskoeffizienten der Typen mit Standardfehler geschätzt und so auf Signifikanz überprüft.

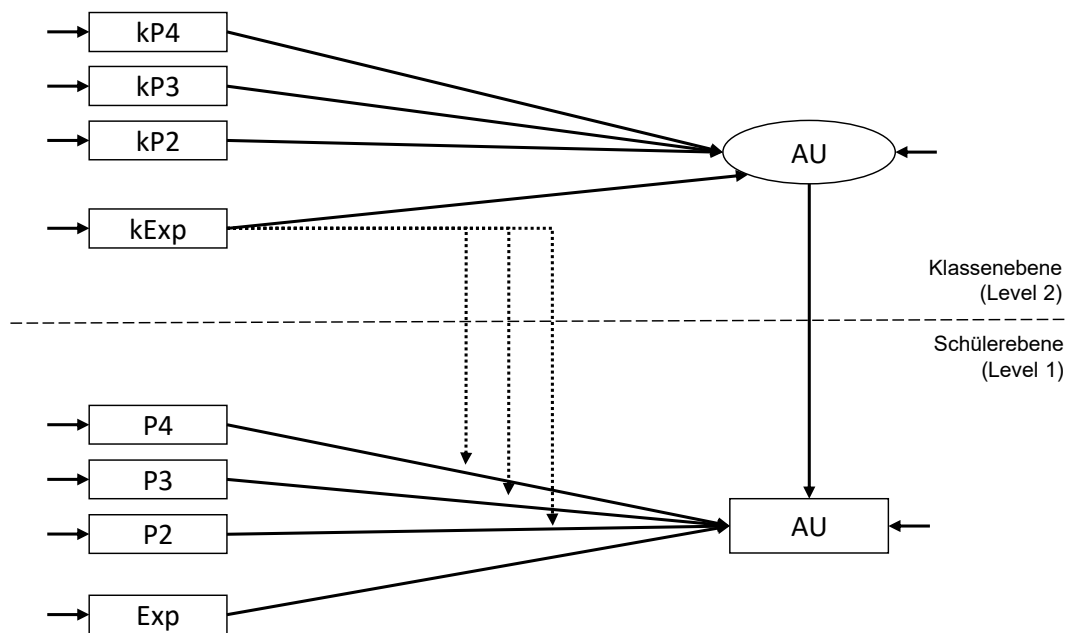


Abbildung 15. Modell III mit Cross-Level-Interaktionen als Pfadmodell (gestrichelte Linien: Cross-Level-Interaktionen; AU = Autonomieunterstützung; P2, P3 und P4 = Zugehörigkeitswahrscheinlichkeiten zu Typ 2, Typ 3 und Typ 4 (Typ 2 = interessierte HP, Typ 3 = leistungszuversichtliche HP, Typ 4 = motivational-affektiv beteiligte HP); kP2, kP3 und kP4 = Klassenmittelwerte der Zugehörigkeitswahrscheinlichkeiten zu Typ 2, Typ 3 und Typ 4; Typ 1 (motivational-affektiv indifferente HP) wird als Referenzkategorie verwendet; Exp = Experimente; kExp = Klassenmittelwert Experimente)

Sowohl Profile als auch Unterrichtsmerkmale wurden in der Mehrebenenanalyse aus inhaltlichen und methodischen Gründen am Klassenmittelwert zentriert (*group-mean*). In den verwendeten Analysen sind grundsätzlich die Zusammenhänge auf Schüler-ebene bei Zentrierung am Stichproben (*grandmean*) und am Klassenmittelwert (*group-mean*) algebraisch ineinander überführbar und daher äquivalent (Marsh et al., 2012; Preacher et al., 2016). Nur die Zentrierung am Klassenmittelwert erlaubt jedoch, Zusammenhänge auf Schüler- und Klassenebene im Sinne von Unterrichtsklimaeffekten zu trennen (Lüdtke et al., 2008; Marsh et al., 2012). Sie ermöglicht es, Zusammenhänge auf Schülerebene unabhängig von der Klassenebene, d. h. ohne enthaltene Varianz auf Klassenebene, zu betrachten und entspricht inhaltlich der Fragestellung nach Zusammenhängen über Klassen hinweg und Zusammenhängen innerhalb der Klassen (Lüdtke et al., 2011; Marsh et al., 2012). Darüber hinaus erfordern die Analysen der Interaktionen im jeweiligen Modell III die Verwendung einer Zentrierung am Klassenmittelwert (Enders & Tofighi, 2007; Preacher et al., 2016), da eine Zentrierung am Stichprobenmittelwert statistisch Interaktionseffekte ergeben kann, obwohl sie in der Stichprobe nicht existieren (Hoffmann, Krapp et al., 1998). Die Zentrierung am Klassenmittelwert ermöglicht darüber hinaus, die (Cross-Level-)Interaktion zwischen Unterrichtsaktivität und Profilen im Zusammenhang mit der Wahrnehmung der Lehr-Lernbedingungen methodisch korrekt und ohne ungewollte Anteile der Interaktion von Unterrichtsaktivitäten und Profilen auf Klassenebene zu modellieren und entspricht dem Vorschlag von Preacher et al. (2016). Die Zentrierung am Klassenmittelwert ist auch bei bivariaten Variablen möglich und die Interpretation der Effekte entspricht unabhängig davon, ob die Variablen dummy- oder effektkodiert sind, der Interpretation bei kontinuierlichen Variablen (Enders & Tofighi, 2007).

Die Formel, mit der die abhängigen Variablen in der Mehrebenenanalyse beschrieben werden können, wird in Anhang A.2 beschrieben. Die verwendete Zentrierung am Klassenmittelwert bestimmt nachfolgend die Interpretation der Ergebnisse. Zusammenhänge auf Klassenebene geben an, inwieweit High Performer in Klassen mit einem höheren Anteil an High Performern eines Selbstkonzept-Interesse-Typs bessere bzw. schlechtere Lehr-Lernbedingungen wahrnehmen. Zusammenhänge auf Schülerebene zeigen dagegen, inwieweit innerhalb der Klassen Mikro-Lernumgebungen existieren, indem sich High Performer unterschiedlicher Selbstkonzept-Interesse-Typen innerhalb einer Klasse (unter Kontrolle des Zusammenhangs auf Klassenebene) zusätzlich in ihrer

Wahrnehmung der Lehr-Lernbedingungen unterscheiden. Zusammenhänge auf Schü-
lerebene sind daher als reine Zusammenhänge innerhalb Klassen zu interpretieren, bei
denen jeweils Abweichungen der unabhängigen Variablen zum Klassenmittelwert mit
Abweichungen der abhängigen Variable zum Klassenmittelwert in Verbindung gebracht
werden. Da beim Vergleich der Profile mehrere paarweise Vergleiche angestellt wurden,
wurde in den Analysen neben dem üblichen Signifikanzniveau von $p < .05$ jeweils ein
Bonferroni-korrigiertes Signifikanzniveau angesetzt und nicht-signifikante Werte in den
Tabellen jeweils gekennzeichnet.

Einschub: Fächer bei der Erfassung motivationsunterstützender Lehr- Lernbedingungen

Da die Skalen zu den motivationsunterstützenden Lehr-Lernbedingungen von
allen High Performern nur für ein Fach beantwortet und in den Analysen zur Ein-
schätzung der Lehr-Lernbedingungen in Naturwissenschaften allgemein zusam-
mengenommen wurden, wurde vorab untersucht, ob die Verteilung der jeweiligen
Fächer in den Profilen vergleichbar war. Tabelle 46 zeigt die Verteilung der ge-
wählten Fächer für die Beantwortung der Skalen innerhalb der Profile. Statistisch

Tabelle 46.

*Verteilung der gewählten Fächer bei der Beantwortung der Skalen zu den Lehr-Lernbedin-
gungen*

Typ	Biologie		Physik		Chemie		Naturw.	
	M (%)	(SE)	M (%)	(SE)	M (%)	(SE)	M (%)	(SE)
Motiv.-aff. indiff. HP	42.98	(4.31)	26.58	(4.72)	29.09	(4.08)	1.35	(1.21)
Interessierte HP	42.77	(3.81)	24.80	(3.71)	29.87	(4.15)	2.56	(1.65)
Leistungszuv. HP	31.94	(3.56)	32.62	(3.75)	33.68	(4.33)	1.75	(1.06)
Motiv.-aff. beteil. HP	30.05	(2.73)	30.49	(2.24)	35.30	(2.46)	4.16	(1.25)

signifikant unterschieden sich nur der Anteil der Antworten in Biologie zwischen
den interessierten und leistungszuversichtlichen High Performern (Typ 2 und 3;
 $M = 10.8\%$, $SE = 5.2\%$, $p < .05$) sowie den motivational-affektiv indifferenten
und leistungszuversichtlichen (Typ 1 und 3; $M = 11.0\%$, $SE = 5.6\%$, $p < .05$)
bzw. motivational-affektiv beteiligten High Performer (Typ 1 und 4; $M = 12.9\%$,
 $SE = 5.1\%$, $p < .05$). Insgesamt ist die Verteilung der Fächer innerhalb der Typen
jedoch in Verlauf und Höhen vergleichbar. Ein Vergleich der Lehr-Lernbedingun-

gen wurde daher als vertretbar angesehen.

Autonomieunterstützung

Die Ergebnisse für die drei Lehr-Lernbedingungen Autonomieunterstützung, Kompetenzunterstützung und inhaltliche Relevanz in Kombination mit den Unterrichtsaktivitäten „Experimentieren“ und „Schülerinnen und Schüler forschen“ sind jeweils in einer eigenen Tabelle dargestellt (Tabelle 47 bis 52). Die Tabellen 47 und 48 fassen die Ergebnisse für die wahrgenommene Autonomieunterstützung zusammen. Modell I untersuchte den Unterschied zwischen den Selbstkonzept-Interesse-Typen in der wahrgenommenen Autonomieunterstützung. Die Ergebnisse zeigen, dass sich High Performer mit unterschiedlichen Selbstkonzept-Interesse-Typen auf Klassenebene nicht in ihrer Wahrnehmung der Autonomieunterstützung unterscheiden. In Schulklassen mit höherem Anteil an interessierten, leistungszuversichtlichen und motivational-affektiv beteiligten High Performern wurde damit im Mittel ein ähnlich hohes Niveau an Autonomieunterstützung berichtet als in Schulklassen mit geringeren Anteilen bzw. mehr motivational-affektiv indifferenten High Performern. Innerhalb der Klassen dagegen berichteten interessierte, leistungszuversichtliche und motivational-affektiv indifferente High Performer signifikant mehr Autonomieunterstützung als motivational-affektiv indifferente High Performer. Untereinander unterschieden sich diese Typen jedoch nicht in ihrer wahrgenommenen Autonomieunterstützung. Insgesamt nahmen daher motivational-affektiv indifferente High Performer innerhalb einer Schulklasse ihre Lernumgebung hinsichtlich der Autonomieunterstützung ungünstiger wahr als High Performer anderer Typen. Über Schulklassen hinweg ließen sich jedoch keine Unterschiede beobachten. Die Annahme, dass sich High Performer, die sich durch Typen mit höherem Interesse auszeichnen, mehr autonomiegestützt fühlen, bestätigte sich nur für den Vergleich der beiden Typen mit niedrigem Selbstkonzept. Zwischen den beiden Typen mit hohem Selbstkonzept bestand dagegen kein signifikanter Unterschied. Die Zusammenhänge zeigten sich darüber hinaus lediglich innerhalb der Schulklassen, nicht über Schulklassen hinweg.

Modell II untersuchte, inwieweit die Häufigkeit der Schülerexperimente die Wahrnehmung der Autonomieunterstützung von High Performern auf Klassen- bzw. Schülerebene vorhersagt. Sowohl auf Klassen- als auch auf Schülerebene konnten Schülerexperimente als signifikanter Prädiktor identifiziert werden. Dies bedeutet, in Klassen

Tabelle 47.

Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interesse-Typen mit Autonomieunterstützung bei unterschiedlich häufigen Experimenten im Unterricht

	Autonomieunterstützung (AU)					
	Modell I		Modell II		Modell III	
	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)
<i>Schülerebene</i>						
Geschlecht	0.08	(0.07)			0.10	(0.07)
Profile						
2: Interessierte HP (sw2)	0.31**	(0.13)				
3: Leistungszuv. HP (sw3)	0.35**	(0.12)				
4: Motiv.-aff. beteil. HP (sw4)	0.43***	(0.10)				
Experimente (Exp.)			0.26***	(0.04)	0.25***	(0.04)
Residualvarianz AU	0.86***	(0.04)	0.83***	(0.04)	0.81***	(0.04)
Differenzen						
sw2-sw3	-0.04	(0.11)				
sw3-sw4	-0.08	(0.08)				
sw2-sw4	-0.12	(0.09)				
<i>Klassenebene</i>						
Profile						
2: Interessierte HP (sb2)	0.23	(0.55)			0.24	(0.55)
3: Leistungszuv. HP (sb3)	0.43	(0.53)			0.13	(0.54)
4: Motiv.-aff. beteil. HP (sb4)	0.79	(0.41)			0.50	(0.41)
Experimente (Exp.)			0.30***	(0.07)	0.27***	(0.08)
Cross-Level-Interactions						
Exp. x Interessierte HP (sw2)					-0.26	(0.24)
Exp. x Leistungszuv. HP (sw3)					-0.28	(0.20)
Exp. x Mot.-aff. beteil. HP (sw4)					-0.14	(0.18)
Korrelationen						
Exp. u cb2					0.00	(0.01)
Exp. u cb3					0.01	(0.01)
Exp. u cb4					0.03	(0.02)
MW						
Exp.					-0.04	(0.05)
2: Interessierte HP					0.13***	(0.01)
3: Leistungszuv. HP					0.19***	(0.01)
4: Motiv.-aff. beteil. HP					0.54***	(0.02)
Intercepts						
AU	-0.51	(0.36)	0.05	(0.04)	-0.29	(0.36)
s2					0.23	(0.12)
s3					0.28* ¹	(0.12)
s4					0.34**	(0.10)
Varianz						
AU	0.08**	(0.03)				
Exp.					0.29***	(0.05)
cb2					0.01***	(0.00)
cb3					0.02***	(0.00)
cb4					0.04***	(0.00)
Residualvarianz						
AU			0.08**	(0.02)	0.07**	(0.02)
Differenzen						
sb2-sb3	-0.21	(0.59)				
sb3-sb4	-0.36	(0.34)				
sb2-sb4	-0.57	(0.47)				
R ² Level 1	0.02***	(0.01)	0.05***	(0.01)		
R ² Level 2	0.17*	(0.15)	0.24*	(0.11)		

Anmerkung. * p<.05, ** p<.01, *** p<.001. ¹ nach Bonferroni-Korrektur nicht signifikant.

mit häufigeren Schülerexperimenten nahmen High Performer insgesamt im Sinne eines Unterrichtsklimaeffektes (vgl. Lüdtke et al., 2009; Marsh et al., 2012) ein höheres Niveau an Autonomieunterstützung wahr. Zusätzlich berichteten High Performer, die die Häufigkeit von Schülerexperimenten höher als ihre Mitschüler in der Klasse einschätzten, eine höhere Autonomieunterstützung als diese Mitschüler. Offen bleibt dabei, ob sich hinter der unterschiedlich wahrgenommenen Häufigkeit der Experimente innerhalb der Klasse eine positivere Einschätzung von High Performern mit positiven Merkmalen (vgl. Wittwer, 2008) verbirgt, High Performer einer Schulklasse tatsächlich unterschiedlich viele Gelegenheiten für Schülerexperimente bekommen (beispielsweise durch unterschiedliche Naturwissenschaftszweige innerhalb einer Schulklasse) oder in der subjektiven Einschätzung der Jugendlichen entsprechend ihren Bedürfnissen unterschiedliche Häufigkeiten von Experimenten als „wenig“ bzw. „viel“ gelten.

In Modell III wurde schließlich untersucht, ob sich der Zusammenhang von häufigeren Schülerexperimenten mit höher wahrgenommener Autonomieunterstützung zwischen High Performern unterschiedlicher Selbstkonzept-Interesse-Typen unterscheidet. Die Interaktion von Schülerexperimenten und Typen ist eine Interaktion einer Variable auf Klassenebene (Unterrichtsaktivität „Experimentieren“) mit einer Variable auf Schülerebene (Typen) im Zusammenhang mit einer Variable auf Schülerebene (Autonomieunterstützung). In der Notation von Preacher et al. (2016) wird diese Interaktion als $2 \times (1 \rightarrow 1)$ Interaktion bezeichnet. Sie wird häufig mit einem *random slope* auf Klassenebene ohne Zentrierung am Klassenmittelwert modelliert und stellt dann jedoch eine Mischung aus einer Interaktion auf Klassenebene und einer tatsächlichen Cross-Level-Interaktion dar (Preacher et al., 2016). Die Zentrierung am Klassenmittelwert erlaubt die Aufteilung in eine Cross-Level-Interaktion ohne Verfälschung durch andere Varianzanteile und einer $2 \times (2 \rightarrow 2)$ Interaktion der Varianzen der jeweiligen Variablen auf Klassenebene. Da in dieser Arbeit die Fragestellung verfolgt wird, inwieweit Mikro-Lernumgebungen für High Performer unterschiedlicher Selbstkonzept-Interesse-Typen innerhalb einer Schulklasse existieren, wird nur der Cross-Level-Interaktionsteil aus mittlerer Häufigkeit der Schülerexperimente auf Klassenebene mit Typ auf Schülerebene im Zusammenhang mit wahrgenommener Autonomieunterstützung auf Schülerebene untersucht.⁵

⁵Auch der $2 \times (2 \rightarrow 2)$ Teil der Interaktion wurde zur Kontrolle in einer weiteren Analyse modelliert. Bei allen abhängigen Variablen (AU, KU und IR) sowie allen Unterrichtsaktivitäten (Exp. und For.) ergaben sich jedoch keine signifikanten Zusammenhänge

Tabelle 48.

Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interesse-Typen mit Autonomieunterstützung bei unterschiedlich häufigen forschenden Unterrichtsaktivitäten

	Autonomieunterstützung (AU)					
	Modell I		Modell II		Modell III	
	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)
<i>Schülerebene</i>						
Geschlecht	0.08	(0.07)			0.10	(0.07)
Profile						
2: Interessierte HP (sw2)	0.31**	(0.13)				
3: Leistungszuv. HP (sw3)	0.35**	(0.12)				
4: Motiv.-aff. beteil. HP (sw4)	0.43***	(0.10)				
SuS forschen (For.)			0.23***	(0.03)	0.25***	(0.04)
Residualvarianz AU	0.86***	(0.04)	0.84***	(0.05)	0.81***	(0.04)
Differenzen						
sw2-sw3	-0.04	(0.11)				
sw3-sw4	-0.08	(0.08)				
sw2-sw4	-0.12	(0.09)				
<i>Klassenebene</i>						
Profile						
2: Interessierte HP (sb2)	0.23	(0.55)			0.24	(0.55)
3: Leistungszuv. HP (sb3)	0.43	(0.53)			0.13	(0.54)
4: Motiv.-aff. beteil. HP (sb4)	0.79	(0.41)			0.50	(0.41)
SuS forschen (For.)			0.32***	(0.09)	0.27***	(0.08)
Cross-Level-Interactions						
For. x Interessierte HP (sw2)					-0.26	(0.24)
For. x Leistungszuv. HP (sw3)					-0.28	(0.20)
For. x Mot.-aff. beteil. HP (sw4)					-0.14	(0.18)
Korrelationen						
For. u cb2					0.00	(0.01)
For. u cb3					0.01	(0.01)
For. u cb4					0.03	(0.02)
MW						
For.					-0.04	(0.05)
2: Interessierte HP					0.13***	(0.01)
3: Leistungszuv. HP					0.19***	(0.01)
4: Motiv.-aff. beteil. HP					0.54***	(0.02)
Intercepts						
AU	-0.51	(0.36)	0.05	(0.04)	-0.29	(0.36)
s2					0.23	(0.12)
s3					0.28* ¹	(0.12)
s4					0.34**	(0.10)
Varianz						
AU	0.08**	(0.03)				
For.					0.29***	(0.05)
cb2					0.01***	(0.00)
cb3					0.02***	(0.00)
cb4					0.04***	(0.00)
Residualvarianz						
AU			0.08***	(0.02)	0.07**	(0.02)
Differenzen						
sb2-sb3	-0.21	(0.59)				
sb3-sb4	-0.36	(0.34)				
sb2-sb4	-0.57	(0.47)				
R ² Level 1	0.02***	(0.01)	0.04**	(0.01)		
R ² Level 2	0.17*	(0.15)	0.20*	(0.09)		

Anmerkung. * p<.05, ** p<.01, *** p<.001. ¹ nach Bonferroni-Korrektur nicht signifikant.

Wie die Ergebnisse von Modell III (Tabelle 47) zeigen, bestand keine Interaktion zwischen Experimenten und Typen im Zusammenhang mit der wahrgenommenen Autonomieunterstützung. Das bedeutet, die gefundenen Zusammenhänge zwischen Typen und wahrgenommener Autonomieunterstützung innerhalb einer Klasse waren in ihrer Höhe unabhängig davon, wie viel Gelegenheit High Performer in der Klasse für Experimente hatten und widerlegte die angenommene Hypothese. Insgesamt nahmen High Performer der drei betrachteten Selbstkonzept-Interesse-Typen mehr Autonomieunterstützung wahr als motivational-affektiv indifferente High Performer und es gab einen positiven Zusammenhang zwischen Experimenten und Autonomieunterstützung über Klassen hinweg. Es konnte jedoch kein Einfluss der Häufigkeit von Experimenten auf die unterschiedliche Wahrnehmung der Autonomieunterstützung der Typen beobachtet werden.

Neben der Unterrichtsaktivität zu Experimenten wurden Modell II und III auch für „Schülerinnen und Schüler forschen“ – die zweite Unterrichtsaktivität, die forschend-entdeckenden Unterricht beschreibt – berechnet. Ebenso wie bei Experimenten bestätigte sich, dass in Klassen mit häufigeren forschenden Unterrichtsaktivitäten mehr Autonomieunterstützung berichtet wird und darüber hinaus High Performer, die innerhalb einer Klasse mehr forschende Elemente angeben, sich im Vergleich zu ihren Mitschülern stärker in ihrer Autonomie unterstützt fühlen. Dieser Zusammenhang blieb auch unter Berücksichtigung der Interaktion der Häufigkeit forschender Elemente auf Klassenebene mit der Profiltugehörigkeit bestehen, es zeigte sich jedoch kein Interaktionseffekt. Analog zum Ergebnis für Experimente konnte auch für den Zusammenhang von „Schülerinnen und Schüler forschen“ mit Autonomieunterstützung kein Unterschied zwischen den Profilen beobachtet werden. Auch hier wurde demnach die angenommene Interaktionshypothese widerlegt.

Inhaltliche Relevanz

Wie für die wahrgenommene Autonomieunterstützung wurden auch für die wahrgenommene inhaltliche Relevanz drei Modelle mit den Zugehörigkeitswahrscheinlichkeiten zu den Typen sowie Experimenten bzw. forschenden Elemente als Prädiktoren berechnet (Tabelle 49 und 50). Modell I zeigt, dass wie bei der Autonomieunterstützung interessierte, leistungszuversichtliche sowie motivational-affektiv beteiligte High Performer innerhalb einer Klasse den Naturwissenschaftsunterricht im Vergleich zu motivational-

affektiv indifferenten High Performern als inhaltlich relevanter erlebten. Im Gegensatz zur Autonomieunterstützung wurde hier jedoch auch ein Unterschied zwischen den Typen sowie ein Zusammenhang auf Klassenebene gefunden: Motivational-affektiv beteiligte High Performer berichteten innerhalb der Schulklassen nicht nur mehr inhaltliche Relevanz als motivational-affektiv indifferente High Performer, sondern auch mehr inhaltliche Relevanz als interessierte High Performer. Das bedeutet, High Performer mit einem Profil mit hohem Selbstkonzept und hohem Interesse unterschieden sich von Profilen mit niedrigem Selbstkonzept, nicht jedoch von Profilen mit hohem Selbstkonzept und niedrigem Interesse, in ihrer wahrgenommenen inhaltlichen Relevanz. Diese günstigere Wahrnehmung motivational-affektiv beteiligter High Performer ließ sich auch auf Klassenebene beobachten. In Klassen mit höherem Anteil an motivational-affektiv beteiligten High Performern wurde eine höhere inhaltliche Relevanz berichtet als in Klassen mit einem geringeren Anteil. Die Ergebnisse bestätigten die Annahme, dass Profile mit höherem Interesse mehr inhaltliche Relevanz wahrnehmen als die Profile mit niedrigem Selbstkonzept auf Schülerebene. Zwischen den Typen mit hohem Selbstkonzept und unterschiedlich hohem Interesse war der Unterschied jedoch nicht signifikant. Auf Klassenebene zeigte sich ein Unterschied zwischen High Performern mit niedrigem Selbstkonzept und Interesse und High Performern mit hohem Selbstkonzept und Interesse. Zusätzlich zu den angenommenen Zusammenhängen ergab die Analyse, dass die Typen mit unterschiedlich hohem Selbstkonzept bei gleichzeitig hohem Interesse sich in ihrer wahrgenommenen inhaltlichen Relevanz unterschieden. Dies deutet einen Zusammenhang zwischen inhaltlicher Relevanz und Selbstkonzept bei hohem Interesse an, der möglicherweise z. B. auf höhere Lernerfolge bei hohem Interesse zurückzuführen ist.

Modell II für die inhaltliche Relevanz zeigte darüber hinaus, dass Experimente sowohl auf Schülerebene als auch auf Klassenebene einen positiven Zusammenhang mit der wahrgenommenen inhaltlichen Relevanz aufwiesen. Häufigere Experimente gehen demnach für High Performer über Klassen hinweg mit einer höher wahrgenommenen inhaltlichen Relevanz einher und sind auch innerhalb Schulklassen mit einer günstigeren Einschätzung der inhaltlichen Relevanz assoziiert. Auch hier können individuelle Unterschiede innerhalb der Schulklasse bedeutsame Wahrnehmungsunterschiede aber auch unterschiedlich zutreffende Einschätzungen sein.

Wie Modell III zeigt, wurde analog zur Autonomieunterstützung auch der Zu-

7. Ergebnisse

Tabelle 49.

Mehrebenenanalyse zum Zusammenhang der Selbstkonzept-Interesse-Typen mit inhaltlicher Relevanz bei unterschiedlich häufigen Experimenten im Unterricht

	Inhaltliche Relevanz (IR)					
	Modell I		Modell II		Modell III	
	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)
<i>Schülerebene</i>						
Geschlecht	0.05	(0.07)			0.07	(0.06)
Profile						
2: Interessierte HP (sw2)	0.35***	(0.13)				
3: Leistungszuv. HP (sw3)	0.48***	(0.12)				
4: Motiv.-aff. beteil. HP (sw4)	0.68***	(0.11)				
Experimente (Exp.)			0.23***	(0.04)	0.21***	(0.04)
Residualvarianz IR	0.87***	(0.04)	0.88***	(0.04)	0.84***	(0.04)
Differenzen						
sw2-sw3	-0.13	(0.11)				
sw3-sw4	-0.20*	(0.09)				
sw2-sw4	-0.33***	(0.10)				
<i>Klassenebene</i>						
Profile						
2: Interessierte HP (sb2)	0.04	(0.49)			-0.01	(0.50)
3: Leistungszuv. HP (sb3)	0.59	(0.48)			0.36	(0.48)
4: Motiv.-aff. beteil. HP (sb4)	1.02*	(0.40)			0.82* ¹	(0.40)
Experimente (Exp.)			0.21***	(0.06)	0.13	(0.07)
Cross-Level-Interactions						
Exp. x Interessierte HP (sw2)					0.00	(0.23)
Exp. x Leistungszuv. HP (sw3)					0.00	(0.21)
Exp. x Mot.-aff. beteil. HP (sw4)					0.03	(0.19)
Korrelationen						
Exp. u cb2					0.00	(0.01)
Exp. u cb3					0.01	(0.01)
Exp. u cb4					0.03	(0.02)
MW						
Exp.					-0.04	(0.05)
2: Interessierte HP					0.13***	(0.01)
3: Leistungszuv. HP					0.19***	(0.01)
4: Motiv.-aff. beteil. HP					0.54***	(0.02)
Intercepts						
IR	-0.59	(0.34)	0.10**	(0.04)	-0.43	(0.34)
s2					0.35* ¹	(0.15)
s3					0.45***	(0.13)
s4					0.65***	(0.12)
Varianz						
IR	0.05*	(0.02)				
Exp.					0.29***	(0.05)
cb2					0.01***	(0.00)
cb3					0.02***	(0.00)
cb4					0.04***	(0.00)
Residualvarianz						
IR			0.07**	(0.02)	0.05* ¹	(0.02)
Differenzen						
sb2-sb3	-0.55	(0.42)				
sb3-sb4	-0.43	(0.32)				
sb2-sb4	-0.98**	(0.36)				
R ² Level 1	0.05***	(0.01)	0.04**	(0.01)		
R ² Level 2	0.38*	(0.20)	0.15	(0.09)		

Anmerkung. * p<.05, ** p<.01, *** p<.001. ¹ nach Bonferroni-Korrektur nicht signifikant.

Tabelle 50.

Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interesse-Typen mit inhaltlicher Relevanz bei unterschiedlich häufigen forschenden Unterrichtsaktivitäten

	Inhaltliche Relevanz (IR)					
	Modell I		Modell II		Modell III	
	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)
<i>Schülerebene</i>						
Geschlecht	0.05	(0.07)			0.09	(0.06)
Profile						
2: Interessierte HP (sw2)	0.35***	(0.13)				
3: Leistungszuv. HP (sw3)	0.48***	(0.12)				
4: Motiv.-aff. beteil. HP (sw4)	0.68***	(0.11)				
SuS forschen (For.)			0.23***	(0.04)	0.22***	(0.04)
Residualvarianz IR	0.87***	(0.04)	0.88***	(0.04)	0.84***	(0.04)
Differenzen						
sw2-sw3	-0.13	(0.11)				
sw3-sw4	-0.20*	(0.09)				
sw2-sw4	-0.33***	(0.10)				
<i>Klassenebene</i>						
Profile						
2: Interessierte HP (sb2)	0.04	(0.49)			-0.02	(0.48)
3: Leistungszuv. HP (sb3)	0.59	(0.48)			0.47	(0.48)
4: Motiv.-aff. beteil. HP (sb4)	1.02*	(0.40)			0.92* ¹	(0.40)
SuS forschen (For.)			0.18	(0.09)	0.07	(0.09)
Cross-Level-Interactions						
For. x Interessierte HP (sw2)					0.04	(0.34)
For. x Leistungszuv. HP (sw3)					0.17	(0.34)
For. x Mot.-aff. beteil. HP (sw4)					-0.01	(0.29)
Korrelationen						
For. u cb2					-0.01	(0.01)
For. u cb3					0.01	(0.01)
For. u cb4					0.02	(0.02)
MW						
For.					-0.03	(0.04)
2: Interessierte HP					0.13***	(0.01)
3: Leistungszuv. HP					0.19***	(0.01)
4: Motiv.-aff. beteil. HP					0.54***	(0.02)
Intercepts						
IR	-0.59	(0.34)	0.09***	(0.04)	-0.50	(0.34)
s2					0.41**	(0.14)
s3					0.46***	(0.13)
s4					0.66***	(0.11)
Varianz						
IR	0.05*	(0.02)				
For.					0.21***	(0.03)
cb2					0.01***	(0.00)
cb3					0.02***	(0.00)
cb4					0.04***	(0.00)
Residualvarianz						
IR			0.07***	(0.02)	0.05*	(0.02)
Differenzen						
sb2-sb3	-0.55	(0.42)				
sb3-sb4	-0.43	(0.32)				
sb2-sb4	-0.98**	(0.36)				
R ² Level 1	0.05***	(0.01)	0.04**	(0.01)		
R ² Level 2	0.38*	(0.20)	0.08	(0.08)		

Anmerkung. * p<.05, ** p<.01, *** p<.001. ¹ nach Bonferroni-Korrektur nicht signifikant.

sammenhang von Experimenten auf Klassenebene und wahrgenommener inhaltlicher Relevanz innerhalb der Klassen nicht durch die Zugehörigkeit zu Selbstkonzept-Interesse-Typen moderiert. Motivational-affektiv beteiligte High Performer erleben bei häufigeren Experimenten den Ergebnissen zufolge nicht mehr inhaltliche Relevanz als motivational-affektiv indifferente High Performer bei mehr Experimenten. Die Differenz der wahrgenommenen inhaltlichen Relevanz von High Performern unterschiedlicher Typen innerhalb einer Klasse war nicht höher oder niedriger ausgeprägt, wenn in der Klasse mehr Gelegenheiten für Experimente geboten wurden. Analog zur Autonomieunterstützung muss auch hier die Annahme der Interaktion verworfen werden. Die Intercepts und Steigungskoeffizienten in Modell III zeigen vielmehr, dass die Zusammenhänge von Selbstkonzept-Interesse-Typen und inhaltlicher Relevanz auch unter Kontrolle der Häufigkeit von Experimenten unverändert waren.

Auch für die inhaltliche Relevanz wurden die Zusammenhänge zusätzlich für die Häufigkeit der Unterrichtsaktivität „Schülerinnen und Schüler forschen“ betrachtet (Tabelle 50). Analog zu Experimenten standen auch häufigere forschende Unterrichtsaktivitäten sowohl auf Klassen- als auch auf Schülerebene in einem positiven Zusammenhang mit der wahrgenommenen inhaltlichen Relevanz von High Performern. Es zeigten sich jedoch – ebenfalls in Einklang mit den Ergebnissen zu Experimenten – keine signifikanten Interaktionseffekte. Die Häufigkeit forschender Elemente beeinflusste diesem Ergebnis zufolge nicht, wie inhaltlich relevant High Performer unterschiedlicher Selbstkonzept-Interesse-Typen ihren Unterricht innerhalb der Klassen erleben. Auch hier bestätigen die Steigungskoeffizienten und Intercepts, dass die Kontrolle der Häufigkeit forschender Elemente zu einem unveränderten Ergebnis führt und die Interaktionshypothese verworfen werden muss.

Kompetenzunterstützung

Bei der Kompetenzunterstützung stand vor allem der Vergleich der Typen mit hohem und niedrigem Selbstkonzept bei gleichzeitig hohem bzw. niedrigem Interesse im Fokus, d. h. ein Vergleich von motivational-affektiv indifferente und leistungszuversichtlichen High Performern. Tabelle 51 zeigt die Ergebnisse der Mehrebenenanalyse mit den motivational-affektiv indifferenten High Performern als Referenzklasse. Alle drei betrachteten motivational-affektiven Typen berichteten eine signifikant höher wahrgenommene Kompetenzunterstützung als die Referenzgruppe der motivational-affektiv indifferenten

High Performer. Auch leistungszuversichtliche High Performer nahmen damit mehr Kompetenzunterstützung wahr als motivational-affektiv indifferente High Performer. Dies bedeutet, High Performer mit ähnlichem Interesse aber höherem Selbstkonzept erleben wie in der Hypothese angenommen mehr Kompetenzunterstützung. Die Koeffizienten der betrachteten Typen unterschieden sich darüber hinaus statistisch bedeutsam: Je höher ausgeprägt das Selbstkonzept und Interesse eines motivational-affektiven Typs war, desto höher war der Unterschied in der Wahrnehmung der Kompetenzunterstützung. Leistungszuversichtliche aber uninteressierte High Performer erreichten dabei höhere Werte als interessierte aber wenig leistungszuversichtliche High Performer. Auch auf Klassenebene wurden signifikante Unterschiede gefunden. Je höher der Anteil an motivational-affektiv beteiligten High Performern einer Klasse war, desto mehr Kompetenzunterstützung berichteten die High Performer in dieser Klasse.

Die Betrachtung des Zusammenhangs von Experimenten und Kompetenzunterstützung in Modell II ergab, dass in Klassen mit häufigeren Experimenten mehr Kompetenzunterstützung berichtet wurde, d. h. die Häufigkeit von Experimenten über Klassen hinweg vorhersagte, wie kompetenzunterstützt High Performer sich fühlten. Ähnlich wie bei Autonomieunterstützung und inhaltlicher Relevanz bestand auch hier ein Zusammenhang auf Schülerebene zwischen der Abweichung der wahrgenommenen Häufigkeit von Experimenten innerhalb der Klasse mit der wahrgenommenen Kompetenzunterstützung im Vergleich zu den Mitschülern.

Auch für die Kompetenzunterstützung ergab die Berücksichtigung der Interaktionsterme in Modell III keine signifikanten Cross-Level-Interaktionen zwischen Häufigkeit von Experimenten auf Klassenebene und Selbstkonzept-Interesse-Typen. Das bedeutet, wie sehr sich High Performer unterschiedlicher Typen in ihrer wahrgenommenen Kompetenzunterstützung innerhalb einer Schulklasse unterscheiden, war – anders als in den Hypothesen vermutet – unabhängig davon, wie viele Gelegenheiten für Experimente in dieser Schulklasse bestanden. Die Kontrolle der Häufigkeit von Experimenten minderte den Zusammenhang von Typen und Kompetenzunterstützung nicht.

