

Die Bedeutung der Aufgaben für
das Beteiligungsverhalten der Schüler -
Eine Videostudie zur Wirksamkeit des Unterrichtsprozesses

Inauguraldissertation

zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Erziehungswissenschaften
der Johann-Wolfgang-Goethe-Universität
zu Frankfurt am Main

vorgelegt von

Johannes Appel
aus Bad Soden - Salmünster

2015

(Einreichungs- und Erscheinungsjahr)

1. Gutachter: Prof. Dr. Udo Rauin

2. Gutachter: Prof. Dr. Eckhard Klieme

Danksagung

An dieser Stelle möchte ich einigen Menschen meinen Dank aussprechen, die auf unterschiedliche Weise zum Gelingen dieser Arbeit beigetragen haben.

Besonderer Dank gilt Herrn Professor Udo Rauin für die Ermöglichung dieser Arbeit und die Unterstützung zu jedem Zeitpunkt im Forschungsprozess. Durch den anregenden Austausch zu verschiedensten Fragen des Themenkomplexes hat er dazu beigetragen, das Interesse für die Sache immer wieder neu zu wecken.

Herrn Professor Eckhard Klieme danke ich für die Bereitschaft, als Zweitgutachter für diese Arbeit zur Verfügung zu stehen

Bei meinen Kolleg(inn)en Nora Boutaoui, Matthias Herrle, Jette Horstmeyer, Tom Schep und Benjamin Tolweth bedanke ich mich für Hilfen unterschiedlichster Art, in der letzten Phase besonders die konstruktive Kritik und die wertvollen Hinweise zum Manuskript.

Des Weiteren danke ich Alina Pergande und Sebastian Breitenbach für die gewissenhafte und engagierte Mitarbeit bei der Aufgabenanalyse. Sebastian Breitenbach sei darüber hinaus für die sachkundige und kreative Hilfe bei der technischen Aufbereitung des Datenmaterials gedankt.

Großer Dank gilt meinem Bruder Christian für die Hilfe bei der Korrektur des Manuskripts. Darüber hinaus sei meiner ganzen Familie für die anhaltende Unterstützung „im Hintergrund“ gedankt. Am meisten danke ich meiner Frau Daniela und meinem Sohn Jonas – für die unermüdliche Aufmunterung und das Verständnis für die vielen Stunden am Schreibtisch.

Inhaltsverzeichnis

1. Einleitung	1
2. Schülerbeteiligungsverhalten als Indikator für Unterrichtswirksamkeit.....	6
2.1 Das Unterrichtsverständnis dieser Arbeit	6
2.2 Theoretische Ansätze zur Wirksamkeit schulischen Unterrichts	10
2.3 Forschungsstand zum Schülerbeteiligungsverhalten im Unterricht.....	18
2.3.1 Schülerbeteiligungsverhalten und Lernerfolg	19
2.3.2 Bedingungen des Schülerbeteiligungsverhaltens.....	26
2.3.3 Forschungsmethodische Aspekte	36
2.3.4 Zusammenfassung des Forschungsstandes und Schlussfolgerungen	40
2.4 Konstrukte und Begrifflichkeiten im Bereich des Schülerbeteiligungsverhaltens ..	42
3. Zeit als empirische Metrik zur Beschreibung von Unterrichtsprozessen.....	53
3.1 Unterrichtszeit aus theoretischer Sicht	53
3.2 Forschung zu Zeitfaktoren des Unterrichts.....	55
3.3 Die Bedeutung des zeitbasierten Forschungsansatzes für die vorliegende Studie ..	61
4. Aufgaben als Einflussgröße für das Schülerbeteiligungsverhalten.....	62
4.1 Die Bedeutung von Lernaufgaben für den Unterricht.....	63
4.2 Das Aufgabenverständnis dieser Arbeit	65
4.3 Ansätze der Aufgabenanalyse	68
4.3.1 Analyseverfahren mit fachdidaktischem Hintergrund.....	69
4.3.2 Analyseverfahren mit lernpsychologischem Hintergrund	73
4.3.3 Analyseverfahren mit allgemeindidaktischem Hintergrund.....	75
4.4 Zusammenfassung und Schlussfolgerungen für diese Studie	77
5. Fragestellungen und Ziele der Studie	80
5.1 Zusammenfassung der Grundüberlegungen und des Desiderats	80
5.2 Ziele der Studie und Einordnung des Untersuchungsansatzes.....	81
5.3 Fragestellungen.....	83
6. Methodische Anlage der Untersuchung	86
6.1 Videoanalysen in der Unterrichtsforschung	86
6.2 Kontext der Untersuchung	90
6.3 Stichprobenbeschreibung	91
6.4 Vorgehensweise bei den Datenerhebungen.....	92
6.5 Testinstrument zur Erfassung der fachlichen Leistung der Schüler.....	95
6.6 Instrument zur beobachtungs-basierten Kodierung des Schülerbeteiligungsverhaltens	96
6.6.1 Voraussetzungen für die Konstruktion des eigenen Kodiersystems	97
6.6.2 Analytischer Zwischenschritt: Basiskodierung der Sozialformen	99
6.6.3 System zur Kodierung des Schülerbeteiligungsverhaltens	101
6.7 Instrument zur beobachtungs-basierten Kodierung und Kategorisierung der Aufgaben.....	106
6.7.1 Kodierung der Analyseeinheit „Aufgabe“	106
6.7.2 Kategoriensystem zur Beschreibung der Aufgabenmerkmale	108
7. Ergebnisse	122
7.1 Personenmerkmale der Schüler.....	123
7.1.1 Sprachlicher Hintergrund in der Familie	123
7.1.2 Geschlecht.....	124
7.1.3 Leistungsstand im Unterrichtsfach.....	125
7.2 Beteiligungsverhalten der Schüler im Unterricht.....	128

7.2.1 Stichprobe	128
7.2.2 Deskriptive Ergebnisse zum Schülerbeteiligungsverhalten	130
7.2.3 Einflüsse von Personenmerkmalen auf die Ausprägung des individuellen Beteiligungsverhaltens	134
7.2.4 Zeitliche Stabilität des Schülerbeteiligungsverhaltens	137
7.3 Aufgaben im Unterricht	141
7.3.1 Stichprobe	141
7.3.2 Deskriptive Ergebnisse zu den Aufgabenmerkmalen auf Ebene der Einzelaufgabe..	142
7.3.3 Deskriptive Ergebnisse zur Variabilität der Aufgabenmerkmale auf Ebene der Unterrichtsstunden	155
7.4 Beziehungen zwischen den Eigenschaften der Aufgaben und dem Beteiligungsverhalten der Schüler im Unterricht	161
7.4.1 Merkmale der Einzelaufgabe und gleichzeitige on task-Rate.....	161
7.4.2 Stundenspezifische Variabilität der Aufgaben und durchschnittliche on task-Rate der Unterrichtsstunde.....	171
8. Zusammenfassung und Diskussion	187
8.1 Zusammenfassung und Einordnung der Ergebnisse	188
8.2 Grenzen der Untersuchung	202
8.3 Ausblick	205
Literaturverzeichnis.....	208
Abbildungsverzeichnis	229
Tabellenverzeichnis	230
Anhang A: Schriftlicher Leistungstest im Unterrichtsfach Englisch	I
Anhang B: Zusammenhänge zwischen Indikatoren der stundenspezifischen Aufgabenvariabilität	V
Anhang C: Mittlere on task-Raten für alle Aufgabenmerkmalsausprägungen.....	XI

1. Einleitung

Verstärkt durch die Ergebnisse internationaler Leistungsvergleichsstudien wie TIMSS¹ und PISA² werden vermehrt erhöhte Erwartungen an die Schule formuliert. Daraus erwachsen zwangsläufig auch Ansprüche an das, was der Kernbestandteil schulischen Geschehens, der Unterricht, zu leisten habe. Während Fragen der Qualität von Schule und Unterricht seitdem auch in der öffentlichen Debatte ähnlich intensiv beachtet werden wie zuletzt zu Zeiten der deutschen Bildungsreform der 1970er Jahre (Helsper & Böhme, 2008), genießt der Themenkomplex um die Effektivität von Unterricht in der Forschung seit Jahrzehnten eine unvermindert hohe Aufmerksamkeit. So ist die ursprünglich in den USA aufgeworfene Debatte um grundsätzliche Zweifel an der Wirksamkeit von Schulen in ähnlicher Weise heute immer noch als aktuell anzusehen (Lüders & Rauin, 2008). Der wissenschaftliche Kenntnisstand in der Schul- und Unterrichtsforschung hat sich indessen erheblich weiter entwickelt, was nicht zuletzt durch verbesserte technische Möglichkeiten wissenschaftlicher Untersuchung befördert wurde. Dies betrifft insbesondere Erhebungstechniken, die auf audiovisuellen Aufzeichnungen von Unterricht basieren. Nach einem anfänglichen Entwicklungsvorsprung in der US-amerikanischen Unterrichtsforschung (Adams & Biddle, 1970) finden so genannte Videostudien in den vergangenen beiden Jahrzehnten auch im deutschsprachigen Raum starke Verbreitung und gelten seitdem allgemein als ein geeignetes Mittel, um Unterricht in seiner Eigenschaft als Geschehen in den Blick zu nehmen und dabei dessen Bestandteile detailliert zu untersuchen (Aufschnaiter & Welzel, 2001; Janik & Seidel, 2009; Klieme, 2006).

Die Übertragung der wissenschaftlichen Befunde der Unterrichtsforschung auf den schulpraktischen „Anwendungshorizont“ ist jedoch laut Ditton (2002, S. 209) als ausbaufähig zu bezeichnen. Für eine Rückführung der Erkenntnisse in das zugrunde liegende Feld sei *„die wechselseitige Verschränkung von Zugangsweisen und Ebenen, die für die Unterrichtsrealität konstitutiv ist, zu reflektieren“* (ebd.). Eine Konsequenz dessen könnte darin liegen, dass in Teilen der Forschung eine empirische Perspektive eingenommen wird, die enge Bezüge zum konkreten Unterrichtshandeln der Beteiligten

¹ TIMSS = *Third International Mathematics and Science Study* (BMBF, 2001)

² PISA = *Programme for International Student Assessment* (PISA-Konsortium Deutschland, 2003)

zulässt und damit an diesen Stellen die gemeinsame Kommunikation über die untersuchten Phänomene erleichtert. Von besonderem Interesse wäre diesbezüglich die Betrachtung von Geschehenskomponenten in ihrer Ausprägung im Unterrichtsverlauf, von denen anzunehmen ist, dass sie von zentraler Bedeutung für die Wirksamkeit des Unterrichtsprozesses sind.

An diesem Punkt setzt die vorliegende Arbeit an, indem im Rahmen einer empirischen Untersuchung natürlichen Unterrichts das Beteiligungsverhalten der Schüler als die zentrale Größe angesetzt wird, an der sich die Wirksamkeit der einzelnen Unterrichtsstunde misst. Diese Entscheidung hat verschiedene Gründe: Das Schülerbeteiligungsverhalten bietet für Lehrkräfte einen wichtigen Orientierungspunkt beim Handeln in der Situation. Im Gespräch mit Lehrkräften zeigt sich oftmals, dass die Einschätzung darüber, ob eine Unterrichtsstunde gelungen sei bzw. „funktioniert habe“, daran festgemacht wird, inwiefern die Schüler³ erkennbar „mitgemacht“ haben oder „bei der Sache“ waren. Auch in der Unterrichtsforschung wird dem Beteiligungsverhalten der Schüler eine hohe Relevanz beigemessen. Wie noch zu zeigen ist, wird dem Bereich in einer Reihe von Modellen zur Wirksamkeit des Unterrichts eine zentrale Position zugewiesen, und einer breiter Bestand an empirischen Ergebnissen belegt dessen Bedeutung für individuelle, unterrichtsbedingte Lernerträge⁴. Im Hinblick auf die wissenschaftliche Untersuchung liegt überdies ein großer Vorteil dieses Bestandteils des Unterrichtsgeschehens darin, dass er individualisiert und in seiner zeitlichen Dimension erfassbar ist. Damit ist es grundlegend möglich, potenziell unterschiedliche Grade der Unterrichtswirksamkeit in Lerngruppen unter Berücksichtigung heterogener Lernvoraussetzungen der Schüler und unterschiedlicher Zeitabschnitte des Unterrichts in den Blick zu nehmen.

Als zweite Prozessvariable neben dem Schülerverhalten werden in dieser Studie die Aufgaben, verstanden als Lernaufgaben innerhalb des Unterrichtsprozesses, betrachtet. Aufgaben sind ein bedeutendes didaktisches und pädagogisches Instrument zur Steuerung des Unterrichtsverlaufs. Sie sind gleichzeitig ein Planungs- und

³ Aus Gründen besserer Lesbarkeit werden in dieser Arbeit Personenbezeichnungen in nur einer Geschlechtsform geführt, wobei – sofern nicht anders gekennzeichnet – in jedem Fall das jeweils andere Geschlecht mit gemeint ist.

⁴ Ausführliche Darstellungen dazu erfolgen im Rahmen von Kap. 2.

Durchführungselement des Unterrichts und bieten Lehrkräften die Möglichkeit, Lernprozesse und Aktivitäten im Unterrichtsverlauf zu strukturieren. An dieser Stelle wird angenommen, dass die Eigenschaften der in den Unterricht eingebrachten Aufgaben einen bedeutenden Einfluss auf das Beteiligungsverhalten ausüben. Für eine Prüfung dieses Zusammenhangs werden beide Größen detailliert auf ihre Ausprägungen im Unterrichtsverlauf hin untersucht und zueinander in Beziehung gesetzt. Als adäquate forschungsmethodische Herangehensweise für eine prozessensible Erfassung dieser Größen werden videobasierte Beobachtungsverfahren eingesetzt. Ziel dieser Untersuchung ist es damit, durch einen Blick in die „*black box*“ der Unterrichtsprozesse (Ditton, 2002, S. 201) Zusammenhänge der Unterrichtswirksamkeit aufzuzeigen, die sich an einem dem Lernerfolg der Schüler – als üblichem Wirksamkeitskriterium der Unterrichtsqualitäts- und Lehr-Lern-Forschung – vorgelagerten Punkt verorten lassen. Damit sieht sich die vorliegende Studie als ein Beitrag zur Beschreibung von Zusammenhängen zwischen zwei zentralen wirksamkeitsrelevanten Geschehenskomponenten des Unterrichts und zugleich zur Erzeugung wissenschaftlicher Ergebnisse, die eine gewisse Nachvollziehbarkeit und Anschlussfähigkeit für die Akteure der Praxis in sich bergen.

Mit Blick auf diese Zielsetzung beginnt die Arbeit in Kap. 2 mit einer Reihe theoretischer Grundlegungen. Auf eine kurze Darlegung des zugrundeliegenden Unterrichtsverständnisses (2.1) folgt in Kap. 2.2 die Beschreibung von in diesem Bereich international als zentral zu erachtenden Unterrichtsmodellen (Carroll, 1963; Slavin, 1994) und deren Einfluss auf vergleichbare Konzeptionen im deutschsprachigen Bereich (Helmke, 2003, 2012; Klieme, Lipowsky, Rakoczy & Ratzka, 2006; Klieme, Schümer & Knoll, 2001). In diesem Zuge erfolgt eine theoretische Verortung des Schülerbeteiligungsverhaltens in den Gesamtzusammenhang der Forschung zur Unterrichtswirksamkeit.

Im nächsten Schritt wird in Kap. 2.3 der Forschungsstand zum Schülerbeteiligungsverhalten umfassend in den Blick genommen und die empirische Befundlage mit Blick auf die vorherigen theoretischen Betrachtungen geprüft. An dieser Stelle wird die für diese Studie relevante Forschungslücke benannt, die in der Untersuchung der Qualität der Aufgaben im Unterricht in ihrem Einfluss auf das Beteiligungsverhalten der Schüler besteht. In Kap. 2.4 werden dann zunächst Differenzierungen hinsichtlich einiger synonym füreinander bzw. eng miteinander

diskutierter Begriffe und Konstrukte im Bereich des Schülerbeteiligungsverhaltens vorgenommen, was mit Blick auf den Forschungsstand und die Entwicklung einer eigenen forschungsmethodischen Perspektive notwendig erscheint.

Darauf folgt in Kap. 3 eine getrennte Diskussion des mit dem vorliegenden Themenkomplex eng verknüpften Aspekts der Unterrichtszeit und der Benennung von dessen Bedeutung für die vorliegende Arbeit.

Kap. 4 wendet sich als nächstes dem Gegenstand der Aufgaben im Unterricht zu. Dabei wird aufgezeigt, dass Aufgaben die geeignete unterrichtliche Prozesskomponente darstellen, in der sich die Hinweise aus der empirischen Forschung mit den theoretischen Annahmen von Slavin (1994) verbinden lassen (4.1). Hinsichtlich der genaueren begrifflichen Auffassung werden „im Rahmen des Unterrichtsverlaufs von der Lehrkraft eingebrachte Lernaufgaben“ als der fokussierte Gegenstand expliziert (4.2). Darauf folgt in Kap. 4.3 ein Überblick über Ansätze zur Analyse von Aufgaben. Diese werden mit Blick auf (vorrangig methodische) Anschlussmöglichkeiten für die vorliegende Untersuchung diskutiert.

In Kap. 5 werden die grundlegenden Betrachtungen der vorangegangenen Abschnitte zusammengefasst und die Ziele und spezifischen (Teil-)Fragestellungen dieser Studie dargelegt. Die darin formulierte Kernfrage dieser Studie lautet, inwiefern sich das mit Blick auf die Individuen vollständige und in den Beobachtungszeiträumen lückenlos erfasste Beteiligungsverhalten der Schüler durch die Eigenschaften der im Unterrichtsverlauf identifizierbaren Aufgaben erklären lässt.

Kap. 6 umfasst die Darstellung der methodischen Anlage der Untersuchung. Die Studie stützt sich im Wesentlichen auf wissenschaftlich-systematische Beobachtungen des Handelns der am Unterricht Beteiligten anhand von Videoaufnahmen, weshalb dieser forschungsmethodische Zugang zu Beginn des Kapitels überblicksartig beschrieben wird (6.1). Die Studie bezieht aber auch schriftliche Test- bzw. Befragungsdaten auf Seiten der Schüler mit ein. Die verwendeten Daten stammen aus dem Forschungsprojekt „Strategien des Unterrichts in heterogenen Klassen und ihre Wirkung auf Schüleraktivität“, welches unter der Leitung von Prof. Dr. Udo Rauin am Institut für Pädagogik der Sekundarstufe am Fachbereich Erziehungswissenschaften und der Akademie für Bildungsforschung und Lehrerbildung der Goethe-Universität Frankfurt am Main durchgeführt wird (6.2). Da die im Rahmen der vorliegenden Arbeit verwendeten Beobachtungsinstrumentarien zur Erfassung und Beschreibung von Schülerbeteiligungsverhalten und Aufgaben im Unterricht zum Großteil auf

Eigenentwicklungen und Adaptionen bestehender Instrumente basieren, werden diese Instrumente in den Abschnitten 6.6 und 6.7 ausführlich dargestellt.

Der sich im Folgekapitel (Kap. 7) anschließende Ergebnisbericht gliedert sich auf in eine Darstellung der Einzelergebnisse zu den verschiedenen Teilinstrumenten der Studie: Kap. 7.1 zeigt die mittels schriftlicher Tests erhobenen Leistungsstände der Schüler im Unterrichtsfach sowie die Ausprägungen der gleichzeitig erfassten Personenmerkmale in der Stichprobe. Kap. 7.2 stellt die Ergebnisse zu den Beobachtungen des Schülerverhaltens dar. Kap. 7.3 zeigt die Befunde der Erfassung der Aufgaben im Unterricht und der Beschreibung ihrer Eigenschaften. Kap. 7.4 beinhaltet schließlich die Hauptbefunde dieser Untersuchung, d.h. die Ergebnisse der Prüfung von Einflüssen der Aufgaben auf das Beteiligungsverhalten der Schüler.

Im abschließenden Kapitel (Kap. 8) werden die gefundenen Ergebnisse zusammengefasst dargestellt, in den theoretischen Hintergrund eingeordnet und bewertet (8.1), worauf eine Darlegung der Grenzen der Untersuchung (8.2) und ein Ausblick auf sich anschließende Fragen und weitere Forschungsperspektiven folgt (8.3).

2. Schülerbeteiligungsverhalten als Indikator für Unterrichtswirksamkeit

In diesem Kapitel wird der spezifische Zugang der vorliegenden Arbeit zur Abbildung eines Wirksamkeitskriteriums innerhalb des Unterrichtsprozesses begründet. Dazu wird zunächst die zugrunde liegende Auffassung des Gegenstandes „Unterricht“ dargelegt und eine Einordnung der Zugangsweise in die Unterrichtsforschung vorgenommen (2.1). Nachfolgend werden verschiedene Modelle zu den Wirkzusammenhängen schulischen Unterrichts diskutiert und auf dieser Basis der Merkmalsbereich des Schülerbeteiligungsverhaltens als prozessimmanentes Kriterium der Unterrichtswirksamkeit identifiziert (2.2). Die Darstellung des Forschungsstandes dient anschließend zur Feststellung möglicher Forschungsdesiderate (2.3), woraus Schlüsse für die Fragestellungen und die Anlage dieser Studie gezogen werden. Abschließend wird eine systematische Betrachtung der gängigen Begrifflichkeiten und Konstrukte in dem Merkmalsbereich geleistet (2.4), die zur Präzisierung der Schwerpunktsetzung dieser Studie dient.

2.1 Das Unterrichtsverständnis dieser Arbeit

Die leitende Frage dieses Abschnitts ist, was unter „Unterricht“ zu verstehen ist und wie dieser Gegenstand mit Blick auf die übergeordneten Ziele der Arbeit (s. Kap. 1) auf angemessene Weise durch Forschung erschlossen werden kann.

„Was U. [Unterricht] ist, scheint auf den ersten Blick intuitiv gewiss. Trotzdem ist U. ein Begriff mit unscharfen Rändern, er ist ein theoretisches Konstrukt, (...)“ (Meyer, 2007, S. 728, eigene Ergänzung in eckigen Klammern). Damit deutet Meyer auf die Tatsache hin, dass zwar ein den Alltagserfahrungen entsprechender Konsens über die Beschaffenheit von Unterricht zu existieren scheint, allerdings gestaltet sich eine genauere Beschreibung des Gegenstands auf wissenschaftlich-differenzierender Ebene deutlich schwieriger. Lüders (2012) hat in einer Untersuchung zur Verwendung des Unterrichtsbegriffs in pädagogischen Nachschlagewerken im Zeitraum zwischen 1949 und 2007 festgestellt, dass zumindest die Schulpädagogik offenbar kein einheitliches, theoretisch begründetes Begriffsverständnis hervorgebracht und dort auch keine durchgängig aufeinander aufbauende Begriffsentwicklung stattgefunden habe. Der

Unterrichtsbegriff werde „in nahezu jedem Artikel mehr oder weniger neu erfunden“ (S. 123). Eine Gemeinsamkeit der meisten Unterrichtsdefinitionen bestehe zudem darin, dass sie durchgängig „an der Praxis der Profession ausgerichtet“ (ebd.) seien, sprich im Grunde die Perspektive von Lehrerinnen und Lehrern widerspiegeln. „In der Schulpädagogik und Didaktik wird Unterricht vorwiegend als eine an die Existenz moderner Bildungssysteme gebundene, institutionalisierte Form des Lehrens und Lernens begriffen“ beschreibt Lüders (2011, S. 342) den Gegenstand in einem lexikalischen Lemma. Zum Vergleich findet sich bei Lipowsky (2009) die Formulierung „*Unterricht kann als langfristig organisierte Abfolge von Lehr- und Lernsituationen verstanden werden, die von ausgebildeten Lehrpersonen absichtsvoll geplant und initiiert werden und die dem Aufbau von Wissen sowie dem Erwerb von Fertigkeiten und Fähigkeiten der Lernenden dienen. Sie finden in der Regel in bestimmten dafür vorgesehenen Institutionen unter regelhaften Bedingungen statt*“ (S. 74). Diese Definition konkretisiert die bei Lüders genannten Aspekte und ordnet Unterricht als eine Sequenz von Situationen ein, benennt als Zielperspektive Wissensaufbau und Fähigkeitserwerb und betont die institutionelle Organisiertheit von Unterricht. Wie auch in vergleichbaren anderen Beiträgen (vgl. Helsper & Keuffer, 1995) werden in den beiden Begriffsauffassungen Lehren und Lernen als zentrale Bezugspunkte herausgestellt. Ungeachtet grundlegender Übereinstimmungen stehen diese Beschreibungen aber keineswegs repräsentativ für einen breiten disziplinären Konsens (vgl. Lüders, 2012). Verlässt man die Ebene globaler Definitionsversuche und betrachtet Unterrichtsbegriffe im Rahmen sozial- bzw. geisteswissenschaftlicher theoretischer Verortungen, werden erneut differente Sichtweisen deutlich. So sehen etwa Meseth, Proske und Radtke (2011) unter Anknüpfung an Kommunikationstheorien und unter Betonung des zentralen Problems der Kontingenz Unterricht als „hoch komplexes, nicht-lineares soziales Geschehen“ (Meseth, Proske & Radtke, 2011, S. 224). Ebenfalls mit Bezug auf das Moment der Kommunikation bzw. der Interaktion entwirft Lüders (2003, 2011) Unterricht in seiner Sprachspieltheorie als ein Geschehen, welches sich in seinen „kommunikativen Praktiken“ (Lüders, 2011, S. 175) von anderen Interaktionsanlässen des sozialen Lebens unterscheiden lässt. Interaktionen zwischen Lehrperson und Schülern seien demgemäß entlang von „sprachlichen Akten und Gesten“ (ebd.) zu beschreiben, die nach bestimmten Regeln, so genannten „Zugmustern“ bzw. „Äußerungsfolgen“ auftreten (ebd., S. 178). Auch andere Autoren gründen ihr Unterrichtsverständnis auf Kommunikationsaspekte, wengleich mit differenten Schwerpunktsetzungen, wie z.B. Caruso mit einer

gesellschaftshistorisch orientierten Sichtweise auf schulischen Unterricht „als eine *Technologie der lernbezogenen Massenhaltung von Menschen*“ (Caruso, 2011, S. 24). Diese Beispiele verdeutlichen die Bandbreite an spezifischen Schwerpunktsetzungen in der Perspektive auf Unterricht, selbst unter vergleichbaren Grundbegriffen. Ansätze wie diese bieten allerdings keine geeigneten Anknüpfungspunkte für die empirische Bearbeitung von Wirksamkeitsannahmen schulischen Unterrichts.

Eine größere Annäherung an diesen Punkt ergibt sich bei einer Veränderung des Blickwinkels hin zur allgemeinen Didaktik, die innerhalb der Erziehungswissenschaft das zentrale Feld der Auseinandersetzung mit Unterricht darstellt und dabei enge Bezüge zur Unterrichtspraxis aufweist (vgl. Kron, 2008). Die allgemeine Didaktik beschäftigt sich neben Aspekten der Unterrichtsplanung und -analyse auch mit der Legitimation von Zielen und Inhalten des Unterrichts unter Berücksichtigung gesellschaftlicher Zusammenhänge, wozu z.B. auch der Lehrplan als ein vermittelndes Element von „*globalen Ziel- und Wertvorstellungen*“ (Helsper & Keuffer, 1995, S. 82) gehört. Indem über den Einbezug solcher Aspekte oftmals „*normative Argumentationen über den moralischen und gesellschaftlichen Zweck von Schule und Unterricht eingewoben sind*“ (Terhart, 2002, S. 78), wird Schulunterricht im Kontext der allgemeinen Didaktik immer auch als Ermöglichung von Bildungsprozessen gesehen. Im Anschluss daran hebt Gruschka solche Zieldimensionen wie Mündigkeit und Urteilskraft der Schüler hervor und betont, Unterricht sei „*mehr als der didaktisch inszenierte Übergang vom Lehren zum Lernen*“ (Gruschka, 2011, S. 131). Innerhalb der allgemeinen Didaktik liegt ein breites Spektrum an spezifischen Ansätzen vor, repräsentiert durch eine Vielzahl an Modellen und Gegenstandstheorien mit teils weit zurück reichender Historie (Kron, 2008). Als ein verbindendes Element dieser Vielzahl an Konzeptionen kann die Orientierung am Prozess der Ausbildung von zukünftigen Lehrkräften gesehen werden (Terhart, 2002). Demzufolge geht es in der allgemeinen Didaktik vornehmlich „*um die Theoretisierung und operative Gestaltung von Lehren und Lernen im Kontext von Ausbildung für den pädagogischen Beruf des Lehrers*“ (S. 80). Das bedeutet, dass neben grundlegenden theoretischen Erwägungen vor allem praxisorientierte Aspekte des Unterrichtshandelns zum Gegenstand der Überlegungen gemacht werden, wobei die allgemeine Didaktik in ihrer grundsätzlichen Ausrichtung eher als ein „*geisteswissenschaftlich geprägtes Nachdenken über Unterricht*“ (Kiel, 2010, S. 781) zu verstehen ist denn als ein Programm zur empirischen Erforschung des selben.

Demgegenüber ist die Teildisziplin der Lehr-Lern-Forschung ebenfalls befasst mit dem Gegenstand des Unterrichts, macht sich in diesen Züge aber die empirisch basierte Erkenntnisgewinnung über Anlässe des Lehrens und Lernens (und damit auch schulischen Unterrichts) explizit zur Aufgabe (Terhart, 2002). In diesem eng an der pädagogischen Psychologie orientierten Forschungsbereich werden zwar nur selten explizite Unterrichtsbegriffe entfaltet (vgl. Eickhorst, 2011), als gemeinsames Grundverständnis ist jedoch der kognitive Lernzuwachs als zentrale Zielkategorie von schulischem Unterricht erkennbar. Exemplarisch dafür formulieren Hasselhorn und Gold in einem Lehrbuch der Pädagogischen Psychologie: *„Unterricht hat demnach die Aufgabe, in systematischer Weise aufsteigend die zum Verstehen eines Sachverhalts erforderlichen kognitiven Prozesse auszulösen und damit das Erreichen der Teilziele sicherzustellen oder zumindest wahrscheinlicher zu machen“* (Hasselhorn & Gold, 2013, S. 233). Der Forschungsstrang der Lehr-Lern-Forschung hat innerhalb der vergangenen fünf Jahrzehnte eine Fülle an empirisch basierten Erkenntnissen über Unterrichtsprozesse, deren Bedingungen und Wirkungen hervorgebracht (im Überblick: Renkl, 2010; Treiber & Weinert, 1982). Dabei nimmt die Lehr-Lern-Forschung eine variablenorientierte Perspektive (Helmke, 2012) auf mitunter kleinste Geschehensbestandteile des Unterrichts ein (s. dazu die Begriffe *„Mikro-Ebene“*, Terhart, 2002, S. 79 und *„Mikroforschung“*, Lüders & Rauin, 2008, S. 729) und untersucht diese verschiedenen Unterrichtsmerkmale auf ihre wechselseitigen Einflüssen und Wechselwirkungen hin, auch im Hinblick auf Zielkategorien der Qualität oder Effektivität des Unterrichts.

Für die vorliegende Arbeit stellt die Forschungsstrategie der Lehr-Lern-Forschung einen aussichtsreichen Ansatz dar. Indem in dieser Untersuchung aber die Perspektive auf konkretes Unterrichtshandeln gerichtet und auf diesem Wege ein Beitrag zur *„Aufklärung und Förderung von Praxis“* (Kron, 2008, S. 42) geleistet werden soll, bestehen somit auch Bezüge zur allgemeinen Didaktik. Aus diesem Grund kann mit Blick auf die vorliegende Studie eher von didaktisch orientierter Lehr-Lern-Forschung gesprochen werden: *„Eine Grundfrage der didaktisch orientierten LLF lautet deshalb, wie effektive Lehr-Lern-Prozesse gestaltet werden können um in der verfügbaren Unterrichtszeit mit den verfügbaren Ressourcen für möglichst viele Schüler möglichst viele erwünschte Ziele unter Vermeidung unerwünschter Nebeneffekte erreichbar zu machen“* (Lüders & Rauin, 2008, S. 730). Mit Bezug auf die eingangs gestellte Frage ist

abschließend festzuhalten, dass das Gegenstandsverständnis dieser Untersuchung einem in der Schulpädagogik verbreiteten, im Wesentlichen „an den praktischen Belangen der Lehrerverberufung“ (Lüders, 2012, S. 109) ausgerichteten Unterrichtsbegriff folgt, während mit Unterricht explizit „der Unterrichtsprozeß in Schulen gemeint ist, wobei das (staatliche) Pflichtschulsystem im Mittelpunkt steht“ (Terhart, 2002, S. 136). Diese Auffassung impliziert – auch im Anschluss an die Forschungslinie der Lehr-Lern-Forschung – die Annahme einer zentralen Beeinflussbarkeit des Geschehens durch die Handlungen der Lehrkraft.

2.2 Theoretische Ansätze zur Wirksamkeit schulischen Unterrichts

Die wohl bedeutendste theoretische Grundlage für Fragen der Wirksamkeit von Schulunterricht ist das *model of school learning* von Carroll (1963). Dieses Modell beschreibt den Prozess schulischen Lernens unter besonderer Berücksichtigung der zeitlichen Dimension des Unterrichts und der Perspektive des einzelnen Schülers. Demnach ist das Lernen im Unterricht bedingt durch den Zeitanteil, der von den Schülern im Unterricht tatsächlich auf lernzielbezogene Aktivitäten verwendet wird (*time actually spent*), im Verhältnis zu der Zeit, die sie eigentlich zum Erreichen dieses Ziels benötigen (*time needed*). Mit anderen Worten geht es um eine Art Soll-Ist-Verhältnis hinsichtlich der mit der Beschäftigung mit Unterrichtsinhalten verbrachten Zeit durch den individuellen Schüler. Die Bestandteile des Modells setzen sich aus Unterfaktoren zusammen, die den jeweiligen Bereich in seiner Ausprägung bedingen. Die Zusammenhänge zwischen den Elementen des Modells lassen sich nach Art einer Formel schematisch darstellen (s. Abb. 1).

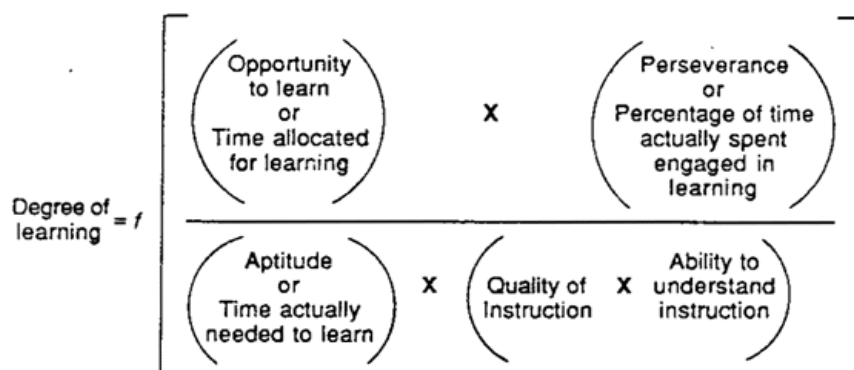


Abb. 1: Schematische Darstellung von Carrolls Modell schulischen Lernens (nach Berliner, 1990, S. 14)

Für die benötigte Lernzeit sind dies (a) das aktuelle Fähigkeitsniveau und die Auffassungsgabe als kognitive Voraussetzungen des Schülers (*aptitude* und *ability to understand instruction*) und (b) die Qualität des Unterrichts als förderlicher bzw. hemmender Faktor in der Beziehung zwischen Lerngegenstand und Schüler (*quality of instruction*), während die tatsächlich aufgewendete Lernzeit bedingt wird durch (a) das Ausmaß der zur Verfügung stehenden Lerngelegenheiten in Form von verfügbarer Lernzeit (*opportunity time for learning*) und (b) der so genannten Ausdauer des Schülers, repräsentiert durch Zeit, die er bereit ist, in Lernaktivitäten zu investieren (*perseverance*). Letzteres ist laut Carroll letztlich Ausdruck der individuellen Lernmotivation. Als wesentliche Bedingung des Lernens im Unterricht wird in diesem Zuge die Beschäftigung der Schüler mit den inhaltlichen Anforderungen des Unterrichts, den *learning tasks* (Carroll, 1963, S. 723), herausgestellt, was wie folgt umrissen wird: „... the time during which the person is oriented to the learning task and actively engaged in learning. In common parlance, it is the time during which he is “paying attention“ and “trying to learn“ (S. 724). Die Zuwendung der Schüler zu den Lerninhalten des Unterrichts wird damit zur Schlüsselgröße eines möglichen, durch den Unterricht initiierten Lernerfolgs. Carrolls Modell lieferte einen in mehreren Punkten einflussreichen Beitrag für die Unterrichtsforschung: Die Stärkung eines individualisierten Blicks auf Lernprozesse im Unterricht und die damit verbundene Idee der Passung von individuellen Lernbedürfnissen und Lernbedingungen, der Ansatz der allgemeinen Beschreibung des schulischen Lernens auf der Basis von Zeitbegriffen und die Bedeutung aktiven Zuwendung der Schüler zu den Aufgaben als grundlegende Lernbedingung sind jene herausragenden Gedanken, die sowohl in Kombination als auch jeweils für sich genommen Gegenstand weitergehender Auseinandersetzungen sein sollten⁵.

Annähernd parallel zu Carroll lieferte Jackson (1968/1990) weitere Impulse zu diesem Themenkomplex, wenn auch ohne expliziten Einbezug der zeitlichen Komponente. Er betont die entscheidende Rolle der „Versunkenheit“ der Schüler in die Lernaufgaben als

⁵ Mit Blick auf die Argumentationslogik soll an dieser Stelle jedoch zunächst der Aspekt des Beteiligungsverhaltens weiter verfolgt werden, während die Diskussion um die Zeitdimension des Unterrichts in Kap. 3 getrennt aufgegriffen wird.

vorrangiges Ziel des Unterrichts: „*Certainly no educational goals are more immediate than those that concern the establishment and maintenance of the student's absorption in the task at hand. Almost all other objectives are dependent for their accomplishment upon the attainment of this basic condition*“ (Jackson 1968/1990, S. 85). Die Aufmerksamkeit bzw. die Beteiligung der Schüler ist demnach eine Grundbedingung für mögliche individuelle Erträge des Unterrichts und zugleich Hauptaufgabe der Unterrichtsgestaltung durch die Lehrkraft: „*The teacher's goal, in other words, is student involvement rather than student learning. It is true, of course, that the teacher hopes the involvement will result in certain beneficial changes in the students, but learning is in this sense a by-product rather than the things about which the teacher is most directly concerned*“ (S. 24). Jackson setzt sich aus sozialpsychologischer Sichtweise mit den Bedingungen auseinander, die der Beteiligung der Schüler zugrunde liegen. In diesem Zuge diskutiert er eine Reihe von Charakteristika der Schule, die in der Summe ein Art inoffizielles, so genanntes „*hidden curriculum*“ (S. 33f.) bilden und dazu führen, dass die Schüler mit der Zeit lernen, Verhaltensstrategien zu entwickeln, um sich in diesem Umfeld erfolgreich zu bewegen. Eine bedeutende Rolle spielt dabei u.a. die Strategie des bewusst gesteuerten, aber von außen nicht beobachtbaren psychologischen Rückzugs der Schüler aus dem Unterrichtsgeschehen („*the strategy of psychological withdrawal*“, S. 37). Möglichkeiten zur Beeinflussung des Beteiligungsverhaltens der Schüler im Unterricht sind laut Jackson vornehmlich in der Klassenführung und -organisation und der Angemessenheit der Unterrichtsinhalte in Bezug auf die Schüler zu suchen (S. 103). Jackson lieferte mit seiner Arbeit einen wichtigen Beitrag zum Verständnis der Verhaltensmechanismen auf Seiten der Schüler einerseits und der Einflussmöglichkeiten der Lehrkraft im Unterricht andererseits. Seine Vorstellungen dessen, was die Wirksamkeit von Unterricht innerhalb seines Geschehens ausmacht, schließen zum einen an die Grundgedanken Carrolls an und lassen sich zum anderen auch in nachfolgenden theoretischen Weiterentwicklungen wiederfinden.

Mit der Absicht, die zentralen Annahmen von Carrolls Modell schulischen Lernens weiter zu entwickeln, ist Slavins Modell zur Effektivität von Unterricht (Slavin, 1984, 1994) entstanden. Die Weiterentwicklungen bestehen zum einen darin, dass in der Formulierung des Modells der zum damaligen Zeitpunkt aktuelle Forschungsstand systematisch Berücksichtigung fand, und zum anderen in der grundlegenden Ausrichtung der Modellelemente an dem Grad ihrer Veränderbarkeit im Unterricht

durch das Handeln der Lehrkraft. Dieses so genannte *QAIT-Modell* basiert auf vier Komponenten effektiven Unterrichts: Unterrichtsqualität (*Quality of instruction*), Angemessenheit (*Appropriate levels of instruction*), Anreizwert (*Incentive*) und Zeit (*Time*, vgl. Abb. 2).

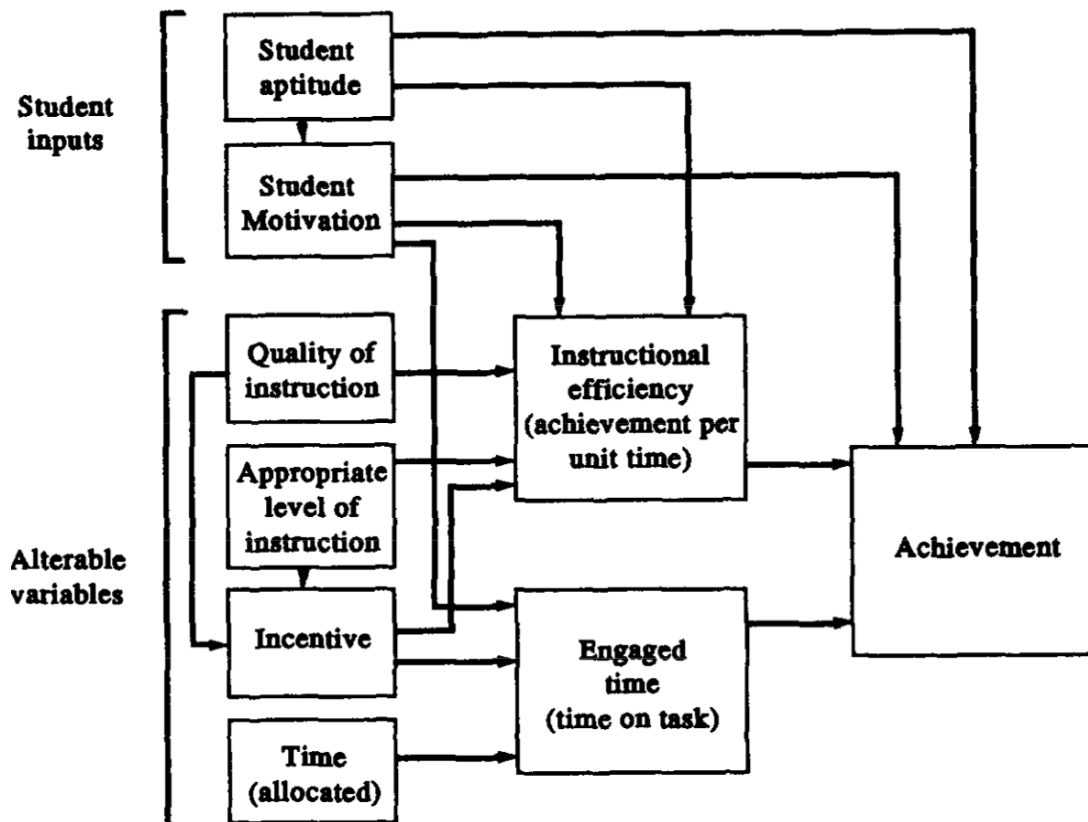


Abb. 2: Das QAIT-Modell effektiven Unterrichts (nach Slavin 1994, S. 149)

Zusätzlich zu den vier innerhalb des Unterrichts angesiedelten Variablen werden Vorbedingungen auf Seiten der Schüler mit berücksichtigt (*student inputs*), repräsentiert durch den aktuellen Stand an Fähigkeiten und Kompetenzen, Vorkenntnisse und die allgemeine Lernmotivation. Der Schwerpunkt des Modells liegt auf den Gestaltungsspielräumen der Lehrkraft innerhalb des Unterrichts und damit auf den vier QAIT-Größen. Unterrichtsqualität umfasst dabei eine klare, konsistente, abwechslungsreiche und attraktive Darstellung der Lehrinhalte und gute Organisation der geplanten Unterrichtsaktivitäten. Die Angemessenheit zielt darauf ab, dass die Schüler die angestrebten Lerninhalte nicht bereits beherrschen, zugleich aber die

nötigen Vorkenntnisse und Fähigkeiten für diesen Schritt besitzen – sprich: der Schwierigkeitsgrad ist angemessen. Der Anreizwert steht für Bemühungen der Lehrkraft, den Erkenntnisdrang der Schüler zu aktivieren und sie damit zu den beabsichtigten Lernaktivitäten motivieren. Dies kann geschehen durch interessante und attraktive Darstellungen durch die Lehrkraft, der Förderung von Eigenverantwortung und Selbständigkeit der Schüler im Lernprozess, die Steigerung der subjektiv wahrgenommenen Sinnhaftigkeit der Inhalte über die Herstellung von Bezügen zur Lebenswelt der Schüler und schließlich durch die Sicherstellung hoher Bewältigungschancen für die Schüler in der Bearbeitung der Aufgaben. Besonders letztgenannter Aspekt, der erneut den angemessenen Schwierigkeitsgrad der Lerninhalte aufgreift, sowie der Aspekt der interessanten Darstellungsweise zeigt eine enge Verschränkung des Anreizwertes mit den Komponenten Unterrichtsqualität und Angemessenheit. Slavin hebt hervor, dass diese drei Elemente (Unterrichtsqualität, Angemessenheit und Anreizwert) letztlich nicht vollständig getrennt voneinander gedacht werden können und besonders der Anreizwert und die Angemessenheit eng miteinander verknüpft sind: *„Incentive is high when the level of instruction is appropriate for a student, so that the student perceives that with effort the material can be mastered, so that the payoff for effort is perceived to be great“* (Slavin, 1994, S. 147). Die letzte Komponente der Unterrichtseffektivität ist die Unterrichtszeit, d.h. die innerhalb des Unterrichts mit der Arbeit an den Inhalten verwendete Zeit. Diese Komponente bildet sozusagen die Quantität des Angebots an Lerngelegenheiten ab, deren Qualität durch die drei übrigen Komponenten beschrieben wird.

Der zentrale Gedanke des Modells ist jener eines multiplikativen Zusammenhangs der vier Effektivitätskomponenten und damit der Abhängigkeit des unterrichtsbasierten Lernerfolgs von einer notwendigen Grundaussprägung aller Größen (*„If any of the elements is at zero, learning will be zero.“*, S. 150). Zwischen den vier durch die Lehrkraft zu beeinflussenden Komponenten effektiven Unterrichts und dem Lernerfolg der Schüler sind zwei vermittelnde Größen positioniert, *instructional efficiency* und *engaged time*. Die Variable *instructional efficiency* führt die Komponenten Unterrichtsqualität, Angemessenheit und Anreizwert des Unterrichts zusammen und stellt sie in Relation zu den persönlichen Voraussetzungen der Schüler, den *student inputs*, womit diese Größe verstanden werden kann als potenzielle Lernwirksamkeit des Unterrichts für den einzelnen Schüler innerhalb eines bestimmten Zeitraums. Zweite zentrale

Moderatorvariable ist die *engaged time* bzw. die *time on task* als zeitbasiert gefasste Beteiligung an bzw. Beschäftigung mit Aufgaben und Unterrichtsinhalten: „*Engaged time is the amount of time students are actually participating in relevant learning activities, such as paying attention to lectures and doing assignments*“ (S. 149). Einflussgrößen dafür sind die zur Verfügung stehende Lernzeit, der Anreizwert des Unterrichts und, als außerhalb des Unterrichtsgeschehens angesiedelte Größe, die grundsätzliche Motivationslage der Schüler.

Ein Verdienst dieses Modells besteht darin, dass es auf sparsame Weise ein konsistentes Bild der Zusammenhänge zwischen den beiden Ebenen des Handelns der Lehrkraft und der Schüler im Unterricht zeichnet. Zugleich wird darin deutlich, dass für eine Annäherung an die Wirksamkeit von Unterricht bzw. einer einzelnen Unterrichtsstunde der Blick auf den Bereich der Schüleraktivitäten gerichtet werden muss. Die Annahmen aus Slavins QAIT-Modell lassen sich in einer Reihe weiterer, nachfolgender Arbeiten zu Fragen der Unterrichtseffektivität wiederfinden. So ist der Grundgedanke der *time on task* als vermittelnde Größe zwischen Schülerperson, Unterrichtseigenschaften und möglichen Lern- bzw. Bildungseffekten auch in einem Modell von Creemers und Kyriakides (2008) zu finden, welches unter dem breiteren Begriff der *educational effectiveness* auch über den Rahmen des Unterrichts hinaus Wirkzusammenhänge zwischen den Ebenen Schule, Unterricht und Schüler thematisiert. Auch die deutschsprachige Unterrichtsforschung wurde von den Ansätzen Slavins nachhaltig beeinflusst. Neben einem breit rezipierten Überblicksartikel von Ditton (2000) sind diese Impulse am deutlichsten abzulesen am Konzept der drei Grunddimensionen der Unterrichtsqualität nach Klieme, Schümer und Knoll (2001) und dem Angebots-Nutzungs-Modell unterrichtlicher Wirkungen nach Helmke (2003).

Das in der Unterrichtsforschung in Variationen wiederholt aufgegriffene Konzept der drei Grunddimensionen der Unterrichtsqualität (vgl. Clausen, Reusser & Klieme, 2003; Klieme, Lipowsky, Rakoczy & Ratzka, 2006; Klieme, & Rakoczy, 2008) ist im Rahmen der TIMS-Studie entstanden und basiert auf videogestützten Beobachtungsanalysen zur Unterrichtsqualität im Mathematikunterricht (Klieme, Schümer & Knoll, 2001). Die daraus erarbeiteten Faktoren sollen „*als Grunddimension der Unterrichtsqualität angesehen werden*“ bzw. „*unverzichtbare Grundbedingungen schulisch organisierten Lernens*“ markieren (S. 52). Diese drei Grunddimensionen sind

- (1) Unterrichts- und Klassenführung (effektive Behandlung von Unterrichtsstörungen, Zeitverschwendung, Sprunghaftigkeit und Regelklarheit der Lehrkraft, Klarheit und Strukturiertheit des Unterrichts, Monitoring durch die Lehrkraft, *time on task*),
- (2) Schülerorientierung (Sozialorientierung, individuelle Bezugsnormorientierung, diagnostische Kompetenz, Interaktionstempo, Leistungsdruck) und
- (3) kognitive Aktivierung (genetisch-sokratisches Vorgehen, anspruchsvolles und repetitives Üben, Motivierungsfähigkeit der Lehrkraft).

Die Systematik dieser drei Faktoren ist auf der Basis empirischer Daten entstanden und erhebt nicht den Anspruch, eine Beschreibung von Zusammenhängen und Verknüpfungen von Unterrichtsfaktoren im Sinne eines vollständigen kausalen Modells zu leisten. Dennoch werden für die verschiedenen Dimensionen Zieleffekte auf Seiten der Schüler benannt: ein Mindestmaß an Aufmerksamkeit für den Bereich der Unterrichts- und Klassenführung, dass „*Schüler sich in die Gruppe integrieren und motiviert mitarbeiten*“ (S. 52) für den Bereich der Schülerorientierung und eine „*aktive kognitive Auseinandersetzung mit dem Lerngegenstand*“ für die kognitive Aktivierung (ebd.). An dieser Stelle werden die Punkte Aufmerksamkeit, Mitarbeit und Auseinandersetzung der Schüler mit dem Gegenstand implizit als Zielkategorien von Unterrichtsqualität formuliert – ein Verständnis, das in Einklang steht mit den Annahmen Carrolls und Slavins.

Ein weiterer, innerhalb der deutschsprachigen Unterrichtsforschung einflussreicher Ansatz, der sich in diese Argumentationslinie einordnen lässt, ist das Angebots-Nutzungs-Modell unterrichtlicher Wirkungen von Helmke (2003, 2012). Dieses Modell greift zum einen die bei Fend (1998) zu findende Angebots-Nutzungs-Logik schulischer Bildungsprozesse auf und versucht zum anderen, Einflussfaktoren auf verschiedenen Ebenen innerhalb und außerhalb des Unterrichts zu integrieren. Das Modell umfasst zum einen Faktoren, die außerhalb des Unterrichtsprozesses anzusetzen sind, wie etwa der persönliche Hintergrund der Schüler (Familie, Lernpotenzial) bzw. der Lehrperson (Persönlichkeitsmerkmale und Kompetenzen) und die für alle Beteiligten geltenden Rahmenbedingungen (Kontext). Zum anderen zielen die Bereiche der Prozessqualität des Unterrichts (zur Beschreibung des Angebots), der Lernaktivitäten der Schüler (als Aspekte der Nutzung dieses Angebots) und der Wirkungen (Zielkriterien und Effekte auf Seiten der Schüler/innen) auf Bestandteile innerhalb des Unterrichts ab. Das Modell

zeigt Zusammenhänge zwischen diesen Faktoren auf und verdeutlicht auf diese Weise das Wirkungsgeflecht, welches den möglichen Bildungserträgen vorausgeht, die durch Unterricht erzielt werden können.

Im Anschluss an Vorschläge zur Anreicherung des bis dato in der empirischen Unterrichts- und Lehr-Lern-Forschung weit verbreiteten Prozess-Produkt-Paradigmas um so genannte Mediationsmerkmale bzw. „Zwischenprozesse“ (Treiber, Weinert & Groeben, 1982, S. 566) berücksichtigt das Modell in besonderer Weise den Punkt der Wahrnehmung der dargebotenen Lerngelegenheiten im Unterricht durch die Schüler. Dadurch wird der Annahme Rechnung getragen, dass Unterricht keine unmittelbare und direkte Wirkung auf die Entwicklung der Schüler zugeschrieben werden könne, sondern *„nur auf dem Wege über individuelle Verarbeitungsprozesse.“* (Helmke, 2012, S. 80), womit Aspekte wie Rezeption und Interpretation, aber auch Kognition, Motivation und Emotionen der Schüler in der Unterrichtssituation gemeint sind. Helmke sieht diesen Bereich der individuellen Verarbeitungsprozesse im Unterricht am besten repräsentiert durch das Aufmerksamkeitsverhalten der Schüler, auch bezeichnet als *on-* bzw. *off task*-Status oder auch aktive Lernzeit (ebd.). In der Logik des Angebots-Nutzungs-Modells wird das Merkmal beeinflusst von den individuellen Eingangsvoraussetzungen der Schüler, wie z.B. ihrer Persönlichkeitsstruktur oder ihrem fachlichen Leistungspotenzial, den Kontextbedingungen, wie z.B. der Klassenzusammensetzung oder dem Klassen- bzw. Schulklima, und den Eigenschaften des Unterrichtsangebots, die sich durch die Prozessqualität des Unterrichts und die Qualität der Lehr-Lern-Materialien beschreiben lassen. Lässt man den Faktor der außerschulischen Lernaktivitäten außer acht, da an dieser Stelle Zusammenhänge auf Ebene des Unterrichtsprozesses im Mittelpunkt stehen, so fungiert in dem Modell die aktive Lernzeit – neben den Einflüssen der allgemeinen Kontextbedingungen – als Voraussetzung für mögliche Erträge des Unterrichts, sprich dessen Wirkungen. Dem Merkmal kommt damit neben einer Vermittlerrolle zugleich die Bedeutung einer zentralen Zielgröße von Unterrichtsprozessen zu.

Auf diesem Wege wird ein Aspekt deutlich, der auch über die Forschungslinien der Unterrichtsqualität und -effektivität hinaus in einer Reihe von Überlegungen zu Fragen der Wirksamkeit von Unterricht thematisiert wird, nämlich die Grenzen der Einflussnahme in der Unterrichtsdurchführung hinsichtlich der Initiierung von Bildungsprozessen auf Seiten der Schüler. Unterricht ist demnach als eine Summe von

Gelegenheiten zur Ermöglichung von Lernen zu verstehen, der sich nicht auf „*lineare Kausalität*“ (Meseth, Proske & Radtke, 2011, S. 223) verlassen bzw. seine Wirkungen nicht „*technologisch herbeiführen*“ (Gruschka, 2011, S. 132) kann. Oder um es aus Sicht der pädagogischen Psychologie zu sagen: „*Lehren hat nicht automatisch Lernen zur Folge*“ (Hasselhorn & Gold, 2006, S. 223). Schulunterricht kann indessen nur plausibel geführt werden „*mit der Unterstellung, durch dessen pädagogische Ausformung sei es möglich, auf allen genannten Ebenen die Schüler einer Klasse zu erreichen*“ (Gruschka, 2011, S. 132). Maßgeblich für Unterrichtswirksamkeit ist demzufolge, „*dass die Lernumgebungen der Schüler einer Gruppe, Klasse oder Schule so gestaltet werden, dass alle eine Gelegenheit zum Lernen wahrnehmen.*“ (Merkens, 2010, S. 116).

Diese Überlegungen knüpfen an die Position von Jackson an, der betont, Lernprozesse seien in gewisser Hinsicht eine Art „Nebenprodukt“ des Unterrichts (Jackson, 1966, S. 24), während das naheliegendste Zielkriterium des Unterrichtshandelns von Lehrkräften zunächst in der Herstellung von Aufmerksamkeit und Beteiligung der Schüler zu suchen sei. In der Zusammenschau der beschriebenen Ansätze zeigt sich, dass der Grundgedanke von Jacksons Position einen gewissen theoretischen Konsens zum Ausdruck bringt hinsichtlich der Rolle dieses Bereichs des Unterrichtsgeschehens als prozessimmanentem Indikator für die Effektivität von Unterricht. Das QAIT-Modell der Unterrichtseffektivität von Slavin (1994) vermag die in diesem Kontext bedeutsamen Zusammenhänge in besonders geeigneter Weise abzubilden, weshalb es im weiteren Verlauf als primärer theoretischer Bezugspunkt herangezogen wird.

2.3 Forschungsstand zum Schülerbeteiligungsverhalten im Unterricht

Im vorangegangenen Abschnitt wurde die theoretische Bedeutung der aktiven Hinwendung der Schüler zu den im Unterricht angebotenen Lerngelegenheiten herausgearbeitet. Anhand einer Reihe von Modellen und Konzeptionen zur Wirksamkeit schulischen Unterrichts konnte aufgezeigt werden, dass dieser Bereich übereinstimmend als Voraussetzung für unterrichtsbedingtes Lernen angesehen und ihm darüber die Bedeutung einer Zielgröße für den Unterrichtsprozess zugeschrieben werden kann. In diesem Abschnitt soll nun die empirische Befundlage zu diesem Bereich dargestellt und diskutiert werden.

Das Verhalten der Schüler im Unterricht im Hinblick auf ihre Lern- und Arbeitsaktivitäten ist seit über hundert Jahren als Gegenstand empirischer Forschung zu finden (Anderson, 1984a; Borg, 1980; Jackson, 1968/1990;). Anderson (1984a, 49 ff.) zeigt Diskussionslinien und Ergebnisse vom Ende des 19. Jahrhunderts bis in die 1930er Jahre hinein auf, als der Bereich von Psychologen und Pädagogen hauptsächlich unter dem Begriff der Schüleraufmerksamkeit bearbeitet wurde. Aufmerksamkeit stand zur damaligen Zeit stellvertretend für vieles, was später unter Begriffen wie etwa *time on task* oder Beteiligung verstanden wurde (Borg, 1980). Die Forschungsergebnisse aus dieser Zeit mündeten in den vorläufigen Konsens, dass die Aufmerksamkeit der Schüler eine Grundvoraussetzung für Lernen im Unterricht sei und dabei Verbindungen sowohl zu Interesse und Motivation der Schüler, als auch zur Unterrichtsführung im Sinne einer Steuerung und Aufrechterhaltung der Schüleraufmerksamkeit bestehen (Anderson, 1984a). Die Forschung in der Phase der 1920er bis 1950er Jahre wird systematisch dargestellt bei Jackson (1968/1990), ebenso fokussiert auf den damals gängigen Begriff der Aufmerksamkeit. Im Vordergrund standen damals die Fragen, inwiefern die Klassengröße, der Führungsstil der Lehrkraft oder bestimmte Eigenschaften der Schüler mit deren Aufmerksamkeit im Unterricht zusammenhängen (S. 88ff.). Verschiedene Forschungsbefunde der 1920er Jahre weisen darauf hin, dass die Lehrperson einen starken Anteil an der Aufmerksamkeitsrate hat, und zwar einen größeren als Faktoren wie etwa die Klassengröße. Darin deuten sich Forschungslinien an, die später ab den 1960er Jahren mit breiteren methodischen Möglichkeiten und der Anbindung an modernere Theorien des Unterrichts wieder aufgegriffen werden sollten (s.u.). Ein häufig bearbeiteter Untersuchungsgegenstand besteht dort im Zusammenhang zwischen Schüleraufmerksamkeit bzw. -beteiligung im Unterricht und Effekten auf den individuellen Lernzuwachs.