Für forschende Elemente als Unterrichtsaktivitäten zeigte sich im Zusammenhang mit der Kompetenzunterstützung ein ähnliches Bild (vgl. Tabelle 52). Auch forschende Elemente waren signifikanter Prädiktor der wahrgenommenen Kompetenzunterstützung von High Performern auf Schüler- und Klassenebene. Sie moderierten jedoch nicht den Zusammenhang von Typen und Kompetenzunterstützung innerhalb Klassen. Wie

7. Ergebnisse

Tabelle 51.

Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interesse-Typen mit Kompetenzunterstützung bei unterschiedlich häufigen Experimenten im Unterricht

	Kompetenzunterstützung (KU)					
	Modell I		Modell II		Modell III	
	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)
<i>Schülerebene</i>						
Geschlecht	0.02	(0.06)			0.04	(0.06)
Profile						
2: Interessierte HP (sw2)	0.41***	(0.11)				
3: Leistungszuv. HP (sw3)	0.81***	(0.11)				
4: Motiv.-aff. beteil. HP (sw4)	1.05***	(0.09)				
Experimente (Exp.)			0.20***	(0.04)	0.17***	(0.03)
Residualvarianz KU	0.73***	(0.03)	0.82***	(0.04)	0.71***	(0.03)
Differenzen						
sw2-sw3	-0.40***	(0.10)				
sw3-sw4	-0.24***	(0.07)				
sw2-sw4	-0.64***	(0.09)				
<i>Klassenebene</i>						
Profile						
2: Interessierte HP (sb2)	0.08	(0.45)			0.10	(0.44)
3: Leistungszuv. HP (sb3)	0.36	(0.36)			0.13	(0.34)
4: Motiv.-aff. beteil. HP (sb4)	1.08***	(0.27)			0.85**	(0.25)
Experimente (Exp.)			0.29***	(0.06)	0.22***	(0.06)
Cross-Level-Interactions						
Exp. x Interessierte HP (sw2)					-0.20	(0.21)
Exp. x Leistungszuv. HP (sw3)					-0.13	(0.17)
Exp. x Mot.-aff. beteil. HP (sw4)					-0.01	(0.15)
Korrelationen						
Exp. u cb2					0.00	(0.01)
Exp. u cb3					0.01	(0.01)
Exp. u cb4					0.03	(0.02)
MW						
Exp.					-0.04	(0.05)
2: Interessierte HP					0.13***	(0.01)
3: Leistungszuv. HP					0.19***	(0.01)
4: Motiv.-aff. beteil. HP					0.54***	(0.02)
Intercepts						
KU	-0.50*	(0.25)	0.18***	(0.03)	-0.33	(0.23)
s2					0.38***	(0.13)
s3					0.79***	(0.11)
s4					1.01***	(0.09)
Varianz						
KU	0.05**	(0.02)				
Exp.					0.29***	(0.05)
cb2					0.01***	(0.00)
cb3					0.02***	(0.00)
cb4					0.04***	(0.00)
Residualvarianz						
KU			0.05**	(0.02)	0.04*	(0.02)
Differenzen						
sb2-sb3	-0.28	(0.49)				
sb3-sb4	-0.72**	(0.28)				
sb2-sb4	-1.00**	(0.34)				
R ² Level 1	0.13***	(0.02)	0.03**	(0.01)		
R ² Level 2	0.43*	(0.15)	0.33*	(0.14)		

Anmerkung. * $p < .05$, ** $p < .01$, *** $p < .001$. ¹ nach Bonferroni-Korrektur nicht signifikant.

Tabelle 52.

Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interesse-Typen mit Kompetenzunterstützung bei unterschiedlich häufigen forschenden Unterrichtsaktivitäten

	Kompetenzunterstützung (KU)					
	Modell I		Modell II		Modell III	
	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)
<i>Schülerebene</i>						
Geschlecht	0.02	(0.06)			0.04	(0.06)
Profile						
2: Interessierte HP (sw2)	0.41***	(0.11)				
3: Leistungszuv. HP (sw3)	0.81***	(0.11)				
4: Motiv.-aff. beteil. HP (sw4)	1.05***	(0.09)				
SuS forschen (For.)			0.15***	(0.04)	0.11**	(0.04)
Residualvarianz KU	0.73***	(0.03)	0.83***	(0.04)	0.72***	(0.03)
Differenzen						
sw2-sw3	-0.40***	(0.10)				
sw3-sw4	-0.24***	(0.07)				
sw2-sw4	-0.64***	(0.09)				
<i>Klassenebene</i>						
Profile						
2: Interessierte HP (sb2)	0.08	(0.45)			0.13	(0.45)
3: Leistungszuv. HP (sb3)	0.36	(0.36)			0.22	(0.34)
4: Motiv.-aff. beteil. HP (sb4)	1.08***	(0.27)			0.91***	(0.29)
SuS forschen (For.)			0.30***	(0.08)	0.20**	(0.08)
Cross-Level-Interactions						
For. x Interessierte HP (sw2)					-0.31	(0.32)
For. x Leistungszuv. HP (sw3)					-0.17	(0.26)
For. x Mot.-aff. beteil. HP (sw4)					-0.11	(0.24)
Korrelationen						
For. u cb2					-0.01	(0.01)
For. u cb3					0.01	(0.01)
For. u cb4					0.02	(0.02)
MW						
For.					-0.03	(0.04)
2: Interessierte HP					0.13***	(0.01)
3: Leistungszuv. HP					0.19***	(0.01)
4: Motiv.-aff. beteil. HP					0.54***	(0.02)
Intercepts						
KU	-0.50*	(0.25)	0.19***	(0.03)	-0.38	(0.25)
s2					0.39***	(0.13)
s3					0.78***	(0.11)
s4					1.02***	(0.09)
Varianz						
KU	0.05**	(0.02)	0.052***	(0.02)		
For.					0.21***	(0.03)
cb2					0.01***	(0.00)
cb3					0.02***	(0.00)
cb4					0.04***	(0.00)
Residualvarianz						
KU					0.04**	(0.02)
Differenzen						
sb2-sb3	-0.28	(0.49)				
sb3-sb4	-0.72**	(0.28)				
sb2-sb4	-1.00**	(0.34)				
R ² Level 1	0.13***	(0.02)	0.02*	(0.01)		
R ² Level 2	0.43*	(0.15)	0.26	(0.13)		

Anmerkung. * p<.05, ** p<.01, *** p<.001. ¹ nach Bonferroni-Korrektur nicht signifikant.

bei Experimenten zeigte Modell III für „Schülerinnen und Schüler forschen“, dass der Zusammenhang von Typzugehörigkeit und Kompetenzunterstützung auf Klassenebene auch unter Kontrolle der Häufigkeit forschender Elemente bestehen blieb und sich nicht statistisch bedeutsam verringerte.

7.3.4. Zusammenfassung

Leitende Fragestellungen dieses Abschnitts waren, welche Selbstkonzept-Interesse-Typen unter den High Performern identifiziert werden können und, ob sich Mädchen und Jungen in Form und Häufigkeit der Typen unterscheiden. Darüber hinaus wurde untersucht, welche Merkmale die Typen aufweisen und wie motivationsunterstützend High Performer unterschiedlicher Profile ihren Unterricht und insbesondere häufigeres forschend-entdeckendes Lernen wahrnehmen. Insgesamt zeigte sich, dass vier verschiedene Selbstkonzept-Interesse-Typen (motivational-affektiv indifferent, interessiert, leistungsversichtlich und motivational-affektiv beteiligt) unter den High Performern vorhanden waren und sich zwischen Mädchen und Jungen in der Häufigkeit, nicht jedoch im Verlauf unterschieden. Jungen gehörten häufiger einem Typ mit hohem Selbstkonzept und/oder Interesse und seltener einem Typ mit niedrigem Selbstkonzept und/oder Interesse an als Mädchen.

Die Untersuchung der Merkmale der Typen ergab, dass innerhalb der Typen vergleichbar häufig Überschneidungen zwischen High Performance in Naturwissenschaften und Mathematik bzw. Lesen auftraten und es keine auffälligen Unterschiede zur Gesamtstichprobe gab. Bei den Notenprofilen wurden im Gegensatz zu den Kompetenzen deutlichere Unterschiede zwischen den Selbstkonzept-Interesse-Typen ausgemacht. High Performer mit hohem Selbstkonzept und Interesse wiesen häufiger ein naturwissenschaftliches Notenprofil auf, High Performer mit niedrigem Selbstkonzept und Interesse häufiger ein sprachliches Notenprofil. Der höchste Anteil an High Performern mit einem ausgeglichenen Notenprofil mit überall guten Noten fand sich beim Selbstkonzept-Interesse-Typ der motivational-affektiv beteiligten High Performer. Es gab keine Hinweise auf signifikante Unterschiede im sozioökonomischen Hintergrund zwischen High Performern unterschiedlicher Selbstkonzept-Interesse-Typen. Auch in der Verteilung der Typen auf Schulklassen wurden keine auffälligen Häufungen gefunden.

Die Analyse der Unterrichtswahrnehmung zeigte, dass motivational-affektiv beteiligte High Performer im Vergleich zu motivational-affektiv indifferenten High Performern sowohl eine höhere Autonomie- und Kompetenzunterstützung als auch inhaltliche Relevanz berichteten. High Performer mit Typen mit unterschiedlich hohem Interesse bei gleichzeitig niedrigem Selbstkonzept unterschieden sich darüber hinaus auf Schüler-ebene in der wahrgenommenen Autonomieunterstützung und inhaltlichen Relevanz. Ein solcher Zusammenhang zwischen Interesse und Lehr-Lernbedingungen konnte bei

Typen mit hohem Selbstkonzept dagegen nicht beobachtet werden. Typen mit unterschiedlich hohem Selbstkonzept nahmen unabhängig vom Interessensniveau wie erwartet unterschiedlich viel Kompetenzunterstützung wahr. Zusätzlich unterschieden sie sich bei niedrigem Interesse jedoch auch in der wahrgenommenen inhaltlichen Relevanz, was zusätzlich zur Kompetenzunterstützung auf einen Zusammenhang von Selbstkonzept und inhaltlicher Relevanz bei den Profilen hinwies.

Bei allen betrachteten Lehr-Lernbedingungen wurden keine Interaktionseffekte durch häufigere Experimente oder forschende Elemente im Naturwissenschaftsunterricht gefunden. Gleichzeitig blieben sämtliche Zusammenhänge zwischen den Selbstkonzept-Interesse-Typen und den Lehr-Lernbedingungen unter Kontrolle der Häufigkeit der Unterrichtsaktivitäten „Experimentieren“ und „Schülerinnen und Schüler forschen“ vergleichbar. Darüber hinaus zeigte sich, dass die Häufigkeit von „Experimentieren“ und „Schülerinnen und Schüler forschen“ vorhersagten, welche mittlere Autonomieunterstützung, Kompetenzunterstützung und inhaltliche Relevanz von High Performern einer Klasse wahrgenommen wurde. Insgesamt implizieren die Analysen daher, dass Experimente und forschende Aktivitäten über Klassen hinweg vorhersagen, auf welchem mittleren Niveau sich alle High Performer der Klasse unterstützt fühlen und häufigere Experimente und forschende Aktivitäten jeweils mit einem höheren Wert verbunden sind. Gleichzeitig vergrößern oder verkleinern häufigere Experimente oder forschend-entdeckende Aktivitäten die Differenzen in der Unterrichtswahrnehmung zwischen High Performern unterschiedlicher Selbstkonzept-Interesse-Typen innerhalb einer Klasse jedoch nicht. Das bedeutet, aus häufiger berichteten Experimenten folgt nicht, dass sich ein Teil motivational-affektiv innerhalb einer Klasse noch unterstützter fühlt und ein anderer noch weniger, häufigere Experimente führen jedoch auch nicht dazu, dass sich unterschiedliche Typen in ihrer Wahrnehmung motivationsunterstützender Bedingungen innerhalb der Klasse annähern.⁶ Innerhalb der Klassen können den Typen – auch unter Berücksichtigung der Unterrichtsaktivitäten in der Klasse – damit unterschiedliche Mikro-Lernumgebungen zugeordnet werden. Wie sehr sich diese Mikro-Lernumgebungen in den wahrgenommenen motivationsunterstützenden Lehr-Lernbedingungen unterscheiden, ist dabei unabhängig von der Häufigkeit von Experimenten und forschend-entdeckenden Elementen, nicht jedoch das mittlere wahrgenommene Niveau der Lehr-

⁶Wie zusätzliche Interaktionsanalysen von Typen x Unterrichtsaktivitäten auf Klassenebene zeigten, ist der Zusammenhang von Typen und wahrgenommenen Lehr-Lernbedingungen auch auf Klassenebene nicht von der Häufigkeit der Unterrichtsaktivitäten abhängig.

Lernbedingungen der Typen insgesamt. Dieses Niveau war umso höher ausgeprägt, umso häufiger Experimente und forschende Aktivitäten in einer Klasse vorkamen.

8. Diskussion

Ziel dieser Arbeit war es, leistungsstarke Jugendliche in Naturwissenschaften zu charakterisieren und dabei herauszufinden, was diese Schülerinnen und Schüler sowie ihr Lernumfeld auszeichnet. Als methodisches Vorgehen wurden dafür Sekundäranalysen der Daten aus PISA 2006 und PISA 2012 gewählt. Damit wurde die Gelegenheit genutzt, eine Subgruppe der Schülerinnen und Schüler in Deutschland mithilfe eines internationalen Large-Scale-Datensatz systematisch zu untersuchen und sowohl individuelle und familiäre als auch institutionelle und unterrichtsbezogene Merkmale in den Blick zu nehmen. Die Arbeit verfolgte damit die Absicht, bestehende Forschungsergebnisse zu leistungsstarken Jugendlichen in Naturwissenschaften zu ergänzen, sowie Stärken und Problemlagen bei der Förderung leistungsstarker Schülerinnen und Schüler in Deutschland zu identifizieren und daraus Hinweise für die pädagogische und politische Praxis abzuleiten.

In diesem Kapitel werden die Ergebnisse inhaltlich, theoretisch und methodisch diskutiert. Im ersten Abschnitt (Kapitel 8.1) werden methodische Limitationen zur Anlage der Studie sowie dem analytischen Vorgehen dargelegt. Im Anschluss daran werden die zentralen Befunde der Arbeit in Kapitel 8.2 zusammengefasst und unter Einbezug ihres theoretischen Hintergrunds diskutiert. In Kapitel 8.3 werden daraufhin Schlussfolgerungen aus den Ergebnissen dieser Arbeit für die pädagogische und politische Praxis gezogen. Abschließend befasst sich Kapitel 8.4 mit offenen Punkten und gibt einen Ausblick auf weiterführende Forschungsfragen.

8.1. Methodische Überlegungen

Als Sekundäranalyse von Daten aus PISA 2006 und PISA 2012 profitierte diese Arbeit von verschiedenen Aspekten eines Large-Scale-Datensatzes wie der großen Stichprobe, der Repräsentativität und den erprobten Erhebungsinstrumenten. Gleichzeitig ergab sich

durch die Verwendung eines Large-Scale-Datensatzes einige systematische methodische Einschränkungen. Nachfolgend werden methodische Überlegungen zu diesen durch den Datensatz bedingten Einschränkungen sowie Limitationen der hier verwendeten Unterrichtserfassung angestellt.

Die Untersuchungen der Arbeit sind auf die im Datensatz enthaltenen Variablen sowie einen Messzeitpunkt festgelegt. Die Ergebnisse liefern daher beschreibendes Wissen zu Merkmalsausprägungen, die leistungsstarke Jugendliche von anderen unterscheiden, sowie zu Zusammenhängen verschiedener Merkmale. Auf Basis diese beschreibenden Wissens werden Stärken und Problemfelder im Kontext leistungsstarker Jugendlicher in Naturwissenschaften identifiziert. Die Ergebnisse können jedoch nicht als Erklärungs- oder Veränderungswissen für kausale Zusammenhänge interpretiert werden, wie z. B. den Einfluss von forschend-entdeckendem Unterricht auf Lernprozesse leistungsstarker Jugendlicher zu bestimmen.

Das querschnittliche Design bedingt auch, dass sich Fragebogenmerkmale und naturwissenschaftliche Kompetenz zum Teil auf einen unterschiedlichen zeitlichen Bezugsrahmen beziehen. Während Schülerinnen und Schüler bei der Unterrichtswahrnehmung nach dem Unterricht des aktuellen Schuljahres gefragt werden, verbirgt sich hinter dem Kompetenzwert das Ergebnis eines kumulativen Lernprozesses über mehrere Schuljahre hinweg und es werden nicht nur die Lernprozesse der neunten Klasse abgebildet. Dies schränkt die Analysen nicht ein, muss jedoch in der Interpretation der Zusammenhänge von Unterricht und Bildungsergebnissen berücksichtigt werden.

Bei der Interpretation der Ergebnisse zu den Unterrichtsaktivitäten müssen zwei weitere methodische Aspekte der Erfassung der Unterrichtsaktivitäten einschränkend beachtet werden: (a) die Messung von Unterricht aus Schülerperspektive und (b) die Erfassung von Häufigkeiten. Die Erfassung des Unterrichts aus Schülersicht hat zunächst den Vorteil, dass abgebildet werden kann, wie Schülerinnen und Schüler den Unterricht einschätzen und erleben. Ausgehend von Angebots-Nutzungs-Modellen bestimmten diese Wahrnehmung der Schülerinnen und Schüler und daraus resultierende vermittelnde Prozesse (z. B. Lernmotivation und Lernaktivitäten), welche Lernergebnisse das Unterrichtsangebot bei den Schülerinnen und Schüler erzielt (Helmke, 2007; Seidel & Prenzel, 2006b; Seidel & Shavelson, 2007). Für Unterrichtsmerkmale wie die motivationsunterstützenden Lehr-Lernbedingungen, die abbilden, wie Jugendliche das Unterrichtsangebot individuell wahrnehmen, ist diese Erfassung über Schülereinschätzungen daher gut ge-

eignet. Bei den Unterrichtsaktivitäten steht jedoch vor allem die tatsächlich vorhandene Häufigkeit im Vordergrund. Hier besteht bei der Unterrichtserfassung aus Schülersicht der Nachteil, dass die Wahrnehmung der Schülerinnen und Schüler nicht zwingend einer objektiven Einschätzung aus der Beobachterperspektive entspricht (Lüdtke et al., 2009; Wittwer, 2008).

Die Werte der Unterrichtsaktivitäten wurden in dieser Arbeit daher auf Klassenebene aggregiert und so die geteilte Wahrnehmung einer ganzen Schulklasse und die individuellen Abweichungen der einzelnen Schüler von diesem Wert getrennt betrachtet. Dabei wurde sichergestellt, dass auf Klassenebene sowohl leistungsstarke als auch andere Jugendliche berücksichtigt wurden. Dennoch weicht auch die geteilte Klassenwahrnehmung unter Umständen von der Einschätzung eines Beobachters ab und ihr Mehrwert liegt vor allem darin, dass sie die Perspektive der Schülerinnen und Schüler abbildet. Für die Unterrichtseinschätzungen auf Individualebene, die in diesem Vorgehen individuelle Abweichungen vom Klassenmittelwert angeben, sind im Fall der Unterrichtsaktivitäten mehrere Interpretationen möglich: Einerseits kann die unterschiedliche Unterrichtswahrnehmung innerhalb der High bzw. Nicht-High-Performer einer Klasse eine subjektiv positivere Beurteilung des Unterrichtsgeschehens darstellen, die auf höhere Leistungen bzw. höheres Interesse zurückzuführen ist (vgl. Wittwer, 2008). Dies würde in der Unterrichtsklimaforschung als Unreliabilität der Messung der Unterrichtsklimavariablen gesehen (Lüdtke et al., 2009; Marsh et al., 2011). Andererseits kann es jedoch auch als Indikator für das unterschiedliche Erleben und Einschätzen der einzelnen Schülerinnen und Schüler gedeutet werden, die beispielsweise unterschiedlich viele Experimente entsprechend ihrer individuellen Bedürfnisse als „häufig“ bewerten.

Ein zweiter methodischer Aspekt, der bei der Interpretation der Ergebnisse zu Unterrichtsaktivitäten beachtet werden muss, liegt darin, dass die Unterrichtsaktivitäten als Häufigkeiten erfasst wurden. Die Unterrichtsaktivitäten „interaktives Lehren und Lernen“, „Experimentieren“, „Schülerinnen und Schüler forschen“ und „Anwendungsbezüge“ wurden mit Skalen im Schülerfragebogen gemessen, bei denen Schülerinnen und Schüler einschätzten, wie oft die angesprochenen Unterrichtsmerkmale in ihrem Naturwissenschaftsunterricht im letzten Schuljahr vorkamen. Die Antwortskalen reichten von „nie oder fast nie“ bis hin zu „in allen Stunden“. Die Zusammenhänge der Skalen zu Unterrichtsaktivitäten und anderen Variablen beschreiben daher keine Zusammenhänge mit einer Unterrichtsaktivität an sich, sondern Zusammenhänge mit dem aus

Schülersicht häufigeren Vorkommen einer Unterrichtsaktivitäten.

8.2. Diskussion zentraler Befunde

Nachdem das Vorgehen in dieser Arbeit methodisch diskutiert wurde, werden die Befunde der Arbeit im nachfolgenden Abschnitt zusammengefasst und in den Forschungsstand eingebettet. Inhaltliche Ausgangslage dieser Arbeit waren die Ergebnisse aus PISA 2006 und PISA 2012, dass Deutschland im OECD-Vergleich zwar einen überdurchschnittlichen Anteil an leistungsstarken Jugendlichen erreichte, der Vergleich mit Staaten wie Neuseeland oder Finnland jedoch zeigt, dass Deutschland hier sein Potential noch nicht ausschöpft (Schiepe-Tiska et al., 2013). Die Ergebnisse in PISA verdeutlichten darüber hinaus, dass ein großer Teil leistungsstarker Jugendlicher sich in Relation zur Leistung nur gering für Naturwissenschaften interessierte. Dies warf die grundlegende Frage auf, durch welche Merkmale sich leistungsstarke Jugendliche in Naturwissenschaften und ihre Lernumgebungen auszeichnen. Ausgehend von dieser Problemstellung hatte diese Arbeit in drei Fragestellungsbereichen das Ziel (a) leistungsstarke Jugendliche in Naturwissenschaften theoriegeleitet in leistungsrelevanten individuellen, familiären und institutionellen Merkmalen zu charakterisieren, (b) ausgehend von Ergebnissen der Lehr-Lernforschung zu untersuchen, welche Zusammenhänge zwischen Merkmalen ihres Naturwissenschaftsunterrichts und ihrer naturwissenschaftlicher Kompetenz sowie ihrem Interesse bestehen sowie (c) herauszufinden, welche motivational-affektiven Typen innerhalb der leistungsstarken Jugendlichen identifiziert werden können und, ob diese Typen ihren Unterricht unterschiedlich motivationsunterstützend wahrnehmen. Diese Struktur mit drei thematischen Abschnitten wird auch in der nachfolgenden Zusammenfassung und Diskussion zentraler Ergebnisse der Arbeit beibehalten.

8.2.1. Individuelle, familiäre und institutionelle Merkmale leistungsstarker Jugendlicher in Naturwissenschaften

Ausgehend von theoretischen Modellen für Merkmale schulischer und außerschulischer Lernumwelten (Doll & Prenzel, 2004b) wurde im ersten Teil der Fragestellungen untersucht, welche Merkmale leistungsstarke Jugendliche auszeichnen. Individuelle, familiäre und institutionelle Merkmale können Prädiktoren hoher Kompetenz in Naturwissenschaften sein. Im Kontext hoher Leistungen in Naturwissenschaften sind sie von Bedeutung, weil sie einen Eindruck davon geben, wie sich leistungsstarke Schülerinnen und Schüler von anderen Jugendlichen unterscheiden und wie heterogen die Gruppe

leistungsstarker Jugendlicher selbst in diesen Merkmalen sind. Wissen dazu zeigt, an welchen Merkmalen leistungsstarke Schülerinnen und Schüler in Naturwissenschaften erkannt bzw. nicht erkannt werden können. Es verdeutlicht darüber hinaus, welche unveränderlichen aber auch veränderlichen Merkmale möglicherweise eine Rolle für hohe Leistung in Naturwissenschaften spielen und in welchen Bereichen Fördermaßnahmen ansetzen können.

Zur Beantwortung dieser Fragestellung wurden individuelle, familiäre und institutionelle Merkmale, die aus theoretischer Sicht mögliche Bedingungsfaktoren für hohe Leistung in Naturwissenschaften sind, untersucht (vgl. Kapitel 7.1). In deskriptiven Analysen mit Mittelwertsvergleichen wurden High mit Nicht-High-Performern verglichen sowie Unterschiede zwischen PISA 2006 und PISA 2012 betrachtet. Für ausgewählte Merkmale wurden die Top Performer als Teilmenge der High Performer und die übrigen High Performern gegenübergestellt.

Individuelle Merkmale

Zu den individuellen Merkmalen, die leistungsstarke Jugendliche in Naturwissenschaften möglicherweise auszeichnen, zählt die Leistung in anderen Fachbereichen wie Mathematik und Lesen bzw. Deutsch. Fähigkeiten in Naturwissenschaften und Mathematik sowie in Naturwissenschaften und Deutsch besitzen gemeinsame Elemente wie Modellieren oder Textverstehen, in denen sie sich überschneiden (Cano et al., 2014; Uhden et al., 2012). Ausgehend davon stellt sich die Frage, inwieweit leistungsstarke Jugendliche in Naturwissenschaften auch in Mathematik oder Deutsch leistungsstärker als andere Jugendliche sind oder ebenfalls zur Spitzengruppe Deutschlands gehören. Auch inwieweit sich dieses Ergebnisse bei den beiden Leistungsmaßen Kompetenzen und Noten zeigt, ist dabei von Interesse für die Identifikation und Förderung leistungsstarker Jugendlicher.

Die Ergebnisse zeigen, dass High Performer in Naturwissenschaften im Vergleich zu Nicht-High-Performern höhere Kompetenzen in Mathematik und Lesen aufwiesen und der Großteil der High Performer in Naturwissenschaften auch in Mathematik oder Lesen bzw. beiden Fächern zu den High Performern gehörte. Auch bei den Schulnoten erreichten High Performer im Vergleich zu den restlichen Schülerinnen und Schüler bessere Werte, sowohl in Naturwissenschaften als auch in Mathematik und Lesen. Es traten dabei allerdings sehr unterschiedliche Notenprofile auf. Nur ein geringer Teil der High Performer besaß ein naturwissenschaftliches Notenprofil, ein großer Teil erzielte

in allen drei Fächern ähnlich gute Noten. Gleichzeitig wurde jedoch auch sichtbar, dass über ein Fünftel der High Performer lediglich die Note „ausreichend“ oder schlechter in Naturwissenschaften erhalten hatte. Sowohl den Ergebnissen zu Kompetenzen als auch zu Noten zufolge ist High Performance in Naturwissenschaften keine isolierte hohe Kompetenz in einer Domäne, sondern geht häufig mit hohen Kompetenzen in Mathematik und Lesen einher.

Aus den Ergebnissen wird deutlich, dass High Performance in Naturwissenschaften nur selten eine Art „Inselbegabung“ darstellt, sondern umgekehrt sehr häufig auch mit hohen Kompetenzen in Mathematik und Lesen einhergeht. Dies bestätigt vorliegende Ergebnisse zu den Top Performern (Jugendliche auf den beiden obersten Kompetenzstufen) unter den 15-Jährigen in PISA. Dort erreichte ebenfalls ein hoher Anteil in mindestens einer der anderen beiden Domänen die beiden obersten Kompetenzstufen (OECD, 2009b). Die Befunde befinden sich auch in Einklang mit der Annahme, dass die naturwissenschaftliche und mathematische Kompetenz sowie die Lesekompetenz einen gewissen Überlappungsbereich in Anforderungen aufweisen bzw. auf gemeinsamen Wissens-elementen und Prozessen beruhen (vgl. Kapitel 2.2). Damit ergänzt die Arbeit Forschung zum Zusammenhang von Mathematik und Naturwissenschaften (Hudson & McIntire, 1977) sowie Lesen und Naturwissenschaften (Cano et al., 2014) und bestätigt die Annahmen des naturwissenschaftlichen Rahmenkonzepts in PISA (Bybee et al., 2009). Mit den Ergebnissen zur Überschneidung von Naturwissenschaften und Lesen reiht sich diese Arbeit darüber hinaus in zunehmend häufiger durchgeführten Untersuchungen zur Rolle von Textverständnis und Diskussionen in Naturwissenschaften (Osborne, 2010; Osborne, Simon, Christodoulou, Howell-Richardson & Richardson, 2013) ein.

Auch das Ergebnis, dass sich Schulnoten und Kompetenzen von High Performern nicht decken, war aus theoretischer Sicht zu erwarten, da Zeugnisnoten auch mündliche Leistungen und das Verhalten im Unterricht abbilden, sich an Inhalten des Lehrplans orientieren und diese eher kurzfristig prüfen. Schulnoten werden darüber hinaus häufig an der Form einer Normalverteilung orientiert (*grading on a curve*) und sind daher zwischen Klassen und Schulen mit unterschiedlichen Leistungsniveaus nur bedingt vergleichbar (Ingenkamp, 1995; Ingenkamp & Lissmann, 2008). Mit dem Maß der naturwissenschaftlichen Kompetenz wurden dagegen weniger lehrplanbezogene und auf ein Schuljahr beschränkte, sondern konsolidierte, anwendbare und über mehrere Jahre hinweg erworbene Inhalte erfasst. Überraschend ist an den Ergebnissen dennoch,

dass ein beträchtlicher Teil der besten 20% der deutschen Schülerinnen und Schüler Naturwissenschaftsnoten im Bereich „ausreichend“ und schlechter, d. h. auch im versetzungsgefährdeten Bereich erzielte. Dies zeigt, dass eine hohe naturwissenschaftliche Kompetenz und das damit verbundene hohe Potential für weitere Lernprozesse in Naturwissenschaften durch Schulnoten nur bedingt abgebildet wird. Damit machen Schulnoten als wichtigstes Kommunikationsinstrument für Schulleistungen naturwissenschaftliche Stärken für Schülerinnen und Schüler jedoch nicht sichtbar und es besteht die Gefahr, dass Jugendliche ihre Fähigkeiten in Naturwissenschaften unterschätzen und seltener naturwissenschaftliche Bildungsgänge oder Berufe wählen (vgl. Taskinen et al., 2008).

Als ein weiteres individuelles Merkmal steht die Intelligenz in einem Zusammenhang zur Schulleistung. Da die Intelligenz ein relativ stabiles individuelles Merkmal ist (Asendorpf & Neyer, 2012), würde eine hohe Varianzaufklärung von hoher Leistung in Naturwissenschaften durch die Intelligenz bedeuten, dass die Entwicklung hoher Leistung in Naturwissenschaften nur bedingt beeinflusst und gefördert werden kann. Die Untersuchung der Subskale „Figurale Analogien“ des Kognitive Fähigkeitstests von Heller und Perleth (2000) als Proxy für Intelligenz ergab jedoch, dass hohe kognitive Fähigkeiten nicht zwingend mit einer hohen Leistung in Naturwissenschaften einhergehen. Umgekehrt erreichten auch Jugendliche mit mittleren kognitiven Fähigkeiten High Performance. Alle leistungsstarken Jugendlichen zeichneten sich jedoch durch ein mindestens durchschnittliches kognitives Fähigkeitsniveau aus. Dieses Ergebnis entspricht im Wesentlichen den Annahmen der Expertiseansatzes für hohe fachspezifische Leistungen (Bloom & Sosniak, 1985; Ericsson, 1996). Die Expertiseforschung legt nahe, dass für hohe fachspezifische Leistungen zwar ein gewisses kognitives Niveau erreicht werden muss (Schwellenhypothese), darüber hinaus jedoch die Gelegenheiten für gezieltes Üben und Feedback bestimmen, ob jemand hohe Leistungen in einem speziellen Bereich erlangt (Ericsson, 1996). Die Befunde zu leistungsstarken Jugendlichen in Naturwissenschaften sind darüber hinaus in Einklang mit Untersuchungen zur Schulleistung, die belegen, dass nur rund ein Viertel der Leistungsvarianz durch die Intelligenz erklärt wird (Schrader & Helmke, 2008, S. 291).

Die Arbeit zeigt damit in Ergänzung zu den Ergebnissen der Expertiseforschung, dass auch für hohe Leistung in Naturwissenschaften zwar ein gewisses Niveau an kognitiven Fähigkeiten notwendig ist. Ob Schülerinnen und Schüler mit einer kognitiven Fähigkeit über diesem Schwellenwert High Performance in Naturwissenschaften errei-

chen, wird jedoch nicht von der Höhe ihrer kognitiven Fähigkeiten, sondern anderen Faktoren bestimmt. Die kognitive Grundfähigkeit als ein relativ stabiles Merkmal spielt daher zwar eine nicht zu vernachlässigende aber begrenzte Rolle für hohe Leistungen in Naturwissenschaften und lässt Raum für mögliche andere variable Einflussmerkmale, mit deren Veränderung hohe Leistung in Naturwissenschaften begünstigt werden kann. Ausgehend von der Expertiseforschung kann eine planvolle, früh beginnende Auseinandersetzung mit naturwissenschaftlichen Inhalten als ein möglicher variabler Einflussfaktor für die Erlangung von High Performance in Naturwissenschaften vermutet werden (W. Schneider, 1997; W. Schneider & Stumpf, 2007).

Abschließend wurde das Geschlecht als individuelles Merkmal leistungsstarker Schülerinnen und Schüler betrachtet. Während sich in PISA in der Gesamtstichprobe der 15-Jährigen keine Geschlechterunterschiede gezeigt hatten (Schiepe-Tiska et al., 2013), wurden in einer Betrachtung der Jugendlichen auf den beiden obersten Kompetenzstufen in PISA Unterschiede zugunsten der Jungen gefunden (OECD, 2009b). Für High Performer als breiter definierte Gruppe leistungsstarker Schülerinnen und Schüler stellte sich daher die Frage, ob sie bei Geschlechterunterschieden mit der Gesamtstichprobe oder der engeren Leistungsspitze der Jugendlichen auf den beiden obersten Kompetenzstufen vergleichbar ist. In den Analysen zeigten sich signifikante Geschlechterunterschiede, indem bei männlichen High Performern eine höhere naturwissenschaftliche Kompetenz gefunden wurde. Damit unterschieden sich die High Performer in den Geschlechterunterschieden von den 15-Jährigen in PISA 2006 und 2012, bei denen Jungen und Mädchen im Mittel vergleichbare Kompetenzen erreichten (Prenzel et al., 2007; Schiepe-Tiska et al., 2013). Dies differenziert die bestehende Forschungslage zu Geschlechterdifferenzen in Naturwissenschaften in PISA dahingehend, dass Mädchen und Jungen sich zwar generell nicht in ihrer naturwissenschaftlichen Kompetenz unterscheiden, Mädchen jedoch seltener die Leistungsspitze erreichen. Einschränkend muss jedoch beachtet werden, dass die Stichprobe der 15-Jährigen nicht nur Neuntklässler, sondern auch Schülerinnen und Schüler der zehnten Klasse enthielt, und deshalb nicht uneingeschränkt vergleichbar ist. Die Ergebnisse deutet damit insgesamt an, dass Geschlechterunterschiede in der Gesamtstichprobe zwar kein Problemfeld mehr sind, für die Leistungsspitze das Potential der Schülerinnen jedoch noch nicht ausgenutzt wurde. Der Anteil leistungsstarker Jugendlicher könnte deshalb möglicherweise auch dadurch erhöht werden, dass mehr Mädchen hin zu hoher Leistung in Naturwissenschaften gefördert werden.

Familiäre Merkmale

Familiäre Merkmale wie finanzielle aber auch die kulturelle und allgemein bildungsrelevante Ressourcen und Einstellungen beeinflussen die Chancen und Impulse, die ein Kind für seine Leistungsentwicklung bekommt (Baumert & Maaz, 2006; Baumert, Stanat & Watermann, 2006a). Aus internationalen Untersuchungen der Top Performer in Naturwissenschaft (OECD, 2009b) sowie allgemein hochkompetenter Jugendlicher (Zimmer et al., 2007) war bereits bekannt, dass leistungsstarke Schülerinnen und Schüler einen höheren sozioökonomischen Status, gemessen mit den quantitativen Indizes HISEI und ESCS aufwiesen. Offen war jedoch, inwieweit diese Unterschiede in einem eindimensionalen quantitativen Index wie dem HISEI auch zwischen High und Nicht-High-Performern zu finden sind und welche Unterschiede sich ergeben, wenn man statt eines eindimensionalen Index eine kategoriale und stärker qualitative Einteilung des sozioökonomischen Status basierend auf den EGP-Klassen (Erikson et al., 1979; Erikson & Goldthorpe, 2002) benutzt wird. In diesem Maß werden Berufe nicht nur in Bezug auf das Einkommen, sondern zusätzlich in der mit dem Beruf verbundenen Weisungsbefugnis, Tätigkeiten und notwendiger Qualifikation eingeteilt (vgl. Kapitel 2.2). Innerhalb eines Einkommensniveaus können damit verbundene Merkmale wie Macht oder gesellschaftliches Ansehen eines Elternberufs Auswirkungen auf das Bildungsumfeld, das Erziehungsverhalten und die Möglichkeiten eines Kindes haben. In dieser Arbeit wurde daher untersucht, wie sich High und Nicht-High-Performer in der Höhe des HISEI und den EGP-Klassen unterscheiden.