2.3.1 Schülerbeteiligungsverhalten und Lernerfolg

Durch die Fokussierung auf Fragen der Wirksamkeit von Schulunterricht stellen sich enge Bezüge zum Forschungsstrang der *educational effectiveness* bzw. *educational productivity research* her (vgl. Creemers & Kyriakides, 2008; Fraser, Walberg, Welch & Hattie, 1987; Scheerens & Bosker, 1997; Seidel & Shavelson, 2007; Wang, Haertel & Walberg, 1993; Wang & Walberg, 1991). Mit dem Fokus auf die „Produktivität“

schulischer Bildung nimmt diese Forschung ihren Ausgangspunkt in der Frage nach Faktoren, die sich im Hinblick auf den Lernerfolg von Schülern in bestimmten Fächern und Kompetenzbereichen als förderlich erwiesen haben (vgl. Fraser, Walberg, Welch & Hattie, 1987). Die zentralen Arbeiten aus diesem Bereich fassen in synthetisierender Weise empirische Befunde der internationalen Unterrichtsforschung aus mehreren Jahrzehnten zusammen, die etwas über Zusammenhänge zwischen Merkmalen der beteiligten Personen (und ihrer Hintergründe), Institutionen und Prozessen bzw. Praktiken einerseits und der Leistungsentwicklung der Schüler andererseits aussagen. Das Ergebnis sind systematische Forschungsüberblicke und Meta-Analysen, die zum einen als Hilfe zur Effektschätzung für Entscheidungsträger in Bildungszusammenhängen (Walberg & Lai, 1999) und zum anderen als Beiträge zur Theoriebildung in der Schul- und Unterrichtsforschung dienen sollen (Wang, Haertel & Walberg, 1993). Die Arbeiten aus diesem Bereich deuten durchgängig auf positive Zusammenhänge zwischen hohen Ausprägungen in Variablen des Schülerbeteiligungsverhaltens im Unterricht, wozu mehrere, teilweise deckungsgleiche Konstrukte wie z.B. *time on task* oder *engagement*⁶ gehören, und dem Lernerfolg der Schüler hin.

In einer Synthese aus 19 Überblicksarbeiten stellt Walberg (1986) fest, dass das „*engagement in learning*“ der Schüler im Unterricht als eines von fünf Konstrukten identifiziert werden kann, die substantiell mit der Schülerlernleistung zusammenhängen (S. 218). Diesen Befunden stellt Walberg eine Meta-Analyse multivariat angelegter Studien gegenüber. Dabei kommt er zu dem Ergebnis, dass Variablen der Unterrichtsquantität, darunter *time on task* als Variante des Schülerbeteiligungsverhaltens, durchgängig positive Einflüsse auf die Lernleistung aufweisen, betont in diesem Zuge aber gleichzeitig die insgesamt eher niedrigen Effekte (standardisierte Regressionsgewichte zwischen 0,05 und 0,25), was aber zum Teil auf das methodisch anspruchsvolle Auswertungsverfahren zurückzuführen sein könnte (S. 225). In einer späteren Überblicksarbeit verbinden Wang, Haertel und Walberg (1993) mehrere Analysestrategien der empirischen Synthese mit einem theoretischen Rahmen, der zuvor als Ergebnis eines systematischen Vergleichs von psychologischen Modellen schulisch bedingter Leistungsentwicklung gewonnen worden war (Haertel, Walberg &

⁶ Für eine differenzierende Darstellung der Konstrukte und Begriffe in diesem Bereich s. Kap. 2.4

Weinstein, 1983). Als Ergebnis (a) einer Inhaltsanalyse von 179 einschlägigen Handbuchartikeln, (b) einer schriftlichen Befragung von 61 amerikanischen Bildungsforscher(inne)n hinsichtlich ihrer Einschätzung zur Relevanz von 228 Variablen für das Schülerlernen und (c) der Meta-Analyse von Fraser, Walberg, Welch und Hattie (1987) wurden dreißig Einflussfaktoren für schulisches Lernen herausgearbeitet und hinsichtlich ihrer Effekte auf die Leistungsentwicklung in eine Rangfolge gebracht. Der Bereich *quantity of instruction*, der auch dort unter anderem die Variable *time on task* enthält, rangiert in den Ergebnissen mit einem T-Score⁷ von 53,7 auf Platz neun der dreißig Einflussfaktoren, wobei *time on task* in der Inhaltsanalyse die höchste Wirksamkeitseinstufung aller Zeitvariablen erhält (Wang & Walberg, 1991, S. 89).

In einer weiteren Überblicksarbeit zeigen Walberg und Lai (1999) für die Variable *time on task* in 89 % der untersuchten Studien einen positiven Zusammenhang zur Lernleistung, bei Zusammenhangsmaßen und Effektstärken im mittleren Bereich (S. 450). Ähnliche Ergebnisse sind in einer Meta-Analyse von Scheerens und Bosker (1997) zu finden, wo innerhalb des Bereichs der Unterrichtsbedingungen (*instructional conditions*) der *time on task* eine positive Wirkung bei mittleren bis schwachen Zusammenhangsmaßen (0,19) bescheinigt wird (S. 17). Auch in jüngeren Überblicksarbeiten setzt sich dieses Bild fort, etwa in einer Synthese aus 800 Meta-Analysen von Hattie (2009), die auf über 50.000 Einzelstudien basiert und für die *time on task* ein mittelstarkes Effektmaß von 0,38 (Cohens d) anzeigt. Gleichzeitig wird an anderer Stelle derselben Studie das *student engagement*, verstanden als Lernverhalten und -einstellung der Schüler, kombiniert mit ihrer Konzentrationsleistung, als Faktor mit noch stärkerer Effektschätzung von 0,48 ausgewiesen.

In der Summe zeichnen all diese Befunde ein einstimmiges, wenn auch nur grobes Bild bezüglich des Effektes von Variablen wie *time on task* und *student engagement in learning* im Unterricht auf die Lernentwicklung von Schülern. Es sind schwache bis mittelstarke, aber durchgängig positive Effekte und Zusammenhänge erkennbar, was dafür spricht, dass ein höherer Anteil aktiver Lernzeit und ein höheres Maß an Mitarbeit und Beteiligung der Schüler im Unterricht mit höheren Lernerträgen zusammenhängt.

⁷ Der T-Score ist eine aus den Rohdaten transformierte, einheitlich skalierte Effektgröße mit einem arithmetischen Mittel von 50 und einer Standardabweichung von 10.

Es ist allerdings davon auszugehen, dass durch die – in Meta-Analysen notwendige – Zusammenfassung von Variablen innerhalb übergeordneter Analysekatoren die spezifischen Effekte der einzelnen Konstrukte verdeckt werden. Insbesondere die Größe *time on task* ist davon insofern betroffen, dass sie meist mit anderen zeitbasierten Größen aus dem Bereich der Unterrichtsquantität kombiniert wird (vgl. Walberg, 1986; Wang, Haertel & Walberg, 1993). So belegt die Meta-Analyse von Seidel und Shavelson (2007) auf Grundlage von 112 Einzelstudien für den Bereich der Lernzeit schwache bis mittelstarke Effekte (Cohens d 0,08 - 0,32, S. 473), wobei diese Komponente neben der schülernahen Variable *time on task* gleichzeitig zwei weitere Größen umfasst: *opportunity to learn*, eine in erster Linie von den Lehrpersonen beeinflusste Maß unterrichtlicher Zeitnutzung, und *homework* als eine Variable der außerhalb des Unterrichts stattfindenden Lernleistung (S. 468). Daher fällt es an dieser Stelle schwer, die Ergebnisse gezielt auf Effekte der *time on task* hin zu interpretieren.

Ein Blick in Einzelstudien liefert mehr Aufschluss zum Verhältnis von Schülerbeteiligungsverhalten im Unterricht und Lernerfolg. Brophy und Evertson (1980) berichten im Rahmen des *Texas Teacher Effectiveness Project* von Beobachtungsstudien an 50 Grundschulen über zwei Jahre hinweg (zweites und drittes Schuljahr)⁸. Insgesamt wurden ca. 40 Stunden pro Klasse mittels direkter Beobachtung im Klassenraum und unter Verwendung von niedrig- und hoch-inferenten Beobachtungssystemen untersucht (S. 191). Gegenstand war unter anderem die Beobachtung von Schüleraktivitäten während Arbeitsaufgaben und die erkennbare Aufmerksamkeit. Im Vergleich mit nachfolgend durchgeführten Leistungstests ergeben sich für die Einschätzung der Arbeitsaktivitäten hohe Zusammenhänge mit der Lernleistung der Schüler, während die Beurteilungen der Aufmerksamkeit nur sehr gering mit den Leistungsindikatoren korrelieren (S. 77). Die Autoren ziehen daraus den Schluss, dass Aufmerksamkeit aufgrund der Gefahr möglichen Täuschungsverhaltens der Schüler als kein zuverlässiger Indikator für individuelle Verstehens- oder Verarbeitungsprozesse gewertet werden kann, sondern dazu eher offensichtliches Arbeitsverhalten wie beispielsweise Schülerantworten geeignet sind⁹.

⁸ Der genaue Umfang der Beobachtungsstichprobe ist aus den heute verfügbaren Dokumenten nicht zweifelsfrei erkennbar.

⁹ Dieser Aspekt wird in den Abschnitten 2.3.3 und 2.4.1 eigens diskutiert.

Im gleichen Zeitraum wurde innerhalb der *Beginning Teacher Evaluation Study* (Denham & Lieberman, 1980) das Lernverhalten der Schüler im Unterricht und dessen Wirkungen auf die Leistungsentwicklung untersucht. Die Stichprobe bestand aus 46 Klassen der Jahrgangsstufen zwei und fünf, deren Unterricht innerhalb eines Jahres zu 20 Zeitpunkten von externen Beobachtern besucht und beurteilt wurde. Die Beobachtungsdaten wurden in der Auswertung zu Pre- und Post-Testdaten zur Leistung der Schüler in den Bereichen Mathematik und Lesen in Beziehung gesetzt. Im Rahmen der Feldbeobachtungen wurde das Schülerverhalten im Unterricht anhand eines sequenziellen, rotierenden Beobachtungsplans erfasst. Mit diesem Verfahren wurden sechs zufällig ausgewählte Schüler pro Klasse (insgesamt 261) jeweils nacheinander in festen Zeitabständen von drei bis sechs Minuten beobachtet und ihre Aktivitäten im Hinblick auf die selbständige Beteiligung an den Aufgaben anhand der folgenden Kategorien und Unterkategorien beurteilt (Dishaw, 1978; Fisher, Berliner, Filby & Marliave, 1980; Romberg, 1980):

(1) beteiligt (*engaged*), und zwar durch (a) Schreiben (*written response*), (b) Sprechen (*oral response*), Denken bzw. Zuhören (*covert response*) oder mit Anweisungen der Lehrperson (*engaged directions*)

(2) nicht beteiligt (*not engaged*), und zwar durch (a) Arbeitsorganisation und Phasenübergänge (*interim activity*), (b) Warten auf Unterstützung (*waiting for help*) oder (c) Gespräche, Träumen, Störverhalten (*off task*)

Die Studie erfuhr in der Folgezeit, besonders im englischsprachigen Raum, eine breite Rezeption (Berliner, 1990; Borich, 2010; Karweit, 1989), was sowohl auf die großzügige Anlage der Untersuchung zurückzuführen ist als auch die zum damaligen Zeitpunkt innovativen Konzepte der *academic learning time* und der *engagement rate*¹⁰. Letztere wurde auf Grundlage der beobachtungsbasierten Einschätzungen als zeitbasiertes Maß des Schülerbeteiligungsverhaltens konstruiert und steht für den Anteil der von Schülern mit Lernaktivitäten genutzten Zeit an der zur Verfügung stehenden Unterrichtszeit (Fisher, Berliner, Filby & Marliave, 1980, S. 14). Ein zentrales Ergebnis der Studie ist der durchgängig positive Zusammenhang zwischen *engaged rate* und Lernerfolg, was von den Autoren mit dem Satz „*Students who pay attention more learn more*“ (S. 16)

¹⁰ Näheres dazu unter Kap. 3.

zusammengefasst wird. Abgesehen davon zeigen die Befunde große Varianzen der *engagement rate* sowohl zwischen als auch innerhalb der Klassen (Rosenshine, 1980), was als deutlicher Hinweis auf weiteren Forschungsbedarf zu den Bedingungen und Folgen dieser Unterschiede in der Klassensituation gewertet wurde.

Die grundsätzlich auch an sozialen und emotionalen Aspekten des Erlebens und Handelns von Schülern im Unterricht orientierte Forschung zum *student engagement* (vgl. Christenson, Reschly & Wylie, 2012) konnte in einigen Studien ähnliche Effekte nachweisen (z.B. Finn & Zimmer, 2012; Lam, Wong, Yang & Liu, 2012). Fredricks, Blumenfeld und Paris (2004) kommen in ihrer Überblicksarbeit zu dem Ergebnis, dass die Forschung der vorausgegangenen fünfzehn Jahre konsistent positive Zusammenhänge zwischen dem *student engagement*, darunter speziell dem Schülerbeteiligungsverhalten (*behavioral engagement*), und Lernleistung zeigt. Dabei seien aber deutliche Schwankungen der Effekte in Abhängigkeit von den Operationalisierungen und Erhebungsverfahren festzustellen (S. 70). Bei der Betrachtung einzelner Studien werden diese Unterschiede deutlich.

Finn, Pannozzo und Voelkl (1995) untersuchten beispielsweise den Zusammenhang zwischen dem Unterrichtsverhalten von über tausend Schülern der vierten Jahrgangsstufe und ihren Leistungen in einem fachlich breit angelegten, standardisierten Test. Das Schülerverhalten wurde über schriftliche Befragungen von 187 Lehrkräften zu den Unterrichtsaktivitäten von zufällig ausgewählten Schülern ihrer Klasse in den zurückliegenden zwei bis drei Monaten erfasst. Die Untersuchung fokussiert die Gruppe der *students at risk*, d.h. Schüler mit drohendem Abbruch der Schulkarriere, was sich – so die übergeordnete Hypothese – durch bestimmte Verhaltensweisen im Schulkontext vorhersagen lassen sollte. Die Auswertung konzentriert sich daher auf Kategorien negativen Unterrichtsverhaltens, das heißt häufig auftretender Unaufmerksamkeit oder der Neigung zur Unterrichtsstörung. In den Ergebnissen zeigen Schüler mit besonders problematischen Verhaltenseinstufungen¹¹ deutlich schlechtere Testleistungen als jene Schüler, deren Unterrichtsverhalten sich als angepasst oder bemüht beschreiben lässt (ebd., S. 426/427). Die Verhaltensklassifizierungen ergaben in der Stichprobe insgesamt 67 % unauffällige bzw.

¹¹ Das bedeutet, mit Werten von mehr als einer Standardabweichung oberhalb des Stichprobenmittels in den Items negativen Verhaltens (Pannozzo & Voelkl, 1995, S. 426).

angepasste, 14 % störende, 9 % unaufmerksame und 9 % unaufmerksame und gleichzeitig störende Schüler, sodass demzufolge in einer durchschnittlichen Klasse mit dreißig Schülern bei ungefähr acht Schülern von einem erhöhtem Risiko bezüglich ihrer Leistungsentwicklung auszugehen wäre. Die Ergebnisse sind jedoch mit Zurückhaltung zu interpretieren, da das Schülerverhalten durch die Lehrkräfte rückblickend für einen längeren Zeitraum eingeschätzt wurde und damit Verzerrungstendenzen in den Urteilen nicht auszuschließen sind.

Im Unterschied dazu wählte Stipek (2002) in einer Studie zum Einfluss von Formen der Schulmotivation auf die fachliche Leistung den methodischen Zugang der Feldbeobachtung im Unterricht, kombiniert mit schriftlichen Angaben der Lehrpersonen. Die Schüler in 173 Klassen der zweiten und dritten Jahrgangsstufe wurden jeweils über einen ganzen Schultag hinweg hinsichtlich ihres unterrichtlichen Arbeitsverhaltens (Aufmerksamkeit, Beharrlichkeit in Arbeitsphasen) und ihres damit verbunden Gemütszustands (Enthusiasmus) auf einer fünfstufigen Skala eingeschätzt (S. 320), zusätzlich wurde das Urteil der Lehrkräfte zum *engagement*-Status von ein bis zwei Schülern in der Klasse eingeholt. Gleichzeitig wurde von den Beobachtern die Qualität des Unterrichts anhand des Auftretens bestimmter didaktischer Strategien in der Unterrichtsführung beurteilt. Indikatoren schulischer Leistung waren schriftliche Tests zu drei Zeitpunkten über die beiden Schuljahre hinweg und zusätzlich eine Leistungseinstufung durch die Lehrkräfte. In den Ergebnissen zeigen sich deutlich positive Effekte des *student engagement* im Unterricht auf den Leistungszuwachs (Regressionskoeffizient $\beta = 0,28$ ($p < 0,001$) für die Jahrgangsstufe zwei; $\beta = 0,18$ ($p < 0,05$) für die Jahrgangsstufe drei; S. 326). Stipek und Kollegen ziehen daraus den Schluss: „*The importance of academic engagement is indicated by the finding that our measure of student engagement did, in fact, predict how much children learned*“ (ebd.). Einschränkend muss jedoch angemerkt werden, dass dieses Ergebnis alleine auf den Lehrereinschätzungen des Schülerbeteiligungsverhaltens basiert, für die Beobachtungsmaße werden keine Befunde berichtet. Eine weitere Erkenntnis der Studie ist überdies, dass die verschiedenen Unterrichtsqualitätsaspekte, wie beispielsweise das kognitive Niveau der Auseinandersetzung mit den fachlichen Inhalten, Formen eigenverantwortlichen Lernens oder die Tiefe der Ergebnisdiskussion in der Lerngruppe, signifikante Korrelationen zum *student engagement* in einer Größenordnung zwischen 0,2 und 0,4 aufweisen (S. 323). Somit wird in der Studie

einerseits die Bedeutung des Schülerbeteiligungsverhaltens für die Lernleistung belegt, andererseits werden Hinweise auf Möglichkeiten der Einflussnahme auf das Schülerbeteiligungsverhalten durch die Unterrichtsgestaltung geliefert.

Die Vielzahl der in ihrer Tendenz gleichartigen Befunde über die Jahrzehnte hinweg spricht insgesamt für einen positiven Zusammenhang zwischen dem Beteiligungsverhalten der Schüler im Unterricht und ihrer Lernleistung. Dies lässt sich wie gezeigt sowohl in Meta-Analysen und Forschungssynthesen aus der Zeit zwischen den Jahren 1980 und 2010 erkennen, als auch in den Arbeiten aus dem Bereich der *student engagement*-Forschung der letzten 25 Jahre. Letztgenannter Forschungszweig bietet zudem Hinweise auf Lerneffekte, die über den rein kognitiv-fachlichen Bereich hinausgehen und auch soziale und emotionale Lerndimensionen betreffen, wie beispielsweise soziales Bewusstsein, Selbstwahrnehmung oder Emotionsregulation (Reschly & Christenson 2012). Die Befundlage lässt zwar an keiner Stelle eine formale Überprüfung der im Abschnitt 2.2 dargestellten Modelle zur Unterrichtswirksamkeit erkennen, aber dennoch bestätigt die allgemeine Tendenz der Ergebnisse die theoretische Annahme eines positiven Zusammenhangs zwischen Schülerbeteiligungsverhalten im Unterricht und Lernerfolgen.

2.3.2 Bedingungen des Schülerbeteiligungsverhaltens

Im Anschluss an die vorangegangenen Betrachtungen stellt sich die Frage nach Bedingungen und Voraussetzungen des Schülerbeteiligungsverhaltens, um daraus Möglichkeiten zur Einflussnahme auf die Unterrichtswirksamkeit durch eine geeignete Unterrichtsgestaltung ableiten zu können. Einschlägige Unterrichtsmodelle machen in dieser Hinsicht unterschiedliche Erklärungsangebote (s. Kap. 2.2). Zusammengefasst lassen diese sich unterscheiden nach Hintergrundvariablen der beteiligten Personen, wie etwa den individuellen (fachlichen oder allgemeinen kognitiven) Fähigkeiten und Persönlichkeitsmerkmalen des einzelnen Schülers, und Prozessvariablen des Unterrichts, wie z.B. der Angemessenheit der Inhalte im Hinblick auf die Ausgangslagen der Schüler, den Motivierungs- und Aktivierungsgrad des Unterrichts oder die Effektivität der Klassenführung (Helmke, 2003; Klieme, Schümer & Knoll, 2001; Slavin, 1994). Die genannten Ansätze deuten an, dass es sich zwar um ein komplexes Wirkgeflecht von Einflussgrößen handelt, doch wie zu zeigen ist, haben nur wenige

empirische Arbeiten mehrere der in Frage kommenden Größen zugleich untersucht und zueinander in Beziehung gesetzt.

Helmke, Schneider und Weinert (1986) untersuchten im Rahmen eines Gesamtmodells zugleich persönliche Hintergrundvariablen der Schüler und Prozessvariablen des Unterrichts in ihrer Wirkung auf die Lernergebnisse. Als Erhebungsverfahren kamen sowohl schriftliche Tests und Befragungen der Schüler als auch Feldbeobachtungen durch das Forscherteam zum Einsatz. Die Stichprobe setzte sich zusammen aus 39 Hauptschulklassen der Jahrgangsstufe fünf. Untersucht wurde Unterricht im Fach Mathematik über zwei Schuljahre hinweg, also bis zum Ende der Jahrgangsstufe sechs. Die Hintergrundvariablen der Schüler, aufgeteilt in eine kognitive und affektive Komponente, wurden repräsentiert durch Eingangstests (Pre-Tests) in allgemeinen kognitiven und mathematischen Fähigkeiten und Befragungen zur Lernmotivation im Fach Mathematik. Beide Komponenten wurden zum zweiten Messzeitpunkt (Post-Tests) auch als Indikatoren für Lerneffekte verwendet. Die Prozessvariablen des Unterrichts fokussieren die beiden Bereiche Effizienz der Klassenführung, operationalisiert durch Unterrichtsstörungen und effektive Nutzung der Unterrichtszeit, und Unterrichtsqualität, zusammengesetzt aus den Teilaspekten verstehensorientierte Hilfestellungen der Lehrkraft und Klarheit der Instruktionen (ebd., S. 7). Die Komponenten der Unterrichtsqualität wurden beide sowohl beobachtungsbasiert erfasst als auch durch schriftliche Befragungen der Schüler. Das Schülerbeteiligungsverhalten, bezeichnet als *student engagement*, wurde erhoben in einer Kombination aus schriftlichen Selbstauskünften der Schüler zum Grad der eigenen Unaufmerksamkeit einerseits und Einschätzungen des *on task*-Verhaltens durch die Feldbeobachter andererseits, wobei jeweils acht zufällig ausgewählte Schüler pro Klasse sechs mal in der Stunde hinsichtlich verschiedener Verhaltensfacetten (z.B. Meldeverhalten, Blickrichtung, abgelenkt sein etc.) beurteilt wurden. Die Ergebnisse der Studie sind in einem nach der PLS-Methode (*Partial Least Squares*, ebd., S. 2) berechneten kausalen Pfadmodell dargestellt. Gemessen an der Ausprägung der Pfadkoeffizienten ist der augenscheinlichste Befund der starke Zusammenhang zwischen den Pre- und Post-Tests in den kognitiven und affektiven Eingangsmerkmalen der Schüler (kognitiv: 0,62; affektiv: 0,40; S. 14). Dieses Teilergebnis ist jedoch nur wenig überraschend, da es zunächst nur eine gewisse Stabilität der individuellen Werteneiveaus bezüglich der kognitiven und affektiven Eigenschaften der Schüler im

Klassenvergleich über den Untersuchungszeitraum hinweg andeutet. Der Einfluss des *student engagement* im Unterricht auf die individuellen Lernerfolge zeigt für den Bereich der kognitiven Leistung im Gegensatz zum bisherigen allgemeinen Stand der Empirie (s. Kap. 2.3.1) keinerlei Zusammenhang (kein Kennwert berichtet, s. Helmke, Schneider & Weinert, 1986, S. 16). Die Autoren bieten diesbezüglich verschiedene Erklärungen an. Zum einen sei dies möglicherweise auf die Struktur der Inhalte und Lernmaterialien zurückzuführen, die im untersuchten Unterrichtsfach in dieser Klassenstufe üblicherweise größtenteils auf Wissenskonsolidierung und Übung abzielen. Dementsprechend würde eine intensive Auseinandersetzung der Schüler mit den Unterrichtsinhalten, indiziert durch eine hohe *time on task*, nur begrenzten Lernzuwachs erzeugen, sodass auch kein entsprechender Effekt erkennbar werden kann. Eine weitere Ursache für einen fehlenden Effekt könnte in der Analysemethode gesucht werden, da klassenweise aggregierte Werte des *student engagement* möglicherweise differenzielle Effekte für einzelne Schüler oder verschiedene Schülergruppen maskieren, die aufgrund ihrer individuellen Ausgangslagen unterschiedliche Konstellationen an Unterrichtsmerkmalen für einen produktiven Lernprozess benötigen. Beide Argumente spielen auf ein Problem an, das von den Autoren als „*unnecessary or even dysfunctional time on task*“ (ebd.) umschrieben wird, sprich einer prinzipiell möglichen Auseinandersetzung der Schüler mit Lerngegenständen, die für sie keinen potenziellen Lerngewinn in sich bergen. Neben diesem Ergebnis, welches den Zusammenhang zwischen *student engagement* und Lernzuwachs betrifft, werden in der Studie auch Einflüsse für das *student engagement* aufgezeigt: Das Ausmaß der allgemeinen Lernmotivation der Schüler und ihr kognitives Ausgangsniveau weisen geringe, aber nicht zu vernachlässigende Zusammenhänge zum Schülerbeteiligungsverhalten im Unterricht auf (Pfadkoeffizienten 0,20 bzw. -0,24). Der negative Zusammenhang zwischen kognitiven Eingangsvoraussetzungen der Schüler und ihrem Beteiligungsverhalten im Unterricht wird dabei als Hinweis auf eine mögliche Unterforderung von Teilen der Schülerschaft im Unterricht gewertet. Deutlich stärker aber als die persönlichen Hintergrundvariablen wirken Merkmale des Unterrichtsprozesses auf das Schülerbeteiligungsverhalten. Besonders die Klassenführung, welche sowohl auf Beobachtungen externer Forscher als auch Schülerauskünften basiert, erweist sich als aussagekräftig (0,46 bzw. 0,34). Die Aspekte der Unterrichtsqualität ergeben hingegen ein uneinheitliches Bild: Die Schülerurteile in diesem Bereich ergeben einen moderat positiven (0,28), die beobachtungsbasierten

Größen hingegen einen negativen mittleren Zusammenhang (-0,47) zum *student engagement*. Eine Interpretation dieses Teilergebnisses bieten die Autoren allerdings nicht an.

Der Einfluss effizienter Klassenführung auf das Schülerbeteiligungsverhalten im Unterricht ist in empirischer und theoretischer Hinsicht vielfach bearbeitet worden (Anderson, 1984b; Brophy, 2006; Doyle, 1986; Ophardt & Thiel, 2008). Die durch Forschungsergebnisse weitgehend bestätigten Kernannahmen bestehen diesbezüglich darin, dass ein geeignetes Arrangement der materiellen Lernumgebung, der sozialen Ordnung in der Schülergruppe und der Unterrichtsabläufe in einen insgesamt optimalen Lernkontext und damit auch höhere Lern- und Beteiligungsaktivitäten der Schüler resultiert. Jüngere Interventionsstudien, die auf der quasi-experimentellen Umsetzung von Trainingskonzepten mit dem Ziel der Verbesserung von Handlungskompetenzen von Lehrkräften im Bereich der Klassenführung aufbauen, belegen diese Zusammenhänge erneut. In der Studie von De Sa Maini (2011) wurde gezeigt, dass die Teilnahme von Lehrkräften an einem vierstündigen Workshop zu Aspekten des *classroom behaviour management* zu einer signifikanten Reduktion der von den Lehrern wahrgenommenen Unaufmerksamkeit von Schülern und von durch Beobachter berichtetem Stör- bzw. aufgabenfernem Verhalten der Schüler führt. Auch Piwowar, Thiel und Ophardt (2013) kommen bei der Evaluation eines selbst entwickelten, aus fünf Lerneinheiten mit einem Gesamtumfang von 23 Stunden bestehenden Trainingsprogramm für Lehrkräfte zu „Kompetenzen des Klassenmanagements“ (KODEK) zu dem Ergebnis, dass neben einer Zunahme des deklarativen Wissens im Themenkomplex des Klassenmanagements und der teilweisen Steigerung von Handlungskompetenzen der Lehrpersonen in diesem Bereich ein großer Effekt auf Seiten der Schüler festzustellen ist, und zwar in Form einer Steigerung des *student engagements* im Unterricht (Prüfung von Gruppenunterschieden zwischen Kontroll- und Interventionsgruppe via ANOVA mit Messwiederholung: $F_{(1,30)} = 15,55$; $p < 0,001$; $\eta^2 = 0,34$; ebd., S. 7).

Die Untersuchung von Effekten effizienter Klassenführung im Rahmen eines vergleichenden Modells mit mehreren Unterrichtsfaktoren wurde in einer Studie von Helmke und Renkl (1993) geleistet, wo neben der Klassenführung die Zusammensetzung der Klasse, die Adaptivität des Unterrichts und die Ausprägung des

Sozialklimas in ihrem Einfluss auf das Aufmerksamkeitsverhalten der Schüler verglichen wurden. Die Stichprobe umfasst knapp über tausend Schüler in 52 deutschen Grundschulklassen.

Die Datenerhebung geschah zum einen Teil mittels Feldbeobachtungen im Unterricht. Die Variablen der Unterrichtsbedingungen (Klassenführung, Adaptivität, Sozialklima) wurden mittels hoch-inferenter¹² Einschätzungsverfahren erzeugt. Das so genannte Aufmerksamkeitsverhalten der Schüler wurde mit einem eigenen niedrig-inferenten Beobachtungsverfahren, dem Münchener Aufmerksamkeitsinventar (MAI, Helmke, 1988; Helmke & Renkl, 1992) erfasst. Im Rahmen des MAI, welches später u.a. im Rahmen der SCHOLASTIK-Studie (Weinert & Helmke, 1997) eingesetzt wurde, steht Aufmerksamkeit synonym für die als Bezugspunkt herangezogenen Konstrukte *time on task* bzw. *active learning time* (Helmke, 1988, S. 3). Die Datenerfassung mit diesem Inventar wurde auf Grundlage eines Zeitstichprobenverfahrens in Fünf-Sekunden-Intervallen und mit rotierender Personenstichprobe vorgenommen (S. 12). Für die Kategorisierung des Schülerverhaltens sind zwei Schritte vorgesehen: Erstens wird das Schülerverhalten einer der beiden Kategorienstufen *on* oder *off task* zugeordnet: „*Verhaltensweisen von Schülern gelten dann als on-task, wenn sie sich als Nutzung der jeweiligen unterrichtlichen Lerngelegenheit interpretieren lassen; andernfalls wird off-task angenommen*“ (S. 13). Zweitens wird durch die Unterscheidung von Unterstufen des *on/off task*-Verhaltens die Qualität des Aufmerksamkeitsverhaltens erfasst. Die Grundidee ist dabei eine Differenzierung in passive und aktive Formen des Verhaltens. Im Falle von *off task* bezieht sich dies auf eine passive Nicht-Nutzung von Lerngelegenheiten einerseits und das aktive Stören des Unterrichts andererseits, im Falle von *on task* umfassen die Unterkategorien einerseits die passiv-unauffällige Teilnahme an der Aufgabe und andererseits die besondere Beteiligung „*über die Mindestanforderungen hinaus*“ (S. 14). Gegenüber der ersten, eher „*lehrerzentrierten*“ Bewertungsebene (S. 13) ist das Ziel der zweiten Ebene eine stärkere Berücksichtigung der Schülerperspektive, der „*Registrierung mutmaßlicher Lernaktivitäten*“ (S. 35). In der Auswertung der Studie wurde ein mithilfe des MAI gewonnener „*summativer Kennwert für die Aufmerksamkeit*“ (Helmke & Renkl, 1993, S. 191) verwendet. Neben den

¹² „Hoch-“ bzw. niedrig-inferent“ beschreibt den Grad der erforderlichen Schlussfolgerungen bzw. Interpretationen bei der beobachtungsbasierten Beurteilung eines Merkmals (vgl. Lotz, Gabriel & Lipowsky, 2013).

Beobachtungsdaten flossen auch Merkmale der Kontextbedingungen aus alternativen Quellen ein: die Klassengröße und das Geschlecht aus administrativen Daten der Schule, die Muttersprache aus einer schriftlichen Schülerbefragung und die kognitiven Eigenschaften der Schüler aus einem sprachfreien Intelligenztest (S. 192).

Die Ergebnisse der Studie zeigen zunächst eine hohe Varianz hinsichtlich des Aufmerksamkeitsniveaus zwischen den Klassen (min. 57 %, max. 94 %, Stichprobenmittelwert bei ca. 80 %). In der Analyse von Zusammenhängen zwischen den Unterrichtsvariablen und dem Aufmerksamkeitsverhalten der Schüler stellt sich die Klassenführung als dominanter Faktor heraus. In zwei der drei untersuchten Schuljahre, den Klassen drei und vier, zeigt sich dies in Form des jeweils höchsten Korrelationskoeffizienten unter allen einbezogenen Variablen. Die Korrelationen bewegen sich dabei in einem Wertebereich zwischen 0,42 und 0,67 und erweisen sich als überwiegend hoch signifikant. Nur die Adaptivität des Unterrichts, das heißt die Angemessenheit der Unterrichtsinhalte in Bezug zu den individuellen Ausgangslagen der Schüler, zeigt im zweiten Schuljahr einen noch stärkeren Zusammenhang zum Aufmerksamkeitsverhalten der Schüler (0,54). Dieser Effekt fällt aber in den beiden folgenden Schuljahren gravierend ab (0,40 im dritten und 0,22 im vierten Schuljahr). Für diesen abnehmenden Effekt der Unterrichtsadaptivität auf das Aufmerksamkeitsverhalten über die Grundschulzeit hinweg finden die Autoren keine Erklärung, äußern jedoch Vermutungen zu möglicherweise verantwortlichen Schwierigkeiten der Validität und Operationalisierung des Konstrukts der Adaptivität (S. 201). Neben den Ergebnissen zu Klassenführung und Adaptivität erweist sich das Sozialklima als wenig bedeutsam für das Unterrichtsverhalten der Schüler (Korrelationskoeffizienten zwischen 0,05 bis 0,29), ebenso ergeben die Kontextbedingungen lediglich schwache Zusammenhänge (zwischen 0,03 und 0,23). Zusammengefasst deutet die Untersuchung darauf hin, dass Faktoren der Klassenzusammensetzung vergleichsweise wenig, Aspekte der Unterrichtsführung hingegen mehr Einfluss auf das Schülerbeteiligungsverhalten zu haben scheinen. Als besonders bedeutsam stellen sich dabei insbesondere die Art der Klassenführung durch die Lehrkraft und in Teilen die Adaptivität der Unterrichtsgestaltung dar.

Eine vergleichende Prüfung der Einflüsse von Merkmalen der Unterrichtsführung einerseits und der Personenhintergründe der Schüler andererseits auf das

Beteiligungsverhalten im Unterricht wurde auch in einer Studie von Yair (2000) mit 856 Schülern an 33 US-amerikanischen Elementar- und Sekundarschulen vorgenommen. Leitender Gedanke war die Annahme, dass Charakteristika des Unterrichts als vermittelnde Größe zwischen den individuellen sozialen Einflussfaktoren der Schüler und ihrem Beteiligungsverhalten im Unterricht stehen. Bemerkenswert ist in dem Zusammenhang die Verwendung der *experience sampling method (ESM)* - einer Erhebungstechnik, die mit Hilfe von speziell programmierten digitalen Armbanduhren arbeitet und die Erfassung situationsspezifischer Angaben aus Sicht der Schüler ermöglicht (S. 252). Dazu trugen die an der Studie beteiligten Schüler eine Woche lang diese Geräte mit sich und wurden acht mal am Tag (zwischen 7:30 Uhr und 22:30 Uhr) durch einen Signalton aufgefordert, Fragen zu ihrer momentanen Tätigkeit, ihren aktuellen Gedanken und Stimmungen per Eingabe in das Gerät zu beantworten. Auf diese Weise wurden insgesamt über 28.000 solcher personenspezifischen Situationsbeschreibungen gesammelt, was einem Rücklauf von 71 % entspricht. Die Angaben wurden schließlich auf Tageszeiten des Schulunterrichts reduziert und die in nicht standardisierter Form vorliegenden Schülerantworten Aktivitätscodes zugeordnet, die nachträglich induktiv gebildet wurden (S. 253). Auf die Ausprägung des *student engagements* wurde daraufhin geschlossen, indem eine Kombination der Angaben zur momentanen Umgebung der Person (*place*) und ihren Gedanken (*thought*) gebildet wurden: Eine Angabe wie beispielsweise „*place: mathematics class*“ und „*thought: my girlfriend*“ wurde dabei als Anzeichen von Ablenkung (*alienation*) gewertet (ebd.).

Die Ergebnisse belegen zunächst ein verhältnismäßig niedriges Beteiligungsniveau der Schüler/innen von durchschnittlich 54 % der Unterrichtszeit (S. 254). Regressionsanalysen zeigen daran anschließend, dass der ethnische Hintergrund der Schüler, aber auch das Alter ihr Beteiligungsverhalten bzw. -erleben zu beeinflussen scheint: *hispanic* und *african americans* sind signifikant weniger dem Unterricht zugewandt als weiße und asiatisch-stämmige Schüler. Auch Schüler der höheren Jahrgangsstufen weisen niedrigere Grade des *engagement* auf, während das Geschlecht keinerlei Effekte zeigt (S. 256). Bei Betrachtung der Unterrichtsmerkmale (u.a. Unterrichtsfach, Arbeits- und Sozialformen, Lebensweltbezug der Inhalte, Niveau der fachlichen Herausforderung) als Bedingungsfaktoren ist ein höheres *student engagement* in Phasen der Gruppenarbeit, experimentellen Unterrichts und Diskussionsphasen festzustellen, ebenso auch bei hohem inhaltlichen Lebensweltbezug

des Unterrichts und einem hohem Niveau fachlicher Herausforderung für die Schüler. Persönliche Hintergrundvariablen der Schüler wie Zeugnisnoten, die allgemeine emotionale Verfassung oder das Freizeitverhalten sind in ihrem Einfluss auf das *student engagement* hingegen zwar mäßig ausgeprägt, dennoch sind in diesem Bereich die Variablen mit emotionalem Bezug (Synonyme von Kontrollüberzeugungen, intrinsischer Motivation und Selbstwirksamkeitserwartungen, S. 254) über alle Regressionsmodelle hinweg durchgängig signifikant. Im direkten Vergleich der untersuchten Variablen in ihrem Einfluss auf das *student engagement* ist festzustellen, dass die Unterrichtsmerkmale in den Kennwerten zwar den größeren Effekt zeigen ($\beta = 0,55$ bis $1,02$) als die Hintergrundvariablen der Schüler ($0,10$ bis $0,44$), die letztgenannten - darunter besonders der ethnische Hintergrund und der allgemeine emotionale Status - durch die Unterrichtsführung allerdings nicht vollständig kompensierbar zu sein scheinen. Dies wird durch Teilergebnisse unterstützt, die belegen, dass nicht alle Schüler innerhalb einer Klasse (in diesem Falle festgemacht am Beispiel des ethnischen Hintergrunds) in gleichem Maße auf Variationen der Unterrichtsmerkmale reagieren (S. 261). Laut Autoren findet damit die Metapher eines „Tauziehens“, sprich eines Widerstreits der Kräfte zwischen Charakteristika des sozialen Hintergrunds der Schüler und Eigenschaften der Unterrichtsgestaltung um den größeren Einfluss auf das Schülerbeteiligungsverhalten, Bestätigung (S. 260).

Bezogen auf den Einfluss der individuellen Hintergründe und Dispositionen von Schülern auf ihr Aufmerksamkeits- oder Beteiligungsverhalten im Unterricht lassen sich darüber hinaus wiederholt Hinweise auf differenzielle Effekte finden, die auf Unterschiede in den Unterrichtsbedingungen zurückführbar sind: Tendenziell sind es die fachlich oder kognitiv schwächeren Schüler in einer Klasse, die eine höhere Neigung zu Unaufmerksamkeit oder *off task*-Verhalten zeigen und zugleich sensibler auf Variationen der Unterrichtsbedingungen reagieren (Beckerman & Good, 1978; Blatchford, Bassett & Brown, 2011; Treiber, Weinert & Groeben, 1982). Wie in einigen der dargestellten Studien ersichtlich, gibt es offenbar Gruppierungskriterien auf Ebene der Schülerpersonen, die zur Erklärung von Verhaltenstendenzen unter bestimmten Unterrichtsbedingungen betragen, wie beispielsweise der fachliche Leistungsstand, allgemeine kognitive Fähigkeiten oder auch der ethnische Hintergrund (s. zuletzt genannte Studien, zuzüglich Helmke, 1997; Yair, 2000). Die Abwägung von Effektgewichten zwischen den beiden Bereichen der Schülerpersonen und der

Unterrichtsmerkmale fällt insgesamt nur in Einzelfällen deutlich zugunsten der Schülerhintergründe (z.B. Pauli & Lipowsky, 2007) und meist eher zugunsten der Unterrichtsbedingungen aus (z.B. Buff, Reusser & Pauli, 2010; Sacher, 1995). Das Ausmaß, mit dem sich Schüler im Unterricht den Lerngelegenheiten zuwenden, lässt sich also weder alleine, noch zum überwiegenden Teil durch die individuellen Personeneigenschaften der Schüler erklären. Vielmehr muss davon ausgegangen werden, dass die Charakteristika des jeweiligen Unterrichts letztendlich entscheidend dafür sind, ob die Schüler ihre persönlichen Tendenzen in Form von Aufmerksamkeit oder aktiver Beteiligung zur Entfaltung bringen¹³. An diesem Punkt scheint die Befundlage damit die grundlegende Plausibilität der Annahmen zum Zusammenspiel von *student inputs* und *alterable variables of instruction* aus Slavins QAIT-Modell (Slavin, 1994, s. Kap. 2.2) zu untermauern.

Von besonderem Interesse ist in der Folge, inwiefern unterschiedliche Unterrichtskonfigurationen dazu beitragen können, verschiedenartige Schüler einer Klasse in ihrem Beteiligungsverhalten positiv zu beeinflussen. Die Vielfalt bzw. Heterogenität der Schülerschaft hinsichtlich ihrer individuellen Voraussetzungen schlägt sich beispielsweise in Form von fachbezogenen Leistungsunterschieden nieder. Einige Untersuchungen haben sich ausgehend davon mit der Frage auseinandergesetzt, inwiefern sich das Verhältnis von der Schwierigkeit der Lerninhalte und -materialien zum Fähigkeitsniveau der Schüler auf deren Lernverhalten im Unterricht auswirkt.

Gickling und Armstrong (1978) stellen in einer quasi-experimentellen Untersuchung des aufgabenbezogenen Schülerverhaltens unter drei unterschiedlichen Bedingungen des fachlichen Schwierigkeitsniveaus fest, dass bei der Konfrontation mit einem Schwierigkeitsgrad, der sich durch einen Anteil von etwa 70 - 85 % bekannter Elemente in den Aufgaben beschreiben lässt, nicht nur die höchste Rate an *on task*-Verhalten zu beobachten ist, sondern auch die meisten Aufgaben verstanden und vollständig bearbeitet werden (S. 564). Sowohl eine Über- als auch Unterforderung der Schüler führte hingegen zu einem Abfallen des aufgabenorientierten Schülerverhaltens. Der Befund hoher Arbeits- und Beteiligungsaktivitäten von Schülern unter der Bedingung

¹³ Helmke und Renkl (1993) sprechen diesbezüglich von der „Plausibilitätsannahme einer kausalen Prädominanz des Unterrichts“ (S. 198).

eines optimalen Herausforderungsniveaus konnte in ähnlicher Weise auch in einer Untersuchung von Burns und Dean (2005) gezeigt werden. Für die Abschätzung des Herausforderungsgrades der Aufgaben wird hier die Methode des so genannten *curriculum-based assessment* angewandt, die eine Kalibrierung zwischen der fachbezogenen Lernausgangslage der einzelnen Schüler, den Lernaufgaben und entsprechend daran anschließenden Leistungstests erlaubt (S. 273). Derartige Untersuchungen finden somit notwendigerweise unter relativ stark regulierten Unterrichtsbedingungen (s. Gickling & Armstrong, 1978) oder gar in gänzlich anderen Lernumgebungen statt (s. Burns & Dean, 2005), die sich nicht mit herkömmlichem Alltagsunterricht vergleichen lassen.

Doch auch in methodisch an natürlichem Unterricht ausgerichteten Studien konnten ähnliche Zusammenhänge bestätigt werden: Lam, Wong, Yang & Liu (2012) führten schriftliche Befragungen bei 822 Schülern und ihren Lehrkräften in drei chinesischen Großstädten durch. Bei den Schülern wurden neben kognitiven, affektiven und verhaltensbezogenen Facetten des *student engagements* auch Einschätzungen zu Merkmalen des Unterrichts und persönliche sozial-emotionale Hintergrundbedingungen erfragt, während die Lehrkräfte Auskunft gaben über schulische Leistungen ihrer Schüler. Dabei ließ sich zeigen, dass verschiedene Aspekte eines von den Schülern wahrgenommenen „motivierenden Unterrichtskontextes“ (optimal herausfordernde Aufgaben, Wecken von Neugier, ausgeprägter Lebensweltbezug der Unterrichtsinhalte u.a.) allesamt hoch signifikant mit dem selbst berichteten Schülerbeteiligungsverhalten zusammenhängen (Korrelationskoeffizienten zwischen 0,24 und 0,38; S. 412). Im Untersuchungskontext sticht jedoch in besonderem Maße der Faktor des Lebensweltbezugs (*real-life significance*) hervor, der sowohl bezüglich des korrelativen Zusammenhangs zum Schülerbeteiligungsverhalten ($r = 0,38$) als auch des prognostischen Zusammenhangs ($\beta = 0,33$) zur globalen *student engagement*-Kategorie die stärksten Effekte aufweist (ebd.). Damit schließt diese Arbeit an Ergebnisse aus einer Studie von Marks (2000) an. Auf Basis von standardisierten schriftlichen Befragungen von 3.669 Schülern in 143 Klassen der Jahrgangsstufen fünf, acht und zehn an 24 unterschiedlichen Schulen wurden kontrastierend verschiedene Einflussfaktoren (Hintergrundmerkmale der Schüler, strukturelle und soziale Eigenschaften der Schulumgebung, Charakteristika des Unterrichts) in ihrer Wirkung auf das *student engagement* untersucht. Letzteres wurde erhoben durch Selbstauskünfte der Schüler zu

ihrer Leistungsanstrengung, ihrem allgemeinen Aufmerksamkeitsniveau und ihrem Arbeitsverhalten im Unterricht. Die Ergebnisse zeigen eine herausragende Bedeutung der individuell wahrgenommenen Authentizität des Unterrichts (*authentic instructional work*), die sich aus den Aspekten (a) Interessantheit von Fragen durch die Lehrkraft, (b) Vollzug eines tiefen Gegenstandsverständnisses, (c) Herstellung eines lebensweltlichen Anwendungs- und Problembezugs für die gegebenen Lerninhalte und (d) Spielräume zur Diskussion eigener Ideen im Unterricht zusammensetzt, für das *student engagement* (S. 169). Im Rahmen der multivariaten Auswertung wird zudem deutlich, dass dieser Effekt im Vergleich mit den Einflüssen durch die persönlichen Hintergrundvariablen (sozio-ökonomischer Hintergrund und Leistungsniveau in Mathematik, Lesen, Schreiben) deutlich überwiegt.

Neben dieser Studie nennen Fredricks, Blumenfeld und Paris (2004) in ihrem Forschungsüberblick einige weitere Studien, in denen sowohl das Herausforderungsniveau bzw. der Schwierigkeitsgrad von Aufgaben als auch das Ausmaß des darin angesprochenen Lebensweltbezugs positiv mit dem Schülerbeteiligungsverhalten im Unterricht assoziiert ist (S. 79). Die Autoren bekräftigen in diesem Zuge zwar die Auffassung, dass dem so genannten *behavioral engagement* der Schüler (als spezifische Ausprägung des breiter angelegten *school engagement*-Konzepts, S. 60ff.) die Bedeutung einer veränderbaren Größe im Wirkgeflecht schulischen Lernens und damit eines aussichtsreichen Ansatzpunkts zur Einflussnahme für Lehrkräfte und Schuladministration zukommt und das *engagement* darüber hinaus als zentrale Mediatorgröße zwischen den Eigenschaften von Individuum und Kontext auf der einen und schülerbezogenen Leistungsparametern auf der anderen Seite gelten kann. Zugleich heben die Autoren aber hervor, dass noch Forschungsbedarf darin bestehe, herauszuarbeiten, in welcher Weise die Personeneigenschaften der Schüler und die Aufgabeneigenschaften in ihrem Einfluss auf das *student engagement* interagieren.

2.3.3 Forschungsmethodische Aspekte

In der Forschung zum Schülerbeteiligungsverhalten sind einige Schwierigkeiten zu erkennen, die die Methodologie in Bezug auf den Gegenstand betreffen. Es wurde ersichtlich, dass der Gegenstandsbereich mehrere unterschiedliche Facetten des auf die

Lerngelegenheiten des Unterrichts gerichteten Verhaltens und Erlebens der Schüler umfasst. Wie der Forschungsüberblick zeigt, arbeiten einige Studien mit Befragungen der Schüler oder der Lehrkräfte, um Daten zu den inneren Zuständen der Schüler zu erheben, aber auch, um Einschätzungen zu deren Unterrichtsverhalten einzuholen (Helmke, 1997; Finn, Pannozzo & Voelkl, 1995; Stipek, 2002). Dabei deuten die Befunde mancher Studien auf mögliche Urteilsverzerrungen in den Befragungsdaten hin, wie etwa jene von Stipek (2002), die unterschiedliche Zusammenhänge zwischen *student engagement* und Schulleistung nachweisen konnte, je nachdem, ob Daten auf Basis von Einschätzungen der Lehrkräfte über ihre Schüler oder Beobachtungsdaten der Forschergruppe zur Auswertung herangezogen wurden (S. 326), oder auch die Untersuchung von Helmke, Schneider und Weinert (1986).

Dort wurde deutlich, wie sehr die Ergebnisse zum Zusammenhang zwischen dem *student engagement* und anderen Merkmalsbereichen des Unterrichts differieren können, wenn diese auf unterschiedliche Weise erhoben wurden: Aspekte der Unterrichtsqualität, eingeschätzt von externen Beobachtern, ergaben in der genannten Studie eine Korrelation zum *student engagement* in Höhe von -0,47, basierend auf Urteilen der Schüler hingegen einen Wert von 0,28 (S. 14). Das Phänomen der teilweise geringen Übereinstimmung zwischen Urteilen aus unterschiedlichen Personenperspektiven in Bezug auf Unterrichtsmerkmale ist durchaus bekannt und wurde zum Beispiel in einer Arbeit von Clausen (2002) behandelt. Daraus weiß man, dass die Übereinstimmung zwischen den Perspektiven mit zunehmendem Komplexitätsgrad des zu bewertenden Merkmals sinkt. Bei Selbstauskünften von Schülern besteht allgemein die Gefahr, dass sie die erfragten Aspekte nicht so auffassen, wie es aus Forschungsperspektive intendiert ist. Aber auch bei Befragungen der Lehrkräfte zum Verhalten ihrer Schüler sind Ergebnisse mit Zurückhaltung zu interpretieren, da sie in der Situation, zu der sie Auskunft geben und die sie bewerten sollen, selbst als agierende Person eingebunden sind. Es ist damit nicht auszuschließen, dass sich Eindrücke überlagern oder retrospektiv verändern. Weiterhin ist zu beachten, dass befragungsbasierte Urteile, sofern sie nicht von den Betroffenen selbst und direkt aus der Situation heraus erhoben werden, wie beispielsweise mit Hilfe der *experience sampling method* (Yair, 2000), immer ein summatives und retrospektives Bild über den Unterricht abgeben und keine verlaufsspezifischen, prozesssensiblen Informationen liefern. Besonders für den Gegenstand des Schülerverhaltens stellt das eine gewisse

Einschränkung dar: Welche Personen sich wann, wie oft und wie lange auf bestimmte Art und Weise verhalten, ist durch nachträgliche Befragungen prinzipiell nicht zu erschließen.

Aufgrund dessen erscheinen Verfahren systematischer Beobachtung als am besten geeignete Datenerhebungsmethode für eine möglichst präzise Erfassung des Schülerverhaltens im Unterricht. In der Geschichte der Erforschung von Schülerverhalten im Unterricht der vergangenen hundert Jahre stellen Beobachtungen durch Forscher/-innen oder Lehrkräfte entsprechend auch die häufigste Erhebungsmethode dar (siehe Überblicke bei Anderson, 1984a; Borg, 1980; Helmke, 1988; Jackson, 1968/1990). Aus der Anwendung von Beobachtungsmethoden auf diesen Forschungsgegenstand ergeben sich allerdings gewisse Herausforderungen. Nahezu alle auf Unterrichtsbeobachtungen basierenden Studien in diesem Feld berücksichtigen lediglich eingeschränkte Ausschnitte der Lerngruppe oder des Geschehens, indem entweder repräsentativ einige Schüler aus der Lerngruppe ausgewählt (z.B. sechs Schüler je Klasse bei Houtveen, Booij, de Jong & van de Grift, 1999) oder zeitlich limitierte Momentaufnahmen jedes Schülers in sequenziellen, rotierenden Personenstichproben (z.B. Dishaw, 1978; Helmke, 1988) festgehalten werden. Auf dieser Basis wird in den meisten Fällen der prozentuale Wert kategorialer Variablenausprägungen pro Unterrichtsstunde berechnet und anschließend in Form einer anteiligen Rate (z.B. als Aufmerksamkeits- oder *on task*-Rate) für einzelne oder mehrere Unterrichtsstunden berichtet. Das bedeutet, mögliche zeitabhängige Variationen des Schülerverhaltens im Stundenverlauf, aber auch individuelle Verhaltensprofile in einer Unterrichtsstunde können durch derlei Verfahrensweisen nicht berücksichtigt werden. Zum anderen sehen sich Beobachtungsverfahren in diesem Bereich mit gewissen validitätsbedingten Grenzen konfrontiert. Der breite Bestand an Forschungsbefunden zeigt relativ übereinstimmend ein hohes Aufmerksamkeits- bzw. Beteiligungsniveau der Schüler im Unterricht auf. Zum Beispiel liegen die bei Jackson (1968/1990) zusammengetragenen Befunde aus der Forschung zur Schüleraufmerksamkeit zwischen 1920 und 1960 ausnahmslos in einem Bereich zwischen 80 und 95 % (S. 88ff.). Eine Reihe Studien zu aufgabenbezogenem Verhalten und Beteiligungsverhalten (*on task behavior*, *time on task* und *engagement rate*) der 1970er und 1980er Jahre zeigen ein ähnlich hohes Niveau von durchschnittlich um die 80 % (Fredrick & Walberg, 1980; Helmke, Schneider & Weinert, 1986; Rosenshine,

1980; Wang & Walberg, 1983). Arbeiten mit dem Münchner Aufmerksamkeitsinventar von Helmke und Renkl (1988, 1992) setzen dieses Bild mit *on task*-Raten von 80 - 85 % fort, ebenso die Ergebnisse einer Londoner Arbeitsgruppe (Blatchford, Bassett, & Brown, 2011: 80 - 86 %). Diesen Ergebnissen, die allgemein den Eindruck vermitteln, als seien Schüler die meiste Zeit des Unterrichts den Lerngelegenheiten zugewandt, stehen Erkenntnisse aus Studien gegenüber, die die Divergenz zwischen beobachtbarer Aufmerksamkeit oder Beteiligung und dem Grad tatsächlicher geistiger Auseinandersetzung mit den Inhalten zum Gegenstand haben. Eine bei Borg (1980, S. 55) berichtete Studie von Bloom aus dem Jahre 1976 zeigt mit Hilfe der Methode des *stimulated recall*, einem Befragungsverfahren zu Gedanken und Wahrnehmungen einer Person in einer zurückliegenden Situation anhand von erinnerungsstimulierenden Reizen (in diesem Falle Tonaufnahmen des Unterrichts), dass lediglich 55 - 65 % der selbstberichteten Gedanken der Schüler einen Unterrichtsbezug hatten. Ebenfalls unter Anwendung dieser Methode, allerdings ergänzt um Schülerbefragungen und Unterrichtsbeobachtungen, fanden Peterson, Swing, Stark und Waas (1984) heraus, dass das beobachtete aufgabenbezogene Verhalten von Schülern keinen Zusammenhang zum Grad ihrer selbstberichteten Aufmerksamkeit oder ihrer kognitiven Prozesse während des Unterrichts aufweist ($r = -0,04$ bzw. $0,08$; S. 500), womit Sie ähnliche Ergebnisse aus einer eigenen Vorstudie bestätigen (*off task*-Verhalten korreliert mit selbstberichteter Aufmerksamkeit: $r = 0,09$; Peterson, & Swing, 1982). Teilweise untermauert werden diese Befunde durch die bereits vorgestellte Studie von Yair (2000), in der mittels *experience sampling method* die Gedanken der Schüler im Unterricht erfasst und dabei eine verhältnismäßig niedrige „tatsächliche“ *engaged rate* von durchschnittlich 54 % der Unterrichtszeit herausgefunden wurde.

Diese Ergebnisse illustrieren ein Phänomen, welches als verdeckte Aufmerksamkeit (Helmke, 1988, S. 7) bezeichnet wird. Demnach ist der äußerlich wahrnehmbare Eindruck eines Schülers nicht gleichzusetzen mit seinen tatsächlichen inneren Zuständen¹⁴. Beobachtungsbasierte Untersuchungen in diesem Bereich stehen damit vor einem Validitätsproblem, welches in die Anforderung resultiert, Aspekte innerer kognitiver Prozesse und des sichtbaren Verhaltens von Schülern möglichst klar

¹⁴ Zu diesem Aspekt s. auch Kap. 2.4.1.

voneinander abgrenzen (z.B. in Form getrennter Erhebungsinstrumentarien oder theoretisch differenzierter Konstrukte).

2.3.4 Zusammenfassung des Forschungsstandes und Schlussfolgerungen

Im vorherigen Abschnitt wurde entlang von zwei Forschungslinien, die sich intensiv mit verschiedenen Formen der aktiven Zuwendung der Schüler zu den Inhalten und Aufgaben des Unterrichts auseinandersetzen, der Stand der Forschung zum Schülerbeteiligungsverhalten im Unterricht dargestellt: zum einen die Forschung zur *educational effectiveness* bzw. *educational productivity*, die zugleich theoretisch bedeutsame Zugänge für die Ziele dieser Studie liefert, und zum anderen in ergänzender Weise die Forschung zum *student engagement*. Einen insgesamt beständig diskutierten Gegenstand stellt der Zusammenhang zwischen Schülerbeteiligungsverhalten im Unterricht und durch Unterricht induzierten Lerneffekten dar. Diesbezüglich bestätigt die teils sehr weit zurückreichende Empirie die Annahme, dass ein hohes Maß an Zuwendung der Schüler zum Unterrichtsgegenstand – sei es durch aktive Beteiligung oder bloße Aufmerksamkeit – zu besseren Lernerfolgen in den betreffenden Inhalten (und in Teilen auch überfachlichen Lernbereichen) beiträgt. Damit präsentiert sich dieser Bereich im Anschluss an die eingangs dargestellten theoretischen Annahmen als wichtige Vorbedingung von Lern- oder auch anderen Effekten, die durch schulischen Unterricht initiiert werden sollen, und damit gleichzeitig als legitimer Zielindikator für die Wirksamkeit von Unterrichtsprozessen. Hinsichtlich der Effektstärken zeichnet sich zwar trotz durchgehend positiver Zusammenhänge kein vollständig konsistentes Bild ab, doch lässt sich das möglicherweise zurückführen auf Schwierigkeiten der Vergleichbarkeit und Synthese von Ergebnissen aufgrund der Vielfalt von Operationalisierungen und methodischer Zugänge.

Mit Blick auf die Frage, wie das Beteiligungsverhalten der Schüler zu beeinflussen ist, zeigt die Forschung, dass die individuellen Eingangsvoraussetzungen der Schüler, z.B. in Form ihres persönlichen (ethnischen) Hintergrunds, ihres allgemeinen oder fachspezifischen Fähigkeitsniveaus oder ihrer allgemeinen Lernmotivation, zunächst einen gewissen Rahmen liefern für den Einflussspielraum innerhalb des Unterrichts. Die Befundlage zeigt, dass diese Hintergrundvariablen zwar in individuell unterschiedlich

ausgeprägte Verhaltenstendenzen resultieren, diese allerdings letztlich durch Effekte der Eigenschaften der Unterrichtssituation überwogen werden.

Bei der sich anschließenden Frage, wie denn Unterricht im Hinblick darauf am besten zu gestalten sei, erweisen sich aus empirischer Sicht eine effektive Klassenführung und eine angemessene Gestaltung der Inhalte und Aktivitäten im Hinblick auf die Voraussetzungen der einzelnen Schüler als die bedeutsamsten Unterrichtscharakteristika. Diese Befundlage schließt an die eingangs dargestellten theoretischen Ansätze (vgl. Kap. 2.2) an und bestätigt dabei besonders die Kernannahmen von Jackson (1968/1990) und Slavin (1994). In der Summe spricht damit vieles dafür, dass auf der einen Seite eine effiziente Klassenführung das Einbringen von Lerngelegenheiten in den Unterricht durch die Lehrkraft und deren Wahrnehmung durch die Schüler überhaupt erst ermöglicht, was in der Folge die Wahrscheinlichkeit des Beteiligungsverhaltens erhöht. Dieser Effekt wurde wie gezeigt bereits mehrfach bearbeitet und gilt weitgehend als wissenschaftlich abgesichert. Auf der anderen Seite scheint die Qualität der Lerngelegenheiten, dabei insbesondere der Schwierigkeitsgrad (bzw. die Art der Anforderungen) und die Attraktivität der Aufgaben, darüber zu entscheiden, ob sich die Schüler dem Unterricht zuwenden, wobei noch wenig darüber bekannt ist, wie sich dieser Zusammenhang auf der konkreten Prozessebene des Unterrichts vollzieht und welche Rolle dabei die persönlichen Ausgangslagen der Schüler spielen.

Vor diesem Hintergrund besteht unter der Zielperspektive der Unterrichtswirksamkeit eine zentrale Frage darin, wie die Eigenschaften der in den Unterricht implementierten Aufgaben dazu beitragen können, dass möglichst viele Schüler mit ihren je unterschiedlichen Lernvoraussetzungen sie letztendlich ergreifen. Von großer Bedeutung wäre diesbezüglich eine individualisierte Sicht auf das Schülerverhalten im Unterricht, ebenso wie ein umfassender Einblick in den Unterrichtsverlauf, um differenzielle Effekte der Lernanlässe auf die Schüler zu unterschiedlichen Zeitpunkten des Unterrichts erkennen zu können. Die Untersuchung der Beziehung zwischen den Eigenschaften der Aufgaben und dem Beteiligungsverhalten der Schüler innerhalb des Unterrichtsprozesses unter Berücksichtigung interindividueller Unterschiede bildet den zentralen Ansatzpunkt für die vorliegende Studie. Um im weiteren Verlauf eine Präzisierung dieses Untersuchungsziels und der forschungsmethodischen Konzeption vornehmen zu können, soll zunächst aber mit Blick auf die tragende Rolle des

Schülerbeteiligungsverhaltens für diese Studie eine differenzierte Betrachtung der unterschiedlichen Konstrukte und Begrifflichkeiten in der Forschung zu diesem Bereich vorgenommen werden.

2.4 Konstrukte und Begrifflichkeiten im Bereich des Schülerbeteiligungsverhaltens

Bei den vorangegangenen Darstellungen wurden unterschiedliche, teils synonym verwendete Bezeichnungen deutlich, die für die aktive Nutzung bzw. Hinwendung der Schüler zu den Lerngelegenheiten des Unterrichts stehen, weshalb in den bisherigen Ausführungen des Öfteren die Rede von einem Merkmalsbereich war. Konstrukte wie *engagement*, *time on task* oder Aufmerksamkeit bringen verschiedene Komponenten schülerseitigen Verhaltens und Erlebens im Unterricht zum Ausdruck, die sich in einigen Teilen überschneiden, aber dennoch nicht deckungsgleich sind und damit teils unterschiedliche theoretische, besonders aber methodische Schwerpunktsetzungen erfordern. Zur Einordnung des in dieser Arbeit verfolgten Ansatzes sollen daher einige differenzierende Betrachtungen der in diesem Zusammenhang relevantesten Begriffe vorgenommen werden.

Aufmerksamkeit

In einer Reihe theoretischer und empirischer Arbeiten wird Aufmerksamkeit bedeutungsgleich mit *on task*-Verhalten oder *engagement* diskutiert und in der Folge zumeist als Indikator für Lernaktivitäten von Schülern im Unterricht angesehen (vgl. Carroll, 1963; Fisher, Berliner, Filby & Marliave, 1980; Karweit, 1989; Kreuzberger, 2002; Rosenshine & Berliner 1978). Etwa im Rahmen des Angebots-Nutzungs-Modells von Helmke sind Lernaktivitäten von Schülern im Unterricht mit Aufmerksamkeitsverhalten gleichgesetzt (Helmke 2003). Zur Beobachtung des Aufmerksamkeitsverhaltens schlägt der Autor das Münchner Aufmerksamkeitsinventar (MAI) vor, in dessen Beschreibung die Termini Aufmerksamkeits- und Lernverhalten synonym verwendet werden (Helmke, 1988; Helmke & Renkl, 1993, s. Kap. 2.3.2). In diesem Zuge wird implizit die Annahme formuliert, dass der Aufmerksamkeitsstatus als Indikator für Lernprozesse angesehen werden kann, die im jeweiligen Moment stattfinden oder initiiert werden („*Registrierung mutmaßlicher Lernaktivitäten*“, Helmke,

1988, S. 35). Betrachtet man den Begriff der Aufmerksamkeit aus pädagogischer und psychologischer Perspektive genauer, erscheint diese Verknüpfung allerdings nicht sehr zuverlässig. So wird in einem pädagogischen Grundlagenband Aufmerksamkeit definiert als *„Die Leistung des Bewusstseins, aus der Vielfalt des (innerlich und äußerlich) Wahrnehmbaren etwas als Inhalt auszuwählen, wird als Aufmerksamkeit bezeichnet“* (Dinkelaker, 2011, S. 175). Aus Sicht der pädagogischen Psychologie fassen Schmidt-Atzert, Büttner & Bühner (2004) Aufmerksamkeit in Übereinstimmung dazu als das *„selektive Beachten relevanter Reize oder Informationen“* (S. 5), also als kognitive Reizselektion auf. Aufmerksamkeit ist demnach zunächst eine Wahrnehmungsleistung und grenzt sich von mentalen Prozessen daran anschließender Informationsverarbeitung und damit auch des Lernens ab. Die Ausrichtung der Wahrnehmung der Schüler auf das inhaltsbezogene Geschehen kann unter dieser Perspektive zwar als eine notwendige, aber keine hinreichende Vorbedingung für unterrichtliches Lernen gesehen werden. In einer Reihe von Arbeiten wird zudem hervorgehoben, dass Aufmerksamkeit keinen zuverlässigen Indikator für die Ausrichtung des Schülerverhaltens oder -erlebens an den Unterrichtsinhalten darstelle und daher nur wenige Rückschlüsse auf unterrichtliches Lernen zulasse. Jackson (1968/1990) bringt dies mit der Formulierung *„all eyes on the teacher does not necessarily mean all thoughts on the topic at hand“* (S. 102) zum Ausdruck. Daran schließen weitere Überlegungen zur Differenzierung zwischen Aufmerksamkeit und tiefer liegenden Formen der Zuwendung zu den Unterrichtsinhalten an:

„... attention and involvement are not the same conditions and the teacher would do well to keep the distinction in mind. Even though he might labor to control attention and though he may be forced to rely on signs of alertness as indicators of involvement, it is the latter condition, rather than the former, that he is seeking to cultivate.“ (S. 111).

Demgemäß ist Aufmerksamkeit im Sinne eines Zustandes der „Wachheit“ der Schüler kein ausreichender Indikator für die Wirksamkeit des Unterrichts. Denn nach Jacksons Ansicht bestehe die Möglichkeit einer Inszenierung von Aufmerksamkeit, um sich innerlich dem Unterrichtsgeschehen und infolgedessen auch einem möglichen Lernen durch die gebotenen Inhalte zu entziehen. Diese Strategie des „psychologischen Rückzugs“ (S. 34) besteht in einer graduellen Reduzierung der persönlichen Teilnahme und des eigenen Einbezogenenseins in das Unterrichtsgeschehen bis zu dem Punkt, an dem keine offiziellen oder formalen Verhaltenserfordernisse an das Individuum, sprich

Verhaltensregeln im Unterricht, für Außenstehende spürbar enttäuscht werden. Diese als formal zu bezeichnenden Erfordernisse sind auf der sichtbaren Ebene des Geschehens fassbar und betreffen offensichtliche Indikatoren von Aufmerksamkeit, während sich die inneren Rückzugsstrategien auf der Ebene des so genannten *hidden curriculum* (ebd., s. auch Kap. 2.2) abspielen.

Beeinflusst von diesen Überlegungen haben sich einige weitere Arbeiten aus unterschiedlichen Perspektiven mit dem Thema der Inszenierung von Aufmerksamkeit befasst. Heinze (1980) beleuchtet unter Rückgriff auf den sozialpsychologischen Ansatz von Jackson und die Soziologie Goffmans verschiedene Formen so genannter Schülertaktiken. Dies sind Verhaltensweisen, die unterschiedliche Konstellationen des Zusammenspiels von innerer und sichtbarer Zuwendung zum Unterricht, von „*subjektivem*“ und „*anlassgemäßem Engagement*“ (S. 86) bezeichnen. Im Vordergrund steht dabei allerdings mehr das Ergründen schulischer Anpassungsmechanismen im Sinne von Bestandteilen sozialisatorischer Wirkungen in der sozialen Klassensituation als Fragen der Unterrichtsführung. Markowitz (1986) diskutiert überdies aus Sicht der Systemtheorie ebenfalls das Schülerverhalten im Schulunterricht und benennt in diesem Zuge ähnliche Phänomene als Probleme pädagogischen Handelns. Aufmerksamkeit vollzieht sich seiner Ansicht nach in „*einzelnen Akten des Aufmerkens*“ (S. 57), als „*abtastendes Umherschweifen*“. Dieses ist von außen schwer nachvollziehbar und birgt das Potenzial für „*verdeckte Engagements*“ in sich, einer „*Vielfalt von Aktivitäten hinter der glatten Fassade der Aufmerksamkeit*“ (S. 81). Im Anschluss daran arbeitet Breidenstein (2006) in einer ethnographisch-erziehungswissenschaftlichen Studie zu Praktiken der Teilnahme am Unterricht eine Konstellation des Schülerverhaltens heraus, die er „*Schülerjob*“ nennt. Gemeint ist damit die formale Teilnahme am Unterricht und gleichzeitige Distanznahme vom inhaltlichen Geschehen als Form des Umgangs mit dem gegebenen institutionellen Setting, indem die Schüler mit einer gewissen Routine die Verhaltensanforderungen in verschiedenen Unterrichtssituationen erfüllen.

All dies deutet darauf hin, dass der nach außen signalisierte Aufmerksamkeitsstatus und das Maß des tatsächlichen inneren Aufmerkens mitunter stark differieren können. Die Entscheidung darüber, inwiefern sie sich dem Geschehen zuwenden und – in der Terminologie des Angebots-Nutzungs-Konzepts von Unterricht ausgedrückt – das Unterrichtsangebot annehmen und nutzen, treffen die Schüler letztlich selbst. Diesen

Aspekt der Selbststeuerung beschreiben Autoren wie Anderson (1984a, S. 47) oder eben auch Breidenstein (2006, S. 66) mit der Metapher des „Ein-“ oder „Abschaltens“¹⁵. Pädagogisches Handeln im Unterricht versucht nun aber, auf die Aufmerksamkeit Einfluss zu nehmen, sie zu lenken, um darüber unterrichtliche Lerngelegenheiten zu platzieren (Dinkelaker, 2011). Damit stellt sich sowohl für die Unterrichtspraxis als auch die Erforschung von Unterricht das Problem, dass Schüleraufmerksamkeit zwar angestrebt wird, der tatsächliche Aufmerksamkeitsstatus der Schüler aber nicht ohne Weiteres von außen erkennbar ist. Dessen Beurteilung muss sich auf wenige offensichtliche Verhaltensindikatoren beschränken, wie zum Beispiel Blickkontakt, Ausrichtung des Körpers, Vollzug bestimmter Handlungen oder verbale Äußerungen (ebd.).

Engagement

Ähnlich wie bei der Aufmerksamkeit wird auch *engagement* häufig in einem Zuge oder gar synonym mit angrenzenden Konstrukten, wie z.B. *on task*-Verhalten, Beteiligung (*involvement*) oder eben auch Aufmerksamkeit, genannt (z.B. Berliner, 1990; Blatchford, Bassett, & Brown, 2011; Helmke, Schneider & Weinert, 1986; Karweit, 1989; Yair, 2000). Kennzeichnend dafür formuliert Borg (1980) in einem Literaturüberblick: „*Engaged time is essentially synonymous to time on task, attention, and participation – all of which are found in the research literature*“ (S. 55). Während allerdings Aufmerksamkeit wie gezeigt als ein Indiz der rezeptiven Hinwendung der Schüler zu einem Aspekt des Unterrichtsgeschehens gewertet werden kann, integriert der seitdem besonders in der englischsprachigen Forschung gebräuchliche Begriff des *engagement* darüber hinaus unterschiedliche Komponenten der Auseinandersetzung der Schüler mit dem Unterricht. So umschreibt Borich (2010) in einem Lehrbuch zur Unterrichtsmethodik *student engagement* als Schülerverhalten im Unterricht, das sich kennzeichnet durch „*actively thinking about, working with, or using what is being presented*“ (S. 12). Nystrand & Gamoran (1991) hingegen beschreiben zum einen *procedural engagement* als Formen korrekten Betragens und Mitarbeitens im Unterricht und zum anderen *substantive engagement* als innere Verbundenheit mit den Fachinhalten auf der anderen Seite. Im

¹⁵ „*Students ,tune in' to certain classroom events, and then either ,tune in' to other events or literally ,tune in(ward)' to spend some time ,day-dreaming'.*“ (Anderson, 1984a, S. 47)

Rahmen empirischer Forschung finden sich derartige Konzepte ausformuliert als konkrete Untersuchungsvariablen: Zum Beispiel in der deutschen SCHOLASTIK-Studie (Weinert & Helmke, 1997) wird das so genannte Schülerengagement durch die Häufigkeit des Meldeverhaltens, Schülerfragen und Mitarbeit operationalisiert. Walberg & Lai (1999) definieren *engagement* in einem Forschungsüberblick als aktive und anhaltende Schülerbeteiligung, die sich in unterschiedlichen Aktivitäten des Schülers, wie etwa angemessenem Verhalten im Klassenraum, der Konzentration auf Aufgaben oder Wortbeiträgen zeigen kann. An dieser Stelle wird die Breite an Auffassungen über konkrete Ausformungen des Schülerengagements deutlich.

Als hilfreich erweist sich in diesem Zusammenhang der Überblicksartikel von Fredricks, Blumenfeld & Paris (2004), in dem die Autoren die Bedeutung des *school engagement*-Konzepts herausarbeiten und versuchen, dessen Begriffsraum zu systematisieren. Dabei wird *engagement* zunächst als ein breites Konzept gesehen, welches sowohl die Unterrichts- als auch die Schulebene betrifft und auf unterschiedliche Weise das „*commitment to education*“ der Schüler, d.h. ihren Einsatz für und ihre persönliche Bindung an den Bildungsgang beschreibt (S. 59). In einer theoretischen und empirischen Bestandsaufnahme kommen die Autoren zu der Feststellung, dass ein grundsätzlicher Bedarf zur konzeptionellen Schärfung der beteiligten Begriffe und Konstrukte besteht: „*Our concern is that, although engagement has considerable practical benefit as an umbrella that synthesizes a broad range of research, it suffers from being everything to everybody*“ (S. 84). Aus diesem Grund schlagen sie eine Systematisierung des Konzepts in drei Komponenten vor, die jeweils eine Facette des *engagements* abbilden und sich in unterschiedlichen Ausprägungen des Denkens, Fühlens und Verhaltens von Schülern niederschlagen:

1. *behavioral engagement*

normenkonformes Verhalten (in Klassenraum und Schule), Arbeits- und Aufmerksamkeitsverhalten im Unterricht und Teilnahme an Aktivitäten im Schulkontext wie beispielsweise der Schülervertretung

2. *emotional engagement*

Bindung an die Institution und Beziehung zu ihren Akteuren, Lernemotionen, Einstellungen zur Schule, Motivation und Interesse

3. *cognitive engagement*

Bemühungen und Leistungsbereitschaft, Selbststeuerung und metakognitive Strategien

Diese in der Fachdiskussion bis heute weitgehend als konsensfähig geltende Systematisierung (s. verschiedene Beiträge in Christenson, Reschly & Wylie 2012) macht deutlich, dass das Spektrum des *engagement*-Begriffs in Teilen über den Klassenraum hinaus reicht. Aber auch unter Fokussierung der Unterrichtsebene zeigt sich das *student engagement* als ein relativ breites Konzept, das sowohl offensichtliches Schülerverhalten umfasst wie zum Beispiel regelkonformes Betragen, aktive Beteiligung oder Mitarbeit, als auch weniger offensichtliche oder innere Zustände wie Aufmerksamkeit oder Aspekte der Motivation und Volition, sprich des Wollens und Bemühens. Je nach Interessensbereich sind daher genauere Eingrenzungen nötig, um den *student engagement*-Begriff für spezifische Fragestellungen nutzbar zu machen. In der deutschsprachigen Unterrichtsforschung ist *engagement* über wenige Ausnahmen hinaus (s. Buff, Reusser & Pauli, 2010; Weinert & Helmke, 1997;) nicht als theoretisches Konstrukt etabliert, was möglicherweise auch mit den sprachlich bedingten, unterschiedlichen Begriffskonnotationen des englischsprachigen *engagement* und des deutschsprachigen Begriffs „Engagement“ zusammenhängt.

Beteiligung

Bezogen auf Unterricht bringt der im Deutschen gebräuchliche Begriff der Beteiligung verschiedene Aspekte zum Ausdruck. Einige Arbeiten beschäftigen sich in diesem Zusammenhang mit der Perspektive der Teilhabe von Schülern am Unterricht im Sinne von Mitbeteiligung oder Partizipation (Böhme & Kramer, 2001). Dabei geht es im Wesentlichen um den Einbezug von Schülern in Entscheidungsprozesse zur Unterrichtsgestaltung, womit entweder auf eine „*Partizipationskultur*“ zur Förderung eines schülerbezogenen Selbstbestimmungsideals abgezielt wird (Kötters, Schmidt & Ziegler, 2001), oder auf didaktischer Ebene unter dem Vorzeichen der so genannten „*Lernerautonomie*“ das selbstbestimmte und selbstgesteuerte Lernen der Schüler Betonung findet (Meyer, 2001). Abgesehen von dieser Fokussierung auf Mitbestimmung und Mitgestaltung durch die Schüler markiert eine andere Auslegung des Beteiligungsbegriffs den Aspekt der Teilnahme von Schülern am Unterricht bzw. ihr

Einbezogenheit in die Auseinandersetzung mit dem Unterrichtsgegenstand. Beteiligung kann demzufolge als Involviertheit von Schülern in die fachlich orientierte Kommunikation, als Meldeverhalten der Schüler oder Aufrufverhalten der Lehrkraft (bzw. auch einer Kombination aus diesen beiden) verstanden werden (Helmke, 1997; Lipowsky, Pauli, & Rakoczy, 2008; Pauli & Lipowsky, 2007; Sacher, 1995). Auf diese Weise gerät sowohl das „Sich-Beteiligen“ im Sinne aktiver Mitarbeit der Schüler als auch das „Beteiligt-Werden“ in Form von Aufrufen durch die Lehrkraft in den Blick. Schülerbeteiligung präsentiert sich damit als eine mögliche Bezeichnung sowohl für Schüler- als auch Lehrerhandeln oder einer Kombination aus beidem. Lipowsky, Pauli und Rakoczy (2008) sehen in diesem Zusammenhang die Beteiligung von Schülern an der didaktischen Kommunikation als einen Indikator für die aktive Nutzung von unterrichtlichen Lerngelegenheiten. Die möglichst gleichmäßige Beteiligung von Schülern am inhaltsbezogenen Unterrichtsgespräch sei demnach zu verstehen *„als ein Gradmesser für die potenzielle Aufmerksamkeit der Schülerinnen und Schüler und für die Auseinandersetzung mit dem Unterrichtsgegenstand“* (S. 69). Durch wen diese Teilnahme am Unterrichtsgespräch letztlich initiiert wurde, ob der sprachliche Beitrag also auf eine Meldung des Schülers oder einen Aufruf der Lehrperson zurückgeht, wird dabei allerdings außer Acht gelassen. Dass Schülerbeteiligung auch als ein abstraktes Merkmal ohne konkreten Urheber gesehen werden kann, steht bezeichnend für die Breite des Begriffs.

Trotz fehlender Spezifität des Beteiligungsbegriffs wird im vorliegenden Kontext der Ausdruck des Schülerbeteiligungsverhaltens als deutschsprachiger Begriff für den übergeordneten Merkmalsbereich verwendet. Dieser vermag zum einen Abgrenzungen deutlich zu machen gegenüber (1) einer Mitbeteiligungs- und Partizipationsperspektive auf die Rolle von Schülern im Unterricht, (2) dem Ansatz, Schüler- und Lehrerhandeln in einer gemeinsamen Beteiligungsvariable zu integrieren, und (3) vom Verhalten losgelöste Komponenten des Wahrnehmens und Erlebens der Schüler.

Aufgabenbezogenes bzw. on task-Verhalten

In der im vorangegangenen Abschnitt erwähnten Arbeit von Lipowsky, Pauli und Rakoczy (2008) wurde der Versuch deutlich, über den Begriff der Schülerbeteiligung eine Verknüpfung zwischen offenkundig wahrnehmbarer Teilnahme der Schüler am

Unterrichtsgespräch und ihren inneren Prozessen herzustellen, die die Auseinandersetzung mit den Unterrichtsinhalten markieren. Diese Bezogenheit des Schülerverhaltens auf die Unterrichtsinhalte wird teilweise auch mit dem Begriff des *on task*-Verhaltens gekennzeichnet.

Im Rahmen der Unterrichtseffektivitätsforschung und einer damit verbundenen geführten Auseinandersetzung mit dem Faktor Zeit als Ressource für den Unterricht wird *on task*-Verhalten seit den 1970er Jahren besonders im englischsprachigen Bereich oft in Verbindung mit dem Konstrukt der *time on task*¹⁶ verwendet (Bloom, 1974; Myers, 1990). So warnt Anderson (1984b) im Zuge ihres Überblicksartikels zur *time on task* vor der zu jener Zeit erkennbaren Tendenz, in der Forschung den *task*-Aspekt zugunsten des *time*-Aspekts zu vernachlässigen (S. 160). Allerdings konnte sich das Konzept des *on task*-Verhaltens auch losgelöst von dem Zeitaspekt als eigenständiges Konstrukt etablieren, wobei sich die Frage stellt, inwiefern sich eine dazu passende, eigenständige Verhaltenskonstellation in Abgrenzung zu anderen, angrenzenden Konstrukten (s.o.) benennen lässt.

In einer Reihe von empirischen Arbeiten werden Anzeichen der Aufmerksamkeit (Burns & Dean, 2005), Meldeverhalten und Wortbeiträge (Helmke, Schneider & Weinert, 1986; Helmke, 1997), themenbezogene Interaktionen oder Mitarbeit (Blatchford, Bassett, & Brown, 2011) oft in gleicher Weise als Indikatoren sowohl für *engagement* als auch *on task*-Verhalten gewertet. So sind beispielsweise in einem Beobachtungssystem von Emmer, Sanford, Clements & Martin (1982) zur Beurteilung von Schüleraktivitäten im Unterricht alle diese Verhaltenskomponenten einbezogen: Unter dem Titel „*Student Engagement Rating*“ dient es zur Verhaltenskategorisierung anhand unterschiedlicher Formen von *on*- und *off task*-Verhalten. Gegenstand sind sichtbare Aktivitäten von Schülern, wie beispielsweise an einer Aufgabe arbeiten, mit den Blicken einem Vortrag folgen, sich regelkonform betragen oder aufgabenorganisatorische Tätigkeiten verrichten. Während dies Aktivitätsaspekte sind, die üblicherweise auch im Kontext des *student engagement* diskutiert werden (s. Kap. 2.4.2), wird das Spezifikum des *on task*-Verhaltens, also die Orientierung der genannten Schüleraktivitäten an den Aufgaben im Unterricht, daran geknüpft wird, ob die gezeigten Verhaltensweisen im Rahmen definierter Erwartungen der Lehrkraft in Bezug auf eine bestimmte Aufgabe liegen (vgl.

¹⁶ dazu s. Kap. 3.2

auch Gickling & Armstrong, 1978). Im Rahmen anderer Arbeiten wird der Aufgabenbezug des Verhaltens auf ähnliche Weise festgestellt: etwa daran, ob der Unterricht zu diesem Zeitpunkt überhaupt eine Aufgabe anbietet und die Verhaltensweisen der Schüler sich zugleich in irgendeiner Weise „als Nutzung der jeweiligen unterrichtlichen Lerngelegenheiten interpretieren lassen“ (Helmke, 1988, S. 13), oder ob zusätzlich zu den genannten Hinweisen die Interaktionspartner des Schülers (Lehrkraft, Mitschüler) aktuell in eine Aufgabe verwickelt sind (Blatchford, Bassett, & Brown, 2005, 2011).

Insgesamt lassen sich also solche Verhaltensweisen als *on task* bezeichnen, die erkennbar auf eine gerade präsente inhaltliche Aufgabe im Unterricht ausgerichtet sind, wobei sich diese Orientierung darüber definiert, ob die Personen oder Objekte, die der Schüler zum Gegenstand seiner Wahrnehmung macht, oder die Tätigkeiten und Handlungen, die er vollzieht, den Anforderungen und Erwartungen, die sich aus der Aufgabenstellung ergeben, entsprechen. Auf diese Weise findet die Abhängigkeit der Schüleraktivitäten von der Verfügbarkeit unterrichtlicher Lerngelegenheiten in Form von Aufgaben besondere Betonung. Aufgrund der engen Bindung des Schülerverhaltens an konkrete inhaltliche Unterrichtsangebot wird dem Konzept des *on task*-Verhaltens immer wieder bescheinigt, ein sichtbares Anzeichen für Lernprozesse im Unterricht bzw. ein Indikator für „Lernverhalten“ zu sein (vgl. Harris & Yinger, 1977, zit. nach Myers, 1990; S. 16; Helmke, 2012). Demzufolge kann das Konzept des *on task*-Verhaltens als größtmögliche Annäherung an die in theoretischen Modellen so genannten „individuellen Verarbeitungsprozesse“ (Helmke, 2012, S. 80) auf Seiten der Schüler angesehen werden, die als vermittelnde Größe zwischen den unterrichtlichen Lerngelegenheiten und der Initiierung von Lernprozessen bei den Schülern stehen.

Zusammenfassung und Schlussfolgerungen zur Betrachtung der Konstrukte und Begrifflichkeiten im Bereich des Schülerbeteiligungsverhaltens

Bei dem Versuch, die zur Beschreibung von unterrichtlichen Schüleraktivitäten häufig synonym verwendeten Begriffe Aufmerksamkeit, *engagement*, Beteiligung und *on task*-Verhalten konzeptionell voneinander abzugrenzen, wird deutlich, dass keine vollständig trennscharfen Definitionen gefunden werden können. Es lassen sich allenfalls theoretische Schwerpunktsetzungen hinsichtlich dessen erkennen, wie durch diese

Konstrukte die Zuwendung der Schüler zum Unterrichtsangebot beschrieben wird. Im Hinblick auf das Ziel dieser Untersuchung, zum Zwecke der Analyse der Wirksamkeit von Unterrichtsprozessen ein individualisiertes Bild der unterrichtlichen Aktivitäten von Schülern über den Zeitverlauf zu erhalten, sind diese Schwerpunktlegungen unterschiedlich zu bewerten. Aufmerksamkeit ist als die Ausrichtung der Wahrnehmung einer Person auf selektierte Objekte und Geschehnisse in ihrer Umgebung zu verstehen. Übertragen auf Schulunterricht bedeutet das üblicherweise, den Ausführungen der Lehrkraft oder des Arbeitspartners oder Wortbeiträgen der Mitschüler zu folgen beziehungsweise das Zentrum des Geschehens zu beobachten. Einerseits markiert Aufmerksamkeit eine notwendige Grundbedingung für im Unterricht initiierte Lernprozesse, andererseits ist Aufmerksamkeit für Fragen der Prozessanalyse von Unterricht aber nur schwer zugänglich, da die dazugehörigen kognitiven Vorgänge von außen kaum nachvollziehbar sind. Unter dem Begriff des *engagement* werden im Kontext von Unterricht verschiedene Facetten gefasst, wozu neben Aufmerksamkeit und Konzentration auch jegliche Formen von Arbeitsaktivitäten, die Befolgung sozialer Konventionen im Klassenraum und emotionale und motivationale Aspekte zählen. Aus methodischer Sicht sind Faktoren wie emotionale Zustände, situationsbedingtes Interesse und Motivation nur schwer kontinuierlich über den Unterrichtsverlauf erfassbar und mit sichtbaren Schüleraktivitäten in Zusammenhang zu bringen, womit das Konzept für die Ziele dieser Untersuchung in Gänze zu breit und zu schwierig operationalisierbar erscheint. Ähnliches gilt für den im deutschsprachigen Bereich häufig verwendeten Begriff der Beteiligung, der sich durch eine ähnliche Breite der damit assoziierten Phänomene auszeichnet wie das *engagement*, allerdings ohne auf einem vergleichbaren wissenschaftlichen Begriffskonsens aufzubauen. Der Begriff der Beteiligung bedarf demnach allgemein weiterer Spezifizierungen, die genauer darlegen, wer sich konkret an was beteiligt und worin sich das zu zeigen habe. Das Konstrukt des *on task*-Verhaltens spricht Schüleraktivitäten an, die an Aufgaben im Unterricht ausgerichtet sind, wobei sich diese Ausrichtung in der Erfüllung von Verhaltens- und Handlungserwartungen aus Sicht der Lehrkraft oder im Sinne der Aufgabenanforderungen niederschlägt. Mit diesem Ansatz ist allerdings die Anforderung verbunden, konkrete Schüleraktivitäten auf ein bestimmendes Konstrukt von Aufgaben beziehen zu können. Mit anderen Worten: Man muss zunächst wissen, was die Schüler tun sollen, um entscheiden zu können, ob das, was sie tun, aufgabenbezogen ist.

Das bedeutet, dass *on task*-Verhalten mit variierenden Charakteristika des jeweiligen Unterrichts oder der jeweiligen Unterrichtssituation unterschiedlich zu definieren ist, es damit aber auch eine höhere Situationsspezifität besitzt als die anderen dargestellten Konzepte. Zudem stützt sich das Konzept weitgehend auf von außen beurteilbare, sichtbare Verhaltensweisen, sodass dieses Konstrukt sich in besonderem Maße zur Verwendung in Analysen von Unterrichtsprozessen anbietet. Die Verknüpfung von *on task*-Verhalten mit der Zeitdimension des Unterrichts ist ein Gegenstand der Betrachtungen des folgenden Abschnitts (Kap. 3), der sich mit der Rolle der Unterrichtszeit im Zusammenhang mit Fragen der Wirksamkeit von Unterricht auseinandersetzt.

3. Zeit als empirische Metrik zur Beschreibung von Unterrichtsprozessen

In diesem Kapitel wird die besondere Stellung von Zeitfaktoren in der Forschung zur Wirksamkeit von schulischem Unterricht dargelegt und zentrale Diskussionen überblicksartig nachgezeichnet. Abschließend werden daraus Konsequenzen für die Anlage dieser Studie abgeleitet.

3.1 Unterrichtszeit aus theoretischer Sicht

Die Unterrichtsforschung, dabei insbesondere jene Bereiche, die sich mit der Wirksamkeit bzw. Effektivität von Schulunterricht beschäftigen, ist bereits seit mehreren Jahrzehnten eng mit Aspekten der Unterrichtszeit verknüpft. Ausgangspunkt dessen ist das bereits eingeführte Modell schulischen Lernens von Carroll (1963, s. Kap. 2.2), dessen Logik wesentlich auf Zeit als einheitlicher Grundmetrik für alle beteiligten Variablen basiert. Demnach lassen sich zum Beispiel die individuellen Lernvoraussetzungen eines Schülers beschreiben durch die hypothetisch benötigte Zeit zur Bewältigung einer bestimmten Aufgabe („*time needed in learning*“, S. 725), was bei einem zu hohen Schwierigkeitsgrad für das jeweilige Individuum unter Umständen auch in eine unendliche hohe benötigte Lernzeit resultieren kann (Carroll, 1984). Auch die Menge und Qualität der im Unterricht verfügbaren Lerngelegenheiten so wie das Ergreifen dieser Lerngelegenheiten durch die Schüler werden nach dieser Sichtweise als Zeitannteile formuliert (*time allowed for learning* und *time actually spent*, Carroll, 1963). Im Mittelpunkt steht dabei die Idee, dass im Grunde jeder Schüler jede Lernaufgabe bewältigen kann, sofern dafür genügend Zeit zur Verfügung steht und die Eigenschaften des Unterrichts währenddessen begünstigend wirken.

Der daraus ableitbare Gedanke individualisierbarer Lernzeiten lieferte Impulse für die Entwicklung verschiedener unterrichtspraktischer Konzepte wie zum Beispiel dem *mastery learning* (Bloom, 1968). Dessen konsequente Umsetzung ist allerdings mit einer hochgradigen Individualisierung von Lernmaterialien und Lernbegleitung durch die Lehrkraft bzw. mit einer graduellen Auflösung des Klassenverbands verbunden, sodass dieses Konzept schnell an unterrichtsorganisatorische Grenzen stößt, die im herkömmlichen institutionellen Rahmen kaum realistisch einzulösen sind (Merkens,

2010, S. 100). Innerhalb des als Weiterentwicklung von Carrolls Modell schulischen Lernens angelegten QAIT-Modells der Unterrichtseffektivität von Slavin (1994) spielt die zeitliche Dimension des Unterrichts ebenfalls eine zentrale Rolle. Während Zeit in diesem Modell als eine der vier gleichberechtigten Größen (*quality of instruction, appropriate levels of instruction, incentive* und *time*) dargestellt wird, ist sie genau genommen als separat zu sehende Basisgröße konzipiert, die den Umfang des anhand der drei übrigen Komponenten beschriebenen Unterrichts abbildet¹⁷. In dieser Kombination wird in gewisser Weise das Zusammenspiel von Qualität und Quantität des Unterrichts charakterisiert.

Der theoretische Mehrwert einer systematischen Berücksichtigung der zeitlichen Dimension des Unterrichts wurde in der Unterrichtsforschung unterschiedlich eingeschätzt. Auf der einen Seite wurde diesem Denkansatz die Rolle eines wissenschaftlichen Basisprinzips von wegweisender Bedeutung bescheinigt: *„The fact is that instructional time has the same scientific status as the concept of homeostasis in biology, reinforcement in psychology, or gravity in physics“* (Berliner, 1990, S. 1). Zeit als allgemeine Abbildungsmetrik für alle Unterrichtskomponenten biete die Vorteile der Einheitlichkeit (und damit Vergleichbarkeit) von empirischen Daten, der gesteigerten Verständlichkeit bzw. Anschaulichkeit von Variablenausprägungen und eines hohen Praxisbezugs erzeugter Befunde. Zeit sei zugleich die beste Möglichkeit zur Übersetzung abstrakter normativer Konzepte in konkretes Geschehen, welches sich durch Dauer und Lage im Zeitablauf spezifizieren lässt (S. 27). Dem steht die eher skeptischere Auffassung Treibers (1982) gegenüber, dass zeitbasierte Ansätze in der Unterrichtsforschung hauptsächlich zurückzuführen seien auf das Bestreben, *„schulisches Lernen auf notwendige Grundbedingungen zu beziehen, die empirisch einfach zu erheben, theoretisch sparsam zu begründen und zudem leicht zu beeinflussen sind.“* (S. 12). Das Forschungsprogramm gründe sich auf unterrichtsbezogene Zeitvariablen, die insgesamt *„nur wenig gehaltvoll und deshalb um so deutungsbedürftiger“* seien und dabei *„häufig wechselnde Indikatorfunktionen für weitaus komplexere Erklärungsbedingungen und -mechanismen“* übernehmen (S. 32). Deshalb sei anzuraten, die damit verbundenen

¹⁷ *„More time spent teaching a subject does not always translate into additional learning, but if instructional quality, appropriateness of instruction, and incentives for learning are all high, then more time on instruction is likely to pay off in greater learning.“* (Slavin, 1994, S. 148)

Grundannahmen „zu enttrivialisieren und theoretisch anzureichern“ (S. 25). Zeitbasierte Ansätze in der empirischen Unterrichtsforschung würden insgesamt von einer gewissen Attraktivität profitieren, deren Wirkung sich auch nicht durch damit verbundene forschungsmethodische Herausforderungen verringere (ebd.). Es ist allerdings nicht zu bestreiten, dass die Zeitdimension des Unterrichts in pragmatischer Hinsicht als ein pädagogisches Problem anzusehen ist. In einer theoretischen Studie zum Stellenwert des Zeitbegriffs in der Pädagogik betont Lüders (1995) unter anderem, dass Zeit als ein „*technologischer*“ Aspekt der Unterrichtspraxis (S. 9) seine Bedeutung vorrangig aus der Einbettung des Geschehens in institutionelle Organisationsschemata erhält, indem im Rahmen der Ordnung moderner Schulen die „*Fixierung von Terminen*“ notwendig wurde (S. 76); auf dieser Grundlage sei Zeit aber als legitimer Gegenstand pädagogischer Forschung zu Fragen der Didaktik und des Lehrerhandelns im Unterricht zu verstehen (S. 162).

3.2 Forschung zu Zeitfaktoren des Unterrichts

Mit Blick auf die internationale Schul- und Unterrichtsforschung ist indessen festzustellen, dass die Grundfigur des zeitbasierten Denkens über Unterrichtswirksamkeit sich insgesamt als sehr einflussreich erwiesen hat (Creemers & Kyriakides, 2008; Scheerens & Bosker, 1997; Walberg, 1988; Wang, Haertel & Walberg, 1993). Im Rahmen einer intensiven Debatte um die Wirksamkeit öffentlicher Schulen in den USA in den frühen 1970er Jahren wurde unter anderem die Frage aufgeworfen, ob die Effektivität von Schulen durch mehr Unterricht verbessert werden könne (Wiley & Harnischfeger, 1974). Die theoretischen Annahmen Carrolls wurden dabei als hilfreiche konzeptionelle Grundlage wahrgenommen, ließen im Rahmen der empirischen Umsetzung aber schnell die Notwendigkeit einer Präzisierung erkennen, was Wiley und Harnischfeger dazu veranlasste, ein „verfeinertes“ Modell schulischen Lernens vorzuschlagen (ebd.).

In einer Reihe daran anknüpfender Arbeiten wurden Differenzierungen zwischen verschiedenen schul- und unterrichtsbezogenen Zeitebenen und ihre Bedeutung für die letztendliche Effektivität von Schule und Unterricht diskutiert. Gemäß einer zusammenfassenden Darstellung von Treiber (1982) sind diese unterschiedlichen Zeitebenen zu bezeichnen als (1) nominale (d.h. offiziell angesetzte) Unterrichtszeit, (2)

tatsächliche (d.h. real vollzogene) Unterrichtszeit, (3) nutzbare Instruktionszeit (d.h. effektiv für inhaltliche Arbeit genutzte Unterrichtszeit), (4) die Anwesenheitszeit eines Schülers und schließlich (5) die aktive Lernzeit, d.h. die durch einen Schüler mit Unterrichtsbeteiligung und Aufmerksamkeit verbrachte Unterrichtszeit. Im Rahmen der auf diesen Unterscheidungen aufbauenden Forschungsdiskussion wurden unter anderem empirische Divergenzen der Zeitanteile zwischen nominaler, tatsächlicher und nutzbarer Unterrichtszeit deutlich (Karweit, 1989; Karweit & Slavin, 1981). Dies löste eine vertiefte Auseinandersetzung mit möglichen Zeitverlusten im Unterricht aus, deren Ursachen überwiegend in der Unterrichtsführung und -organisation bzw. der Effektivität des *classroom management* durch die Lehrkraft gesucht wurden (vgl. Anderson, 1984c; Karweit, 1989). Bis heute wird davon ausgehend die Maximierung verfügbarer Lernzeit im Unterricht als Zielhorizont von Klassenmanagement bzw. -führung formuliert (vgl. Helmke, 2007; Ophardt & Thiel, 2008).

Als ein wichtiges Ergebnis der Forschung zur Unterrichtszeit ist die Herausarbeitung der besonderen Rolle der aktiven Lernzeit, auch bezeichnet als *engaged time*, *time on task* oder *active learning time*, festzuhalten (Anderson, 1984c; Myers, 1990; Rosenshine & Berliner, 1978). Dieses Segment der Unterrichtszeit umfasst verschiedene Aspekte von Schüleraktivitäten zur Abbildung der Orientierung an den Inhalten und Aufgaben des Unterrichts, wie etwa Beteiligung, Aufmerksamkeit, aktive Mitarbeit oder auch geistige Auseinandersetzung mit dem Unterrichtsgegenstand¹⁸. Das im internationalen Maßstab der Schul- und Unterrichtsforschung besonders unter dem Namen *time on task* etablierte Konstrukt wurde begrifflich von Bloom (1974) geprägt („*the amount of time the learner is actively engaged in learning – the time on task*“, S. 685), der sich dabei weitgehend auf die theoretischen Überlegungen von Carroll (1963) stützt. Die herausragende empirische Rolle der *time on task* in diesem Themenfeld zeigt sich bei einem Blick auf den Forschungsstand.

In zahlreichen Studien wurden Zusammenhänge zwischen Variationen verschiedener, unterrichtsbezogener Zeitsegmente und dem Lernzuwachs von Schülern untersucht. Forschungssynthesen und Meta-Analysen belegen diesbezüglich wiederholt positive

¹⁸ Für eine Diskussion der verschiedenen Komponenten schülerseitigen Verhaltens und Erlebens siehe Kapitel 2.4.

Zusammenhänge, was Walberg (1986) zu der vorläufigen Schlussfolgerung führt: „*Time counts. No matter how measured, time is consistently associated with learning gains. This association is perhaps the most consistent finding in educational research*“ (S. 452). Diese Deutung des Forschungsstandes wurde auch zu späteren Zeitpunkten und von anderen Autoren aufrecht erhalten (Anderson, 2004; Scheerens & Bosker, 1997; Walberg & Lai, 1999). Allerdings werden in den Forschungsüberblicken die unterschiedlichen Zeitsegmente des Unterrichts nicht immer systematisch getrennt voneinander berücksichtigt (s. Seidel & Shavelson, 2007), sodass die gefundenen Effekte nicht differenziert deutlich werden. Bei einer genaueren Prüfung der empirischen Befundlage im Kontext der zeitbasierten Forschung zur Effektivität von Unterricht wird indessen evident, dass die stärksten Zusammenhänge zu Indikatoren der Schülerlernleistung auf der Grundlage des *time on task* - Konstrukts erzeugt wurden. In einem auf 35 Einzelstudien basierenden Forschungsüberblick zu den Effekten unterschiedlicher, schulbezogener Zeitanteile (Schuljahre, Schultage, Schulstunden, verschiedene Anteile der Unterrichtszeit) auf die Lernerfolge von Schülern kommen Fredrick und Walberg (1980) zu dem Ergebnis, dass die Stärke der gefundenen Zusammenhänge am höchsten ist, wenn (möglichst präzise) Varianten der *time on task* als Zeitindikator verwendet werden und darüber hinaus enge inhaltliche Bezüge zwischen den Inhaltsbereichen der nachträglichen Lernleistungstests und den Inhalten der untersuchten Unterrichtszeit bestehen. Damit wird die zuvor von Rosenshine und Berliner (1978) getroffene Feststellung unterstrichen, *time on task* sei eine durch kein anderes Konstrukt zu ersetzende Größe von essenzieller Bedeutung (S. 12). Auch spätere Forschungsüberblicke bekräftigen die Erkenntnis, dass die deutlichsten Zusammenhänge zu Indikatoren der Schülerlernleistung unter Verwendung der *time on task*-Variable erkennbar sind (Anderson, 1984c; Myers, 1990). Laut Treiber (1982) sind die empirischen Vorteile dieser Größe begründet durch den engen „*Erklärungsabstand des jeweils verknüpften Variablenpaares*“ (S. 21), sprich der Nähe der *time on task* zu den individuellen Lernprozessen der Schüler, was im Zusammenhang der verschiedenen Zeitkonzeptionen als entscheidender forschungsmethodischer Vorteil anzusehen ist.

Im Laufe der weiteren Diskussion wurden jedoch auch die Grenzen des *time on task*-Konstrukts hervorgehoben. In einer Beobachtungsstudie in 18 Grundschulklassen thematisieren Karweit und Slavin (1981) die Herausforderungen der Erfassung verschiedener, lernrelevanter Zeitanteile des Unterrichts und damit

zusammenhängenden Konsequenzen für die Effektstärke auf die gemessene Lernleistung. Im direkten Vergleich führen auch in dieser Studie Indikatoren der schülerseitigen *engaged time* (einem Synonym für *time on task*) zu deutlicheren Effekten auf die Post-Testergebnisse der Lernleistung als die Größen der nominalen und tatsächlichen Unterrichtszeit. Daneben zeigen die Ergebnisse aber auch unterschiedliche Ausprägungen der Zusammenhänge zwischen dieser Größe und der individuellen Lernleistung in Abhängigkeit von der Pre-Testleistung der Schüler. Dies wird als Hinweis dafür gesehen, dass die *engaged time* nicht eine gleichermaßen positive Qualität für alle Schüler haben muss, weil diese sich – unter Bezugnahme auf die Annahmen Carrolls – schließlich in ihrem individuellen Lernzeitbedarf unterscheiden (Karweit & Slavin, 1981, S. 171). Unter Berücksichtigung weiterer Forschungsentwicklungen bekräftigt Karweit in einem späteren Beitrag daran anschließend die Notwendigkeit der Berücksichtigung dieses Verhältnisses („*time needed and time spent are out of harmony*“, Karweit, 1989, S. 91). Als praktische Implikationen für den Unterricht wird daraus abgeleitet, dass nicht etwa eine allgemeine Erhöhung der *engaged time* angestrebt werden müsse, sondern lediglich für bestimmte Schüler in bestimmten Phasen des Unterrichts, sozusagen als individuell angemessene Lernzeit.

Das damit verbundene Bestreben, das Konstrukt der *time on task* weiter zu entwickeln und im Sinne einer „*time-on-the-right-tasks*“ (Berliner, 1990, S. 16) noch näher an die individuellen unterrichtlichen Lernprozesse der Schüler zu rücken, wird an zwei Konzepten besonders deutlich. So betont die bei Karweit (1989, S. 71) zu findende Unterscheidung zwischen *engaged time* und *appropriate engaged time* den Unterschied zwischen jenen Zeitanteilen, in denen die Schüler sich den Lerngelegenheiten des Unterrichts zuwenden, und den Zeitanteilen, in denen den verfügbaren Lerngelegenheiten ein tatsächliches individuelles Lernpotenzial innewohnt, da sie auf einem angemessen inhaltlichen Niveau angesiedelt sind. Dieser Gedanke greift das im Rahmen der *Beginning Teacher Evaluation Study* (Denham & Lieberman, 1980) entstandene Konzept der *academic learning time* auf, welches konzipiert ist als Indikator für Schülerlernprozesse im Unterricht: „*To summarize, time spent by a student engaged on a task that s/he can perform with high success and that is directly relevant to an academic outcome constitutes a measure of student classroom learning*“ (S. 10). Leitender Grundgedanke dafür ist die Frage, ob aufgabenbezogenes Verhalten für jeden Schüler zu jedem Zeitpunkt des Unterrichts dieselbe Wertigkeit für den potenziellen Lerngehalt

aufweist. Die Ergebnisse dieser Studie¹⁹ belegen einen positiven Zusammenhang zwischen *academic learning time* und Lernleistung, was die Autoren schließlich zu der Einschätzung führt, dass diese Variable als geeigneter Maßstab für die Effektivität von Unterricht und damit auch als Instrument für dessen Evaluation gelten könne (S. 30).

Der Ansatz dieser Studie gilt zumindest in der praxisbezogenen Literatur bis heute als sehr einflussreich (vgl. Borich, 2010; Weinstein, Romano & Mignano, 2010), weil damit Aspekte wie Individualisierung, Lernzielorientierung und Unterrichtszeit gleichermaßen anschaulich gemacht werden. In empirischer Hinsicht erweist sich diese Konzeption hingegen rückblickend als weniger ertragreich. Da in der Untersuchung die Angemessenheit von Aufgaben an der beobachteten und eingeschätzten Korrektheit von Schülerantworten gemessen wurde (*success rate*), sind damit mutmaßliche Lerneffekte auch dann diagnostiziert worden, wenn Schüler die Aufgabe schon beherrschen, sich aber trotzdem erfolgreich daran beteiligen. Mögliche Verzerrungen durch solche Unterforderungseffekte werden von den Autoren mit der niedrigen Jahrgangsstufe und den darin üblichen Lernzielen (Konsolidierung und Übung in kleinschrittiger Sequenz) begründet (Denham & Lieberman, 1980, S. 17; dazu auch in kritischer Weise: Berliner, 1990, S. 16), womit der Geltungsbereich der *academic learning time*-Konzeption auf die Elementarstufe beschränkt wird. Weiterhin wurden in der Beobachtungsanalyse nur solche Aufgaben im Unterricht berücksichtigt, deren Inhalte auch in den Pre- und Post-Leistungstests enthalten sind, um die Effekte der inhaltsbezogenen Schülererfolgsrate abschätzen zu können (Romberg, 1980, S. 74). Es erscheint fraglich, inwiefern durch eine solche Beobachtungsstichprobe valide Daten erzeugt und damit die Bedingungen natürlichen Unterrichts adäquat beschrieben werden können. Das Konstrukt der *academic learning time* erscheint somit aus forschungsmethodischer Sicht insgesamt nur eingeschränkt nutzbar als ein Unterrichtseffektivitätsindikator, der einen Mehrwert gegenüber der *time on task* besitzt. Möglicherweise ist dadurch zu begründen, dass dieser Ansatz einen insgesamt eher geringen Einfluss auf die empirische Forschung ausgeübt hat (z.B. Gettinger & Seibert, 2002).

¹⁹ Eine ausführlichere Darstellung dieser Studie und ihrer Ergebnisse ist zu finden in Kap. 2.3.1.

Vergleichbare Grundüberlegungen finden sich auch in einem von Walberg, Niemiec und Frederick (1994) vorgestellten Ansatz wieder: „*Productive time is only a fraction of engaged time since lessons are seldom suitable to all and since individual study is often inappropriate for what is being studied*“ (S. 88). Die so genannte *productive curriculum time* wird beschrieben als ein Anteil der mit Lernaktivitäten verbrachten Unterrichtszeit, in der die Schüler je angemessene Lerngelegenheiten vorfinden, sodass dieser Zeit ein hohes Produktivitätspotenzial hinsichtlich des individuellen Lerngewinns innewohnt. Trotz des erhofften theoretischen Mehrwerts im Zusammenhang zeitbasierter Ansätze zur Erklärung der Effektivität von Unterricht fand dieses Konzept jedoch keine Realisierung im Rahmen der empirischen Forschung und verblieb eher auf der Ebene einer praxisleitenden Denkfigur. Als langfristiger Ertrag dieser Versuche zur Weiterentwicklung des Konstrukts der *time on task* kann somit festgehalten werden, dass dadurch zwar auf theoretischer Ebene der Blick geöffnet wurde für – simplifiziert ausgedrückt – die Qualität der Quantität des Unterrichts, die Empirie dadurch aber nur geringe, über die *time on task* hinausgehende Fortschritte verzeichnen konnte.

Fasst man den Forschungsstand zu Zeitfaktoren des Unterrichts zusammen, ist festzustellen, dass die Befundlage zwar durchgängig positive Effekte des Umfangs verschiedener unterrichtsbezogener Zeitanteile auf die Lernleistung von Schülern belegt, das Ausmaß der gefundenen Effekte aber teilweise stark divergiert. Einige Überblicksarbeiten machen deutlich, dass dies sowohl durch Unterschiede in der Untersuchungsmethodik als auch durch uneinheitliche theoretische Erklärungsansätze bedingt sein kann. Als nützlichstes Ergebnis ist in diesem Kontext die Herausarbeitung der zentralen Rolle der *time on task* als Konstrukt zur zeitlich quantifizierten Erfassung des Schülerbeteiligungsverhaltens im Unterricht zu konstatieren. Indem die Aussagekraft dieser Variable in Bezug auf unterrichtsbedingte Lernerfolge der Schüler in einer Vielzahl von Studien belegt wird, erhalten die in den vorangegangenen Abschnitten dargelegten Erkenntnisse zur Bedeutung des Schülerbeteiligungsverhaltens (s. Kap. 2) an dieser Stelle Bestätigung.

3.3 Die Bedeutung des zeitbasierten Forschungsansatzes für die vorliegende Studie

Mit Blick auf das übergeordnete Bestreben dieser Arbeit, das Zusammenspiel von Prozessmerkmalen des Unterrichts im Hinblick auf ihre Wirksamkeit zu untersuchen, ist die Komponente der zeitlichen Quantifizierung als ein bedeutender Zugewinn anzusehen: Das in diesem Zusammenhang sowohl in theoretischer als auch empirischer Hinsicht herausragende *time on task*-Konstrukt erfordert die Beschreibung konkreter Verhaltensweisen anhand der Häufigkeit, der Dauer und des Zeitpunkts ihres Auftretens innerhalb des Unterrichtsgeschehens, was zu einer Steigerung der empirischen Präzision und damit zu einer geeigneten Erschließung der Mikro-Ebene von Unterrichtsprozessen beiträgt. Das Konstrukt der *time on task* wird aus diesen Gründen als konkreter Operationalisierungsansatz für das Schülerbeteiligungsverhalten aufgegriffen. Die Berücksichtigung der zeitlichen Dimension soll aber gleichermaßen für jene Unterrichtskomponenten Anwendung finden, die als Erklärvariablen für das Schülerbeteiligungsverhalten herangezogen werden – mit Bezug auf das in Kap. 2.3 identifizierte Desiderat repräsentiert durch die Eigenschaften der in den Unterricht implementierten Lerngelegenheiten. Diese Überlegung basiert zum einen darauf, dass die Logik von Zeit als allgemeiner empirischer Abbildungsmetrik konsequenterweise für alle beteiligten Variablen des Unterrichts zu gelten habe (s. Kap. 3.1); zum anderen finden damit jene Erkenntnisse Berücksichtigung, denen zufolge die Aussagekraft von Forschungsergebnissen von einer sinnvollen wechselseitigen Abstimmung der Variablenpaare abhängig ist (s. Kap. 3.2).

4. Aufgaben als Einflussgröße für das Schülerbeteiligungsverhalten

In den bisherigen Ausführungen wurde deutlich, dass das Beteiligungsverhalten von Schülern im Unterricht – wissenschaftlich am präzisesten zu fassen über das Konstrukt der *time on task* – als ein prozessimmanenter Wirksamkeitsindikator für die Durchführung von Unterricht verstanden werden kann (vgl. Kap. 2 und 3). Wie die Forschung zeigt, kommen dafür als bedingende Größen unterschiedliche Faktoren in Betracht. Neben den individuellen Personenhintergründen der Schüler spielen dabei die Prozesseigenschaften des Unterrichts, darunter die Qualität der Lerngelegenheiten, eine entscheidende Rolle: In empirischen Studien erweisen sich ein optimales Herausforderungsniveau bzw. ein angemessener Schwierigkeitsgrad, ein ausgeprägter inhaltlicher Lebensweltbezug und eine allgemeine Interessantheit von Inhalten und Aktivitäten im Unterricht wiederholt als bedeutsame Faktoren für das Schülerbeteiligungsverhalten (vgl. Kap. 2.3.2); in theoretischer Hinsicht knüpfen diese Ergebnisse an die im Unterrichtseffektivitätsmodell von Slavin (1994, s. Kap. 2.2) postulierten Zusammenhänge an, wo diese Aspekte als Angemessenheit (*appropriateness*), Anreizwert (*incentive*) und Qualität (*quality*) des Unterrichts beschrieben werden.

Wie in der Darstellung des Forschungsstandes deutlich wurde, ist jedoch kaum etwas über die diesbezüglichen Zusammenhänge auf der Prozessebene des Unterrichts bekannt. Das folgende Kapitel soll aufzeigen, dass die durch die Lehrkraft in den Unterricht eingebrachten Aufgaben als jene Komponente des Unterrichtsgeschehens anzusehen sind, die eine Untersuchung dieser Merkmale auf einer zum Konstrukt der *time on task* analogen Analysestufe erlaubt. Überdies sollen Wege zur wissenschaftlichen Untersuchung von Aufgaben im Unterrichtsprozess aufgezeigt werden. Dies beinhaltet die Darlegung der theoretischen Rolle von Aufgaben im Kontext von Schulunterricht, die Klärung einiger relevanter Begriffsdifferenzierungen in diesem Feld, inklusive der Einordnung des begrifflichen Verständnisses dieser Arbeit, und schließlich eine Auseinandersetzung mit Ansätzen zur Charakterisierung bzw. Analyse von Aufgaben.

4.1 Die Bedeutung von Aufgaben für den Unterricht

Zum Zusammenhang zwischen Schülerbeteiligungsverhalten und Aufgaben im Unterricht formuliert Anderson (1984a, S. 63) „... *there must be an academic task to which students are expected to attend, become engaged, and ultimately accomplish if time-on-task is to be meaningful.*“ Die Qualität der Unterrichtserfahrung des einzelnen Schülers hängt demzufolge davon ab, dass eine bedeutungsvolle Lerngelegenheit vorliegt. Im Zusammenhang mit Unterricht würden Aufgaben aber oftmals fälschlicherweise gleichgesetzt mit den präsentierten Lernmaterialien (z.B. Abschnitte im Schulbuch) oder den formalen Unterrichtsaktivitäten (z.B. Einzelarbeitsphasen, Tests, Gruppendiskussionen), die für sich genommen im Hinblick auf das Lernpotenzial des Unterrichts bedeutungslos seien und in Anlehnung an Doyle eher die Verknüpfung zu einer *time on activity* denn einer wirklichen *time on task* zuließen (zitiert nach Anderson, 1984a, S. 65).

Eine Auffassung dazu, was unter einer Aufgabe im Unterricht im engeren Sinne zu verstehen ist, findet sich im Rahmen von Carrolls Theorie schulischen Lernens: „*The learner's task of going from ignorance of some specified fact or concept to knowledge or understanding of it, or of proceeding from incapability of performing some specified act to capability of performing it, is a learning task*“ (Carroll, 1963, S. 723). Als zentrales Kennzeichen von als *Lernaufgaben* bezeichneten Aufgaben ist dementsprechend die Absicht (oder zumindest die Möglichkeit) zu sehen, dass der Schüler durch dessen Absolvierung einen Fortschritt erzielt, der sich auf ein bestimmtes Lernziel beziehen lässt, unabhängig davon, ob es um Komponenten des Wissens oder Verstehens geht oder um die Fähigkeit, bestimmte Handlungen zu vollziehen. Nach diesem Verständnis ist eine wesentliche Funktion von Lernaufgaben – im Gegensatz etwa zu „*Arbeitsaufgaben*“ bzw. „*lebenspraktischen Aufgaben*“ (Gagné, 1980, S. 243) – die Initiierung, Steuerung und Organisation von Lernprozessen (Seel, 1981, S. 14).

Daran wird die Abgrenzung zu anderen didaktischen Aufgabenfunktionen deutlich: Nach Kiper (2010) hilft die Lernaufgabe bei der Überwindung einer lernbezogenen „*Differenz von IST und SOLL*“ (S. 45), während diagnostisch relevante Aufgaben die Feststellung des Vorkenntnisstands von Schülern zum Gegenstand haben und Prüfaufgaben auf die Messung des Lernzuwachses abzielen. Lernaufgaben sind somit nicht als Leistungsanlässe zu sehen, sondern als von einer Lehrkraft geplante und

vorgegebene Stationen auf einem individuellen, gegenstandsspezifischen Lernweg (Leisen, 2010). In der neueren, deutschsprachigen Literatur wird deutlich, dass im Zusammenhang mit Unterricht der Aufgabenbegriff mehrheitlich stellvertretend für das verwendet wird, was im engeren Sinne durch Lernaufgaben zu bezeichnen wäre. Der Grund dafür liegt vermutlich darin, dass Leistungs-, Prüf- oder Diagnoseaufgaben nicht zwangsläufig bzw. nur selten Bestandteil des Unterrichtsgeschehens sind. Lernaufgaben hingegen kommt eine konstituierende Rolle für den Unterrichtsprozess zu: Durch die Generierung konkreter Anlässe *„zur Anregung der Aktivitäten der Schülerinnen und Schüler“* (Neubrand, 2002, S. 51) wird durch sie eine Steuerungs- und Organisationsfunktion im Hinblick auf die allgemeine Strukturierung des Unterrichts ausgeübt. Im Unterricht wird diese Funktion durch *„Arbeitsaufträge der Lehrenden an die Lernenden beobachtbar, mit denen eine Lerner selbsttätigkeit in Gang gesetzt und in Gang gehalten werden soll“* (Müller, 2010, S. 84), bzw. als Fragestellungen oder andere Anforderungsformen, durch die die Schüler zur *„Auseinandersetzung mit einem bestimmten Unterrichtsinhalt angeregt werden“* (Blömeke & Müller, 2008, S. 241). So gesehen wirken Aufgaben *„als Orientierungs- und Kristallisationspunkte für die Lehrer-Schüler-Interaktion“*, um die herum sich der Unterrichtsablauf aufbaut (Bromme, Seeger & Steinbring, 1990, S. 2).

Diese Auffassung findet sich auch bei Doyle (1983) wieder, der Aufgaben als *„basic treatment unit in classrooms“* (S. 162) sieht. Doyle schreibt den Aufgaben nicht nur die Funktion eines so genannten Aktionsprogramms zu, sondern auch jenes Geschehensbestandteils des Unterrichts, an dem sich die Lerninhalte festmachen lassen (Doyle, 1986, S. 406). Seiner Ansicht nach geschieht dies in Form von Arbeitsaktivitäten, die erstens auf verfügbaren Anweisungen und vorbereiteten Lernmaterialien basieren, zweitens in Ergebnisse (z.B. Antworten oder schriftliche Beiträge der Schüler) resultieren, und damit letztlich zur Herstellung von Bedeutung im Unterricht dienen (Doyle, 1979). Demnach sind Aufgaben zugleich Strukturgeber für den Unterrichtsprozess und Repräsentant der Lerninhalte. Oder wie Jordan et al. (2006, S. 11) es ausdrücken, definieren sie *„die Grundstruktur potenzieller Lerngelegenheiten“* und stellen auf diese Weise ein *„pädagogisch wirksames Instrument“* (S. 13) für die Lehrkraft dar. Daran wird deutlich, dass Aufgaben einen geeigneten Ansatzpunkt zur Untersuchung jener Faktoren des Unterrichts bieten, die wie zuvor dargelegt sowohl aus

theoretischer (vgl. Kap. 2.2) als auch empirischer (2.3) Sicht als prozessimmanente Einflussfaktoren für das Schülerbeteiligungsverhalten in Frage kommen.