Den Analysen zufolge hatten High Performer sowohl 2006 als auch 2012 einen höheren sozioökonomischen Hintergrund als Nicht-High-Performer, die Unterschiede waren 2012 jedoch geringer. Dies würde zunächst andeuten, dass die Unterschiede im sozioökonomischen Hintergrund zwischen High und Nicht-High-Performern abnahmen. In 2012 wurde jedoch eine neue Kodierung der Berufe (ISCO-08 statt ISCO-88) verwendet, auf die der Rückgang möglicherweise zurückzuführen ist (vgl. Müller & Ehmke, 2013). Diese Ergebnisse entsprechen den Befunden von Zimmer et al. (2007) für 15-Jährige, die sowohl in Naturwissenschaften als auch Mathematik und Lesen zu den besten 10% gehören. Auch in PISA 2006 zeigte sich für die Top Performer (OECD, 2009b), dass sie einen höheren sozioökonomischen Status aufwiesen. Den Ergebnissen von TIMSS/IGLU zufolge unterschieden sich soziale Schicht und Bildungsabschluss der Eltern bereits zwischen leistungsstarken und weniger leistungsstarken Viertklässlern (Wendt et al.,

2013).

Für die EGP-Klassen zeigten die Ergebnisse dieser Arbeit darüber hinaus, dass High Performer häufiger als Nicht-High-Performer aus Elternhäusern stammten, bei denen der Beruf des Vaters den beiden oberen EGP-Klassen (z. B. freie akademische Berufe, Hochschul- oder Gymnasiallehrer oder technische Angestellte mit nicht-manueller Tätigkeit) zugerechnet wird und seltener aus Elternhäusern, in denen der Beruf des Vaters zu den Facharbeitern und ungelerten Arbeitern gezählt wird. Dies ist in Einklang mit dem Befund aus PISA 2012, dass Familien in den beiden oberen EGP-Klassen mehr Ressourcen und lernrelevante Besitztümer aufweisen, die Eltern höhere Bildungsabschlüsse erreichen, Kinder häufig mindestens ein Jahr den Kindergarten besuchen sowie mit höherer Wahrscheinlichkeit ein Gymnasium besuchen (Müller & Ehmke, 2013). Einschränkend muss allerdings berücksichtigt werden, dass die ISCO-Kodierung das vergleichsweise hohe Bildungsniveau des dualen Ausbildungssystems in Deutschland nicht berücksichtigt und daher nichtakademische Elternhäuser in ihrem Bildungsniveau tendenziell unterschätzt. Insgesamt bestätigen die Ergebnisse zum sozioökonomischen Hintergrund leistungsstarker Schülerinnen und Schüler die bestehende Forschungslage und erweitern sie um Ergebnisse zu den EGP-Klassen sowie der Entwicklung innerhalb der High Performer zwischen 2006 und 2012. Die Ergebnisse zeigen, dass das familiäre Umfeld, gemessen durch den sozioökonomischen Hintergrund, leistungsstarke Jugendliche nach wie vor auszeichnet. Der sozioökonomische Hintergrund kann von der pädagogischen Praxis nicht verändert werden und eine Benachteiligung von Jugendlichen aus besseren sozialen Lagen ist nicht anzustreben. Die Ergebnisse in dieser Arbeit verdeutlichen deshalb vor allem, dass in der Entwicklung von und Forschung zu Fördermaßnahmen für leistungsstarke Jugendliche mögliche familiäre Einflüsse berücksichtigt werden sollten und für Chancengerechtigkeit und einen höheren Anteil leistungsstarker Jugendlicher geeignete Kompensationsmaßnahmen für sozioökonomisch schlechter gestellte Schülerinnen und Schüler getroffen werden sollten.

Vergleich von Top und High Performern

Mit den Top Performern (Schülerinnen und Schüler auf den beiden obersten Kompetenzstufen) bestand in den internationalen Vergleichen in PISA bereits eine Definition für eine enge Leistungsspitze in Naturwissenschaften (vgl. OECD, 2009b). Für die breiter definierte Gruppe der High Performer (beste 20%) stellte sich daher die Frage, inwieweit sich

die darin enthaltenen Teilgruppen der Top Performer und übrigen High Performer, die die beiden obersten Kompetenzstufen noch nicht erreicht haben, in ihren individuellen und familiären Merkmalen gleichen. Wie der Vergleich dieser beiden Gruppen innerhalb der High Performer zeigte, stimmen beide Gruppen leistungsstarker Jugendlicher in den Merkmalen, die sie von den restlichen Jugendlichen auszeichnen, überein. Lediglich in einzelnen Merkmalen wie der Kompetenz in Mathematik und Lesen erreichten Top Performer noch höhere Werte. Dies ist jedoch ebenfalls konform zur Annahme, dass die Kompetenzen der unterschiedlichen Domänen Überlappungsbereiche besitzen. Damit liefert diese Arbeit auch Hinweise darauf, dass weitere bereits bekannte Ergebnisse zu Top Performern (vgl. OECD, 2009b) möglicherweise in höherem Umfang auf High Performer übertragen werden können.

Institutionelle Merkmale

Neben den individuellen und familiären Merkmalen wurden die Verteilung auf Schularten und Klassen sowie die Klassenkompositionen als institutionelle Rahmenbedingungen, in denen sich High Performer befinden, im Vergleich zu Nicht-High-Performern untersucht. Wie Baumert und Köller (1998b) aus Ergebnissen in TIMSS schlussfolgerten, lernen Schülerinnen und Schüler, die ein Gymnasium besuchen, vergleichsweise mehr als Schülerinnen und Schüler mit gleichen Lernvoraussetzungen, die eine andere weiterführende Schulart besuchen. Da das Gymnasium den Bildungsauftrag hat, leistungsstarke Jugendliche zu fördern (Heller, 2002c), stellt sich insbesondere für diese Schülergruppe die Frage, inwieweit sie von dieser Schulart stammt. Als zwei weitere Aspekte der institutionellen Bedingungsfaktoren haben die Verteilung auf Schulklassen sowie die Klassenkomposition möglicherweise Einfluss auf die Entwicklung hin zu hoher Leistung in Naturwissenschaften (OECD, 2009b). Forschungsfragen dieser Arbeit waren daher, wie sich leistungsstarke im Vergleich zu den anderen Jugendlichen auf Schularten und Klassen verteilen und wie leistungsstark und heterogen Schulklassen mit unterschiedlichen Anteilen an leistungsstarken Jugendlichen sind.

Die Untersuchungen zeigten, dass über drei Viertel der High Performer ein Gymnasium besuchten und insgesamt knapp die Hälfte der Schülerinnen und Schüler einer Gymnasialklasse zu den High Performern gehörte. Den zweithöchsten Anteil an High Performern wiesen die Realschulen auf, während jeweils weniger als 10% der High Performer Schulen mit mehreren Bildungsgängen und Integrierte Gesamtschulen be-

suchten. Zwischen 2006 und 2012 gab es dabei keine nennenswerten Veränderungen in der Verteilung der High Performer auf Schularten. Dies ist angesichts der Tatsache, dass von 2006 auf 2012 der Anteil an Top Performern nicht gestiegen ist und sich auch das allgemeine Naturwissenschaftsniveau in Deutschland nicht verändert hat (Schiepe-Tiska et al., 2013), plausibel. Überraschend ist dieses Ergebnis jedoch im Vergleich zur Entwicklung der Gymnasialbeteiligung. In der Gesamtstichprobe der 15-Jährigen hat sich die Gymnasialbeteiligung von 2006 auf 2012 um fünf Prozent erhöht (Ehmke & Baumert, 2007; Sälzer et al., 2013), unter den High Performern ist sie jedoch unverändert geblieben. Einschränkend muss auch hier bei der Interpretation berücksichtigt werden, dass in der Stichprobe der 15-Jährigen ein Teil der Schülerinnen und Schüler die zehnte Klasse besucht. Dennoch kann insgesamt aus den Ergebnissen geschlossen werden, dass Deutschland durch eine Erhöhung der Gymnasialbeteiligung unter den High Performern (ein Viertel der High Performer besucht nicht-gymnasiale Schularten) möglicherweise weiteres Potential für Leistungssteigerungen dieser Gruppe ausschöpfen könnte.

Neben der Schulartverteilung wurde in den Analysen in dieser Arbeit herausgefunden, dass sich High Performer an Gymnasien relativ gleichmäßig auf Schulklassen verteilen. Darüber hinaus zeigte sich, dass sich Schulklassen mit vergleichsweise hohen High-Performer-Anteilen nicht durch eine stärkere Homogenität auszeichnen. Diese Ergebnisse widersprechen der Ansicht, dass leistungsstarke Jugendliche vor allem in homogenen leistungsstarken Schulklassen anzutreffen sind.

Verglichen mit bestehenden Forschungsergebnissen zu leistungsstarken Jugendlichen bestätigen die Befunde zu institutionellen Merkmalen, dass leistungsstarke Jugendliche zum Großteil an Gymnasien zu finden sind (vgl. Zimmer et al., 2007). In Ergänzung zu Ergebnissen internationaler Untersuchungen (z. B. OECD, 2009b) wurden in dieser Arbeit die länderspezifischen Schularten in Deutschland betrachtet, eine klassenbasierte Stichprobe verwendet sowie die Klassenkomposition berücksichtigt.

Insgesamt wurde in der Charakterisierung der High Performer in dieser Arbeit einerseits deutlich, dass High Performer eine Gruppe mit besonderen Merkmalen sind und sich von Nicht-High-Performern in ihrer kognitiven Grundfähigkeit, den Leistungen in Mathematik und Lesen bzw. Deutsch, dem soziökonomischen Hintergrund und auch der Verteilung auf Schularten unterscheiden. Andererseits zeigte sich aber auch, dass die High Performer selbst in vielen Merkmalen wie der kognitiven Grundfähigkeit, den Kompetenz- und Notenprofilen und ihren sozialen Schichten eine durchaus heterogene

Gruppe sind. Sie repräsentieren damit vor allem den oberen Abschnitt der Verteilung individueller, familiärer und institutioneller Merkmale. Dies bedeutet zwar zum einen, dass die Gruppe nicht pauschal aufgrund ihrer individuellen, familiären und institutionellen Merkmalausprägungen identifiziert werden kann, sondern differenzierter gesehen und behandelt werden muss. Zum anderen eröffnet die Tatsache, dass sich leistungsstarke Schülerinnen und Schüler am oberen Ende eines Kontinuums verschiedener veränderbarer Merkmale befindet, Ansatzpunkte und Möglichkeiten für Forschungs- und Förderbemühungen, beispielsweise in der Förderung mathematischer Kompetenzen oder dem Übertritt auf ein Gymnasium. Schließlich konnten mit den Geschlechterunterschieden und der fehlenden Steigerung der Gymnasialbeteiligung zwei Problemlagen in der Förderung leistungsstarker Jugendlicher in Naturwissenschaften identifiziert werden, für die weiterer Forschungs- und Handlungsbedarf in der pädagogischen Praxis besteht.

8.2.2. Unterrichtsaktivitäten im Zusammenhang mit der naturwissenschaftlichen Kompetenz und dem Interesse von leistungsstarken Jugendlichen in Naturwissenschaft

Wie in der Einleitung dargestellt, besteht ein großes Interesse daran, leistungsstarke Jugendliche nicht nur in ihrer Leistung, sondern auch ihren motivational-affektiven Merkmalen zu fördern. Der Schule als institutioneller Kontext kommt dabei besondere Bedeutung zu, da hier alle Schülerinnen und Schüler Gelegenheit für systematische Erfahrungen in Naturwissenschaften haben. Im ersten Ergebnisteil wurde bereits untersucht, in welchen Schularten und Klassenkompositionen High Performer zu finden sind. Daran schließt sich die Frage an, in welchem Zusammenhang der Unterricht innerhalb dieser Rahmenbedingungen mit der naturwissenschaftlichen Kompetenz und dem Interesse leistungsstarker Jugendlicher in Naturwissenschaften steht. Im zweiten Ergebnisteil wurde daher untersucht, inwieweit das Vorkommen der Unterrichtsaktivitäten „interaktives Lehren und Lernen“, „forschend-entdeckender Unterricht“ und „Anwendungen“ mit der naturwissenschaftlichen Kompetenz und dem Interesse von High Performern zusammenhängt und welche Unterschiede zu Nicht-High-Performern bestehen.

Dies wurde in Mehrebenenanalysen mit einem Multigruppendedesign untersucht. Für die Unterrichtsmerkmale wurden auf Klassenebene jeweils die Werte aller Schü-

lerinnen und Schüler einer Klasse, auf Individualebene jeweils nur die Angaben der High bzw. Nicht-High-Performer verwendet. Der Zusammenhang zwischen Unterricht und Kompetenz bzw. Interesse wurde auch auf Schülerebene modelliert, da die individuelle Abweichung der Unterrichtswahrnehmung innerhalb einer Klasse inhaltlich sinnvoll interpretiert werden konnte (vgl. Lüdtke et al., 2009, 2011). Damit sichergestellt werden konnte, dass ein Vergleich von High und Nicht-High-Performern mit den verwendeten Skalen zulässig ist, wurde zunächst die Messinvarianz der Skalen zwischen beiden Gruppen nachgewiesen. In den darauf aufbauenden Mehrebenenanalysen wurden die Zusammenhänge von Unterrichtsaktivitäten und Kompetenz sowie Unterrichtsaktivitäten und Interesse untersucht und zwischen beiden Leistungsgruppen verglichen.

Die Ergebnisse im Zusammenhang mit der naturwissenschaftlichen Kompetenz ergaben, dass bei keiner der Unterrichtsaktivitäten die betrachtete Häufigkeit die naturwissenschaftliche Kompetenz von High Performern vorhersagte, weder zwischen Klassen, noch innerhalb Klassen. Für Nicht-High-Performer sagte die Häufigkeit der Anwendungen in einer Schulklasse dagegen vorher, wie leistungstark diese Schülergruppe einer Klasse war. Je häufiger diese Jugendlichen im Unterschied zu ihren Mitschülern forschend-entdeckende Elemente berichteten, desto niedriger war ihre naturwissenschaftliche Kompetenz im Vergleich zu den Mitschülern. Ein Unterschied zwischen leistungsstarken und den restlichen Schülerinnen und Schülern wurde nur im Zusammenhang der Häufigkeit von Experimenten auf Schülerebene und naturwissenschaftlicher Kompetenz beobachtet.

Auch zwischen der wahrgenommenen Häufigkeit von Unterrichtsaktivitäten und dem Interesse von High Performern wurde kein Zusammenhang gefunden. Dies impliziert, dass die naturwissenschaftliche Kompetenz und das Interesse, das High Performer und Nicht-High-Performer im Mittel in einer Klasse erreichen, unabhängig von der Häufigkeit der Unterrichtsaktivitäten sind, die in der Klasse wahrgenommen werden. Berichteten High Performer bzw. Nicht-High-Performer innerhalb ihrer Klasse jedoch häufiger Anwendungsbezüge als ihre Mitschüler der gleichen Leistungsgruppe, so waren sie interessierter als ihre Mitschüler. Insgesamt war keines der Unterrichtsmerkmale sowohl mit naturwissenschaftlicher Kompetenz als auch Interesse in einer der beiden Leistungsgruppen assoziiert.

Der Vergleich von High und Nicht-High-Performern verdeutlichte, dass nur im Zusammenhang der Häufigkeit von Experimenten auf Schülerebene und naturwissenschaft-

licher Kompetenz sowie im Zusammenhang von Geschlecht und Interesse Unterschiede bestehen. Bei Nicht-High-Performern wurden insgesamt mehr Prädiktoren gefunden und alle für High Performer identifizierten positiven Prädiktoren galten ebenso für Nicht-High-Performer.

Da in dieser Arbeit Häufigkeiten einzelner Unterrichtsaktivitäten (z. B. das Vorkommen von Schülerexperimenten) anstatt einzelner Unterrichtsformen im Vergleich zu anderen (z. B. forschend-entdeckender Unterricht im Vergleich zu traditionellem Unterricht) betrachtet wurde, geben die Ergebnisse systematische Zusammenhänge zwischen Häufigkeitsunterschieden in den Unterrichtsaktivitäten und Unterschieden in der Höhe der naturwissenschaftlichen Kompetenz und des Interesses an. Sie bedeuten daher nicht, dass leistungsstarke Schülerinnen und Schüler von diesen Unterrichtsaktivitäten nicht profitierten.

Im Gegenteil fand die Forschung der letzten Jahrzehnte zahlreiche Hinweise dafür, dass die betrachteten Unterrichtsaktivitäten sowohl bei High als auch bei Nicht-High-Performern in einem Zusammenhang mit Leistung und Interesse in Naturwissenschaften stehen. Interaktives bzw. kooperatives Lehren und Lernen geht demnach allgemein (Scheerens & Bosker, 2011; Seidel & Shavelson, 2007) sowie bei leistungsstarken Schülerinnen und Schülern (Slavin, 1990a, 1990b) mit höherer Leistung einher. Forschend-entdeckender Unterricht unterstützt vor allem mit motivational-affektiven Merkmale und bei kognitiv aktivierender Gestaltung auch die Leistungsentwicklung (George & Kaplan, 1998; Kjærnsli & Lie, 2011; Seidel & Shavelson, 2007; White & Frederiksen, 1998). High Performer sollten aufgrund ihres höheren Vorwissens und ihrer vermutlich ausgefeilteren Lernstrategien in der Elaboration und Organisation von Lerninhalten weniger Schwierigkeiten haben, ein höheres Lerntempo aufweisen und daher insgesamt mehr von den Unterrichtsaktivitäten profitieren. Auch direkt für die in dieser Arbeit betrachteten Unterrichtsskalen wurde in früheren Untersuchungen nachgewiesen, dass Unterrichtsmuster mit häufigem interaktivem Lehren und Lernen vor allem in Zusammenhang mit Leistung stehen, Unterrichtsmuster mit zahlreichen forschend-entdeckenden Elementen vor allem im Zusammenhang mit Interesse assoziiert sind (Kobarg et al., 2011). Den stärksten Zusammenhang mit naturwissenschaftlicher Kompetenz und Interesse zeigte dagegen ein Unterrichtsmuster, das zahlreiche Diskussionen und gezielt eingesetzte Experimente enthält (Kobarg et al., 2011). Darüber hinaus wurde für die in dieser Arbeit betrachteten Unterrichtsmerkmale bereits gezeigt, dass bei Jugendlichen des oberen

Leistungsquartils in Naturwissenschaften ein höheres Interesse mit häufiger berichteten Anwendungsbezügen einhergeht und zu den anderen Unterrichtsmerkmalen kein Zusammenhang bestand (Prenzel & Schütte, 2007). Dieser Befund wurde in dieser Arbeit bestätigt und mit Mehrebenenanalysen repliziert. Die Mehrebenenanalysen in dieser Arbeit bilden im Unterschied zu den genannten Studien ab, ob ein häufigeres Auftreten eines Unterrichtsmerkmals mit einer höheren Ausprägung einer abhängigen Variable einhergeht (vgl. Kapitel 8.1). Wenn beispielsweise einzelne Unterrichtsaktivitäten in ihrer Häufigkeit vor allem bis zu einem Schwellenwert wirksam sind – wie es etwa bei Kobarg et al. (2011) das Unterrichtsprofil mit gezielten Experimenten andeutet –, so ist das in den variablenzentrierten Mehrebenenanalyse nicht unbedingt sichtbar.

Die Ergebnisse dieser Arbeit stellen die bisherigen Befunde zur generellen Wirksamkeit der Unterrichtsaktivitäten daher nicht in Frage. Sie zeigen jedoch, dass High Performer unabhängig davon, wie häufig die einzelnen Unterrichtsaktivitäten aus Schülersicht vorkommen, Lernerfolge erzielen. Ihre hohe Kompetenz befähigt sie möglicherweise dazu, von verschiedenen Unterrichtsaktivitäten zu profitieren, und schützt sie möglicherweise vor Auswirkungen, die mit dem unterschiedlich häufigen Vorkommen der Unterrichtsaktivitäten verbunden sein können. Es ist beispielsweise denkbar, dass sich leistungsstarke Jugendliche aus wenigen Schülerexperimenten durch ihre hohe Vorleistung ähnlich viele Lernimpulse holen können wie aus häufigeren Experimenten oder selten vorkommende Anwendungsbezüge sie bereits dazu anregen, bei fehlenden Anwendungsbezügen selbst über mögliche Anwendungen nachzudenken oder sich damit zu beschäftigen.

Eine solche Interpretation der Befunde wird durch Ergebnisse der Forschung zu Aptitude-Treatment-Interaction (ATI, Brünken & Leutner, 2005) unterstützt. Dort zeigt sich insgesamt die Tendenz, dass der Design-Effekt eines Treatments bei Lernenden mit höherem Vorwissen geringer ausfällt als bei Lernenden mit geringerem Vorwissen. Leistungsstarke Jugendliche erzielten relativ unabhängig von der Methode Lernzuwächse, bei leistungsschwächeren Jugendlichen konnten deutlichere Unterschiede zwischen Methoden ausgemacht werden (vgl. Hasebrook & Brünken, 2010). Schülerinnen und Schüler mit hohem Vorwissen konnten wenig förderliche Designs damit kompensieren. In Einklang mit den Ergebnissen der ATI-Forschung sowie den Befunden dieser Arbeit steht auch das Ergebnis von White und Frederiksen (1998) für forschend-entdeckenden Unterricht in Naturwissenschaften: Leistungsschwächere Schülerinnen und Schüler

erzielten bei forschend-entdeckendem Unterricht höhere Leistungszuwächse als leistungsstarke, beide profitierten jedoch von der Lernumgebung. In der ATI-Forschung zeigte sich jedoch auch, dass für Jugendliche mit hohem Vorwissen instruktionale Designs, die für Jugendliche mit niedrigem Vorwissen lernförderlich und kompensierend wirken, ab einem gewissen Vorwissensniveaus nicht nur nicht nötig sind, sondern sogar lernhinderlich wirken (Kalyuga, Ayres, Chandler & Sweller, 2003). Dies wird als *Expertise-Reversal Effect* bezeichnet. Häufig ergibt sich für den Einfluss eines Treatments in Abhängigkeit des Vorwissens insgesamt eine U-Form, indem Jugendliche mit extremen Vorwissensausprägungen am wenigsten profitieren (Hasebrook & Brünken, 2010). Dies könnte auch erklären, warum nicht nur bei den High Performern, sondern auch den Nicht-High-Performern wenig Zusammenhänge gefunden wurden, denn sie setzen sich in dieser Arbeit sowohl aus leistungsschwachen Jugendlichen als auch Jugendlichen mit mittlerem Leistungsniveau zusammen.

Bei den Mehrebenenanalysen in dieser Arbeit wurde darüber hinaus deutlich, dass sich die Häufigkeiten der Unterrichtsaktivitäten zwischen den Gymnasialklassen nur gering unterschieden (vgl. Kapitel 7.2). Diese geringe Varianz des Naturwissenschaftsunterrichts zwischen Klassen einer Schulart ist zwar in Einklang mit bestehenden Ergebnissen wie beispielsweise aus der IPN-Videostudie für den Physikunterricht (Seidel, 2003). Gleichzeitig wird jedoch ein beträchtlicher Teil der Varianz der naturwissenschaftlichen Kompetenz und des Interesses leistungsstarker Jugendlicher durch die Unterrichtsaktivitäten zwischen und innerhalb Klassen nicht aufgeklärt. Dies könnte darauf hinweisen, dass andere Merkmale des Unterrichts im Gegensatz zu der Häufigkeit von Unterrichtsaktivitäten Varianz erklären. In der Unterrichtsforschung der letzten Jahrzehnte zeigte sich, dass proximale Merkmale und Merkmale der Unterrichtsqualität die Wirksamkeit von Unterricht stärker beeinflussen als strukturelle Unterrichtsmerkmale (Seidel & Shavelson, 2007). Die Orchestration der Unterrichtsaktivitäten und die qualitative Gestaltung von interaktivem Lehren und Lernen sowie forschend-entdeckendem Unterricht sind unter anderem zwei Merkmalsbereiche, auf die die Ergebnisse dieser Arbeit möglicherweise hinweisen (vgl. Kobarg et al., 2011; Slavin, 1990a; Seidel, Prenzel, Rimmele et al., 2007; Tesch & Duit, 2004). Dies wurde unter anderem im dritten Ergebnisteil der Arbeit aufgegriffen.

8.2.3. Selbstkonzept-Interesse-Typen und mit ihnen verbundene Unterschiede in der Unterrichtswahrnehmung

Die Befunde, dass nur wenig Zusammenhänge zwischen Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz sowie Interesse leistungsstarker Jugendlicher belegt werden konnten, werfen die Frage auf, wie motivationsunterstützend naturwissenschaftlicher Unterricht gestaltet ist. Gleichzeitig zeigen Untersuchungen zu individuellen Schülervoraussetzungen, dass innerhalb aller Schülerinnen und Schüler Typen mit unterschiedlichen Merkmalskombinationen aus Selbstkonzept und Interesse gefunden werden können und Jugendliche unterschiedlicher Typen ihren Unterricht unterschiedlich wahrnehmen (z. B. Seidel, 2006). Übertragen auf leistungsstarke Jugendliche stellt sich dadurch die Frage, welche motivational-affektiven Typen in dieser Gruppe identifiziert werden können und inwieweit eine unterschiedliche Unterrichtswahrnehmung von Jugendlichen unterschiedlicher Typen erklären kann, warum ein Teil der leistungsstarken Jugendlichen bei gleichem Naturwissenschaftsunterricht motivational-affektiv indifferent bleibt. Erkenntnisse dazu würden näher beschreiben, inwieweit motivational-affektiv indifferente leistungsstarke Jugendliche eine Problemlage darstellen und welche Merkmale sie auszeichnen. Zusammenhänge zwischen unterschiedlichen motivational-affektiven Typen und ihrer Unterrichtswahrnehmung würden darüber hinaus Hinweise darauf geben, wie leistungsstarke Jugendliche zukünftig motivational-affektiv besser gefördert werden können.

In dieser Arbeit wurde daher untersucht, (a) welche Selbstkonzept-Interesse-Typen unter den High Performern identifiziert werden können, (b) welche Merkmale diese Typen aufweisen und (c) ob sie ihren Naturwissenschaftsunterricht unterschiedlich motivationsunterstützend erleben. Zusätzlich wurde untersucht, (d) ob eine unterschiedliche motivationsunterstützende Wahrnehmung der Typen davon abhängt, wie häufig forschend-entdeckende Unterrichtsaktivitäten erfolgen. Letztere Frage verbindet den letzten Fragestellungsteil mit dem vorherigen Abschnitt, in dem sich – obwohl theoretisch anders erwartet – ergab, dass häufigere forschend-entdeckende Unterrichtsaktivitäten in keinem Zusammenhang zum Interesse von High Performern standen. Zur Beantwortung dieser Fragestellungen wurden mit Hilfe einer Latenten Profilanalyse (LPA) unterschiedliche Selbstkonzept-Interesse-Typen unter den High Performern identifiziert. Im Anschluss daran wurden Unterschiede in der Unterrichtswahrnehmung in

einer Mehrebenenanalyse untersucht und schließlich eine Moderationsanalyse für die *cross-level*-Interaktionen von forschend-entdeckendem Unterricht und Typen in der Vorhersage von Interesse angestellt.

Mit Hilfe der latenten Profilanalysen wurden vier Selbstkonzept-Interesse-Typen unter den High Performern identifiziert: (1) motivational-affektiv indifferente High Performer mit niedrigem Selbstkonzept und niedrigem Interesse, (2) interessierte High Performer mit niedrigem Selbstkonzept aber hohem Interesse, (3) leistungszuversichtliche High Performer mit hohem Selbstkonzept aber niedrigem Interesse und (4) motivational-affektiv beteiligte High Performer mit hohem Selbstkonzept und hohem Interesse. Sowohl bei Mädchen als auch Jungen wurden ähnliche Typen gefunden, die sich nicht in ihrer Profilverlauf hinsichtlich Selbstkonzept und Interesse, jedoch in ihrem relativen Anteil innerhalb der Geschlechter unterschieden. Jungen gehörten häufiger als Mädchen dem motivational-affektiv beteiligten Typ an. Der Typ der motivational-affektiv beteiligten High Performer trat unter Jungen und Mädchen am häufigsten auf, ungefähr 8% der Jungen und 18% der Mädchen gehörten jedoch zu den motivational-affektiv indifferenten High Performern. Damit bestätigt diese Arbeit mit einer personen-zentrierten Analyse das Ergebnis von Schütte et al. (2007), dass ein beträchtlicher Teil leistungsstarker Jugendlicher in Naturwissenschaften im Vergleich zu ihrer Leistung nur gering interessiert ist. In den personen-zentrierten Analysen zu Typen aus kognitiven und motivational-affektiven Merkmalen aller Schülerinnen und Schüler in Physik von Seidel (2006) resultierten zwei Typen, die als leistungsstark bezeichnet werden könnten. Ein Typ zeichnete sich gleichzeitig durch ein hohes Selbstkonzept und Interesse aus, das zweite durch ein vergleichsweise niedriges Interesse und insbesondere niedriges Selbstkonzept. Diese beiden Typen wurden auch in der LPA in dieser Arbeit identifiziert. Zusätzlich wurden jedoch noch zwei weitere Typen mit lediglich hohem Selbstkonzept bzw. lediglich hohem Interesse mit Anteilen von ungefähr 10-20% gefunden. Diese Typen wurden bei Seidel (2006) lediglich für Jugendliche mit mittlerem Leistungsniveau identifiziert. Die beobachteten Geschlechterunterschiede in den Häufigkeiten der Typen decken sich ebenfalls mit bestehender Forschung (vgl. Seidel, 2006) und ergänzen sie um das Ergebnis, dass zwischen Mädchen und Jungen auch keine unterschiedlichen Profilverläufe in den Selbstkonzept-Interesse-Typen zu finden sind, ähnlich wie es Lau und Roeser (2008) in Typen aus Leistung und verschiedenen motivational-affektiven Merkmalen wie Zielorientierung, Emotionen oder Werte fand.

Die Selbstkonzept-Interesse-Typen unterschieden sich nicht in ihrem Kompetenzprofil in Naturwissenschaften, Mathematik und Lesen, zeigten jedoch unterschiedliche Notenprofile in Naturwissenschaften, Mathematik und Deutsch. Motivational-affektiv beteiligte High Performer hatten häufiger ein naturwissenschaftliches Notenprofil, motivational-affektiv indifferente High Performer häufiger ein sprachliches Notenprofil. Dies entspricht der Annahme dimensionaler Vergleiche, dass Schülerinnen und Schüler in dem Fach, in dem sie im Vergleich zu anderen Fächern höhere Leistungen erreichen, ein höheres Selbstkonzept zeigen (z. B. Möller & Marsh, 2013). Allerdings wurde das Selbstkonzept bzw. Interesse in den anderen Fachbereichen nicht erfasst und es konnte daher nicht untersucht werden, ob das Selbstkonzept und Interesse in anderen Fachbereichen entsprechend der Theorie der dimensional Vergleiche höher oder niedriger ausgeprägt war. Den höchsten Anteil an High Performern mit guten Noten in allen Fächern, d. h. High Performern, bei denen ein dimensionaler Vergleich der Leistungen nicht möglich bzw. wirkungslos ist, hatte der motivational-affektiv beteiligte Typ.

Im sozioökonomischen Hintergrund dagegen unterschieden sich High Performer unterschiedlicher motivational-affektiver Typen nicht. Eine Art Habitus zugunsten des Selbstkonzepts und Interesse aus einem höheren sozioökonomischen Hintergrund kennzeichnet die Typen daher nicht (vgl. Archer et al., 2012). Allerdings vermag der verwendete Index HISEI auch nur, den sozioökonomischen Status eindimensional auf Grundlage des Berufs der Eltern abzubilden, weitere Ergebnisse würden daher gegebenenfalls mehrdimensionale Indizes wie der ESCS liefern, die beispielsweise kulturelle Besitztümer oder andere Bildungsressourcen berücksichtigen.

Im Hinblick auf die Unterrichtswahrnehmung leistungsstarker Schülerinnen und Schüler unterschiedlicher motivational-affektiver Typen wurde gezeigt, dass motivational-affektiv beteiligte High Performer mehr Autonomie- und Kompetenzunterstützung sowie mehr inhaltliche Relevanz berichteten als motivational-affektiv indifferente High Performer. Mit einem höheren Interesse war bei den Typen nur bei niedrigem Selbstkonzept eine höher wahrgenommene Autonomieunterstützung und inhaltliche Relevanz assoziiert, bei hohem Selbstkonzept wurde dagegen kein Unterschied beobachtet. Dies impliziert, dass das Niveau des Selbstkonzepts den Zusammenhang zwischen Interesse und Autonomieunterstützung bzw. inhaltlicher Relevanz von Selbstkonzept-Interesse-Typen leistungsstarker Jugendlicher moderiert. Inhaltlich bedeutet das, dass die genannten Lehr-Lernbedingungen möglicherweise nur bei niedrigem Selbstkonzept Auswirkungen

auf das Interesse von High Performern haben, das Interesse bei hohem Selbstkonzept möglicherweise aber nicht beeinflusst wird. Ein Nachweis dieser Moderation und dabei stattfindender Wirkmechanismen würde jedoch weitere Analysen und ein experimentelles Studiendesign erfordern. Bei Typen mit unterschiedlichem Selbstkonzept wurde dagegen unabhängig vom Interessenniveau eine unterschiedlich hohe Kompetenzunterstützung und damit keine Moderation gefunden. Insgesamt bestätigten sich die theoretischen Annahmen, dass Autonomieunterstützung und inhaltliche Relevanz mit Interesse, Kompetenzunterstützung mit Selbstkonzept assoziiert waren, allerdings beim Interesse nur für ein niedriges Selbstkonzeptniveau. Damit konnten analog zu ähnlichen Untersuchungen (Seidel, 2006; Huber et al., 2015) innerhalb einer Klasse unterschiedliche Mikro-Lernumgebungen identifiziert werden. Im Unterschied zu bestehenden Analysen ergaben sich in dieser Arbeit für High Performer jedoch differenziertere Zusammenhänge zwischen Typen und Lehr-Lernbedingungen, indem Selbstkonzept und Interesse unterschiedliche Unterrichtsmerkmale zugeordnet werden konnten sowie sich Unterschiede in Abhängigkeit des Selbstkonzeptniveaus zeigten.

Da forschend-entdeckender Unterricht aus theoretischer Sicht die Erfüllung motivationsunterstützender Lehr-Lernbedingungen erleichtern sollte und im Zusammenhang mit Interesse steht (vgl. Furtak & Kunter, 2012; Kobarg et al., 2011), stellt sich die Frage, inwieweit die Zusammenhänge von Typen und Wahrnehmung motivationsunterstützender Bedingungen unabhängig von der Häufigkeit forschend-entdeckenden Lernens sind. Die Analysen zeigten, dass häufigere forschend-entdeckende Unterrichtselemente bei High Performern positiv mit dem Gesamtniveau der Lehr-Lernbedingungen zusammenhängt und gleichzeitig jedoch nicht die Unterschiede in den Lehr-Lernbedingungen innerhalb der High Performer einer Klasse verstärkt. Je mehr forschend-entdeckende Unterrichtsaktivitäten in einer Klasse stattfanden, desto höher nahmen leistungsstarke Jugendliche aller motivational-affektiven Typen die Autonomieunterstützung, Kompetenzunterstützung und inhaltliche Relevanz in der Klasse im Vergleich zu anderen Klassen mit weniger forschend-entdeckendem Unterricht wahr. Gleichzeitig vergrößerten oder verkleinerten diese häufigeren forschend-entdeckende Unterrichtsaktivitäten die Schere zwischen den in einer Klasse existierenden Micro-Teaching-Learning-Umgebungen von High Performern offenbar nicht.

Nachdem häufiger wahrgenommene Elemente forschend-entdeckender Unterrichtszugänge im zweiten Ergebnisteil dieser Arbeit keinen Zusammenhang zu Interesse

von High Performern zeigten, ergaben die Analysen des dritten Teils nun ein differenzierteres Bild und belegen, dass häufigerer forschend-entdeckender Unterricht über Klassen hinweg von High Performer motivationsunterstützend wahrgenommen wird. Darüber hinaus implizieren die Ergebnisse, dass dieser Zusammenhang von häufiger wahrgenommenen forschend-entdeckenden Unterrichtselementen und Lehr-Lernbedingungen für alle High Performer gleichermaßen gilt und die Unterschiede zwischen High Performern unterschiedlicher Selbstkonzept-Interesse-Typen innerhalb der Klassen nicht beeinflusst werden. Aufgrund des querschnittlichen Designs können aus diesen Ergebnissen nur systematische gerichtete Zusammenhänge abgeleitet werden. Sie geben jedoch Hinweise auf mögliche Ursache-Wirkungs-Beziehungen, die in anschließenden experimentellen Forschungsdesigns untersucht werden können.