4.2 Das Aufgabenverständnis dieser Arbeit

In Abhängigkeit vom bereichs- bzw. fragespezifischen Kontext bestehen offensichtlich unterschiedliche Auffassungen darüber, was im Unterrichtskontext genau als Aufgabe zu verstehen ist. Ein Teil der Überlegungen in diesem Feld beschäftigt sich mit dem Maßstab der Betrachtung, d.h. mit der Frage, welche Geschehenseinheiten innerhalb des Unterrichtsverlaufs als Aufgaben gesehen werden sollen. So nutzt beispielsweise Renkl (1991) in einer Studie zu Auswirkungen der Aufgaben- und Rückmeldungsgestaltung im Mathematikunterricht auf unterschiedliche Arten der Mathematikleistung von Schüler(inne)n einen vergleichsweise fein auflösenden Aufgabenbegriff, der auch einzelne Aufforderungen und Lehrerfragen im lehrerzentrierten Klassengespräch umfasst (S. 63). Im Kontrast dazu orientieren sich Bohl, Kleinknecht, Batzel und Richey (2012), bei denen der Vergleich von Aufgabenkulturen unterschiedlicher Schulformen mit besonderem Augenmerk auf Merkmale der kognitiven Aktivierung im Vordergrund steht, an einem Aufgabenbegriff, der Fragen und Aufforderungen im Klassenunterricht explizit nicht als Aufgaben zählt (S. 27). Die Unterschiede zwischen enger und weiter gefassten Perspektiven auf Aufgaben werden bei Neubrand (2002, S. 16ff.) am Beispiel einer Textaufgabe im Mathematikunterricht diskutiert. Dessen Bearbeitung im öffentlichen Klassengespräch würde z.B. nach Renkls Auffassung mehrere Aufgaben beinhalten, weil dazu in der Regel eine Abfolge mehrerer Lehrerfragen benötigt wird. Neubrand dagegen sieht die Textaufgabe als eine einzelne Aufgabe, da die Einheit „Aufgabe“ ihrer Auffassung nach jeweils an ein bestimmtes Beispiel für einen mathematischen Sachverhalt geknüpft ist und dies in der Regel durch eine Textaufgabe abgedeckt wird.

Ein etwas weiter gefasstes Aufgabenverständnis findet sich demgegenüber bei Stein, Grover und Henningsen (1996). Ihnen zufolge ist eine Aufgabe eine Unterrichtsaktivität, die mit einer bestimmten fachlichen Idee bzw. einem fachlichen Konzept verknüpft sein muss: „... a mathematical task is defined as a classroom activity, the purpose of which is to focus students' attention on a particular mathematical idea. An activity is not classified as a different or new task unless the underlying mathematical idea toward which the activity

is oriented changes“ (S. 460). Mit dieser Sichtweise besteht eine übliche Mathematikstunde meist aus zwei bis vier Aufgaben. Als ein verhältnismäßig breiter Aufgabenbegriff ist schließlich jener von Doyle (1983) zu nennen. Dieser knüpft das Phänomen „Aufgabe“ (bezeichnet als *academic task*, S. 160) wesentlich an die Aufforderung der Lehrkraft an die Schüler, ein bestimmtes Produkt zu erzeugen, wozu der Vollzug gewisser Tätigkeiten bzw. Lösungswege und die Nutzung vorgegebener Informationen und Ressourcen erforderlich sind. Unter dieser Perspektive können jedoch unterrichtliche Handlungseinheiten sehr unterschiedlichen zeitlichen Ausmaßes als Aufgabe angesehen werden, wie z.B. einzelne Frage-Antwort-Sequenzen, aber auch länger andauernde Phasen wie etwa die Produktion eines Aufsatzes.

Neben solchen Aspekten des Auflösungsgrades bei der Betrachtung des Phänomens „Aufgabe“ betonen andere Begriffe die inhaltliche Qualität von Aufgaben, so zum Beispiel die Unterscheidung zwischen Lernaufgaben und Problemen. Seel (1981) macht die Differenzierung folgendermaßen deutlich: *„Wird der Lernende in einer spezifischen Situation mit einer Lernaufgabe befaßt, die er mit den ihm augenblicklich verfügbaren Informationen oder Lösungsmethoden nicht direkt bewältigen kann, wird die Lernaufgabe als Problem beurteilt, (...)“* (S. 104). Konstitutiv für ein Problem ist demzufolge die *„Inkongruenz von Zielen und verfügbaren Mitteln“* (Baumert et al., o.J., S. 3), sodass mit Problemen in der Regel höhere kognitive Anforderungen für den Lerner verbunden sind als bei Lernaufgaben, die diese Voraussetzung nicht erfüllen. Diese spezifische Konzeption von Aufgaben ist unter dem Begriff der Problemorientierung besonders in naturwissenschaftlichen Fächern verbreitet (Vollrath, 2000). Schabram (2007) weist diesbezüglich darauf hin, dass der Mangel an Mitteln bzw. das Fehlen der zur Lösung notwendigen Vorkenntnisse oder Fähigkeiten individuell variieren kann: Innerhalb einer Lerngruppe verfügen möglicherweise manche Personen über bereits zuvor gemachte Erfahrungen mit vergleichbaren Anforderungssituationen und andere nicht, sodass eine bestimmte Aufgabe unter Umständen nur für einen Teil der Schüler einen tatsächlichen Problemcharakter hat (S. 7). Die Einstufung einer Lernaufgabe als Problem hängt demnach von hinreichenden Kenntnissen über die individuellen Voraussetzungen der Schüler ab. Damit sind gewisse Schwierigkeiten verbunden, zum einen für die Unterrichtspraxis, da der gezielte Einsatz von Problemen als spezielle Art der Lernaufgabe an das nicht immer ausreichend kontrollierbare Verhältnis zwischen Voraussetzungen der Schüler und Anforderungen der Aufgabe geknüpft ist, und zum

anderen für die empirische Forschung, weil für eine valide Einstufung als Aufgabe mit Problemcharakter umfangreiche Hintergrundinformationen über die Schüler vorliegen müssen (ebd.).

Eine im Bereich des Fremdsprachenlernens bedeutsame und damit aufgrund der Stichprobe der vorliegenden Studie²⁰ erwähnenswerte Sichtweise auf Lernaufgaben besteht im Rahmen der *task based instruction* (Skehan, 1998). Nach diesem Ansatz qualifizieren sich Unterrichtsaktivitäten unter anderem dadurch als Aufgaben, dass die Bewältigungsanforderungen in der Lernsituation sich möglichst auf reale Anforderungssituationen übertragen lassen: „... a task, if it is to be worth the name, will make meaning primary, and have a relationship with the real world by giving learners something worthwhile to do under conditions which relate to real life“ (Skehan, 1998, S. 100).

Die dargestellten Ansätze verdeutlichen exemplarisch, dass das konkrete Aufgabenverständnis vom jeweiligen fachlichen Kontext, der didaktischen Auffassung und der wissenschaftlichen Fragestellung abhängt²¹. Mit Blick auf die Ziele der vorliegenden Untersuchung erscheint es sinnvoll, in der empirischen Konzeption der Geschehenseinheit „Aufgabe“ Aspekte der inhaltlichen Qualität, wie sie der problemorientierte bzw. handlungsorientierte Ansatz fordern, außer Acht zu lassen, da es in dieser Studie nicht um die Überprüfung eines spezifischen didaktischen oder lerntheoretischen Ansatzes geht. Vielmehr wird ein Aufgabenbegriff angesetzt, der einen hohen analytischen Auflösungsgrad (d.h. die Betrachtung kleiner Geschehenseinheiten im Unterrichtsverlauf) erlaubt, um die Struktur der Lerngelegenheiten möglichst umfassend abbilden zu können. Daher dient der bei Renkl (1991) entfaltete Aufgabenbegriff als Orientierung. Dieser umfasst „Fragen, Aufforderungen etc., das heißt alle „öffentlichen“ Leistungsanforderungen an Schüler“ (Renkl, 1991, S. 63). Indem Aufgaben mit diagnostischem Hintergrund (z.B. Testaufgaben in Lernkontrollen oder Klassenarbeiten) oder außerhalb des regulären Unterrichtskontextes verortete Aufgaben (z.B. Hausaufgaben) im gegebenen Kontext keine Rolle spielen, steht im weiteren

²⁰ Das zugrundeliegende Videomaterial wurde durchgängig im Englischunterricht aufgezeichnet (s. Kap. 6.2).

²¹ Einen anschaulichen Einblick in das Verhältnis von allgemeindidaktischen und fachdidaktischen Sichtweisen am Beispiel der Analyse konkreter Aufgaben mithilfe eines überfachlichen Kriterienkatalogs bietet der Band von Kleinknecht, Bohl, Maier und Metz (2013).

Verlauf der Begriff „Aufgabe“ – sofern nicht anders gekennzeichnet – stellvertretend für innerhalb des Unterrichtsverlaufs stattfindende Lernaufgaben. Nach der Explikation des Gegenstandsverständnisses besteht der nächste Schritt darin, Wege zur Charakterisierung jener Aufgabeneigenschaften zu suchen, die für die Ziele der vorliegenden Studie relevant sind. Daher werden im folgenden Abschnitt Ansätze zur Beschreibung bzw. Verfahrensweisen zur Analyse von Aufgaben beleuchtet.

4.3 Ansätze der Aufgabenanalyse

Bei der Analyse sowohl von Lern- als auch Diagnoseaufgaben ist eine grundsätzliche Unterscheidung zu treffen zwischen rationaler und empirischer Aufgabenanalyse (Bromme, Seeger & Steinbring, 1990): *„Während die rationale Aufgabenanalyse unter einer pädagogischen oder didaktischen Perspektive die ideale Performanz eines mittleren, idealen Aufgabenlösers (re-)konstruiert, versucht die empirische Aufgabenanalyse herauszufinden, wie Aufgaben tatsächlich gelöst werden oder worden sind“* (S. 6). Diese Unterscheidung kommt besonders zum Tragen bei der Entwicklung und Auswertung von Testaufgaben, wo – vorrangig zu Zwecken der Validierung – eine der Testanwendung vorausgehende Abschätzung der Schwierigkeitsgrade einzelner Aufgaben oder deren Einordnung in spezifische Kompetenzstufenmodelle angestrebt wird (z.B. im Rahmen von PISA: Prenzel, Häußler, Rost & Senkbeil, 2002; DESI: Hartig, 2007; oder dem Gemeinsamen Europäischen Referenzrahmen für Sprachen: Alderson et al., 2006). Da die Qualität der Aufgabenlösungen durch die Schüler nicht Bestandteil der Fragestellungen dieser Studie ist bzw. die *„subjektive Aufgabenwahrnehmung“* (Bromme, Seeger & Steinbring, 1990, S. 7) der Schüler auf Basis der vorliegenden Daten forschungsmethodisch nicht zugänglich gemacht werden kann (vgl. Kap. 6.2), wird hier der Ansatz der rationalen Aufgabenanalyse verfolgt. Das heißt, es werden solche Analyseverfahren fokussiert, die eine Beschreibung von *„objektiven“* Aufgabenmerkmalen erlauben, die *„bereits aus der gestellten Aufgabe heraus zu erfassen sind“* (Neubrand, 2002, S. 93).

Solche Systeme basieren in der Regel auf einem Katalog von Aufgabenmerkmalen, die auf Grundlage der Begutachtung der Aufgabe anhand des jeweils vorliegenden Datenbestands (Aufgabenmaterialien, Videoaufnahmen, etc.) hinsichtlich ihrer Ausprägung eingeschätzt werden. Instrumente zur systematischen Beschreibung von

Aufgabeneigenschaften sind sowohl in verschiedenen fachdidaktischen Bereichen als auch in der Psychologie und der allgemeinen Didaktik erarbeitet worden und werden im Folgenden überblicksartig beschrieben. Auch wenn der Fokus der vorliegenden Studie auf Lern- bzw. Unterrichtsaufgaben liegt, finden dabei ebenfalls Verfahren zur Analyse von Aufgaben mit diagnostischem Hintergrund Berücksichtigung, weil aus diesem Feld wertvolle Vorarbeiten (z.B. zu Aspekten der Aufgabenschwierigkeit) vorliegen, die auch im gegebenen Kontext Berücksichtigung finden können.

4.3.1 Analyseverfahren mit fachdidaktischem Hintergrund

„Aufgrund der inhaltlichen Nähe eines Faches bzw. einer Disziplin zur Entwicklung von Aufgaben wird die Analyse und Diskussion um Aufgaben eher als fachdidaktische Domäne betrachtet“ konstatieren Maier, Bohl, Kleinknecht und Metz (2013, S. 9) in einem Band zur fächerübergreifenden Analyse von Lern- und Leistungsaufgaben. Dementsprechend umfangreich zeigt sich der Bestand an Arbeiten zur Aufgabenanalyse in den verschiedenen Fachdidaktiken. Aus allgemein bildungswissenschaftlicher Sicht haben sich diesbezüglich, u.a. aufgrund der Beteiligung an breit rezipierten Studien wie TIMSS, COACTIV oder DESI, besonders Arbeiten aus dem Bereich der Mathematik (Klieme, Pauli & Reusser, 2006; Jordan et al., 2006; Neubrand, 2002), und der Sprachdidaktik (Alderson et al., 2006; Bachman & Palmer, 2010; Hartig, 2007) als herausragend erwiesen.

Im Rahmen der DESI-Studie (Deutsch Englisch Schülerleistungen International, DESI-Konsortium, 2008) wurde der Unterricht in den beiden genannten Fächern mit dem Ziel in den Blick genommen, die Effekte von Merkmalen der Unterrichtsqualität auf den Kompetenzzuwachs der Schüler zu untersuchen. Die Aufgaben der fachbezogenen Leistungstests waren dabei in mehrfacher Hinsicht Gegenstand systematischer Begutachtungen: Erstens mit dem Ziel der Zuordnung der Aufgaben zu fachspezifischen Kompetenzbereichen im Rahmen der Bildungsstandards, zweitens hinsichtlich ihrer eingeschätzten Relevanz für ein erfolgreiches Absolvieren des mittleren Schulabschlusses (Dubberke & Harks, 2008), und drittens im Rahmen einer Analyse *„anforderungsrelevanter Aufgabenmerkmale“* (Hartig, 2007, S. 88) zur Bestimmung des voraussichtlichen Schwierigkeitsgrades beim Bearbeiten und Lösen der Aufgaben. Diese Merkmale lassen sich in folgende Bereiche zusammenfassen (ebd.): (1) Zur Lösung

notwendige kognitive Operationen (z.B. Grad der Informationsverarbeitung, Verstehensniveau), (2) Kriterien der inhaltlich begründeten Schwierigkeit (z.B. Wortschatz eines Lesetextes), (3) spezifische Phänomene des jeweiligen Leistungsbereichs (z.B. Grammatikformen) und (4) Format der Aufgaben (z.B. geschlossen vs. offen). Die Bereiche umfassen jeweils mehrere, anhand bestimmter Ausprägungsstufen strukturierte Items (z.B. Item „Verstehen: Informationsverarbeitung“: 0 - Erkennen von Informationen, 1 - Inferieren eindeutiger Umschreibungen, 2 - Inferieren impliziter Informationen; Nold & Rossa, S. 204). Die Items wurden innerhalb von Teiluntersuchungen teilweise mehrfach und in modifizierten Versionen verwendet, weshalb von getrennten Kodieranweisungen und möglicherweise auch Beurteilerschulungen ausgegangen werden kann (Schneider, 2007).

Auch im Kontext des in Bezug auf Fremdsprachenerwerb gängigen Kompetenzstufenmodells im europäischen Raum, des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GERS; Europarat, 2001), wurden Aufgaben zum Gegenstand systematischer Beschreibungen. So finden sich innerhalb des GERS Erläuterungen zu Schwierigkeitsfaktoren von Aufgaben im Schulunterricht (S. 157ff.). Dabei wird betont, dass sich die Schwierigkeit von Aufgaben nicht exakt vorhersagen lasse, es aber dennoch wichtig sei, *„dass die Aufgabenparameter variiert werden, um die Aufgabe den Bedürfnissen und Fähigkeiten des Lernenden anzupassen“* (S. 155). Dazu werden relevante Einflüsse und Bedingungen vorgestellt, allerdings nicht in Form eines wissenschaftlichen Analysesystems, sondern als ein Orientierungsrahmen für die Unterrichtspraxis. Berücksichtigt werden dabei unter anderem Merkmale des Textes (z.B. sprachliche Komplexität, Textsorte, Diskursstruktur, Länge, Lebensweltbezug), der Rezeptionshilfen (z.B. Vorwissenaktivierung, Verständnisfragen), der geforderten Reaktion oder Antwort (z.B. non-verbale vs. verbale Leistung, Bearbeitungszeit), aber auch der materiellen Rahmenbedingungen (z.B. Lärm) oder Eigenschaften der Dialogteilnehmer (Kooperationsbereitschaft, Kompetenzen). Das Ergebnis kann als ein umfassender und differenzierter Kriterienkatalog mit Anregungen zur Erstellung möglicher Instrumentarien verstanden werden. Als ein Instrument mit direktem Bezug auf den GERS wurde das so genannte *Dutch Grid* entwickelt, welches Testentwicklern eine Anpassung ihrer Aufgaben an die GERS-Kompetenzmodellierung erleichtern soll (Alderson et al., 2006). Die darin zur Beurteilung herangezogenen Eigenschaften *„stellen*

einerseits schwierigkeitsbestimmende Merkmale dar, andererseits Merkmale, die sich vorwiegend eignen, um die Breite der in den Tests repräsentierten Inhalte oder Kontexte zu dokumentieren“ (Nold, 2007, S. 301). Dazu gehören Eigenschaften (1) des Aufgabentextes (z.B. sprachliche Komplexität, inhaltliche und kontextuelle Ausrichtung), (2) der Bearbeitung (z.B. kognitive Operationen, Zeitlimit), (3) des Antwortformats (z.B. Grad der geforderten Eigenkonstruktion) und (4) der fachlichen Phänomene (z.B. textlicher Diskurstyp). Aus der Summe der Merkmalseinschätzungen und unter Abgleich mit Deskriptoren der GERS-Kompetenzstufen soll eine Einschätzung des GERS-Niveaus der jeweiligen Aufgabe abgeleitet werden können (Alderson et al., 2006). Das Erreichen dieser Anwendungsziele wird allerdings aufgrund fraglicher Validität kritisch gesehen, sowohl aus externer Sicht (Leucht, Harsch, Pant & Köller, 2012), als auch von den Autoren selbst (Alderson et al., 2004).

Als letztes Beispiel der Aufgabenbeschreibung aus dem Feld der Fremdsprachendidaktik ist jenes der bereits unter Kap. 4.2 erwähnten *task-based instruction* zu nennen. Nach diesem Ansatz besteht die Rolle von Aufgaben im Unterricht darin, einen starken Bezug zu tatsächlichen Sprachanforderungen in konkreten Lebenssituationen herzustellen, wobei sich dies in einem Realitätsbezug sowohl der Inhalte als auch den Bearbeitungsbedingungen der Aufgabe selbst niederschlagen sollte (vgl. Bachman & Palmer, 1996; Skehan, 1998). Die Erfüllung dieser Ansprüche bei der Gestaltung des Aufgabenangebots erfordert folglich einen genauen und differenzierten Blick auf deren Charakteristika. Zusammenstellungen entsprechender Beschreibungskriterien, wie beispielsweise jene von Skehan (1998) oder Bachman und Palmer (1996, 2010), sind dabei weniger zur Anwendung innerhalb wissenschaftlicher Studien ausgelegt, sondern vielmehr als Planungshilfe für die Aufgabengestaltung in fremdsprachlichen Lernsettings. Hervorzuheben wäre diesbezüglich das „*Framework of Language Task Characteristics*“ von Bachman und Palmer (1996, 2010). Dieses System soll es ermöglichen, Testaufgaben (*language test tasks*) und reale Sprachanforderungen (*language use task*) analog klassifizieren und damit die Aufgabeneigenschaften in der Testsituation durch bewusste Variation an die Testnehmer anpassen zu können (Bachman, 1996, S. 45ff.). Dabei werden – neben pragmatisch-physischen Rahmenbedingungen – Charakteristika (1) des Tests insgesamt (z.B. Aufbau, Qualität der Instruktionen, Bewertungskriterien), (2) des Aufgabentextes (Art und Weise der Präsentation, sprachliche Charakteristiken, Lebensweltbezüge), (3) des geforderten

Ergebnisses (Merkmale analog zu jenen des Aufgabentextes) und (4) der Bearbeitung (Interaktionserfordernisse, thematische Offenheit, Herstellung externer Bezüge) abgedeckt. Vorteil dieses Systems ist die umfassende Berücksichtigung zahlreicher Aspekte, auch hinsichtlich des Einbringens der Aufgabe in die Unterrichtssituation, wie z.B. die Frage, über welche Kanäle (visuell, auditiv, sprachbasiert, grafisch), in welchem Umfang (Länge der Ausführungen) und mit welcher Qualität (Klarheit der Ausführungen, unterstützende Beispiele) die Aufgabe präsentiert wird.

Eine der bekanntesten und umfassendsten Auseinandersetzungen mit der Analyse von Aufgaben im Unterricht findet sich bei Neubrand (2002). Diese Untersuchung zielt ab auf einen Vergleich von Aufgaben in Schülerarbeitsphasen des Mathematikunterrichts in Japan, den USA und der BRD und greift dabei auf Daten aus der TIMS-Studie zurück. Datengrundlage für die Untersuchung sind so genannte Stundenablaufpläne, d.h. anhand von Videoaufnahmen aufbereitete, detaillierte Verlaufsbeschreibungen der Unterrichtsstunden, die im Rahmen der TIMS-Videostudie angefertigt wurden (S. 199ff.; vgl. auch Stigler, Gonzales, Kawanaka, Knoll & Serrano, 1999). Nur in Einzelfällen wurde diese Datenbasis durch Eigenbeobachtungen am Videomaterial ergänzt. Bei der darauf basierenden Klassifikation der Aufgaben (n = 1.143, Neubrand, 2002, S. 208) wurden neben mathematikdidaktischen Spezifika, die unter dem Begriff des Aufgabenkerns²² (z.B. Art des mathematischen Wissens, strukturelle fachliche Tiefe, zu vollziehende mathematische Tätigkeiten) und den Wissensseinheiten²³ gefasst sind, auch allgemeine Aufgabenmerkmale in den Blick genommen. Dazu gehören (1) die Aufgabenperipherie mit Aspekten der Aufgabenpräsentation (z.B. Quelle der Aufgabe, eingesetzte Medien und Visualisierungen, einbringende Person) und der Art der Anweisungen, (2) die Aufgabenbearbeitung und Implementierung (z.B. Sozialformen, Anzahl der Lösungswege oder technische Hilfsmittel, kognitiver Anspruch im Löseprozess) und (3) Aspekte der Aufgabenvernetzung, d.h. Maßnahmen der inhaltlichen oder unterrichtsmethodischen Verknüpfung zwischen Aufgaben im Unterrichtsverlauf. Das System zählt sicherlich zu den umfassendsten, differenziertesten und zugleich am

²² „Als *Aufgabenkern* werden diejenigen Aspekte einer Aufgabe zusammengefasst, die die Grundorientierung des jeweiligen Mathematikunterrichts reflektieren.“ (Neubrand 2002, S. 94)

²³ „Die von einem Experten in Hinblick auf die Anforderung der jeweiligen Aufgabe aktivierten Wissensbestandteile werden hier als (die zur Lösung der Aufgabe notwendigen) *Wissenseinheiten* bezeichnet.“ (Neubrand 2002, S. 95)

besten dokumentierten Instrumentarien zur Aufgabenanalyse in der Unterrichtsforschung, wenngleich durch die enge Anbindung an den spezifischen fachlichen Kontext die Übertragbarkeit der Analysekatoren als eingeschränkt gesehen werden muss.

In ähnlicher Weise stellt sich das im Rahmen des Forschungsprogramms *COACTIV* verwendete Klassifikationsschema dar (Jordan et al., 2006; Jordan et al., 2008; Neubrand et al., 2011), welches in vielen Punkten an das Instrument von Neubrand (2002) anknüpft. Unterschiede bestehen zum einen in konzeptioneller Hinsicht, indem das System den Analysefokus auf den Bereich der kognitiven Aktivierung von Aufgaben unter Aussparung von „komplexen Effekten ihrer Implementierung in den Unterricht“ legt (Jordan et al. 2006, S. 14). Die enthaltenen Analysebereiche beziehen sich dementsprechend vornehmlich auf Aspekte des Verstehens und Bearbeitens fachlicher Phänomene, wie z.B. den mathematischen Stoffbereich, die Art der Lösungsprozesse bzw. Typen mathematischen Arbeitens oder mathematische Grundvorstellungen (vgl. Neubrand et al., 2011). Zum anderen bezieht sich das Instrument auf eine andere Datengrundlage: Gegenstand der Analyse waren von Lehrkräften dokumentierte Aufgaben aus Klassenarbeiten, Hausaufgaben und Einstiegsaufgaben im Unterricht (Jordan et al., 2008). Die Aufgabenstichprobe für das Analysesystem beinhaltet somit keine durch Unterrichtsbeobachtungen, sondern allein materialgebundene Aufgabendaten, wodurch eine außerordentlich große Stichprobe (47.573 Aufgaben zu zwei Messzeitpunkten) in den Blick genommen werden konnte. Das System kann damit als eines der empirisch am besten erprobten Aufgabenanalyseinstrumente gelten.

4.3.2 Analyseverfahren mit lernpsychologischem Hintergrund

In der pädagogischen Psychologie ist die Beschreibung von Aufgaben im Kontext von schulischem Lernen fest verbunden mit der Definition von Lernzielen. Verdeutlichen lässt sich dies an der psychologisch begründeten Didaktik von Gagné (1980). Nach Gagné stellt die Bestimmung der angestrebten Lernergebnisse den ersten Schritt der Unterrichtsplanung dar, um daraufhin geeignete Bedingungen herzustellen, „die das erwartete Ergebnis garantieren“ (S. 242). Verknüpft ist dies mit einer Analyse der Lernaufgaben, die den Weg dorthin definieren: „Unter dem Aspekt des Lernens stellt eine Aufgabenbeschreibung eine Beschreibung der „angezielten“ Lernergebnisse dar; das heißt,

sie beschreibt die Leistungen, die ein Lernender erreichen soll, wenn er das Lernprogramm durchlaufen hat“ (S. 248). Diese Auffassung schließt an die Grundidee der Lernzieltaxonomie von Bloom, Engelhart, Furst, Hill und Krathwohl (1956, zitiert nach Anderson et al., 2001) an, dem wohl bekanntesten Konzept zur Beschreibung von Lernzielen und den damit verbundenen kognitiven Anforderungen von Aufgaben. Die Taxonomie beinhaltet sechs hierarchisch angeordnete Klassen von Lernzielen: (1) Wissen, (2) Verstehen, (3) Anwendung, (4) Analyse, (5) Synthese und (6) Bewertung. Die hierarchische Struktur impliziert, dass „Bewertung“ mit der höchsten, „Wissen“ hingegen mit der niedrigsten kognitiven Komplexität verbunden ist. Diese Anordnung wurde in einer Überarbeitung durch Anderson et al. (2001) in eine zweidimensionale und weiter ausdifferenzierte Struktur überführt (s.u.). Hauptzweck der Taxonomie ist es, Lehrkräften ein Instrument an die Hand zu geben, um Lernzielsetzung, Unterrichtsgestaltung und Leistungsmessung leichter in Analogie bringen zu können (*concept of alignment*, S. 10). Dazu wird das jeweilige Lernziel, die Lern- oder Prüfungsaufgabe in ein zweidimensionales, hierarchisch angelegtes Klassifikationssystem eingeordnet, welches die Art und das Niveau der zur Bewältigung erforderlichen kognitiven Anforderungen für die lernende Person abbildet. Die beiden Dimensionen und ihre Hierarchiestufen schlüsseln sich folgendermaßen auf:

I. Wissen

- A. Fakten (z.B. Begrifflichkeiten, gegenstandsspezifische Details)
- B. Konzepte (z.B. Kategorien, Prinzipien, Theorien)
- C. Prozeduren (z.B. Algorithmen, Techniken, Methoden)
- D. Metakognition (z.B. Strategien, Selbstwissen)

II. Kognitive Prozesse

- 1. Erinnern (z.B. Wiedererkennen, Abrufen von Informationen)
- 2. Verstehen (z.B. Deuten, Klassifizieren, Folgern, Zusammenfassen, Vergleichen)
- 3. Anwenden (z.B. Vollziehen, Ausführen)
- 4. Analysieren (z.B. Unterscheiden, Ordnen)
- 5. Evaluieren (z.B. Prüfen, Kritisieren)
- 6. Gestalten (z.B. Entwerfen, Planen, Herstellen)

Durch die hohe Anzahl an Unterkategorien bzw. „Sub-Typen“ (11 bei der Wissensdimension, 19 bei der kognitiven Prozessdimension, s. Beispiele in Klammern)

sind bei der Zuordnung relativ feine Differenzierungen möglich. Bestimmte Kombinationen aus kognitiven Prozess- und Wissenskategorien (erinnern von Fakten, verstehen von Konzepten und anwenden von Prozeduren) dürfen zwar als besonders häufig angenommen werden (ebd., S. 107), doch zeigt zweidimensionale Struktur der Taxonomie, dass diese Verknüpfungen nicht als zwingend anzusehen sind und prinzipiell die Kombination aller Stufen beider Ebenen möglich erscheint.

Die Taxonomie ist per se nicht als wissenschaftliches Instrument angelegt. Dennoch wird sie in der Forschung – mitunter in Form der ursprünglichen Taxonomie von Bloom – häufig verwendet, wenn auch meist als Vorlage bzw. Impulsgeber für die Konstruktion eigener, teils fachspezifischer und komplexitätsreduzierter Analysekategorien (vgl. Blömeke, Risse, Müller, Eichler & Schulz, 2006; Maier, Kleinknecht, Metz, & Bohl, 2010; Niegemann & Stadler, 2001; Renkl, 1991; Schabram, 2007). Wie die Darstellung des Forschungsstands zu den beiden Lernzieltaxonomien bei Schabram (2007, S. 52ff.) zeigt, steht die wissenschaftliche Verwendung der Systeme in der Kritik: Besonders die postulierte Hierarchielogik, d.h. die Annahme der aufsteigenden Schwierigkeit der Niveaustufen und die Faktorenstruktur, bspw. hinsichtlich der mangelnden Trennschärfe der beiden höchsten kognitiven Prozessniveaus, werden bezweifelt. In der Summe deutet der Forschungsüberblick aber darauf hin, dass zumindest die Struktur der ersten vier Kategorienstufen kognitiver Prozesse als plausibel angesehen werden kann (S. 54), und abgesehen von dieser laufenden Diskussion um die empirische Qualität des Systems werde es doch überwiegend als sinnvolles und nützliches Instrument anerkannt.

4.3.3 Analyseverfahren mit allgemeindidaktischem Hintergrund

Auch aus Sicht der allgemeinen Didaktik werden Aufgaben zum Gegenstand systematischer Beschreibungen. So legen Blömeke, Risse, Müller, Eichler und Schulz (2006) aus der Absicht heraus, den in dieser Teildisziplin bisher „*eher vernachlässigten Forschungsbereich*“ (S. 331) zu stärken, ein Analysemodell der Aufgabenqualität vor. Das lerntheoretisch, motivationspsychologisch und allgemeindidaktisch begründete Modell setzt sich aus neun Kriterien zusammen, die Anforderungen an „*lernprozessanregende Aufgaben*“ (S. 330) repräsentieren. Zusammengefasst sind dies: (1) Gesellschaftliche Relevanz des Aufgabeninhalts, (2) Ansprache von Schülerbedürfnissen, (3) kognitive

Aufgabenqualität, (4) Neuigkeitswert von Unterrichtsmethoden und Inhalten, (5) Verständlichkeit und realistische individuelle Bewältigungschancen, (6) Potenzial zur inneren Differenzierung, (7) Komplexität und Reichweite der inhaltlichen Situierung, (8) Problemlösepotenzial bzw. Offenheit der Aufgabenstellung und (9) Einbezug sozialer Interaktionen im Bearbeitungsprozess. Die Ausprägung dieser Merkmale wird in drei getrennten, auf unterschiedlichen Informationsbeständen zur jeweiligen Aufgabe basierenden Analyseschritten beurteilt: (1) In der Potenzialanalyse wird der Frage nachgegangen, welches didaktische und fachliche Potenzial der Aufgabe objektiv innewohnt. Beurteilungsgrundlage dafür sind schriftliche Fixierungen zur Aufgabe, wie z.B. Arbeitsblätter oder Tafelanschriften. (2) Die Analyse intendierter Anforderungen stellt dar, welche Vorstellungen die Lehrkraft von Unterricht generell und der Lerngruppe im Speziellen hat. Die Beurteilung dessen findet auf Basis von Interviewaussagen der jeweiligen Lehrkraft statt. (3) Abschließend steht die Analyse der realisierten Aufgabenqualität. Damit wird geprüft, inwieweit das objektive Potenzial und die Intentionen der Lehrkraft in der Unterrichtssituation realisiert wurden. Dieser Schritt wird differenziert nach den Aufgabenphasen Aufgabenstellung, Schülerarbeitsphase und Auswertung bzw. Reflexion. Datengrundlage bilden Videoaufnahmen des Unterrichts. Das System ist zur wissenschaftlichen Verwendung angelegt, wurde aber – so weit bekannt – nicht im größeren Rahmen empirisch erprobt. Wie eine exemplarische Fallstudie der Autoren zeigt, orientiert sich die forschungsmethodische Realisierung der Analyse an qualitativen Auswertungsverfahren. Eine Anwendung auf größere Datenbestände, etwa mittels kategorienbasierter Auswertung, wäre daher an umfangreiche Modifikationen des Systems geknüpft, was allerdings durch die geringe Operationalisierungstiefe der Analysekriterien erschwert würde. Zudem erfordert das Analysemodell die Verfügbarkeit unterschiedlicher Datenbestände (Aufgabendokumente, Interviews mit Lehrkräften, Videoaufnahmen), was mit einem hohen forschungstechnischen Aufwand verbunden ist.

Auch Maier, Kleinknecht, Metz und Bohl (2010) sehen den Bedarf eines allgemeindidaktischen Systems zur Aufgabenanalyse und setzen dabei den Fokus auf das kognitive Potenzial. Übergeordnetes Ziel ist es dabei, Kriterien zu entwickeln, „mit denen sich auch fachübergreifend Aufgaben hinsichtlich ihrer kognitiven Komplexität zuverlässig einschätzen lassen“ (Kleinknecht, Maier, Metz, & Bohl, 2011, S. 330). Unter

Rückgriff auf die fachdidaktisch orientierten Arbeiten von Neubrand (2002), Jordan et al. (2006) und die Taxonomie von Anderson et al. (2001) schlagen sie ein Kategoriensystem aus sieben Faktoren vor, jeweils repräsentiert durch drei bis vier Ausprägungsstufen: (1) Angesprochene Wissensarten (z.B. Fakten, Prozeduren), (2) kognitive Prozesse (z.B. Reproduktion, Transfer), (3) Anzahl der Wissenseinheiten (eine, bis vier, mehr als vier), (4) Grad der Offenheit der Aufgabenstellung (z.B. konvergent, divergent), (5) Lebensweltbezug (z.B. konstruiert, real), (6) sprachlogische Komplexität der Aufgabeninformationen (niedrig, mittel, hoch) und (7) wissensbezogene Repräsentationsformen (eine, Integration, Transformation). Mit diesem Aufbau kommen die Autoren ihrem Anspruch nach, bedeutsame Aufgabenfaktoren „möglichst sparsam“ (Maier, Kleinknecht, Metz & Bohl, 2010, S. 85) zu erfassen. Die Anwendungsziele des Instruments liegen hauptsächlich im Kontext der Lehrerbildung (z.B. zur Förderung diagnostischer Kompetenz von Lehramtsstudierenden oder Lehrkräften in Fortbildung) und der Unterrichts- und Schulpraxis, dabei z.B. als Hilfe zur Konstruktion und Modifikation von Aufgaben oder der Vor- und Nachbereitung von Unterricht (vgl. Maier, Bohl, Kleinknecht & Metz, 2013). Das Kategoriensystem wurde vielfach erprobt, sowohl in empirischen Erprobungsstudien (Kleinknecht, Maier, Metz, & Bohl, 2011) als auch in einer Reihe von unterrichtsfachspezifischen Fallstudien (Maier, Bohl, Kleinknecht & Metz, 2013). Außerdem ist es umfangreich dokumentiert: Neben den zahlreichen Publikationen gibt es eine ausführliche Internetseite mit Beispielanalysen und Downloads²⁴. Die Ergebnisse der Erprobungen deuten darauf hin, dass das System insgesamt gut anwendbar ist und zu reliablen Ergebnissen führt, auch wenn aufgrund seines Querschnittscharakters gewisse Limitierungen in der Aussagekraft für unterschiedliche fachdidaktische Kontexte bestehen (Maier, Bohl, Kleinknecht & Metz, 2013).

4.4 Zusammenfassung und Schlussfolgerungen für diese Studie

Die vorangegangenen Ausführungen haben die Bedeutung von Aufgaben für die Strukturierung und Lernwirksamkeit von Unterricht aufgezeigt und verschiedene Ansätze zu ihrer Beschreibung vorgestellt. Es wurde deutlich, dass sich in Aufgaben –

²⁴ www.erziehungswissenschaft.uni-tuebingen.de/abteilungen/schulpaedagogik/aufgabenanalyse.html, zuletzt aufgerufen am 27.11.2014

verstanden als Lernaufgaben innerhalb des Unterrichtsprozesses – Merkmale der Lerninhalte verbinden mit der Art und Weise ihrer Repräsentation im Unterricht (s. Kap. 4.1). Damit bieten sie einen geeigneten Ansatzpunkt für die Ziele dieser Studie: Aufgaben sind Träger einer Reihe von Prozessmerkmalen, die aus theoretischer Sicht als Bedingung für das Schülerbeteiligungsverhalten in Frage kommen. In Anknüpfung an das Modell effektiven Unterrichts nach Slavin (1994) sind diese zu sehen in interessanten und zugleich klaren und verständlichen Darstellungen der Unterrichtsinhalte und -aktivitäten, der Herstellung von Bezügen zur Lebenswelt der Schüler und der Sicherstellung eines optimalen Schwierigkeitsniveaus in den inhaltlichen Anforderungen. Wie gezeigt, bestätigt sich der Einfluss dieser Aspekte auch in empirischen Forschungsergebnissen (s. Kap. 2.3.2), allerdings nehmen die entsprechenden Studien eine hauptsächlich summative Perspektive auf Unterrichtsstunden, -reihen oder ganze Schul(halb)jahre ein, sodass die Prozessebene des Unterrichts in diesen Fragen bislang weitgehend unbeleuchtet bleibt. Aufgaben liefern hingegen die Möglichkeit, diese Prozessmerkmale anhand konkreter Ereignisse im Geschehensverlauf zu verorten und sie dadurch mit dem selben analytischen Auflösungsgrad zu erschließen wie die hier als zentral erachtete Bezugsgröße, die *time on task* als Operationalisierung für das Schülerbeteiligungsverhalten.

Der Betrachtungsmaßstab der vorliegenden Arbeit impliziert daher ein Aufgabenverständnis, welches auch kleine Geschehensbestandteile (wie z.B. Fragen) als Aufgaben ansieht und die Legitimierung einer Geschehenseinheit als Aufgabe nicht an bestimmte qualitative Voraussetzungen knüpft, wie z.B. den Grad des Problemlösens oder der Übertragbarkeit auf reale Anforderungssituationen (s. Kap. 4.2). Im letzten Unterkapitel wurden schließlich verschiedene, als besonders relevant zu erachtende Ansätze zur systematischen Beschreibung von Aufgaben aus den Fachdidaktiken, der pädagogischen Psychologie und der allgemeinen Didaktik dargestellt, wobei der Schwerpunkt auf Verfahrensweisen der rationalen bzw. objektiven Aufgabenanalyse lag (Bromme, Seeger & Steinbring, 1990; Neubrand, 2002, s. Kap. 4.3). Nicht alle dieser Instrumente sind auf eine beobachtungsbasierte Analyse von Aufgaben im Unterricht ausgelegt (wie z.B. jene von Neubrand, 2002, oder Blömeke, Risse, Müller, Eichler & Schulz, 2006), doch ermöglichen viele der Systeme grundsätzlich die Übertragung auf einen solchen Untersuchungsansatz bzw. schließen ihn nicht explizit aus (z.B. Anderson et al., 2001; Maier, Kleinknecht, Metz & Bohl, 2010; Bachman & Palmer, 1996, 2010).

Es wurde auch deutlich, dass sich die Verfahren trotz unterschiedlicher Analyseschwerpunkte und disziplinärer Hintergründe in der Berücksichtigung mancher Aufgabencharakteristika überschneiden: So ist die Art der Aufgabenpräsentation im Unterricht in ähnlicher Weise sowohl in einem System der Fremdsprach- (*Input-Format/Instructions* bei Bachman & Palmer, 1996, 2010) als auch der Mathematikdidaktik (Aspekte der Aufgabenperipherie bei Neubrand, 2002) enthalten. Einige Facetten finden sich zudem in der überwiegenden Mehrzahl der Instrumente wieder, wie etwa der Lebenswelt- und Alltagsbezug der Aufgabeninhalte, das kognitive Anspruchsniveau oder die Offenheit der geforderten Lösungen bzw. Antworten. Einerseits spricht dies für den Ansatz überfachlich ausgerichteter, allgemeindidaktischer Analysesysteme. Andererseits soll dabei nicht außer Acht gelassen werden, dass fachspezifisch ausgerichtete Konzepte durch die größere Nähe zu den Inhalten das Potenzial zu tieferen und differenzierteren Analysen innewohnt. Mit Blick auf die vorliegende Studie lässt sich daraus ableiten, dass für die hier angestrebte Untersuchung von Aufgaben eine Mischkonstruktion mit Elementen aus verschiedenen Instrumentarien sinnvoll erscheint, die unter Abwägung der Möglichkeiten zur Übertragbarkeit der jeweiligen Analysekomponente auf den gewählten empirischen Zugang und den Anschluss an die theoretischen Fragen vorzunehmen ist.

5. Fragestellungen und Ziele der Studie

In diesem Kapitel werden zunächst die zentralen Erkenntnisse der vorangegangenen Ausführungen zusammengefasst (5.1). Daraus werden die übergeordneten Ziele und der methodische Ansatz der Untersuchung abgeleitet (5.2) und abschließend die konkreten Forschungsfragen für die Untersuchung formuliert (5.3).

5.1 Zusammenfassung der Grundüberlegungen und des Desiderats

Ausgangspunkt dieser Studie ist die Frage nach Zusammenhängen zwischen Prozesskomponenten des Unterrichts im Hinblick auf dessen Wirksamkeit (vgl. Kap. 1). Im Anschluss an effektivitätsorientierte Forschungszugänge zum Feld und ausgehend von einem praxisorientierten Unterrichtsverständnis wird der Grundannahme gefolgt, dass durch die Teilnahme an regulärem Schulunterricht bei den Schülern Lern- und Bildungsprozesse angestoßen und gefördert werden können, indem die Lehrkraft Lerngelegenheiten plant, vorbereitet und in den Unterricht einbringt (vgl. Kap. 2.1). Um aber die intendierten Wirkungen erzielen zu können, muss sichergestellt werden, dass der Adressatenkreis innerhalb dieses Settings auch erreicht wird, d.h. die angezielten und in den Prozess eingebrachten Lern- und Bildungsinhalte für die Schüler zum Gegenstand der Auseinandersetzung werden. Daher wird in dieser Arbeit in Anknüpfung an klassische Modelle effektiven Unterrichts (Carroll, 1963; Slavin; 1994; Kap. 2.2) und unter Berücksichtigung eines breiten Bestands an Forschungsarbeiten aus den beiden Feldern der Unterrichtseffektivitätsforschung und der Forschung zum *school engagement* (Kap. 2.3) das Schülerbeteiligungsverhalten in Form des Konstrukts der *time on task* (Kap. 3) als prozessimmanenter Indikator für wirksamen Unterricht angesetzt.

Als zentrales Desiderat ist diesbezüglich festzustellen, dass trotz des Vorliegens entsprechender theoretischer Annahmen und einer Reihe empirischer Hinweise auf der Ebene des Unterrichtsprozesses noch kaum etwas über die Zusammenhänge zwischen der Qualität der Aufgaben und dem Beteiligungsverhalten der Schüler bekannt ist (Kap. 2.2/2.3). Wie gezeigt bewegen sich bisher dazu vorliegende und auf entsprechende Zusammenhänge hindeutende Forschungsergebnisse bislang überwiegend auf der Ebene summativer Beschreibungen ganzer Unterrichtsstunden oder -reihen oder

sprechen eher global gefasste Unterrichtsfacetten an, die die Eigenschaften des Aufgabenangebots lediglich „mit“ abdecken, aber nicht zentral in den Blick nehmen. Beobachtungsstudien, die Einblicke in die Zusammenhänge auf Prozessebene gewähren, liegen hingegen nicht vor. Dies verwundert auch insofern, dass sich Aufgaben als zentrales inhalts- und strukturgebendes Element des Unterrichts zuverlässig an konkrete Anlässe innerhalb des Unterrichtsverlaufs knüpfen lassen und damit die Möglichkeit bieten, auf einem zur Bezugsgröße *time on task* äquivalenten Spezifikationsgrad abgebildet und damit für entsprechende Analysen verfügbar gemacht werden zu können (Kap. 4.1/4.2). An dieser Forschungslücke setzt die vorliegende Studie an.

5.2 Ziele der Studie und Einordnung des Untersuchungsansatzes

Auf der Grundlage der dargestellten Überlegungen verfolgt die vorliegende Studie das Ziel, Aussagen über Wirksamkeitszusammenhänge im Unterrichtsprozess zu treffen, indem versucht wird, das Schülerbeteiligungsverhalten durch die Eigenschaften der in den Unterricht eingebrachten Aufgaben zu erklären. Grundgedanke dafür ist die Annahme, dass der Grad der Hinwendung der Schüler zu den verfügbaren Lerngelegenheiten mit der Beschaffenheit derselben zusammenhängen sollte.

Im Anschluss an den zentralen theoretischen Bezugsrahmen dieser Arbeit (s. Kap. 2.2) wird das Schülerbeteiligungsverhalten als Mediatorgröße zwischen Prozessmerkmalen des Unterrichts, Personeneigenschaften der Schüler und außerhalb des Prozesses liegenden Wirkungen, d.h. möglichen Lern- und Bildungserträgen, gesehen. Diese letztendlich erhofften und angezielten Unterrichtserträge sind allerdings nicht Gegenstand dieser Untersuchung, denn an dieser Stelle steht die Suche nach einem prozessinternen „Produkt“ im Vordergrund. Dennoch ist die vorliegende Studie damit gemäß Doyle (1977) nicht dem Prozess-Produkt-Paradigma der Unterrichtsforschung zuzuordnen, sondern dem *mediating process*-Paradigma, da die übergeordneten Wirkannahmen die Funktion einer zentralen Mediatorgröße voraussetzen und hier ein Ausschnitt daraus betrachtet wird. Das Vorhaben ist damit im engeren Sinne als Prozess-Prozess-Studie zu verstehen (Doyle, 1977, S. 173).

Für den hier fokussierten Bereich der unterrichtsinternen Wirkzusammenhänge sind dabei jene Verbindungen von besonderer Bedeutung, die nach Slavin (1994) zwischen

den Komponenten Q-A-I (*quality of instruction, appropriate levels of instruction, incentives*) und dem auf zeitlicher Basis beschriebenen Schülerbeteiligungsverhalten, der *time on task*, bestehen. Kristallisationspunkt dieser Komponenten unterrichtlicher Prozessqualität ist der Bereich der *incentives*, der so genannte Anreizwert des Unterrichts, in dem sich die Merkmale der *quality of instruction*, d.h. der Unterrichtsqualität, und der *appropriateness*, d.h. der Angemessenheit des Lernangebots, niederschlagen und auf das individuelle Schülerbeteiligungsverhalten im Unterricht wirken. Zentrale Bestandteile der dieser Prozessqualitätskomponenten lassen sich am Aufgabenangebot festmachen (z.B. klare, abwechslungsreiche und interessante Darstellung und angemessener Schwierigkeitsgrad der Lerninhalte, Ermöglichung subjektiv wahrgenommener Sinnhaftigkeit der Inhalte durch die Herstellung von Bezügen zur Lebenswelt der Schüler, s. Kap. 4). Manche dieser Aspekte implizieren ein Verhältnis zu den als *student inputs* (ebd.) bezeichneten persönlichen Voraussetzungen der Schüler, wie etwa das inhaltliche Anforderungsniveau zum inhaltspezifischen Leistungsvermögen oder der Lebensweltbezug der Lerninhalte zum jeweiligen alltäglichen Erfahrungsraum des Schülers. Aus diesem Grund soll in dieser Arbeit zusätzlich untersucht werden, ob bei einer Differenzierung nach Personenmerkmalen in der Schülergruppe unterschiedliche Effekte im Unterrichtsprozess feststellbar sind.

Um die fokussierten Merkmalsbereiche in einer für diese Absichten geeigneten Präzision abbilden zu können (z.B. ein zeitlich umfassendes Bild des Beteiligungsverhaltens der Schüler im Klassenraum), ist eine Mikroperspektive auf den Unterrichtsverlauf und seine Geschehenbestandteile sowie das Handeln aller Beteiligten erforderlich. Als geeigneter forschungsmethodischer Ansatz für diese Zwecke sind videobasierte Beobachtungen zu sehen (ausführlicher zu dieser Methodik: Kap. 6.1). Die bisherigen Ausführungen haben gezeigt, dass weder im Bereich der Erfassung und Beschreibung von Schülerbeteiligungsverhalten (vgl. Kap. 2.3/2.4) noch von Aufgaben im Unterricht (Kap. 4.3) Untersuchungsinstrumente vorliegen, die eine direkte Übertragung auf diesen Untersuchungskontext erlauben, sodass Adaptionen bestehender Instrumentarien bzw. teilweise Eigenkonstruktionen erforderlich sind (s. Kap. 6).

Die Datengrundlage zur Bearbeitung der Untersuchungsziele mithilfe dieser Methode liefert das an der Akademie für Bildungsforschung und Lehrerbildung und dem Institut

für Pädagogik der Sekundarstufe am Fachbereich Erziehungswissenschaften der Goethe-Universität Frankfurt am Main angesiedelte Forschungsprojekt „Strategien des Unterrichts in heterogenen Klassen und ihre Wirkung auf Schüleraktivität“. Das Untersuchungsdesign und die Datenerhebungen dieses Projekts werden ausführlicher dargestellt in Kap. 6.2. Über Videoaufzeichnungen von vollständigen Unterrichtsstunden hinaus wurden in dem genannten Projektkontext auch schriftliche Schülerbefragungen durchgeführt, aus denen Angaben zu Personenmerkmalen vorliegen, die eine gewünschte Differenzierung nach Untergruppen der Schülerstichprobe erlauben (Leistungsniveau im Fach, familiärer Sprachhintergrund und Geschlecht, vgl. Kap. 6.5). Durch die Abbildung natürlichen, d.h. durch keine untersuchungsbedingten Vorgaben oder Interventionen beeinflussten Unterrichts, bietet diese Stichprobe den Vorteil einer hohen ökologischen bzw. externen Validität (vgl. Bortz & Döring, 2006). Das heißt sie lässt Bezüge von den zu erzielenden Ergebnissen zur Wirklichkeit alltäglichen Unterrichts zu.

Unter anderem bedingt durch die Art des bearbeiteten Forschungsdesiderats liegen zu dieser spezifischen Fragestellung, die mithilfe des genannten forschungsmethodischen Ansatzes und auf Grundlage authentischen Unterrichts bearbeitet werden soll, kaum Forschungsergebnisse vor, die einen direkten Vergleich zulassen. Aus diesem Grund können keine genau spezifizierten Hypothesen, z.B. Erwartungen zu Ausprägungen und Verteilungen der Variablen oder Effektstärken, sondern mit Bezug auf die theoretischen Annahmen lediglich vereinzelt erwartbare Effektrichtungen formuliert werden (s. Forschungsfragen im nächsten Abschnitt); die Untersuchung ist somit von überwiegend explorativem Charakter.

5.3 Fragestellungen

Die übergeordnete Fragestellung dieser Studie, ob und welche Zusammenhänge sich zwischen den Aufgaben im Unterricht und dem Beteiligungsverhalten der Schüler erkennen lassen, wurde in den vorangegangenen Ausführungen dargelegt und begründet. Daraus werden die nachfolgenden spezifischen Forschungsfragen abgeleitet:

1. Wie ist das unterrichtliche Schülerbeteiligungsverhalten in der Stichprobe ausgeprägt?

- a. Wie stark variiert das Schülerbeteiligungsverhalten im Unterricht zwischen den Individuen? Sind interindividuelle Varianzen durch verfügbare Personenmerkmale der Schüler erklärbar?
- b. Wie stark schwankt das Schülerbeteiligungsverhalten zwischen aufeinanderfolgenden Unterrichtsstunden und auch innerhalb der Unterrichtsstunden? Gibt es Hinweise darauf, dass die individuelle Verhaltenstendenz durch Prozessbedingungen des Unterrichts zu erklären ist?
2. Wie ist das Aufgabenangebot in den untersuchten Unterrichtsstunden zu beschreiben?
 - a. Wie sind die untersuchten Aufgabenmerkmale in der Stichprobe ausgeprägt?
 - b. Wie stark unterscheiden sich die Unterrichtsstunden hinsichtlich quantitativer (Häufigkeit und Dauer der gestellten Aufgaben) und qualitativer Merkmale der Aufgaben (Verteilung der einzelnen Merkmalsfacetten)?
 - c. Wie ist die Variabilität des Aufgabenangebots in den Unterrichtsstunden ausgeprägt?
3. Sind Zusammenhänge zwischen dem Aufgabenangebot und dem Beteiligungsverhalten der Schüler im Unterricht festzustellen – und wenn ja, wie sind diese zu beschreiben?
 - a. Ist das Schülerbeteiligungsverhalten während einzelner Aufgaben im Unterrichtsprozess durch die Ausprägung bestimmter Merkmale (oder Merkmalskombinationen) der Aufgaben erklärbar?
 - b. Ist das durchschnittliche Schülerbeteiligungsverhalten in einer Unterrichtsstunde durch zusammenfassende Kennzeichen des stundenspezifischen Aufgabenangebots, wie z.B. die Häufigkeit von Aufgaben oder die Variabilität bestimmter Merkmale in der Stunde, erklärbar?
 - c. Wie stellen sich die Zusammenhänge zwischen Aufgabenangebot und Schülerbeteiligungsverhalten dar, wenn bei der Schülerstichprobe (auf Grundlage der verfügbaren Personenmerkmale) nach Untergruppen differenziert wird?

Eine Formulierung exakter Forschungshypothesen ist an dieser Stelle aufgrund fehlender Spezifikationen im zugrunde gelegten theoretischen Modell (Slavin, 1994) und eines Mangels an vergleichbar vorgehenden wissenschaftlichen Studien nicht

möglich und erscheint mit Blick auf den explorativen Charakter der Studie auch nicht sinnvoll. Unter Berücksichtigung der in den vorangegangenen Kapiteln dargestellten, allgemeinen Forschungserkenntnisse und unter Rückgriff auf den Bezugsrahmen des QAIT-Modells (ebd.) lassen sich jedoch für die zentrale Frage der Studie (Frage 3) die Erwartungen formulieren, dass im Falle (1) eines starken Bezugs der Aufgabeninhalte zur Lebenswelt der Schüler, (2) einer hohen Variabilität an Aufgabenmerkmalen und einer damit einhergehenden Anpassung an unterschiedliche Lernvoraussetzungen der Schüler und (3) einer als „Interessantheit“ bzw. „Attraktivität“ der Inhalte und Aktivitäten zu interpretierenden Ausprägung der Aufgaben ein höheres Niveau des Schülerbeteiligungsverhaltens zu vermuten ist.

6. Methodische Anlage der Untersuchung

Inhalt dieses Kapitels ist die Darstellung der Erhebungsverfahren und der verwendeten Methoden zur Analyse der erhobenen Daten. Wegen seiner zentralen Bedeutung für die forschungsmethodische Realisierung dieser Studie wird dazu zunächst der Ansatz videobasierter Analysen in der Unterrichtsforschung dargestellt (6.1). Darauf folgt eine Einordnung der vorliegenden Studie in den Kontext des Forschungsprojekts, in dessen Rahmen die Daten erhoben wurden (6.2), sowie eine Beschreibung der Stichprobe (6.3) und der Vorgehensweise bei den Datenerhebungen (6.4). Danach werden die verwendeten Forschungsinstrumente dargestellt (6.5 - 6.8).

6.1 Videoanalysen in der Unterrichtsforschung

In der empirischen Unterrichtsforschung gehören neben Befragungen, Tests und Dokumentenanalysen Verfahren der systematischen Beobachtung, speziell unter Nutzung audiovisueller Aufnahmen, zum gängigen Methodenrepertoire (Aufschnaiter & Welzel, 2001; Böhm-Kasper & Weishaupt, 2004; Janik & Seidel, 2009; Klieme, 2006). Reusser, Pauli und Waldis (2010, S. 7) kommen diesbezüglich zu der Feststellung: *„Der Einbezug von systematischen Videoanalysen und damit von Daten zur didaktischen Inhalts- und Prozessstruktur des Unterrichts erscheint in der aktuellen Unterrichtsforschung mittlerweile als nahezu selbstverständlich und ist nicht mehr wegzudenken.“* Eine Veranschaulichung der Rolle dieses methodischen Ansatzes in der Unterrichtsforschung liefern auch Seidel und Shavelson (2007) in einer theorie- und methodenvergleichenden Meta-Analyse zu Faktoren effektiven Unterrichts, wo sie belegen, dass video- bzw. beobachtungsbasierte Unterrichtsstudien im Vergleich zu unterschiedlichen Formen der Befragung insgesamt die höchsten Effekte in den Forschungsergebnissen erzielen (S. 478).

Im deutschsprachigen Raum hat die Verbreitung dieser Erhebungs- und Analysemethodik in den letzten Jahren stark zugenommen, unter anderem verstärkt durch Forschungsprojekte wie TIMSS (Stigler, Gonzales, Kawanaka, Knoll & Serrano, 1999), die schweizerisch-internationale Videostudie (Reusser, Pauli & Waldis, 2010), die IPN-Videostudie (Seidel et al., 2006) oder PERLE (Lotz, Lipowsky & Faust; 2013). Ihre Anfänge hat die videogestützte Unterrichtsforschung in den 1970er Jahren, als

Technologien zum Anfertigen und Abspielen von Videoaufnahmen so weit entwickelt waren, dass Videostudien als zugleich ökonomisch vertretbare und praktikable Möglichkeit zur Erforschung von Unterricht in Betracht kamen. Zu den Pionieren dieses Zugangs zählen beispielsweise Adams und Biddle (1970), die bei ihrer Arbeit zum Interaktionsgeschehen zwischen Lehrern und Schülern die zeitgenössische Empirie als unbefriedigend empfanden und daher mit Videoaufnahmen arbeiteten. Oder Kounin (1976), der zur Untersuchung des *classroom managements* nach Möglichkeiten zur Erweiterung der damals gängigen Beobachtung im Klassenraum suchte und aus diesem Grund so genannte „Video-Recorder-Studien“ durchführte.

Die von Beginn an gesehenen Vorteile der wissenschaftlichen Verwendung von Videoaufnahmen des Unterrichts werden im Wesentlichen bis heute angeführt (vgl. Aufschnaiter & Welzel, 2001; Biddle, 1967; Dinkelaker & Herrle, 2009; Jacobs et al., 2003; Pauli & Reusser, 2006; Petko, Waldis, Pauli & Reusser, 2003): Ein großer Gewinn ergibt sich aus der Tatsache, dass Videodaten einen wiederholten Zugriff auf das audiovisuelle Geschehen der Klassensituation zulassen. Somit werden prinzipiell beliebig viele Analysen des Materials mit unterschiedlichen analytischen Zugängen und Fragestellungen ermöglicht. Dies erscheint zum einen unter methodologischen Gesichtspunkten interessant, weil auf diese Weise die Erprobung und der Vergleich von Beobachtungsinstrumentarien vereinfacht wird, und zum anderen Datenbestände aufgebaut werden können, die sich zur Nutzung in Forschungsverbänden oder für Sekundäranalysen anbieten. Erfüllen die Aufnahmen bestimmte Standards der Bild- und Tonqualität, besteht zudem ein weiterer Nutzen darin, dass die Daten durch Techniken wie Vor- und Zurückspulen, Zeitlupe oder Zoom tieferen Analysen zugänglich gemacht werden, als es dem menschlichen Beobachter unter normalen Bedingungen möglich wäre. Das bezieht sich sowohl auf meist unbemerkte Ereignisse im Klassenraum (z.B. Briefchen unter dem Tisch, Gesten von Schülern), als auch auf die Gleichzeitigkeit verschiedener Handlungen in der Unterrichtssituation, die sich in ihrer Fülle und Komplexität der Aufmerksamkeit eines einzelnen Beobachters zwangsläufig entziehen. Videoaufnahmen ermöglichen somit ein relativ hohes Niveau an Vollständigkeit und Genauigkeit der Daten zum Unterrichtsgeschehen, wenngleich durch unvermeidbare Selektionsentscheidungen bei der Aufzeichnung (z.B. Wahl des Ausschnitts oder der Kameraperspektive) immer einige Aspekte der abzubildenden Situation unberücksichtigt bleiben müssen, wie z.B. Geschehnisse außerhalb des

Aufzeichnungsbereichs oder Sinnerfahrungen zur physischen Umgebung (Dinkelaker & Herrle, 2009, S. 15). Audiovisuelle Daten bieten aber auch vielfältige Punkte zur Verknüpfung mit alternativ erhobenen Datenbeständen, wie z.B. Lehrerbefragungen oder Schülertests, um die Aussagekraft von beobachtungsbasierten Analysen durch Angaben über den Kontext oder die beteiligten Personen aufwerten zu können (vgl. Appel & Rauin, 2015; Petko, Waldis, Pauli & Reusser, 2003; Wild, 2003). Ein weiterer Vorteil dieses forschungsmethodischen Zugangs besteht darin, dass durch die Trennung der apparativ basierten Datenerhebung und der zu einem späteren Zeitpunkt von den Forschenden vollzogenen Analyse eine klare Trennung zwischen Protokollierung und Interpretation der Ereignisse möglich wird, was zunächst die Objektivität der Daten erhöht. Zusätzlich ist auch eine erhöhte Genauigkeit der Einschätzungen bei gleichzeitiger Reduzierung von Fehlentscheidungen möglich, weil beim Vorgang der Identifizierung und Bewertung der Ereignisse weder besonderer Zeitdruck besteht, noch kognitive Einschränkungen des Beobachters aufgrund von Ablenkungen oder „einem schlechten Tag“ zu befürchten sind.

Neben diesen Vorteilen weiß man aber mittlerweile auch um einige Herausforderungen, die mit wissenschaftlichen Videoanalysen verbunden sind. So ist bereits im Vorfeld die datenschutzrechtliche Absicherung zu beachten, zu der schriftliche Einverständniserklärungen aller abzubildenden Personen bzw. ihrer Erziehungsberechtigten, inkl. der zuständigen Behörde (Schulleitung, Schulamt) einzuholen sind. Solche Vorgänge können durch bereits bestehende Kooperationen zu den Schulen bzw. Schulträgern unter Umständen erheblich erleichtert werden, sollten jedoch hinsichtlich Aufwands und Bedeutung nicht unterschätzt werden. Eine weitere Herausforderung besteht in der Realisierung eines tragfähigen Untersuchungsdesigns. So kann die Qualität der Stichprobe - mit Blick auf die jeweilige Fragestellung - durch organisatorische Einschränkungen der Datenerhebung beeinträchtigt werden (z.B. Verteilung von Personen- oder Gruppenmerkmalen in den zugänglichen Lerngruppen, mangelnde Verfügbarkeit der gewünschten Klassenstufe oder des Unterrichtsfachs). Nach Abschluss der Vorarbeiten erfordert die Anfertigung der Videoaufnahmen eine Reihe von Erwägungen, die sich aus dem Zusammenspiel technisch-organisatorischer und methodischer Gesichtspunkte ergeben. Denn möchte man eine möglichst umfassende Abbildung des Geschehens bei gleichzeitig hoher Qualität der Video- und Audioaufnahmen erzielen, ist der Einsatz entsprechend hochwertiger Geräte in

ausreichender Anzahl notwendig. Für ein annähernd vollständiges Bild der Klasse sind beispielsweise in der Regel zwei Kamera- und mindestens ebenso viele Mikrofonperspektiven erforderlich, was im Falle variierender Raum- oder Gruppenanordnungen und in Abhängigkeit von der Fragestellung gegebenenfalls zu erweitern ist (vgl. Müller, Eichler & Blömeke, 2006; Seidel, Dalehefte & Meyer, 2005).

Mit dem Anstieg des technischen Erhebungsaufwandes verstärkt sich allerdings die Gefahr unerwünschter Beeinflussung des zu erforschenden Feldes und der darin agierenden Personen durch so genannte Invasivitätseffekte (vgl. Bortz & Döring, 2006; Petko, Waldis, Pauli & Reusser, 2003), sodass die Datengrundlage zwar technisch an Qualität gewinnen mag, zugleich aber auf Grund möglicher Verzerrungen die Validität leiden kann. Abschließend kann die Aufbereitung der Daten je nach Aufnahmequalität, verfügbarer Hard- und Software und erwünschten Zieleigenschaften des Datenbestandes mit einem immensen Aufwand verbunden sein. Somit lässt sich der Zugang videogestützter Beobachtungen von (Schul-) Unterricht aus planungs- und organisationstechnischer Hinsicht zusammenfassend als voraussetzungsreich, gleichzeitig aber hinsichtlich sowohl methodischer als auch theoretischer Möglichkeiten vielversprechend bewerten.

Für die vorliegende Untersuchung erscheint diese methodische Herangehensweise aus verschiedenen Gründen besonders geeignet: (1) Der für die vorliegende Arbeit bedeutsame Prozesscharakter des Unterrichts lässt sich auf der Basis von Video-Audio-Daten gut abbilden. (2) Von zentraler Bedeutung ist dafür das lückenlose Protokollieren der Handlungs- und Verhaltensweisen von Lehrkraft (Steuerung des Aufgabengeschehens) und Schülern (Beteiligungsverhalten), was mit anderen Arten von Datenbeständen kaum in der erforderlichen Tiefe möglich wäre. (3) Die – im vorliegenden notwendige – Konstruktion und Erprobung von Beobachtungsinstrumentarien ist durch wiederholte Analysen der Videodaten relativ problemlos realisierbar. (4) Der Untersuchungsgegenstand, Situationen realen Alltagsunterrichts, soll durch das Datenmaterial möglichst umfassend abgebildet werden. Besonders in Verbindung mit flankierenden Datenbeständen, wie z.B. schriftlichen Schülerbefragungen oder administrativen Daten, sind audiovisuelle Daten in ihrer „Authentizität und Ganzheitlichkeit“ (Pauli & Reusser, 2006, S. 787) im Stande, dies zu leisten.

6.2 Kontext der Untersuchung

Die Daten der vorliegenden Studie stammen aus dem Forschungsprojekt „Strategien des Unterrichts in heterogenen Klassen und ihre Wirkung auf Schüleraktivität“, welches unter der Leitung von Prof. Dr. Udo Rauin am Institut für Pädagogik der Sekundarstufe am Fachbereich Erziehungswissenschaften und der Akademie für Bildungsforschung und Lehrerbildung der Goethe-Universität Frankfurt am Main durchgeführt wird. Ziel dieser als Pilotstudie angelegten Untersuchung ist es, geeignete Instrumentarien zu entwickeln und zu erproben, die in größer angelegten Folgestudien Aufschluss über Effekte adaptiver Unterrichtsstrategien von Lehrkräften auf das Lernverhalten der Schüler in Abhängigkeit von der Größe und Heterogenität der Klasse und der diagnostischen Kompetenzen der Lehrkraft liefern können. Die Datengrundlage der Studie setzt sich aus folgenden Bestandteilen zusammen:

- (1) Videoaufnahmen von vollständigen Unterrichtsstunden zum Zwecke der Beschreibung der Handlungsstrategien von Lehrkräften und des Verhaltens der Schüler im Unterricht,
- (2) Leitfadeninterviews zur Erfassung der diagnostischen Kompetenz der Lehrkräfte,
- (3) ein schriftlicher Test mit den Schülern zur Abbildung des individuellen Leistungsstandes im Unterrichtsfach und zur Erfassung von Personenmerkmalen,
- (4) schriftliche Befragungen der Schüler jeweils nach Ende der aufgezeichneten Unterrichtsstunden zu ihrer Wahrnehmung des erlebten Unterrichts und
- (5) eine Dokumentation der verwendeten Lernmaterialien durch die Lehrkräfte sowie Klassenlisten und Sitzpläne der Klassen für eine Zuordnung der Schüler im Raum.

Den Mittelpunkt der Studie bilden die Videoaufnahmen der Unterrichtsstunden, die durch eine umfassende technische Postproduktion und flankierende Datenaufbereitungen (z.B. Transkription der verbalen Kommunikation in der Klasse sowie automatische Vorkodierungen des Sprachgeschehens) für unterschiedliche Analysen vorbereitet wurden. Die vorliegende Arbeit behandelt einen Ausschnitt aus dieser Rahmenstudie, sodass an dieser Stelle nicht alle der genannten Datenbestände von Bedeutung sind. Daher werden in den nachfolgenden Abschnitten nur die an dieser Stelle relevanten Datenbestände und Instrumente aus diesem Kontext dargestellt: die

Videoaufzeichnungen und die entsprechenden Beobachtungsinstrumente, der Leistungstest und ergänzend dazu die Klassenlisten, Sitzpläne und Dokumentationen der Lernmaterialien.

6.3 Stichprobenbeschreibung

Die Stichprobe setzt sich zusammen aus sechs Klassen der Jahrgangsstufe sechs, verteilt auf fünf Schulen in städtischen Gebieten der hessischen Rhein-Main-Region. Je zwei Klassen entstammen den Schulformen Hauptschule, Realschule und Gymnasium. Die Klassen wurden von sechs unterschiedlichen Lehrkräften unterrichtet. Um eine Vergleichbarkeit hinsichtlich der allgemeinen didaktischen Orientierung und des grundlegenden Spektrums der erwartbaren Sozial- und Arbeitsformen zwischen den aufgezeichneten Unterrichtsstunden zu erreichen, wurden die Videoaufzeichnungen überall im selben Unterrichtsfach, konkret dem Fach Englisch, durchgeführt. Von jeder Klasse wurde der Englischunterricht von vier aufeinanderfolgenden Unterrichtsstunden aufgezeichnet, sodass – unter Abzug einer ausgefallenen Unterrichtsstunde – Videoaufnahmen von insgesamt 23 Unterrichtsstunden vorliegen.

Um die Authentizität des aufgezeichneten Unterrichts nicht zu beeinträchtigen, wurden von Seiten der Forschenden keinerlei Vorgaben gemacht oder Wünsche geäußert hinsichtlich der Gestaltung und Durchführung des zu untersuchenden Unterrichts. Die Teilnahme an der Studie geschah auf eigenen Wunsch der Lehrkräfte hin, nachdem unter Zustimmung der zuständigen Dienstbehörde Anfragen an die beteiligten Schulen gestellt worden waren. Zur rechtlichen Absicherung der Datenerhebungen wurden im Vorfeld der Untersuchung Einverständniserklärungen von den Erziehungsberechtigten aller Schüler eingeholt. Lediglich zwei Schüler mussten aufgrund fehlenden Einverständnisses von den Erhebungen ausgenommen werden.

Einen Überblick zur Stichprobenzusammensetzung bietet Tab. 1. Die beteiligten Klassen umfassen insgesamt 145 Schüler. Diese Zahl konnte allerdings in der letztendlichen Untersuchungsstichprobe nicht vollständig ausgeschöpft werden, was neben der bereits erwähnten Verweigerung der Untersuchungsteilnahme im Falle von zwei Schülern (s.o.) auf Schülerabwesenheiten bei Erhebungsterminen (Leistungstest und/oder Videoaufzeichnungen) oder im Falle der Beobachtungsdaten auf eine zeitweise

eingeschränkte Sichtbarkeit einzelner Schüler bei den Videoaufnahmen zurückzuführen ist. Am schriftlichen Leistungstest nahmen 132 Schüler teil, was einer Beteiligung von 91 % entspricht. Der Anteil anwesender und beobachtbarer Schüler in den Klassen zu den Terminen der Videoaufzeichnung lag bei durchschnittlich 85 %. Die schwankende Zusammensetzung der Beobachtungsstichprobe stellt insofern kein schwerwiegendes Problem für diese Studie dar, dass für die unterschiedlichen Teilfragestellungen zusammengefasste Gruppenmittelwerte die Grundlage bilden. Erläuterungen dazu werden bei den Einführungen der jeweiligen Auswertungsschritte in Kap. 7 geliefert.

Wie Tab. 1 zeigt, sind die Geschlechter in der Stichprobe trotz vereinzelter Unterschiede zwischen klassenspezifischen Verteilungen (z.B. Klassen B und D) insgesamt in etwa gleich verteilt und entsprechen damit den bundesweiten Verhältnissen an allgemeinbildenden Schulen im Erhebungszeitraum (Statistisches Bundesamt, 2010). Der Anteil an Schülern, die zu Hause Deutsch sprechen, liegt bei durchschnittlich 22 % und ist als relativ niedrig anzusehen²⁵.

Tab. 1: Überblick über die Schülerstichprobe

Klasse	Schulform ^a	Klassen-größe	Stichprobe Leistungstest	Beobachtungsstichprobe ^b	Geschlecht (Anteil weiblich)	Umgangssprache in der Familie (Anteil Deutsch)
A	RS	19	16	18	58 %	25 %
B	HS	20	17	16	30 %	18 %
C	HS	22	17	17	41 %	24 %
D	GY	27	26	22	70 %	31 %
E	RS	29	28	25	41 %	18 %
F	GY	28	28	26	39 %	18 %
Durchschnitt		24	22	21	47 %	22 %
Summe		145	132	-	-	-

^a RS = Realschule, HS = Hauptschule, GY = Gymnasium ^b Durchschnitt über Erhebungszeitpunkte (Unterrichtsstunden) hinweg

6.4 Vorgehensweise bei den Datenerhebungen

Die Datenerhebungen fanden im Laufe des Jahres 2009 statt. Die Im Februar 2009 wurden die schriftlichen Leistungstest durchgeführt. Die Videoaufzeichnungen in den

²⁵ Ausführlichere Ergebnisdarstellungen zu diesem Merkmal sind in Kap. 7.1 zu finden.

Klassen wurden im Zeitraum zwischen Februar und April 2009 ausgeführt. Die hier nicht näher dargestellten Leitfadeninterviews mit den Lehrkräften wurden im September 2009 durchgeführt. Ihren Abschluss fanden die Datenerhebungen im Oktober 2009 mit Eingang der letzten Lernmaterialien, Klassenlisten und Sitzpläne von den Lehrkräften. Die Art und Weise der Organisation und Durchführung der unterschiedlichen Erhebungsbestandteile wird nachfolgend beschrieben.

Leistungstest im Unterrichtsfach und Erhebung von Geschlecht und Sprachhintergrund: Die Administration der schriftlichen Tests, in deren Rahmen gleichzeitig das Geschlecht und der familiäre Sprachhintergrund erhoben wurden (s. Kap. 6.5), oblag den Lehrkräften und geschah in der jeweils letzten Unterrichtsstunde vor Beginn der Videoaufzeichnungen in dem Fach bzw. bei der betreffenden Lehrkraft. Um im Sinne der Durchführungsobjektivität (Moosbrugger & Kelava, 2008) die Testdurchführung soweit wie möglich zu standardisieren, wurden den Lehrkräften einheitliche Anweisungen zur Kommunikation der mit dem Test verbundenen Absichten gegenüber den Schülern sowie zur Durchführung in der Klassensituation an die Hand gegeben.

Aufzeichnung der audiovisuellen Daten: Es wurden sowohl Video- als auch Tonaufzeichnungen der zu untersuchenden Unterrichtsstunden angefertigt, beides nach einer möglichst einheitlichen Vorgehensweise, um eine Vergleichbarkeit des Materials zu gewährleisten. Für eine möglichst vollständige visuelle Abbildung des Geschehens im Klassenraum wurden in Anlehnung an die Verfahrensweisen von Videostudien, die vergleichbare Anforderungen an die Daten stellten (Jacobs et al., 2003; Klieme, Pauli & Reusser, 2006; Seidel, Dalehefte, & Meyer, 2005), zwei Kameraperspektiven verwendet. Die so genannte Lehrerkamera (LK) war auf einem Stativ im hinteren Teil des Raumes platziert (Abb. 3). Die Aufnahmeperspektive dieser Kamera wurde durch eine eigens dafür abgestellte Person flexibel geführt, sodass die Lehrkraft in ihren Bewegungen durch den Klassenraum verfolgt und bei besonderen Vorkommnissen der betreffende Bereich fokussiert werden konnte (vgl. Prinzipien „*Folge immer der Lehrperson*“ und „*Zone der Interaktion*“, Petko, Waldis, Pauli & Reusser, 2003, S. 271). Die so genannte Schülerkamera (SK) war mit einem Weitwinkelobjektiv ausgestattet und wurde in einer der vorderen Ecken des Raumes mit fester Perspektive auf die Schüler positioniert (s. Abb. 3). Durch die Kombination von statischem und dynamischem Bild wurde somit ein Kompromiss zwischen „*Standardisierung und Offenheit*“ (ebd.) der Erhebungssituation erreicht. Auch die zusätzlich angefertigten Tonaufnahmen geschahen über mehrere

Kanäle: Erstens mit Hilfe eines statischen Stereo-Raummikrofons im hinteren Teil des Raums (Abb. 3), positioniert direkt neben der Lehrerkamera, zweitens mit einem Mikrofon, welches an der Kleidung (z.B. dem Hemdkragen) der Lehrperson befestigt war, und drittens zusätzlich durch die integrierten Mikrofone der Kameras.

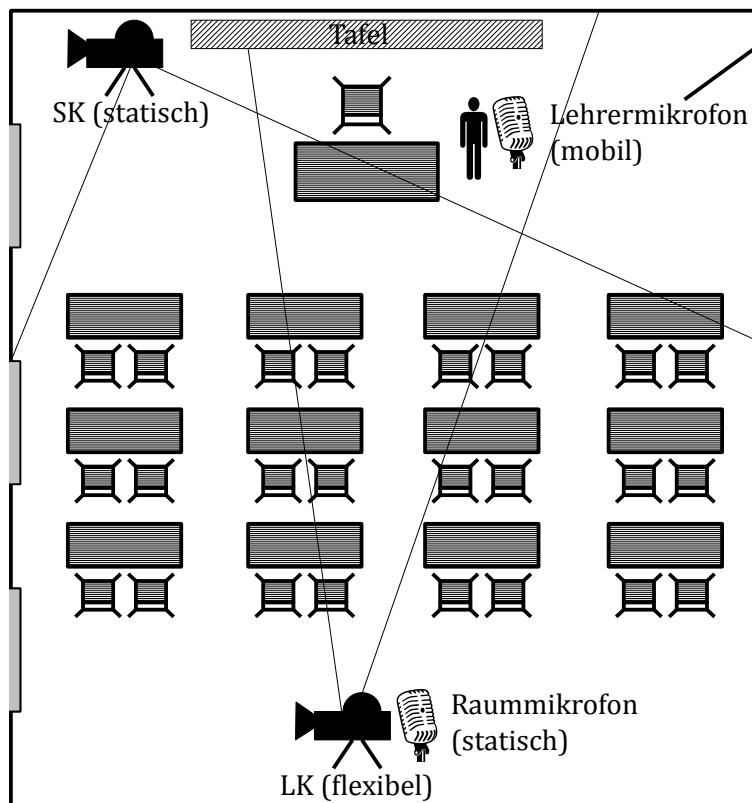


Abb. 3: Skizze des Geräteaufbaus zur Erhebung der Video- und Audiodaten im Klassenraum

Auf diese Weise war eine Abbildung des Geschehens im Klassenraum in einem den Studienzielen angemessenen Umfang möglich, ohne die natürliche Klassensituation allzu stark zu beeinflussen. Um mögliche invasive bzw. reaktive Effekte (vgl. Bortz & Döring, 2006; Petko, Waldis, Pauli & Reusser, 2003) auf das Untersuchungsfeld, beispielsweise in Form sozial erwünschten Verhaltens der Schüler oder „idealisierten“ Lehrerhandelns (Stigler, Gonzales, Kawanaka, Knoll & Serrano, 1999, S. 6), zu minimieren, wurden die teilnehmenden Klassen im Vorfeld der Aufzeichnungen vom wissenschaftlichen Personal besucht, die Absichten der Studie vorgestellt und probeweise Aufzeichnungen durchgeführt, um die Beteiligten an die Personen, die Geräte und die gesamte Erhebungssituation zu gewöhnen. Zudem wurde versucht, die Präsenz der Forscher im Klassenraum möglichst dezent zu gestalten (Petko, 2006), indem die Aufnahmen in den

Klassen von lediglich zwei Personen durchgeführt wurden, die sich in der Klassensituation um ausgeprägt zurückhaltendes und unauffälliges Verhalten bemühten. Zudem geschahen Aufbau und Einrichtung der Geräte für die Aufzeichnungen stets vor Eintritt der Klasse in den Raum.

Lernmaterialien und Klassendaten: Um detaillierte Untersuchungen der Unterrichtsinhalte und -aktivitäten zu ermöglichen, wurden von den Lehrkräften nach Abschluss der Datenerhebungen in den Klassen Kopien der in den aufgezeichneten Unterrichtsstunden verwendeten Arbeitsmaterialien (z.B. Arbeitsblätter oder Projektorfolien) oder Angaben zu den genutzten Lehrwerken (z.B. Lehrbuch oder Übungsheft, inkl. Seitenangaben) an die Forschungsgruppe übergeben. Mit dem Ziel einer Zuordnung der Beobachtungs- zu den Befragungs- bzw. Testdaten der einzelnen Schüler wurden auch die Klassenlisten und Sitzpläne der Forschungsgruppe zugänglich gemacht.

6.5 Testinstrument zur Erfassung der fachlichen Leistung der Schüler

Mit dem Ziel, die Schüler hinsichtlich ihrer Leistungsstände in dem Unterrichtsfach differenzieren zu können, wurde im Vorfeld der Videoerhebungen ein schriftlicher standardisierter Test zur Bearbeitung in die Klassen gegeben. Der dazu verwendete Englischtest wurde weitgehend aus der Evaluation eines Projekts zur Einführung von Englischunterricht in der Primarstufe in 6. Klassen des Schweizer Kantons Appenzell Innerrhoden übernommen (Bader & Schaer, 2005). Die Wahl fiel aus verschiedenen Gründen auf diesen Test: Zum einen war aufgrund der Beteiligung von Klassen unterschiedlicher Schulformen eine hohe fachspezifische Leistungsheterogenität in der Stichprobe zu erwarten. Da in den Schweizer Untersuchungen, wo die Schüler sich in der 6. Klasse im letzten Jahr der Primarstufe befinden, eine ähnliche Ausgangslage der Stichprobe besteht, sollte das Instrument diese Anforderungen erfüllen können. Zum anderen mussten die fachlichen Inhalte des Tests möglichst allen beteiligten Klassen bekannt und damit curricular abgesichert sein. Nach Sichtung der Lehrpläne und Absprachen mit den beteiligten Lehrern erschien der vorliegende Test auch in dieser Hinsicht geeignet. Um verschiedene Felder des stufenspezifischen Englischunterrichts anzusprechen, deckt das Testinstrument drei Bereiche ab (s. Anhang A): Hörverstehen (*Listening*, 5 Items), Leseverstehen (*Reading*, 3 Items) und Grammatik (*Grammar*, 15

Items). Die Reliabilität des Leistungstests, gemessen an der inneren Konsistenz mittels Cronbachs α , fällt für die Untertests ausreichend bis zufriedenstellend (Hörverstehen: 0,76; Leseverstehen: 0,54; Grammatik: 0,83) und für den Gesamttest gut aus (0,88). Da im weiteren Verlauf der Untersuchung nicht nach Teilleistungen differenziert werden soll, ist an dieser Stelle die Reliabilität des Gesamttests maßgeblich.

Innerhalb des Testfragebogens wurde gleichzeitig das Geschlecht und der familiäre sprachliche Hintergrund der Schüler erfragt („Ich bin ein Mädchen./Ich bin ein Junge.“, „Zu Hause spreche ich Deutsch./Ich spreche zu Hause eine andere Sprache./Welche?“). In der Schul- und Unterrichtsforschung wird die Umgangssprache in der Familie, auch „Familiensprache“ genannt, als wichtiger Hinweis auf einen möglichen Migrationshintergrund der Schüler gesehen (vgl. Hesse, Göbel und Hartig 2008). Im Rahmen des übergeordneten Forschungsprojekts (s. Kap. 6.2) wurde die Erfassung dieses Merkmals gegenüber einer differenzierten Erfragung des Migrationshintergrunds – etwa über das Geburtsland der Elternteile oder den Zeitpunkt der Zuwanderung der Familie (ebd.) – aufgrund der besseren Verständlichkeit für die Schüler als vorteilhaft angesehen und geschah mit dem vorrangigen Ziel, in Kombination mit den Leistungsdaten die Heterogenität der Lernvoraussetzungen in der jeweiligen Klasse als zusammenfassenden Index abzubilden. In der vorliegenden Studie wird das Merkmal hingegen als eigenständige Variable behandelt und dient neben dem beschriebenen Leistungsstand im Unterrichtsfach und dem Geschlecht der Schüler im weiteren Verlauf dazu, die zu untersuchenden Effekte zwischen den Aufgaben und dem Schülerbeteiligungsverhalten im Unterricht nach Untergruppen in der Schülerschaft zu differenzieren (vgl. Kap. 5.3, Fragestellung 3c).

6.6 Instrument zur beobachtungs-basierten Kodierung des Schülerbeteiligungsverhaltens

Das Schülerbeteiligungsverhalten bildet in dieser Untersuchung die Zielgröße, an der die Effektivität des Unterrichts innerhalb von dessen Prozess gemessen werden soll. Damit sind Anforderungen an das zur Erfassung erforderliche Beobachtungsverfahren verbunden. Aufgrund der zentralen Rolle dieser Variable im vorliegenden Kontext sollen daher zunächst die Voraussetzungen und erwünschten Eigenschaften des

Kodierverfahrens expliziert werden (6.6.1), bevor in den beiden darauf folgenden Abschnitten auf das selbst konstruierte System eingegangen wird (6.6.2/6.6.3).

6.6.1 Voraussetzungen für die Konstruktion des eigenen Kodiersystems

Hinsichtlich der Ziele der Untersuchung sollte mit dem System eine ausreichend große Beobachtungsstichprobe angestrebt werden, die viele Beobachtungszeitpunkte im Unterrichtsverlauf umfasst. Denn nur auf diese Weise ist eine angemessene Berücksichtigung sowohl des Prozesscharakters des Unterrichts als auch der Zusammensetzung der Lerngruppe zu erreichen. Wie gezeigt, erscheint dafür aus konzeptioneller Sicht das Konstrukt der *time on task* als besonders geeignet (vgl. Kap. 3). Wie bereits im Rahmen des Forschungsüberblicks angedeutet, erscheint jedoch kein bestehendes Instrument für die spezifischen Untersuchungsziele dieser Studie nutzbar (s. Kap. 2.3.4). So werden Variablen wie *time on task*, *on/off task*-Verhalten, *engagement* usw. in der Forschungspraxis meist auf Basis von ausgewählten (und in ihrem Umfang meist recht geringen) Teilstichproben der Schülergruppe und des Beobachtungszeitraums erhoben und die Daten in der Regel auf Stundenebene aggregiert²⁶. Diese Herangehensweise hat den Vorteil, dass sich der Erhebungsaufwand in Grenzen hält, beschränkt damit aber die Aussagekraft der Ergebnisse auf die gesamte Unterrichtsstunde (vgl. Treiber, 1982). Einige Instrumente setzen zudem innerhalb der Verhaltenskategorien inhaltliche Schwerpunkte, die für die vorliegende Fragestellung nicht angemessen sind, wie z.B. die Art der Interaktion zwischen den Unterrichtsbeteiligten (Blatchford, Bassett & Brown, 2005, 2011), die Involviertheit in Warte- und Übergangszeiten (Romberg, 1980), oder umfassen relativ komplexe „Verhaltensbündel“, die aufgrund von Sonderfällen und Prioritätenregelungen für die Erhebung größerer Beobachtungsstichproben wenig praktikabel sind (Emmer, Sanford, Clements & Martin, 1982; Helmke, 1988). Andere Instrumente sind nicht sehr umfassend dokumentiert oder nicht für größere empirische Studien konzipiert (Bessoth, 2010; Houtveen, Booij, de Jong & van de Grift, 1999; Trautmann, 2008).

Mit Blick auf diese Voraussetzungen erscheint die Eigenkonstruktion eines Beobachtungsinstruments notwendig. Da eine große Beobachtungsstichprobe angestrebt wird und dies mit einem großen Umfang an Kodierentscheidungen

²⁶ s. Kap. 2.3 und 2.4

verbunden ist, sollte dabei eine hohe Praktikabilität in der Handhabung bei gleichzeitiger Präzision der Urteile gegeben sein. Daher sollte die Kodierung des Beteiligungsverhaltens als niedrig-inferentes Verfahren realisiert werden, d.h. durch die Gestaltung der Beobachtungskategorien einen möglichst niedrigen Anteil an Schlussfolgerungen bzw. Interpretationen seitens des Beobachters erfordern (vgl. Lotz, Gabriel & Lipowsky, 2013; Pauli, 2012). Urteilsverzerrungen könnte zudem vorgebeugt werden durch die Verwendung von Verhaltensindikatoren als Entscheidungsgrundlage für die letztendliche Kategorisierung des Verhaltens, wie etwa in einem Beobachtungssystem von Erhardt, Findeisen, Marinello und Reinartz-Wenzel (1981, zitiert nach Büttner & Schmidt-Atzert, 2004), in dem die individuelle Schüleraufmerksamkeit als ein Summenwert auf Grundlage von getrennt voneinander eingeschätzten Indikatoren (z.B. Blickrichtung und Körperhaltung zum Unterrichtsmittelpunkt und Ausübung von aufgabenbezogenen Tätigkeiten) erhoben wird. Indem sich die Erfassung des Schülerbeteiligungsverhaltens bzw. der *time on task* auf tatsächlich beobachtbares Verhalten beschränkt und innere Prozesse des Wahrnehmens und Erlebens der Schüler, wie z.B. Denken oder Träumen (Dishaw, 1978), Enthusiasmus (Stipek, 2002) oder Aufmerksamkeit und Lernen (Helmke, 1988) vermeidet, werden zudem die in Kap. 2.3.3 beschriebenen, allgemeinen Validitätsprobleme in diesem Bereich umgangen.

Bei einer solch detaillierten Sichtweise auf das Beteiligungsverhalten der Schüler muss berücksichtigt werden, dass verschiedene Sozial- und Arbeitsformen des Unterrichts unterschiedliche Beteiligungsmöglichkeiten für die Schüler implizieren (Blömeke & Müller, 2008), sodass ein auf alle denkbaren Unterrichtsbedingungen anwendbares System eine immense Vielzahl von Verhaltenskonstellationen zu berücksichtigen hätte. Daher erscheint es sinnvoll, sich an dieser Stelle auf eine in quantitativer Hinsicht möglichst repräsentative Sozialform des Unterrichts zu beschränken. Wie eine Reihe von Ergebnissen aus der Unterrichtsforschung in unterschiedlichen Fächern und Schulformen zeigt, bietet sich dafür die Konstellation des lehrerzentrierten Klassenunterrichts, auch bezeichnet als Klassengespräch, Unterrichtsgespräch, Frontalunterricht oder öffentlicher Unterricht, an (Bohl, Kleinknecht, Batzel & Richey, 2012; Götz, Lohrmann, Ganser & Haag, 2005; Hage et al., 1985; Kleinknecht, 2010; Seidel et al., 2006). Im nächsten Abschnitt wird daher zunächst die Vorgehensweise zur Selektion entsprechender Unterrichtsphasen in der vorliegenden

Untersuchungsstichprobe mithilfe eines getrennten Kodiersystems beschrieben, bevor in Abschnitt 6.6.3 das System zur Kodierung des Schülerbeteiligungsverhaltens dargestellt wird.

6.6.2 Analytischer Zwischenschritt: Basiskodierung der Sozialformen

Da sich das Beobachtungssystem zur Erfassung des Schülerbeteiligungsverhaltens wie zuvor dargelegt auf lehrerzentrierte Sozialformen des Unterrichts beschränken soll, besteht eine Vorstufe der Analysen darin, die entsprechenden Phasen des Unterrichts zu identifizieren. Dazu wurde mit einem an Systeme wie jenes der Sichtstrukturen aus der IPN-Videostudie (Seidel, 2003a) oder der Sozialformen aus der schweizerisch-deutschen Videostudie (Hugener, 2006) angelehntes Kategoriensystem eine Basiskodierung der aufgezeichneten Unterrichtsstunden hinsichtlich der Sozialformen durchgeführt. Das Kodierverfahren wurde anhand einer Ereignisstichprobe durchgeführt (vgl. Pauli, 2012; Seidel & Prenzel, 2010). Die Sozialformen wurden in diesem Zuge durch Setzen einer Anfangs- und Endzeit in ihrem tatsächlichen Auftreten innerhalb des Unterrichts protokolliert (*timed-event recording*, Bakeman & Quera, 2011, S. 29). Die Kategorien für die Sozialformen sind in Kurzform folgendermaßen zu beschreiben:

- *Öffentlicher Unterricht (ÖU)*: Der primäre Aufmerksamkeitsfokus der Klasse liegt auf der Lehrperson, einem anderen legitimen Sprecher oder dem Unterrichts- bzw. Präsentationsmaterial. Das Geschehen wird zentral gesteuert durch die Lehrkraft.
- *Einzelarbeit (EA)*: Die Schüler beschäftigen sich alleine mit einer Aufgabe oder einem Arbeitsauftrag. Abgesehen von Äußerungen der Lehrkraft oder Wortmeldungen der Schüler (z.B. mit Fragen zur Aufgabe) ist in diesen Phasen keine öffentliche Kommunikation im Klassenraum vorgesehen.
- *Partnerarbeit (PA)*: Hier gelten die gleichen Bedingungen wie bei der Einzelarbeit, nur dass die Schüler paarweise mit Arbeitsaufträgen beschäftigt sind und dabei in der Regel die mündliche Kommunikation zwischen den Arbeitspartnern erlaubt bzw. erwünscht ist.
- *Gruppenarbeit (GA)*: Für Phasen der Gruppenarbeit gelten wiederum ähnliche Bedingungen wie für die Partnerarbeit, nur dass die Gruppen hierbei aus mehr als zwei Schülern bestehen.

- *Uneinheitliche Sozialform (US)*: Mischformen wie z.B. parallel stattfindende Einzel- und Partnerarbeit.
- *Keine Sozialform (KS)*: Es ist keine Sozialform, d.h. keine von der Lehrkraft eingeleitete und gesteuerte Interaktions- und Arbeitsstruktur in der Klasse erkennbar, was in der Konsequenz das Fehlen eines „offiziellen“ Unterrichtsgeschehens bedeutet.

Durchführung der Kodierungen: Ausführliche Ereignis- und Kategorienbeschreibungen wurden in einer schriftlichen Kodieranleitung für die Beobachter festgehalten. Zusätzlich fand eine Schulung zur Durchführung der Kodierungen statt. Die Kodierungen wurden von vier Personen mithilfe der Videoanalysesoftware *Interact 9* (Mangold International, Arnstorf)²⁷ vorgenommen. Dafür wurden beide verfügbaren Kameraperspektiven (Lehrer- und Schülerkamera, s. Kap. 6.4) genutzt. Nach einem ersten Kodierdurchgang für alle 23 Unterrichtsstunden wurde eine Nachschulung vorgenommen, in der Anwendungsschwierigkeiten diskutiert und unter Berücksichtigung erster Erkenntnisse zur Beobachterübereinstimmung Modifikationen der Vorgehensweise beschlossen wurden, bevor die Daten in einem zweiten Durchgang korrigiert und komplettiert wurden.

Beobachterübereinstimmung: So wie bei anderen sozialwissenschaftlichen Verfahren der Datenerhebung sollte auch bei Beobachtungsverfahren eine ausreichende Qualität der Daten gewährleistet werden. Zentrales Gütekriterium ist dabei die Beobachterübereinstimmung, auch Interrater-Reliabilität genannt (vgl. Bortz & Döring, 2006). Zur Prüfung der Beobachterübereinstimmung bieten sich je nach Dateneigenschaften unterschiedliche Kennwerte an (Wirtz & Caspar, 2002). Im Falle der Basiskodierung ist diesbezüglich sowohl die Übereinstimmung der Beobachter hinsichtlich der zeitlichen Lage und Dauer der Ereignisse als auch der Kategorienzuordnung von Bedeutung. Im Hinblick auf die zeitliche Segmentierung wurden in Anlehnung an die Vorgehensweise der IPN-Videostudie (Seidel, 2003b) die „Übereinstimmungen in der Markierung von In-Points (Beginn eines Ereignisses) und Out-Points (Ende eines Ereignisses)“ geprüft (S. 105), wobei ein Übereinstimmungsbereich von dreißig Sekunden als Toleranzgrenze gesetzt wurde. Die auf diese Weise berechnete

²⁷ Die Kodierung aller von diesem Punkt an dargestellten Beobachtungsdaten geschah mithilfe dieses Programms.

durchschnittliche paarweise Übereinstimmungsrate liegt mit 91 % (Min.: 83 %; Max.: 95 %) für die In- bzw. Out-Points auf einem guten Niveau²⁸. Die sich daraus ergebende Übereinstimmung hinsichtlich des zeitlichen Umfangs der Phasen ist mit 99 % (Min.: 94 %; Max.: 99 %) als sehr gut zu bezeichnen, ebenso wie die Übereinstimmung hinsichtlich der Zuordnung der Kategorien (97 %, Min.: 86 %; Max.: 100 %).

6.6.3 System zur Kodierung des Schülerbeteiligungsverhaltens

Die niedrig-inferente Kodierung des Schülerbeteiligungsverhaltens erfolgte anhand eines Zeitstichprobenplans (*time-sampling*, vgl. Bortz & Döring, 2006; Pauli, 2012), welcher die Beobachtung in vordefinierte Zeitabschnitte aufteilt. Die Länge der Zeitintervalle wurde auf eine Minute festgesetzt. Der Vorteil dieser als Intervallkodierung (*interval recording*, Bakeman & Quera, 2011, S. 30ff.) zu bezeichnenden Vorgehensweise liegt darin, dass das Schülerverhalten zeitlich lückenlos und – verglichen mit Ereignisstichproben – auf relativ praktikable Weise abgebildet werden kann. Zudem erhalten die Verhaltensdaten dadurch eine einheitliche Struktur, welche besonders bei großen Datenmengen spätere Auswertungen erheblich erleichtert (ebd.). Wie beschrieben wurde das Beobachtungssystem ausschließlich auf zuvor identifizierte Phasen des öffentlichen Unterrichts angewendet und die Schüler in diesen Phasen individualisiert beobachtet. Das bedeutet, dass alle sichtbaren Schüler einzeln und in jedem der einminütigen Beobachtungsintervalle hinsichtlich ihres Verhaltens kategorisiert wurden. Grundlage für die Beobachtungen war die Perspektive der Schülerkamera (s. Kap. 6.4), wobei in Fällen eingeschränkter Sichtbarkeit auf die Lehrerkamera zurückgegriffen werden konnte.

Bei der Ausgestaltung des Kodiersystems dienten die Arbeiten von Helmke und Kollegen (Helmke, 1988; Helmke & Renkl, 1992) zur methodologischen Orientierung. Entgegen dem Ansatz des in jenem Kontext entstandenen „Münchener Aufmerksamkeitsinventars“ (ebd.)²⁹ waren im vorliegenden Forschungszusammenhang (vgl. Kap. 6.2) verschiedene Intensitäten des Schülerbeteiligungsverhaltens allerdings

²⁸ Zum Vergleich: Die von der IPN-Forschergruppe angesetzte Mindestrate an übereinstimmenden Kodierungen von In- und Out-Points liegt bei 85 % (Seidel, 2003b, S. 105).

²⁹ Siehe auch Beschreibung des Instruments in Kap. 2.3.2.

nicht von unmittelbarem Interesse, sodass auch mit Blick auf Erkenntnisse zur Validität und Anwendbarkeit von Beobachtungssystemen zur Erfassung von Schülerverhalten im Unterricht an dieser Stelle das so genannte *molecular composite*-Konzept (Hoge, 1985) gewählt wurde. Das Schülerbeteiligungsverhalten innerhalb eines Beobachtungsintervalls wurde dementsprechend auf Grundlage von fünf Verhaltensindikatoren kodiert, deren Ausprägung getrennt voneinander eingeschätzt wurde und am Ende durch eine additive Verknüpfung in ein Urteil zum Verhalten des Schülers (*on/off task*) resultierte (vgl. Tab. 2). Die Kodierung der Verhaltensindikatoren geschah zu diesem Zweck dichotom (0/1), wobei aufgrund der wechselnden Positiv-/Negativrichtung der Indikatoren³⁰ der Wert „0“ für eine Tendenz hin zu *off* und „1“ in Richtung *on task* stand. Die Indikatoren wurden in Anlehnung an das System von Erhardt, Findeisen, Marinello und Reinartz-Wenzel (1981, zitiert nach Büttner & Schmidt-Atzert, 2004) zusammengesetzt und auf Basis der methodologischen Erkenntnisse von Helmke (1988) und Helmke und Renkl (1992) modifiziert.

Tab. 2: Aufbau des Kodiersystems zur Erfassung des Schülerbeteiligungsverhaltens

Verhaltensindikatoren	Urteil
1. Blickkontakt zum legitimen Sprecher oder Objekt	0/1
2. Aktive Beteiligung an der Aufgabe	0/1
3. Ausübung anderer Tätigkeiten	0/1
4. Motorische Unruhe	0/1
5. Themenferne Kommunikation	0/1
Additives Urteil zum Schülerbeteiligungsverhalten	überwiegend 1 (≥ 3) → <i>on task</i> überwiegend 0 (< 3) → <i>off task</i>

Maßgeblich für die Kodierentscheidungen zu den Verhaltensindikatoren war die überwiegende Ausprägung des jeweiligen Verhaltensaspekts innerhalb des Beobachtungsintervalls (*predominant activity sampling*, Bakeman & Quera, 2011, S. 33). Die Verhaltensindikatoren lassen sich wie folgt beschreiben:

³⁰ Das Vorliegen der mit den Indikatoren Nr. 1 und 2 bezeichneten Verhaltensweisen kann als „positiv“, d.h. im Sinne von *on-task* verstanden werden, während das Vorliegen des durch die Indikatoren Nr. 3 - 5 beschriebenen Verhaltens eher für einen *off-task*-Status spricht.

1. *Blickkontakt*: Der Schüler hält Blickkontakt zum aktuell legitimen Sprecher, welcher außer der Lehrkraft auch ein Mitschüler (bspw. während eines Wortbeitrags oder der Bearbeitung einer Aufgabe an der Tafel) oder andere Person (z.B. weitere Lehrkräfte oder Gäste mit Bezug zum Unterrichtsthema) sein kann, oder einem als themenrelevant zu verstehenden Objekt. Dies umfasst – sofern in der Situation vorhanden – sowohl das Arbeitsmaterial (Bücher, Hefte, Arbeitsblätter etc.), das Präsentationsmedium (Tafelbild, Projektionsfolie, Filmausschnitt etc.) als auch jegliche Art von Anschauungsgegenstand (in dieser Stichprobe z.B. Uhren, Kalender, Werbeprospekte).
2. *Aktive Beteiligung*: Der Schüler beteiligt sich durch aufgabenspezifisch geforderte Aktivitäten am Unterrichtsgeschehen. Die jeweilige Beteiligungserwartung wird dabei durch den Arbeitsauftrag oder eine Anweisung der Lehrkraft definiert. Dazu können Wortmeldungen und -beiträge, das Mitsprechen von Vokabeln im Chor oder das Vorlesen von Texten zählen. In Phasen, in denen eine aktive Beteiligung nicht explizit gefordert oder nicht möglich ist (z.B. Lehrervorträge, öffentliche Dialoge), wurde dieser Verhaltensindikator bereits durch eine „passive Beteiligung“ des Schülers, d.h. die Aufrechterhaltung des Blickkontakts bei gleichzeitigem Fehlen von aufgabenfernen Verhaltenskomponenten (z.B. andere Tätigkeiten oder themenferne Kommunikation, s.u.), als erfüllt angesehen.
3. *Ausübung anderer Tätigkeiten*: Der Schüler führt eine Tätigkeit aus, von der anzunehmen ist, dass sie die Aufmerksamkeit für die aktuelle Aufgabe deutlich reduziert oder gar verhindert, wie z.B. Essen, Nagelpflege, Comics zeichnen³¹ oder auch der Gang zum Abfalleimer. Als beiläufig zu interpretierende Verhaltensweisen (z.B. Fingerspielereien mit dem Stift) wurden hingegen nicht dazu gezählt. Sobald der Vollzug der Tätigkeit (oder die Summe verschiedener, aufeinander folgender Handlungen dieser Art) die überwiegende Zeit des Beobachtungsintervalls in Anspruch nahm, wurde das Vorliegen einer anderen Tätigkeit registriert.

³¹ Der Großteil der Videoaufnahmen liegt im HDV-Format (*High Definition Video*) vor und bietet damit eine ausreichend hohe Bildqualität, um gemeinsam mit Vergrößerungsoptionen des Videoanalyseprogramms ausreichend viele Details der Schülertätigkeiten erfassen zu können.

4. *Motorische Unruhe*: Dieser Indikator betrifft wiederholte körperliche Bewegungen, deren Intensität (oder Lautstärke) ein beiläufiges Maß überschreitet und von denen anzunehmen ist, dass sie die Hinwendung des Schülers zur aktuellen Lerngelegenheit beeinträchtigen, wie z.B. „Stuhlkippen“, „Zappeln“, Tischklopfen. Im Unterschied zur „Ausübung anderer Tätigkeiten“ (Indikator 3) wird hierbei keine als zielgerichtet zu beschreibende Handlung vollzogen.
5. *Themenferne Kommunikation*: Kommunizieren Schüler sicht- oder hörbar miteinander, obwohl es für sie durch die aktuelle Interaktionsstruktur nicht vorgesehen ist, wurde diese Kommunikation als themenfern gewertet. In Situationen mit unklaren Vorgaben seitens der Lehrkraft hinsichtlich der erwarteten bzw. geduldeten Kommunikationsarten in der Klasse wurde auch der Inhalt der Äußerungen, sofern dieser aus den Aufzeichnungen zu entnehmen war³², oder das sichtbare Verhalten der Schüler während des Kommunikationsvorgangs (z.B. offensichtliches Scherzen oder Provokationen) als Hinweis auf den Themenbezug der Kommunikation herangezogen. Berücksichtigt wurden auch erfolglose Versuche der Kommunikationsaufnahme, sofern sie zeitlich lange genug anhielten, ebenso wie Kommunikation mit Hilfe von Gestik und Mimik.

Übergreifende Kodierregeln: In Situationen, die trotz wiederholter Begutachtung des Beobachtungsintervalls und nach Ausschöpfen aller technischen Möglichkeiten (Abgleich beider Kameraperspektiven, Zoom-Funktion, Zeitlupe, Verstärkung der Audiospur etc.) keine zweifelsfreie Kodierentscheidung zuließen, wurde nach dem so genannten „in dubio pro discipulo“-Gebot vorgegangen, d.h. zugunsten des Schülers entschieden und der jeweilige Verhaltensindikator mit *on task*-Tendenz beurteilt. Dem liegt das Bestreben zugrunde, die Rate an fälschlicherweise getroffenen *off task*-Urteilen möglichst gering zu halten³³. Eine weitere Kodierregel betrifft die Sichtbarkeit der

³² Die Audiospuren der Aufzeichnungen wurden im Rahmen der technischen Postproduktion so bearbeitet, dass die Lautstärke von den mittels Raummikrofon (vgl. Kap. 6.4) aufgezeichneten Schüleräußerungen angehoben wurde. In Kombination mit der Verwendung von Studiokopfhörern ließen sich damit viele vermeintlich unverständliche Schüleräußerungen erschließen.

³³ Im Gegensatz zum Münchner Aufmerksamkeitsinventar, welches in solchen Fällen das *off-task*-Verhalten priorisiert und damit mögliche, durch offene Beobachtung im Unterricht verursachte Effekte sozial erwünschten Schülerverhaltens ausgleichen soll (Helmke 1988, S. 29).

Schüler: Aus verschiedenen Gründen (z.B. Wechsel der Sitzpositionen) waren manche Schüler zeitweise nicht ausreichend sichtbar. Die Kodierung des Beteiligungsverhaltens wurde dabei nur für jene Zeitintervalle vorgenommen, in denen ein Schüler zu mehr als 30 Sekunden am Stück beobachtbar war. Die Kodierung wurde zudem ausgesetzt, wenn keine auf Unterrichtsinhalte bezogene Verhaltenserwartung an die Schüler zu erkennen war oder außerordentliche Störsituationen oder Unterbrechungen des themenbezogenen Geschehens vorkamen (z.B. Besuch des Hausmeisters, Verletzungen von Schülern).

Durchführung der Kodierungen: Vor der Datenerfassung durchlief das Kodiersystem eine Erprobungsphase, an der das gesamte Forschungsteam beteiligt war. Auf Grundlage eines Kodieranweisungspapiers und einer Beobachterschulung wurde die erste Version des Systems auf einen Stichprobenausschnitt angewendet, erste Kennwerte der Urteilsübereinstimmung (s.u.) berechnet und Optimierungsbedarfe identifiziert. Daraufhin wurden die Kodiervorschriften vervollständigt (z.B. durch Regeln für Sonderfälle, Ankerbeispiele, technische Anleitungen für die Analysesoftware) und das für die Hauptanalysen vorgesehene Beobacherteam erneut geschult. Die letztendlich gültige, vollständige Kodierung der 23 Unterrichtsstunden wurde schließlich von einem Team aus sieben Beobachtern durchgeführt.

Beobachterübereinstimmungen: Aufgrund ungleicher Häufigkeitsverteilungen der beiden Verhaltenskategorien *on* bzw. *off task* (s. Kap. 7.2.2) wurde unter Bezugnahme auf Wirtz und Caspar (2002) eine Sonderform der kategorienspezifischen prozentualen Übereinstimmung, die so genannte bedingte Prozentzahlerwartung (P_{SF+} bzw. P_{SF-}) als Übereinstimmungsmaß herangezogen (S. 88). Bei selten auftretenden Ereignissen – im vorliegenden Kontext die Kategorie *off task* – wird empfohlen, diese Größen dem ansonsten gängigen Kennwert Cohens κ vorzuziehen, da dessen Höhe „gerade bei sehr schiefen Randverteilungen irreführend sein“ kann (S. 91). Die vorliegenden, auf Basis von ca. 1.600 doppelt kodierten Beobachtungsintervallen³⁴ berechneten Werte veranschaulichen dies: Während Cohens κ nach allgemeinen Konventionen für eine

³⁴ Dies entspricht ca. 13 % der gesamten Beobachtungsstichprobe (vgl. Kap. 7.2.1) und wurde ermittelt anhand sieben paarweiser Übereinstimmungen, unter Einbezug jeder Kodierperson in zwei verschiedenen Paaren. Die Prüfwerte ließen an keiner Stelle die Notwendigkeit zur Herausnahme einer Kodierperson aus den Analysen erkennen.

lediglich als akzeptabel zu bezeichnende Übereinstimmung zwischen den Beobachtern spricht (ebd., S. 59), fallen die alternativen Größen PÜ und PÜ_{SF} zufriedenstellend bis gut aus (Tab. 3).

Tab. 3: Kennwerte der Beobachterübereinstimmung zur Kodierung des Schülerbeteiligungsverhaltens

Übereinstimmungsmaße	Md^a	Min./Max.
Cohens κ	0,46	0,44 / 0,49
PÜ	85 %	72 % / 93 %
PÜ _{SF}	90 %	77 % / 96 %

^a Md = Median über alle Übereinstimmungspaare

6.7 Instrument zur beobachtungs-basierten Kodierung und Kategorisierung der Aufgaben

Dieser Abschnitt beschreibt die Vorgehensweise zur Analyse der Aufgaben im Unterricht. Um eine differenzierte Beschreibung von Aufgabencharakteristika vornehmen zu können, ist zuerst die Identifizierung der Analyseeinheit „Aufgabe“ erforderlich. Daher erfolgt in Kap. 6.7.1 die Darstellung der Vorgehensweise zur Erfassung jener Geschehenseinheiten im Unterricht, die entsprechend der Ausführungen in Kap. 4 als Aufgaben zu verstehen sind. Die Beschreibung eines für die hier verfolgten Untersuchungszwecke geeigneten Kategoriensystems und dessen Konstruktion ist Gegenstand von Kapitel 6.7.2.

6.7.1 Kodierung der Analyseeinheit „Aufgabe“

Um das Angebot an Aufgaben im Unterricht möglichst genau und umfassend abzubilden und dabei die zeitliche Dimension angemessen zu berücksichtigen, wurde wie auch bei der Basiskodierung der Sozialformen (s. Kap. 6.6.2) die Verfahrensweise der kodierenden Beobachtung anhand einer Ereignisstichprobe gewählt. Dabei werden die vorab festgelegten Ereignisse durch die Setzung von Anfangs- und Endzeitpunkten in ihrer Lage und Ausdehnung im Geschehensstrom festgehalten. In Anschluss an den unter Kap. 4.2 dargestellten Aufgabenbegriff wurden folgende Ereignisse als Aufgaben definiert:

Als Aufgaben gelten alle an die Schüler gerichteten öffentlichen Denk- und Handlungsaufforderungen, die sich durch einen Bezug zu den Unterrichtsinhalten

und eine Zielgerichtetheit, d.h. die geforderte Erzeugung eines Produkts oder den Vollzug von Verhaltens- und Handlungsweisen, auszeichnen.

Diese weitgehend an dem Verständnis von Renkl (1991) orientierte und um Aspekte der Auffassung von Doyle (1979) und ferner Jatzwauk, Rumann und Sandmann (2008) ergänzte Definition umfasst damit auch kleinste Geschehensbestandteile des Unterrichts, wie etwa Lehrerfragen. Bezug zu den Unterrichtsinhalten bedeutet, dass die Aufgabe sich im Rahmen des aktuellen Unterrichtsthemas oder eines damit in Verbindung stehenden Themas, wie z.B. einem (möglicherweise auch fächerübergreifenden) Exkurs oder einer Rück- oder Vorschau der Unterrichtsinhalte, bewegt. War dieser Bezug nicht erkennbar, wurde das Ereignis nicht als Aufgabe festgehalten. Die Aufgaben wurden bei der Kodierung in die beiden Phasen Aufgabenstellung und Aufgabenbearbeitung aufgeteilt. Dies hat den Zweck, in der Auswertung zeitliche Verhältnisse dieser Teile berücksichtigen zu können.

Durchführung der Kodierungen: Auf Basis dieser Vorgaben wurde eine Kodiervorschrift erstellt, die genaue Angaben zur Identifizierung der Ereignisse und ihrer Anfangs- und Endzeiten enthielt. So wurde z.B. das Ende der Aufgabenstellung und der Beginn der Aufgabenbearbeitung gesetzt, wenn die Formulierung der Anweisungen und die Präsentation der zur Bearbeitung nötigen Informationen abgeschlossen ist, indem die Lehrkraft dies durch eine erkennbare Pause in ihren Ausführungen und eine wartende bzw. fragende Haltung zum Ausdruck bringt und sie gleichzeitig die Aufnahme der Bearbeitung durch die Schüler akzeptiert bzw. erste Wortmeldungen annimmt. Maßgeblich für die Kodierung der Aufgaben war das Handeln der Lehrkraft, sodass die Perspektive der Lehrerkamera die primäre Datengrundlage für diesen Analyseschritt bildete, und nur in Einzelfällen die Perspektive der Schülerkamera als Ergänzung herangezogen werden musste. Die Ereigniskodierung der Aufgaben, durchgeführt von drei Personen, umfasste drei Durchgänge: (1) Im ersten Durchgang wurde das gesamte Videomaterial anhand der beschriebenen Ereignisdefinition und nach einer einmaligen Beobachterschulung durchkodiert, wobei zum Zwecke der Prüfung von Übereinstimmungen ein Drittel des Materials zur Parallelkodierung durch zwei oder drei Personen vergeben wurde. Darauf folgte ein Abgleich der Ergebnisse und die Prüfung von Übereinstimmungen, woraus Rückschlüsse für die Präzisierung der Kodiervorschriften und Korrekturen der Daten gezogen wurden. (2) Der zweite Durchgang diente der Prüfung aller Kodierungen und – sofern nötig – der Korrektur der

Daten. (3) Der dritte Durchgang hatte eine weitere Präzisierung der Kodierungen zum Inhalt und fand im Rahmen der Kategorisierung der Aufgabenmerkmale statt (s. Kap. 6.6.2). Ziel dessen war die Trennung aufeinanderfolgender Aufgabenereignisse, sobald bei mindestens einem der kategorialen Aufgabenmerkmale eine Ausprägungsvariation (z.B. Medienwechsel oder Erhöhung des Lebensweltbezugs) oder eine bedeutsame Veränderung der Informationsgrundlage durch die Lehrkraft zu erkennen war. Sofern keine dieser Voraussetzungen erfüllt war, wurden z.B. wiederholte Lehrerfragen als Fortsetzung derselben Aufgabe und damit als ein Ereignis gewertet.

Beobachterübereinstimmungen: Im vorliegenden Falle, wo eine einzelne Ereigniskategorie im *event-sampling*-Verfahren kodiert wird, ist mit Blick auf das Gütekriterium der Interrater-Reliabilität von vorrangiger Bedeutung, ob die Beobachter die Ereignisse übereinstimmend denselben Zeitpunkten zuordnen. Daher wurde an dieser Stelle zur Übereinstimmungsprüfung die bereits bei der Kodierung der Sozialformen (s. Kap. 6.6.2) angewendete Vorgehensweise genutzt. Dabei wurde unter Berücksichtigung eines zeitlichen Toleranzbereichs von zehn Sekunden der Anteil an Übereinstimmungen bei der Setzung der Anfangs- und Endzeiten der Ereignisse geprüft. Das Ergebnis zeigt eine sehr gute prozentuale Übereinstimmungsrate von 93 %.

6.7.2 Kategoriensystem zur Beschreibung der Aufgabenmerkmale

Die zuvor dargestellte Ereigniskodierung der Aufgaben bildet die Grundlage für den nächsten Analyseschritt, in dessen Rahmen die Eigenschaften der Aufgaben anhand eines Kategoriensystems zu beschreiben waren. Wie gezeigt, liegt in der Unterrichtsforschung eine Reihe an Instrumentarien zur Analyse von Aufgaben vor, von denen sich allerdings aus verschiedenen Gründen keines zur direkten Verwendung in dieser Studie anbietet (vgl. Kap. 4.3/4.4). Für diese Untersuchung wird ein Instrument benötigt, welches ausgewählte, angesichts der angezielten Desiderate und Fragestellungen geeignete Aufgabeneigenschaften in den Blick nimmt. Unter Berücksichtigung des Standes der wissenschaftlichen Diskussion zu unterrichtlichen Bedingungsfaktoren des Schülerbeteiligungsverhaltens (vgl. Kap. 2.3.4) sind diese zu sehen in der Verständlichkeit, Klarheit und Interessanztheit der Anweisungen und inhaltlichen Darstellungen, der Herstellung von Bezügen zur Lebenswelt der Schüler und der Sicherstellung eines optimalen Anforderungsniveaus. Weiterhin sollte das System auf die hier als Aufgabe verstandenen Analyseeinheiten – sprich auch kleinste

Geschehensbestandteile wie z.B. Lehrerfragen – anwendbar sein. Außerdem sollte durch die Art der Ausformulierung der Analysekategorien eine ausreichende Standardisierung der Daten gewährleistet sein, was mit Blick auf die quantitativ orientierte Auswertungsstrategie von Bedeutung ist.

Vor diesem Hintergrund wurde auf Grundlage bereits bestehender Instrumentarien ein eigenes Kategoriensystem zusammengestellt. Die Vorgehensweise bei der Konstruktion, Erprobung und Anwendung geschah entlang der folgenden Schritte:

(1) Zuerst wurde ein systematischer Vergleich von ausgewählten Instrumenten zur Aufgabenanalyse hinsichtlich der zuvor dargelegten Anforderungen dieser Studie vorgenommen. Dies geschah mithilfe einer Synopse der Analysekategorien bzw. -bestandteile und umfasste acht der in Kap. 4.3 beschriebenen Systeme.

(2) Auf dieser Basis wurde eine 34 Aufgabenmerkmale umfassende Vorversion konstruiert, von drei Kodierpersonen auf einen Stichprobenausschnitt von sechs Unterrichtsstunden angewendet und aufgrund der Anwendungserfahrungen und Urteilsübereinstimmungswerte reduziert. So konnten beispielsweise zum Merkmalsbereich der sprachlichen Qualität von Anweisungen und Aufgabenstellung mit keiner der vorliegenden Kategorisierungsvarianten (vgl. z.B. „sprachlogische Komplexität“, Maier, Kleinknecht, Metz & Bohl, 2010; *input language*, Bachman & Palmer, 2010) reliable Urteile erzielt werden, weshalb dieser Bereich gestrichen wurde. Eine zweite, 22 Aufgabenmerkmale umfassende Vorversion des Systems wurde auf die gleiche Weise erprobt und nochmals reduziert.

(3) Die daraus generierte, letztendliche Zusammenstellung von 14 Analysekategorien wurde im nächsten Schritt unter Beibehaltung des Kodierpersonals auf die 16 Unterrichtsstunden umfassende Stichprobe der Aufgabenanalyse angewendet³⁵. Tab. 4 liefert einen Überblick über das Kategoriensystem³⁶.

³⁵ Die Darstellung und Begründung dieser Stichprobenauswahl erfolgt in Kap. 7.3.1.

³⁶ Obgleich vier der aufgeführten Aspekte keine kategoriale Erfassung der Merkmalsausprägung vorsehen (verbaler Umfang der Aufgabenstellung, Anzahl an Hilfestellungsstrategien in der Aufgabenstellung, verbaler Umfang der Medieninhalte, zeitlicher Umfang der Aufgabenstellung und -bearbeitung), wird das Instrument an dieser Stelle aufgrund seiner Grundausrichtung trotzdem als Kategoriensystem bezeichnet.