Anschließend an bisherige Untersuchungen zu Mikro-Lernumwelten (vgl. Seidel et al., 2016) zeigt die Arbeit, dass innerhalb von Schulklassen unterschiedliche Mikro-Lernumwelten identifiziert werden können, in denen Jugendliche ihren Unterricht unterschiedlich motivationsunterstützend erleben. Die Befunde belegen darüber hinaus, dass die Unterschiede in der Unterrichtswahrnehmung zwischen den Mikro-Lernumwelten einer Schulklasse unabhängig davon sind, wie häufig die Schülerinnen und Schüler der Schulklasse forschend-entdeckende Unterrichtsaktivitäten berichten. Schließlich zeigt sie, dass die Häufigkeit von forschend-entdeckendem Lernen – obwohl sie in keinem Zusammenhang zu den Unterschieden von Mikro-Lernumwelten steht – in den Schulklassen positiv mit dem mittleren Niveau aller Mikro-Lernumwelten zusammenhängt. Das bedeutet, häufigere Experimente in Schulklassen gehen über Mikro-Lernumwelten hinweg mit höher wahrgenommenen Lehr-Lernbedingungen in Schulklassen einher. In bisherigen Untersuchungen wurde zwar die Mehrebenenstruktur berücksichtigt, Zusammenhänge von Typen, strukturellen Unterrichtsmerkmalen und Lehr-Lernbedingungen wurden auf Level 2 jedoch nicht modelliert (z. B. Jurik et al., 2014; Huber et al., 2015).

8.3. Implikationen für die pädagogische und bildungspolitische Praxis

Nachdem in den vorherigen Abschnitten methodische Einschränkungen der Arbeit dargestellt und die Ergebnisse zusammengefasst und diskutiert wurden, wird im folgenden Kapitel erläutert, welche Implikationen diese Arbeit für die pädagogische Praxis und darauf abzielende bildungspolitische Entscheidungen hat. Die Schlussfolgerungen können in die drei Formen evidenzbasierten Wissens eingeordnet werden, die in der empirischen Bildungsforschung unterschieden werden (vgl. Bromme, Prenzel & Jäger, 2014; Prenzel, 2012): Beschreibungswissen, Erklärungswissen und Veränderungswissen. Im Hinblick auf Handlungsempfehlungen für die Praxis bauen die drei Wissensarten häufig aufeinander auf, indem zur Identifikation der Problemstellung und Situation beschreibendes Wissen notwendig ist, Erklärungswissen mögliche Gründe belegt und Veränderungswissen zeigt, mit welchen Maßnahmen ausgehend von der identifizierten Problemstellung zukünftig Verbesserungen erreicht werden können. Die Wissensbereiche beziehen sich jeweils auf Voraussetzungen, Prozesse oder Ergebnissen von Bildung bzw. deren Zusammenhänge in verschiedenen Kontexten.

Möglichkeiten für die Anwendung und Nutzung der Ergebnisse dieser Arbeit in der pädagogischen und politischen Praxis liegen im Bereich des Beschreibungswissens sowie des Vorhersagewissens. Sie beziehen sich sowohl auf Voraussetzungen als auch Ergebnisse von Bildung und berücksichtigen in Form von Unterrichtsmerkmalen auch Merkmale, die Rahmenbedingungen des Lernprozesses beschreiben. Durch die Stichprobenbeziehung in PISA ermöglicht die Arbeit, beschreibendes Wissen repräsentativ für die Jugendlichen der neunten Klassen in Deutschland zu generieren und dadurch Stärken aber auch Problemlagen in der Förderung leistungsstarker Jugendlicher in Deutschland aufzuzeigen. Die Implikationen der Arbeit werden nachfolgend den drei Teilbereichen (a) Identifikation leistungsstarker Jugendlicher in Naturwissenschaft, (b) Förderung der naturwissenschaftlichen Kompetenz und (c) Förderung von Interesse zugeordnet.

Identifikation leistungsstarker Jugendlicher in Naturwissenschaft

Im Beschluss der Kultusministerkonferenz zur Förderstrategie für leistungsstarke Jugendliche werden Schülerinnen und Schüler, die „bereits sehr gute beobachtbare Leistungen erbringen, ebenso wie Schülerinnen und Schüler, deren Potenziale es zu erkennen

und durch gezielte Anregung und Förderung zu entfalten gilt“ (KMK, 2015, S. 3) als Zielgruppe beschrieben. Die Diagnose dieser Jugendlichen soll mit Hilfe geschulter Beratungslehrer, externer Diagnosestellen und den unterrichtenden Lehrkräften erfolgen. Dies lässt weitgehend offen, auf welchen Kriterien die Identifikation erfolgen soll und lässt implizit beispielsweise sowohl die Diagnose auf Basis von Schulnoten als auch des Eindrucks der Lehrkräfte oder Messinstrumenten von Beratungsstellen für Hochbegabung zu.

Die vorliegende Arbeit liefert einige Erkenntnisse dazu, durch welche besonderen individuellen, familiären und institutionellen Merkmale sich Jugendliche mit einem hohen Potential in Naturwissenschaften auszeichnen und an welchen Merkmalen sie darauf aufbauend in der pädagogischen Praxis identifiziert aber auch nicht erkannt werden können. Im Hinblick auf individuelle Merkmale zeigen die Ergebnisse, dass leistungsstarke Schülerinnen und Schüler in Naturwissenschaften selten nur allein in Naturwissenschaften ein hohes Potential haben, sondern vielmehr sowohl im Kompetenztest als auch den Noten häufig auch in anderen Fächern gute Leistungen erzielen. Für die Identifikation leistungsstarker Jugendlicher durch Lehrkräfte aber auch für bildungspolitische Vorgaben zur Identifikation kann aus diesem beschreibenden Wissen gefolgert werden, dass man ein hohes Potential in Naturwissenschaften nicht unbedingt an Charakteristika, die lediglich auf ein naturwissenschaftliches Profil abzielen, festmachen sollte. Vielmehr sollte in der Identifikation berücksichtigt werden, dass ein beträchtlicher Teil in anderen, z. B. sprachlichen Fächern, ähnlich leistungsstark ist wie in Naturwissenschaften.

Wie die Arbeit ebenfalls zeigt, eignen sich Schulnoten für die Identifikation nur eingeschränkt. Gut ein Fünftel der leistungsstarken Jugendlichen erhielt über alle naturwissenschaftlichen Fächer hinweg nur eine Durchschnittsnote im Bereich „ausreichend“ und schlechter. Schulnoten bilden daher nur bedingt ab, welche Jugendlichen eine hohe naturwissenschaftliche Kompetenz und damit hohes Potential für weitere Lernprozesse besitzen. Für eine stärkenorientierte Förderung sind in der pädagogischen Praxis daher zusätzliche Identifikationsinstrumente nötig, mit denen auch Schülerinnen und Schülern erkannt werden, die zwar ein hohes Potential besitzen aber keine guten Noten vorweisen. Hier empfiehlt sich – ähnlich wie bereits in der Förderstrategie empfohlen –, auch darauf zu achten, welches Interessenprofil Jugendliche besitzen, welche Stärken Eltern von ihren Kindern berichten und wie sich die Jugendlichen im Unterricht sowie in außerunterrichtlichen Lernumgebungen zu Naturwissenschaften, z. B. Schülerlaboren

oder Wettbewerben, verhalten.

Auch an ihrer kognitiven Grundfähigkeit sind sie – basierend auf den Ergebnissen des hier verwendeten Proxys – nur bedingt erkennbar. Dies gibt Anlass dazu, Identifikationsinstrumente, die lediglich auf der kognitiven Grundfähigkeit bzw. Intelligenz beruhen, kritisch zu hinterfragen.

Neben individuellen Merkmalen beschreibt die Arbeit auch institutionelle Kontexte, in denen leistungsstarke Jugendliche gefunden werden können. Weder in der Förderstrategie (KMK, 2015) noch in der gemeinsamen Initiative von Bund und Ländern zur Förderung leistungsstarker Jugendlicher (KMK, 2016) wird eine Schulart speziell in den Fokus genommen und vielmehr eine Förderung in allen Schularten angestrebt, ohne Größenordnungen zu nennen. Die Ergebnisse der Arbeit zeigen, dass sich ein Großteil der leistungsstarken Jugendlichen in Naturwissenschaften am Gymnasium befindet und aber auch die Realschulen beträchtliche Anteile aufweisen. Hervorzuheben ist, dass in einer Gymnasialklasse durchschnittlich knapp die Hälfte der Schülerinnen und Schüler zu den besten 20% in Naturwissenschaften in Deutschland gehören und Zielgruppe für eine Förderung leistungsstarker Jugendlicher in Naturwissenschaften sein könnten.

Förderung der naturwissenschaftlichen Kompetenz leistungsstarker Jugendlicher

Neben Schlussfolgerungen für die Identifikation liefern die Ergebnisse dieser Arbeit auch Hinweise für eine Förderung leistungsstarker Jugendlicher im Hinblick auf naturwissenschaftliche Kompetenz. Im institutionellen Bereich liegt ein Ansatzpunkt zunächst im Übertritt bzw. dem Zugang zu am Gymnasium vorhandenen Lernmöglichkeiten für leistungsstarke Jugendliche. Wie die Ergebnisse in dieser Arbeit zeigen, besucht ein Fünftel der leistungsstarken Jugendlichen in Naturwissenschaft nicht das Gymnasium. Studien argumentieren jedoch, dass Jugendliche mit vergleichbaren Leistungsvoraussetzungen an Gymnasien höhere Lernprozesse erzielen (Baumert & Köller, 1998b). Hier könnte die pädagogische Praxis und Bildungspolitik möglicherweise durch die Gestaltung des Übertritts auf ein Gymnasium bzw. entsprechenden Bildungsgang oder alternativ der Sicherstellung einer angemessenen Förderung in einer anderen Schulart weitere Verbesserungen erzielen. Im Hinblick auf die Klassenkomposition finden sich dabei in den Analysen keine Hinweise darauf, dass für die Förderung leistungsstarker Jugendlicher am Gymnasium homogene Klassen notwendig sind bzw. unterstützend wirken. Schulklassen mit einem hohen Anteil leistungsstarker Jugendlicher waren vielmehr weder auffällig

homogen noch auffällig heterogen.

Im Bereich des Naturwissenschaftsunterrichts ist aufbauend auf der Lehr-Lernforschung davon auszugehen, dass die Unterrichtsaktivitäten „interaktives Lehren und Lernen“, „forschend-entdeckender Unterricht“ und „Anwendungen“ die naturwissenschaftliche Kompetenz leistungsstarker Jugendlicher fördern. In den Analysen zeigte sich jedoch, dass die wahrgenommene Häufigkeit dieser Unterrichtsaktivitäten keinen Zusammenhang zur naturwissenschaftlichen Kompetenz leistungsstarker Schülerinnen und Schüler hatte. Es ist davon auszugehen, dass leistungsstarke Jugendliche möglicherweise unabhängig von der Häufigkeit einzelner Unterrichtsaktivitäten erfolgreich lernen und ihr hohes Vorwissen sie von unterschiedlichen Unterrichtsaktivitäten profitieren lässt.

Förderung von Selbstkonzept und Interesse leistungsstarker Jugendlicher in Naturwissenschaft

Ein dritter Bereich der Schlussfolgerungen für die pädagogische und politische Praxis betrifft das Selbstkonzept und Interesse der leistungsstarken Jugendlichen.

In dieser Arbeit wurden innerhalb der leistungsstarken Jugendlichen vier unterschiedliche Selbstkonzept-Interesse-Typen identifiziert. Darunter befand sich ein beträchtlicher Teil an leistungsstarken Jugendlichen, die ihre Fähigkeiten im Vergleich zu ihrer Leistung unterschätzen und sich nur wenig für Naturwissenschaften interessierten. Rund ein Zehntel der Jungen und ein Fünftel der Mädchen gehörte einem Selbstkonzept-Interesse-Typ an, der sich sowohl durch ein niedriges Selbstkonzept als auch ein niedriges Interesse auszeichnete. Hier wurde erneut ein Problemfeld aufgezeigt, auf das bereits vorherige Studien hingewiesen haben (vgl. Prenzel & Schütte, 2007; Seidel, 2006). Dieser Befund ist besorgniserregend, denn auf Basis bestehender Forschungsergebnisse zur Bildungs- und Berufswahl (Taskinen et al., 2013) besteht die Gefahr, dass diese Jugendlichen ungeachtet ihres hohen Potentials in Naturwissenschaften keine weitere naturwissenschaftliche Ausbildung bzw. Berufsbildung anstreben und beispielsweise naturwissenschaftliche Kurse bzw. Fächer in der Oberstufe abwählen. Hier bedarf es weiterer Anstrengungen, das Potential dieser Jugendlichen auch motivational-affektiv auszuschöpfen. Dies kann sowohl im als auch außerhalb des Unterrichts erfolgen.

Für die motivational-affektive Förderung leistungsstarker Jugendlicher im Naturwissenschaftsunterricht verdeutlichen die Ergebnisse der Arbeit die wichtige Rolle

der motivationsunterstützenden Lehr-Lernbedingungen „Autonomieunterstützung“, „Kompetenzunterstützung“ und „inhaltliche Relevanz“. In den Analysen in dieser Arbeit zeigte sich ein Zusammenhang zwischen Selbstkonzept-Interesse-Typen und motivationsunterstützenden Lehr-Lernbedingungen, der andeutet, dass eine höhere Autonomieunterstützung und inhaltliche Relevanz möglicherweise ein höheres Interesse bei niedrigem Selbstkonzept unterstützt. Darüber hinaus ist es auf Basis der Ergebnisse für die Erhöhung des Selbstkonzepts von High Performern empfehlenswert, die wahrgenommene Kompetenzunterstützung von Schülerinnen und Schülern im Unterricht zu fördern.

Es zeigte sich darüber hinaus, dass häufigere forschend-entdeckende Unterrichtselemente über Klassen hinweg mit mehr Autonomieunterstützung, Kompetenzunterstützung und inhaltlicher Relevanz assoziiert sind. Für die pädagogische Praxis bedeutet dies, dass Lehrkräfte häufigere Schülerexperimente dazu nutzen könnten, die wahrgenommene Autonomieunterstützung, Kompetenzunterstützung und inhaltliche Relevanz unter leistungsstarken Jugendlichen zu erhöhen und dadurch ihre Motivation zu fördern. Gleichzeitig waren die Unterschiede in der Unterrichtswahrnehmung von High Performern unterschiedlicher Selbstkonzept-Interesse-Typen innerhalb der Klassen unabhängig von der Häufigkeit forschend-entdeckenden Unterrichts. Insgesamt legen die Befunde damit für den Unterricht nahe, dass motivationsunterstützende Lehr-Lernbedingungen bei High Performern möglicherweise Selbstkonzept und Interesse fördern sowie forschend-entdeckender Unterricht wiederum diese Lehr-Lernbedingungen über Klassen hinweg unterstützt.

Da leistungsstarke Mädchen noch häufiger als Jungen motivational-affektiv indifferent gegenüber Naturwissenschaft bleiben, sollte eine Förderung von Selbstkonzept und Interesse leistungsstarker Jugendlicher auch gezielt bei Mädchen ansetzen. Dies kann beispielsweise dadurch angestrebt werden, dass Anwendungsbereiche naturwissenschaftlicher Phänomene an für Mädchen und Jungen interessanten Themen orientiert werden (Häussler, Bündler, Duit, Gräber & Mayer Jürgen, 1998). Dies kann im Unterricht und außerunterrichtlichen Aktivitäten beispielsweise erreicht werden, indem nicht nur technische Anwendungsbezüge (z. B. Pumpton), sondern auch gesellschaftlich relevante (z. B. Radioaktivität, Energieerzeugung) und humanbiologische Anwendungsbezüge (z. B. Sehen, Ergonomie, Hebelgesetz am Beispiel der Armmuskulatur) behandelt werden.

Auch Lehrkräfte sollten daher darin unterstützt werden, unterschiedliche moti-

vational-affektive Typen unter leistungsstarken Jugendlichen zu erkennen. Neben ihrer Diagnosekompetenz sollten sie darin gefördert werden, ihren Naturwissenschaftsunterricht und insbesondere forschend-entdeckende Lernumgebungen kognitiv aktivierend und motivationsunterstützend zu gestalten.

Auch außerunterrichtliche Fördermaßnahmen wie Schülerlabore oder Wettbewerbe eignen sich, das Selbstkonzept und Interesse von leistungsstarken Jugendlichen zu fördern. Ausgehend vom in den Analysen gefundenen Zusammenhang forschend-entdeckenden Unterrichts und motivationsunterstützender Lehr-Lernbedingungen bei leistungsstarken Jugendlichen, ist anzunehmen, dass die forschend-entdeckenden Lernmöglichkeiten in Schülerlaboren die Interessenentwicklung leistungsstarker Jugendlicher unterstützen könnten. Insbesondere in Unterrichtsthemen eingebettet, können Schülerlabore Interesse wecken und verstärken und bieten zahlreiche Möglichkeiten für Erfolgserlebnisse in der aktiven Auseinandersetzung mit naturwissenschaftlichen Inhalten (Engeln, 2004; Guderian & Priemer, 2008). Wettbewerbe sind traditionell Teil der Maßnahmen zur Förderung leistungsstarker Jugendlicher und sind auch weiterhin bildungspolitisch vorgesehen (vgl. KMK, 2016). Auch sie ermöglichen häufig eine forschend-entdeckende Auseinandersetzung mit naturwissenschaftlichen Themen und fördern damit ausgehend von den Ergebnissen dieser Arbeit sehr wahrscheinlich auch eine höhere Autonomieunterstützung, Kompetenzunterstützung und inhaltliche Relevanz.

8.4. Ausblick

Die Ergebnisse dieser Arbeit werfen verschiedene Anschlussfragestellungen auf, die sich dadurch kennzeichnen lassen, dass sie entweder das vorhandene Wissen in Richtung Erklärungs- oder Veränderungswissen erweitern oder weitere Merkmale im Bereich der Voraussetzungen, Prozesse und Ergebnisse von Bildung betrachten. Nachfolgend werden mögliche Fragestellungen im Bereich (a) der Definition und Identifikation leistungsstarker Jugendlicher, (b) ihrer motivational-affektiven Merkmale und zukünftigen Berufswege sowie (c) der Förderung leistungsstarker Jugendlicher diskutiert.

Ein erster Bereich für Anschlussfragestellungen betrifft die Definition hoher naturwissenschaftlicher Kompetenz und ihre Vorhersagekraft. In aktuellen Förderstrategien werden leistungsstarke Jugendliche relativ offen definiert (KMK, 2015) und weitere Forschung zu Vorgehensweisen und einer wissenschaftlichen Fundierung unterstützt (KMK, 2016). In dieser Arbeit wurde mit den besten 20% im Kompetenztest in Naturwissenschaften in PISA eine mögliche Definition gewählt und charakterisiert. Im Anschluss daran stellt sich die Frage, in welchem Verhältnis diese Definition zu alternativen Definitionen leistungsstarker Jugendlicher steht und inwieweit eine in dieser Arbeit definierte hohe Leistung in Naturwissenschaften zukünftige Lernprozesse und Leistungen vorher sagen kann. Erkenntnisse darüber, welche Abschlussnote, welchen Berufserfolg oder welche spätere Kompetenz dieser Jugendlichen erreichen, würden eine differenzierte Einschätzung dieser Definition erlauben und wichtige Hinweise darauf geben, wie leistungsstarke Jugendliche in der Mittelstufe identifiziert werden könnten. Dies könnte untersucht werden, indem die in PISA identifizierten leistungsstarken Jugendlichen zu einem späteren Zeitpunkt erneut befragt und getestet werden. Ein solcher zweiter Messzeitpunkt für ein Oversample leistungsstarker Jugendlichen wäre beispielsweise eingebettet in eine Panelstudie wie dem Nationalen Bildungspanel (NEPS) möglich.

Ausgehend von den Ergebnissen zu besonderen Merkmalen leistungsstarker Jugendlicher wie einer hoher mathematischen Kompetenz oder dem sozioökonomischen Hintergrund stellt sich zudem die Frage, inwieweit diese Merkmale bereits in einer früheren Alterstufe vorhersagen bzw. kausal erklären können, ob Jugendliche in der neunten Klasse eine hohe naturwissenschaftliche Kompetenz erreichen. Auch inwieweit die leistungsstarken Jugendlichen in der neunten Klasse schon seit mehreren Jahrgangsstufen die Gruppe der leistungsstarken Jugendlichen bilden oder Verschiebungen erfolgen, ist

von Interesse. An Einflussfaktoren, die dort identifiziert werden, könnte die bildungspolitische und pädagogische Praxis bereits in früheren Jahrgangsstufen konkret mit Fördermaßnahmen ansetzen. Für eine solche Untersuchung wäre ein zweiter Messzeitpunkt in einer früheren Jahrgangsstufe nötig. Dies könnte beispielsweise erfolgen, indem leistungsstarke Jugendliche, die in der vierten Jahrgangsstufe in TIMSS/IGLU teilgenommen haben, als Oversample in der neunten Jahrgangsstufe in PISA aufgenommen werden. Auch ein Oversample, das auf Teilnehmer des Nationalen Bildungspanels (NEPS) zurückgeht, ist denkbar. Bei beiden Varianten müsste durch entsprechende Testaufgaben und Skalierungsmaßnahmen sichergestellt werden, dass die Leistungsmaße zu beiden Messzeitpunkten verglichen werden können.

Des Weiteren stellt sich die Frage, was die hohe Kompetenz der leistungsstarken Jugendlichen im Vergleich zu niedrigeren Kompetenzniveaus im Detail auszeichnet und wie die Lösungsstrategien und kognitiven Prozesse dieser Jugendlichen näher beschrieben werden können. Dies wäre sowohl in der Identifikation leistungsstarker Jugendlicher als auch in der gezielten Förderung hin zu hoher Kompetenz in Naturwissenschaften nutzbar. Eine genauere Beschreibung der mit der hohen naturwissenschaftlichen Kompetenz verbundenen Lösungsstrategien könnten mit detaillierten Analysen der offenen Antworten und des Vorgehens der Jugendlichen in den interaktiven (Simulations-)Naturwissenschaftsaufgaben in PISA generiert werden. Obwohl Vorgehensweisen in den interaktiven Aufgaben und Lösungswege bei offenen Antworten bereits mit den Daten von PISA 2015 untersucht werden können, wäre für eine detaillierte Analyse der Kompetenz leistungsstarker Jugendlicher eine Erweiterungsstichprobe, in die vermehrt leistungsstarke Jugendliche oder beispielsweise Gymnasialklassen aufgenommen werden, sinnvoll. Dort könnte mit einem eigenen Testheft eine höhere Anzahl schwieriger Aufgaben in den Test aufgenommen oder ein adaptives Testvorgehen implementiert und so im oberen Leistungsbereich besser differenziert werden. Auch zusätzlich eingesetzte Testformate wie *Concept Maps* oder Lautes Denken an einem zweiten Testtag wären in einer solchen Erweiterungsstichprobe einsetzbar und könnten dazu genutzt werden, die Fähigkeiten der leistungsstarken Jugendlichen in Naturwissenschaften genauer zu beschreiben und zu identifizieren, wie sich ihre Lösungen von denen leistungsschwächerer Jugendlicher unterscheiden.

Ein zweiter Bereich möglicher Anschlussfragestellungen betrifft die Untersuchung weiterer motivational-affektiver Merkmale sowie zukünftige Bildungs- und Be-

rufswege leistungsstarker Jugendlicher. Neben den hier untersuchten Merkmalen Selbstkonzept und Interesse existieren zahlreiche weitere motivational-affektive Merkmale wie die Selbstwirksamkeit, die zukunftsorientierte oder die instrumentelle Motivation in Naturwissenschaften und wurden in der Mehrzahl auch in PISA 2006 und PISA 2015 erfasst (vgl. Frey & Asseburg, 2009; Schiepe-Tiska, Roczen, Müller, Prenzel & Osborne, 2016). In weiteren Forschungsfragen könnte untersucht werden, welche Merkmalsausprägungen leistungsstarke Jugendliche allgemein und Jugendliche aus verschiedenen Selbstkonzept-Interesse-Typen darin haben. Auch wie diese weiteren motivational-affektiven Merkmale leistungsstarker Jugendlicher mit Unterrichtsmerkmalen zusammenhängen wurde bisher nicht untersucht. Beide Fragestellungen könnten mit deskriptiven bzw. mehrebenenanalytischen Untersuchungen der Datensätze aus PISA 2006 und 2015 beantwortet werden.

Im Kontext der hier untersuchten motivational-affektiven Merkmale Selbstkonzept und Interesse stellt sich zudem die Frage, welche Bildungs- und Berufswege leistungsstarke Jugendliche einschlagen möchten und inwiefern dies durch Selbstkonzept und Interesse oder andere motivational-affektive Merkmale vorhergesagt wird. Wissen darüber, inwiefern leistungsstarke Jugendliche unterschiedlicher Selbstkonzept-Interesse-Typen naturwissenschaftliche Karrierewege einschlagen würde weitere Informationen zur Validität der gefundenen Selbstkonzept-Interesse-Typen liefern und die bekannten Erkenntnisse zum Zusammenhang individueller Lernvoraussetzungen mit abhängigen Variablen wie Lernstrategien, Lernmotivation und Unterrichtswahrnehmung (Jurik et al., 2013, 2014) ergänzen. Der Zusammenhang der Selbstkonzept-Interesse-Typen mit einem Berufswunsch im Bereich Naturwissenschaften in der neunten Klasse kann bereits mit bestehenden Daten aus PISA 2006 und PISA 2015 untersucht werden. Die entscheidenden Erkenntnisse würde jedoch ein zweiter Messzeitpunkt ermöglichen, in dem leistungsstarke Jugendliche, die an einem PISA Durchgang mit Hauptdomäne Naturwissenschaften teilgenommen haben, ca. 5 Jahre später zu ihrer Studien- und Ausbildungswahl oder ca. 15 Jahre später zu ihrem Beruf erneut befragt werden würden. Die Studie von Parker, Marsh, Ciarrochi, Marshall und Abduljabbar (2014) zum Einfluss von Selbstwirksamkeit und Selbstkonzept für zukünftige Bildungswege 15-Jähriger in Australien wäre ein Beispiel für ein solches Vorgehen. Neben motivational-affektiven Merkmalen könnten auch Variablen wie das Kompetenzprofil leistungsstarker Jugendlicher, außerschulische Aktivitäten oder ihr sozioökonomische Hintergrund in derarti-

ge Untersuchungen aufgenommen werden und möglicherweise vorhersagen, welchen Ausbildungs- und Karriereweg leistungsstarke Jugendliche wählen.

Für den dritten Themenkomplex zu den Unterrichtszusammenhängen lassen sich ebenfalls weiterführende Fragen aus dieser Arbeit ableiten. Zu jeder der drei in dieser Arbeit betrachteten Unterrichtsaktivitäten lassen sich ein oder zwei Gestaltungsmerkmale nennen, deren Betrachtung die Ergebnisse dieser Arbeit vertiefen und Erklärungswissen generieren könnten. Daneben stellt sich bei allen drei Unterrichtsaktivitäten die Frage nach dem Veränderungswissen darüber, inwieweit durch Interventionen mit oder zu diesen Merkmalen eine höhere Leistung oder ein höheres Interesse bei leistungsstarken Jugendlichen erreicht werden könnte.

Wichtige Gestaltungskriterien beim „interaktiven Lehren und Lernen“ sind die Qualität der Lehrer-Schüler-Interaktion, beispielsweise im Hinblick auf Fragen und Feedback der Lehrperson (vgl. Jurik et al., 2014), sowie die tatsächlich kooperativ und kognitiv aktivierende Gestaltung von Lernprozessen in heterogenen Lerngruppen (vgl. Slavin, 1987, 1991). Bei beiden wurde für leistungsstarke Jugendliche noch nicht abschließend untersucht, inwieweit diese Aspekte beeinflussen, wie effektiv interaktiver Unterricht ihre Leistung und ihre Interesse fördert. Fragen nach Gestaltungskriterien von Aufgabenstellungen, Beurteilungsvarianten, Feedbackformen und Lernbegleitung könnten hier weitere wichtige Erkenntnisse dazu liefern, wie kooperativer Unterricht für leistungsstarke Jugendliche gestaltet sein sollte.

Auch für die zweite betrachtete Unterrichtsaktivität des forschend-entdeckenden Lernens („Experimentieren“ und „Schülerinnen und Schüler forschen“) schließen sich verschiedene Fragen im Bereich der Gestaltung an die Ergebnisse dieser Arbeit an. Ausgehend von theoretischen Überlegungen zu Basismodellen von Unterricht (Oser & Patry, 1990) und Ergebnissen der IPN-Videostudie zum Physikunterricht (Seidel, Prenzel, Rimmel et al., 2007) können als wichtige Gestaltungsbedingungen für forschend-entdeckendes Lernen angenommen werden, dass es eingebettet in den Unterricht erfolgt und neben dem Durchführen von Experimenten auch das Generieren von Fragestellungen und Hypothesen sowie die Interpretation der Ergebnisse enthält. Eine an die Ergebnisse der Arbeit anschließende Fragestellung ist daher, welchen Einfluss die dahingehende Gestaltung von forschend-entdeckendem Unterricht auf die Leistung und das Interesse leistungsstarker Jugendlicher hat und inwieweit Unterschiede zu den anderen Jugendlichen bestehen.

Dabei stellt sich anschließend an die Ergebnisse zum positiven Zusammenhang von häufigerem forschend-entdeckendem Unterricht und wahrgenommenen motivationsunterstützenden Bedingungen in dieser Arbeit die Frage, ob der Einfluss einzelner Gestaltungsmerkmale auf die Leistung und das Interesse leistungsstarker Jugendlicher von den wahrgenommenen motivationsunterstützenden Lehr-Lernbedingungen mediiert wird (vgl. auch Furtak & Kunter, 2012). Dies würde den Wirkmechanismus von forschend-entdeckendem Unterricht auf die Kompetenz und das Interesse von leistungsstarken Jugendlichen näher erklären. Forschend-entdeckende Lernumgebungen, mit denen leistungsstarke Jugendliche im Naturwissenschaftsunterricht oder auch in außerschulischen Lernorten wie Schülerlaboren gefördert werden sollen, könnten in Anlehnung an diese Ergebnisse gestaltet und verbessert werden.

Bei Anwendungsbezügen im Naturwissenschaftsunterricht liegen mögliche Anschlussfragen darin, dass verschiedene Anwendungsbereiche differenziert untersucht werden. Mädchen und Jungen interessieren sich im Physikunterricht unterschiedlich für technisch-physikalische, gesellschaftliche und biologische Anwendungsbereiche (vgl. Häussler, Hoffmann et al., 1998). Bisher ist unklar, inwieweit sich unter den leistungsstarken Jugendlichen in Naturwissenschaften ein ähnliches Bild zeigt und verschiedene Anwendungsbereiche das Interesse weiblicher und männlicher leistungsstarker Jugendlicher unterschiedlich fördern. In sowohl querschnittlichen oder längsschnittlichen Fragebogendesigns könnte untersucht werden, wie verschiedene Anwendungsbereiche bei High Performern mit der inhaltlichen Relevanz sowie unterschiedlichen motivational-affektiven Typen zusammenhängen und, ob das Geschlecht als Moderator auftritt.

Die genannten Fragestellungen zu Gestaltungsmerkmalen interaktiven Lehren und Lernens, forschend-entdeckenden Unterrichts und Anwendungsbezügen könnten unter anderem mithilfe einer Erweiterungsstichprobe in PISA untersucht werden, für die zusätzlich ein zweiter Messzeitpunkt sowie eine Erfassung des Unterrichts mithilfe von Unterrichtsvideos in die Studie implementiert wird. Der zweite Messzeitpunkt wäre dabei am Ende der zehnten Klasse möglich. Alternativ könnte die Erweiterungsstichprobe jedoch auch bereits ein Jahr vor der PISA-Testung in der achten Klasse als ersten Messzeitpunkt getestet werden. Dies hätte im Vergleich zu bisherigen Untersuchungen (z. B. Kuger, Klieme, Lüdtke, Schiepe-Tiska & Reiss, 2017) den Vorteil, dass die Unterrichtserfassung in der neunten Klasse sich genau auf das Jahr zwischen beiden Messzeitpunkten beziehen würde. Alternativ könnte der Unterricht auch in der zeh-

ten Jahrgangsstufe erneut erfasst werden. Die Erfassung des Unterrichts mit Videos zwischen den beiden Messzeitpunkten würde erlauben, (a) Gestaltungsmerkmale wie die Lehrer-Schüler-Interaktion oder unterschiedliche Phasen forschend-entdeckenden Unterrichts verlässlich beobachten und einschätzen zu können, (b) sie mit Unterrichtseinschätzungen der leistungsstarken Schülerinnen und Schüler in Verbindung zu bringen sowie (c) deren Verhalten zu messen. Darüber hinaus könnten die Selbstkonzept-Interesse-Kombinationen, die in dieser Arbeit gefunden wurden, den Schülerinnen und Schülern in den Videos zugeordnet werden und Zusammenhänge zu ihrem Verhalten in interaktiven aber auch forschend-entdeckenden Lernumgebungen untersucht werden. Die beiden Messzeitpunkte würden erlauben, die in Videos und Fragebögen erfassten Gestaltungskriterien von Unterricht sowie Verhaltensmerkmale mit Veränderungen in der Leistung oder dem Interesse in Zusammenhang zu bringen. In Metaanalysen könnte zudem näher untersucht werden, inwieweit diese Gestaltungsmerkmale in bisherigen Studien zu kooperativem Lernen und forschend-entdeckendem Lernen bei leistungsstarken Jugendlichen erfüllt wurden. Unterschiede in den Designs können möglicherweise bestehende Widersprüche in den Ergebnissen auflösen (vgl. Neber et al., 2001).

Im Anschluss daran stellt sich die Frage, wie Lehrkräfte darin unterstützt werden können, ihren Naturwissenschaftsunterricht förderlich für leistungsstarke Jugendliche zu gestalten. Veränderungswissen dazu könnte aus Interventionsstudien erlangt werden, in denen Lehrkräfte Fortbildungen zur Gestaltung interaktiver oder forschend-entdeckender Lernumgebungen für heterogene Lerngruppen erhalten. In Studien wurde bereits gezeigt, dass Fortbildungsmaßnahmen den interaktiven Unterricht von Lehrkräften verbessern und das Lernverhalten sowie die Lernmotivation beeinflussen (Kiemer, Gröschner, Pehmer & Seidel, 2015; Pehmer, Gröschner & Seidel, 2015). Eine Untersuchung der Leistungsentwicklung und Motivation leistungsstarker Jugendlicher in solchen Interventionsstudien könnte weiteres Veränderungswissen dazu liefern, durch welche Maßnahmen die Förderung leistungsstarker Jugendlicher im interaktiven Unterricht verbessert werden könnte.

Schließlich wurden in dieser Arbeit mögliche außerschulische Einflussfaktoren und informelle Lernumgebungen wie eine frühe Förderung im Elternhaus und der Besuch von Schülerlaboren, naturwissenschaftlichen Freizeitaktivitäten etc. nicht berücksichtigt. Hier bleiben weitere Forschungsfragen dazu offen, welchen Zusammenhang bzw. welchen Einfluss sie auf das Erreichen hoher Leistung in Naturwissenschaften haben.

Auch hier bieten sich Large-Scale-Assessments für eine beschreibende Untersuchung und Paneldesigns für kausale Analysen auf Basis eines quasi-experimentellen Designs an.

Ein weiterer Forschungsbereich, der in dieser Arbeit nicht berücksichtigt wurde, sind Merkmale der Lehrperson. Die Persönlichkeitsmerkmale und die Expertise und Einstellung einer Lehrperson in Bezug auf ihren Unterricht sind zentrale Einflussfaktoren auf die Effektivität von Unterricht (Shulman, 1986, 2001). Mit der Diagnosekompetenz, der Lehrer-Schüler-Interaktion und dem Vorschlag, Veränderungswissen in Fortbildungen von Lehrkräften zu generieren, wurde die Rolle der Lehrkräfte bereits an einigen Stellen angesprochen. In diesem Bereich existieren jedoch zahlreiche weitere offene Fragen zur Förderung von Leistung und Interesse leistungsstarker Schülerinnen und Schüler. Auch im Hinblick auf die Förderinitiative der Kultusministerkonferenz (KMK, 2015) stellt sich beispielsweise die Frage, welche Einstellung Lehrkräfte an Gymnasien zur Förderung leistungsstarker Jugendlicher haben, welche Ziele sie verfolgen und welche Best-Practice-Ansätze sie nutzen. Dies könnte deskriptiv in Surveys, beispielsweise auch durch Skalen im Lehrerfragebogen in PISA, oder mithilfe von Interviews untersucht werden.