Tab. 4: Überblick über das Kategoriensystem zur Beschreibung der Aufgaben im Unterricht

Aufgabenmerkmal	Ausprägungen
<i>Einführung und Präsentation der Aufgabe</i>	
Sprache	Deutsch – Englisch – Kombination – andere
verbaler Umfang	Anzahl der Wörter
Hilfestellungsmaßnahmen	Anzahl der Maßnahmen: Wiederholung der Anweisungen – Verständnisfragen – Vorwissenaktivierung – Hinweise zum Arbeitsablauf – Nennung von Beispielen – Lösungshinweise – Offenlegung der Bewertungskriterien
Medien: Darstellungsform	Text – Tabelle/Diagramm – Abbildung/Grafik – Film – Hörbeispiel – Gegenstand – Kombination – keine – andere
Medien: Gerät/Material	Tafel – Projektionsfolie – Arbeitsblatt – Buch – PC – Kombination – keines – andere
Bekanntheit des Mediums	bereits bekannt – simultan eingeführt – angekündigt
Umfang der Medieninhalte	Anzahl der Wörter
<i>Aufgabeninhalte</i>	
Lebensweltbezug	ohne – konstruiert – authentisch – real
Abstraktheit	konkret/alltäglich – abstrakter/nicht alltäglich – abstrakt
<i>Anforderungen der Bearbeitung</i>	
Kognitiver Prozess	Erinnern – Verstehen – Anwenden – Analysieren – Evaluieren – Gestalten
Wissensdimension	Fakten – Konzepte – Prozeduren – metakognitive Strategien
Format des geforderten Produkts	Verarbeitung von Vorgaben – eingeschränkte Konstruktion – erweiterte Konstruktion
Art des geforderten Produkts	Sprache – Schrift – Grafik – Kombination – andere
Zeitlicher Umfang	Zeitlicher Umfang von Aufgabenstellung und -bearbeitung in Sekunden
Reflexions-/Auswertungsfunktion	Keine Reflexions-/Auswertungsfunktion – Refl./Ausw. einer vorherigen Unterrichtsaufgabe – Refl./Ausw. einer Hausaufgabe

Die Aufgabenmerkmale sind bis auf zwei Ausnahmen den übergeordneten Bereichen *Einführung und Präsentation der Aufgabe*, *Aufgabeninhalte* und *kognitiver Anspruch der Bearbeitung* zugeordnet und werden im Folgenden beschrieben.

Merkmalsbereich: Einführung und Präsentation der Aufgabe

Die Analyseaspekte innerhalb dieses Bereichs betreffen die Art und Weise, wie die Aufgabe von der Lehrkraft in den Unterricht eingebracht wird, z.B. durch die Formulierung der Arbeitsanweisungen und die Präsentation der Informationen. An dieser Stelle wird in Anlehnung an Slavin (1994) davon ausgegangen, dass diese Komponenten die Qualität der Darstellung der Lerngelegenheiten repräsentieren und damit Anreize zur individuellen Beteiligung setzen.

- *Sprache der Einführung und Präsentation (Quelle/Anlehnung an: Bachman & Palmer, 2010)*: Da es sich bei der untersuchten Stichprobe um Englischunterricht handelt, wird es als bedeutsam angesehen, in welcher Sprache die Lehrkraft die Aufgabe in den Unterricht einführt bzw. präsentiert, da dies Einfluss darauf nimmt, inwiefern die Schüler mit ihren unterschiedlichen Leistungsständen in der Lage sind, die Aufgabenstellung zu verstehen und dementsprechend auch erfolgreich zu bearbeiten. Mit dieser von Bachman und Palmer (2010, S. 74) auf den gegebenen Kontext angepassten Kategorie wurde erfasst, ob die Aufgabenstellung und die Anweisungen in der allgemeinen Verkehrssprache (Deutsch) oder der unterrichteten Fremdsprache (Englisch) formuliert wurde. Bei einer gemischten Sprachverwendung war entscheidend, welche Sprache in der Äußerung quantitativ überwiegt (gemessen an der Anzahl der Wörter, unter Zuhilfenahme der Verbaltranskriptionen) oder – bei annähernder Gleichgewichtung der Sprachverwendung – in welcher Sprache der Bedeutungskern der Aussage getroffen wurde. Die Ausprägung *Kombination* wurde somit nur dann vergeben, wenn die Äußerung ihren Bedeutungsgehalt zu gleichen Teilen aus englischsprachigen und deutschsprachigen Bestandteilen zog (z.B. „Als nächstes öffnet ihr bitte your textbooks on page 38.“³⁷). Eine Äußerung wie „And now möchte ich mit euch eine Diskussion führen.“ hingegen wurde als Deutsch kategorisiert. Die Kategorie *Andere* beinhaltet alle anderen Sprachen außer Deutsch und Englisch.
- *Verbaler Umfang der Einführung und Präsentation (Quelle/Anlehnung an: Bachman & Palmer, 2010)*: Die Länge der Anweisungen zur Aufgabe beeinflusst das Ausmaß der bei der Rezeption zu leistenden Informationsverarbeitung für

³⁷ Nicht näher gekennzeichnete Zitate dieser Art entstammen den Verbaltranskriptionen.

die Schüler. Da Aufgaben in der hier fokussierten Sozialform des öffentlichen Unterrichts zumeist mündlich eingeführt werden, wird durch dieses Analysemerkmal die Anzahl der Wörter während der Einführung und Erklärung der Aufgabenstellung durch die Lehrkraft erfasst. In der Kodierung wurden kurze verbale Exkurse der Lehrkraft, wie z.B. Ermahnungen von Schülern oder organisatorische Zwischeninformationen (z.B. „Ja, lass das Buch ruhig draußen.“) mitgezählt, da in dieser Situation auch alle anderen Schüler diese Äußerungen in der Erwartung auf relevante Informationen mit anhören und verarbeiten müssen. Ausgedehnte Abschweifungen von mehr als einem Satz wurden hingegen als Themenwechsel angesehen und aus der Wortzählung herausgenommen. Wie bereits bei dem Merkmal *Sprache der Einführung und Präsentation der Aufgabe* wurde für eine exakte Zählung auch hier zusätzlich auf die Verbaltranskriptionen zurückgegriffen.

- *Hilfestellungsmaßnahmen (Eigenkonstruktion)*: Dieser Aspekt fokussiert unterstützende Elemente innerhalb der Anweisungen und Erläuterungen der Lehrkraft, die den Schülern das Verständnis der Aufgabenstellung und die darauf folgende Bearbeitung erleichtern sollen. Das können Verweise der Lehrkraft auf bereits bekannte Sachverhalte oder analoge Themen sein (Vorwissenaktivierung, vgl. Europarat, 2001), Hilfen zur Strukturierung des Arbeitsprozesses, Hinweise zum Lösungsweg, die Offenlegung der Bewertungskriterien (d.h. Konkretisierung der Ergebniserwartungen zur Einengung des Lösungsspektrums, vgl. Bachman und Palmer, 2010), eine Wiederholung bzw. Paraphrasierung der Aufgabenstellung und die Möglichkeit für die Schüler, Verständnisfragen zu stellen (Europarat, 2001; Bachman und Palmer, 2010). Da es sich hierbei nicht um disjunkte Ausprägungsstufen handelt, wurde die Anzahl der Hilfestellungsstrategien bzw. -hinweise in den Erläuterungen und Anweisungen der Lehrkraft erfasst (Wertebereich 0 - 7).
- *Medieneinsatz (Quelle/Anlehnung an: Neubrand, 2002)*: Dieser Bereich spricht durch mehrere Teilaspekte jegliche Art des Medieneinsatzes – ob zur Unterstützung der verbalen Ausführungen der Lehrkraft oder als Träger der Aufgabeninhalte – in der Aufgabenstellung an. Dabei wurde zum einen die vermittelte *Darstellungsform* festgehalten (z.B. Text oder Abbildungen, s. Tab. 4), weil davon ausgegangen wird, dass dies sowohl den kognitiven Anspruch der Rezeption als auch die Attraktivität der Anschauung beeinflusst, und zum

anderen das jeweilige *Gerät bzw. Material*, welches diese Darstellungsform trägt (z.B. Tafel oder Arbeitsblatt), weil damit die physische Verfügbarkeit innerhalb des Klassenraums variieren kann (z.B. Verstehbarkeit von Hörbeispielen, Sichtbarkeit von Projektionsfolien).

In diesem Zusammenhang wird auch die *Bekanntheit des Mediums* zum Zeitpunkt der Aufgabenstellung als potenziell bedeutsam angesehen. Dieser Teilaspekt ergibt sich aus der kleinteiligen Sichtweise auf die als Aufgaben angesehenen Analyseeinheiten und berücksichtigt den Umstand, dass nicht alle Medien simultan zu einer neuen Aufgabenstellung in den Unterricht eingebracht werden (z.B. Auflegen der Projektionsfolie während den mündlichen Erläuterungen durch die Lehrkraft). Manche Medien sind Gegenstand mehrerer, aufeinanderfolgender Aufgaben oder werden zu einem späteren Zeitpunkt im Unterrichtsverlauf erneut aufgegriffen und thematisiert, wohingegen andere in der Aufgabenstellung angekündigt und erst im Verlauf der Bearbeitung offenbart werden (z.B. Hörbeispiel von CD, Tafelrückseite).

Da bei Einbringen eines Mediums die dargestellten Inhalte für die Schüler zusätzlich zu den verbalen Ausführungen der Lehrkraft zum Gegenstand einer Verstehensleistung werden müssen, wurde auch der *verbale Umfang der Medieninhalte* mittels Wortzählung erfasst.

Merkmalsbereich: Aufgabeninhalte

Dieser Bereich fokussiert die durch die Aufgabe angesprochenen Inhalte, so wie sie in den Erläuterungen der Lehrkraft oder den vorhandenen schriftlichen, visuellen oder auditiven Informationen (s.o., Merkmalsbereich Medieneinsatz) zu erkennen sind. In Anlehnung an den Grundgedanken des Anreizwerts des Unterrichts bei Slavin (1994, S. 147) ist anzunehmen, dass diesbezüglich besonders die von den Schülern subjektiv wahrgenommene Sinnhaftigkeit der Inhalte eine zentrale motivierende Wirkung auf deren Beteiligungsverhalten ausübt.

- *Lebensweltbezug (Quelle/Anlehnung an: Maier, Kleinknecht, Metz & Bohl, 2010):* Diese Kategorie erfasst den Grad der Herstellung von Bezügen der Aufgabeninhalte zum lebensweltlichen Erfahrungsraum der Schüler. Aufgaben *ohne Lebensweltbezug* verbleiben meist auf der Ebene rein fachlicher Phänomene

(z.B. Fragen nach verschiedenen Winkelarten in der Mathematik oder die Identifikation von Wortarten im Fach Deutsch, ebd.). Aufgaben mit *konstruiertem Lebensweltbezug* stellen zwar entsprechende Verknüpfungen her, diese sind aber künstlich hergestellt, wirken mitunter wie „an den Haaren herbeigezogen“ (Maier, Kleinknecht, Metz & Bohl, 2010, S. 37) und lassen Analogien der Inhalte zu tatsächlichen Erfahrungen der Schüler relativ unwahrscheinlich erscheinen (z.B. Texte über Erlebnisse fiktiver jugendlicher Charaktere in Lehrbüchern des Fachs Englisch). Aufgaben mit *authentischem Lebensweltbezug* sind deutlicher am tatsächlichen Erfahrungshorizont der Schüler orientiert, indem die Inhalte auf den (mutmaßlichen) Alltag der Schüler übertragbar oder diesem gezielt entnommen sind (z.B. Gespräche über persönliche Vorlieben, Erlebnisberichte über Ferienaktivitäten). Bei Aufgaben mit *realem Lebensweltbezug* hingegen ist kaum eine Differenz zwischen Aufgabe und Lebenswelt erkennbar, indem die Schüler sich mit einer wirklichen Problemstellung auseinandersetzen (z.B. das Führen eines Klassenkontos, ebd., S. 37, oder mit Blick auf das Unterrichtsfach der vorliegenden Stichprobe der Schriftwechsel mit einer englischsprachigen Brieffreundin).

- *Abstraktheit (Quelle/Anlehnung an: Nold & Rossa, 2007)*: Dieser aus dem Kontext der DESI-Studie übernommene Analyseaspekt beschreibt den Grad der inhaltlichen Abstraktheit der Aufgabe. Das Merkmal weist damit eine gewisse Nähe zum Lebensweltbezug auf, thematisiert aber im Unterschied dazu nicht die Wahrscheinlichkeit möglicher Verknüpfungen zum Erfahrungsraum der Schüler, sondern die Art der Verankerung der Inhalte im alltäglichen Leben. *Konkrete* Aufgaben weisen Bezüge zu greifbaren bzw. beobachtbaren Aspekten des Alltags auf, wie z.B. bestimmte Ereignisse oder persönliche Erlebnisse. *Leicht abstrakte* Aufgaben bewegen sich dagegen auf der Ebene nicht alltäglicher Ereignisse (z.B. Fiktives, Mystisches, seltene Ereignisse) oder wenig greifbarer Bestandteile des Alltags, wie z.B. emotionale Reaktionen. Aufgaben mit gänzlich *abstrakten Inhalten* beziehen sich auf im Alltag nicht direkt wahrnehmbare Aspekte wie z.B. Meinungen, Einstellungen oder logische Strukturen.

Merkmalsbereich: Anforderungen der Bearbeitung

Dieser Bereich umfasst verschiedene Aspekte der anhand der Aufgabenstellung erkennbaren Anforderungen an die Schüler bei der Bearbeitung der Aufgabe. Unter Rückgriff auf Slavin (1994, S. 147) wird davon ausgegangen, dass diese überwiegend auf kognitiver Ebene angesiedelten Faktoren die Bewältigungschancen für den einzelnen Schüler beeinflussen und damit eine Wirkung auf dessen individuelle Bereitschaft zur Partizipation an der Lerngelegenheit ausüben.

- *Kognitiver Prozess (Quelle/Anlehnung an: Anderson et al., 2001)*: Dieser Aspekt bildet neben den Wissensdimensionen (s.u.) eine von zwei Dimensionen der Lernzieltaxonomie von Anderson et al. (2001) und erfasst unterschiedliche kognitive Prozesse, die zum Erreichen des Aufgabenziels bzw. zur Erzeugung des geforderten Produkts zu vollziehen sind. Die Ausprägungsstufe *erinnern* umfasst Aufgaben, bei denen das Wiedererkennen und Abrufen von Informationen aus dem Gedächtnis im Vordergrund steht (z.B. „*Wer kann jetzt mal von zehn bis hundert nur die Zehnerzahlen, zehn, zwanzig, dreißig, sagen?*“). *Verstehen* hingegen beinhaltet Facetten der Herstellung oder Extraktion von Bedeutung in Bezug auf die gegebenen Informationen, wie etwa Paraphrasierungen und Zusammenfassungen, Klassifizierungen, Schlussfolgerungen, Vergleiche (z.B. die Zusammenfassung einer Geschichte oder der Vergleich von Verbformen). Prozesse des *Anwendens* zielen auf die Nutzung und Ausführung von Prozeduren (z.B. Durchführung von Operationen zur Bestimmung von Satzgliedern, vgl. Bremerich-Vos, 2008), wozu im vorliegenden Kontext auch die richtige Aussprache und grammatikalisch korrekte Bildung von Sätzen gezählt wurde (ebd.). *Analysieren* umfasst die Unterscheidung verschiedener Bestandteile eines Sachverhalts, deren Beziehungen zueinander und zu einem möglichen übergeordneten Zusammenhang (z.B. die Rekonstruktion des zeitlichen Geschehensablaufs in einer Geschichte und die Selektion bestimmter Informationen daraus vor dem Hintergrund einer Fragestellung). Die Ausprägung *evaluieren* betrifft Aufgaben, bei denen Bewertungen vorzunehmen bzw. Urteile vor dem Hintergrund bestimmter Kriterien zu fällen sind (z.B. die Bewertung der Vorgehensweise eines fiktiven Charakters im Lehrbuch bei der Findung von preisgünstigen Verbindungen des öffentlichen Nahverkehrs im Urlaub). Die Ausprägungsstufe *gestalten* repräsentiert schließlich Prozesse, bei denen

eigenständig ein bisher in dieser Zusammenstellung nicht vorliegendes, sinnvolles Produkt generiert wird (z.B. die Konzeption einer Werbeanzeige für ein selbst ausgedachtes Produkt, neue Aufgaben für die Mitschüler ausdenken). Entscheidend für die Kodierung dieses Merkmals waren alleine die aus den Äußerungen der Lehrkraft zu erkennenden Absichten hinsichtlich des offenbar beabsichtigten kognitiven Prozesses.

- *Wissensdimension (Quelle/Anlehnung an: Anderson et al., 2001)*: Diese ebenfalls aus der Lernzieltaxonomie von Anderson et al. (2001) übernommene Kategorie beschreibt die Art der Wissensbestände, die zur Bearbeitung der Aufgabe zu aktivieren sind. *Faktenwissen* umfasst dabei basale Begriffe oder einzelne Elemente eines Gegenstandsbereichs (z.B. Vokabeln, Details einer Geschichte). *Konzeptwissen* beinhaltet Kenntnisse über Strukturen, Schemata und Modelle und deren funktionale Zusammensetzung. Bezogen auf die Untersuchungsstichprobe betrifft dies z.B. grammatische Regeln oder Wortarten (vgl. Bremerich-Vos, 2008), aber auch in den Anschauungsinhalten repräsentierte Konzepte wie etwa das Prinzip des Wechselkurses („*And the exchange rate, can you explain it in german?*“). Die Ausprägungsstufe *prozedurales Wissen* steht für Kenntnisse über Teilschritte und Anwendungsvoraussetzungen von Methoden und Techniken, wie z.B. die Schreibung und Aussprache von Ordnungszahlen („*Was wird auch gesprochen, aber nicht geschrieben?*“). *Metakognitives Wissen* schließlich umfasst Kenntnisse über Denkprozesse im Allgemeinen und das Bewusstsein der eigenen Denkprozesse im Speziellen, darunter auch strategisches Wissen über Lernvorgänge (z.B. „*Und wenn du ein Wort nicht weißt im Text, was machst du da?*“). Es bleibt zu erwähnen, dass mit Blick auf die konzeptionelle Nähe zum ebenfalls auf Anderson et al. (2001) zurückgehenden Merkmal der kognitiven Prozesse die zweiseitigen Ausprägungskombinationen *erinnern von Fakten, verstehen von Konzepten* und *anwenden von Prozeduren* zwar als nahe liegend, nicht aber zwingend zu betrachten sind (ebd., S. 107).
- *Format des geforderten Produkts (Quelle/Anlehnung an: Bachman & Palmer, 2010)*: Dieses Merkmal weist Ähnlichkeit zum Konstrukt der Offenheit bzw. Konvergenz und Divergenz von Aufgabenstellung und Lösungswegen auf (vgl. Blömeke, Risse, Müller, Eichler & Schulz, 2006; Maier, Kleinknecht, Metz & Bohl, 2010), ist demgegenüber aber stärker ausgelegt auf die Betrachtung kleinteiliger

Aufgabeneinheiten. Das Merkmal beschreibt das Format des von den Schülern im Rahmen der Aufgabenbearbeitung zu erzeugenden Produkts (z.B. einer Antwort) anhand der Art der erforderlichen eigenen Konstruktionsleistung. Die Ausprägungsstufe *Verarbeitung von Vorgaben* umfasst Aufgaben, die eine Auswahl vorgegebener Elemente mit einer bestimmten Zielsetzung (z.B. Kennzeichnung von wahren Aussagen oder grammatikalisch korrekten Formulierungen), eine Zuordnung bzw. Sortierung vorgegebener Elemente nach bestimmten Kriterien (z.B. Ausfüllen von Lückentexten mit Wortvorgaben, chronologische Anordnung von Ereignissen einer Geschichte) oder eine wörtliche Wiedergabe von Informationen (z.B. Zitierung von Textstellen, Nennung von Daten oder Namen) verlangen. Bei der *eingeschränkten Konstruktion* hingegen liegen keine zu übernehmenden Elemente vor, sondern es ist die eigenständige Konstruktion eines Produkts erforderlich, wobei der Umfang auf ein Wort, eine Formulierung oder einen Satz beschränkt ist (z.B. „Which date do we have today?“). Die *erweiterte Konstruktion* unterscheidet sich davon durch einen geforderten Produktumfang von mehr als einem Wort, Satz oder einer Formulierung. In der Regel betrifft dies Aufträge zur Zusammenfassung, Nacherzählung oder Stellungnahme bzw. Diskussionsbeiträge jeglicher Art („Now, what about yesterday for you? What was it like?“).

- *Art des geforderten Produkts (Quelle/Anlehnung an: Bachman & Palmer, 2010):* Durch diese Kategorie wurde erfasst, auf welche Art und Weise das geforderte Produkt geäußert bzw. fixiert werden sollte. Dies erscheint im vorliegenden Kontext insofern bedeutsam, dass damit unterschiedliche Handlungsformen und somit auch Fähigkeitsbereiche angesprochen werden, aus denen interindividuell differente Herausforderungsniveaus für die Schüler resultieren. Das Merkmal beinhaltet die Ausprägungsstufen *Sprache* (z.B. mündliche Antworten oder Dialoge), *Schrift* (z.B. Tafelanschriften oder Notizen), *Grafik* (z.B. Zeichnungen oder Skizzen), *anderen Formen* (z.B. Handzeichen oder Bewegungen im Raum zum Zwecke von Gruppenzuordnungsspielen) und *Kombination*, was bedeutet, dass mindestens zwei der genannten Arten gemeinsam gefordert werden (z.B. eine Zeichnung an der Tafel anfertigen und kommentieren).

Weitere Merkmale:

- *Zeitlicher Umfang der Aufgabe (Eigenkonstruktion):* Durch die Erfassung der Aufgaben in ihrem tatsächlichen Auftreten innerhalb des Unterrichtsgeschehens mittels Ereignisstichprobe (s. Kap. 6.7.1) liegen zu jeder Aufgabe sekundengenaue Angaben zu ihrer zeitlichen Lage (Anfangs- und Endzeitpunkt) und Dauer vor, und dies differenziert nach den Phasen Aufgabenstellung und Aufgabebearbeitung. Die zeitlichen Eigenschaften sind keinem der oben genannten Aufgabenmerkmalsbereiche zugeordnet, da jene Bereiche qualitative Eigenschaften der Lerngelegenheiten darstellen und mit Slavin (1994) theoretisch begründet sind als Faktoren mit potenziell motivierender bzw. aktivierender Wirkung auf das Schülerbeteiligungsverhalten, wohingegen der zeitliche Umfang die Quantität der Aufgaben abbildet. Dennoch werden die zeitlichen Aspekte mit berücksichtigt, um im weiteren Verlauf die anderen Aufgabenmerkmale in Kombination mit der zeitlichen Dimension auswerten bzw. interpretieren zu können.
- *Reflexions- bzw. Auswertungsfunktion der Aufgabe (Eigenkonstruktion):* Dieses Merkmal wurde im Rahmen der Erprobungsphase in das Instrument aufgenommen, um die Identifikation von solchen Aufgaben zu ermöglichen, die die Funktion der Reflexion oder Auswertung einer bereits in der Vergangenheit bearbeiteten Aufgabe (Hausaufgabe oder zurückliegende unterrichtliche Lernaufgabe) erfüllen. Dieser Aspekt ist insofern bedeutsam für die vorliegende Untersuchung, dass die Anforderungen und Erwartungen an die Schüler in einer solchen Situation nicht völlig getrennt von jenen der zugrunde liegenden Anforderungssituation in der Vergangenheit und der dazwischen liegenden Bearbeitung durch die Schüler, inklusive der damit verbundenen individuellen Erfolgs- oder Misserfolgserfahrungen, beurteilbar sind. Mit Rückgriff auf Slavin (1994) gesprochen besteht in solchen Fällen die Frage, welcher der beiden Anlässe – die ursprüngliche oder die gegenwärtige Anforderungssituation – mit seinen jeweiligen Charakteristika den für das in der aktuellen Situation beobachtbare Beteiligungsverhalten entscheidenden Anreizwert gesetzt hat. Weil dies im gegebenen Rahmen nicht ausreichend geklärt werden kann, wurden entsprechende Aufgaben in Anlehnung an Blömeke, Risse, Müller, Eichler und Schulz (2006) als Aufgaben mit *Reflexions-/Auswertungsfunktion* gekennzeichnet

und in den Ausprägungsstufen differenziert nach dem Anlass der Bezugnahme (*Unterrichtsaufgabe* oder *Hausaufgabe*). Die solcherart gekennzeichneten Aufgaben wurden in der Folge nur hinsichtlich der zweifellos beurteilbaren Aufgabenmerkmale analysiert, was in der Regel Aspekte der verwendeten Medien und der Aufgabeninhalte umfasste.

Übergreifende Analyse- und Kodierregeln: Alle genannten Aufgabeneigenschaften, bis auf die Aspekte des Zeitumfangs, wurden mit einer lehrerzentrierten Perspektive kodiert bzw. beurteilt. Das heißt, dass die jeweils erkennbaren (z.B. über konkrete verbale Äußerungen) oder mutmaßlichen, d.h. auf Grundlage der vorliegenden Informationen eingeschätzten Intentionen der Lehrkraft zur jeweiligen Aufgabe maßgeblich waren für die Einordnung einer Merkmalsausprägung. Aus den Äußerungen und dem Handeln der Lehrkraft und der Art der von ihr präsentierten Materialien ging allerdings nicht immer ausreichend hervor, welche Erwartungen an die Aufgaben geknüpft sind, wodurch besonders Merkmale wie das Format des geforderten Produkts oder die erforderlichen kognitiven Prozesse schwierig zu beurteilen waren. In solchen Fällen war für die betreffenden Analyseaspekte der Code *nicht bestimmbar* zu vergeben.

Die Qualität der Aufgabenbearbeitung durch die Schüler oder deren individuelle subjektive Wahrnehmung der Aufgaben waren nicht Gegenstand der Analysen und wurden in der Durchführung der Kodierungen ausgeblendet. Wie bereits in Kap. 4.3 erwähnt stand eine so genannte rationale (Bromme, Seeger & Steinbring, 1990) bzw. objektive (Neubrand, 2002) Beschreibung von solchen Aufgabenmerkmalen im Vordergrund, die sich aus der Aufgabenstellung heraus erschließen lassen. Die genannten Analysemerkmale wurden dementsprechend alleine auf Grundlage der in der Ereigniskodierung definierten Sequenzen der Aufgabenstellung eingeschätzt. An verschiedenen Stellen diskutierte Differenzen zwischen der Qualität der Aufgabenstellung und deren -bearbeitung (Drollinger-Vetter, 2006; Stein, Grover & Henningsen, 1996) stellen für den vorliegenden Kontext keine Beeinträchtigung dar: Durch den qua Aufgabendefinition gewählten Betrachtungsmaßstab dieser Untersuchung gingen nennenswerte Änderungen der Bedingungen einer Aufgabe während der Bearbeitungszeit in der Regel mit der Kodierung eines neuen Aufgabenereignisses einher (vgl. Kap. 6.7.1).

Beobachterübereinstimmungen: Die Wahl einer geeigneten Prüfgröße zur Bestimmung der Beobachterübereinstimmung richtet sich u.a. nach dem Skalenniveau der erhobenen Daten (Wirtz & Caspar, 2002). Bei den dargestellten Aufgabenmerkmalen sind unterschiedliche Skalenniveaus vertreten: Während die Erfassung des verbalen Umfangs von Aufgabenstellung und Medieninhalten mittels Wortzählung als intervallskaliert angesehen werden kann, beinhaltet der überwiegende Teil der Merkmale kategoriale Zuordnungen, die je nach zugrunde liegendem Konstrukt als nominal- oder ordinalskaliert zu verstehen sind³⁸. Aus diesem Grund werden mehrere Übereinstimmungsmaße zur Prüfung herangezogen: Für nominal- und ordinalskalierte Daten die prozentuale Übereinstimmung (PÜ), die den Anteil von Übereinstimmungen an der Gesamtanzahl aller vorgenommenen Kodierungen bezeichnet, für nominalskalierte Daten zusätzlich Cohens κ - Koeffizient, welcher das Verhältnis zwischen zufällig erwartbaren und tatsächlichen Übereinstimmungen berücksichtigt; für ordinalskalierte Merkmale außerdem die Rangkorrelation nach Spearman (Koeffizient r_s), und für die intervallskalierten Merkmale die Intraklassenkorrelation (ICC).

Tab. 5 liefert einen Überblick über die berechneten Übereinstimmungswerte zu den unterschiedlichen Aufgabenmerkmalen. Gemessen an den Orientierungswerten aus der Methodenliteratur (ebd.) zeigen die berechneten Übereinstimmungsmaße überwiegend gute bis sehr gute Werte. Lediglich die Ergebnisse zu den beiden Merkmalen *Anzahl der Hilfestellungsmaßnahmen* und *Format des geforderten Produkts* liegen auf einem Niveau, welches in der Zusammenschau der Kennwerte nur knapp als zufriedenstellend zu bezeichnen ist, im gegebenen Kontext aber als akzeptabel gelten kann.

³⁸ Das Merkmal „Anzahl der Hilfestellungsmaßnahmen“ kann aus messtheoretischer Sicht trotz Erfassung mittels Zählung lediglich als ordinalskaliertes Merkmal gelten (vgl. Bortz, 2010; Bortz & Döring, 2006).

Tab. 5: Kennwerte der Beobachterübereinstimmung für die beurteilten Aufgabenmerkmale

	PÜ	Cohens κ	r_s	ICC _{one-way, random}
<i>Einführung und Präsentation der Aufgabe</i>				
Sprache	92 %	0,91	-	-
verbaler Umfang	-	-	-	0,97
Hilfestellungsmaßnahmen	66 %	0,43	0,73	-
Medien: Darstellungsform	84 %	0,82	-	-
Medien: Gerät/Material	87 %	0,83	-	-
Bekanntheit des Mediums	90 %	0,86	-	-
Umfang der Medieninhalte	-	-	-	1,00
<i>Aufgabeninhalte</i>				
Lebensweltbezug	83 %	0,74	0,72	-
Abstraktheit	87 %	0,73	0,75	-
<i>Anforderungen der Bearbeitung</i>				
Kognitiver Prozess	81 %	0,77	0,67	-
Wissensdimension	84 %	0,76	0,69	-
Format des geforderten Produkts	75 %	0,43	0,67	-
Art des geforderten Produkts	90 %	0,72	-	-
Reflexions-/Auswertungsfunktion	95 %	0,78	-	-

Anm.: Werte stellen die Mediane über alle Übereinstimmungspaare dar; Grundlage: 6 doppelt und 3 dreifach kodierte Unterrichtsstunden; zu keinem der Merkmale liegen mehr als 10 % nichtkategorisierte Objekte vor (vgl. Empfehlung für nominale Kategoriensysteme bei Wirtz & Caspar 2002, S. 50); für Ergebnisse der Übereinstimmungsprüfung zu den Zeitmerkmalen s. Kap. 6.7.1.

7. Ergebnisse

Nachfolgend werden die Ergebnisse dieser Studie dargestellt. Der Aufbau der Darstellungen orientiert sich an den in Kap. 5.3 genannten Forschungsfragen. Dementsprechend wird als Erstes das in der vorliegenden Stichprobe beobachtete Schülerbeteiligungsverhalten beschrieben und Einflüsse der schriftlich erfassten Personeneigenschaften auf dieses Merkmal geprüft. Dazu werden zunächst die deskriptiven Ergebnisse zu den erhobenen Personenmerkmalen der Schüler (Leistungsstand im Unterrichtsfach, familiärer Sprachhintergrund und Geschlecht) berichtet (7.1). Über die Verteilung der Merkmalsausprägungen soll ein Eindruck über die Heterogenität der Lerngruppen gewonnen werden. Daraufhin erfolgt die Beschreibung der Ergebnisse zum Beteiligungsverhalten (7.2), die neben der Stichprobencharakterisierung (7.2.1) und der Wiedergabe der deskriptiven Befunde (7.2.2) eine Betrachtung von Einflüssen der Personenmerkmale auf die individuelle Verhaltenstendenz (7.2.3) und eine Beleuchtung der zeitlichen Stabilität des Merkmals, sowohl innerhalb der Unterrichtsstunden als auch über die Unterrichtsstunden hinweg (7.2.4), umfasst. Die Ergebnisse dieses Unterkapitels dienen sowohl der Prüfung der Varianz des Schülerbeteiligungsverhaltens in interindividueller und zeitlicher Hinsicht als auch der Abhängigkeit des Merkmals von den Personeneigenschaften der Schüler. In diesem Zusammenhang wird auch die so genannte *on task*-Rate als zentrale Auswertungs- und Berechnungsgröße für das Schülerbeteiligungsverhalten definiert.

Im darauf folgenden Abschnitt werden die Ergebnisse der ebenfalls beobachtungsbasierten Erfassung und Charakterisierung der Aufgaben im Unterricht dargelegt (7.3). Neben der Beschreibung der Stichprobe für die Aufgabenanalyse (7.3.1) beinhaltet dieser Teil deskriptive Ergebnisse sowohl zu den Eigenschaften der Aufgaben im Unterrichtsverlauf (7.3.2) als auch der stundenspezifischen Variabilität der Aufgabenmerkmale (7.3.3). Insgesamt sollen die Teilergebnisse dieses Abschnitts Aufschluss über die Ausprägungen und Verteilungen der erfassten Aufgabeneigenschaften geben.

Das abschließende Unterkapitel dient zur Beantwortung der zentralen Frage dieser Studie, inwieweit sich Ausprägungsvariationen des Schülerbeteiligungsverhaltens durch Eigenschaften der Aufgaben erklären lassen (7.5). Dazu wird im ersten Schritt geprüft, ob das Beteiligungsniveau innerhalb der Klasse zu verschiedenen Zeitpunkten des

Unterrichtsverlaufs durch die gleichzeitige Ausprägung der einzelnen Aufgabenmerkmale (und auch Kombinationen von Merkmalsausprägungen) beeinflusst wird (7.5.1). Im zweiten Schritt erfolgt eine Untersuchung der Zusammenhänge zwischen verschiedenen Indikatoren der stundenspezifischen Variabilität der Aufgabeneigenschaften und dem durchschnittlichen Schülerbeteiligungsniveau in der jeweiligen Stunde (7.5.2). Die Ergebnisse beider Auswertungsschritte werden differenziert dargestellt nach generellen Effekten, die für alle Schüler der Stichprobe feststellbar sind, und Effekten für Untergruppen der Schülerstichprobe, die anhand der vorliegenden Personenmerkmale (s.o.) unterschieden werden.

Zur Auswertung kamen unterschiedliche statistische Verfahren zur Anwendung, wie z.B. deskriptive Statistiken zur Variablenbeschreibung, Varianzanalysen zur Prüfung von Gruppenunterschieden und verschiedene Korrelationsverfahren zur Prüfung von Merkmalszusammenhängen, wobei je nach Erfordernis parametrisch oder nonparametrisch ausgerichtete Verfahrensweisen verwendet wurden (Bortz, 2010; Bühner & Ziegler, 2009; Leonhart, 2004; Rasch, Friese, Hofmann & Naumann, 2006, 2010). Die Berechnungen wurden mithilfe der Statistiksoftware SPSS (Versionen 20/21) und in einzelnen Fällen Microsoft Excel (2008, Version 12.3.6) durchgeführt. Beschreibungen der einzelnen Verfahrensweisen erfolgen an den entsprechenden Stellen des nachfolgenden Ergebnisberichts.

7.1 Personenmerkmale der Schüler

7.1.1 Sprachlicher Hintergrund in der Familie

Tab. 6 zeigt die Ergebnisse der Erfassung des familiären sprachlichen Hintergrunds der Schüler im Rahmen des Testfragebogens (vgl. Kap. 6.5). Nur etwa ein Viertel der Stichprobe (22 %) verwendet zu Hause Deutsch als einzige Sprache, ein weiteres Viertel (24 %) spricht zu Hause eine andere Sprache als Deutsch. Diese „anderen“ Sprachen wurden über ein offenes Antwortfeld erfragt und lassen eine breite Ausdifferenzierung erkennen: Insgesamt sind 26 nicht-deutsche Familiensprachen in der Stichprobe festzustellen, darunter Türkisch (27 Schüler), Marokkanisch (12 Schüler), Kroatisch (8 Schüler), Portugiesisch (6 Schüler) und Englisch bzw. Polnisch (jeweils 4 Schüler), um

beispielhaft die fünf häufigsten zu nennen. Den größten Anteil bilden Schüler, die zu Hause Deutsch und zusätzlich eine andere Sprache sprechen (54 %).

Tab. 6: Häufigkeiten der familiären Umgangssprache, differenziert nach Klassen

Klasse	Umgangssprache in der Familie		
	Deutsch	andere Sprache als Deutsch	Deutsch + andere Sprache
A	4	1	11
B	3	9	5
C	4	4	9
D	8	2	16
E	5	11	12
F	5	5	18
Häufigkeit gesamt	29	32	71
Anteil an Stichprobe	22 %	24 %	54 %

Der Anteil von Schülern mit einem ausschließlich deutschsprachigen familiären Hintergrund erscheint verglichen mit bisherigen Forschungsergebnissen, z.B. im Rahmen von PISA (Walter, 2008) oder DESI (Hesse, Göbel und Hartig, 2008), sehr gering, lässt sich aber vermutlich durch regionale Spezifika der Stichprobe erklären: Der Ballungsraum Frankfurt am Main weist mit 65 % einen im Bundesvergleich außerordentlich hohen Anteil an Kindern und Jugendlichen mit Migrationshintergrund auf (Autorengruppe Bildungsberichterstattung, 2010). Bezüglich der schulklassenspezifischen Verteilung der Umgangssprache in der Familie ist lediglich ein auffällig hoher Anteil von Schülern mit nicht-deutscher Familiensprache in den Klassen B und E erkennbar.

7.1.2 Geschlecht

Die Erhebung des Geschlechts über die Klassenlisten (n = 145) ergibt ein Verhältnis von 68 Schülerinnen zu 77 Schülern (47 % weiblich, 53 % männlich). Diese Ergebnisse weichen nur unwesentlich von den mittels schriftlicher Befragung (im Rahmen des Leistungstests, n = 132) erfassten Daten ab (48 % weiblich, 52 % männlich). Diese Verteilung repräsentiert, bei vereinzelt klassenspezifischen Abweichungen (s. Stichprobenbeschreibung, Kap. 6.3, Tab. 1), insgesamt gut die durchschnittliche

Geschlechterverteilung an deutschen allgemeinbildenden Schulen im Erhebungszeitraum (Statistisches Bundesamt, 2010).

7.1.3 Leistungsstand im Unterrichtsfach

Der individuelle Leistungsstand der Schüler im für die vorliegende Stichprobe einheitlichen Unterrichtsfach Englisch (vgl. Kap. 6.3) wurde mithilfe eines auszugsweise von Bader und Schaer (2005, vgl. Kap. 6.5) übernommenen, schriftlichen Leistungstests erhoben. Tab. 7 bietet einen Überblick zu den gemessenen Testleistungen, differenziert dargestellt nach unterschiedlichen Leistungsbereichen (Subtests) und der Gesamtestleistung.

Tab. 7: Deskriptive Kennwerte der Leistungen im Englischtest

Subtest (n = 132)	M	SD	Min	Max
Hörverstehen	3,6	1,6	0	5
Leseverstehen	1,8	1,0	0	3
Grammatik	9,2	3,8	0	15
Summenscore Gesamtest	14,6	5,6	1	23

Die vorliegende Stichprobe zeigt sich im Vergleich zu den bei Bader und Schaer (ebd.) getesteten Schülern etwas leistungsschwächer (durchschnittlicher Anteil korrekter Items diese Stichprobe: 63 %, Stichprobe Bader und Schaer: 74 %). Die für die weiteren Berechnungen maßgebliche Testsumme ist nicht normalverteilt (Kolmogorov-Smirnov-Test: $p < 0,001$) und zeigt sich mit einer Schiefe von -0,44 leicht rechtssteil. Abb. 4 bietet einen Überblick zur Verteilung der Testleistung in den beteiligten Klassen, differenziert nach Schulformen³⁹.

³⁹ Darstellung als *Box-and-Whisker-Plots* (vgl. Janssen & Laatz, 2010, S. 688): Senkrechte Kästchen umfassen 50 % der Fälle; Querstrich kennzeichnet Median; *whiskers* („Fühler“) markieren Minimal-/Maximalwerte bis zum 1,5-fachen Interquartilsabstand.; darüber hinausgehende Extremwerte sind als Punkte dargestellt.

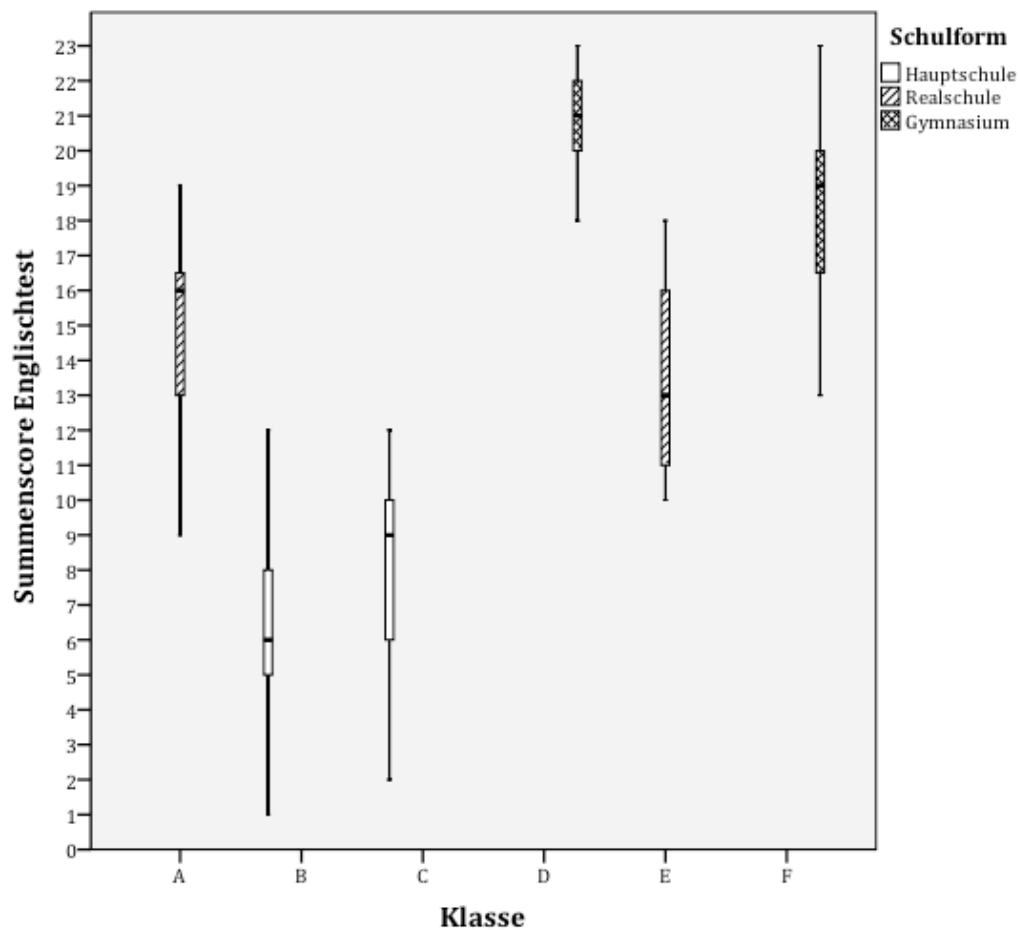


Abb. 4: Streuungs- und Lagemaße der Englischtestleistungen nach Klassen und Schulformen

Unmittelbar ersichtlich wird die Ausschöpfung des gesamten Wertebereichs in der Stichprobe, zugleich sind Leistungsdifferenzen zwischen den Schulformen erkennbar: Die beiden Hauptschul- und Gymnasialklassen decken einen Leistungsbereich in der jeweils unteren bzw. oberen Hälfte des Wertebereichs ab, während sich die beiden Realschulklassen auf einem (erhöhten) mittleren Niveau bewegen. Zudem sind große Leistungsspannen innerhalb der Klassen zu konstatieren, die zu teilweise deutlichen Überschneidungen zwischen den Leistungsbereichen der Schulformen führen, wie etwa der Vergleich von Gymnasialklasse F mit den beiden Realschulklassen A und E zeigt. Lediglich Klasse D, als leistungstärkste Lerngruppe der Stichprobe, weist einen vergleichsweise engen Leistungsbereich von fünf Punkten auf. Davon abgesehen lässt sich die Stichprobe zusammenfassend sowohl hinsichtlich der klassenspezifischen Verteilungen als auch mit Blick auf die Gesamtverteilung als sehr leistungsheterogen bezeichnen.

Einflüsse von Geschlecht und Sprachhintergrund auf die Testleistung:

Mit dem Ziel, mögliche Interdependenzen zwischen den vorliegenden Personenmerkmalen der Schüler zu identifizieren, wurde der Frage nachgegangen, ob interindividuelle Unterschiede in den Testleistungen auf das Geschlecht oder den familiären Sprachhintergrund (oder eine Kombination aus diesen beiden Faktoren) zurückzuführen sind. Dazu wurde mithilfe einer zweifaktoriellen, univariaten Varianzanalyse geprüft, ob sich die Leistungsmittelwerte der anhand der Ausprägungsstufen differenzierbaren Gruppen (z.B. männlich vs. weiblich) signifikant voneinander unterscheiden. Zwar ist eine der notwendigen statistischen Voraussetzungen für das Verfahren – die Normalverteilung der abhängigen Variable (Bortz, 2010; Bühner & Ziegler, 2009) – an dieser Stelle nicht erfüllt, doch wird diesbezüglich Empfehlungen aus der Methodenliteratur gefolgt, besonders bei mindestens mittlerer Stichprobengröße parametrische Verfahren und damit auch die Varianzanalyse anzuwenden (vgl. Bortz, 2010; Bühner & Ziegler, 2009; Rasch, Frieze, Hofmann & Naumann, 2006). Grundlage für die Prüfung sind die in Tab. 8 dargestellten Gruppenmittelwerte.

Tab. 8: Gruppenmittelwerte der Englischtestleistung nach Geschlecht und familiärem Sprachhintergrund (Standardabweichungen in Klammern)

		Geschlecht	
		männlich (n = 69)	weiblich (n = 63)
Umgangssprache in der Familie			
		13,7 (5,7)	15,5 (5,3)
Deutsch (n = 29)	15,2 (5,5)	14,8 (5,9)	15,6 (5,4)
andere Sprache als Deutsch (n = 32)	12,6 (5,8)	11,2 (6,0)	14,4 (5,2)
Deutsch + andere Sprache (n = 71)	15,2 (5,3)	14,5 (5,2)	15,9 (5,4)

Hinsichtlich des familiären Sprachhintergrunds ist bei den Schülern, die zu Hause kein Deutsch sprechen, eine niedrigere Testleistung festzustellen als bei denjenigen, die zu Hause hauptsächlich Deutsch oder eine Kombination aus Deutsch und mindestens einer anderen Sprache sprechen ($M_{\text{Deutsch}} = 15,2$; $M_{\text{andere Sprache}} = 12,6$; $M_{\text{Deutsch+andere Sprache}} = 15,2$). Der Mittelwertvergleich mittels zweifaktorieller Varianzanalyse weist diese Gruppenunterschiede allerdings als statistisch nicht bedeutsam aus, bei zudem geringer

Effektstärke ($F_{(2;126)} = 2,30$; $p = 0,10$; $\eta^2 = 0,04$)⁴⁰. Auch der Faktor Geschlecht verfehlt die Signifikanzgrenze von 5 % und zeigt einen lediglich schwachen Effekt, der dadurch zustande kommt, dass Mädchen eine leicht höhere mittlere Testleistung zeigen als Jungen ($M_{\text{männlich}} = 13,7$; $M_{\text{weiblich}} = 15,5$; $F_{(1;126)} = 3,21$; $p = 0,08$; $\eta^2 = 0,03$). Eine Wechselwirkung zwischen den beiden Faktoren Geschlecht und familiärer Sprachhintergrund kann mit Blick auf die Kennwerte des Interaktionseffekts ausgeschlossen werden ($F_{(2;126)} = 0,43$; $p = 0,65$; $\eta^2 = 0,01$). Zusammenfassend zeigen die Ergebnisse, dass sich die Testleistung der Schüler trotz geringfügiger gruppenspezifischer Tendenzen nicht in substantieller Weise durch das Geschlecht oder den familiären Sprachhintergrund oder gar eine Kombination aus beiden Variablen erklären lässt, was als ausreichender Beleg dafür betrachtet wird, dass die Personenmerkmale im weiteren Verlauf der Auswertungen als voneinander unabhängige Größen betrachtet werden können.

7.2 Beteiligungsverhalten der Schüler im Unterricht

7.2.1 Stichprobe

Die beobachtungsbasierte Erfassung des Schülerbeteiligungsverhaltens mittels Intervallkodierung (vgl. Kap. 6.6.3) wurde ausschließlich auf solche Unterrichtsphasen angewendet, in denen die Sozialform des so genannten öffentlichen Unterrichts vorlag, sodass zunächst die entsprechenden Zeitabschnitte mittels Ereignisstichprobe zu erfassen waren (vgl. Kap. 6.6.2). Die Ergebnisse dieser Basiskodierung (s. Tab. 9) zeigen wie erwartet ein Vorherrschen des öffentlichen Unterrichts: Mit durchschnittlich 27 Minuten entfallen gut 60 % der nominellen bzw. „offiziellen“ Unterrichtszeit (entspricht 45 Minuten) und 70 % der tatsächlichen Unterrichtszeit (nominelle Unterrichtszeit abzüglich Unterrichtsunterbrechungen, verspätetem Beginn und verfrühtem Ende, s. Kategorie „keine Sozialform“) auf diese Sozialform.

⁴⁰ Interpretation der Effektgröße η^2 in Anlehnung an die auf Cohen (1988, zitiert nach Bühner & Ziegler, 2009, S. 364) zurückgehenden Orientierungsgrößen: kleiner Effekt .01, mittlerer Effekt .06, großer Effekt .14.

Tab. 9: Deskriptive Ergebnisse der Basiskodierung der Sozialformen

Unterrichtsstunde	Zeitlicher Umfang der Sozialformen (in Min., gerundet) ^a				
	Öffentlicher Unterricht	Einzelarbeit	Partnerarbeit	Gruppenarbeit	Keine Sozialform
A 1103	24	0	0	13	8
A 1903I	26	15	0	0	4
A 1903II	40	0	0	0	5
A 0204	28	0	6	0	11
B 0903	32	0	0	0	13
B 3103	26	14	0	0	5
B 0204I	27	8	0	0	10
B 0204II	19	10	0	7	9
C 0303	23	16	0	0	7
C 1103	30	0	15	0	0
C 1603	37	8	0	0	0
D 1602	18	2	0	20	5
D 0603	27	17	0	0	2
D 1303I	37	6	0	0	2
D 1303II	24	0	15	0	6
E 0503	15	25	0	0	5
E 1003	31	2	6	0	7
E 1203	22	20	0	0	3
E 1703	20	15	0	0	10
F 1203	28	7	0	0	10
F 2703	35	0	0	0	10
F 3003I	36	5	0	0	4
F 3003II	19	19	0	0	8
Mittelwert	27	8	2	2	6
Standardabweichung	7	8	5	5	4

^a Kategorie „uneinheitliche Sozialform“ wegen fehlenden Auftretens nicht aufgeführt

Die Befunde belegen damit erneut eine aus der Forschung bekannte Tendenz zu lehrerzentrierten Sozialformen (Bohl, Kleinknecht, Batzel & Richey, 2012; Götz, Lohrmann, Ganser & Haag, 2005; Hage et al., 1985; Kleinknecht, 2010; Seidel et al., 2006), wobei deren Gewichtung in dieser Stichprobe vergleichsweise stark ausfällt: Der Anteil öffentlichen Unterrichts an der tatsächlichen Unterrichtszeit liegt bei keiner der

23 untersuchten Unterrichtsstunden – mit Ausnahme von Stunde E 0503 (38 %, s. Tab. 9) – niedriger als 40 %. Die Kodierung des Beteiligungsverhaltens der Schüler bezieht sich somit auf den überwiegenden Zeitanteil des aufgezeichneten Unterrichts. Mit Blick auf den Umfang an zu beobachtenden Personen ist festzustellen, dass bedingt durch die Schüleranwesenheit bei den Aufzeichnungsterminen und die Sichtbarkeit der Schüler im Videomaterial durchschnittlich 85 % der Schüler einer Klasse in die Verhaltenskodierungen einbezogen werden konnten. Aus dem zeitlichen Umfang der einbezogenen Unterrichtsphasen und der Anzahl an beobachtbaren Personen resultiert damit eine letztendliche Beobachtungstichprobe von insgesamt 12.895 Beobachtungsintervallen zu je einer Minute.

7.2.2 Deskriptive Ergebnisse zum Schülerbeteiligungsverhalten

Wie in Kap. 6.6.3 beschrieben, wurde das Beteiligungsverhalten aller sichtbaren Schüler innerhalb der Phasen öffentlichen Unterrichts entlang eines intervallbasierten Zeitstichprobenplans erfasst (Intervalllänge: eine Minute) und als dichotome Merkmalsausprägung (*on/off task*) festgehalten. Auf diese Weise entstand eine Datenmatrix, deren kleinste Einheit den Verhaltensstatus eines Schülers in einer bestimmten Minute des Unterrichts als zweistufigen Wert (0/1) abbildet. Basierend auf Anteilswerten der *on task*-Ausprägung lassen sich damit sowohl personen- als auch situationsspezifische Kennwerte für das Schülerbeteiligungsverhalten generieren: einerseits der Anteil des als *on task* zu bezeichnenden Verhaltens einer bestimmten Person im Laufe eines ausgewählten Zeitraums (Unterrichtssequenz, -stunde oder -reihe), und andererseits der Anteil von Schülern einer Klasse, die zu einem bestimmten Zeitpunkt, d.h. während eines ausgewählten Beobachtungsintervalls, *on task*-Verhalten zeigen. Diese so genannten *on task*-Raten bilden die Berechnungsgrundlage für alle weiteren Auswertungsschritte.

Prozessbezogene on task-Rate: Einen Überblick zu den Ergebnissen der Verhaltenskodierung liefert Tab. 10 anhand der über die Beobachtungsintervalle aggregierten *on task*-Rate der Klasse im Stundendurchschnitt. Am Beispiel der Stunde A 1103 ist daran abzulesen, dass zu einem beliebigen Zeitpunkt im Verlauf dieser

Unterrichtsstunde durchschnittlich ca. 94 % der Schüler der Klasse ein als *on task* zu bezeichnendes Verhalten erkennen ließen. In dem Beobachtungsintervall mit der niedrigsten *on task*-Rate dieser Stunde (s. Minimalwert) lag der Anteil von Schülern mit *on task*-Verhaltensstatus in der Klasse bei 63 %.

Tab. 10: Deskriptive Ergebnisse zu den klassenweise aggregierten *on task*-Raten in den Unterrichtsstunden

Unterrichtsstunde	on task-Rate (in %) ^a			
	M	SD	Min	Max
A 1103	94,1	9,4	63,2	100,0
A 1903I	94,5	7,8	70,6	100,0
A 1903II	91,7	7,3	72,2	100,0
A 0204	89,4	8,9	66,7	100,0
B 0903	88,6	10,3	68,8	100,0
B 3103	90,1	8,2	73,3	100,0
B 0204I	81,9	17,1	40,0	100,0
B 0204II	72,2	15,9	41,7	100,0
C 0303	97,5	4,7	86,7	100,0
C 1103	70,1	14,1	38,5	100,0
C 1603	81,0	16,9	33,3	100,0
D 1602	92,3	12,1	62,5	100,0
D 0603	91,1	8,9	66,7	100,0
D 1303I	83,9	12,9	43,8	100,0
D 1303II	79,6	16,1	42,1	100,0
E 0503	73,8	16,1	36,4	94,1
E 1003	69,0	16,5	28,6	95,5
E 1203	73,6	17,3	39,1	95,0
E 1703	77,7	11,6	55,6	95,2
F 1203	84,9	12,4	55,6	100,0
F 2703	90,6	10,7	56,0	100,0
F 3003I	75,1	10,8	42,9	90,9
F 3003II	74,8	18,0	37,5	100,0
Durchschnitt	83,4	12,3	53,1	98,7

^a N = Anzahl der Beobachtungsintervalle gemäß Kodierung der Phasen öffentlichen Unterrichts (s. Tab. 9)

Über alle Unterrichtsstunden hinweg zeigen die Ergebnisse ein durchschnittliches *on task*-Niveau von 83,4 %. Die Höhe dieses Stichprobenmittelwerts kann durch die Eigenschaften des Konstrukts und die Erhebungsmethodik begründet werden: Wie in Kap. 2.3.3 dargelegt, unterliegt die beobachtungsbasierte Erfassung von als Zuwendung zu den Lerngelegenheiten zu interpretierendem, sichtbarem Schülerverhalten gewissen validitätsbedingten Einschränkungen, die grundsätzlich mit relativ hohen Messwerten verbunden sind. Zusätzlich kommt in der vorliegenden Untersuchung eine als *in dubio pro discipulo* bezeichnete Kodierregel zum Tragen, bei der in Zweifelsfällen zugunsten der *on task*-Ausprägung entschieden wurde (s. Kap. 6.6.3). In diesem Zusammenhang betrachtet schließen die dargestellten Ergebnisse direkt an den bisherigen Forschungsstand zum Beteiligungsverhalten an, welcher durchgängig *on task*- bzw. *engagement*-Raten in der Spanne zwischen 60 % und 95 %, mit besonderem Gewicht auf den Bereich oberhalb der 80 %, belegt (vgl. Kap. 2.3.3). Wie die Standardabweichung in Höhe von 12,3 andeutet, sind jedoch teils deutliche Differenzen zu diesem allgemeinen Wertenniveau festzustellen. Dies zeigt sich zum einen in Form von einigen vergleichsweise niedrigen Stundendurchschnittswerten (z.B. B 0204II: 72,2 %; C 1103: 70,1 %; E 1003: 69 %; E 1203: 73,6 %), und zum anderen durch eine Reihe stundenspezifischer Minimalwerte, die auf einen erheblichen Abfall der *on task*-Rate zu mindestens einem Zeitpunkt innerhalb des Stundenverlaufs hinweisen (z.B. C 1603: 33,3 %; E 1003: 28,6 %; F 3003II: 37,5 %). Am Beispiel von Stunde E 1003 veranschaulicht bedeutet das, dass im Verlauf des öffentlichen Unterrichts dieser Stunde durchschnittlich 31 % und zu einem Zeitpunkt sogar 71,4 % der Schüler offen erkennbar von den Lerngelegenheiten abgewandt waren.

Individuenbezogene on task-Rate: Richtet man den Blick auf die personenweise über die Unterrichtsstunden hinweg aggregierte *on task*-Rate⁴¹, zeigt sich wie zu erwarten ein der prozessbezogenen Aggregation (s.o.) nahezu gleiches Gesamtniveau des Beteiligungsverhaltens von 82,9 % *on task*. Die minimale Differenz von 0,5 Prozentpunkten zwischen den beiden Aggregationsweisen kommt zustande durch die Auswirkungen fehlender Werte (mangelnde Beobachtbarkeit einzelner Schüler in vereinzelten Beobachtungsintervallen) auf die unterschiedlichen Summenbildungen. Da bei der auf Individualwerten basierenden *on task*-Rate Extremwerte der Einzelpersonen

⁴¹ Berechnungsgrundlage: Schüler mit mindestens zwei Unterrichtsstunden Anwesenheit (n = 135)

bei der Datenzusammenfassung nicht durch das allgemeine Niveau der Klasse ausgeglichen werden, zeigt die Verteilung hierbei eine leicht höhere Streuung bzw. Spannweite (SD: 13,2; Min: 29,7; Max: 100). Demnach war der durch den Minimalwert repräsentierte Schüler zu durchschnittlich 29,7 % der untersuchten Unterrichtszeit von außen erkennbar mit der aktuellen Aufgabe beschäftigt⁴². Die Werte sind nicht normalverteilt (Kolmogorov-Smirnov-Test: $p < 0,01$), und die Verteilung ist mit einer Schiefe von -1,2 erkennbar rechtssteil (vgl. Abb. 5). Die auf Aggregation von Individualwerten über die Messzeitpunkte basierende *on task*-Rate dient als rechnerische Grundlage für die im nächsten Abschnitt folgende Prüfung möglicher Einflüsse der Personenmerkmale auf das individuelle Beteiligungsverhalten (Kap. 7.2.3), bevor im nächsten Schritt das Ausmaß verlaufsbedingter Schwankungen der *on task*-Rate innerhalb und zwischen den Unterrichtsstunden ausführlicher beleuchtet wird (Kap. 7.2.4).

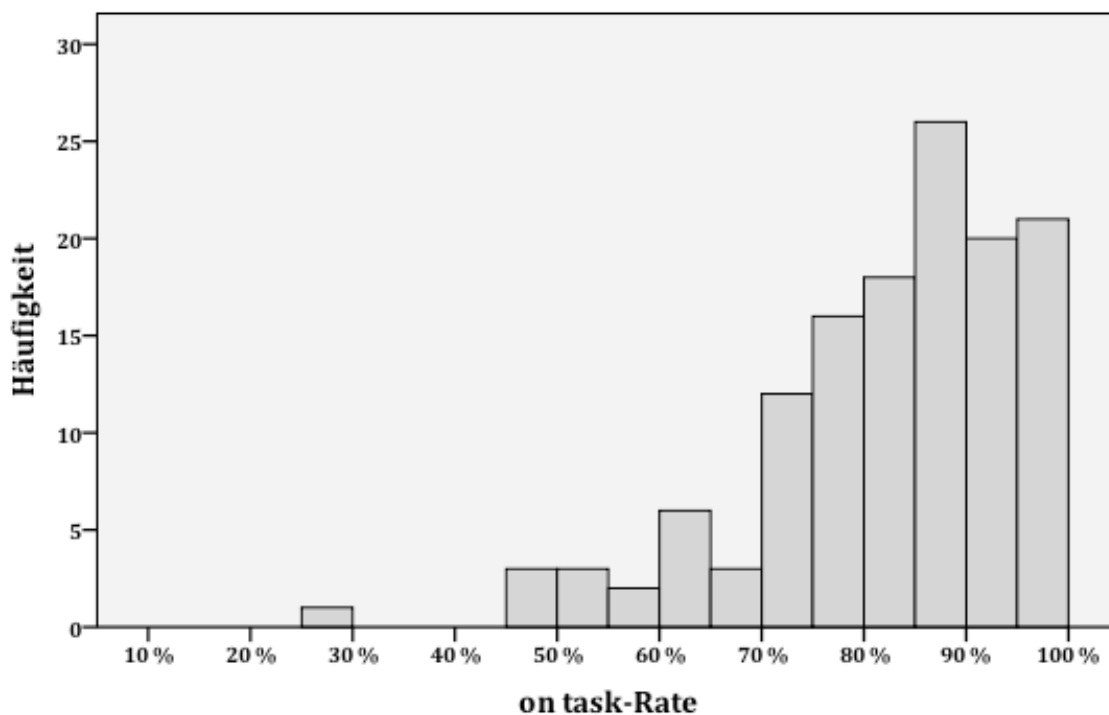


Abb. 5: Häufigkeitsverteilung der durchschnittlichen individuenbezogenen *on task*-Rate über die untersuchten Unterrichtsstunden hinweg (n = 135)

⁴² Die niedrigste gemessene *on task*-Rate eines Schülers in einer einzelnen Stunde beträgt 4,2 %, d.h. diese Person ließ in einem von 24 untersuchten Beobachtungsintervallen innerhalb dieser Stunde (D 1303II) ein als Beteiligung zu wertendes Verhalten erkennen.

7.2.3 Einflüsse von Personenmerkmalen auf die Ausprägung des individuellen Beteiligungsverhaltens

Das Hauptziel dieser Studie besteht darin, mögliche Einflüsse des Aufgabenangebots im Unterricht auf das Beteiligungsverhalten der Schüler zu identifizieren und zu beschreiben, wobei differenzielle Effekte für verschiedene Untergruppen der Schülerstichprobe mit aufgezeigt werden sollen. Zur Differenzierung der Untergruppen werden drei personenbezogene Merkmale eingesetzt: der familiäre Sprachhintergrund, das Geschlecht und der Leistungsstand im Unterrichtsfach (vgl. Kap. 6.5 und 7.1). Bevor aber eine Untersuchung von Effekten innerhalb des Unterrichtsprozesses erfolgt, ist zu klären, in welchem Ausmaß die individuelle Verhaltenstendenz der Schüler durch ihre Personenmerkmale, also außerhalb des Unterrichtsprozesses begründete, individuelle Bedingungsgrößen (*student inputs*, Slavin, 1994), zu erklären ist.

Einflüsse der Personenmerkmale: Zu diesem Zweck wird die individuenbezogene *on task*-Rate, d.h. der durchschnittliche Anteil von als *on task* zu bezeichnendem Beteiligungsverhalten eines Schülers innerhalb des Untersuchungszeitraums (s. vorheriger Abschnitt) als Prüfgröße herangezogen und mithilfe von Mittelwertvergleichen geprüft, wie stark sich die anhand der Personenmerkmale differenzierbaren Schülergruppen in der Ausprägung dieses Werts unterscheiden. Als Prüfverfahren wird dazu eine mehrfaktorielle, univariate Varianzanalyse angewendet. Auch an dieser Stelle (vgl. Kap. 7.1.3) wird in Anlehnung an die Empfehlungen bei Bühner und Ziegler (2009) und Rasch, Friese, Hofmann und Naumann (2006) trotz Verletzung der Normalverteilungsvoraussetzung auf die Nutzung eines nonparametrischen Alternativverfahrens (z.B. Kruskal-Wallis-Test) verzichtet. Zum Zwecke der Gruppenbildung wurden die intervallskalierten Testdaten zum Leistungsstand im Unterrichtsfach überführt in eine kategoriale Variable, die eine Zuordnung des individuellen Leistungsstands im Klassenvergleich anhand von drei Niveaustufen erlaubt (niedrig - mittel - hoch). Zu diesem Zweck wurden klassenweise drei Perzentilgruppen gebildet. Tab. 11 zeigt die zusammenfassenden Prüfergebnisse des Mittelwertvergleichs.

Tab. 11: Ergebnisse der mehrfaktoriellen, univariaten Varianzanalyse mit der individuellen durchschnittlichen *on task*-Rate als abhängige Variable und Personenmerkmalen der Schüler als Faktoren

	df	F	p	η^2
<i>Haupteffekte</i>				
Sprachhintergrund	2/106	0,61	0,55	0,01
Geschlecht	1/106	10,59	0,00	0,09
Leistungsniveau	2/106	0,21	0,81	0,00
<i>Interaktionseffekte</i>				
Sprachhintergrund * Geschlecht	2/106	0,23	0,79	0,00
Sprachhintergrund * Leistungsniveau	4/106	0,41	0,81	0,02
Geschlecht * Leistungsniveau	2/106	0,01	0,99	0,00
Sprachhintergrund * Geschlecht * Leistungsniveau	4/106	1,73	0,15	0,06

n = 124 (Schüler mit mindestens zwei Unterrichtsstunden Anwesenheit als Grundlage für die individuenbezogene *on task*-Rate und vollständigen Befragungs-/Testdaten)

Die Kennwerte des Mittelwertvergleichs zeigen keine nennenswerten Unterschiede der individuellen durchschnittlichen *on task*-Rate in Abhängigkeit vom familiären Sprachhintergrund ($F_{(2;106)} = 0,61$; $p = 0,55$; $\eta^2 = 0,01$) oder dem fachlichen Leistungsstand der Schüler ($F_{(2;106)} = 0,21$; $p = 0,81$; $\eta^2 = 0,00$). Eine getrennte Prüfung des Zusammenhangs zwischen dem Summenwert des Leistungstests und der *on task*-Rate mittels Produkt-Moment-Korrelation bestätigt durch den niedrigen, nicht signifikanten Korrelationseffizienten das Ausbleiben eines Effektes für den Faktor fachlicher Leistungsstand ($r = 0,12$; $p > 0,05$). Für das Geschlecht hingegen ist ein höchst signifikanter, mittelstarker Effekt festzustellen ($F_{(1;106)} = 10,59$; $p = 0,001$; $\eta^2 = 0,09$), der durch einen separat gerechneten (zweiseitigen) t-Test bestätigt wird ($t_{(122)} = -3,5$; $p = 0,001$; $d = 0,6$). Mädchen zeigen eine deutlich höhere durchschnittliche *on task*-Rate als Jungen ($M_{\text{männlich}} = 78,7\%$; $M_{\text{weiblich}} = 87,3\%$). Von den untersuchten Interaktionseffekten erreicht keiner annähernd die Signifikanzgrenze von 5 %, sodass Wechselwirkungen zwischen den einzelnen Personenmerkmalen ausgeschlossen werden können. Für das Ausbleiben von leistungs- bzw. herkunftsbedingten Effekten auf das durchschnittliche Niveau des individuellen Beteiligungsverhaltens finden sich in der bisherigen Forschung zum unterrichtlichen Schülerbeteiligungsverhalten (oder verwandten Konstrukten, s. Kap. 2.4) kaum vergleichbare Ergebnisse, weil meist andere Hintergrundvariablen der Schüler herangezogen werden, wie z.B. die ethnische Herkunft, der sozialökonomische Status der Familie oder die elterliche Lernunterstützung (vgl. Blatchford, Bassett &

Brown, 2011; Lam, Wong, Yang & Liu, 2012; Yair, 2000), oder die Personenmerkmale auf Klassenebene aggregiert ausgewertet werden (zum Sprachhintergrund z.B. Helmke & Renkl, 1993). Aus theoretischer Sicht erscheint es jedoch interessant, dass insbesondere der individuelle Leistungsstand im Unterrichtsfach, welcher im Sinne von Slavins (1994) Modell effektiven Unterrichts als Bestandteil der auf das Beteiligungsverhalten einwirkenden Lernvoraussetzungen der Schüler angesehen werden kann, keinerlei erkennbaren Einfluss auf das unterrichtliche Beteiligungsverhalten zeigt. Das Ergebnis zum Unterschied zwischen den Geschlechtern untermauert Forschungsergebnisse, die ebenfalls höhere Beteiligungsraten bei Mädchen belegen (z.B. Marks, 2000; Rimm-Kaufman, Baroody, Larsen, Curby & Abry, 2014). Da die diesbezügliche Befundlage aber insgesamt sehr uneinheitlich ausfällt – mitunter sind die Effekte nur schwach (Helmke & Renkl, 1993), manche Studien stellen keinerlei Geschlechterunterschiede fest (Blatchford, Bassett & Brown, 2011; Finn, Pannozzo & Voelkl, 1995), und an anderen Stellen zeigen eher sogar Jungen höhere Beteiligungsraten als Mädchen (Yair, 2000) – war der vorliegende Effekt in dieser Form nicht zu erwarten.

Einflüsse von Schulform und Klasse: Zusätzlich zum Einfluss der drei Personenmerkmale wurde auch geprüft, inwiefern Variationen der individuellen mittleren *on task*-Rate durch die Zugehörigkeit zur jeweiligen Schulklasse oder die Schulform bedingt sind. Während zwischen Schülern unterschiedlicher Schulformen keine nennenswerten Mittelwertunterschiede festzustellen sind ($F_{(1,121)} = 0,86$; $p = 0,43$; $\eta^2 = 0,01$), zeigt sich für den Faktor Schulklassenzugehörigkeit ein höchst signifikanter, starker Effekt ($F_{(1,118)} = 6,14$; $p = 0,000$; $\eta^2 = 0,21$). Bei einem Blick auf die Mehrfachvergleiche (Post Hoc-Tests nach Scheffé) wird ersichtlich, dass dieser Effekt hauptsächlich durch das im Stichprobenvergleich besonders niedrige allgemeine *on task*-Niveau in Klasse E (vgl. Tab. 10, Kap. 7.2.2.) verursacht wird. Dennoch bedeutet dieses Ergebnis, dass die Zugehörigkeit zu einer bestimmten Schulklasse von großer Bedeutung für die Tendenz des individuellen Beteiligungsverhaltens ist. Um den Schulklasseneffekt in Beziehung setzen zu können zum ebenfalls relativ großen Einfluss des Geschlechts (s.o.), wurde ergänzend eine zweifaktorielle Varianzanalyse mit diesen beiden Faktoren durchgeführt. Im Ergebnis vermindern sich die Effektstärken beider Faktoren bei praktisch unveränderten Signifikanzwerten nur geringfügig (Haupteffekt Geschlecht: $F_{(1,112)} = 7,34$; $p = 0,01$; $\eta^2 = 0,06$; Haupteffekt Schulklasse: $F_{(1,112)} = 5,48$; $p = 0,000$; $\eta^2 =$

0,20); das starke Gewicht des Faktors Schulklassenzugehörigkeit bleibt damit bestehen. Der Interaktionseffekt erweist sich daneben als nicht signifikant, allerdings bei mittlerer Effektstärke ($F_{(1,112)} = 1,28$; $p = 0,28$; $\eta^2 = 0,05$). Diese Wechselwirkung kommt dadurch zustande, dass sich in einer Klasse die *on task*-Mittelwerte von Jungen und Mädchen im Gegensatz zum Stichprobentrend praktisch nicht unterscheiden (Klasse D; $M_{\text{männlich}} = 86,9\%$; $M_{\text{weiblich}} = 86,4\%$). Der Befund zum Schulklasseneffekt schließt an Forschungsergebnisse von Helmke und Renkl (1993) und Helmke und Weinert (1997) an, die ebenfalls eine starke Abhängigkeit der Variationen des Aufmerksamkeitsverhaltens der Schüler von der Schulklasse bzw. der Lehrkraft nachweisen.

Zusammenfassend ist für die Ergebnisse dieses Abschnitts festzuhalten, dass ein Teil der geprüften Hintergrundvariablen keinen relevanten Einfluss auf das Beteiligungsverhalten der Schüler erkennen lässt. Dies trifft auf den familiären Sprachhintergrund und den Leistungsstand im Unterrichtsfach, aber auch auf die Kontextvariable der Schulform zu. Daneben zeigen sich bedeutsame Unterschiede im individuellen *on task*-Niveau in Abhängigkeit vom Geschlecht und in noch größerem Maße von der Zugehörigkeit zur jeweiligen Klasse. Dieses Bild wird als Hinweis darauf gewertet, dass das Beteiligungsverhalten der Schüler stärker durch Prozessmerkmale des Unterrichts in den jeweiligen Lerngruppen bedingt ist als durch außerhalb des Unterrichts begründete Faktoren.

7.2.4 Zeitliche Stabilität des Schülerbeteiligungsverhaltens

Im Anschluss an die Ergebnisse des vorherigen Abschnitts, die eine starke Abhängigkeit des individuellen Beteiligungsverhaltens von der Schulklassenzugehörigkeit belegen, stellt sich die Frage, wie sehr das Schülerbeteiligungsverhalten im Zeitverlauf, d.h. von Stunde zu Stunde, aber auch innerhalb der Stunden, schwankt, da eine große zeitliche Stabilität des Merkmals den Einflussspielraum für Prozessbedingungen des Unterrichts, wie z.B. die hier fokussierten Aufgaben, einschränken würde.

Zeitliche Stabilität zwischen Unterrichtsstunden: Vor diesem Hintergrund soll als nächstes geklärt werden, wie konstant die durchschnittliche *on task*-Rate einer Klasse über aufeinanderfolgende Unterrichtsstunden hinweg bleibt. Als Prüfverfahren für

diesen Auswertungsschritt wird der Friedman-Test verwendet. Dieser Test ist eine nonparametrische Alternative zur Varianzanalyse mit Messwiederholung (Bühner & Ziegler, 2009; Leonhart, 2004) deren Anwendungsvoraussetzung der Normalverteilung an dieser Stelle verletzt wird und im Gegensatz zu den vorherigen Anwendungsfällen (vgl. Kap. 7.1.3/7.2.3) nicht durch die Stichprobengröße ($n = 13 - 25$, s.u.) kompensiert werden kann. Mithilfe des Friedman-Tests wird geprüft, ob die gemessenen Werte über die Messzeitpunkte hinweg statistisch signifikant schwanken. Tab. 12 zeigt die Ergebnisse dieser Prüfung. Zur Veranschaulichung sind die Werteverläufe der einzelnen Klassen in Abb. 6 dargestellt.

Tab. 12: Ergebnisse des Friedman-Tests zur zeitlichen Stabilität der on task-Rate im Klassendurchschnitt über die Unterrichtsstunden hinweg

Klasse	n ^a	df	χ^2	p
A	18	3	4,87	0,182
B	13	3	7,95	0,047
C	15	2	23,33	0,000
D	16	3	14,93	0,002
E	15	3	5,96	0,114
F	25	3	11,45	0,010

^a nur Fälle mit vollständigen Daten über die Messzeitpunkte (Unterrichtsstunden)

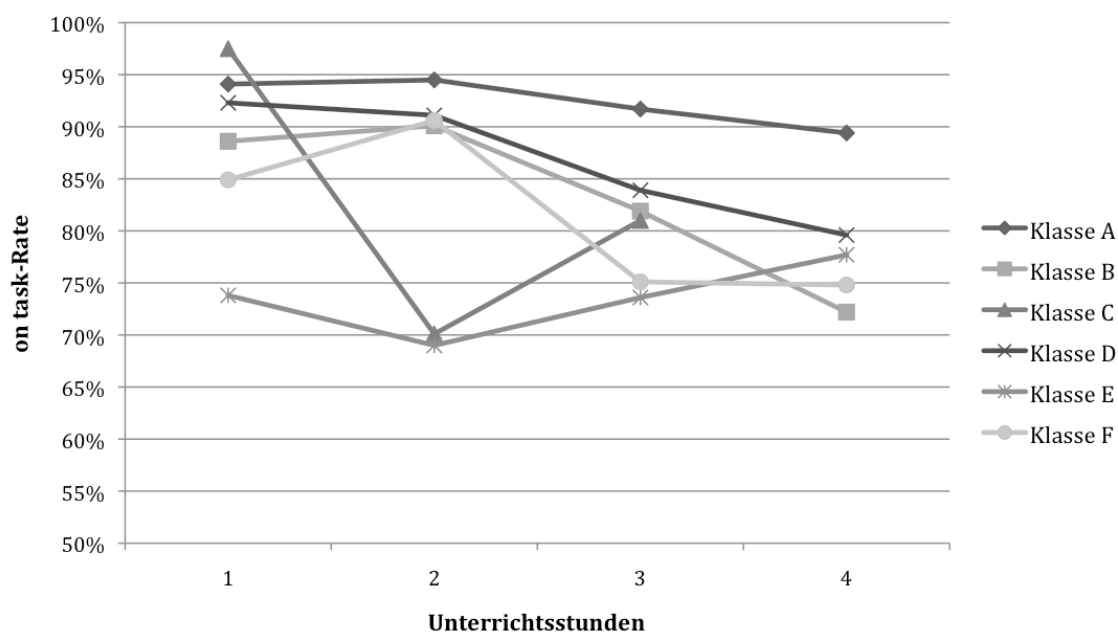


Abb. 6: Mittlere prozentuale on task-Rate der Klassen nach Messzeitpunkten (Unterrichtsstunden)

Wie die Kennwerte des Friedman-Tests belegen (Tab. 12), kann in Klasse A (χ^2 (df = 3, n = 18) = 4,87; p = 0,182) und Klasse E (χ^2 (df = 3, n = 15) = 5,96; p = 0,114) von einem relativ konstanten *on task*-Niveau gesprochen werden. Dem Liniendiagramm (Abb. 6) ist diesbezüglich zu entnehmen, dass die *on task*-Rate bei Klasse A über die Unterrichtsstunden hinweg auf einem beständig hohen (89,4 - 94,5 %) und bei Klasse E eher niedrigen Niveau (69 - 77,7 %) verläuft. In den übrigen vier Klassen hingegen bestehen statistisch signifikante Unterschiede zwischen den aufeinanderfolgenden Messzeitpunkten (Tab. 12) und damit bedeutsame Werteschwankungen über die Zeit. Das Schülerbeteiligungsverhalten kann daher nicht als zeitlich konstanter Unterrichtsfaktor bezeichnet werden, zumindest nicht mit diesem Maßstab des Zeitverlaufs.

Zeitliche Stabilität innerhalb der Unterrichtsstunden: Mit Blick auf die Ziele dieser Studie ist es jedoch auch wichtig zu klären, ob das Schülerbeteiligungsverhalten innerhalb des Unterrichtsverlaufs nennenswerten Schwankungen unterliegt, um diese in einem späteren Schritt durch andere Prozesseigenschaften des Unterrichts – das Aufgabenangebot – erklären zu können. Das Auswertungsverfahren ist komplementär zum vorangegangenen Schritt angelegt, d.h. die Prüfung erfolgt mithilfe des Friedman-Tests⁴³. Wie aus Tab. 13 zu entnehmen ist, sind für den Verlauf aller aufgezeichneten Unterrichtsstunden höchst signifikante Schwankungen der *on task*-Rate festzustellen. Das Schülerbeteiligungsverhalten erweist sich demnach auch mit Blick in den Unterrichtsprozess als zeitlich instabiles Merkmal. Auch wenn an diesem Punkt keine Differenzierung zwischen unterschiedlichen Verlaufscharakteristiken oder der Intensitäten der stundenspezifischen Wertschwankungen deutlich werden, kann dies dennoch als ausreichender Hinweis auf erklärungsbedürftige Variationen des Schülerbeteiligungsverhaltens im Unterrichtsprozess gewertet werden.

⁴³ Um an dieser Stelle Verzerrungen der Prüfergebnisse durch Lücken in den personenspezifischen Verlaufsdatenreihen (z.B. wegen eingeschränkter Sichtbarkeit des Schülers in einzelnen Beobachtungsintervallen) zu vermeiden, wurden die Daten für diese Analyse durch das Datenimputationsverfahren „Zeitreihen-Mittelwert“ komplettiert. Dabei wird der fehlende Wert durch den Mittelwert der Klasse zu dem jeweiligen Zeitpunkt ersetzt.

Tab. 13: Ergebnisse des Friedman-Tests zur zeitlichen Stabilität der on task-Rate im Stundenverlauf

Unterrichtsstunde	n ^a	df	χ^2	p
A 1103	19	23	76,74	0,000
A 1903I	18	25	62,27	0,000
A 1903II	18	37	67,63	0,002
A 0204	18	26	51,43	0,002
B 0903	18	30	100,30	0,000
B 3103	18	22	55,77	0,000
B 0204I	15	24	82,68	0,000
B 0204II	13	19	38,63	0,005
C 0303	17	22	62,61	0,000
C 1103	20	30	114,28	0,000
C 1603	15	31	109,00	0,000
D 1602	16	15	58,26	0,000
D 0603	24	27	144,11	0,000
D 1303I	20	38	137,26	0,000
D 1303II	26	21	111,67	0,000
E 0503	15	25	117,54	0,000
E 1003	26	24	86,07	0,000
E 1203	25	19	95,50	0,000
E 1703	21	19	55,29	0,000
F 1203	25	26	174,85	0,000
F 2703	27	34	159,47	0,000
F 3003I	25	29	91,94	0,000
F 3003II	26	18	97,63	0,000

^a nur Fälle mit vollständigen Daten über die Messzeitpunkte (Beobachtungsintervalle)

7.3 Aufgaben im Unterricht

7.3.1 Stichprobe

Die Aufgaben innerhalb des Unterrichtsverlaufs wurden in einem auf Ereignisstichproben basierenden Beobachtungsverfahren in ihrer zeitlichen Lage und Ausdehnung kodiert (vgl. Kap. 6.7.1) und anschließend ihre Eigenschaften mithilfe eines mehrdimensionalen Kategoriensystems eingeschätzt (vgl. Kap. 6.7.2). Da nicht alle aufgezeichneten Unterrichtsstunden hinsichtlich des Aufgabenangebots analysiert werden konnten, wurde aufgrund einer kombinierten Betrachtung der Stundenstrukturen und der Verlaufscharakteristik des Schülerbeteiligungsverhaltens eine Auswahl an Unterrichtsstunden getroffen: Häufige Sozialformwechsel, ein geringer Zeitanteil der Sozialform öffentlichen Unterrichts und das Fehlen markanter Wertschwankungen der *on task*-Rate im Unterrichtsverlauf (gemäß Sichtprüfung) bildeten die Ausschlusskriterien bei der Stundenauswahl. Darauf basierend wurden 16 Unterrichtsstunden in die Aufgabenanalyse einbezogen, wobei alle Klassen der Stichprobe beteiligt sind (vgl. Tab. 14). Die daraus resultierende Stichprobe umfasst insgesamt 165 Aufgaben, von denen 144 in Phasen des öffentlichen Unterrichts stattfanden. Im Durchschnitt waren 10 Aufgaben pro Stunde festzustellen, wobei einzelne Stunden deutlich davon abweichen. Zum Vergleich: Die Untersuchung von Neubrand (2002, S. 213), in der von einem breiteren Aufgabenverständnis ausgegangen wird als im vorliegenden Kontext (vgl. Kap. 4.2), kommt auf durchschnittlich 12,4 Aufgaben pro Unterrichtsstunde (nur deutsche Schulen). Jatzwauk, Rumann und Sandmann (2008, S. 272), in deren Rahmen nur Aufgaben in Schülerarbeitsphasen berücksichtigt wurden, weisen im Mittel 6,07 Aufgaben nach, während Bohl, Kleinknecht, Batzel und Richey (2012, S. 67) unter Verwendung einer ähnlichen Aufgabendefinition, aber unter Einbezug aller Unterrichtsphasen, zwischen 11 und 17 Aufgaben pro Stunde feststellen. Renkl (1991, S. 109) hingegen kommt mit einer sehr kleinteiligen, an Fragen orientierten Aufgabendefinition auf durchschnittlich ca. 53 Aufgaben pro Stunde. Und auch Schabram (2007), am Beispiel des Physikunterrichts, zeigt eine vergleichsweise hohe Aufgabenanzahl von durchschnittlich ca. 32 pro Unterrichtsstunde.

Zwischen der Aufgabenanzahl und der Verteilung der Sozialformen in den Unterrichtsstunden in den vorliegenden Daten besteht kein nennenswerter

Zusammenhang (Rangkorrelation nach Spearman zwischen der stundenspezifischen Aufgabenanzahl und dem Zeitanteil der ÖU-Phasen bzw. der Arbeitsphasen in der Stunde: $r_s = 0,10/0,16$, beide nicht signifikant).

Tab. 14: Aufgabenanzahl nach Unterrichtsstunden

Unterrichtsstunde	Anzahl Aufgaben	
	gesamte Stunde	innerhalb Sozialform ÖU
A 1103	12	11
A 1903I	12	10
A 1903II	6	6
A 0204	8	7
B 0903	11	11
B 3103	8	6
B 0204I	9	6
C 0303	15	14
C 1103	11	10
C 1603	13	12
D 1303I	18	16
D 1303II	5	4
E 1003	10	8
E 1703	9	5
F 1203	8	8
F 2703	10	10
Durchschnitt/Standardabweichung	10,3/3,3	9/3,3
Summe	165	144

7.3.2 Deskriptive Ergebnisse zu den Aufgabenmerkmalen auf Ebene der Einzelaufgabe

Einen Überblick zu den Ergebnissen der Aufgabenanalyse bieten Tab. 15 und Tab. 16. Dargestellt sind die Häufigkeitsverteilungen der Ausprägungsstufen der kategorialen Merkmale und deskriptive Statistiken für die metrischen Merkmale.

Tab. 15: Häufigkeitsverteilungen der kategorial erfassten Aufgabeneigenschaften (n = 165)

Aufgabenmerkmale und Ausprägungsstufen	Beobachtete Häufigkeit der Ausprägungsstufen	
	absolut	prozentual
<i>Merkmale der Aufgabenstellung/-einführung/-präsentation</i>		
Sprache (n = 163)		
Deutsch	48	29 %
Englisch	104	64 %
Kombination	11	7 %
andere	0	0 %
Anzahl Hilfestellungsmaßnahmen (n = 162)		
0	61	38 %
1	65	40 %
2	21	13 %
3	9	6 %
4	5	3 %
Medien: Darstellungsform (n = 163)		
Text	91	56 %
Tabelle/Diagramm	6	4 %
Abbildung/Grafik	11	7 %
Film	0	0 %
Hörbeispiel	11	7 %
Gegenstand	1	1 %
Kombination	18	11 %
kein	24	15 %
andere	1	1 %
Medien: Gerät/Material (n = 163)		
Tafel	40	25 %
Projektionsfolie	3	2 %
Arbeitsblatt	15	9 %
Buch	55	34 %
PC	0	0 %
Kombination	8	5 %
kein	24	15 %
andere	18	11 %
Bekanntheit d. Mediums (n = 138)		
bereits bekannt	92	67 %
simultan eingeführt	35	25 %
angekündigt	11	8 %

(Fortsetzung der Tabelle siehe nächste Seite)

(Fortsetzung von Tabelle 15)

Aufgabenmerkmale und Ausprägungsstufen	Beobachtete Häufigkeit der Ausprägungsstufen	
	absolut	prozentual
<i>Merkmale der Aufgabeninhalte</i>		
Lebensweltbezug (n = 161)		
ohne	39	24 %
konstruiert	76	47 %
authentisch	46	29 %
real	0	0 %
Abstraktheit (n = 161)		
konkret/alltäglich	92	57 %
abstrakter/nicht alltäglich	35	22 %
abstrakt	34	21 %
<i>Merkmale der Bearbeitungsanforderungen</i>		
Kognitiver Prozess (n = 142)		
Erinnern	27	19 %
Verstehen	51	36 %
Anwenden	42	30 %
Analysieren	15	11 %
Evaluieren	6	4 %
Gestalten	1	1 %
Wissensdimension (n = 141)		
Fakten	44	31 %
Konzepte	28	20 %
Prozeduren	67	48 %
Metakognitive Strategien	2	1 %
Format d. Produkts (n = 141)		
Verarbeitung von Vorgaben	66	47 %
eingeschränkte Konstruktion	48	34 %
erweiterte Konstruktion	27	19 %
Art d. Produkts (n = 157)		
Sprache	125	80 %
Schrift	18	12 %
Grafik	2	1 %
Kombination	4	3 %
andere	8	5 %
Reflexions-/Auswertungsfunktion (n = 165)		
keine	142	86 %
Refl./Ausw. einer Unterrichtsaufgabe	13	8 %
Refl./Ausw. einer Hausaufgabe	10	6 %

Tab. 16: Deskriptive Kennwerte zu den metrischen skalierten Aufgabeneigenschaften

Aufgabenmerkmale	M	Md ^a	SD	Min	Max
Verbaler Umfang d. Aufgabenstellung ^b (n = 163)	36,6	22,0	43,3	2	325
Verbaler Umfang d. Medieninhalte ^b (n = 126)	69,5	39,0	88,9	0	432
Zeitlicher Umfang (in mm:ss; n = 165)					
Aufgabenstellung	00:28	00:12	00:48	00:01	05:43
Aufgabenbearbeitung	02:49	01:55	02:50	00:03	15:34
Gesamte Aufgabe	03:18	02:21	03:17	00:04	17:47

^a Md = Median ^b gemessen an der Anzahl an Wörtern

Von der Gesamtstichprobe abweichende Stichprobenumfänge bei den einzelnen Merkmalen (s. Werte in Klammern) kommen durch die jeweilige Häufigkeit des Codes *nicht bestimmbar* zustande (vgl. Kap. 6.7.2). Es ist anzumerken, dass in der gesamten Stichprobe lediglich eine Aufgabe identifiziert werden konnte, bei der von der Lehrkraft verschiedene Varianten des Arbeitsauftrages für unterschiedliche Teilgruppen der Klasse formuliert wurden und diese parallel zueinander zu bearbeiten waren – sozusagen als Maßnahme zur Differenzierung des Unterrichts. Dieser Fall wurde aufgrund der problematischen Kategorisierbarkeit mithilfe des vorliegenden Systems aus der Stichprobe entfernt.

Beschreibung der Ausprägungshäufigkeiten: Mit Blick auf die Ergebnisse zu den Aspekten der Einführung und Präsentation der Aufgabe in den Unterricht ist festzustellen, dass die Lehrkräfte die Aufgabenstellungen überwiegend in englischer Sprache formulierten (64 % der Aufgaben, s. Tab. 15), während nur knapp ein Drittel der Aufgaben (29 %) auf Deutsch eingeführt und erläutert wurde. Der Anteil englischer Sprache fällt damit niedriger aus als jener im Rahmen der DESI-Studie, wo in einer Teiluntersuchung mittels videobasierter Beobachtungen im Englischunterricht der 9. Jahrgangsstufe ein Anteil von 84 % englischsprachiger Äußerungen der Lehrkräfte (allerdings im gesamten Unterricht) herausgefunden wurde (Helmke, Helmke, Schrader, Wagner, Nold & Schröder, 2008).

Der Umfang der verbalen Ausführungen der Lehrkraft bei der Präsentation der Aufgabe, gemessen an der Wortanzahl (Tab. 16), beträgt durchschnittlich 70 Wörter, differiert aber sehr stark und zeigt erhebliche Extremwerte: Die Messwerte zwischen 100 und

dem Maximum von 325 Wörtern entfallen auf lediglich neun Fälle, während bei 75 % der Aufgaben von der Lehrkraft maximal 43 Wörter zur Einführung genutzt wurden.

Das allgemeine Niveau an beobachteten Hilfestellungsmaßnahmen in den Anweisungen und Erläuterungen der Lehrkraft fällt eher niedrig aus. Bei gut einem Drittel der Aufgaben wurden keinerlei unterstützenden Hinweise in die Formulierungen integriert. 90 % der Aufgabenstellungen enthielten maximal zwei unterschiedliche Arten der Hilfestellung durch die Lehrkraft – in den meisten Fällen eine Wiederholung bzw. Paraphrasierung der Aufgabenstellung oder eine Nennung von Beispielen.

Ein Großteil der untersuchten Aufgaben (85 %, s. Tab. 15) steht in Verbindung mit einem Unterrichtsmedium bzw. stützt sich auf mediengebundene Informationen. In den meisten Fällen betrifft dies textförmige Informationen (56 %), seltener grafische bzw. schematische Darstellungen (Tabellen, Diagramme, Grafiken, Abbildungen: 11 %), eine Kombination aus mindestens zwei Darstellungsformen (11 %) oder Hörbeispiele (7 %), die trotz der geringen Gesamthäufigkeit in fast allen Lerngruppen – in vier der sechs untersuchten Klassen – zu finden waren. Träger dieser Informationen bzw. Darstellungen waren in den meisten Fällen das Lehrbuch (34 %) oder die Tafel (25 %). Wie die Ausprägungshäufigkeiten des Merkmals „Medienbekanntheit“ zeigen, wurden etwa zwei Drittel der Medien (67 %) über mehrere Aufgaben hinweg als Bezugspunkt verwendet. Die Kodierung eines neuen Mediums erfolgte auch beim Wechsel des Aufgabenblocks im Buch oder auf dem Arbeitsblatt, da sich dabei in der Regel mehrere Bedingungen für die Aufgabe ändern (z.B. Art und Umfang der Abbildungen, kognitive Anforderungen der Bearbeitung). Dementsprechend lässt sich feststellen, dass durchschnittlich bei jeder vierten Aufgabe das Medium, d.h. die mediengebundene Informationsgrundlage, gewechselt wurde (simultane Einführung des Mediums bei Aufgabenstellung: 25 %). Bei einem kleinen Teil der Aufgaben (8 %) wurden von der Lehrkraft in der Aufgabenstellung medienbasierte Informationen angekündigt, das Medium aber erst im Laufe der Aufgabenbearbeitung eingeführt. Der Umfang der textförmigen Informationen in den verwendeten Medien (s. Tab. 16) variiert im Bereich zwischen 0 (entspricht rein grafischen Darstellungen) und 432 Wörtern und liegt im Durchschnitt bei 70 Wörtern. Drei Viertel der in den Medien abgebildeten Texte übersteigen nicht einen Umfang von 100 Wörtern.

Im Bereich der Eigenschaften der Aufgabeninhalte zeigt sich, dass knapp ein Viertel der analysierten Aufgaben (24 %, s. Tab. 15) keinerlei inhaltlichen Bezug zur Lebenswelt der Schüler aufweist. Aufgaben dieser Art beziehen sich in der Regel auf rein fachliche Phänomene, wie z.B. Grammatikregeln. Ein weitaus größerer Anteil der Aufgaben (47 %) stellt konstruierte Lebensweltbezüge her, meist in Form von Beispielsituationen aus dem Leben fiktiver Personen. Bezüge zur tatsächlichen Lebenswelt der Schüler (Ausprägungsstufe „authentisch“) sind in 29 % der Fälle erkennbar, wohingegen keine der untersuchten Aufgaben einen „realen“ Lebensweltbezug aufweist. Diese Ergebnisse decken sich mit der bei Kleinknecht, Maier, Metz und Bohl (2011) auf Basis einer Erprobungsstudie berichteten Ausprägungstendenz dieser Variable, wo von „*selten auftretenden höheren Kategorien*“ (S. 340) gesprochen wird (ohne Angabe von Kennwerten).

Hinsichtlich der Abstraktheit der Aufgabeninhalte bleibt mehr als die Hälfte der Aufgaben (57 %) auf der Ebene konkret beobachtbarer Aspekte des Alltags, z.B. in Form der Beschreibung von Ereignissen oder Objekten, während jeweils etwa 20 % auf die Ausprägungsstufen „abstrakter“ (z.B. Emotionen fiktiver Charaktere) und „abstrakt“ (z.B. grammatische Strukturen) entfallen.

Mit Blick auf die Anforderungen bei der Bearbeitung der Aufgabe ist für das Merkmal der kognitiven Prozesse eine quantitative Dominanz der Ausprägungsstufen „Verstehen“ und „Anwenden“ zu konstatieren, die zusammengefasst 66 % der Aufgaben abdecken. Die beiden in der Lernzieltaxonomie von Anderson et al. (2001) hierarchisch am höchsten angesiedelten kognitiven Prozesse „Evaluieren“ und „Gestalten“ nehmen dagegen gemeinsam nur 5 % der untersuchten Aufgaben ein. In Analogie zu diesen Ergebnissen zeigt sich für die zur Bearbeitung erforderlichen Wissensdimensionen der mehrheitliche Teil der Aufgaben im Bereich des Wissens über Prozeduren (48 %), gefolgt von Faktenwissen (31 %) und dem Wissen über Konzepte (20 %). Die mit 1 % als marginal zu bezeichnende Häufigkeit der Ausprägungsstufe „Wissen über metakognitive Strategien“ schließt an die Erfahrungen von Kleinknecht, Maier, Metz und Bohl (2011) an, die im Rahmen einer Erprobungsstudie zu ihrem Aufgabenanalyzesystem ebenfalls von einer sehr geringen Häufigkeit dieser Ausprägung berichten (ohne Angabe von Kennwerten). Die Häufigkeit von Anwendungsprozessen bzw. dem geforderten Wissen über Prozeduren in den vorliegenden Ergebnissen steht im Kontrast zu Befunden zum Physikunterricht von Schabram (2007), wo

Anwendungsaufgaben mit ca. 6 % die Minderheit darstellen und dafür so genannte Wissensaufgaben (ca. 60 %) klar im Vordergrund stehen, oder aus dem Bereich des Biologieunterrichts, wo auch eher Wissens- und Verstehensaufgaben überwiegen (vgl. Jatzwauk, Rumann & Sandmann, 2008).

Ein weiteres Kennzeichen der Bearbeitungsanforderungen ist das Format des geforderten Produkts. Die Ergebnisse dazu zeigen, dass bei fast der Hälfte der Aufgaben (47 %) die Anforderung darin bestand, vorgegebene Elemente zu verarbeiten, d.h. vorliegende Informationen wiederzugeben, zu ordnen oder zu selektieren. Etwa ein Drittel der Aufgaben (34 %) erforderte die Eigenkonstruktion eines Wortes oder Satzes, 19 % der Aufgaben die Eigenkonstruktion ausführlicherer Produkte wie z.B. freie Erzählungen. Sehr ähnliche Verteilungen der geforderten Antwortformate werden nachgewiesen in einer Studie von Jatzwauk, Rumann und Sandmann (2008).

Hinsichtlich der Art des geforderten Produkts überwiegen ganz deutlich sprachförmige Äußerungen (80 %); schriftliche Produkte (z.B. Mitschriften) wurden in nur 12 % der Fälle gefordert, was aufgrund der großen Anteile öffentlichen Unterrichts in den untersuchten Stunden nicht überrascht.

Die mittels Ereigniskodierung sekundengenau erfasste Dauer der Aufgaben zeigt sowohl für Aufgabenstellungen als auch Aufgabenbearbeitungen eine große Spanne des Wertebereichs (s. Tab. 16). Wie ein Überblick über die Verteilungen zeigt, sind diese Ergebnisse allerdings stark durch Extremwerte beeinflusst (vgl. Abb. 7). Die Verteilungen sind zudem erkennbar linkssteil und weichen deutlich von der Normalverteilung ab (Kolmogorov-Smirnov-Test: $p < 0,001$). Das Hauptgewicht der Dauer der Aufgabenstellungen liegt im Bereich unterhalb einer halben Minute ($M = 00:28^{44}$, $Md = 00:12$; 75 % - Perzentil = 00:31), jenes der Aufgabenbearbeitungen bei knapp vier Minuten ($M = 02:49$; $Md = 01:55$; 75 % - Perzentil = 03:49), und daraus resultierend für die Gesamtaufgabendauer unterhalb von fünf Minuten ($M = 03:18$; $Md = 02:21$; 75 % - Perzentil = 04:14).

⁴⁴ Zeitangaben im Format mm:ss

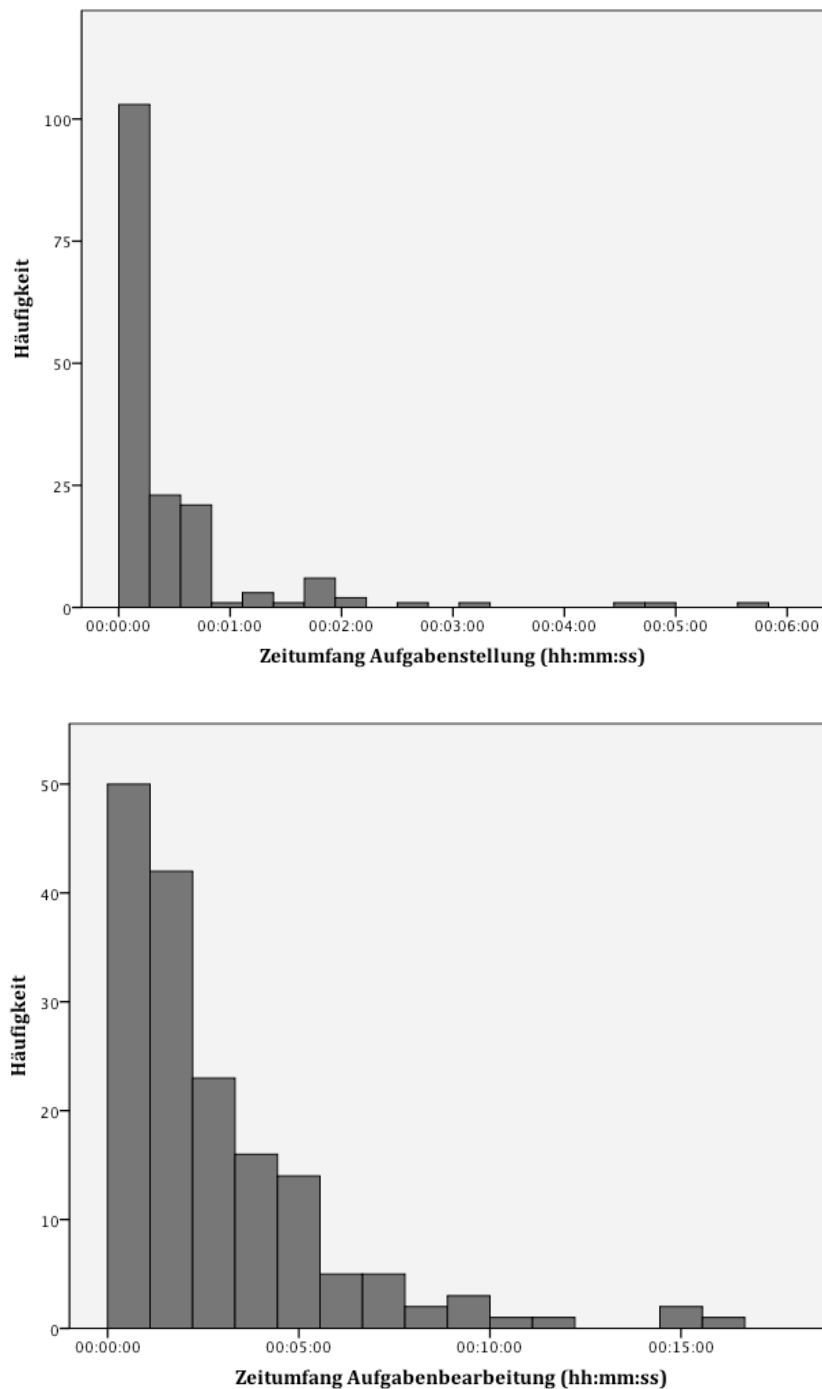


Abb. 7: Häufigkeitsverteilung des zeitlichen Umfangs von Aufgabenstellung und -bearbeitung

Die im Rahmen der Instrumentenkonstruktion zusätzlich eingeführte Identifizierung einer möglichen Reflexions- bzw. Auswertungsfunktion (vgl. Kap. 6.7.2) zeigt, dass sich zusammengefasst 23 (entspricht 14 %, s. Tab. 15) der untersuchten Aufgaben auf eine Hausaufgabe oder eine bereits zuvor innerhalb des Unterrichts bearbeitete Aufgabe beziehen, indem z.B. Lösungsvorschläge einzelner Schüler bewertet,

Bearbeitungsschwierigkeiten diskutiert oder bestimmte Aufgabenbestandteile gemeinsam im Plenum gelöst werden. Aufgrund der damit verbundenen eingeschränkten Beurteilbarkeit mancher Aufgabeneigenschaften resultieren aus den derart kategorisierten Aufgaben teilweise merkmalspezifische fehlende Werte, wie z.B. im Bereich der Bearbeitungsanforderungen. Darüber hinaus soll die Variable aber aufgrund des explorativen Charakters der Untersuchung im weiteren Verlauf der Auswertungen mit berücksichtigt werden.

Unterschiede der Aufgabenmerkmalsausprägungen zwischen den Unterrichtsstunden: Im nächsten Schritt soll geklärt werden, ob bzw. wie stark sich die zuvor überblicksartig beschriebenen Ausprägungen der Aufgabenmerkmale zwischen den Unterrichtsstunden unterscheiden. Für die nominalskalierten Merkmale wird diese Frage mithilfe eines χ^2 -Unabhängigkeitstests beantwortet (Bortz, 2010). Erweisen sich die beiden Variablen „Aufgabenmerkmal“ und „Unterrichtsstunde“ in dieser Prüfung als voneinander abhängig, kann dies als Beleg für nennenswerte Unterschiede der Aufgabenmerkmalsausprägungen zwischen den Unterrichtsstunden gewertet werden (ebd., S. 137)⁴⁵. Für die Merkmale auf metrischem Skalenniveau erfolgt die Prüfung der Ausprägungsunterschiede mithilfe einfaktorieller, univariater Varianzanalysen, mit den Unterrichtsstunden als 16-stufigem Faktor und dem jeweiligen Aufgabenmerkmal als abhängige Variable.

Von diesem Punkt an werden aus Gründen der Vergleichbarkeit der Prüfergebnisse alle kategorialen – und damit auch die ordinalskalierten – Aufgabenmerkmale als nominalskalierte Variablen behandelt. Da bei manchen Merkmalen die Rangcharakteristik der Skalierung nicht als gesichert angesehen werden kann (z.B. Aspekte der kognitiven Anforderungen nach Anderson et al., 2001, oder „Format des geforderten Produkts“) und in dieser Studie bezüglich der Aufgabeneigenschaften ohnehin keine gerichteten Hypothesen geprüft werden, stellt diese Herabsetzung des Messniveaus für die Bearbeitung der verfolgten Fragestellungen keinen Verlust dar. In diesem Zusammenhang wurden auch die Variablen „verbaler Umfang der Einführung

⁴⁵ Da bei den meisten Merkmalen mehr als 20 % der Zellen der Tabellenmatrix erwartete Häufigkeiten von weniger als fünf aufweisen, wurde der exakte Test nach Fisher durchgeführt (vgl. Bortz, 2010, S. 141; Leonhart, 2004, S. 159).

und Präsentation der Aufgabe“, „Anzahl der Hilfestellungsmaßnahmen“ und „verbaler Umfang der Medieninhalte“ mittels Bildung von Perzentilgruppen umkodiert in drei- bzw. vierfach gestufte Variablen.

Tab. 17: Ergebnisse der Prüfung auf Unterschiede in den Aufgabenmerkmalsausprägungen zwischen den Unterrichtsstunden

Nominalskalierte Aufgabenmerkmale	χ^2 -Unabhängigkeitstests			
	N	df	χ^2	p
<i>Aufgabenstellung/-einführung/-präsentation</i>				
Sprache	163	30	102,8	0,000
Verbaler Umfang	163	45	67,5	0,016
Anzahl Hilfestellungsmaßnahmen	162	30	23,7	0,784
Medium: Darstellungsform	163	105	237,2	0,000
Medium: Gerät/Material	163	90	216,0	0,000
Bekanntheit d. Mediums	138	45	88,2	0,000
Verbaler Umfang d. Medieninhalte	126	45	112,1	0,000
<i>Aufgabeninhalte</i>				
Lebensweltbezug der Inhalte	161	30	114,6	0,000
Abstraktheit der Inhalte	161	30	138,2	0,000
<i>Bearbeitungsanforderungen</i>				
Kognitive Prozesse der Bearbeitung	142	75	111,3	0,004
Erforderliche Wissensdimensionen	141	45	109,6	0,000
Format d. Produkts	141	30	81,4	0,000
Art d. Produkts	157	60	94,6	0,003
Reflexions-/Auswertungsfunktion	165	30	62,4	0,000
Metrisch skalierte Aufgabenmerkmale	Varianzanalysen ^a			
	N	df	F	p
Zeitumfang gesamte Aufgabe	165	15/149	1,5	0,180
Zeitumfang Aufgabenstellung	165	15/149	0,8	0,665
Zeitumfang Aufgabenbearbeitung	165	15/149	1,5	0,118

^a einfaktoriell, univariat

Wie die Ergebnisse der Prüfung (Tab. 17) zeigen, fallen die χ^2 -Tests auf Seiten der nominalskalierten Aufgabenmerkmale in fast allen Fällen hoch oder höchst signifikant aus ($p < 0,01$; außer „verbaler Umfang der Aufgabenstellung“: $p < 0,05$). Diesen Ergebnissen zufolge unterscheiden sich die Unterrichtsstunden statistisch bedeutsam in

den Ausprägungen dieser Aufgabeneigenschaften. Die einzige Ausnahme bildet das Merkmal „Anzahl der Hilfestellungsmaßnahmen“ ($p = 0,784$). Der enge Wertebereich dieses Merkmals (vgl. deskriptive Ergebnisse, Tab. 15) lässt offenbar keinen Spielraum für nennenswerte Wertdifferenzen zwischen den Unterrichtsstunden. Im Unterschied zum Großteil der kategorialen Aufgabenmerkmale zeigen die Aspekte des zeitlichen Aufgabenumfangs keine bedeutsamen Unterschiede zwischen den Unterrichtsstunden: Keine der Varianzanalysen unterschreitet nur annähernd das 5 %-Signifikanzniveau. Somit ist festzuhalten, dass jene Merkmale, die die qualitativen Eigenschaften der Aufgaben abbilden (vgl. Kap. 6.7.2), in einem bedeutsamen Ausmaß zwischen den untersuchten Unterrichtsstunden variieren, während Attribute der Aufgabenquantität – die zeitlichen Merkmale – dies nicht erkennen lassen.

Wechselseitige Zusammenhänge zwischen den Aufgabenmerkmalen: Um zu erfahren, in welchem Ausmaß die Ausprägungen bestimmter Aufgabenmerkmale gemeinsam variieren, wurden mithilfe unterschiedlicher Kennwerte die Zusammenhänge zwischen allen Aufgabenmerkmalen geprüft (nominalskaliert: Cramers V , intervallskaliert: Produkt-Moment-Korrelation nach Pearson; nominalskaliert * intervallskaliert: η^2). Die Ergebnisse (s. Tab. 18) zeigen eine Reihe mittlerer und auch einige große Zusammenhänge, die bei genauerer Betrachtung (z.B. der von der erwarteten Häufigkeit abweichende Zellen in der Kreuztabellierung bei Cramers V bzw. den Mittelwertunterschieden bei η^2) in der überwiegenden Mehrheit erwartbare Effekte belegen. So zeigen die Korrelationen zwischen den Zeitmerkmalen ein angesichts der Häufigkeitsverteilungen (vgl. Tab. 16) abzusehendes Bild, indem die Dauer der Aufgabenbearbeitung mit der Dauer der Aufgabenstellung hoch ($r = 0,45$; $p \leq 0,01$; s. Tab. 18) und mit der Gesamtdauer der Aufgabe jedoch annähernd vollständig korreliert ($r = 0,98$; $p \leq 0,01$). In ebenfalls erwartbarer Weise steigt mit zunehmendem Zeitumfang der Aufgabe – insbesondere der Aufgabenstellung – der Umfang der verbalen Ausführungen der Lehrkraft in der Aufgabenstellung ($\eta^2 = 0,25$; $p \leq 0,001$) sowie die Anzahl an gebotenen Hilfestellungshinweisen ($\eta^2 = 0,18$; $p \leq 0,001$). Weitere, leicht zu erklärende Zusammenhänge sind im Bereich der Medien festzustellen: Textförmige Darstellungen werden überwiegend an der Tafel oder auf einem Arbeitsblatt dargeboten, Kombinationen der Darstellungsformen (z.B. Text mit Abbildung oder Foto) hingegen überwiegend im Lehrbuch (Cramers $V_{\text{Mediendarstellungsform} * \text{Mediengerät/-material}} = 0,59$;

$p \leq 0,001$); die Nutzung der Tafel impliziert meist eine „simultane Einführung“, während Abbildungen auf Arbeitsblättern oder im Lehrbuch eher über mehrere Aufgaben hinweg genutzt werden (Cramers $V_{\text{Mediengerät/-material*Medienbekanntheit}} = 0,63$; $p \leq 0,001$). An dieser Stelle ist festzuhalten, dass auffällig viele Aufgabenmerkmale in nennenswertem Ausmaß mit Aspekten der Mediennutzung zusammenhängen. Wird im Rahmen der Aufgabe keinerlei Medium eingesetzt, was für eine rein auf Gesprächsebene vollzogene Aufgabe spricht, so ist der Lebensweltbezug der Aufgabeninhalte häufiger „authentisch“ und das Format des geforderten Produkts entspricht einer „eingeschränkten Konstruktion“ (Cramers $V_{\text{Mediengerät/-material*Lebensweltbezug}} = 0,34$; $p \leq 0,001$; Cramers $V_{\text{Mediengerät/-material*Format d. Produkts}} = 0,36$; $p \leq 0,001$;) – eine Kombination, die sich in Form von Fragen wie z.B. „*When you go to the shopping centre, what do you buy there?*“ zeigt. Weitere im Bereich des Erwartbaren liegende Zusammenhänge zeigen sich zudem zwischen den beiden aus der Lernzieltaxonomie von Anderson et al. (2001) entnommenen Merkmalen zu den kognitiven Bearbeitungsanforderungen, die ein überproportional häufiges Auftreten der Kombinationen „erinnern von Fakten“, „verstehen von Konzepten“ und „anwenden von Prozeduren“ aufweisen (Cramers $V_{\text{Kognitiver Prozess*Wissensdimension}} = 0,44$; $p \leq 0,001$), und den beiden Merkmalen im Bereich der Aufgabeninhalte, wo häufig die Ausprägungskombinationen „konkret“ mit „konstruiertem Lebensweltbezug“, „leicht abstrakt“ mit „authentischem Lebensweltbezug“ und „abstrakt“ „ohne Lebensweltbezug“ zu finden sind (Cramers $V_{\text{Lebensweltbezug*Abstraktheit}} = 0,46$; $p \leq 0,001$). Ein weiterer erwähnenswerter Zusammenhang zeigt sich zwischen der Sprache der Aufgabenstellung und der Abstraktheit der Aufgabeninhalte: Aufgaben mit abstrakten Inhalten (z.B. Grammatikaufgaben) werden deutlich häufiger in deutscher Sprache eingeführt, solche mit konkreten Inhalten dagegen deutlich mehr auf Englisch (Cramers $V_{\text{Sprache Aufgabenstellung*Abstraktheit}} = 0,43$; $p \leq 0,001$).

Tab. 18: Zusammenhänge zwischen den Aufgabenmerkmalen

Aufgabenmerkmal	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>- nominalskaliert</i>																
1 Sprache Aufgabenstellung																
2 Verbaler Umfang Aufgabenstellung	0,22*															
3 Hilfestellungen bei Aufgabenstellung	0,16	0,57***														
4 Medium: Darstellungsform	0,35***	0,26	0,26													
5 Medium: Gerät/Material	0,17	0,25*	0,16	0,59***												
6 Bekanntheit Medium	0,11	0,29***	0,28***	0,64***	0,63***											
7 Verbaler Umfang Medieninhalte	0,09	0,17	0,11	0,32***	0,27*	0,25*										
8 Lebensweltbezug Aufgabeninhalte	0,29***	0,11	0,17	0,37***	0,34***	0,28***	0,33***									
9 Abstraktheit Aufgabeninhalte	0,43***	0,20*	0,07	0,30*	0,22	0,14	0,26**	0,46***								
10 Kognitiver Prozess Bearbeitung	0,32***	0,23	0,23	0,28*	0,29***	0,36***	0,25	0,21	0,22							
11 Wissensdimension Bearbeitung	0,16	0,21*	0,14	0,28*	0,25	0,18	0,21	0,13	0,32***	0,44***						
12 Format d. Produkts	0,13	0,10	0,15	0,32**	0,36***	0,32***	0,14	0,18	0,19*	0,24	0,23*					
13 Art d. Produkts	0,25*	0,22*	0,21	0,27*	0,35***	0,22*	0,21	0,09	0,11	0,27**	0,16	0,20				
14 Reflexions-/Auswertungsfunktion	0,15	0,07	0,14	0,27	0,29**	0,23**	0,20	0,13	0,17	0,32**	0,16	0,22**	0,22**			
<i>- intervallskaliert</i>																
15 Dauer gesamte Aufgabe	0,01	0,13***	0,08***	0,05	0,07	0,08**	0,12***	0,02	0,04	0,04	0,03	0,00	0,27***	0,00		
16 Dauer Aufgabenstellung	0,01	0,25***	0,18***	0,04	0,03	0,12***	0,05	0,01	0,03	0,05	0,05	0,01	0,14***	0,01	0,64**	
17 Dauer Aufgabebearbeitung	0,01	0,08**	0,04*	0,05	0,07	0,05*	0,12**	0,02	0,03	0,05	0,02	0,00	0,25***	0,01	0,98**	0,45**

Koeffizienten:

nominal * nominal: Cramers-V (kleiner Effekt = 0,10; mittlerer Effekt = 0,30; großer Effekt = 0,50)

intervall * intervall: Produkt-Moment nach Pearson (kleiner Effekt = 0,10; mittlerer Effekt = 0,30; großer Effekt = 0,50)

nominal * intervall: η^2 (kleiner Effekt = 0,01; mittlerer Effekt = 0,06; großer Effekt = 0,14)Signifikanzkennzeichnung: * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; keine Kennzeichnung $p > 0,05$ (zweiseitig)

7.3.3 Deskriptive Ergebnisse zur Variabilität der Aufgabenmerkmale auf Ebene der Unterrichtsstunden

Die Ergebnisse der vorangegangenen beiden Abschnitte fokussieren die Eigenschaften der einzelnen Aufgabe. Dies geschieht mit der Absicht, Aufgaben aus einer Mikroperspektive heraus als Prozessqualitätskomponente des Unterrichts zu beschreiben und im nächsten Schritt den Einfluss der Ausprägungsvariationen bei Einzelaufgaben auf das Beteiligungsverhalten der Schüler zu prüfen. Auf dieser Grundlage sind allerdings solche Effekte schwierig zu berücksichtigen, die sich möglicherweise aus dem Aufgabengefüge der gesamten Unterrichtsstunde heraus ergeben, beispielsweise durch die Abfolge bestimmter Aufgabenmerkmalsausprägungen oder die besondere Gewichtung spezieller Ausprägungen in der Chronologie der Stunde. Im Hinblick auf den zentralen theoretischen Bezugspunkt dieser Untersuchung, das QAIT-Model effektiven Unterrichts nach Slavin (1994), ist diesbezüglich von besonderem Interesse, inwiefern das Aufgabenangebot einer Unterrichtsstunde Anreize (*incentives*, ebd.) setzt, die die Wahrscheinlichkeit der Beteiligung der Schüler an den Lerngelegenheiten erhöhen, zum Beispiel durch eine gewisse Interessantheit in Form eines Abwechslungsreichtums der Aufgabeneigenschaften oder durch gleiche Bewältigungschancen für Schüler mit unterschiedlichen Lernvoraussetzungen mittels Variabilität der Anforderungen.

Generierung von Variabilitätsindikatoren: Vor diesem Hintergrund wurden auf Grundlage der Daten zu den Eigenschaften der Einzelaufgaben drei verschiedene Indikatoren zur Abbildung der stundenspezifischen Variabilität des Aufgabenangebots gebildet:

1. *Variationsbreite*

Dieser Indikator gibt die Anzahl der beobachteten Ausprägungsstufen eines Merkmals in der betreffenden Stunde wieder. Liegen beispielsweise drei unterschiedliche kognitive Prozesse der Aufgabenbearbeitung in einer Stunde vor, wird für den Indikator der Wert „3“ erfasst (s. Datenbeispiel, Tab. 19). Der Indikator bildet damit die vorhandene Spannweite eines Merkmals in einer Unterrichtsstunde ab. Der Wertebereich dieses Indikators definiert sich durch die Anzahl an Ausprägungsstufen des jeweiligen Merkmals.

2. Gleichförmigkeit

Zur Berücksichtigung des Verhältnisses der Variationsbreite zur Auftretenshäufigkeit der Merkmalsstufen wurde zudem ein Gleichförmigkeitsindikator erzeugt. Dieser entspricht dem Häufigkeitsanteil des Modalwerts an der Aufgabenmenge in der jeweiligen Stunde, womit der Wertebereich dieses Indikators sich zwischen 0 und 1 bewegt (s. Datenbeispiel in Tab. 19: sechs von zwölf Aufgaben in der Kategorie „Anwenden“ = 0,5). Anhand dieses Indikators ist abzulesen, wie stark eine bestimmte Ausprägungsstufe das Aufgabenmerkmal über den Stundenverlauf hinweg dominiert bzw. ob es Monotonie erkennen lässt.

3. Wechselhaftigkeit

Mithilfe dieses Indikators wird berücksichtigt, wie häufig die Ausprägungen eines Aufgabenmerkmals innerhalb einer Unterrichtsstunde variiert werden. Dazu wird die Anzahl an Ausprägungsveränderungen von einer zur nächsten Aufgabe innerhalb der Unterrichtsstunde erfasst (s. Datenbeispiel, Tab. 19). Der Wertebereich dieses Indikators ist begrenzt durch die Anzahl an Aufgaben in der jeweiligen Unterrichtsstunde.