Literaturverzeichnis

- Adams, R. J. (2009). *PISA 2006 technical report*. Paris: OECD.
- Aebli, H. (1980). *Denken: Das Ordnen des Tuns: Band I*. Stuttgart: Klett-Cotta.
- Aebli, H. (1981). *Denken: Das Ordnen des Tuns: Band II*. Stuttgart: Klett-Cotta.
- Aebli, H. (2006). *Zwölf Grundformen des Lehrens: Eine allgemeine Didaktik auf psychologischer Grundlage ; Medien und Inhalte didaktischer Kommunikation, der Lernzyklus* (13. Aufl.). Stuttgart: Klett-Cotta.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principles. *IEEE Transactions on Automatic Control*, 267–281.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 716–723.
- Alfieri, L., Brooks, P. J., Aldrich, N. J. & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, 103 (1), 1.
- Amelang, M. & Bartussek, D. (1990). *Differentielle Psychologie und Persönlichkeitsforschung* (3. Aufl.). Stuttgart: Kohlhammer.
- Ames, C. & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80 (3), 260.
- Anderman, E. M. & Maehr, M. L. (1994). Motivation and Schooling in the Middle Grades. *Review of Educational Research*, 64 (2), 287. doi: 10.2307/1170696
- Andersen, L. & Chen, J. A. (2016). Do High-Ability Students Disidentify with Science? A Descriptive Study of U.S. Ninth Graders in 2009. *Science Education*, 100 (1), 57–77.
- Anderson, C., Kim J. & Keller, B. (2014). Multilevel modeling of categorical response variables. In M. von Davier & L. Rutkowski (Hrsg.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (S. 481–519). New York: Taylor & Francis Group.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89 (4), 369.

- Andersson-Bakken, E. & Klette, K. (2016). Teachers' use of questions and responses to students' contributions during whole class discussions: Comparing language arts and science classrooms. In K. Klette, O. K. Bergem & A. Roe (Hrsg.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (Bd. 12, S. 63–85). Cham: Springer.
- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B. & Wong, B. (2012). Science aspirations, capital, and family habitus: How families shape children's engagement and identification with science. *American Educational Research Journal*, 49 (5), 881–908.
- Asendorpf, J. B. & Neyer, F. J. (2012). *Psychologie der Persönlichkeit: Mit 110 Tabellen* (5. Aufl.). Berlin: Springer. doi: 10.1007/978-3-642-30264-0
- Asparouhov, T. & Muthén, B. (2012). *Multiple Group Multilevel Analysis* (Nr. 16). Zugriff am 16.06.2016 auf <https://www.statmodel.com/examples/webnotes/webnote16.pdf>
- Ausubel, D. P., Novak, J. D. & Hanesian, H. (1981). Psychologische und pädagogische Grenzen des entdeckenden Lernens. In H. Neber (Hrsg.), *Entdeckendes Lernen* (S. 30–44). Weinheim: Beltz.
- Bacher, J., Pöge, A. & Wenzig, K. (2010a). Analyse latenter Klassen für nominale, ordinale und gemischtskalierte Variablen. In J. Bacher, A. Pöge & K. Wenzig (Hrsg.), *Clusteranalyse* (S. 377–393). München: Oldenbourg.
- Bacher, J., Pöge, A. & Wenzig, K. (Hrsg.). (2010b). *Clusteranalyse: Anwendungsorientierte Einführung in Klassifikationsverfahren* (3. Aufl.). München: Oldenbourg.
- Barker, V. & Millar, R. (2000). Students' reasoning about basic chemical thermodynamics and chemical bonding: what changes occur during a context-based post-16 chemistry course? *International Journal of Science Education*, 22 (11), 1171–1200.
- Baumert, J., Bos, W., Brockmann, J., Gruehn, S., Klieme, E., Köller, O., ... Watermann, R. (2000). *TIMSS/III-Deutschland: Der Abschlussbericht: Zusammenfassung ausgewählter Ergebnisse der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur mathematischen und naturwissenschaftlichen Bildung am Ende der Schullaufbahn*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Baumert, J. & Köller, O. (1998a). Interest research in secondary level I: An overview. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender* (S. 241–256). Kiel: IPN.

- Baumert, J. & Köller, O. (1998b). Nationale und internationale Schulleistungsstudien: Was können sie leisten, wo sind ihre Grenzen? *Pädagogik*, 50 (6), 12–18. Zugriff auf http://www.mpib-berlin.mpg.de/en/institut/dok/full/Baumert/bjnuip_/bjnuip_.htm
- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., ... Neubrand, J. (1997). *TIMSS - Mathematics and science achievement in international comparison: Summary descriptive results*. Berlin: Max Planck Institute for Human Development and Education.
- Baumert, J. & Maaz, K. (2006). Das theoretische und methodische Konzept von PISA zur Erfassung sozialer und kultureller Ressourcen der Herkunftsfamilie: Internationale und nationale Rahmenkonzeption. In J. Baumert (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit: Vertiefende Analysen im Rahmen von PISA 2000* (S. 11–29). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J., Schnabel, K. U. & Lehrke, M. (1998). Learning math in school: Does interest really matter? In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender* (S. 327–336). Kiel: IPN.
- Baumert, J., Stanat, P. & Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In J. Baumert (Hrsg.), *PISA 2000* (S. 15–68). Opladen: Leske + Budrich.
- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.). (2006a). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit: Vertiefende Analysen im Rahmen von PISA 2000*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J., Stanat, P. & Watermann, R. (2006b). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit: Vertiefende Analysen im Rahmen von PISA 2000* (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J., Trautwein, U. & Artelt, C. (2003). Schulumwelten - institutionelle Bedingungen des Lehrens und Lernens. In J. Baumert et al. (Hrsg.), *PISA 2000 - Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 261–331).

- Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bayrhuber, H., Bögeholz, S., Elster, D., Hammann, M., Hössle, C., Lücken, M., ... Sandmann, A. (2007). Biologie im Kontext: Ein Programm zur Kompetenzförderung durch Kontextorientierung im Biologieunterricht und zur Unterstützung von Lehrerprofessionalisierung. *Mathematische und Naturwissenschaftliche Unterricht*, 60 (5), 282–286.
- Büchner, P. (2003). Stichwort: Bildung und soziale Ungleichheit. *Zeitschrift für Erziehungswissenschaft*, 6 (1), 5–24. doi: 10.1007/s11618-003-0002-9
- Beerman, L., Heller, K. A. & Menacher, P. (1992). *Mathe: nichts für Mädchen? Begabung und Geschlecht am Beispiel von Mathematik, Naturwissenschaft und Technik* (1. Aufl.). Bern: Huber.
- Beller, M. & Gafni, N. (1996). The 1991 International Assessment of Educational Progress in Mathematics and Sciences: The gender differences perspective. *Journal of Educational Psychology*, 88 (2), 365–377.
- Benbow, C. P., Lubinski, D. & Hyde, J. S. (1997). Mathematics: Is biology the cause of gender differences in performance? In M. R. Walsh (Hrsg.), *Women, men, & gender: Ongoing debates* (S. 271–287). New Haven, CT, US: Yale University Press.
- Bennett, J. & Lubben, F. (2006). Context-based Chemistry: The Salters approach. *International Journal of Science Education*, 28 (9), 999–1015. doi: 10.1080/09500690600702496
- Berg, C. A. R., Bergendahl, V. C. B., Lundberg, B. & Tibell, L. (2003). Benefiting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus an open-inquiry version of the same experiment. *International Journal of Science Education*, 25 (3), 351–372. doi: 10.1080/09500690210145738
- Bergman, L. & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology*, 9 (2), 291–319. doi: 10.1017/S095457949700206X
- Berland, L. K. & McNeill, K. L. (2010). A Learning Progression for Scientific Argumentation: Understanding Student Work and Designing Supportive Instructional Contexts. *Science Education*, 94 (5), 765–793.
- Blanchard, M. R., Southerland, S. A., Osborne, J. W., Sampson, V. D., Annetta, L. A. & Granger, E. M. (2010). Is inquiry possible in light of accountability?: A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory

- instruction. *Science Education*, 94 (4), 577–616.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. Kozlowski (Hrsg.), *Multilevel theory, research, and methods in organizations* (S. 349–381). San Francisco: Jossey-Bass.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. Longmans: McKay.
- Bloom, B. S. & Sosniak, L. A. (Hrsg.). (1985). *Developing talent in young people*. New York: Ballantine Books.
- Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M. & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26 (3-4), 369–398.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. S. 183-198. *Soziale Ungleichheiten. Soziale Welt. Sonderband, 2*.
- Bourdieu, P. (1998). *Die feinen Unterschiede: Kritik der gesellschaftlichen Urteilskraft* (10. Aufl., Bd. 164). Frankfurt am Main: Suhrkamp.
- Brandt, A., Möller, J. & Kohse-Höinghaus, K. (2008). Was bewirken außerschulische Experimentierlabors? *Zeitschrift für Pädagogische Psychologie*, 22 (1), 5–12. doi: 10.1024/1010-0652.22.1.5
- Bransford, J. D. (2000). *How people learn: Brain, mind, experience, and school* (9. Aufl.). Washington, DC: National Academy Press.
- Bransford, J. D. & Donovan, M. S. (2005). Scientific inquiry and how people learn. In M. S. Donovan & J. D. Bransford (Hrsg.), *How students learn: history, mathematics, and science in the classroom* (S. 397–420). Washington, DC: National Academy Press.
- Bransford, J. D. & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24, 61–100.
- Britner, S. L. (2008). Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes. *Journal of Research in Science Teaching*, 45 (8), 955–970.
- Brünken, R. & Leutner, D. (2005). Individuelle Unterschiede beim Lernen mit neuen Medien–neue Wege in der ATI-Forschung? In S. R. Schilling, J. R. Sparfeldt & C. Pruisken (Hrsg.), *Aktuelle Aspekte pädagogisch-psychologischer Forschung* (S.

- 25–40). Münster: Waxmann.
- Bromme, R. (2014). *Von der Forschung zur evidenzbasierten Entscheidung: Die Darstellung und das öffentliche Verständnis der empirischen Bildungsforschung* (Bd. 27). Wiesbaden: Springer.
- Bromme, R., Prenzel, M. & Jäger, M. (2014). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik: Eine Analyse von Anforderungen an die Darstellung, Interpretation und Rezeption empirischer Befunde. *Zeitschrift für Erziehungswissenschaft*, 17 (S4), 3–54. doi: 10.1007/s11618-014-0514-5
- Bronfenbrenner, U. (1981). *Die Ökologie der menschlichen Entwicklung: Natürliche und geplante Experimente* (1. Aufl.). Stuttgart: Klett-Cotta.
- Brophy, J. E. (1999). Toward a model of the value aspects of motivation in education: Developing appreciation for particular learning domains and activities. *Educational Psychologist*, 34 (2), 75–85. doi: 10.1207/s15326985ep3402{\textunderscore}1
- Brophy, J. E. & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, 61 (5), 365–374. doi: 10.1037/h0029908
- Brotman, J. S. & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching*, 45 (9), 971–1002. doi: 10.1002/tea.20241
- Brown, J. S., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18 (1), 32–42.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2. Aufl.). New York, NY: Guilford Press.
- Buccheri, G., Gürber, N. A. & Brühwiler, C. (2011). The Impact of Gender on Interest in Science Topics and the Choice of Scientific and Technical Vocations. *International Journal of Science Education*, 33 (1), 159–178. doi: 10.1080/09500693.2010.518643
- Bybee, R. W. & McCrae, B. (Hrsg.). (2009). *PISA science 2006: Implications for science teachers and teaching*. Arlington, Va: NSTA Press.
- Bybee, R. W. & McCrae, B. (2011). Scientific literacy and student attitudes: Perspectives from PISA 2006 science. *International Journal of Science Education*, 33 (1), 7–26. doi: 10.1080/09500693.2010.518644
- Bybee, R. W., McCrae, B. & Laurie, R. (2009). PISA 2006: An assessment of scientific literacy. *Journal of Research in Science Teaching*, 46 (8), 865–883. doi: 10.1002/tea.20333

- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge Taylor & Francis Group.
- Byrne, B. M., Shavelson, R. J. & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105 (3), 456.
- Cano, F., García, A., Berbén, A. & Justicia, F. (2014). Science learning: A path analysis of its links with reading comprehension, question-asking in class and science achievement. *International Journal of Science Education*, 36 (10), 1–23. doi: 10.1080/09500693.2013.876678
- Carstensen, C. H., Frey, A., Walter, O. & Knoll, S. (2007). Technische Grundlagen des dritten internationalen Vergleichs. In M. Prenzel et al. (Hrsg.), *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 367–390). Münster: Waxmann.
- Chan, L. K. S. (1996). Motivational Orientations and Metacognitive Abilities of Intellectually Gifted Students. *Gifted Child Quarterly*, 40 (4), 184–193. doi: 10.1177/001698629604000403
- Chase, W. G. & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4 (1), 55–81. doi: 10.1016/0010-0285(73)90004-2
- Chi, M. T., Feltovich, P. J. & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, 5 (2), 121–152.
- Chin, C. & Chia, L.-G. (2006). Problem-based learning: Using ill-structured problems in biology project work. *Science Education*, 90 (1), 44–67. doi: 10.1002/sce.20097
- Chin, C. & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, 44 (1), 1–39.
- Chiu, M.-S. (2012). The internal/external frame of reference model, big-fish-little-pond effect, and combined model for mathematics and science. *Journal of Educational Psychology*, 104 (1), 87–107. doi: 10.1037/a0025734
- Christ, O. & Schlüter, E. (2012). *Strukturgleichungsmodelle mit Mplus: Eine praktische Einführung*. München: Oldenbourg.
- Clausen, M. (2002). *Unterrichtsqualität: eine Frage der Perspektive?: Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität* (Bd. 29). Münster: Waxmann.
- Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational evaluation and policy analysis*, 12 (3), 311–329.

- Colangelo, N., Assouline, S. G., Cole, V., Cutrona, C. & Maxey, J. E. (1996). Exceptional Academic Performance. Perfect Scores on the PLAN. *Gifted Child Quarterly*, 40 (2), 102–110.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American journal of sociology*, 94, 95-S120.
- Collins, A., Brown, J. S. & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Hrsg.), *Knowing, learning, and instruction* (S. 453–494). Hillsdale, NJ: Erlbaum.
- Corno, L. & Snow, R. E. (1986). Adapting teaching to individual differences among learners. *Handbook of research on teaching*, 3, 605–629.
- Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. Oxford, England: Irvington.
- Cross, T. L. & Coleman, L. J. (1992). Gifted High School Students' Advice to Science Teachers. *Gifted Child Today (GCT)*, 15 (5), 25–26.
- Csikszentmihalyi, M. & Schiefele, U. (1993). Die Qualität des Erlebens und der Prozess des Lernens. *Zeitschrift für Pädagogik*, 39 (2), 207–221.
- Dalehefte, I. M. (2001). *Lernmotivation im Physikunterricht: Eine Videostudie zur Untersuchung motivationsunterstützender Bedingungen im Unterrichtsverlauf. Unveröffentlichte Diplomarbeit*. Kiel: IPN.
- Daniels, Z. (2008). *Entwicklung schulischer Interessen im Jugendalter: Berlin, Univ., Diss.–Potsdam, 2004* (Bd. 69). Münster: Waxmann.
- Dar, Y. & Resh, N. (1986). Classroom intellectual composition and academic achievement. *American Educational Research Journal*, 23 (3), 357–374.
- Deci, E. L. (1998). The relation of interest to motivation and human needs - The self-determination viewpoint. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender* (S. 146–162). Kiel: IPN.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39 (2), 223–238.
- Deci, E. L. & Ryan, R. M. (2000). The what and why of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11 (4), 227–268. doi:

10.1207/S15327965PLI1104{\textunderscore}01

- Deci, E. L. & Ryan, R. M. (Hrsg.). (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.
- de Jong, R. & Westerhof, K. J. (2001). The quality of student ratings of teacher behavior. *Learning Environments Research*, 4 (1), 51–85. doi: 10.1023/A:1011402608575
- Denissen, J., Zarrett, N. R. & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development*, 78 (2), 430–447. doi: 10.1111/j.1467-8624.2007.01007.x
- Dewey, J. (1913). *Interest and effort in education*. Boston: Houghton Mifflin Co.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich* (Bd. 41, S. 73–92). Weinheim: Beltz.
- Ditton, H. (2002a). Lehrkräfte und Unterricht aus Schülersicht. Ergebnisse einer Untersuchung im Fach Mathematik. *Zeitschrift für Pädagogik*, 48 (2), 262–286.
- Ditton, H. (2002b). Unterrichtsqualität: Konzeptionen, methodische Überlegungen und Perspektiven. *Unterrichtswissenschaft : Zeitschrift für Lernforschung*, 30 (3), 197–212.
- Dixon, F. A. (1998). Social and academic self-concepts of gifted adolescents. *Journal for the Education of the Gifted*, 22 (1), 80–94.
- Dochy, F., Segers, M., van den Bossche, P. & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, 13 (5), 533–568. doi: 10.1016/S0959-4752(02)00025-7
- Doll, J. & Prenzel, M. (Hrsg.). (2004a). *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung*. Münster: Waxmann.
- Doll, J. & Prenzel, M. (2004b). Das DFG-Schwerpunktprogramm 'Bildungsqualität von Schule (BIQUA): Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen'. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung* (S. 9–23). Münster: Waxmann.

- Donovan, S. & Bransford, J. (2005). *How students learn: Science in the classroom*. Washington, D.C.: National Academies Press.
- Drechsel, B., Prenzel, M. & Seidel, T. (2015). Nationale und internationale Schulleistungsstudien. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 343–368). Berlin, Heidelberg: Springer.
- Dresel, M., Ziegler, A. & Broome, P. (1998). Gender differences in science education: the double-edged role of prior knowledge in physics. *Roeper Review*, 21 (2), 101–106.
- Dörner, D. (1979). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.
- Duit, R. & Treagust, D. F. (1998). Learning in science - from behaviorism towards social constructivism and beyond. In B. J. Fraser & K. Tobin (Hrsg.), *The international handbook of science education* (S. 3–25). Dordrecht: Kluwer Academic Publishers.
- Dunkin, M. J. & Biddle, B. J. (1974). *The study of teaching*. Oxford: Holt, Rinehart & Winston.
- Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the eccles et al. model of achievement-related choices. *Psychology of Women Quarterly*, 18 (4), 585–609. doi: 10.1111/j.1471-6402.1994.tb01049.x
- Ehmke, T. & Baumert, J. (2007). Soziale Herkunft und Kompetenzerwerb: Vergleiche zwischen PISA 2000, 2003 und 2006. In M. Prenzel et al. (Hrsg.), *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 309–336). Münster: Waxmann.
- Ehmke, T., Hohensee, F., Heidemeier, H. & Prenzel, M. (2004). Soziale Herkunft. In M. Prenzel et al. (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (S. 225–254). Münster: Waxmann.
- Eijkelhof, H. & Lijnse, P. (1988). The role of research and development to improve STS education: Experiences from the PLON project. *International Journal of Science Education*, 10 (4), 464–474.
- Enders, C. K. & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12 (2), 121–138. doi: 10.1037/1082-989X.12.2.121
- Engeln, K. (2004). *Schülerlabors: authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken*. Dissertation (Bd. 36). Berlin: Logos-Verlag.

- Engeln, K. & Euler, M. (2005). Experimentieren im Schülerlabor: Forschen statt Pauken. In K. Engeln & M. Euler (Hrsg.), *Physikunterricht modernisieren* (S. 67–90). Kiel: IPN.
- Ericsson, K. A. (1996). The acquisition of expert performance. In K. A. Ericsson (Hrsg.), *The road to excellence. The acquisition of expert performance in the arts and sciences, sports, and games* (S. 1–50). Mahwah, NJ: Erlbaum.
- Ericsson, K. A., Krampe, R. T. & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100 (3), 363–406. doi: 10.1037/0033-295X.100.3.363
- Erikson, R. & Goldthorpe, J. H. (2002). Intergenerational inequality: A sociological perspective. *Journal of Economic Perspectives*, 16, 31–44.
- Erikson, R., Goldthorpe, J. H. & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, 30, 341–415.
- Etkina, E., Matilsky, T. & Lawrence, M. (2003). Pushing to the edge: Rutgers astrophysics institute motivates talented high school students. *Journal of Research in Science Teaching*, 40 (10), 958–985. doi: 10.1002/tea.10118
- Euler, M. (2001). Lernen durch Experimentieren. In U. Ringelband (Hrsg.), *Lernort Labor* (S. 13–42). Kiel: IPN.
- Fend, H. (1980). *Theorie der Schule*. München: Urban & Schwarzenberg.
- Fend, H. (2001). *Qualität im Bildungswesen: Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung* (2. Aufl.). Weinheim: Juventa-Verlag.
- Fensham, P. J. (2009). Real world contexts in PISA science: Implications for context-based science education. *Journal of Research in Science Teaching*, 46 (8), 884–896.
- Fey, A., Gräsel, C., Puhl, T. & Parchmann, I. (2004). Implementation einer kontextorientierten Unterrichtskonzeption für den Chemieunterricht. *Unterrichtswissenschaft*, 32 (3), 238–256.
- Fischer, H. E., Labudde, P., Neumann, K. & Viiri, J. (Hrsg.). (2014). *Quality of Instruction in Physics: Comparing Finland, Switzerland and Germany* (1. Aufl.). Münster: Waxmann.
- Fischer, H. E., Neumann, K., Labudde, P. & Viiri, J. (2014). Theoretical Framework. In H. E. Fischer, P. Labudde, K. Neumann & J. Viiri (Hrsg.), *Quality of Instruction in Physics* (S. 13–30). Münster: Waxmann.

- Fonseca, J. R. S. & Cardoso, M. G. (2007). Mixture-model cluster analysis using information theoretical criteria. *Intelligent Data Analysis*, 11 (2), 155–173.
- Fraser, B. J. (1998). Science learning environments: Assessment, effects and determinants. In B. J. Fraser & K. Tobin (Hrsg.), *The international handbook of science education* (S. 527–564). Dordrecht: Kluwer Academic Publishers.
- Fraser, B. J. & Walberg, H., Welch, W., & Hattie, J. (1987). Syntheses of educational productivity research. *International Journal of Educational Research*, 11, 145–427. doi: 10.1016/0883-0355(87)90035-8
- Freeman, J. (2004). Cultural Influences on Gifted Gender Achievement. *High Ability Studies*, 15 (1), 7–23.
- Frey, A. & Asseburg, R. (Hrsg.). (2009). *PISA-2006-Skalenhandbuch: Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Furtak, E. M. & Kunter, M. (2012). Effects of autonomy-supportive teaching on student learning and motivation. *Journal of Experimental Education*, 80 (3), 284–316. doi: 10.1037/t00805-000.
- Furtak, E. M., Seidel, T., Iverson, H. & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, 82 (3), 300–329. doi: 10.3102/0034654312457206
- Gagne, R. (1985). *The Conditions of Learning and Theory of Instruction Robert Gagné*. New York, NY: Holt, Rinehart and Winston.
- Gallagher, S. A., Stepien, W. J. & Rosenthal, H. (1992). The effects of problem-based learning on problem solving. *Gifted Child Quarterly*, 36 (4), 195–200.
- Ganzeboom, H. B. G., de Graaf, P. M., Treiman, D. J. & de Leeuw, J. (1992). A Standard International Socio-Economic Index of Occupational Status. *Social Science Research*, 1–56.
- Ganzeboom, H. B. G. & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research*, 25, 201–239.
- Gardner, P. L. (1998). The development of males' and females' interests in science and technology. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender* (S. 41–57). Kiel: IPN.
- Garn, A. C. & Jolly, J. L. (2014). High ability students' voice on learning motivation. *Journal*

- of Advanced Academics*, 25 (1), 7–24. doi: 10.1177/1932202X13513262
- Garner, R., Gillingham, M. G. & White, C. . S. (1989). Effects of 'Seductive Details' on Macroprocessing and Microprocessing in Adults and Children. *Cognition and Instruction*, 6 (1), 41–57. doi: 10.1207/s1532690xci0601{\textunderscore}2
- Gebhardt, M., Sälzer, C., Mang, J., Müller, K. & Prenzel, M. (2015). Performance of Students With Special Educational Needs in Germany: Findings From Programme for International Student Assessment 2012. *Journal of Cognitive Education and Psychology*, 14 (3), 343–356.
- Geiser, C. (2011). *Datenanalyse mit Mplus: Eine anwendungsorientierte Einführung* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- George, R. & Kaplan, D. (1998). A structural model of parent and teacher influences on science attitudes of eighth graders: Evidence from NELS: 88. *Science Education*, 82 (1), 93–109. doi: 10.1002/(SICI)1098-237X(199801)
- Good, T. L. & Brophy, J. E. (1990). *Educational psychology: A realistic approach* (4. Aufl.). New York and London: Longman.
- Good, T. L., Cooper, H. M. & Blakey, S. L. (1980). Classroom interaction as a function of teacher expectations, student sex, and time of year. *Journal of Educational Psychology*, 72 (3), 378–385. doi: 10.1037/0022-0663.72.3.378
- Gottfried, A. E., Fleming, J. S. & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology*, 93 (1), 3–13. doi: 10.1037//0022-0663.93.1.3
- Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W. & Oliver, P. H. (2009). A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation. *Journal of Educational Psychology*, 101 (3), 729–739. doi: 10.1037/a0015084
- Gräber, W. & Nentwig, P. (2002). Scientific Literacy - Naturwissenschaftliche Grundbildung in der Diskussion. In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Hrsg.), *Scientific Literacy* (S. 7–20). Opladen: Leske + Budrich.
- Greeno, J. G., Moore, J. L. & Smith, D. R. (1993). Transfer of situated learning. In D. K. Detterman R. J. Sternberg (Hrsg.), *Transfer on trial: Intelligence, cognition, and instruction* (S. 99–167). Westport, CT, US: Ablex Publishing.
- Gruber, H. (2007). Bedingungen von Expertise. In K. A. Heller & A. Ziegler (Hrsg.), *Begabt sein in Deutschland* (S. 93–112). Berlin: Lit Verlag.

- Gruber, H., Renkl, A. & Schneider, W. (1994). Expertise und Gedächtnisentwicklung. Längsschnittliche Befunde aus der Domäne Schach. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 26 (1), 53–70.
- Gruber, H. & Stamouli, E. (2015). Intelligenz und Vorwissen. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 25–44). Berlin, Heidelberg: Springer.
- Guay, F., Larose, S. & Boivin, M. (2004). Academic self-concept and educational attainment level: A ten-year longitudinal study. *Self and identity*, 3 (1), 53–68.
- Guderian, P. & Priemer, B. (2008). Intressenförderung durch Schülerlaborbesuche - eine Zusammenfassung der Forschung in Deutschland. *Physik und Didaktik in Schule und Hochschule* (7), 27–36.
- Gustafson, J.-E. & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner & R. C. Calfee (Hrsg.), *Handbook of educational psychology* (S. 186–242). New York: MacMillan.
- Hannover, B. (1997). Die Bedeutung des Pubertären Reifestatus für die Herausbildung informeller Interaktionsgruppen in koedukativen Klassen und in Mädchenschulklassen. *Zeitschrift für Pädagogische Psychologie*, 11 (1), 3–13.
- Hannover, B. (1998). The development of self-concept and interests. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender* (S. 105–125). Kiel: IPN.
- Hannover, B. & Kessels, U. (2002). Challenge the science-stereotype. Der Einfluss von Technik-Freizeitkursen auf das Naturwissenschaften-Stereotyp von Schülerinnen und Schülern. In M. Prenzel & J. Doll (Hrsg.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (Bd. 45, S. 341–358). Weinheim: Beltz.
- Hannover, B. & Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learning and Instruction*, 14 (1), 51–67. doi: 10.1016/j.learninstruc.2003.10.002
- Hannover, B. & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze 1 Dieser Beitrag wurde unter der geschäftsführenden Herausgeberschaft von Jens Möller angenommen. *Zeitschrift für Pädagogische Psychologie*, 25 (2), 89–103. doi: 10.1024/1010-0652/a000039
- Hardy, I., Jonen, A., Möller, K. & Stern, E. (2006). Effects of instructional support within

- constructivist learning environments for elementary school students' understanding of 'floating and sinking'. *Journal of Educational Psychology*, 98 (2), 307.
- Harlen, W. (1999). *Effective teaching of science: A review of research* (Bd. 21). Glasgow: Scottish Council for Research in Education.
- Hartinger, A. (2005). Verschiedene Formen der Öffnung von Unterricht und ihre Auswirkungen auf das Selbstbestimmungsempfinden von Grundschulkindern. *Zeitschrift für Pädagogik*, 51 (3), 397–414.
- Hasebrook, J. & Brünken, R. (2010). Aptitude-Treatment-Interaktion. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 23–29). Weinheim: Beltz.
- Hattie, J. A. C. (2010). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Haug, B. S. & Ødegaard, M. (2014). From Words to Concepts: Focusing on Word Knowledge When Teaching for Conceptual Understanding Within an Inquiry-Based Science Setting. *Research in Science Education*, 44 (5), 777–800. doi: 10.1007/s11165-014-9402-5
- Hedges, L. & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269 (5220), 41–45. doi: 10.1126/science.7604277
- Heine, J.-H., Sälzer, C., Borchert, L., Sibberns, H. & Mang, J. (2013). Technische Grundlagen des fünften internationalen Vergleichs. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012* (S. 309–346). Münster: Waxmann.
- Heller, K. A. (Hrsg.). (2002a). *Begabtenförderung im Gymnasium*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Heller, K. A. (Hrsg.). (2002b). *Begabtenförderung im Gymnasium: Ergebnisse einer zehnjährigen Längsschnittstudie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Heller, K. A. (2002c). Zum Bildungsauftrag des Gymnasiums unter besonderer Berücksichtigung der Begabtenförderung. In K. A. Heller (Hrsg.), *Begabtenförderung im Gymnasium* (S. 11–36). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Heller, K. A. (2004). Identification of gifted and talented students. *Psychology Science*, 46 (3), 302–323.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4.-12. Klassen: Revision*. Göttingen: Hogrefe.
- Helmke, A. (2007). *Unterrichtsqualität erfassen, bewerten, verbessern: Dieses Buch ist*

- Franz-Emanuel Weinert gewidmet* (6. Aufl.). Seelze: Klett Kallmeyer.
- Helmke, A. (2010). Unterrichtsqualität. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 886–895). Weinheim: Beltz.
- Helmke, A. & Schrader, F.-W. (2010). Determinanten der Schulleistung. In D. H. Rost (Hrsg.), *Handwörterbuch der Pädagogischen Psychologie* (S. 90–102). Weinheim: Beltz.
- Helmke, A. & van Aken, M. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology, 87* (4), 624–637. doi: 10.1037/0022-0663.87.4.624
- Helmke, A. & Weinert, F. E. (1997). Unterrichtsqualität und Leistungsentwicklung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 241–251). Weinheim: Beltz.
- Henderson, J. B., MacPherson, A., Osborne, J. & Wild, A. (2015). Beyond Construction: Five arguments for the role and value of critique in learning science. *International Journal of Science Education, 37* (10), 1668–1697. doi: 10.1080/09500693.2015.1043598
- Herbart, J. F. (1806). *Allgemeine Pädagogik, aus dem Zweck der Erziehung abgeleitet*. Düsseldorf: Küpper.
- Herrmann, T. (1976). *Lehrbuch der empirischen Persönlichkeitsforschung* (3. Aufl.). Göttingen: Hogrefe.
- Hidi, S. & Harackiewicz, J. M. (2000). Motivating the Academically Unmotivated: A Critical Issue for the 21st Century. *Review of Educational Research, 70* (2), 151–179. doi: 10.3102/00346543070002151
- Hidi, S. & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist, 41* (2), 111–127. doi: 10.1207/s15326985ep41024
- Hmelo-Silver, C. E., Duncan, R. G. & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42* (2), 99–107.
- Hänze, M. & Berger, R. (2007). Cooperative learning, motivational effects, and student characteristics: An experimental study comparing cooperative learning and direct instruction in 12th grade physics classes. *Learning and Instruction, 17* (1), 29–41. doi: 10.1016/j.learninstruc.2006.11.004

- Hofer, M. (2010). Adolescents' Development of Individual Interests: A product of multiple goal regulation? *Educational Psychologist*, 45 (3), 149–166. doi: 10.1080/00461520.2010.493469
- Hoffmann, L. (2002). Promoting girls' interest and achievement in physics classes for beginners. *Learning and Instruction*, 12 (4), 447–465. doi: 10.1016/S0959-4752(01)00010-X
- Hoffmann, L., Krapp, A., Renninger, K. A. & Baumert, J. (Hrsg.). (1998). *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender*. Kiel: IPN.
- Hoffmann, L., Lehrke, M. & Häussler, P. (1998). *Die IPN-Interessenstudie Physik* (Bd. 158). Kiel: IPN.
- Hofstein, A. & Kesner, M. (2006). Industrial chemistry and school chemistry: Making chemistry studies more relevant. *International Journal of Science Education*, 28 (9), 1017–1039.
- Hofstein, A. & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, 88 (1), 28–54. doi: 10.1002/sce.10106
- Hoge, R. D. & Renzulli, J. S. (1993). Exploring the link between giftedness and self-concept. *Review of Educational Research*, 63 (4), 449–465. doi: 10.3102/00346543063004449
- Hong, Z. (2010). Effects of a Collaborative Science Intervention on High Achieving Students' Learning Anxiety and Attitudes toward Science. *International Journal of Science Education*, 32 (15), 1971–1988. doi: 10.1080/09500690903229304
- Hornstra, L., van der Veen, I., Peetsma, T. & Volman, M. (2013). Developments in motivation and achievement during primary school: A longitudinal study on group-specific differences. *Learning and Individual Differences*, 23, 195–204. doi: 10.1016/j.lindif.2012.09.004
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2. Aufl.). New York, NY: Routledge.
- Huang, C. (2011). Self-concept and academic achievement: a meta-analysis of longitudinal relations. *Journal of School Psychology*, 49 (5), 505–528. doi: 10.1016/j.jsp.2011.07.001
- Huber, S. A., Häusler, J., Jurik, V. & Seidel, T. (2015). Self-underestimating students in physics instruction: Development over a school year and its connection to internal learning processes. *Learning and Individual Differences*, 43, 83–91. doi: 10.1016/

j.lindif.2015.08.021

- Hudson, H. & McIntire, W. (1977). Correlation Between Mathematical Skills and Success in Physics. *American Journal of Physics*, 45 (5), 470–471.
- Hugener, I., Pauli, C., Reusser, K., Lipowsky, F., Rakoczy, K. & Klieme, E. (2009). Teaching patterns and learning quality in Swiss and German mathematics lessons. *Learning and Instruction*, 19 (1), 66–78. doi: 10.1016/j.learninstruc.2008.02.001
- Hulleman, C. S., Godes, O., Hendricks, B. L. & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*, 102 (4), 880.
- Hulleman, C. S. & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, 326 (5958), 1410–1412.
- Häussler, P. (1987). Measuring students' interest in physics - design and results of a cross-sectional study in the Federal Republic of Germany. *International Journal of Science Education*, 9 (1), 79–92. doi: 10.1080/0950069870090109
- Häussler, P., Bündler, W., Duit, R., Gräber, W. & Mayer Jürgen. (1998). *Naturwissenschafts-didaktische Forschung - Perspektiven für die Unterrichtspraxis*. Kiel: IPN.
- Häussler, P. & Hoffmann, L. (2000). A curricular frame for physics education: Development, comparison with students' interests, and impact on students' achievement and self-concept. *Science Education*, 84 (6), 689–705. doi: 10.1002/1098-237X(200011)84:6<689::AID-SCE1>3.0.CO;2-L
- Häussler, P., Hoffmann, L., Langeheine, R., Rost, J. & Sievers, K. (1998). A typology of students' interest in physics and the distribution of gender and age within each type. *International Journal of Science Education*, 20 (2), 223–238. doi: 10.1080/0950069980200207
- Hyde, J. S., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107 (2), 139–155.
- IAB = Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit. (2015). *IAB-Jahresbericht 2015*. Nürnberg.
- Ingenkamp, K. (1995). *Die Fragwürdigkeit der Zensurengebung: Texte und Untersuchungsberichte* (9. Aufl.). Weinheim: Beltz.
- Ingenkamp, K. & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik* (6. Aufl.). Weinheim: Beltz.
- International Labour Office. (1990). *International Standard Classification of Occupations*

- ISCO-88*. Genf: ILO.
- International Labour Office. (2012). *International Standard Classification of Occupations ISCO-08*. Genf: ILO.
- Ireson, J. & Hallam, S. (2009). Academic self-concepts in adolescence: Relations with achievement and ability grouping in schools. *Learning and Instruction, 19* (3), 201–213. doi: 10.1016/j.learninstruc.2008.04.001
- IW = Institut der deutschen Wirtschaft Köln. (2016). *MINT-Frühjahrsreport 2016: Herausforderungen der Digitalisierung*. Köln.
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S. & Wigfield, A. (2002). Changes in Children's Self-Competence and Values: Gender and Domain Differences across Grades One through Twelve. *Child Development, 73* (2), 509.
- Janos, P. M. & Robinson, N. M. (1985). The performance of students in a program of radical acceleration at the university level. *Gifted Child Quarterly, 29* (4), 175–179.
- Jansen, M. (2015). *Akademische Selbstkonzepte in den naturwissenschaftlichen Fächern: Ausdifferenzierung, Geschlechtsunterschiede und Effekte dimensionaler Vergleiche* (Unveröffentlichte Dissertation). Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät, Berlin.
- Jansen, M., Scherer, R. & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology, 41*, 13–24. doi: 10.1016/j.cedpsych.2014.11.002
- Jansen, M., Schroeders, U. & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences, 30*, 11–21. doi: 10.1016/j.lindif.2013.12.003
- Jansen, M., Schroeders, U. & Stanat, P. (2013). Motivationale Schülermerkmale in Mathematik und den Naturwissenschaften. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 348–365). Münster: Waxmann.
- Johnson, D. W. & Johnson, R. T. (1989). *Cooperation and competition: Theory and research* (2. Aufl.). Edina, Minn.: Interaction Book Co.
- Johnson, D. W., Johnson, R. T. & Taylor, B. (1993). Impact of Cooperative and Individualistic Learning on High-Ability Students' Achievement, Self-Esteem, and

- Social Acceptance. *The Journal of Social Psychology*, 133 (6), 839–844. doi: 10.1080/00224545.1993.9713946
- Jurik, V., Gröschner, A. & Seidel, T. (2013). How student characteristics affect girls' and boys' verbal engagement in physics instruction. *Learning and Instruction*, 23, 33–42. doi: 10.1016/j.learninstruc.2012.09.002
- Jurik, V., Gröschner, A. & Seidel, T. (2014). Predicting students' cognitive learning activity and intrinsic learning motivation: How powerful are teacher statements, student profiles, and gender? *Learning and Individual Differences*, 32, 132–139. doi: 10.1037/t08580-000.
- Jurik, V., Häusler, J., Stubben, S. & Seidel, T. (2015). Interaction: Erste Ergebnisse einer vergleichenden Videostudie im Deutsch- und Mathematikunterricht. *Zeitschrift für Pädagogik*, 61 (5), 692–711.
- Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, 38 (1), 23–31. doi: 10.1207/S15326985EP3801{\textunderscore}4
- Kaufmann, H. & Pape, H. (1984). Clusteranalyse. In L. Fahrmeir & A. Hamerle (Hrsg.), *Multivariate statistische Verfahren* (S. 371–472). Berlin und New York: deGruyter.
- Kerschensteiner, G. (1922). Der Interessenbegriff in der Pädagogik. *Pädagogische Blätter*, 51, 349–354.
- Kessels, U. & Hannover, B. (2004). Entwicklung schulischer Intressen als Identitätsregulation. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung* (S. 398–412). Münster: Waxmann.
- Kiemer, K., Gröschner, A., Pehmer, A.-K. & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and Instruction*, 35, 94–103. doi: 10.1016/j.learninstruc.2014.10.003
- Kim, E. S., Yoon, M., Wen, Y., Luo, W. & Kwok, O.-m. (2015). Within-Level Group Factorial Invariance With Multilevel Data: Multilevel Factor Mixture and Multilevel MIMIC Models. *Structural Equation Modeling: A Multidisciplinary Journal*, 22 (4), 603–616. doi: 10.1080/10705511.2014.938217
- Kim, Y. & Muthén, B. O. (2009). Two-Part Factor Mixture Modeling: Application to an Aggressive Behavior Measurement Instrument. *Structural Equation Modeling: A*

- Multidisciplinary Journal*, 16 (4), 602–624. doi: 10.1080/10705510903203516
- King, D. & Ritchie, S. M. (2012). Learning science through real-world contexts. In B. J. Fraser, K. Tobin & C. McRobbie (Hrsg.), *Second international handbook of science education* (Bd. 24, S. 69–79). Dordrecht: Springer.
- Kircher, E., Girwidz, R. & Häußler, P. (2009). *Physikdidaktik: Theorie und Praxis*. Berlin: Springer-Verlag Berlin Heidelberg. Zugriff auf <http://dx.doi.org/10.1007/978-3-642-01602-8>
- Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41 (2), 75–86.
- Kjærnsli, M. & Lie, S. (2011). Students' Preference for Science Careers: International comparisons based on PISA 2006. *International Journal of Science Education*, 33 (1), 121–144. doi: 10.1080/09500693.2010.518642
- Klahr, D. & Dunbar, K. (1988). Dual Space Search During Scientific Reasoning. *Cognitive Science*, 12 (1), 1–48. doi: 10.1207/s15516709cog1201{\textunderscore}1
- Klahr, D. & Nigam, M. (2004). The equivalence of learning paths in early science instruction: effect of direct instruction and discovery learning. *Psychological Science*, 15 (10), 661–667. doi: 10.1111/j.0956-7976.2004.00737.x
- Klauer, K. J. (1985). Framework for a theory of teaching. *Teaching and Teacher Education*, 1 (1), 5–17.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3. Aufl.). New York: Guilford Press.
- Köller, O. (1998). *Zielorientierungen und schulisches Lernen*. Münster: Waxmann.
- Köller, O., Baumert, J. & Schnabel, K. U. (2001). Does Interest Matter? The Relationship between Academic Interest and Achievement in Mathematics. *Journal for Research in Mathematics Education*, 32 (5), 448–470. doi: 10.2307/749801
- Köller, O., Daniels, Z., Schnabel, K. U. & Baumert, J. (2000). Kurswahlen von Mädchen und Jungen im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse. *Zeitschrift für Pädagogische Psychologie*, 14 (1), 26–37. doi: 10.1024//1010-0652.14.1.26
- Köller, O., Schnabel, K. U. & Baumert, J. (2000). Der Einfluß der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse.

- Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 32 (2), 70–80. doi: 10.1026//0049-8637.32.2.70
- Köller, O., Trautwein, U., Lüdtke, O. & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für Pädagogische Psychologie*, 20 (1/2), 27–39.
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik. (2010). *Förderstrategie für leistungsschwächere Schülerinnen und Schüler: Beschluss der Kultusministerkonferenz vom 04.03.2010*. Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder. Zugriff auf http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_04_03-Foerderstrategie-Broschuere.pdf
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik. (2015). *Förderstrategie für leistungsstarke Schülerinnen und Schüler*. Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder.
- KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik. (2016). *Gemeinsame Initiative von Bund und Ländern zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler*. Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder.
- Kobarg, M., Prenzel, M., Seidel, T., Walker, M., McCrae, B., Cresswell, J. & Wittwer, J. (2011). *An international comparison of science teaching and learning: Further results from PISA 2006*. Münster: Waxmann.
- Kobarg, M. & Seidel, T. (2007). Prozessorientierte Lernbegleitung - Videoanalysen im Physikunterricht der Sekundarstufe I. *Unterrichtswissenschaft : Zeitschrift für Lernforschung*, 35 (2), 148–168.
- Krajewski, K. & Schneider, W. (2006). Mathematische Vorläuferfertigkeiten im Vorschulalter und ihre Vorhersagekraft für die Mathematikleistungen bis zum Ende der Grundschulzeit. *Psychologie in Erziehung und Unterricht*, 53, 246–262.
- Kramer, K. (2002). *Die Förderung von motivationsunterstützendem Unterricht: Ansatzpunkte und Barrieren* (Dissertation). Christian-Albrechts-Universität, Kiel.
- Krapp, A. (2002a). An educational-psychological theory of interest and its relation to self-determination theory. In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 405–427). Rochester, NY: University of Rochester Press.
- Krapp, A. (2002b). Structural and dynamic aspects of interest development: theoretical

- considerations from an ontogenetic perspective. *Learning and Instruction*, 12 (4), 383–409. doi: 10.1016/S0959-4752(01)00011-1
- Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instruction*, 15 (5), 381–395. doi: 10.1016/j.learninstruc.2005.07.007
- Krapp, A., Hidi, S. & Renninger, K. A. (1992). Interest, learning and development. In K. A. Renninger, S. Hidi & A. Krapp (Hrsg.), *The Role of interest in learning and development* (S. 3–25). Hillsdale, N.J.: L. Erlbaum Associates.
- Krapp, A. & Prenzel, M. (Hrsg.). (1992). *Interesse, Lernen, Leistung*. Münster: Aschendorff.
- Krapp, A. & Prenzel, M. (2011). Research on Interest in Science: Theories, methods, and findings. *International Journal of Science Education*, 33 (1), 27–50. doi: 10.1080/09500693.2010.518645
- Kuger, S., Klieme, E., Lüdtke, O., Schiepe-Tiska, A. & Reiss, K. (2017). Mathematikunterricht und Schülerleistung in der Sekundarstufe: Zur Validität von Schülerbefragungen in Schulleistungsstudien. *Zeitschrift für Erziehungswissenschaft (Sonderheft 33)*. doi: 10.1007/s11618-017-0750-6
- Kulik, J. A. & Kulik, C. C. (1987). Effects of ability grouping on student achievement. *Equity and Excellence in Education*, 23 (1-2), 22–30.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht* (Bd. 51). Münster: Waxmann.
- Kunter, M. & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, 9 (3), 231–251.
- Larkin, J., McDermott, J., Simon, D. P. & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208 (4450), 1335–1342.
- Lau, S. & Roeser, R. W. (2008). Cognitive abilities and motivational processes in science achievement and engagement: A person-centered analysis. *Learning and Individual Differences*, 18 (4), 497–504. doi: 10.1016/j.lindif.2007.11.002
- Lavonen, J. & Laaksonen, S. (2009). Context of teaching and learning school science in Finland: Reflections on PISA 2006 results. *Journal of Research in Science Teaching*, 46 (8), 922–944. doi: 10.1002/tea.20339
- Lazarowitz, R., Baird, J. H., Bowlden, V. & Hertz-Lazarowitz, R. (1996a). Teaching biology in a group mastery learning mode: high school students' academic achievement and affective outcomes. *International Journal of Science Education*, 18 (4), 447–462.

- doi: 10.1080/0950069960180404
- Lazarowitz, R., Baird, J. H., Bowlden, V. & Hertz-Lazarowitz, R. (1996b). Teaching biology in a group mastery learning mode: high school students' academic achievement and affective outcomes. *International Journal of Science Education*, 18 (4), 447–462. doi: 10.1080/0950069960180404
- Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, S. A. Lazarsfeld, J. A. Star & J. A. Clausen (Hrsg.), *Studies in social psychology in world war II* (S. 362–412). Princeton/N.J.: Princeton University Press.
- Lazarsfeld, P. F. & Henry, N. W. (1968). *Latent structure analysis*. New York: Houghton Mifflin.
- Lazonder, A. W. & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, 86 (3), 681–718. doi: 10.3102/0034654315627366
- Lüdtke, O., Marsh, H. W., Robitzsch, A. & Trautwein, U. (2011). A 2×2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods*, 16 (4), 444–467. doi: 10.1037/a0024376
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T. & Muthen, B. (2008). The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13 (3), 203–229. doi: 10.1037/a0012869
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, 34 (2), 120–131. doi: 10.1016/j.cedpsych.2008.12.001
- Lee, V. E. & Burkam, D. T. (1996). Gender differences in middle grade science achievement: Subject domain, ability level, and course emphasis. *Science Education*, 80 (6), 613–650.
- Lehtinen, E. (1992). Lern- und Bewältigungsstrategien im Unterricht. In H. Mandl & H. F. Friedrich (Hrsg.), *Lern- und Denkstrategien* (S. 125–149). Göttingen: Hogrefe.
- Leman, P. J., Skipper, Y., Watling, D. & Rutland, A. (2016). Conceptual Change in Science Is Facilitated Through Peer Collaboration for Boys but Not for Girls. *Child*

- Development*, 87 (1), 176–183. doi: 10.1111/cdev.12481
- Levy, J., Wubbels, T., den Brok, P. & Brekelmans, M. (2003). Students' perceptions of interpersonal aspects of the learning environment. *Learning Environments Research*, 6 (1), 5–36. doi: 10.1023/A:1022967927037
- Lewalter, D., Krapp, A., Schreyer, I. & Wild, K. P. (1998). Die Bedeutsamkeit des Erlebens von Kompetenz, Autonomie und sozialer Eingebundenheit für die Entwicklung berufsspezifischer Interessen. *Zeitschrift für Berufs-und Wirtschaftspädagogik*, 14, 143–168.
- Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction*, 11 (4), 357–380.
- Linnenbrink-Garcia, L., Patall, E. A. & Messersmith, E. E. (2013). Antecedents and Consequences of Situational Interest. *British Journal of Educational Psychology*, 83 (4), 591–614.
- Linnenbrink-Garcia, L., Pugh, K. J., Koskey, Kristin L. K. & Stewart, V. C. (2012a). Developing conceptual understanding of natural selection: The role of interest, efficacy, and basic prior knowledge. *Journal of Experimental Education*, 80 (1), 45–68. doi: 10.1080/00220973.2011.559491
- Linnenbrink-Garcia, L., Pugh, K. J., Koskey, Kristin L. K. & Stewart, V. C. (2012b). Developing conceptual understanding of natural selection: The role of interest, efficacy, and basic prior knowledge. *The Journal of Experimental Education*, 80 (1), 45–68. doi: 10.1080/00220973.2011.559491
- Lipowsky, F., Rakoczy, K., Pauli, C., Reusser, K. & Klieme, E. (2007). Gleicher Unterricht- gleiche Chancen für alle? Die Verteilung von Schülerbeiträgen im Klassenunterricht. *Unterrichtswissenschaft*, 35 (2), 125–147.
- Lohman, D. F. (2005). An Aptitude Perspective on Talent: Implications for Identification of Academically Gifted Minority Students. *Journal for the Education of the Gifted*, 28 (3/4), 333–360.
- Lonning, R. A. (1993). Effect of cooperative learning strategies on student verbal interactions and achievement during conceptual change instruction in 10th grade general science. *Journal of Research in Science Teaching*, 30 (9), 1087–1101. doi: 10.1002/tea.3660300907
- Lopez, F. G., Lent, R. W., Brown, S. D. & Gore, P. A. (1997). Role of social-cognitive expectations in high school students' mathematics-related interest and performance.

- Journal of Counseling Psychology*, 44 (1), 44–52. doi: 10.1037/0022-0167.44.1.44
- Luke, D. A. (2010). *Multilevel modeling* (Bd. 143). Thousand Oaks, Calif.: Sage Publications.
- Maaz, K., Baumert, J. & Trautwein, U. (2011). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In H.-H. Krüger, U. Rabe-Kleberg, R.-T. Kramer & J. Budde (Hrsg.), *Bildungsungleichheit revisited: Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule* (S. 69–102). Wiesbaden: VS Verlag für Sozialwissenschaften. doi: 10.1007/978-3-531-93403-7{\textunderscore}5
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23 (1), 129–149. doi: 10.3102/00028312023001129
- Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal*, 28 (2), 445–480. doi: 10.3102/00028312028002445
- Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *Zeitschrift für Pädagogische Psychologie*, 19 (3), 119–127.
- Marsh, H. W., Byrne, B. M. & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80 (3), 366–380.
- Marsh, H. W. & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective. Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, 1 (2), 133–163. doi: 10.1111/j.1745-6916.2006.00010.x
- Marsh, H. W. & Hau, K.-T. (2003). Big-Fish–Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, 58 (5), 364–376. doi: 10.1037/0003-066X.58.5.364
- Marsh, H. W., Köller, O. & Baumert, J. (2001). Reunification of East and West German school systems: Longitudinal multilevel modeling study of the big-fish-little-pond effect on academic self-concept. *American Educational Research Journal*, 38 (2), 321–350.
- Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, Alexandre J. S., Abduljabbar,

- A. S. & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist, 47* (2), 106–124. doi: 10.1080/00461520.2012.670488
- Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B. O. & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44* (6), 764–802. doi: 10.1080/00273170903333665
- Marsh, H. W., Liem, G., Martin, A. J., Morin, A. & Nagengast, B. (2011). Methodological measurement fruitfulness of exploratory structural equation modeling (ESEM): New approaches to key substantive issues in motivation and engagement. *Journal of Psychoeducational Assessment, 29* (4), 322–346. doi: 10.1177/0734282911406657
- Marsh, H. W. & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: unidimensional and multidimensional perspectives of self-concept. *Personality & social psychology bulletin, 34* (4), 542–552. doi: 10.1177/0146167207312313
- Marsh, H. W. & Rowe, K. J. (1996). The Negative Effects of School-Average Ability on Academic Self-Concept: An Application of Multilevel Modelling. *Australian Journal of Education, 40* (1), 65–87. doi: 10.1177/000494419604000105
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Development, 76* (2), 397–416. doi: 10.1111/j.1467-8624.2005.00853.x
- Marsh, H. W. & Yeung, A. S. (1997a). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89* (1), 41–54. doi: 10.1037/0022-0663.89.1.41
- Marsh, H. W. & Yeung, A. S. (1997b). Coursework Selection: Relations to Academic Self-Concept and Achievement. *American Educational Research Journal, 34* (4), 691–720. doi: 10.3102/00028312034004691
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47* (2), 149–174.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59* (1), 14–19.

- doi: 10.1037/0003-066X.59.1.14
- McCoach, D. B. & Siegle, D. (2003). The structure and function of academic self-concept in gifted and general education students. *Roeper Review*, 25 (2), 61–65. doi: 10.1080/02783190309554200
- Meece, J. L. & Holt, K. (1993). A pattern analysis of students' achievement goals. *Journal of Educational Psychology*, 85 (4), 582.
- Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality*, 60 (1), 117–174. doi: 10.1111/j.1467-6494.1992.tb00269.x
- Meier, E., Vogl, K. & Preckel, F. (2014). Motivational characteristics of students in gifted classes: The pivotal role of need for cognition. *Learning and Individual Differences*, 33, 39–46. doi: 10.1016/j.lindif.2014.04.006
- Merrill, M. D. (1987). The new Component Design Theory: Instructional design for courseware authoring. *Instructional Science*, 16 (1), 19–34. doi: 10.1007/BF00120003
- Mähler, C. & Stern, E. (2010). Transfer. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 859–869). Weinheim: Beltz.
- Mikelskis-Seifert, S. & Duit, R. (2007). Physik im Kontext. *Mathematische und Naturwissenschaftliche Unterricht*, 60 (5), 265.
- Minner, D. D., Levy, A. J. & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, 47 (4), 474–496. doi: 10.1002/tea.20347
- Miserandino, M. (1996). Children who do well in school: Individual differences in perceived competence and autonomy in above-average children. *Journal of Educational Psychology*, 88 (2), 203–214. doi: 10.1037/0022-0663.88.2.203
- Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, 85 (3), 424–436. doi: 10.1037/0022-0663.85.3.424
- Möller, J. & Köller, O. (2001). Frame of reference effects following the announcement of exam results. *Contemporary Educational Psychology*, 26 (2), 277–287.
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte: Effekte dimensionaler und sozialer Vergleiche. *Psychologische Rundschau*, 55 (1), 19–27. doi: 10.1026/0033-3042.55.1.19
- Möller, J. & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*,

120 (3), 544.

- Möller, J., Streblow, L., Pohlmann, B. & Köller, O. (2006). An extension to the internal/external frame of reference model to two verbal and numerical domains. *European Journal of Psychology of Education*, 21 (4), 467–487.
- Müller, K. & Ehmke, T. (2013). Soziale Herkunft als Bedingung der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012* (S. 245–274). Münster: Waxmann.
- Müller, K., Prenzel, M., Seidel, T., Schiepe-Tiska, A. & Kjærnsli, M. (2016). Science teaching and learning in schools – Theoretical and empirical foundations to investigate classroom level processes in the context of PISA 2015. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing context of learning world-wide– Extended Context Assessment Framework and Documentation of Questionnaire Material*. New York: Springer.
- Müller, K., Prenzel, M., Sälzer, C., Mang, J. & Gebhardt, M. (2017). Wie schneiden Schülerinnen und Schüler an Sonder- und Förderschulen bei PISA ab? Analysen aus der PISA 2012-Zusatzerhebung zu Jugendlichen mit sonderpädagogischem Förderbedarf. *Unterrichtswissenschaft*, 45 (2), 185–192.
- Mönks, F.-J. (2006). Begabung und Hochbegabung – Zum aktuellen Bestand der Begabungsforschung und Begabtenförderung. In C. Fischer (Hrsg.), *Begabtenförderung als Aufgabe und Herausforderung für die Pädagogik* (Bd. 22, S. 15–29). Münster: Aschendorff.
- Morin, A. J., Marsh, H. W., Nagengast, B. & Scalas, L. F. (2014). Doubly latent multilevel analyses of classroom climate: An illustration. *The Journal of Experimental Education*, 82 (2), 143–167. doi: 10.1080/00220973.2013.769412
- Murphy, P. & Whitelegg Elizabeth. (2006). *Girls in the physics classroom: A review of the research on the participation of girls in physics*. London.
- Muthen, B. & Muthen, L. K. (15.03.17). *Chi-Square Difference Testing Using the Satorra-Bentler Scaled Chi-Square*. Zugriff am 15.03.17 auf <https://www.statmodel.com/chidiff.shtml>
- Muthen, B. O. & Satorra, A. (1995). Complex Sample Data in Structural Equation Modeling. *Sociological Methodology*, 25, 267. doi: 10.2307/271070
- Muthen, L. K. & Muthén, B. O. (2012). *Mplus: Statistical analysis with latent variables : User's guide* (7. Aufl.). Los Angeles, CA: Muthén & Muthén.

- Muthén, L. K. & Muthén, B. O. (1998-2015). *Mplus user's guide: Seventh Edition*. Los Angeles, CA: Muthén & Muthén.
- Naglieri, J. A. & Rojahn, J. (2001). Gender differences in planning, attention, simultaneous, and successive (PASS) cognitive processes and achievement. *Journal of Educational Psychology*, 93 (2), 430–437. doi: 10.1037//0022-0663.93.2.430
- Neber, H., Finsterwald, M. & Urban, N. (2001). Cooperative learning with gifted and high-achieving students: A review and meta-analyses of 12 studies. *High Ability Studies*, 12 (2), 199–214. doi: 10.1080/13598130120084339
- Neber, H. & Schommer-Aikins, M. (2002). Self-regulated science learning with highly gifted students: The role of cognitive, motivational, epistemological, and environmental variables. *High Ability Studies*, 13 (1), 59–74. doi: 10.1080/13598130220132316
- Nehring, A., Nowak, K. H., zu Belzen, A. U. & Tiemann, R. (2015). Predicting students' skills in the context of scientific inquiry with cognitive, motivational, and sociodemographic variables. *International Journal of Science Education*, 37 (9), 1343–1363. doi: 10.1080/09500693.2015.1035358
- Neumann, K., Kauertz, A. & Fischer, H. E. (2012). Quality of instruction in science education. In B. J. Fraser, K. Tobin & C. McRobbie (Hrsg.), *Second international handbook of science education* (Bd. 24, S. 247–258). Dordrecht: Springer. doi: 10.1007/978-1-4020-9041-7
- Neumann, M., Milek, A., Maaz, K. & Gresch, C. (2010). Zum Einfluss der Klassenzusammensetzung auf den Übergang von der Grundschule in die weiterführenden Schulen. In K. Maaz, J. Baumert, C. Gresch & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführenden Schulen* (S. 229–252). Bonn: BMBF.
- Nieswandt, M. (2007). Student affect and conceptual understanding in learning chemistry. *Journal of Research in Science Teaching*, 44 (7), 908–937. doi: 10.1002/tea.20169
- O'Donnell, A. M. (2006). The Role of Peers and Group Learning. In P. A. Alexander & P. Winne (Hrsg.), *Handbook of Educational Psychology* (S. 781–802). Mahwah, NJ: Erlbaum.
- OECD = Organisation for Economic Co-operation and Development. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris: OECD.
- OECD = Organisation for Economic Co-operation and Development. (2007). *PISA 2006*:

- Science competencies for tomorrow's world* (Bd. 2). Paris: OECD. doi: 10.1787/9789264040151-en
- OECD = Organisation for Economic Co-operation and Development. (2009a). *PISA data analysis manual: SAS* (2. Aufl.). Paris: OECD.
- OECD = Organisation for Economic Co-operation and Development. (2009b). *Top of the class: High performers in science in PISA 2006*. Paris: OECD.
- OECD = Organisation for Economic Co-operation and Development. (2013a). *Was Schülerinnen und Schüler wissen und können: Schülerleistungen in Lesekompetenz, Mathematik und Naturwissenschaften* (Bde. / OECD, Programme for International Student Assessment ; Bd. 1). Bielefeld: Bertelsmann.
- OECD = Organisation for Economic Co-operation and Development. (2013b). *What students know and can do: Student performance in mathematics, reading and science* (Bde. / OECD, Programme for International Student Assessment ;). Paris: OECD.
- OECD = Organisation for Economic Co-operation and Development. (2014). *Education policy outlooks: Netherlands*. Zugriff auf www.oecd.org/edu/policyoutlook.htm
- Ornstein, A. (2006). The Frequency of Hands-On Experimentation and Student Attitudes Toward Science: A Statistically Significant Relation (2005-51-Ornstein). *Journal of Science Education and Technology*, 15 (3), 285–297. doi: 10.1007/s10956-006-9015-5
- Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, 328 (5977), 463–466.
- Osborne, J. (2012). The role of argument: Learning how to learn in school science. In B. J. Fraser, K. Tobin & C. McRobbie (Hrsg.), *Second international handbook of science education* (Bd. 24, S. 933–949). Dordrecht: Springer.
- Osborne, J., Simon, S., Christodoulou, A., Howell-Richardson, C. & Richardson, K. (2013). Learning to argue: A study of four schools and their attempt to develop the use of argumentation as a common instructional practice and its impact on students. *Journal of Research in Science Teaching*, 50 (3), 315–347.
- Osborne, J., Simon, S. & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, 25 (9), 1049–1079. doi: 10.1080/0950069032000032199
- Oser, F. & Baeriswyl, F. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Hrsg.), *Handbook of research on teaching* (S. 1031–

- 1065). Washington, D.C.: American Educational Research Association.
- Oser, F. & Patry, J.-L. (1990). *Choreographien unterrichtlichen Lernens: Basismodelle des Unterrichts*. Universität Fribourg.
- Ostermeier, C., Prenzel, M. & Duit, R. (2010). Improving science and mathematics instruction: The SINUS project as an example for reform as teacher professional development. *International Journal of Science Education*, 32 (3), 303–327. doi: 10.1080/09500690802535942
- Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology*, 21 (4), 325–344. doi: 10.1006/ceps.1996.0025
- Parchmann, I., Gräsel, C., Baer, A., Nentwig, P., Demuth, R. & Ralle, B. (2006). “Chemie im Kontext”: A symbiotic implementation of a context-based teaching and learning approach. *International Journal of Science Education*, 28 (9), 1041–1062. doi: 10.1080/09500690600702512
- Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S. & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, 34 (1), 29–48. Zugriff auf <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1029130&site=ehost-live> doi: 10.1080/01443410.2013.797339
- Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U. & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: A multicontext study. *Developmental Psychology*, 48 (6), 1629.
- Pastor, D. A., Barron, K. E., Miller, B. J. & Davis, S. L. (2007). A latent profile analysis of college students’ achievement goal orientation. *Contemporary Educational Psychology*, 32 (1), 8–47. doi: 10.1016/j.cedpsych.2006.10.003
- Patrick, L., Care, E. & Ainley, M. (2010). The relationship between vocational interests, self-efficacy, and achievement in the prediction of educational pathways. *Journal of Career Assessment*, 19 (1), 61–74. doi: 10.1177/1069072710382615
- Pehmer, A.-K., Gröschner, A. & Seidel, T. (2015). How teacher professional development regarding classroom dialogue affects students’ higher-order learning. *Teaching and Teacher Education*, 47, 108–119. doi: 10.1016/j.tate.2014.12.007
- Perry, N. E., Turner, J. C. & Meyer, D. K. (2006). Classrooms as contexts for motivating learning. In P. A. Alexander & P. Winne (Hrsg.), *Handbook of Educational Psychology*

- (S. 327–348). Mahwah, NJ: Erlbaum.
- Petko, D., Waldis, M., Pauli, C. & Reusser, K. (2003). Methodologische Überlegungen zur videogestützten Forschung in der Mathematikdidaktik. *Zentralblatt für Didaktik der Mathematik*, 35 (6), 265–280. doi: 10.1007/BF02656691
- Piaget, J. (1974). *Theorien und Methoden der modernen Erziehung*. Frankfurt: Fischer.
- Piaget, J., Aebli, H. & Seiler, B. (1996). *Das Erwachen der Intelligenz beim Kinde* (4. Aufl., Bd. 1). Stuttgart: Klett.
- Plucker, J. A., Robinson, N. M., Greenspon, T. S., Feldhusen, J. F., McCoach, D. B. & Subotnik, R. F. (2004). It's not how the pond makes you feel, but rather how high you can jump. *American Psychologist*, 59 (4), 268–269.
- Plucker, J. A. & Stocking, V. B. (2001). Looking outside and inside: Self-Concept Development of Gifted Adolescents. *Exceptional Children*, 67 (4), 535–548. doi: 10.1177/001440290106700407
- Preacher, K. J., Zhang, Z. & Zyphur, M. J. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological Methods*, 21 (2), 189.
- Precht, C. (Hrsg.). (2015). *Bildung. Mehr als Fachlichkeit: Gutachten* (1. Aufl.). Münster: Waxmann.
- Prenzel, M. (1988). *Die Wirkungsweise von Interesse: Ein pädagogisch-psychologisches Erklärungsmodell: Zugl.: München, Univ., Habil.-Schr., 1987* (Bd. 13). Opladen: Westdeutsche Verlag.
- Prenzel, M. (1992). Selective persistence of interest. In K. A. Renninger, S. Hidi & A. Krapp (Hrsg.), *The Role of interest in learning and development* (S. 71–98). Hillsdale, N.J.: L. Erlbaum Associates.
- Prenzel, M. (1993). Autonomie und Motivation im Lernen Erwachsener. *Zeitschrift für Pädagogik*, 39 (2), 239–253.
- Prenzel, M. (1997). Sechs Möglichkeiten, Lernende zu demotivieren. In H. Gruber & A. Renkl (Hrsg.), *Wege zum Können: Determinanten des Kompetenzerwerbs* (S. 32–44). Bern: Huber.
- Prenzel, M. (2000). Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts. *Unterrichtswissenschaft*, 28, 103–126.
- Prenzel, M. (Hrsg.). (2001). *Erhebungs- und Auswertungsverfahren des DFG-Projekts 'Lehr-Lern-Prozesse im Physikunterricht - eine Videostudie'*. Kiel: IPN.

- Prenzel, M. (Hrsg.). (2007). *Studies on the educational quality of schools: The final report on the DFG priority programme*. Münster: Waxmann.
- Prenzel, M. (2012). Empirische Bildungsforschung morgen: Reichen unsere bisherigen Forschungsansätze aus? In M. Gläser-Zikuda, T. Seidel, C. Rohlf, A. Gröschner & S. Ziegelbauer (Hrsg.), *Mixed Methods in der empirischen Bildungsforschung* (S. 273–286). Münster: Waxmann Verlag.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., ... Schiefele, U. (2006). *PISA 2003: Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Münster: Waxmann.
- Prenzel, M., Häussler, P., Rost, J. & Senkbeil, M. (2002). Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorhersagen? *Unterrichtswissenschaft : Zeitschrift für Lernforschung*, 30 (1), 120–135.
- Prenzel, M., Kramer, K. & Drechsel, B. (2001). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung - Ergebnisse eines Forschungsprojektes. In K. Beck & V. Krumm (Hrsg.), *Lehren und Lernen in der beruflichen Erstausbildung: Grundlagen einer modernen kaufmännischen Berufsqualifizierung* (S. 37–61). Opladen: Leske + Budrich.
- Prenzel, M., Kramer, K. & Drechsel, B. (2002). Self-determined and interested learning in vocational education. In K. Beck (Hrsg.), *Teaching-learning processes in vocational education: Foundations of modern training programmes* (Bd. 5, S. 43–68). Frankfurt am Main: Lang.
- Prenzel, M., Krapp, A. & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik*, 32 (2), 163–173.
- Prenzel, M. & Lankes, E.-M. (2013). Was können Schülerinnen und Schüler über ihren Unterricht sagen? Ein Blick in die Schülerfragebogen von internationalen Vergleichsstudien. In N. McElvany & H. G. Holtappels (Hrsg.), *Empirische Bildungsforschung: Theorien, Methoden, Befunde und Perspektiven: Festschrift für Wilfried Bos* (S. 91–105). Münster: Waxmann.
- Prenzel, M., Reiss, K. & Hasselhorn, M. (2009). Förderung der Kompetenzen von Kindern und Jugendlichen. In J. Milberg (Hrsg.), *Förderung des Nachwuchses in Technik und Naturwissenschaft: Beiträge zu den zentralen Handlungsfeldern* (S. 15–60). Berlin: Springer.
- Prenzel, M., Rost, J., Senkbeil, M., Häußler, P. & Klopp, A. (2001). Naturwissenschaftliche

- Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert (Hrsg.), *PISA 2000* (S. 191–248). Opladen: Leske + Budrich.
- Prenzel, M., Schöps, K., Rönnebeck, S., Senkbeil, M., Walter, O., Carstensen, C. H. & Hammann, M. (2007). Naturwissenschaftliche Kompetenz im internationalen Vergleich. In M. Prenzel et al. (Hrsg.), *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 61–124). Münster: Waxmann.
- Prenzel, M. & Schütte, K. (2007). Interesse an den Naturwissenschaften. In M. Prenzel et al. (Hrsg.), *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 95–106). Münster: Waxmann.
- Prenzel, M. & Schütte, K. (2008). Interesse an den Naturwissenschaften. In M. Prenzel (Hrsg.), *PISA 2006 in Deutschland: Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 95–106). Münster and , New York and NY and München and Berlin: Waxmann.
- Rakoczy, K. (2008). *Motivationsunterstützung im Mathematikunterricht: Unterricht aus der Perspektive von Lernenden und Beobachtern: Zugl.: Frankfurt/Main, Univ., Diss., 2006* (Bd. 65). Münster: Waxmann.
- Rakoczy, K., Klieme, E. & Pauli, C. (2008). Die Bedeutung der wahrgenommenen Unterstützung motivationsrelevanter Bedürfnisse und des Alltagsbezugs im Mathematikunterricht für die selbstbestimmte Motivation. *Zeitschrift für Pädagogische Psychologie*, 22 (1), 25–35. doi: 10.1024/1010-0652.22.1.25
- Ramsden, J. M. (1997). How does a context-based approach influence understanding of key chemical ideas at 16+? *International Journal of Science Education*, 19 (6), 697–710. doi: 10.1080/0950069970190606
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: Univ. of Chicago Pr.
- Raudenbush, S. W. & Bryk, A. S. (2010). *Hierarchical linear models: Applications and data analysis methods* (2. Aufl., Bd. 1). Thousand Oaks, Calif.: Sage Publ.
- Reis, S. M. & Park, S. (2001). Gender differences in high-achieving students in math and science. *Journal for the Education of the Gifted*, 25 (1), 52–73. doi: 10.1177/016235320102500104
- Reis, S. M. & Renzulli, J. S. (2010). Is there still a need for gifted education? An examination of current research. *Learning and Individual Differences*, 20 (4), 308–317. doi: 10.1016/j.lindif.2009.10.012

- Renkl, A. (1996). Vorwissen und Schulleistung. In J. Möller & O. Köller (Hrsg.), *Emotionen, Kognitionen und Schulleistung* (S. 175–190). Weinheim: Beltz.
- Renkl, A. (2010). Träges Wissen. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 854–858). Weinheim: Beltz.
- Renkl, A. (2015). Wissenserwerb. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 3–24). Berlin, Heidelberg: Springer.
- Renkl, A. & Atkinson, R. K. (2007). Interactive Learning Environments: Contemporary Issues and Trends. An Introduction to the Special Issue. *Educational Psychology Review*, 19 (3), 235–238. doi: 10.1007/s10648-007-9052-5
- Renkl, A. & Stern, E. (1994). Die Bedeutung von kognitiven Eingangsvoraussetzungen und schulischen Lerngelegenheiten für das Lösen von einfachen und komplexen Textaufgaben. *Zeitschrift für Pädagogische Psychologie*, 8 (1), 27–39.
- Renninger, K. A., Hidi, S. & Krapp, A. (Hrsg.). (1992). *The Role of interest in learning and development*. Hillsdale, N.J: L. Erlbaum Associates.
- Renzulli, J. S. (1976). The enrichment triad model: A guide for developing defensible programs for the gifted and talented. *The Gifted Child Quarterly*, 20 (3), 303–306.
- Renzulli, J. S. (1984). The triad/ revolving door system: A research-based approach to identification and programming for the gifted and talented. *Gifted Child Quarterly*, 28 (4), 163–171. doi: 10.1177/001698628402800405
- Reusser, K. & Pauli, C. (2003). Unterrichtsskripts im schweizerischen und im deutschen Mathematikunterricht. *Unterrichtswissenschaft : Zeitschrift für Lernforschung*, 31 (3), 238–272.
- Reyer, T., Trendel, G. & Fischer, H. E. (2004). Was kommt beim Schüler an? - Lehrerintentionen und Schülerlernen im Physikunterricht. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung* (S. 195–211). Münster: Waxmann.
- Rheinberg, F. (1998). Theory of interest and research on motivation to learn. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender* (S. 126–145). Kiel: IPN.
- Rheinberg, F. (2004). *Motivationsdiagnostik* (Bd. 5). Göttingen: Hogrefe.
- Rheinberg, F. & Vollmeyer, R. (2000). Sachinteresse und leistungsthematische Herausforderung - zwei verschiedenartige Motivationskomponenten und ihr Zusammenwir-

- ken beim Lernen. In U. Schiefele & K.-P. Wild (Hrsg.), *Interesse und Lernmotivation: Untersuchungen zu Entwicklung, Förderung und Wirkung* (S. 145–161). Münster: Waxmann.
- Robinson, A. (1990). Cooperation or exploitation? The argument against cooperative learning for talented students. *Journal for the Education of the Gifted*, 14 (1), 9–27.
- Robinson, A. (2003). Cooperative learning and high ability students. In N. Colangelo & G. A. Davis (Hrsg.), *Handbook of gifted education* (S. 282–292). Boston: Allyn and Bacon.
- Robinson, A., Shore, B. M. & Enersen, D. L. (2006a). Multiple criteria for identification. In A. Robinson & B. M. Shore (Hrsg.), *Best Practices in Gifted Education* (S. 235–246). Naperville: Prufrock Press Inc.
- Robinson, A., Shore, B. M. & Enersen, D. L. (2006b). Science in the classroom. In A. Robinson & B. M. Shore (Hrsg.), *Best Practices in Gifted Education* (S. 163–169). Naperville: Prufrock Press Inc.
- Roeser, R. W., Eccles, J. S. & Sameroff, A. J. (1998). Academic and emotional functioning in early adolescence: Longitudinal relations, patterns, and prediction by experience in middle school. *Development and Psychopathology*, 10 (2), 321–352.
- Rost, D. H. (2009a). Grundlagen, Fragestellungen, Methoden. In D. H. Rost (Hrsg.), *Hochbegabte und hochleistende Jugendliche* (Bd. 72, S. 1–91). Münster: Waxmann.
- Rost, D. H. (Hrsg.). (2009b). *Hochbegabte und hochleistende Jugendliche: Befunde aus dem Marburger Hochbegabtenprojekt* (2. Aufl., Bd. 72). Münster: Waxmann.
- Rost, D. H. (2009c). *Intelligenz: Fakten und Mythen* (1. Aufl.). Weinheim: Beltz.
- Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion* (2. Aufl.). Bern: Huber.
- Rost, J., Sievers, K., Häußler, P., Hoffmann, L. & Langeheine, R. (1999). Struktur und Veränderung des Interesses an Physik bei Schülern der 6. bis 10. Klassenstufe. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31 (1), 18–31.
- Rost, J., Walter, O., Carstensen, C. H., Senkbeil, M. & Prenzel, M. (2004). Naturwissenschaftliche Kompetenz. In M. Prenzel (Hrsg.), *PISA 2003* (S. 111–146). Münster: Waxmann.
- Roth, K. J., Druker, S. L., Garnier, H. E., Lemmens, M., Chen, C., Kawanaka, T., ... Okamoto, Y. (2006). Teaching Science in Five Countries: Results From the TIMSS 1999 Video Study. Statistical Analysis Report. NCES 2006-011. *National Center for Education*

Statistics.

- Rubin, R. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley & Sons.
- Ryan, R. M. & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25 (1), 54–67. doi: 10.1006/ceps.1999.1020
- Ryan, R. M. & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55 (1), 68.
- Ryan, R. M. & Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 3–33). Rochester, NY: University of Rochester Press.
- Ryu, E. (2014). Factorial invariance in multilevel confirmatory factor analysis. *British Journal of Mathematical and Statistical Psychology*, 67 (1), 172–194.
- Sadeh, I. & Zion, M. (2012). Which type of inquiry project do high school biology students prefer: Open or guided? *Research in Science Education*, 42 (5), 831–848. doi: 10.1007/s11165-011-9222-9
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock & A. Satorra (Hrsg.), *Innovations in Multivariate Statistical Analysis* (Bd. 36, S. 233–247). Boston, MA: Springer.
- Satorra, A. & Bentler, P. M. (2010). Ensuring Positiveness of the Scaled Difference Chi-square Test Statistic. *Psychometrika*, 75 (2), 243–248. doi: 10.1007/s11336-009-9135-y
- Scantlebury, K. & Baker, D. (2007). Gender issues in science education research: Remembering where the difference lies. In S. K. Abell & N. G. Lederman (Hrsg.), *Handbook of research on science education* (S. 257–286). Mahwah and N.J.: Lawrence Erlbaum Associates.
- Scheerens, J. & Bosker, R. J. (2011). *The foundations of educational effectiveness* ([Nachdr. der Ausg.] 1997 Aufl.). Bingley: Emerald.
- Scheerens, J., Luyten, H., Steen, R. & Luyten-Thouars, Y. d. (2007). *Review and meta-analyses of school and teaching effectiveness*. Enschede: Universität Twente.
- Scheerens, J., Seidel, T., Witziers, B., Hendriks, M. & Doornekamp, G. (2005). *Positioning and Validating the Supervision Framework*. Enschede/Kiel: University of Twente,

- Department of Educational Organisation and Management.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8 (2), 23–74.
- Schiefele, H. (1978). *Lernmotivation und Motivlernen: Grundzüge einer erziehungswissenschaftlichen Motivationslehre* (2. Aufl.). München: Ehrenwirth.
- Schiefele, H., Krapp, A., Prenzel, M., Heiland, A. & Kasten, H. (1983). *Principles of an educational theory of interest*.
- Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, 26 (3-4), 299–323. doi: 10.1080/00461520.1991.9653136
- Schiefele, U. (1996). *Motivation und Lernen mit Texten: Zugl.: München, Univ. der Bundeswehr, Habil-Schr., 1994*. Göttingen: Hogrefe.
- Schiefele, U. (1998). Individual interest and learning - What we know and what we don't know. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seeon Conference on Interest and Gender* (S. 91–104). Kiel: IPN.
- Schiefele, U., Krapp, A., Wild, K.-P. & Winteler, A. (1993). Der Fragebogen zum Studieninteresse (FSI). *Diagnostica*, 39 (4), 335–351.
- Schiefele, U., Krapp, A. & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi & A. Krapp (Hrsg.), *The Role of interest in learning and development* (S. 183–212). Hillsdale, N.J.: L. Erlbaum Associates.
- Schiefele, U. & Schreyer, I. (1992). *Intrinsische Lernmotivation und Lernen: Ein Überblick über Theorien und empirische Befunde* (Bd. 24). München: Institut für Erziehungswissenschaft und Pädagogische Psychologie Universität der Bundeswehr München.
- Schiefele, U. & Wild, K.-P. (Hrsg.). (2000). *Interesse und Lernmotivation: Untersuchungen zu Entwicklung, Förderung und Wirkung*. Münster: Waxmann.
- Schiepe-Tiska, A., Roczen, N., Müller, K., Prenzel, M. & Osborne, J. (2016). Science-related outcomes: attitudes, motivation, value beliefs, strategies. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing context of learning world-wide– Extended Context Assessment Framework and Documentation of Questionnaire Material*. New York: Springer.

- Schiepe-Tiska, A., Schöps, K., Rönnebeck, S., Köller, O. & Prenzel, M. (2013). Naturwissenschaftliche Kompetenz in PISA 2012: Ergebnisse und Herausforderungen. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012* (S. 189–215). Münster: Waxmann.
- Schmidt, H. G., de Volder, M. L., de Grave, W. S., Moust, J. H. C. & Patel, V. L. (1989). Explanatory models in the processing of science text: The role of prior knowledge activation through small-group discussion. *Journal of Educational Psychology*, *81* (4), 610–619. doi: 10.1037/0022-0663.81.4.610
- Schneider, T. (2004). Hauptschule, Realschule oder Gymnasium? Soziale Herkunft als Determinante der Schulwahl. In M. Szydlik (Hrsg.), *Generation und Ungleichheit* (Bd. 19, S. 77–103). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schneider, W. (1997). The impact of expertise on performance: Illustrations from developmental research on memory and sports. *High Ability Studies*, *8*, 7–18.
- Schneider, W. & Stumpf, E. (2007). Hochbegabung, Expertise und die Erklärung außergewöhnlicher Leistungen. In K. A. Heller & A. Ziegler (Hrsg.), *Begabt sein in Deutschland* (S. 71–91). Berlin: Lit Verlag.
- Schrader, F.-W. & Helmke, A. (2008). Determinanten der Schulleistung. In M. K. W. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion : Inhaltsfelder, Forschungsperspektiven und methodische Zugänge* (Bd. 24, S. 285–302). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schraw, G. (2006). Knowledge: Structures and processes. In P. A. Alexander & P. Winne (Hrsg.), *Handbook of Educational Psychology* (S. 245–263). Mahwah, NJ: Erlbaum.
- Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T.-Y. & Lee, Y.-H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching*, *44* (10), 1436–1460. doi: 10.1002/tea.20212
- Schütte, K., Frenzel, A. C., Asseburg, R. & Pekrun, R. (2007). Schülermerkmale, naturwissenschaftliche Kompetenz und Berufserwartung. In M. Prenzel et al. (Hrsg.), *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 125–146). Münster: Waxmann.
- Schütte, K. & Köller, O. (2015). 'Discover, Understand, Implement, and Transfer': Effectiveness of an intervention programme to motivate students for science. *International Journal of Science Education*, 1–20. doi: 10.1080/09500693.2015.1077537

- Schunk, D. H. & Swartz, C. W. (1993). Writing strategy instruction with gifted students: Effects of goals and feedback on self-efficacy and skills*. *Roeper Review*, 15 (4), 225–230. doi: 10.1080/02783199309553512
- Schwartz, A. T. (1997). *Chemistry in context: Applying chemistry to society* (2. Aufl.). Dubuque, IA: WCB and McGraw-Hill.
- Schwartz, D., Chase, C., Oppezzo, M. & Chin, D. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, 103 (4), 759–775. doi: 10.1037/a0025140
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6 (2), 461–464.
- sdw = Stiftung der Deutschen Wirtschaft. (2015a). *Begabungsförderung für die Generation X und Y*. Hamburg: Murmann Publishers GmbH.
- sdw = Stiftung der Deutschen Wirtschaft. (2015b). *Ingo Kramer zur Absolventenstudie: Wir müssen Leistungsstarke fördern!/ Essayband veröffentlicht*. Zugriff auf <https://www.sdw.org/news-terminenachrichten/nachricht/ingo-kramer-zur-absolventenstudie-wir-muessen-leistungsstarke-foerdern-essayband-veroeffentlicht>
- Seidel, T. (2003). *Lehr-Lernskripts im Unterricht: Freiräume und Einschränkungen für kognitive und motivationale Lernprozesse - eine Videostudie im Physikunterricht* (Bd. 35). Münster: Waxmann.
- Seidel, T. (2006). The role of student characteristics in studying micro teaching–learning environments. *Learning Environments Research*, 9 (3), 253–271. doi: 10.1007/s10984-006-9012-x
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie: Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, 60 (6), 850–866. doi: 10.3262/ZP1406850
- Seidel, T., Jurik, V., Häusler, J. & Stubben, S. (2016). Mikro-Umwelten im Klassenverband: Wie sich kognitive und motivational-affektive Schülervoraussetzungen auf die Wahrnehmung und das Verhalten im Fachunterricht auswirken. In N. McElvany, W. Bos, H. G. Holtappels, M. M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (Bd. 1, S. 65–87). Münster: Waxmann.
- Seidel, T. & Prenzel, M. (2004). Muster unterrichtlicher Aktivitäten im Physikunterricht. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule: Lehrerprofessionalisie-*

- ...
rung, *Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsentwicklung* (S. 177–194). Münster: Waxmann.
- Seidel, T. & Prenzel, M. (2006a). Stability of teaching patterns in physics instruction: Findings from a video study. *Learning and Instruction*, 16 (3), 228–240. doi: 10.1016/j.learninstruc.2006.03.002
- Seidel, T. & Prenzel, M. (2006b). Teaching and learning of science. In Australian Council for Educational Research (Hrsg.), *Contextual framework for PISA 2006* (S. 57–73). Camberwell: Australian Council of Educational Research (ACER).
- Seidel, T. & Prenzel, M. (2008). Large scale assessments. In J. Hartig (Hrsg.), *Assessment of competencies in educational contexts* (S. 279–304). Cambridge, Mass.: Hogrefe.
- Seidel, T., Prenzel, M. & Kobarg, M. (2005). *How to run a video study*. Münster: Waxmann.
- Seidel, T., Prenzel, M., Rimmele, R., Dalehefte, I. M., Herweg, C., Kobarg, M. & Schwindt, K. (2006). Blicke auf den Physikunterricht. Ergebnisse der IPN Videostudie. *Zeitschrift für Pädagogik*, 52 (6), 799–821.
- Seidel, T., Prenzel, M., Rimmele, R., Herweg, C., Kobarg, M., Schwindt, K. & Dalehefte, I. M. (2007). Science teaching and learning in German physics classrooms: Findings from the IPN Video Study. In M. Prenzel (Hrsg.), *Studies on the educational quality of schools: The final report on the DFG priority programme* (S. 79–100). Münster: Waxmann.
- Seidel, T., Prenzel, M., Wittwer, J. & Schwindt, K. (2007). Unterricht in den Naturwissenschaften. In M. Prenzel et al. (Hrsg.), *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 147–179). Münster: Waxmann.
- Seidel, T., Rimmele, R. & Prenzel, M. (2003). Gelegenheitsstrukturen beim Klassengespräch und ihre Bedeutung für die Lernmotivation. Videoanalysen in Kombination mit Schülerelbsteinschätzungen. *Unterrichtswissenschaft : Zeitschrift für Lernforschung*, 31 (2), 142–165.
- Seidel, T., Rimmele, R. & Prenzel, M. (2005). Clarity and coherence of lesson goals as a scaffold for student learning. *Learning and Instruction*, 15 (6), 539–556. doi: 10.1016/j.learninstruc.2005.08.004
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77 (4), 454–499. doi: 10.3102/0034654307310317
- Shavelson, R. J. & Seidel, T. (2006). Approaches in measuring learning environments.

- Learning Environments Research*, 9 (3), 195–197.
- Shuell, T. J. (1996). Teaching and learning in a classroom context. In D. C. Berliner & R. C. Calfee (Hrsg.), *Handbook of educational psychology* (S. 726–764). New York: MacMillan.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15 (2), 4–14.
- Shulman, L. S. (2001). Paradigms and research programs in the study of teaching: A contemporary perspective. In V. Richardson (Hrsg.), *Handbook of research on teaching* (S. 3–36). Washington, D.C.: American Educational Research Association.
- Silvia, P. J. (2006). *Exploring the psychology of interest*. Oxford and New York: Oxford University Press.
- Sirius Programme. (2008). *The Sirius Programme: Promoting Excellence in Dutch Higher Education*. Zugriff auf www.siriusprogramma.nl/?pid=184
- Sjöberg, L. (1984). Interests, effort, achievement and vocational preference. *British Journal of Educational Psychology*, 54 (2), 189–205.
- Skaalvik, E. M. & Rankin, R. J. (1995). A test of the internal/external frame of reference model at different levels of math and verbal self-perception. *American Educational Research Journal*, 32 (1), 161–184.
- Skaalvik, E. M. & Skaalvik, S. (2002). Internal and external frames of reference for academic self-concept. *Educational Psychologist*, 37 (4), 233–244.
- Skinner, E. A. & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85 (4), 571–581. doi: 10.1037/0022-0663.85.4.571
- Slavin, R. E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57 (3), 293–336.
- Slavin, R. E. (1990a). Ability grouping, cooperative learning and the gifted. *Journal for the Education of the Gifted*, 14 (1), 3–8.
- Slavin, R. E. (1990b). Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis. *Review of Educational Research*, 60 (3), 471–499. doi: 10.3102/00346543060003471
- Slavin, R. E. (1990c). *Cooperative learning: Theory, research, and practice*. Englewood Cliffs, N.J: Prentice Hall.
- Slavin, R. E. (1991). Are cooperative learning and untracking harmful to the gifted.

- Educational Leadership*, 48 (6), 68–71.
- Slavin, R. E. (2010). Cooperative learning. In P. L. Peterson, E. L. Baker & B. McGaw (Hrsg.), *International encyclopedia of education* (S. 177–183). Oxford: Elsevier. doi: 10.1016/B978-0-08-044894-7.00494-2
- Sälzer, C. (2015). *Erkenntnisse aus Large-Scale Assessments als empirischer Bezugspunkt für bildungspolitische Entscheidungen: Habilitation an der TUM School of Education*. München.
- Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Zwischen Grundlagenwissen und Anwendungsbezug: Mathematische Kompetenzen im internationalen Vergleich. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012* (S. 47–97). Münster: Waxmann.
- Sälzer, C., Schiepe-Tiska, A. & Prenzel, M. (2017). Internationale Schulleistungsuntersuchungen. In M. Gläser-Zikuda, M. Haring & C. Rohlf (Hrsg.), *Handbuch Schulpädagogik*. Münster: Waxmann.
- Snow, R. E. (1989). Aptitude-treatment interaction as a framework for research on individual differences in learning. In P. L. Ackerman (Hrsg.), *Learning and individual differences* (S. 13–59). New York: Freeman.
- Snow, R. E., Corno, L. & Belmont, M. J. (1996). Individual differences in affective and conative functions. In D. C. Berliner & R. C. Calfee (Hrsg.), *Handbook of educational psychology* (S. 243–310). New York: MacMillan.
- Stake, J. E. & Mares, K. R. (2001). Science Enrichment Programs for Gifted High School Girls and Boys: Predictors of Program Impact on Science Confidence and Motivation. *Journal of Research in Science Teaching*, 38 (10), 1065–1088.
- Stark, R., Gruber, H., Renkl, A. & Mandl, H. (1998). Instructional effects in complex learning: Do objective and subjective learning outcomes converge? *Learning and Instruction*, 8 (2), 117–129. doi: 10.1016/S0959-4752(97)00005-4
- Stefanou, C. R., Perencevich, K. C., DiCintio, M. & Turner, J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist*, 39 (2), 97–110. doi: 10.1207/s15326985ep3902{\textunderscore}2
- Steinmayr, R. & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, 19 (1), 80–90. doi: 10.1016/j.lindif.2008.05.004

- Steinmetz, H. (2013). Analyzing Observed Composite Differences Across Groups. *Methodology*, 9 (1), 1–12. doi: 10.1027/1614-2241/a000049
- Stemmler, M. & Heine, J.-H. (2015). Person-Centered Developmental Psychology: An Alternative Statistical Approach: Prepared for the Methods and Measures Section of the International Journal of Behavioral Development. *International Journal of Behavioral Development*, 1–47.
- Stern, E. (1994). A microgenetic longitudinal study on the acquisition of word problem solving skills. In Van Luit, Johannes E. H. (Hrsg.), *Research on Learning and Instruction of Mathematics in Kindergarten and Primary School* (S. 229–242).
- Stern, E. (1997). Erwerb mathematischer Kompetenzen: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 157–170). Weinheim: Beltz.
- Stevens, R. J. & Slavin, R. E. (1995). The cooperative elementary school: Effects on students' achievement, attitudes, and social relations. *American Educational Research Journal*, 32 (2), 321–351.
- Stocking, V. B. & Goldstein, D. (1992). Course selection and performance of very high ability students: Is there a gender gap? *Roeper Review: A Journal on Gifted Education*, 15 (1), 48–51.
- Tajfel, H. (1982). Social Psychology of Intergroup Relations. *Annual Review of Psychology*, 33 (1), 1–39. doi: 10.1146/annurev.ps.33.020182.000245
- Taraban, R., Box, C., Myers, R., Pollard, R. & Bowen, C. W. (2007). Effects of active-learning experiences on achievement, attitudes, and behaviors in high school biology. *Journal of Research in Science Teaching*, 44 (7), 960–979.
- Taskinen, P. H. (2010). *Naturwissenschaften als zukünftiges Berufsfeld für Schülerinnen und Schüler mit hoher naturwissenschaftlicher und mathematischer Kompetenz: Eine Untersuchung von Bedingungen für Berufserwartungen* (Dissertation). Christian-Albrechts-Universität, Kiel.
- Taskinen, P. H., Asseburg, R. & Walter, O. (2008). Wer möchte später einen naturwissenschaftsbezogenen oder technischen Beruf ergreifen? Kompetenzen, Selbstkonzept und Motivationen als Prädiktoren der Berufserwartungen in PISA 2006. In M. Prenzel & J. Baumert (Hrsg.), *Vertiefende Analysen zu PISA 2006* (Bd. 10, S. 79–105).
- Taskinen, P. H., Schütte, K. & Prenzel, M. (2013). Adolescents' motivation to select an academic science-related career: The role of school factors, individual interest, and

- science self-concept. *Educational Research and Evaluation*, 19 (8), 717–733. doi: 10.1080/13803611.2013.853620
- Taylor, J. A., Stuhlsatz, M. A. M. & Bybee, R. W. (2009). Windows into high-achieving science classrooms. In R. W. Bybee & B. McCrae (Hrsg.), *PISA science 2006: Implications for science teachers and teaching* (S. 123–130). Arlington, Va: NSTA Press.
- Tesch, M. (2005). *Das Experiment im Physikunterricht: Didaktische Konzepte und Ergebnisse einer Videostudie: Zugl.: Kiel, Univ., Diss., 2005* (Bd. 42). Berlin: Logos-Verlag.
- Tesch, M. & Duit, R. (2004). Experimentieren im Physikunterricht - Ergebnisse einer Videostudie. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 51–69.
- Thissen, D., Steinberg, L. & Mooney, J. A. (1989). Trace lines for testlets: A use of multiple-categorical-response models. *Journal of Educational Measurement*, 26 (3), 247–260.
- Todt, E. (1974). *Untersuchungen über die Motivation zur Beschäftigung mit naturwissenschaftlichen Problemen (Sekundarstufe I : Klassenstufe 5 - 9) (Biologie und Physik)*. Giessen: Justus-Liebig-Universität.
- Todt, E., Drewes, R. & Heils, S. (1994). The development of interests during adolescence: Social context, individual differences, and individual significance. *Adolescence in context: the interplay of family, school, peers and work in adjustment*, 82–95.
- Todt, E. & Schreiber, S. (1998). Development of interests. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Hrsg.), *Interest and learning. Proceedings of the Seon Conference on Interest and Gender* (S. 25–40). Kiel: IPN.
- Tracey, T. J. G. (2002). Development of interests and competency beliefs: A 1-year longitudinal study of fifth- to eighth-grade students using the ICA-R and structural equation modeling. *Journal of Counseling Psychology*, 49 (2), 148–163. doi: 10.1037//0022-0167.49.2.148
- Trautwein, U. & Lüdtke, O. (2005). The big-fish-little-pond effect: Future research questions and educational implications. *Zeitschrift für Pädagogische Psychologie*, 19 (3), 137–140.
- Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology*, 90 (2), 334–349. doi: 10.1037/0022-3514.90.2.334
- Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O. & Baumert, J. (2006). Tracking, grading,

- and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, 98 (4), 788–806. doi: 10.1037/0022-0663.98.4.788
- Travers, R. M. W. (1978). *Children's interests*. Kalamazoo, Mich: Western Michigan University.
- Turner, J. C., Meyer, D. K., Cox, K. E., Logan, C., DiCintio, M. & Thomas, C. T. (1998). Creating contexts for involvement in mathematics. *Journal of Educational Psychology*, 90 (4), 730–745. doi: 10.1037/0022-0663.90.4.730
- Tyler-Wood, T. L., Mortenson, M., Putney, D. & Cass, M. A. (2000). An effective mathematics and science curriculum option for secondary gifted education. *Roeper Review*, 22 (4), 266–269. doi: 10.1080/02783190009554050
- Uhden, O., Karam, R., Pietrocola, M. & Pospiech, G. (2012). Modelling mathematical reasoning in physics education. *Science & Education*, 21 (4), 485–506. doi: 10.1007/s11191-011-9396-6
- Valtin, R., Wagner, C. & Schwippert, K. (2005). Schülerinnen und Schüler am Ende der vierten Klasse: Schulische Leistungen, lernbezogene Einstellungen und außerschulische Lernbedingungen. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU* (S. 187–230). Münster: Waxmann.
- Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K. & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology*, 101 (3), 671–688. doi: 10.1037/a0015083
- VanTassel-Baska, J. (2003). What matters in curriculum for gifted learners: Reflections on theory, research, and practice. In N. Colangelo & G. A. Davis (Hrsg.), *Handbook of gifted education*. Boston: Allyn and Bacon.
- VanTassel-Baska, J., Bass, G., Ries, R., Poland, D. & Avery, L. D. (1998). A national study of science curriculum effectiveness with high ability students. *Gifted Child Quarterly*, 42 (4), 200–211.
- VanTassel-Baska, J. & Wood, S. (2010). The integrated curriculum model (ICM). *Learning and Individual Differences*, 20 (4), 345–357. doi: 10.1016/j.lindif.2009.12.006
- Vermunt, J. K. & Magidson, J. (2002). Latent class cluster analysis. In J. A. Hagenaars & A. L. McCutcheon (Hrsg.), *Applied latent class analysis* (S. 89–106). Cambridge and New York: Cambridge University Press.
- Vitzthum, T. (04.12.13). Schwache gestärkt, Starke gebremst: Die neue Pisa-Studie

- zeigt: Der Zusammenhang zwischen Herkunft und Bildungserfolg verringert sich. Allerdings fehlt es noch immer an der besseren Förderung von Leistungsstarken. *Die Welt*. Zugriff auf http://www.welt.de/print/die_welt/politik/article122529007/Schwache-gestaerkt-Starke-gebremst.html
- von Davier, M. (1997). *Methoden zur Prüfung probabilistischer Testmodelle: Zugl.: Kiel, Univ., Diss., 1996*. Kiel: IPN.
- Wagner, W., Göllner, R., Helmke, A., Trautwein, U. & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction*, 28, 1–11. doi: 10.1016/j.learninstruc.2013.03.003
- Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B. & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, 108 (5), 705–721. doi: 10.1037/edu0000075
- Wang, J., Oliver, J. S. & Staver, J. R. (2008). Self-concept and science achievement: Investigating a reciprocal relation model across the gender classification in a crosscultural context. *Journal of Research in Science Teaching*, 45 (6), 711–725.
- Wang, J.-R., Wang, Y.-C., Tai, H.-J. & Chen, W.-J. (2010). Investigating the Effectiveness of Inquiry-Based Instruction on Students with Different Prior Knowledge and Reading Abilities. *International Journal of Science and Mathematics Education*, 8 (5), 801–820.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63 (3), 249–294. doi: 10.3102/00346543063003249
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54 (3), 427–450.
- Watermann, R. & Baumert, J. (2006). Entwicklung eines Strukturmodells zum Zusammenhang zwischen sozialer Herkunft und fachlichen und überfachlichen Kompetenzen: Befunde national und international vergleichender Analysen. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit: Vertiefende Analysen im Rahmen von PISA 2000* (S. 61–94). Wiesbaden: VS Verlag für Sozialwissenschaften. doi: 10.1007/978-3-531-90082-7{\textunderscore}3

- Watermann, R. & Klieme, E. (2006). Modellierung von Kompetenzstufen mit Hilfe der latenten Klassenanalyse. *Empirische Pädagogik*, 20 (3), 321–336.
- Webb, N. (1985). Student interaction and learning in small groups: A research summary. In R. E. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb & R. Schmuck (Hrsg.), *Learning to cooperate, cooperating to learn* (S. 147–172). New York: Plenum.
- Webb, N. M. (1992). Testing a theoretical model of student interaction and learning in small groups. In R. Hertz-Lazarowitz (Hrsg.), *Interaction in cooperative groups* (S. 102–119). New York: Cambridge University Press.
- Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology*, 79 (1), 1–28.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 17–31). Weinheim: Beltz.
- Weinert, F. E. & Helmke, A. (2010). Learning from wise Mother Nature or Big Brother Instructor: The wrong choice as from an. *Educational Psychologist*, 30 (3), 135–142. doi: 10.1207/s15326985ep3003{\textunderscore}4
- Wendt, H., Willems, A. S., Tarelli, I., Euen, B. & Bos, W. (2013). Ausreichend geförderte Talente? Zu den deutschen Ergebnissen von leistungsstarken Viertklässlerinnen und Viertklässlern in IGLU 2011 und TIMSS 2011. In C. Fischer (Hrsg.), *Schule und Unterricht adaptiv gestalten: Fördermöglichkeiten für benachteiligte Kinder und Jugendliche* (Bd. 29, S. 23–34). Münster u. a.: Waxmann.
- White, B. Y. & Frederiksen, J. R. (1998). Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. *Cognition and Instruction*, 16 (1), 3–118. doi: 10.1207/s1532690xc1601{\textunderscore}2
- Whitehead, A. N. (2008). *The aims of education and other essays*. Princeton, N.J.: Recording for the Blind & Dyslexic.
- Wierstra, R. (1984). A Study on Classroom Environment and on Cognitive and Affective Outcomes of the PLON-Curriculum. *Studies in Educational Evaluation*, 10 (3), 273–282.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25 (1), 68–81. doi: 10.1006/ceps.1999.1015

- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbretton, A. J. A., Freedman-Doan, C. & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, 89 (3), 451–469. doi: 10.1037/0022-0663.89.3.451
- Wild, K.-P. (2000). Die Bedeutung betrieblicher Lernumgebungen für die langfristige Entwicklung intrinsischer und extrinsischer motivationaler Lernorientierungen. In U. Schiefele & K.-P. Wild (Hrsg.), *Interesse und Lernmotivation: Untersuchungen zu Entwicklung, Förderung und Wirkung* (S. 73–93). Münster: Waxmann.
- Wilder, S. (2015). Impact of problem-based learning on academic achievement in high school: A systematic review. *Educational Review*, 67 (4), 414–435. doi: 10.1080/00131911.2014.974511
- Willems, A. S. (2011). *Bedingungen des situationalen Interesses im Mathematikunterricht: Eine mehrebenenanalytische Perspektive: Techn. Univ., Diss.–München, 2010* (Bd. 30). Münster: Waxmann.
- Williams, J. E. & Montgomery, D. (1995). Using Frame of Reference Theory to Understand the Self-Concept of Academically Able Students. *Journal for the Education of the Gifted*, 18 (4), 399–409. doi: 10.1177/016235329501800404
- Wittwer, J. (2008). What influences the agreement among student ratings of science instruction? In M. Prenzel & J. Baumert (Hrsg.), *Vertiefende Analysen zu PISA 2006* (Bd. 10, S. 205–220).
- Wormington, S. V., Corpus, J. H. & Anderson, K. G. (2012). A person-centered investigation of academic motivation and its correlates in high school. *Learning and Individual Differences*, 22 (4), 429–438. doi: 10.1016/j.lindif.2012.03.004
- Yoon, C.-H. (2009). Self-Regulated Learning and Instructional Factors in the Scientific Inquiry of Scientifically Gifted Korean Middle School Students. *Gifted Child Quarterly*, 53 (3), 203–216. doi: 10.1177/0016986209334961
- Ziegler, A. (2000). Die vier Aufgabenfelder der Motivationsförderung von Begabten. In H. Wagner (Hrsg.), *Begabung und Leistung in der Schule* (S. 97–115). Bad Honnef: Verlag Karl Heinrich Bock.
- Ziegler, A., Finsterwald, M. & Grassinger, R. (2005). Predictors of learned helplessness among average and mildly gifted girls and boys attending initial high school physics instruction in Germany. *Gifted Child Quarterly*, 49 (1), 7–18. doi: 10.1177/001698620504900102

- Ziegler, A. & Heller, K. A. (2000a). Approach and avoidance motivation as predictors of achievement behavior in physics instructions among mildly and highly gifted eighth-grade students. *Journal for the Education of the Gifted*, 23 (4), 343–359.
- Ziegler, A. & Heller, K. A. (2000b). Conditions for self-confidence among boys and girls achieving highly in chemistry. *Journal of Advanced Academics*, 11 (3), 144–151. doi: 10.4219/jsge-2000-627
- Ziegler, A. & Heller, K. A. (2000c). Effects of an attribution retraining with female students gifted in physics. *Journal for the Education of the Gifted*, 23 (2), 217–243.
- Ziegler, A., Heller, K. A. & Broome, P. (1996). Motivational preconditions for girls gifted and highly gifted in physics. *High Ability Studies*, 7 (2), 129–143. doi: 10.1080/0937445960070203
- Ziegler, A. & Stoeger, H. (2012). Shortcomings of the IQ-based construct of underachievement. *Roepers Review*, 34 (2), 123–132. doi: 10.1080/02783193.2012.660726
- Zimmer, K., Brunner, M., Lüdtke, O., Prenzel, M. & Baumert, J. (2007). Die PISA-Spitzengruppe in Deutschland: Eine Charakterisierung hochkompetenter Jugendlicher. In K. A. Heller & A. Ziegler (Hrsg.), *Begabt sein in Deutschland* (S. 193–208). Berlin: Lit Verlag.
- Zimmerman, B. J. & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82 (1), 51–59. doi: 10.1037/0022-0663.82.1.51
- Zohar, A. & Aharon-Kravetsky, S. (2005). Exploring the effects of cognitive conflict and direct teaching for students of different academic levels. *Journal of Research in Science Teaching*, 42 (7), 829–855. doi: 10.1002/tea.20075

Abbildungsverzeichnis

1.	Entwicklungsschritte von Interesse (Krapp, 2002b, S. 399, übers. v. Verf.)	31
2.	Profile aus Leistung, Selbstkonzept und Interesse von Seidel (2006, S. 260, leistungsstarke Profile blau markiert)	51
3.	Angebots-Nutzungs-Modell der Unterrichtsforschung (Helmke, 2007, S. 42)	62
4.	Rahmenmodell zur Untersuchung des Naturwissenschaftsunterrichts in PISA 2006 (Seidel & Prenzel, 2006b, S. 60)	64
5.	Überschneidung der High Performer in Naturwissenschaften, Mathematik und Lesen in PISA 2006 und 2012	151
6.	Verteilung der High und Nicht-High-Performer auf Schularten 2006 und 2012	161
7.	Leistungsniveau von Gymnasialklassen mit unterschiedlichen High-Performer-Anteilen (unterdurchschnittliche Leistung: $M_{\text{Klasse}} < M - 1SD$, durchschnittliche Leistung: $M_{\text{Klasse}} \in [M - 1SD, M + 1SD]$, überdurchschnittliche Leistung: $M_{\text{Klasse}} > M + 1SD$)	165
8.	Leistungsheterogenität von Gymnasialklassen mit unterschiedlichen High-Performer-Anteilen (unterdurchschnittliche SD: $SD_{\text{Klasse}} < M - 1SD$, durchschnittliche SD: $SD_{\text{Klasse}} \in [M - 1SD, M + 1SD]$, überdurchschnittliche SD: $SD_{\text{Klasse}} > M + 1SD$)	167
9.	Verlauf der Eigenwerte in der explorativen Mehrebenenfaktorenanalyse	175
10.	Latent-manifestes Modell für naturwissenschaftliche Kompetenz (Modell wurde in einem Multigruppenvergleich für High und Nicht-High-Performer gerechnet)	192
11.	Latent-manifestes Modell für Interesse an den Naturwissenschaften (Modell wurde in einem Multigruppenvergleich für High und Nicht-High-Performer gerechnet)	202

12.	Modell der latenten Profilanalyse getrennt für Mädchen und Jungen (SK1: Selbstkonzept Bundle 1, SK2: Selbstkonzept Bundle 2; Int1: Interesse Bundle 1; Int2: Interesse Bundle2; c: latente Klasse; g: Geschlecht) . . .	211
13.	Selbstkonzept-Interesse-Typen der High Performer getrennt nach Mädchen und Jungen	213
14.	Anteile der Selbstkonzept-Interesse-Typen in den Schulklassen	223
15.	Modell III mit Cross-Level-Interaktionen als Pfadmodell (gestrichelte Linien: Cross-Level-Interaktionen; AU = Autonomieunterstützung; P2, P3 und P4 = Zugehörigkeitswahrscheinlichkeiten zu Typ 2, Typ 3 und Typ 4 (Typ 2 = interessierte HP, Typ 3 = leistungszuversichtliche HP, Typ 4 = motivational-affektiv beteiligte HP); kP2, kP3 und kP4 = Klassenmittelwerte der Zugehörigkeitswahrscheinlichkeiten zu Typ 2, Typ 3 und Typ 4; Typ 1 (motivational-affektiv indifferente HP) wird als Referenzkategorie verwendet; Exp = Experimente; kExp = Klassenmittelwert Experimente)	227

Tabellenverzeichnis

1.	Untersuchungen zu leistungsstarken Jugendlichen in PISA	124
2.	Anteil der Schularten an Stichproben in PISA 2006 und PISA 2012	126
3.	Kennwerte der Kompetenzen	128
4.	Kennwerte der kognitiven Grundfähigkeit sowie der Schulnoten 2006 (1: bester Wert, 6: schlechtester Wert)	128
5.	Kennwerte des „Highest International Socio-Economic Index of Occupa- tional Status (HISEI)“	129
6.	Verteilung der EGP-Klassen	130
7.	Kennwerte der Skalen zu naturwissenschaftsbezogenem Selbstkonzept und Interesse	131
8.	Kennwerte der Skalen zu motivationsunterstützenden Lehr-Lernbedin- gungen	131
9.	Kennwerte der Skalen zu Unterrichtsaktivitäten im Naturwissenschafts- unterricht	132
10.	Vier Modelle zur Berücksichtigung von Mess- und Stichprobenfehlern .	138
11.	Naturwissenschaftliche Teilkompetenzen von High und Nicht-High-Per- formern in PISA 2006	148
12.	Vergleich der mathematischen Kompetenz und Lesekompetenz von High und Nicht-High-Performern in 2006 und 2012	149
13.	Notenprofil von High und Nicht-High-Performern 2006	152
14.	Kognitive Grundfähigkeit (Subskala Figurale Analogien) im Vergleich zwi- schen High und Nicht-High-Performern in 2006	154
15.	Verteilung der kognitiven Grundfähigkeit unter High und Nicht-High- Performern	155
16.	Anteil der High Performer getrennt für Jungen und Mädchen	156

17.	Vergleich des sozioökonomischen Status von High und Nicht-High-Performern anhand des HISEI	158
18.	Verteilung der EGP-Klassen unter High und Nicht-High-Performern in 2006 und 2012	159
19.	Durchschnittlicher High-Performer-Anteil einer Schulklasse getrennt nach Schularten	163
20.	Prozentuales Vorkommen von Schulklassen mit unterschiedlichen High-Performer-Anteilen getrennt nach Schularten	163
21.	Mittlere Klassenleistung von High und Nicht-High-Performern getrennt nach Kategorien zur Beschreibung der High-Performer-Anteile	166
22.	Vergleich der mittleren Kompetenzwerte von High und Nicht-High-Performern zwischen den Kategorien zur Beschreibung der High-Performer-Anteile	167
23.	Mathematische Kompetenz und Lesekompetenz von Top Performern sowie Differenzen zu den Werten der High Performer	169
24.	Modellkennwerte der explorativen Faktorenanalyse von „Schülerexperimente“ und „Schülerinnen und Schüler forschen“	176
25.	Faktorladungen explorative Mehrebenen-Faktorenanalyse von „Schülerexperimenten“ und „Schülerinnen und Schüler forschen“	177
26.	Modellkennwerte der Mehrebenen-CFAs der Unterrichtsmerkmale	179
27.	Modelle zur Testung der Invarianz der Faktorladungen von Interesse und Unterrichtsmerkmalen zwischen Schüler- und Klassenebene	180
28.	Faktorladungen und Intraklassenkorrelation von Unterrichtsmerkmalen und Interesse	181
29.	Messinvarianz der Unterrichtsmerkmale und Interesse zwischen High und Nicht-High-Performern	184
30.	Partielle Messinvarianz für Unterrichtsskalen	184
31.	Kompositionseffekte im Zusammenhang mit naturwissenschaftlicher Kompetenz von High und Nicht-High-Performern	189
32.	Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz von High Performern	194
33.	Zusammenhang von Unterrichtsaktivitäten und naturwissenschaftlicher Kompetenz von Nicht-High-Performern	196

34.	Differenz der Koeffizienten aus Modell IV zwischen High und Nicht-High-Performern	197
35.	Kompositionseffekte und Zusammenhänge von Unterrichtsaktivitäten mit der naturwissenschaftlichen Kompetenz von High und Nicht-High-Performern	199
36.	Unterrichtsaktivitäten im Zusammenhang mit dem Interesse von High Performern an Naturwissenschaften	203
37.	Unterrichtsaktivitäten im Zusammenhang mit dem Interesse von Nicht-High-Performern an Naturwissenschaften	204
38.	Vergleich der Koeffizienten von Modell IV für Interesse zwischen High und Nicht-High-Performern	206
39.	Modellparameter der Latenten Profilanalysen für Mädchen und Jungen mit unterschiedlicher Anzahl latenter Klassen	212
40.	Mittelwerte von Selbstkonzept und Interesse in den latenten Klassen sowie Geschlechterdifferenzen	213
41.	Häufigkeiten und Anteile der Typen innerhalb leistungsstarker Mädchen und Jungen	216
42.	Überschneidung der High Performance in unterschiedlichen Domänen innerhalb der Profile	218
43.	Verteilung der Notenniveaus und Notenprofile innerhalb der Selbstkonzept-Interesse-Typen der High Performer	220
44.	Mittlere HISEI-Werte der unterschiedlichen Selbstkonzept-Interesse-Typen der High Performer	222
45.	Anzahl und Anteil der Klassen mit jeweiligem Anteil eines Selbstkonzept-Interesse-Typs von mindestens 30%	224
46.	Verteilung der gewählten Fächer bei der Beantwortung der Skalen zu den Lehr-Lernbedingungen	229
47.	Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interesse-Typen mit Autonomieunterstützung bei unterschiedlich häufigen Experimenten im Unterricht	231
48.	Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interesse-Typen mit Autonomieunterstützung bei unterschiedlich häufigen forschenden Unterrichtsaktivitäten	233

49. Mehrebenenanalyse zum Zusammenhang der Selbstkonzept-Interes- Typen mit inhaltlicher Relevanz bei unterschiedlich häufigen Experimen- ten im Unterricht	236
50. Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interes- Typen mit inhaltlicher Relevanz bei unterschiedlich häufigen forschenden Unterrichtsaktivitäten	237
51. Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interes- Typen mit Kompetenzunterstützung bei unterschiedlich häufigen Experi- menten im Unterricht	240
52. Mehrebenenanalyse zum Zusammenhang von Selbstkonzept-Interes- Typen mit Kompetenzunterstützung bei unterschiedlich häufigen for- schenden Unterrichtsaktivitäten	241
A.1. CFA-Modell der latent-manifesten Mehrebenenanalyse zu naturwissen- schaftlicher Kompetenz für Nicht-High-Performer	344
A.2. CFA-Modell der latent-manifesten Mehrebenenanalyse zu naturwissen- schaftlicher Kompetenz für High Performer	345
A.3. CFA-Modell der latent-manifesten Mehrebenenanalyse zu Interesse an den Naturwissenschaften für Nicht-High-Performer	347
A.4. CFA-Modell der latent-manifesten Mehrebenenanalyse zu Interesse an den Naturwissenschaften für High Performer	348
A.5. „Naturwissenschaftsbezogenes Fähigkeitsselbstkonzept“ (Frey & Asse- burg, 2009, S. 86)	352
A.6. „Freude und Interesse an Naturwissenschaften“ (Frey & Asseburg, 2009, S. 50)	352
A.7. „Naturwissenschaftsunterricht: Interaktives Lehren und Lernen“ (Frey & Asseburg, 2009, S. 76)	353
A.8. „Naturwissenschaftsunterricht: Experimentieren“ (Frey & Asseburg, 2009, S. 77)	353
A.9. „Naturwissenschaftsunterricht: Schülerinnen und Schüler forschen“ (Frey & Asseburg, 2009, S. 78)	354
A.10. „Naturwissenschaftsunterricht: Naturwissenschaftliche Anwendungen“ (Frey & Asseburg, 2009, S. 80)	354

A.11. „Auswahl eines naturwissenschaftlichen Fachs“ (Frey & Asseburg, 2009, S. 99)	355
A.12. „Unterstützende Lehr-Lern-Bedingungen: Wahrgenommene inhaltliche Relevanz“ (Frey & Asseburg, 2009, S. 100)	355
A.13. „Unterstützende Lehr-Lern-Bedingungen: Wahrgenommene Kompetenzunterstützung“ (Frey & Asseburg, 2009, S. 105)	356
A.14. „Unterstützende Lehr-Lern-Bedingungen: Wahrgenommene Autonomieunterstützung“ (Frey & Asseburg, 2009, S. 106)	356

A. Anhang

A.1. Latent-manifeste Mehrebenenanalysen

A.1.1. Unterrichtsaktivitäten im Zusammenhang mit naturwissenschaftlicher Kompetenz

Tabelle A.1.

CFA-Modell der latent-manifesten Mehrebenenanalyse zu naturwissenschaftlicher Kompetenz für Nicht-High-Performer

	Geschlecht		HISEI		KFT		PV		Inter		Exper		Forsch		Anwend	
	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)
<i>Schülerebene (Level 1)</i>																
Faktorladungen																
Indikator 1			1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Indikator 2									1.12	(0.04)	0.79	(0.06)	1.10	(0.04)	1.17	(0.05)
Indikator 3									1.16	(0.07)	0.96	(0.04)	1.15	(0.04)	1.32	(0.06)
Indikator 4									1.30	(0.08)					1.24	(0.07)
Indikator 5																
Indikator 6																
Kovarianzen, Varianzen und Korrelationen																
Geschlecht	0.24	(0.00)					-0.02	(0.01)								
HISEI			1.00	(0.03)	0.02	(0.03)	0.07	(0.02)	-0.02	(0.02)	-0.02	(0.02)	-0.01	(0.02)	-0.01	(0.02)
KFT			0.02	(0.03)	0.93	(0.05)	0.16	(0.02)	0.00	(0.02)	0.00	(0.02)	-0.04	(0.02)	0.01	(0.01)
PV	-0.07	(0.03)	0.11	(0.03)	0.26	(0.03)	0.38	(0.03)	-0.01	(0.01)	0.00	(0.02)	-0.05	(0.01)	0.00	(0.01)
Inter			-0.04	(0.03)	0.00	(0.03)	-0.04	(0.04)	0.26	(0.02)	0.17	(0.02)	0.20	(0.02)	0.16	(0.02)
Exper			-0.03	(0.04)	0.00	(0.04)	0.01	(0.05)	0.62	(0.04)	0.28	(0.03)	0.22	(0.02)	0.15	(0.02)
Forsch			-0.02	(0.03)	-0.07	(0.03)	-0.14	(0.04)	0.65	(0.03)	0.72	(0.04)	0.35	(0.03)	0.20	(0.02)
Anwend			-0.01	(0.04)	0.01	(0.03)	-0.01	(0.04)	0.63	(0.03)	0.58	(0.04)	0.69	(0.03)	0.24	(0.02)
<i>Klassenebene (Level 2)</i>																
Faktorladungen																
Indikator 1							1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Indikator 2									1.12	(0.04)	0.79	(0.06)	1.10	(0.04)	1.17	(0.05)
Indikator 3									1.16	(0.07)	0.96	(0.04)	1.15	(0.04)	1.32	(0.06)
Indikator 4									1.30	(0.08)					1.24	(0.07)
Indikator 5																
Kovarianzen, Varianzen und Korrelationen																
PV							0.03	(0.01)	-0.01	(0.01)	-0.01	(0.01)	-0.02	(0.01)	0.00	(0.01)
Inter							-0.29	(0.13)	0.08	(0.01)	0.06	(0.02)	0.06	(0.01)	0.03	(0.01)
Exper							-0.18	(0.16)	0.52	(0.08)	0.19	(0.03)	0.09	(0.02)	0.04	(0.01)
Forsch							-0.46	(0.18)	0.70	(0.04)	0.65	(0.05)	0.10	(0.02)	0.04	(0.01)
Anwend							-0.07	(0.18)	0.59	(0.07)	0.51	(0.09)	0.55	(0.07)	0.04	(0.01)

Anmerkung. Signifikante Werte fettgedruckt. Werte über der Diagonalen sind Kovarianzen; Werte auf der Diagonalen Varianzen und Werte unter der Diagonalen Korrelationen. HISEI = Highest Socio Economic Index; KFT = Kognitive Fähigkeit; Inter = Interaktives Lehren und Lernen; Exper = Schülerexperimente; Forsch = Schülerinnen und Schüler forschen; Anwend = Anwendungen.

Tabelle A.2.

CFA-Modell der latent-manifesten Mehrebenenanalyse zu naturwissenschaftlicher Kompetenz für High Performer

	Geschlecht		HISEI		KFT		PV		Inter		Exper		Forsch		Anwend	
	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)
<i>Schülerebene (Level 1)</i>																
Faktorladungen																
Indikator 1			1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Indikator 2									1.12	(0.04)	0.79	(0.06)	1.10	(0.04)	1.17	(0.05)
Indikator 3									1.16	(0.07)	0.96	(0.04)	1.15	(0.04)	1.32	(0.06)
Indikator 4									1.30	(0.08)					1.24	(0.07)
Indikator 5																
Indikator 6																
Kovarianzen, Varianzen und Korrelationen																
Geschlecht	0.25	(0.00)							-0.03	(0.01)						
HISEI			0.95	(0.04)	0.00	(0.03)	0.02	(0.02)	-0.02	(0.01)	-0.02	(0.02)	0.00	(0.02)	0.00	(0.02)
KFT			0.02	(0.03)	0.87	(0.04)	0.09	(0.02)	0.01	(0.02)	0.00	(0.02)	-0.02	(0.02)	0.01	(0.01)
PV	-0.07	(0.03)	0.11	(0.03)	0.26	(0.03)	0.30	(0.02)	-0.01	(0.01)	-0.02	(0.01)	-0.03	(0.01)	0.00	(0.01)
Inter			-0.04	(0.03)	0.00	(0.03)	-0.04	(0.04)	0.24	(0.02)	0.12	(0.02)	0.15	(0.02)	0.13	(0.01)
Exper			-0.03	(0.04)	0.00	(0.04)	0.01	(0.05)	0.62	(0.04)	0.25	(0.02)	0.15	(0.02)	0.12	(0.01)
Forsch			-0.02	(0.03)	-0.07	(0.03)	-0.14	(0.04)	0.65	(0.03)	0.72	(0.04)	0.26	(0.03)	0.15	(0.02)
Anwend			-0.01	(0.04)	0.01	(0.03)	-0.01	(0.04)	0.63	(0.03)	0.58	(0.04)	0.69	(0.03)	0.23	(0.02)
<i>Klassenebene (Level 2)</i>																
Faktorladungen																
Indikator 1							1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Indikator 2									1.12	(0.04)	0.79	(0.06)	1.10	(0.04)	1.17	(0.05)
Indikator 3									1.16	(0.07)	0.96	(0.04)	1.15	(0.04)	1.32	(0.06)
Indikator 4									1.30	(0.08)					1.24	(0.07)
Indikator 5																
Kovarianzen, Varianzen und Korrelationen																
PV							0.01	(0.00)	0.00	(0.01)	0.00	(0.01)	0.00	(0.01)	0.00	(0.00)
Inter							-0.29	(0.13)	0.08	(0.01)	0.06	(0.02)	0.06	(0.01)	0.03	(0.01)
Exper							-0.18	(0.16)	0.52	(0.08)	0.19	(0.03)	0.09	(0.02)	0.04	(0.01)
Forsch							-0.46	(0.18)	0.70	(0.04)	0.65	(0.05)	0.10	(0.02)	0.03	(0.01)
Anwend							-0.07	(0.18)	0.59	(0.07)	0.51	(0.09)	0.55	(0.07)	0.04	(0.01)

Anmerkung. Signifikante Werte fettgedruckt. Werte über der Diagonalen sind Kovarianzen; Werte auf der Diagonalen Varianzen und Werte unter der Diagonalen Korrelationen. HISEI = Highest Socio Economic Index; KFT = Kognitive Fähigkeit; Inter = Interaktives Lehren und Lernen; Exper = Schülerexperimente; Forsch = Schülerinnen und Schüler forschen; Anwend = Anwendungen.

A.1.2. Unterrichtsaktivitäten im Zusammenhang mit Interesse an den Naturwissenschaften

Tabelle A.3.

CFA-Modell der latent-manifesten Mehrebenenanalyse zu Interesse an den Naturwissenschaften für Nicht-High-Performer

	Geschlecht		HISEI		Selbstkonzept		Freude		Inter		Exper		Forsch		Anwend	
	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)
<i>Schülerebene (Level 1)</i>																
Faktorladungen																
Indikator 1			1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Indikator 2					1.01	(0.02)	0.98	(0.02)	1.12	(0.04)	0.79	(0.06)	1.10	(0.04)	1.17	(0.05)
Indikator 3					1.07	(0.02)	0.95	(0.02)	1.16	(0.07)	0.96	(0.04)	1.14	(0.04)	1.31	(0.06)
Indikator 4					1.06	(0.02)	1.06	(0.02)	1.30	(0.08)					1.24	(0.06)
Indikator 5					0.92	(0.02)	1.07	(0.02)								
Indikator 6					0.98	(0.03)										
Kovarianzen, Varianzen und Korrelationen																
Geschlecht	0.24	(0.00)					0.02	(0.01)								
HISEI			1.00	(0.03)			0.02	(0.03)	-0.02	(0.02)	-0.02	(0.02)	-0.01	(0.02)	-0.01	(0.02)
Selbstkonzept					0.52	(0.03)	0.30	(0.02)	0.09	(0.02)	0.09	(0.02)	0.13	(0.02)	0.11	(0.02)
Freude					0.54	(0.03)	0.61	(0.03)	0.06	(0.02)	0.06	(0.02)	0.04	(0.02)	0.09	(0.01)
Inter					0.25	(0.04)	0.15	(0.04)	0.26	(0.02)	0.17	(0.02)	0.20	(0.02)	0.16	(0.02)
Exper					0.24	(0.04)	0.15	(0.04)	0.62	(0.04)	0.28	(0.03)	0.22	(0.02)	0.15	(0.02)
Forsch					0.30	(0.04)	0.09	(0.03)	0.65	(0.03)	0.72	(0.04)	0.35	(0.03)	0.20	(0.02)
Anwend					0.30	(0.04)	0.23	(0.03)	0.63	(0.03)	0.58	(0.04)	0.69	(0.03)	0.25	(0.02)
<i>Klassenebene (Level 2)</i>																
Faktorladungen																
Indikator 1							1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Indikator 2							0.98	(0.02)	1.12	(0.04)	0.79	(0.06)	1.10	(0.04)	1.17	(0.05)
Indikator 3							0.95	(0.02)	1.16	(0.07)	0.96	(0.04)	1.14	(0.04)	1.31	(0.06)
Indikator 4							1.06	(0.02)	1.30	(0.08)					1.24	(0.06)
Indikator 5							1.07	(0.02)								
Kovarianzen, Varianzen und Korrelationen																
Freude							0.02	(0.01)	0.00	(0.01)	0.00	(0.01)	0.00	(0.01)	0.01	(0.01)
Inter							0.15	(0.20)	0.08	(0.01)	0.06	(0.02)	0.06	(0.01)	0.03	(0.01)
Exper							-0.03	(0.21)	0.52	(0.08)	0.19	(0.03)	0.09	(0.02)	0.04	(0.01)
Forsch							0.12	(0.19)	0.70	(0.04)	0.65	(0.05)	0.10	(0.02)	0.04	(0.01)
Anwend							0.60	(0.20)	0.59	(0.07)	0.51	(0.09)	0.55	(0.07)	0.04	(0.01)

Anmerkung. Signifikante Werte fettgedruckt. Werte über der Diagonalen sind Kovarianzen; Werte auf der Diagonalen Varianzen und Werte unter der Diagonalen Korrelationen. HISEI = Highest Socio Economic Index; KFT = Kognitive Fähigkeit; Inter = Interaktives Lehren und Lernen; Exper = Schülerexperimente; Forsch = Schülerinnen und Schüler forschen; Anwend = Anwendungen.

Tabelle A.4.

CFA-Modell der latent-manifesten Mehrebenenanalyse zu Interesse an den Naturwissenschaften für High Performer

	Geschlecht		HISEI		Selbstkonzept		Freude		Inter		Exper		Forsch		Anwend	
	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)
<i>Schülerebene (Level 1)</i>																
Faktorladungen																
Indikator 1			1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Indikator 2					1.01	(0.02)	0.98	(0.02)	1.12	(0.04)	0.79	(0.06)	1.10	(0.04)	1.17	(0.05)
Indikator 3					1.07	(0.02)	0.95	(0.02)	1.16	(0.07)	0.96	(0.04)	1.14	(0.04)	1.31	(0.06)
Indikator 4					1.06	(0.02)	1.06	(0.02)	1.30	(0.08)					1.24	(0.06)
Indikator 5					0.92	(0.02)	1.07	(0.02)								
Indikator 6					0.98	(0.03)										
Kovarianzen, Varianzen und Korrelationen																
Geschlecht	0.25	(0.00)					-0.02	(0.01)								
HISEI			0.96	(0.04)	0.06	(0.02)	0.03	(0.02)	-0.02	(0.01)	-0.02	(0.02)	0.00	(0.02)		
Selbstkonzept			0.08	(0.03)	0.57	(0.03)	0.35	(0.03)	0.06	(0.01)	0.03	(0.02)	0.06	(0.02)	0.00	(0.02)
Freude			0.04	(0.03)	0.61	(0.03)	0.59	(0.03)	0.03	(0.02)	0.03	(0.02)	0.03	(0.02)	0.07	(0.01)
Inter			-0.04	(0.03)	0.16	(0.03)	0.07	(0.05)	0.24	(0.02)	0.12	(0.02)	0.15	(0.02)	0.13	(0.01)
Exper			-0.03	(0.04)	0.09	(0.04)	0.08	(0.04)	0.50	(0.05)	0.25	(0.02)	0.15	(0.02)	0.12	(0.01)
Forsch			-0.01	(0.03)	0.17	(0.04)	0.07	(0.04)	0.61	(0.03)	0.60	(0.05)	0.26	(0.03)	0.15	(0.02)
Anwend			0.00	(0.03)	0.20	(0.04)	0.20	(0.03)	0.57	(0.04)	0.50	(0.05)	0.61	(0.04)	0.23	(0.02)
<i>Klassenebene (Level 2)</i>																
Faktorladungen																
Indikator 1							1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
Indikator 2							0.98	(0.02)	1.12	(0.04)	0.79	(0.06)	1.10	(0.04)	1.17	(0.05)
Indikator 3							0.95	(0.02)	1.16	(0.07)	0.96	(0.04)	1.14	(0.04)	1.31	(0.06)
Indikator 4							1.06	(0.02)	1.30	(0.08)					1.24	(0.06)
Indikator 5							1.07	(0.02)								
Kovarianzen, Varianzen und Korrelationen																
Freude							0.02	(0.01)	0.00	(0.01)	0.01	(0.01)	0.00	(0.01)	0.01	(0.01)
Inter							0.09	(0.18)	0.08	(0.01)	0.06	(0.02)	0.06	(0.01)	0.03	(0.01)
Exper							0.18	(0.19)	0.51	(0.08)	0.19	(0.03)	0.09	(0.02)	0.04	(0.01)
Forsch							0.09	(0.16)	0.70	(0.05)	0.65	(0.05)	0.10	(0.02)	0.03	(0.01)
Anwend							0.37	(0.19)	0.59	(0.07)	0.50	(0.09)	0.55	(0.07)	0.04	(0.01)

Anmerkung. Signifikante Werte fettgedruckt. Werte über der Diagonalen sind Kovarianzen; Werte auf der Diagonalen Varianzen und Werte unter der Diagonalen Korrelationen. HISEI = Highest Socio Economic Index; KFT = Kognitive Fähigkeit; Inter = Interaktives Lehren und Lernen; Exper = Schülerexperimente; Forsch = Schülerinnen und Schüler forschen; Anwend = Anwendungen.

A.2. Mehrebenenanalysen zum Zusammenhang von Selbstkonzept-Interesse-Typen mit Lehr-Lernbedingungen

Modell I (Zusammenhang von Profilen und Lehr-Lernbedingungen)

Level 1:

$$y_{ij} = \beta_{0j} + \beta_{1j}c_{1i} + \beta_{2j}c_{2j} + \beta_{3j}c_{3j} + \epsilon_{ij} \quad (\text{A.1})$$

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \epsilon_{0j} \\ \beta_{1j} &= \gamma_{10} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30} \end{aligned} \quad (\text{A.2})$$

Modell II (Zusammenhang von Unterrichtsaktivitäten und Lehr-Lernbedingungen)

Level 1:

$$y_{ij} = \beta_{0j} + \beta_{1j}f_{1i} + \epsilon_{ij} \quad (\text{A.3})$$

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \epsilon_{0j} \\ \beta_{1j} &= \gamma_{10} \end{aligned} \quad (\text{A.4})$$

Modell III (Interaktion von Unterrichtsaktivitäten und Profilen im Zusammenhang mit Lehr-Lernbedingungen)

Level 1:

$$y_{ij} = \beta_{0j} + \beta_{1j}c_{1i} + \beta_{2j}c_{2j} + \beta_{3j}c_{3j} + \beta_{4j}f_{1i} + \epsilon_{ij} \quad (\text{A.5})$$

Level 2:

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}f_j + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}f_j + u_{1j} \\ \beta_{2j} &= \gamma_{20} + \gamma_{21}f_j + u_{2j} \\ \beta_{3j} &= \gamma_{30} + \gamma_{31}f_j + u_{3j} \\ \beta_{4j} &= \gamma_{40} \end{aligned} \quad (\text{A.6})$$

Insgesamt als reduzierte Formel:

$$y_{ij} = \underbrace{\gamma_{00} + \gamma_{10}c_{1i} + \gamma_{20}c_{2j} + \gamma_{30}c_{3j} + \gamma_{40}f_j + \gamma_{01}f_{1i} + \gamma_{11}c_{1ij}f_j + \gamma_{21}c_{2ij}f_j + \gamma_{31}c_{3ij}f_j}_{\text{fixed part}} + \underbrace{u_{0j} + u_{1j}c_{1ij} + u_{2j}c_{2ij} + u_{3j}c_{3ij} + \epsilon_{ij}}_{\text{random part}} \quad (\text{A.7})$$

Anmerkung: Term $\beta_{4j}f_{1i}$ beschreibt den Einfluss von Unterrichtaktivitäten auf abhängige Variable auf Schülerebene und wurde in dieser Arbeit zusätzlich aufgenommen

Modell III mit ebenenspezifischen Varianzanteilen

Modell mit implizierter *cross-level* und (Level 2)x(Level 2)-Interaktion der Level-2-Anteile der Profile und den Unterrichtsaktivitäten (nach Preacher et al., 2016):

Level 1

$$y_{ij} = \beta_{0j} + \beta_{1j}c_{1i} + \beta_{2j}c_{2j} + \beta_{3j}c_{3j} + \beta_{4j}f_{1i} + e_{ij} \quad (\text{A.8})$$

Level 2

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}c_{1,j} + \gamma_{02}c_{2,j} + \gamma_{03}c_{3,j} + \gamma_{04}f_j + \\ &\quad + \gamma_{05}c_{1,j}f_j + \gamma_{06}c_{2,j}f_j + \gamma_{07}c_{3,j}f_j + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}f_j + u_{1j} \\ \beta_{2j} &= \gamma_{20} + \gamma_{21}f_j + u_{2j} \\ \beta_{3j} &= \gamma_{30} + \gamma_{31}f_j + u_{3j} \\ \beta_{4j} &= \gamma_{40} \end{aligned} \quad (\text{A.9})$$

Da in dieser Arbeit nur die tatsächliche *cross-level* Interaktion berücksichtigt wird (B1 bei Preacher et al. (2016)), wird folgendes Modell verwendet:

Level 1

$$y_{ij} = \beta_{0j} + \beta_{1j}c_{1i} + \beta_{2j}c_{2j} + \beta_{3j}c_{3j} + \beta_{4j}f_{1i} + e_{ij} \quad (\text{A.10})$$

Level 2

$$\begin{aligned}
 \beta_{0j} &= \gamma_{00} + \gamma_{01}c_{1,j} + \gamma_{02}c_{2,j} + \gamma_{03}c_{3,j} + \gamma_{04}f_j + u_{0j} \\
 \beta_{1j} &= \gamma_{10} + \gamma_{11}f_j + u_{1j} \\
 \beta_{2j} &= \gamma_{20} + \gamma_{21}f_j + u_{2j} \\
 \beta_{3j} &= \gamma_{30} + \gamma_{31}f_j + u_{3j} \\
 \beta_{4j} &= \gamma_{40}
 \end{aligned}
 \tag{A.11}$$

Insgesamt als reduzierte Formel:

$$\begin{aligned}
 y_{ij} &= \gamma_{00} + \gamma_{01}c_{1,j} + \gamma_{02}c_{2,j} + \gamma_{03}c_{3,j} + \gamma_{04}f_j + u_{0j} + \\
 &+ \gamma_{10}c_{1i} + \gamma_{20}c_{2i} + \gamma_{30}c_{3i} + \gamma_{11}c_{1i}f_j + \gamma_{21}c_{1i}f_j + \gamma_{31}c_{3i}f_j + \\
 &+ \underbrace{u_{0j} + u_{1j}c_{1i} + u_{2j}c_{2i} + u_{3j}c_{3i} + \epsilon_{ij}}_{\text{random part}}
 \end{aligned}
 \tag{A.12}$$

Cross-level-Interaktionen: $\gamma_{11}, \gamma_{21}, \gamma_{31}$

A.3. Instrumente

A.3.1. Naturwissenschaftsbezogene motivational-affektive Merkmale

Tabelle A.5.

„Naturwissenschaftsbezogenes Fähigkeitsselbstkonzept“ (Frey & Asseburg, 2009, S. 86)

Item	Wortlaut
Wie sehr stimmst du mit den folgenden Aussagen überein?	
ST37Q01	a) Ich glaube, dass ich anspruchsvollen Stoff im naturwissenschaftlichen Unterricht leicht lernen kann.
ST37Q02	b) Normalerweise kann ich Prüfungsfragen im naturwissenschaftlichen Unterricht gut beantworten.
ST37Q03	c) Ich lerne neuen Stoff im naturwissenschaftlichen Unterricht schnell.
ST37Q04	d) Den Stoff im naturwissenschaftlichen Unterricht finde ich einfach.
ST37Q05	e) Wenn ich in Biologie, Chemie, Physik (oder Naturwissenschaften) unterrichtet werde, verstehe ich neue Begriffe leicht.
ST37Q06	f) Es fällt mir leicht, neue Ideen im naturwissenschaftlichen Unterricht zu verstehen.

Anmerkung. Antwortoptionen: 1 = „stimme ganz zu“; 2 = „stimme eher zu“; 3 = „stimme eher nicht zu“; 4 = „stimme gar nicht zu“.

Tabelle A.6.

„Freude und Interesse an Naturwissenschaften“ (Frey & Asseburg, 2009, S. 50)

Item ID	Wortlaut
Wie sehr stimmst du mit den folgenden Aussagen überein?	
ST16Q01	a) Im Allgemeinen macht es mir Spaß, mich mit naturwissenschaftlichen Themen zu befassen.
ST16Q02	b) Ich lese gerne etwas über Naturwissenschaften.
ST16Q03	c) Ich beschäftige mich gerne mit naturwissenschaftlichen Problemen.
ST16Q04	d) Ich eigne mir gerne neues Wissen in den Naturwissenschaften an.
ST16Q05	e) Ich bin interessiert, Neues in den Naturwissenschaften zu lernen.

Anmerkung. Antwortoptionen: 1 = „stimme ganz zu“; 2 = „stimme eher zu“; 3 = „stimme eher nicht zu“; 4 = „stimme gar nicht zu“.

A.3.2. Unterrichtsaktivitäten im Naturwissenschaftsunterricht

Tabelle A.7.

„Naturwissenschaftsunterricht: Interaktives Lehren und Lernen“ (Frey & Asseburg, 2009, S. 76)

Item ID	Wortlaut
Wie sehr stimmst du mit den folgenden Aussagen überein?	
ST34Q01	a) Schülerinnen und Schüler bekommen Gelegenheit, ihre Ideen zu erklären.
ST34Q05	e) Im Unterricht geht es auch um die Meinung der Schülerinnen und Schüler zu den Themen.
ST34Q09	i) Im Unterricht findet ein Klassengespräch statt oder es wird diskutiert.
ST34Q13	m) Schülerinnen und Schüler diskutieren über Themen.

Anmerkung. Antwortoptionen: 1 = „in allen Stunden“, 2 = „in den meisten Stunden“, 3 = „in manchen Stunden“, 4 = „nie oder fast nie“.

Tabelle A.8.

„Naturwissenschaftsunterricht: Experimentieren“ (Frey & Asseburg, 2009, S. 77)

Item ID	Wortlaut
Wie sehr stimmst du mit den folgenden Aussagen überein?	
ST34Q02	b) Die Schülerinnen und Schüler führen praktische Experimente im Labor durch.
ST34Q03	c) Die Schülerinnen und Schüler müssen festlegen, wie eine Fragestellung aus dem naturwissenschaftlichen Unterricht im Schullabor untersucht werden könnte.
ST34Q06	f) Die Schülerinnen und Schüler sollen Schlüsse aus einem Experiment ziehen, das sie durchgeführt haben.
ST34Q14	n) Die Schülerinnen und Schüler führen Experimente durch, indem sie den Anweisungen des Lehrers/der Lehrerin folgen.

Anmerkung. Antwortoptionen: 1 = „in allen Stunden“, 2 = „in den meisten Stunden“, 3 = „in manchen Stunden“, 4 = „nie oder fast nie“.

Tabelle A.9.

„Naturwissenschaftsunterricht: Schülerinnen und Schüler forschen“ (Frey & Asseburg, 2009, S. 78)

Item ID	Wortlaut
Wie sehr stimmst du mit den folgenden Aussagen überein?	
ST34Q08	h) Die Schülerinnen und Schüler dürfen ihre eigenen Experimente entwickeln.
ST34Q11	k) Den Schülerinnen und Schülern wird die Möglichkeit gegeben, ihre eigenen Untersuchungen auszuwählen.
ST34Q16	p) Die Schülerinnen und Schüler sollen eine Untersuchung durchführen, um ihre eigenen Ideen auszutesten.

Anmerkung. Antwortoptionen: 1 = „in allen Stunden“, 2 = „in den meisten Stunden“, 3 = „in manchen Stunden“, 4 = „nie oder fast nie“.

Tabelle A.10.

„Naturwissenschaftsunterricht: Naturwissenschaftliche Anwendungen“ (Frey & Asseburg, 2009, S. 80)

Item ID	Wortlaut
Wie sehr stimmst du mit den folgenden Aussagen überein?	
ST34Q07	g) Der Lehrer/die Lehrerin erklärt, wie ein im Unterricht erlerntes naturwissenschaftliches Prinzip auf eine Reihe verschiedener Phänomene angewendet werden kann (z. B. auf die Bewegung von Objekten oder auf Substanzen mit ähnlichen Eigenschaften).
ST34Q12	l) Der Lehrer/die Lehrerin verwendet den naturwissenschaftlichen Unterricht, um den Schülerinnen und Schülern die Welt außerhalb der Schule verständlich zu machen.
ST34Q15	o) Der Lehrer/die Lehrerin erklärt deutlich die Wichtigkeit von naturwissenschaftlichen Konzepten für unser Leben
ST34Q17	q) Der Lehrer/die Lehrerin zeigt am Beispiel technischer Anwendungen, wie wichtig der Unterricht in den naturwissenschaftlichen Fächern für die Gesellschaft ist.

Anmerkung. Antwortoptionen: 1 = „in allen Stunden“, 2 = „in den meisten Stunden“, 3 = „in manchen Stunden“, 4 = „nie oder fast nie“.

A.3.3. Motivationsunterstützende Lehr-Lernbedingungen im Naturwissenschaftsunterricht

Tabelle A.11.

„Auswahl eines naturwissenschaftlichen Fachs“ (Frey & Asseburg, 2009, S. 99)

Item ID	Wortlaut
Wie sehr stimmst du mit den folgenden Aussagen überein?	
ST45N01	Wähle ein naturwissenschaftliches Fach, in dem du zurzeit Unterricht hast: (Bitte nur ein Kästchen ankreuzen.)

Anmerkung. Antwortoptionen: 1 = „in allen Stunden“, 2 = „in den meisten Stunden“, 3 = „in manchen Stunden“, 4 = „nie oder fast nie“.

Tabelle A.12.

„Unterstützende Lehr-Lern-Bedingungen: Wahrgenommene inhaltliche Relevanz“ (Frey & Asseburg, 2009, S. 100)

Item ID	Wortlaut
Wie sehr stimmst du mit den folgenden Aussagen überein?	
ST46N01	a) Im Unterricht wird deutlich, dass der heutige Lernstoff auch im Alltag wichtig ist.
ST46N02	b) Im Unterricht lernen wir, wo die Inhalte praktisch angewendet werden.
ST46N03	c) Im Unterricht werden wir darüber informiert, dass dieser Stoff wichtig für andere Themen ist.

Anmerkung. 1 = „in fast keiner Unterrichtsstunde“, 2 = „in einigen Unterrichtsstunden“, 3 = „in den meisten Unterrichtsstunden“.

Tabelle A.13.

„Unterstützende Lehr-Lern-Bedingungen: Wahrgenommene Kompetenzunterstützung“ (Frey & Asseburg, 2009, S. 105)

Item ID	Wortlaut
	Wie sehr stimmst du mit den folgenden Aussagen überein?
ST46N13	m) Die Lehrerin/der Lehrer traut mir zu, dass ich die naturwissenschaftlichen Inhalte verstehen kann.
ST46N14	n) Die Lehrerin/der Lehrer traut mir auch anspruchsvolle Dinge zu.
ST46N15	o) Der Lehrer/die Lehrerin traut uns zu, dass wir selbstständig arbeiten können.

Anmerkung. 1 = „in fast keiner Unterrichtsstunde“, 2 = „in einigen Unterrichtsstunden“, 3 = „in den meisten Unterrichtsstunden“.

Tabelle A.14.

„Unterstützende Lehr-Lern-Bedingungen: Wahrgenommene Autonomieunterstützung“ (Frey & Asseburg, 2009, S. 106)

Item ID	Wortlaut
	Wie sehr stimmst du mit den folgenden Aussagen überein?
ST46N16	p) Der Lehrer/die Lehrerin ist offen für unterschiedliche Beiträge der Schüler.
ST46N17	q) Der Lehrer/die Lehrerin ermuntert uns, selbst zu überlegen, wie man am besten vorgeht.
ST46N18	r) Der Lehrer/die Lehrerin gibt uns genügend Möglichkeit, selbstständig zu arbeiten.

Anmerkung. 1 = „in fast keiner Unterrichtsstunde“, 2 = „in einigen Unterrichtsstunden“, 3 = „in den meisten Unterrichtsstunden“.