Tab. 19: Datenbeispiel zur Illustration der Berechnung der Indikatoren der stundenspezifischen Aufgabenvariabilität

Aufgabenabfolge einer beispielhaften Unterrichtsstunde	Beispielhaftes Aufgabenmerkmal: Kognitiver Prozess der Bearbeitung
1	Anwenden
2	Verstehen
3	Verstehen
4	Verstehen
5	Erinnern
6	Erinnern
7	Anwenden
8	Anwenden
9	Anwenden
10	Erinnern
11	Anwenden
12	Anwenden
↓	↓
Variabilitätsindikatoren	Indikatorenausprägung
Variationsbreite	3
Gleichförmigkeit	0,5
Wechselhaftigkeit	5

Ausgenommen von der Indikatorenbildung sind die Merkmale des zeitlichen Umfangs, da sie nicht zu den qualitativen Kennzeichen des Aufgabenangebots zählen (vgl. Kap. 6.7.2), und das Merkmal „Bekanntheit des Mediums“, welches bereits eine unterrichtsverlaufsabhängige Eigenschaft der Aufgaben abbildet und daher die daraus generierten Variabilitätsindikatorenwerte nicht logisch zu interpretieren wären. Grundlage zur Berechnung der Indikatorenwerte waren alle erfassten Aufgaben (n = 165). Einen Überblick zur Lage und Streuung der auf dieser Grundlage generierten Indikatorenwerte in der Stichprobe liefert Tab. 20.

Tab. 20: Deskriptive Kennwerte zur Ausprägung der Indikatoren zur Beschreibung der stundenspezifischen Variabilität (n = 16 Unterrichtsstunden)

Aufgabenmerkmale und Variabilitätsindikatoren	M	Md	SD	Min	Max
<i>Merkmale der Aufgabenstellung/-einführung/-präsentation</i>					
Sprache					
Variationsbreite	2,06	2,00	0,85	1,00	3,00
Gleichförmigkeit	0,80	0,89	0,21	0,38	1,00
Wechselhaftigkeit	2,50	2,00	2,31	0,00	6,00
Verbaler Umfang					
Variationsbreite	3,44	3,00	0,51	3,00	4,00
Gleichförmigkeit	0,47	0,46	0,13	0,27	0,71
Wechselhaftigkeit	6,38	6,00	3,56	2,00	17,00
Hilfestellungsmaßnahmen					
Variationsbreite	3,00	3,00	0,00	3,00	3,00
Gleichförmigkeit	0,48	0,45	0,09	0,38	0,64
Wechselhaftigkeit	6,56	6,00	2,94	2,00	15,00
Medien: Darstellungsform					
Variationsbreite	2,50	2,00	1,27	1,00	5,00
Gleichförmigkeit	0,72	0,75	0,22	0,36	1,00
Wechselhaftigkeit	2,94	3,00	2,11	0,00	7,00
Medien: Gerät/Material					
Variationsbreite	2,94	3,00	0,85	2,00	4,00
Gleichförmigkeit	0,56	0,50	0,18	0,25	0,91
Wechselhaftigkeit	4,00	4,00	1,55	2,00	7,00
Verbaler Umfang d. Medieninhalte					
Variationsbreite	2,50	2,00	0,82	1,00	4,00
Gleichförmigkeit	0,58	0,50	0,16	0,37	1,00
Wechselhaftigkeit	3,13	3,00	2,55	0,00	10,00

- Fortsetzung der Tabelle s. nächste Seite -

- Fortsetzung von Tab. 20 -

Aufgabenmerkmale und Variabilitätsindikatoren	M	Md	SD	Min	Max
<i>Merkmale der Aufgabeninhalte</i>					
Lebensweltbezug					
Variationsbreite	2,19	2,00	0,75	1,00	3,00
Gleichförmigkeit	0,72	0,68	0,16	0,50	1,00
Wechselhaftigkeit	3,19	2,50	2,48	0,00	9,00
Abstraktheit					
Variationsbreite	2,06	2,00	0,77	1,00	3,00
Gleichförmigkeit	0,77	0,78	0,18	0,50	1,00
Wechselhaftigkeit	3,06	2,00	2,98	0,00	10,00
<i>Merkmale der Bearbeitungsanforderungen</i>					
Kognitiver Prozess					
Variationsbreite	3,50	4,00	0,97	2,00	5,00
Gleichförmigkeit	0,51	0,50	0,15	0,33	0,82
Wechselhaftigkeit	4,88	4,00	2,90	1,00	12,00
Wissensdimension					
Variationsbreite	2,25	2,00	0,58	1,00	3,00
Gleichförmigkeit	0,66	0,63	0,15	0,45	1,00
Wechselhaftigkeit	3,19	2,50	2,37	0,00	8,00
Format d. Produkts					
Variationsbreite	2,50	3,00	0,73	1,00	3,00
Gleichförmigkeit	0,67	0,60	0,18	0,50	1,00
Wechselhaftigkeit	3,81	3,00	3,49	0,00	13,00
Art d. Produkts					
Variationsbreite	2,31	2,00	0,70	1,00	4,00
Gleichförmigkeit	0,78	0,82	0,15	0,43	1,00
Wechselhaftigkeit	3,06	2,50	1,91	0,00	7,00
Reflexions-/Auswertungsfunktion					
Variationsbreite	1,94	2,00	0,68	1,00	3,00
Gleichförmigkeit	0,85	0,89	0,16	0,50	1,00
Wechselhaftigkeit	2,13	2,00	1,86	0,00	6,00

Beschreibung der Ausprägungen der Variabilitätsindikatoren: Bei Betrachtung der Ergebnisse zeigen die Werte des Indikators „Variationsbreite“ eine durchschnittliche Variationsbreite von zwei bis vier unterschiedlichen Ausprägungen des jeweiligen Merkmals pro Unterrichtsstunde. Dies ist je nach Spannweite des Wertebereichs unterschiedlich zu bewerten, wie z.B. die durchschnittliche Variationsbreite des Merkmals „Format des geforderten Produkts“ in Höhe von 2,5 bei 3 möglichen Ausprägungsstufen gegenüber der mittleren Variationsbreite des Merkmals „Medien:

Darstellungsform“ in Höhe von ebenfalls 2,5 bei 7 möglichen Ausprägungsstufen zeigt. Im Übrigen vermag der Indikator in zwei Einzelfällen nicht ausreichend zwischen den Unterrichtsstunden zu differenzieren, was an einer sehr geringen bzw. nicht vorhandenen Streuung der Indikatorenwerte zu erkennen ist: Wie der Minimal- und Maximalwert des Merkmals „verbaler Umfang der Aufgabenpräsentation“ zeigen, werden in den untersuchten Unterrichtsstunden entweder drei oder vier unterschiedliche Ausprägungsstufen dieses Merkmals abgedeckt; bei den „Hilfestellungsmaßnahmen bei Aufgabenstellung“ ist überdies in allen Unterrichtsstunden dieselbe Variationsbreite festzustellen.

Die Werte des Indikators „Gleichförmigkeit“ lassen erkennen, dass die Aufgabenmerkmale in unterschiedlichem Ausmaß durch eine Monotonie der Ausprägungen gekennzeichnet sind. Etwa die Sprache der Aufgabenpräsentation (0,80), die Abstraktheit der Aufgabeninhalte (0,77) oder die Art des geforderten Produkts (0,78) werden in einer Unterrichtsstunde tendenziell eher durch eine bestimmte Ausprägung dominiert, was in Anbetracht der Häufigkeitsverteilungen dieser Merkmale aber nicht überrascht (vgl. Kap. 7.3.2, Tab. 15). Demgegenüber stellt sich die Verwendung von Mediengeräten/-materialien, die zur Aufgabenbearbeitung erforderlichen kognitiven Prozesse und die Anzahl an Hilfestellungsmaßnahmen in den Aufgabenstellungen deutlich weniger gleichförmig dar, indem bei diesen Merkmalen durchschnittlich die Hälfte der Aufgaben in einer Unterrichtsstunde auf dieselbe Ausprägung entfallen. An einigen Stellen markieren Maximalwerte in Höhe von 1,00⁴⁶ aber auch, dass das betreffende Aufgabenmerkmal in mindestens einer Unterrichtsstunde durchgängig dieselbe Ausprägung hatte.

Der Indikator „Wechselhaftigkeit“ zeigt, dass bei den meisten Aufgabenmerkmalen drei bis vier Ausprägungswechsel pro Unterrichtsstunde erfolgt sind. Nur der verbale Umfang der Aufgabenpräsentation (6,38), die Anzahl der Hilfestellungsmaßnahmen (6,56) und die zur Aufgabenbearbeitung erforderlichen kognitiven Prozesse (4,88) wurden im Durchschnitt häufiger variiert; die Reflexions- und Auswertungsfunktion der Aufgaben hingegen etwas seltener (2,13) – angesichts der Häufigkeitsverteilungen zu diesem Merkmal erwartungsgemäß (vgl. Kap. 7.3.2, Tab. 15). Zugleich ist an den Streuungen der Indikatorenwerte erkennbar, dass die Wechselfrequenz der

⁴⁶ s. dazu auch Minimalwerte des Indikators Variationsbreite #verständlicher formulieren

Aufgabenmerkmalsausprägungen in den Unterrichtsstunden sehr unterschiedlich und mitunter groß ausfällt (s. Standardabweichungen und Maximalwerte).

Wechselseitige Zusammenhänge zwischen den Variabilitätsindikatoren: Um einen Einblick in die Beziehungen zwischen den Indikatoren der stundenspezifischen Aufgabenvariabilität zu erhalten und deren Verhältnis zu anderen Kennzeichen des Stundenverlaufs (mittlerer Zeitumfang der Aufgaben und Aufgabenanzahl) aufzuzeigen, wurden die Werte der Variabilitätsindikatoren aller Aufgabenmerkmale miteinander korreliert. Die Korrelationsberechnungen wurden wegen der kleinen Stichprobe ($n = 16$ Unterrichtsstunden) und der teilweise starken Abweichungen von der Normalverteilung auf Basis des Rangkorrelationskoeffizienten Spearman r_s durchgeführt. Die Ergebnisse (s. Anhang B) lassen für die drei verschiedenen Variabilitätsindikatoren innerhalb der jeweiligen Aufgabenmerkmale mittlere bis große Zusammenhänge erkennen ($r_s = 0,4 - 0,8$), wobei die Korrelationsrichtung des Indikators „Gleichförmigkeit“ gegenüber den beiden anderen meist negativ ist. Die Spannweite dieser Zusammenhänge zeigt, dass die Variabilitätsindikatoren je nach Aufgabenmerkmal zueinander unterschiedlich trennscharf sind. Auffällig ist zudem, dass der Indikator „Wechselhaftigkeit“ besonders viele hohe Korrelationen aufweist, und zwar sowohl innerhalb der Aufgabenmerkmale über die Variabilitätsindikatoren hinweg als auch zwischen den verschiedenen Aufgabenmerkmalen. Dieses Muster kommt möglicherweise dadurch zustande, dass viele Ausprägungswechsel von Aufgabenmerkmalen die Wahrscheinlichkeit einer hohen Variationsbreite und einer geringen Gleichförmigkeit erhöhen. Die Zusammenhänge der Variabilitätsindikatoren zu den Aspekten des Zeitumfangs der Aufgaben (im Durchschnitt pro Unterrichtsstunde) und der stundenspezifischen Aufgabenanzahl folgen wie zu erwarten in der Gesamttendenz diesem Muster: Je länger die Aufgaben einer Unterrichtsstunde, desto geringer die Aufgabenanzahl und desto geringer die Variabilität der Aufgabenmerkmale.

Neben den Beziehungen zwischen den Variabilitätsindikatoren zeigen sich in den Werten auch einige nennenswerte Zusammenhänge zwischen Aufgabenmerkmalen, die bereits auf Ebene der Einzelaufgabe erkennbar waren (vgl. Kap. 7.3.2), wie z.B. zwischen Aspekten des Medieneinsatzes oder zwischen dem Lebensweltbezug und der Abstraktheit der Inhalte. Allerdings werden auf dieser Ebene auch neue

Zusammenhänge deutlich, wie z.B. jener zwischen der Variabilität von Medien und inhaltlichem Lebensweltbezug: Eine höhere Variationsbreite und Wechselhaftigkeit des Lebensweltbezugs der Aufgabeninhalte in einer Stunde geht einher mit höherer Variationsbreite und Wechselhaftigkeit der Mediendarstellungsformen ($r_s = 0,46 - 0,51$). Durchweg große Zusammenhänge sind zudem zu finden zwischen der zur Bearbeitung erforderlichen Wissensdimension und dem Format des geforderten Produkts, und zwar in allen drei Variabilitätsindikatoren ($r_s = 0,66 - 0,86^{47}$).

7.4 Beziehungen zwischen den Eigenschaften der Aufgaben und dem Beteiligungsverhalten der Schüler im Unterricht

Dieses Kapitel beinhaltet die Darstellung der Ergebnisse zur Kernfrage dieser Studie, ob Zusammenhänge zwischen dem Aufgabenangebot und dem Beteiligungsverhalten der Schüler im Unterricht festzustellen sind. Dazu werden zunächst Effekte auf der Ebene des Unterrichtsprozesses beleuchtet, indem Variationen der *on task*-Rate im Unterrichtsverlauf darauf hin geprüft werden, inwiefern sie durch die Ausprägungen der Merkmale der einzelnen Aufgaben beeinflusst werden. Im darauf folgenden Schritt werden Einflüsse der stundenspezifischen Variabilität des Aufgabenangebots auf das allgemeine *on task*-Niveau in der jeweiligen Unterrichtsstunde betrachtet. In beiden Auswertungsschritten werden die Ergebnisse differenziert dargestellt sowohl für die gesamte Schülerstichprobe als auch die mithilfe der verfügbaren Personenmerkmale unterschiedenen Untergruppen (vgl. Kap. 6.5 und 7.1).

7.4.1 Merkmale der Einzelaufgabe und gleichzeitige *on task*-Rate

In diesem Abschnitt wird der Frage nachgegangen, ob Ausprägungen (oder Ausprägungskombinationen) der Aufgabenmerkmale mit Unterschieden der gleichzeitigen, durchschnittlichen *on task*-Rate der Klasse (oder Schüleruntergruppen innerhalb der Klasse) zusammenhängen. Dazu ist es notwendig, dass zu den Aufgaben zeitgleiche Beobachtungsdaten zum Schülerbeteiligungsverhalten vorliegen, weshalb in diesem Auswertungsschritt nur Aufgaben innerhalb von Phasen des öffentlichen

⁴⁷ nur Betrag, ohne Vorzeichen

Aufgabenmerkmale auf die zeitgleiche *on task*-Rate wurden in Abhängigkeit vom Messniveau des Merkmals mit unterschiedlichen statistischen Verfahrensweisen geprüft.

Zusammenhänge zwischen Zeitmerkmalen der Aufgaben und gleichzeitiger on task-Rate: Für die Eigenschaften des zeitlichen Umfangs der Aufgaben wurden Korrelationen mit der durchschnittlichen *on task*-Rate während des Aufgabenzeitraums berechnet. Da die zu korrelierenden Variablen nicht normalverteilt sind (Kolmogorov-Smirnov-Tests bei allen Zeitmerkmalen der Aufgaben: $p < 0,001$; bei den *on task*-Raten aller Untergruppen: $p < 0,01$) und schiefe Verteilungen aufweisen, wurde dafür die Rangkorrelationen nach Spearman genutzt (vgl. Bühner & Ziegler, 2009; Leonhart, 2004). Nimmt man in den Ergebnissen zuerst die allgemeinen, d.h. für die gesamte Schülergruppe gültigen Effekte in den Blick (s. Tab. 21, Spalte „Alle Schüler“), zeigen die Korrelationswerte, dass eine längere Dauer der Aufgabebearbeitung leicht mit niedrigeren *on task*-Raten zusammenhängt ($r_s = -0,20$).

Tab. 21: Zusammenhänge (Rangkorrelation) zwischen Merkmalen des Zeitumfangs der Aufgaben und der gleichzeitigen *on task*-Rate (aller Schüler und unterschiedlicher Teilgruppen der Schülerstichprobe)

Aufgabenmerkmal	on task-Rate							
	Alle Schüler	Untergruppen						
		Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich		
		D	nicht D	m	w	niedrig	mittel	hoch
Zeitumfang Aufgabenstellung	0,06	0,04	0,04	-0,01	0,03	-0,03	0,02	-0,04
Zeitumfang Aufgabebearbeitung	-0,20*	-0,18*	-0,16	-0,17	-0,24**	-0,21*	-0,15	-0,23**
Zeitumfang gesamte Aufgabe	-0,19*	-0,15	-0,16	-0,16	-0,23**	-0,22*	-0,13	-0,23*

Koeffizient: Spearman r_s ; Signifikanzkennzeichnung (zweiseitig): * $p \leq 0,05$, ** $p \leq 0,01$, *** $p \leq 0,001$, keine: $p > 0,05$; Umgangssprache in der Familie dichotomisiert: „Deutsch“ vs. „andere Sprache/Deutsch+andere Sprache“

Durch den hohen quantitativen Anteil der Bearbeitungsdauer an der Aufgabengesamtdauer (vgl. Kap. 7.3.2, Tab. 16) ist dieser Zusammenhang in leicht schwächerer Form auch bei der Aufgabengesamtdauer festzustellen ($r_s = -0,19$). Bei Aufgaben in Einzel-, Partner- oder Gruppenarbeitsphasen könnte man diesen

Zusammenhang damit erklären, dass die Schüler mit fortschreitender Dauer nach und nach den Auftrag abschließen und die Lehrkraft in Folge davon die Aufgabe irgendwann beendet. Da es sich hierbei allerdings um Aufgaben im Rahmen öffentlichen Unterrichts handelt, ist es denkbar, dass mit fortschreitender Dauer immer mehr Schüler die Aufgabe „geistig abschließen“, nicht weiter aktiv an deren Bearbeitung partizipieren und sie auch nicht weiter verfolgen, sodass sie sich in der Folge nach außen hin sichtbar von der Aufgabe ab- und anderen Dingen zuwenden. In den unterschiedlichen Teilgruppen setzt sich dieser Zusammenhang bei leichten Unterschieden in der Stärke des Zusammenhangs fort. Für die Dauer der Aufgabenstellung ist demgegenüber kein nennenswerter Zusammenhang zu erkennen. Zwischen den Korrelationswerten der unterschiedlichen Teilgruppen zeigen sich überdies keine Unterschiede, die auf differenzielle Effekte des zeitlichen Aufgabenumfangs auf das Beteiligungsverhalten unterschiedlicher Schülergruppen hindeuten.

Einflüsse kategorialer Aufgabenmerkmale auf die gleichzeitige on task-Rate: Als statistisches Verfahren auf Seiten der nominalskalierten Aufgabenmerkmale wurde für diesen Auswertungsschritt die einfaktorielle, univariate Varianzanalyse genutzt, mit dem jeweiligen Aufgabenmerkmal als mehrstufigem Faktor und der durchschnittlichen on task-Rate im jeweiligen Aufgabenzeitraum als abhängige Variable⁴⁸. Die Logik dieses Schrittes ist somit die eines Vergleichs von Gruppenmittelwerten in den on task-Raten, wobei die Ausprägungsstufen eines Aufgabenmerkmals die Gruppen definieren. Im Falle der beiden Aufgabenmerkmale „Medien: Darstellungsform“ und „Medien: Gerät/Material“ wurden dafür aufgrund fehlender oder zu geringer Auftretenshäufigkeiten einzelner Ausprägungsstufen Umkodierungen vorgenommen⁴⁹. Auch diese Prüfungen wurden sowohl für die gesamte Schülerstichprobe als auch für die verschiedenen Untergruppen durchgeführt. Wegen der Vielzahl und Komplexität der Einzelergebnisse werden die entsprechenden Befunde anhand einer Übersicht der

⁴⁸ Wie bereits in Kap. 7.1.3 und 7.2.3 wird an dieser Stelle trotz Verletzung der Normalverteilungsvoraussetzung der abhängigen Variable in Anschluss an Empfehlungen von Bühner und Ziegler (2009) und Rasch, Friese, Hofmann und Naumann (2006) angesichts der Stichprobengröße ($n = 137$ Aufgaben) die Varianzanalyse den non-parametrischen Testvarianten vorgezogen.

⁴⁹ Medien-Darstellungsform: Zusammenfassung der Kategorie „Tabelle/Diagramm“ mit der Kategorie „Abbildung/Grafik“ und Löschung der Kategorien „Gegenstand“ und „Andere“; Gerät/Material: Zusammenfassung der Kategorie „Tafel“ mit der Kategorie „Projektor“ und Löschung der Kategorie „Kombination“. Umkodierungen unter Vorabprüfung bedeutsamer Mittelwertunterschiede in den on task-Raten zwischen den zusammengefassten Kategorien.

signifikanten Mittelwertvergleiche zusammenfassend wiedergegeben (Tab. 22)⁵⁰. Eine Darstellung der Gruppenmittelwerte und Kennwerte im Einzelnen ist zu finden in Anhang C.

Der Überblick zeigt, dass die Prüfungen für etwa die Hälfte der Merkmale ohne signifikante Ergebnisse ausfallen. Zwischen den Ausprägungen dieser Aufgabenmerkmale sind somit keine nennenswerten Unterschiede hinsichtlich des durchschnittlichen gleichzeitigen *on task*-Niveaus festzustellen. Bei einer Reihe anderer Aufgabeneigenschaften hingegen sind Einflüsse auf die *on task*-Rate zu erkennen, wobei nicht alle der gefundenen Effekte für die gesamte Schülergruppe gelten, sondern in Teilen nur für Untergruppen, und die Ergebnisse zudem teilweise durch ungleichmäßige Häufigkeitsverteilungen der Aufgabenmerkmalsausprägungen beeinflusst sein können.

So zeigt sich bei der Sprache der Aufgabeneinführung bzw. -präsentation unter der Ausprägungsbedingung „Kombination von Sprachen“ eine signifikant höhere *on task*-Rate ($M = 93,70\%$) als bei ausschließlich auf Englisch eingeführten Aufgaben ($M = 86,04\%$). Dieser auf Ebene aller Schüler feststellbare Effekt setzt sich aber nicht bei allen Untergruppen fort, sondern nur bei Schülern mit deutschsprachigem Familienhintergrund ($M_{\text{Englisch}} = 83,40\%$; $M_{\text{Kombination}} = 96,82\%$) und auch Schülern mit hohem Leistungsstand ($M_{\text{Englisch}} = 87,01\%$; $M_{\text{Kombination}} = 96,59\%$). Das Ergebnis lässt sich nur schwierig interpretieren, da besonders bei den beiden genannten Teilgruppen zu erwarten wäre, dass eher alleine auf Deutsch oder Englisch formulierte Aufgabenstellungen mit höheren *on task*-Raten verbunden sind, nicht aber eine Mischung aus beiden. Bei diesem Ergebnis ist jedoch auch zu beachten, dass die Ausprägung „Kombination“ nur einen geringen Häufigkeitsanteil an den Ausprägungen dieses Aufgabenmerkmals bildet (Englisch: 89, Deutsch: 38, Kombination: 9), sodass fraglich erscheint, ob dieser Effekt darauf basierenden Verzerrungen unterliegt.

Etwas deutlichere Unterschiede ergeben sich hingegen im Bereich des Medieneinsatzes. Bei Betrachtung der Mediengeräte bzw. -materialien zeigt sich bei Verwendung des Lehrbuchs allgemein eine niedrigere *on task*-Rate als bei den anderen Kategorien ($M_{\text{Lehrbuch}} = 84,07\%$; $M_{\text{Tafel/Projektorbild}} = 88,12\%$; $M_{\text{Arbeitsblatt}} = 88,44\%$; $M_{\text{andere}} = 92,62\%$), wobei die Mittelwertdifferenz zur Kategorie „andere“ als signifikant markiert wird.

⁵⁰ Für Erläuterungen zur mit dieser auswertungsmethodischen Vorgehensweise verbundenen Problematik der Alpha-Fehlerkumulierung siehe Kap. 8.2.

Tab. 22: Übersicht über signifikante Mittelwertunterschiede zwischen Ausprägungsstufen der Aufgabenmerkmale in der gleichzeitigen on task-Rate (aller Schüler und unterschiedlicher Teilgruppen der Schülerstichprobe)

Aufgabenmerkmal	on task-Rate							
	Alle Schüler	Untergruppen						
		Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich		
		D	nicht D	m	w	niedrig	mittel	hoch
<i>Merkmale der Aufgabeneinführung/-präsentation</i>								
Sprache	o	o	-	-	-	-	-	o
Verbaler Umfang	-	-	-	-	-	-	-	-
Anzahl Hilfestellungsmaßnahmen	-	-	-	-	-	-	-	-
Medium: Darstellungsform	-	-	o	-	-	-	-	-
Medium: Gerät/Material	o	-	X	o	o	-	o	-
Umfang Medieninhalte	-	-	-	-	-	-	-	-
Bekanntheit Medium	-	-	-	-	o	-	X	-
<i>Merkmale der Aufgabeninhalte</i>								
Lebensweltbezug	-	-	o	-	o	-	o	-
Abstraktheit	-	-	-	-	-	-	o	-
<i>Merkmale der Bearbeitungsanforderungen</i>								
Kognitiver Prozess	-	-	-	-	-	-	-	-
Wissensdimension	-	-	-	-	-	-	-	-
Format d. Produkts	-	-	-	-	-	-	-	-
Art d. Produkts	o	o	o	o	-	o	o	o
Reflexions-/Auswertungsfunktion	X	-	X	-	X	-	X	-

X = signifikante Varianzanalyse (Signifikanzniveau 5 %, bei Verletzung der Varianzhomogenitätsvoraussetzung gemäß Levene-Test herabgesetzt auf 1 %, s. Bühner & Ziegler, 2009, S. 369f.); **o** = gemäß Post-Hoc-Tests mindestens ein signifikanter Mittelwertunterschied zwischen einzelnen Ausprägungsstufen (Testverfahren bei Varianzheterogenität: Games Howell; bei Varianzhomogenität: Bonferroni); - = keine signifikanten Mittelwertunterschiede

Diese Ausprägung umfasst neben Anschauungsgegenständen zum Beispiel auch Audio-Abspielgeräte, was auch die leichte Tendenz zur höheren *on task*-Rate bei der Verwendung von Hörbeispielen im Merkmal „Medien: Darstellungsform“ erklärt, die allerdings nur bei der Untergruppe der Schüler mit nicht-deutschsprachigem Hintergrund signifikant wird ($M_{\text{Text}} = 88,06\%$; $M_{\text{Abbildungen}} = 86,43\%$; $M_{\text{Hörbeispiel}} = 91,08\%$; $M_{\text{Kombination}} = 81,03\%$). Die für die Mediengeräte und -materialien gefundene Mittelwertdifferenz zieht sich hingegen durch die Mehrzahl der Teilgruppen der Schülerstichprobe, bis auf die Schüler mit deutschsprachigem Hintergrund und die im Klassenvergleich leistungsstarken und leistungsschwachen Schüler, wobei der Effekt besonders deutlich bei den Schülern mit nicht-deutschsprachigem Hintergrund hervortritt ($F_{(4;129)} = 2,96$; $p < 0,05$; $\eta^2 = 0,08$). Dieser differenzielle Effekt ist zunächst schwierig zu erklären, könnte aber als Hinweis genommen werden für weitergehende Untersuchungen, in denen – etwa aus Perspektive der Fachdidaktik – die Inhalte der Lehrbücher und der „anderen“ Darstellungen genauer verglichen und zum Sprachhintergrund der Schüler in Beziehung gesetzt werden. Im Hinblick auf den Aspekt der Bekanntheit des verwendeten Mediums ist als nächstes festzustellen, dass Aufgaben mit einer bereits zuvor eingeführten Mediengrundlage mit einer etwas geringeren durchschnittlichen *on task*-Rate einhergehen als solche mit simultan eingeführtem oder auch angekündigtem Medium; allerdings zeigt sich dieser Effekt alleine bei den Schülern im Leistungsmittelfeld ($F_{(3;131)} = 4,87$; $p < 0,01$; $\eta^2 = 0,10$) und annähernd auch bei den Mädchen, sodass er allenfalls als leichte Tendenz gewertet werden kann. Zudem ist die vergleichsweise geringe Häufigkeit einer Ausprägungsstufe („Medium angekündigt“ = 7) zu beachten. Insgesamt können die Ergebnisse zum Medieneinsatz mit Blick auf das QAIT-Modell effektiven Unterrichts von Slavin (1994) als Effekte einer Interessantheit der Darstellungen durch die Lehrkraft gelesen werden, die zum Anreizwert des Unterrichts und damit der Förderung der individuellen Beteiligungsbereitschaft der Schüler beitragen. Im Bereich der Aspekte, die die Einführung und Präsentation der Aufgabe betreffen, sind darüber hinaus allerdings keine weiteren Effekte festzustellen.

Im Bereich der Aufgabeninhalte liegen kaum nennenswerte, bzw. im Falle des Lebensweltbezugs der Aufgabeninhalte unklare Ergebnisse vor. Eben jenes Merkmal lässt nur für einzelne Untergruppen der Schülerstichprobe Unterschiede der *on task*-Raten in Abhängigkeit von dessen Ausprägungen erkennen: Bei Schülern mit nicht-deutschsprachigem Hintergrund, Mädchen und Schülern im Leistungsmittelfeld fällt die

on task-Rate bei Aufgaben ohne Lebensweltbezug signifikant höher aus als bei Aufgaben mit konstruiertem Lebensweltbezug – ein Ergebnismuster, welches sich in dieser Kombination nur schwierig interpretieren lässt. Während nach Slavin eben jene Herstellung von Bezügen der Unterrichtsinhalte zur Alltagswelt der Schüler allgemein positiv auf die Schülerbeteiligung wirken sollte, da sie die subjektiv wahrgenommene Sinnhaftigkeit der Unterrichtsinhalte steigern (ebd.), kann an dieser Stelle nicht von einem solchen Effekt gesprochen werden. Forschungsergebnisse, die deutliche Einflüsse des Lebensweltbezugs der Inhalte auf das *student engagement* belegen (z.B. Lam, Wong, Yang & Liu, 2012; Marks, 2000; im Überblick Fredricks, Blumenfeld & Paris, 2004; vgl. auch Kap. 2.3.2), finden an dieser Stelle auch keine Bestätigung.

Hinsichtlich der Aufgabenmerkmale, die die Anforderungen der Aufgabenbearbeitung beschreiben, ist lediglich für die Art des geforderten Produkts ein Einfluss auf das Schülerbeteiligungsverhalten erkennbar. Demnach zeigt sich in der gesamten Stichprobe eine deutlich höhere *on task*-Rate, wenn von den Schülern „andere Produkte“ (z.B. der Vollzug von Bewegungen wie Handzeichen oder Bewegungen im Raum) zu erbringen sind, als im Falle der übrigen Kategorien ($M_{\text{Sprache}} = 86,33\%$; $M_{\text{Schrift}} = 81,41\%$; $M_{\text{Kombination}} = 90,97\%$; $M_{\text{andere}} = 97,31\%$). Der Effekt ist in allen Teilgruppen zu finden, allerdings muss dieses Ergebnis vor dem Hintergrund möglicher Verzerrungen durch eine stark ungleichmäßige Häufigkeitsverteilung der Ausprägungsstufen gesehen werden (Sprache: 119, Schrift: 4, Kombination: 2; andere Produkte: 6).

Das Aufgabenmerkmal, welches die stärksten Effekten auf die gleichzeitige *on task*-Rate zeigt, zugleich aber ebenfalls von Verzerrungen der zugrunde liegenden Häufigkeitsverteilungen betroffen ist, ist die Reflexions- bzw. Auswertungsfunktion in Bezug auf zurückliegende Aufgaben. Demgemäß ist bei der Besprechung von Hausaufgaben eine um mehr als 10 % niedrigere *on task*-Rate festzustellen als bei neuen, für sich stehenden Aufgaben oder auch der Besprechung von bereits zurückliegenden Unterrichtsaufgaben ($M_{\text{keine Refl./Ausw.}} = 87,74\%$; $M_{\text{Refl./Aus. Unterrichtsaufgabe}} = 88,94\%$; $M_{\text{Refl./Ausw. Hausaufgabe.}} = 76,39\%$; $F_{(2;134)} = 4,99$; $p < 0,01$; $\eta^2 = 0,07$). Bei Betrachtung der Untergruppen der Schülerstichprobe differenziert sich dieser allgemeine Effekt scheinbar aus, indem die Mittelwertvergleiche nur noch für Schüler mit nicht-deutschsprachigem Hintergrund, Mädchen und Schüler auf mittlerem Leistungsniveau signifikant ausfallen; die allgemeine Tendenz des Ergebnisses bleibt jedoch in allen Gruppen erkennbar. Als Einschränkung dieses Befunds muss wie bereits

angedeutet berücksichtigt werden, dass von den insgesamt neun beobachteten Aufgaben mit Auswertungsfunktion in Bezug auf eine Hausaufgabe fünf in jener Unterrichtsstunde festzustellen waren, die die niedrigste *on task*-Rate aller Unterrichtsstunden aufweist (E 1003, vgl. Kap. 7.2.2, Tab. 10). Aufgrund der damit verbundenen Gefahr möglicher Verzerrungen sollte dieses Ergebnis nicht überinterpretiert werden.

Zusammenfassend bleibt an diesem Punkt festzuhalten, dass nur bei einigen der untersuchten Aufgabenmerkmalen Einflüsse auf das gleichzeitige Beteiligungsverhalten der Schüler erkennbar sind, wobei nicht alle diese Effekte gleichermaßen deutlich hervortreten oder in gleicher Weise für alle Teilgruppen der Schüler gelten. Zudem ist die Interpretierbarkeit einiger Effekte eingeschränkt durch verteilungsbedingte Verzerrungen. In der Zusammenschau scheint der relevanteste und am deutlichsten erkennbare Befund im Einfluss des verwendeten Mediums zu bestehen. Es ist jedoch in Betracht zu ziehen, dass die Perspektive auf isolierte Wirkungen der einzelnen Aufgabenmerkmale Effekte außer Acht lässt, die durch das Zusammenspiel von Aufgabeneigenschaften zustande kommen.

Interaktionseffekte zwischen Aufgabenmerkmalen auf die gleichzeitige on task-Rate: Um mögliche Wechselwirkungen verschiedener Aufgabeneigenschaften auf das Beteiligungsverhalten der Schüler zu untersuchen, wurden daher zusätzlich mehrfaktorielle Varianzanalysen durchgeführt, in denen die Aufgabenmerkmale als unabhängige Variablen kombiniert und deren gemeinsamer bzw. wechselseitiger Einfluss auf die abhängige Variable, die zeitgleiche *on task*-Rate geprüft wurde. Folglich sind an diesem Punkt die so genannten Interaktionseffekte von besonderem Interesse, da diese anzeigen, inwiefern die Einflüsse der unabhängigen Variablen nicht additiv bzw. bildlich gesprochen „nebeneinander her“ wirken, sondern unter bestimmten Ausprägungskombinationen Wechselwirkungen entfalten (vgl. Bortz, 2010; Bortz & Döring, 2006). Während bei dieser Verfahrensweise prinzipiell eine unendliche Anzahl miteinander kombinierter unabhängiger Variablen möglich ist, wird aus Gründen der plausiblen Interpretierbarkeit der Interaktionseffekte und mit Blick auf die Qualität der statistischen Prüfung im Verhältnis zum Informationsgewinn durch die Hinzunahme von Faktoren empfohlen, sich auf zweifaktorielle Modelle zu beschränken (ebd.). Die

Daten der vorliegenden Stichprobe würden zudem keine drei- oder mehrfaktoriellen Modelle erlauben, weil aufgrund der Häufigkeitsverteilungen in vielen Fällen nicht die notwendige Voraussetzung der vollständigen Kombinierbarkeit aller Faktorstufen gegeben wäre. Daher wurde in diesem Schritt eine Serie zweifaktorieller, univariater Varianzanalysen durchgeführt, in der alle Aufgabenmerkmale in getrennten Paarungen miteinander kombiniert und in ihrem Einfluss auf die zeitgleiche *on task*-Rate geprüft wurden. Aufgrund der daraus resultierenden hohen Komplexität an Tests wurde an dieser Stelle keine Differenzierung nach Untergruppen der Schülerstichprobe vorgenommen. Das bedeutet, im Vergleich zu den zuvor berichteten Ergebnissen dieses Abschnitts (Tab. 21 & 22) berücksichtigen die Prüfungen an dieser Stelle nur allgemeine, d.h. für die gesamte Schülerstichprobe feststellbare Effekte. Um mögliche Wechselwirkungen zwischen qualitativen und quantitativen Eigenschaften der Aufgaben berücksichtigen zu können (vgl. Kap. 3, Kap. 6.7.2), wurden in diesen Auswertungsschritt auch die Aspekte des Zeitumfangs der Aufgaben einbezogen. Dazu wurden die intervallskalierten Zeitmerkmale mittels Perzentilgruppenbildung in dreistufige, kategoriale Variablen transformiert.

Auf eine umfassende Darstellung der Ergebnisse kann aus zwei Gründen verzichtet werden: Zum einen replizieren die auf die Wirkungen der Einzelmerkmale zurückgehenden Haupteffekte die bereits zuvor berichteten Ergebnisse, zum anderen ist insgesamt nur ein signifikanter Interaktionseffekt zu verzeichnen⁵¹: Zwischen den beiden Aufgabenmerkmalen „Abstraktheit der Aufgabeninhalte“ und „Format des geforderten Produkts“ besteht eine hoch signifikante Wechselwirkung von mittlerer bis hoher Stärke ($F_{(4;109)} = 4,17$; $p < 0,01$; $\eta^2 = 0,13$). Demnach fällt bei Aufgaben mit konkreten Inhalten (z.B. beobachtbare Ereignisse des Alltags) die Höhe der *on task*-Rate unter der Bedingung „Erzeugung eines Produkts mittels Verarbeitung von Vorgaben“ hoch (90,04 %), unter der Bedingung „Eigenkonstruktion eines Produkts mit eingeschränktem Umfang“ hingegen niedrig (77,29 %) aus. Sind den Schülern also Begriffe oder Abbildungen zu greifbaren Aspekten des alltäglichen Lebens vorgegeben, ist die Beteiligung allgemein höher, als wenn sie eigenständig eine kurze Antwort zu solchen Aspekten formulieren müssen. Eine Beeinträchtigung von Haupteffekten

⁵¹ Bei den beiden Merkmalen „Format“ und „Art des geforderten Produkts“ war eine Testung der Wechselwirkung nicht möglich, weil sich aufgrund der Zellenbesetzungen in der Kreuztabellierung zu wenige der Ausprägungsstufen ausreichend miteinander kombinieren ließen.

besteht durch diesen Befund nicht, da die beiden Merkmale für sich genommen keine bedeutsamen Einflüsse auf die gleichzeitige *on task*-Rate zeigen (s. Tab. 22).

Die Prüfung möglicher Wirkungen von Ausprägungskombinationen der Aufgabenmerkmale auf das Beteiligungsverhalten der Schüler liefert somit keine weiterführenden Erkenntnisse. Selbst unter Einbezug der Merkmale des Zeitumfangs konnten in diesem Schritt keine bedeutsamen Interaktionseffekte verzeichnet werden. Damit ist für diesen Abschnitt insgesamt festzuhalten, dass bei Betrachtung der Einzelaufgabe – sozusagen auf der Mikroebene des Unterrichts – kaum Einflüsse der unterschiedlichen Aufgabeneigenschaften auf das Beteiligungsverhalten der Schüler erkennbar sind. Im nächsten Schritt wird daher der Frage nachgegangen, ob die Einflüsse von Aufgaben weniger in der „Sofortwirkung“ zu suchen sind als im summativen Zusammenwirken der stundenspezifischen Aufgabenabfolge.

7.4.2 Stundenspezifische Variabilität der Aufgaben und durchschnittliche *on task*-Rate der Unterrichtsstunde

Ziel dieses Auswertungsschrittes ist es, das Aufgabenangebot einer Unterrichtsstunde als Ganzheit zu betrachten und auf dieser Basis Zusammenhänge zu den stundenspezifischen *on task*-Raten zu prüfen. Zur Beschreibung des Aufgabengefüges auf Stundenebene werden die unter Kap. 7.3.3 beschriebenen Indikatoren herangezogen, welche auf unterschiedliche Weise die Variabilität der Aufgabenmerkmale über den Zeitverlauf der Unterrichtsstunde hinweg abbilden. Zusätzlich werden zur Berücksichtigung von Zeiteffekten die Komponenten des zeitlichen Umfangs der Aufgaben (im Stundendurchschnitt) und als Kennzeichen der Grundstruktur des stundenspezifischen Aufgabenangebots die Anzahl an Aufgaben berücksichtigt.

Als statistisches Verfahren der Zusammenhangsprüfung wurde aufgrund von Abweichungen der zu korrelierenden Variablen von der Normalverteilung die Rangkorrelation nach Spearman (Koeffizient r_s) genutzt (vgl. Bühner & Ziegler, 2009; Leonhart, 2004). Bei Aufgabenmerkmalen mit einer zu hohen Anzahl an Verbundrängen (> 20 % der Rangplätze, s. Bortz, 2010, S. 179) wurde die von Bortz (ebd.) und Leonhart (2004, S. 208) vorgeschlagene, modifizierte Berechnung von r_s angewendet. Wegen der

Stichprobengröße ($n = 16$ Unterrichtsstunden) wird nachfolgend unter Auslassung der Ergebnisse der Signifikanzprüfungen nur die Höhe der Koeffizienten berichtet und interpretiert (vgl. Bühner & Ziegler, 2009).

Tab. 23 zeigt die Zusammenhänge zwischen den Indikatoren der Variabilität der Aufgabenmerkmale auf Stundenebene und der stundenspezifischen *on task*-Rate, letztere auch hier dargestellt für die gesamte Schülerstichprobe und die unterschiedlichen Teilgruppen. Dabei wird ersichtlich, dass sich die auf Ebene der Gesamtgruppe („Alle Schüler“) erkennbaren Zusammenhänge in den verschiedenen Untergruppen teilweise deutlich ausdifferenzieren, wie etwa im Falle des Merkmals „Sprache der Einführung und Präsentation der Aufgabe“ (Alle Schüler: $r_s = -0,20$; Umgangssprache in der Familie „D“: $r_s = 0,19$; Umgangssprache in der Familie „nicht D“: $r_s = -0,15$). Um solche differenziellen Effekte besser identifizieren zu können und in der Ergebnisbeschreibung berücksichtigen zu können, wurde zusätzlich ein Vergleich der Korrelationskoeffizienten innerhalb der unterschiedlichen Kategorien zur Unterscheidung der Schüler vorgenommen. Abgebildet werden die Differenzen mithilfe der Effektgröße q , welche auf Basis von Z-standardisierten Korrelationskoeffizienten die Beurteilung der Ausprägungen paarweiser Korrelationsdifferenzen erlaubt (Bühner & Ziegler, 2009, S. 609)⁵². Die auf diese Weise berechneten Korrelationsdifferenzen (s. Tab. 24) sollen bei der Beschreibung der Zusammenhänge zwischen den Indikatoren der Variabilität der Aufgabenmerkmale auf Stundenebene und der stundenspezifischen *on task*-Rate berücksichtigt werden.

⁵² Nach Rasch, Friese, Hofmann und Naumann (2010, S. 129) ist Fishers Z-Transformation von Korrelationskoeffizienten auch zulässig für die Rangkorrelation nach Spearman.

Tab. 23: Zusammenhänge (Rangkorrelation Spearman) zwischen Indikatoren stundenspezifischer Variabilität der verschiedenen Aufgabenmerkmale und der stundenweise durchschnittlichen on task-Rate (aller Schüler und unterschiedlicher Teilgruppen der Schülerstichprobe)

Variabilitäts- indikatoren der Aufgabenmerkmale	on task-Rate							
	Alle Schüler	Untergruppen						
		Umgangssprache i. d. Familie		Geschlecht		Leistungsniveau im Klassenvergleich		
		D	nicht D	m	w	niedrig	mittel	hoch
<i>Merkmale der Einführung und Präsentation der Aufgabe</i>								
<i>Sprache</i>								
Variationsbreite	-0,20	0,19	-0,15	-0,22	-0,02	-0,30	0,08	-0,10
Gleichförmigkeit	0,32	-0,24	0,28	0,33	0,10	0,37	0,14	0,06
Wechselhaftigkeit	-0,29	0,07	-0,23	-0,31	-0,07	-0,37	-0,15	0,04
<i>Verbaler Umfang</i>								
Variationsbreite	0,21	0,25	0,21	0,18	0,37	0,10	0,21	0,53
Gleichförmigkeit	-0,41	0,02	-0,35	-0,36	-0,29	-0,29	-0,39	-0,45
Wechselhaftigkeit	0,14	-0,11	0,21	0,00	0,27	0,03	0,30	0,22
<i>Anzahl Hilfestellungen</i>								
Variationsbreite ^A	-	-	-	-	-	-	-	-
Gleichförmigkeit	-0,13	-0,41	0,03	-0,05	-0,16	-0,07	-0,15	-0,18
Wechselhaftigkeit	-0,20	-0,30	-0,04	-0,25	-0,05	-0,19	-0,08	0,15
<i>Medien: Darstellungsform</i>								
Variationsbreite	0,15	0,50	0,18	0,08	0,31	0,07	0,20	0,37
Gleichförmigkeit	-0,11	-0,58	-0,16	-0,06	-0,36	-0,11	-0,10	-0,28
Wechselhaftigkeit	-0,09	0,18	-0,01	-0,22	0,09	-0,19	-0,05	0,30
<i>Medien: Gerät/Material</i>								
Variationsbreite	0,22	-0,04	0,31	0,20	0,32	0,23	0,13	0,27
Gleichförmigkeit	-0,27	-0,42	-0,37	-0,32	-0,39	-0,34	-0,22	-0,21
Wechselhaftigkeit	0,01	0,06	0,17	-0,02	0,18	-0,05	0,12	0,24
<i>Umfang Medieninhalte</i>								
Variationsbreite	-0,09	-0,21	-0,16	-0,13	-0,26	-0,08	0,00	-0,05
Gleichförmigkeit	-0,20	-0,34	-0,13	-0,22	-0,11	-0,16	-0,28	-0,18
Wechselhaftigkeit	-0,02	-0,16	0,02	0,03	-0,10	0,03	0,09	0,04
<i>Merkmale der Aufgabeninhalte</i>								
<i>Lebensweltbezug</i>								
Variationsbreite	0,20	0,20	0,17	0,00	0,29	0,15	0,25	0,23
Gleichförmigkeit	-0,03	-0,32	-0,07	0,10	-0,16	0,02	-0,13	-0,17
Wechselhaftigkeit	-0,15	0,04	-0,13	-0,26	-0,04	-0,18	-0,04	0,11
<i>Abstraktheit</i>								
Variationsbreite	-0,03	0,24	-0,13	-0,15	-0,13	-0,19	0,23	-0,11
Gleichförmigkeit	-0,26	-0,27	-0,22	-0,09	-0,20	-0,11	-0,54	0,12
Wechselhaftigkeit	0,27	0,08	0,20	0,10	0,15	0,08	0,56	-0,01
<i>Merkmale der Bearbeitungsanforderungen</i>								
<i>Kognitiver Prozess</i>								
Variationsbreite	0,24	0,19	0,21	0,15	0,19	0,21	0,41	0,13
Gleichförmigkeit	-0,18	-0,17	-0,07	-0,12	-0,03	-0,15	-0,20	-0,17
Wechselhaftigkeit	0,25	0,29	0,24	0,23	0,22	0,18	0,42	0,33

- Fortsetzung der Tabelle s. nächste Seite -

- Fortsetzung von Tab. 23 -

Variabilitäts- indikatoren der Aufgabenmerkmale	on task-Rate									
	Alle Schüler	Untergruppen						Leistungsniveau im Klassenvergleich		
		Umgangssprache i. d. Familie		Geschlecht				niedrig	mittel	hoch
		D	nicht D	m	w					
<i>Merkmale der Bearbeitungsanforderungen (Fortsetzung)</i>										
Wissensdimension										
Variationsbreite	0,24	0,17	0,15	0,19	0,27	0,08	0,20	0,45		
Gleichförmigkeit	-0,08	-0,33	0,04	-0,01	-0,09	0,09	-0,14	-0,29		
Wechselhaftigkeit	0,37	0,51	0,31	0,36	0,43	0,28	0,44	0,47		
Format d. Produkts										
Variationsbreite	0,27	0,35	0,26	0,37	0,19	0,18	0,38	0,37		
Gleichförmigkeit	-0,19	-0,53	-0,09	-0,20	-0,22	-0,04	-0,24	-0,33		
Wechselhaftigkeit	0,27	0,24	0,26	0,29	0,23	0,12	0,40	0,44		
Art d. Produkts										
Variationsbreite	-0,12	0,20	-0,13	-0,10	0,10	-0,14	-0,15	-0,01		
Gleichförmigkeit	-0,09	-0,66	-0,01	-0,16	-0,21	-0,13	0,02	-0,12		
Wechselhaftigkeit	0,04	0,59	0,02	0,06	0,22	-0,04	0,17	0,01		
<i>Reflexions-/Auswertungsfunktion</i>										
Variationsbreite	-0,05	-0,55	-0,02	-0,07	0,00	0,02	-0,13	0,02		
Gleichförmigkeit	0,07	0,25	0,11	0,02	0,09	-0,07	0,32	-0,16		
Wechselhaftigkeit	0,11	-0,17	0,15	0,07	0,20	0,13	0,05	0,31		
<i>Merkmale des Zeitumfangs der Aufgaben</i>										
Dauer gesamt	-0,24	-0,06	-0,27	-0,13	-0,22	-0,09	-0,46	-0,14		
Dauer Stellung	0,27	0,17	0,15	0,27	0,27	0,36	0,01	0,09		
Dauer Bearbeitung	-0,31	-0,21	-0,30	-0,19	-0,31	-0,14	-0,53	-0,12		
Aufgabenanzahl	0,13	-0,11	0,19	-0,04	0,18	-0,04	0,28	0,22		

n = 16 Unterrichtsstunden; ^A keine Variationsbreite vorhanden für diese Variable (s. Kap. 7.3.3., Tab. 19)

Tab. 24: Korrelationsdifferenzen (Betrag der Effektgröße q) zwischen Untergruppen der Schülerstichprobe, bezogen auf Zusammenhänge zwischen Indikatoren stundenspezifischer Variabilität der Aufgabenmerkmale und der stundenweise durchschnittlichen on task-Rate (aller Schüler und unterschiedlicher Teilgruppen der Schülerstichprobe)

Variabilitätsindikatoren der Aufgabenmerkmale	Korrelationsdifferenzen				
	Geschlecht (m/w)	Umgangssprache i. d. Familie (D/nicht D)	Leistungsniveau im Klassenvergleich		
			niedrig/mittel	mittel/hoch	niedrig/hoch
<i>Merkmale der Einführung und Präsentation der Aufgabe</i>					
Sprache					
Variationsbreite	0,20	0,34	0,39	0,19	0,20
Gleichförmigkeit	0,25	0,53	0,25	0,08	0,33
Wechselhaftigkeit	0,25	0,31	0,23	0,19	0,43
Verbaler Umfang					
Variationsbreite	0,21	-0,05	0,11	0,39	0,50
Gleichförmigkeit	0,07	-0,39	-0,11	-0,07	-0,19
Wechselhaftigkeit	0,27	0,32	0,28	-0,08	0,20
Anzahl Hilfestellungen					
Variationsbreite ^A	-	-	-	-	-
Gleichförmigkeit	-0,12	0,47	-0,08	-0,03	-0,11
Wechselhaftigkeit	0,20	0,27	0,11	0,23	0,34
Medien: Darstellungsform					
Variationsbreite	0,24	0,37	0,13	0,19	0,32
Gleichförmigkeit	0,32	0,82	0,02	0,19	0,17
Wechselhaftigkeit	0,32	0,19	0,15	0,36	0,51
Medien: Gerät/Material					
Variationsbreite	0,13	0,37	0,10	0,14	0,04
Gleichförmigkeit	0,08	0,06	0,13	0,02	0,14
Wechselhaftigkeit	0,20	0,11	0,17	0,12	0,29
Umfang Medieninhalte					
Variationsbreite	-0,14	0,05	0,08	-0,05	0,03
Gleichförmigkeit	0,11	0,22	-0,12	0,11	-0,02
Wechselhaftigkeit	-0,13	0,18	0,06	-0,05	0,01
<i>Merkmale der Aufgabeninhalte</i>					
Lebensweltbezug					
Variationsbreite	0,30	0,03	0,11	0,03	0,08
Gleichförmigkeit	0,26	0,26	0,15	0,04	0,19
Wechselhaftigkeit	0,22	0,17	0,14	0,15	0,28
Abstraktheit					
Variationsbreite	0,02	0,38	0,42	0,34	0,08
Gleichförmigkeit	0,11	0,05	0,72	0,49	0,22
Wechselhaftigkeit	0,06	0,12	0,56	0,65	0,09
<i>Merkmale der Bearbeitungsanforderungen</i>					
Kognitiver Prozess					
Variationsbreite	0,04	0,02	0,22	0,30	0,08
Gleichförmigkeit	0,09	0,10	0,05	0,03	0,02
Wechselhaftigkeit	0,01	0,06	0,26	0,10	0,16

- Fortsetzung der Tabelle s. nächste Seite -

- Fortsetzung von Tab. 24 -

Variabilitätsindikatoren der Aufgabenmerkmale	Korrelationsdifferenzen zwischen den Untergruppen				
	Geschlecht (m/w)	Umgangssprache i. d. Familie (D/nicht D)	Leistungsniveau im Klassenvergleich		
			niedrig/mittel	mittel/hoch	niedrig/hoch
<i>Merkmale der Bearbeitungsanforderungen (Fortsetzung)</i>					
Wissensdimension					
Variationsbreite	0,08	0,02	0,11	0,29	0,40
Gleichförmigkeit	0,07	0,38	0,22	0,16	0,39
Wechselhaftigkeit	0,08	0,24	0,18	0,04	0,23
Format d. Produkts					
Variationsbreite	0,19	0,11	0,21	0,01	0,20
Gleichförmigkeit	0,03	0,50	0,20	0,09	0,30
Wechselhaftigkeit	0,06	0,02	0,31	0,05	0,36
Art d. Produkts					
Variationsbreite	0,19	0,33	0,01	0,14	0,13
Gleichförmigkeit	0,04	0,80	0,16	0,14	0,02
Wechselhaftigkeit	0,16	0,65	0,21	0,16	0,05
Reflexions-/Auswertungsfunktion					
Variationsbreite	0,06	0,64	0,14	0,14	0,00
Gleichförmigkeit	0,08	0,15	0,40	0,50	0,10
Wechselhaftigkeit	0,13	0,47	0,08	0,27	0,19
<i>Merkmale des Zeitumfangs der Aufgaben</i>					
Dauer gesamt	0,09	0,22	0,41	0,35	0,06
Dauer Stellung	0,01	0,02	0,37	0,08	0,29
Dauer Bearbeitung	0,13	0,09	0,73	0,72	0,02
Aufgabenanzahl	0,23	0,30	0,32	0,07	0,26

n = 16 Unterrichtsstunden; ^A keine Variationsbreite vorhanden für diese Variable (s. Kap. 7.3.3., Tab. 19)

Die Beschreibung der Ergebnisse erfolgt zunächst über eine ausführliche Betrachtung der einzelnen Aufgabenmerkmale, bevor am Ende des Kapitels die erkennbaren Ergebnismuster zusammenfassend dargestellt werden.

Beschreibung der Einzelergebnisse: Im Bereich der Merkmale zur Beschreibung der Einführung und Präsentation der Aufgabe zeigt sich zunächst, dass eine höhere Variabilität⁵³ der Sprachverwendung durch die Lehrkraft in den Aufgabenstellungen

⁵³ „Variabilität“ hier und im Folgenden als zusammenfassender Begriff für alle drei Indikatoren (Variationsbreite, Gleichförmigkeit, Wechselhaftigkeit, vgl. Kap. 7.3.3).

allgemein leicht mit niedrigeren *on task*-Raten in den entsprechenden Stunden zusammenhängt, was für alle drei Variabilitätsindikatoren in ähnlichem Ausmaß feststellbar ist ($r_s^{\text{Variationsbreite}} = -0,20$; $r_s^{\text{Gleichförmigkeit}} = 0,32^{54}$; $r_s^{\text{Wechselhaftigkeit}} = -0,29$; s. Tab. 23). Dieser Effekt differenziert sich in den Untergruppen jedoch teilweise deutlich aus, wie es die mittleren bis großen Korrelationsdifferenzen⁵⁵ zeigen: Während sich das Ergebnis bei Schülern mit nicht-deutscher Familiensprache in etwas abgeschwächter Form wiederfindet, zeigen Schüler mit deutscher Familiensprache eher die gegenläufige Zusammenhangstendenz ($q = 0,31 - 0,53$; s. Tab. 24); Schüler mit niedrigem fachlichen Leistungsstand indessen zeigen den allgemeinen Effekt gegenüber den Schülern im Leistungsmittelfeld und im oberen Leistungsbereich in deutlich höherer Intensität ($q = 0,23 - 0,43$). Dieses Ergebnis scheint insofern bemerkenswert, dass bedingt durch die Häufigkeitsverteilung der Ausprägungen dieses Merkmals (s. Kap. 7.3.2, Tab. 15) eine höhere Variabilität mit einem höheren Anteil an Aufgabeneinführungen auf Deutsch einhergeht und man auf Seiten der leistungsschwachen Schüler bei deutschsprachig formulierten Aufgabenstellungen eher höhere Beteiligungen erwarten würde, da sie möglicherweise besser verstehen, was von ihnen gefordert wird.

Der verbale Umfang der Aufgabenstellungen zeigt bei höherer Variationsbreite und geringerer Gleichförmigkeit der Ausprägungen in einer Stunde durchschnittlich höhere *on task*-Raten, und dies in leichtem bis mittlerem Ausmaß ($r_s^{\text{Variationsbreite}} = 0,21$; $r_s^{\text{Gleichförmigkeit}} = -0,41$; s. Tab. 23). Weniger von Bedeutung scheint es aber zu sein, wie oft der Umfang der verbalen Ausführungen durch die Lehrkraft von Aufgabe zu Aufgabe variiert wird ($r_s^{\text{Wechselhaftigkeit}} = 0,14$). Die teilweise mittleren bis großen Korrelationsdifferenzen zwischen den Schüleruntergruppen (s. Tab. 24) lassen nur bei der Einteilung in fachbezogene Leistungsstände ein klares Muster erkennen: Je höher das fachliche Leistungsniveau der Schüler, desto stärker fällt der Zusammenhang zwischen der Variabilität dieses Aufgabenmerkmals und dem Beteiligungsverhalten aus.

Hinsichtlich der Anzahl an Hilfestellungshinweisen der Lehrkraft in den Aufgabenstellungen ist auf Ebene der gesamten Schülerstichprobe eine schwache

⁵⁴ Unterschiedliche Korrelationsrichtungen aufgrund negativer Wirkrichtung des Gleichförmigkeitsindikators (vgl. Kap. 7.3.3).

⁵⁵ Konventionen zur Interpretation der Effektstärke von Korrelationsdifferenzen: $q = 0,10 \rightarrow$ gering; $q = 0,30 \rightarrow$ mittel; $q = 0,50 \rightarrow$ groß (Cohen, 1988, zitiert nach Bühner & Ziegler, 2009, S. 609f.)

Korrelation im Bereich der stundenspezifischen Wechselhaftigkeit zu finden ($r_s^{\text{Wechselhaftigkeit}} = -0,20$; Tab. 23). Der Zusammenhang tritt jedoch bei den Schülern mit deutscher Umgangssprache in der Familie etwas deutlicher hervor ($r_s^{\text{Wechselhaftigkeit}} = -0,30$), was auch für den Indikator der Gleichförmigkeit gilt ($r_s^{\text{Gleichförmigkeit}} = 0,41$). Diese Ergebnisse bedeuten, dass die *on task*-Rate besonders bei diesen Schülern tendenziell niedriger ausfällt, wenn die Anzahl an Hilfestellungshinweisen in den Aufgabenstellungen im Stundenverlauf des Öfteren wechselt, aber auch wenn eine Ausprägungsstufe in der Stunde besonders dominiert – ein insgesamt unklares Ergebnis.

Bei den Darstellungsformen der verwendeten Medien bleiben allgemeine Effekte für die gesamte Schülerstichprobe aus (Korrelationseffizienten auf Ebene „alle Schüler“: $r_s = -0,09$ bis $0,15$; s. Tab. 23). Allerdings sind deutliche Korrelationsdifferenzen in allen Personenunterkategorien festzustellen ($q = 0,24 - 0,82$; s. Tab. 24). Eine höhere Variabilität der medienbasierten Darstellungen hängt demnach besonders bei Schülern mit deutschsprachigem Hintergrund ($r_s^{\text{Variationsbreite}} = 0,50$; $r_s^{\text{Gleichförmigkeit}} = -0,58$), Mädchen ($r_s^{\text{Variationsbreite}} = 0,31$; $r_s^{\text{Gleichförmigkeit}} = -0,36$) und Schülern mit hohem fachlichem Leistungsstand ($r_s^{\text{Variationsbreite}} = 0,37$; $r_s^{\text{Gleichförmigkeit}} = -0,28$) mit höheren *on task*-Raten zusammen (Tab. 23). Die Werte des Wechselhaftigkeitsindikators weichen von diesem Differenzierungsmuster ab und fallen insgesamt auch etwas schwächer aus.

Ein weiteres Kennzeichen des Medieneinsatzes, die als Träger fungierenden Geräte bzw. Materialien, lässt für zwei der drei Variabilitätsindikatoren einen Effekt auf allgemeiner Ebene erkennen, welcher zwar eher schwach ausfällt ($r_s^{\text{Variationsbreite}} = 0,22$; $r_s^{\text{Gleichförmigkeit}} = -0,27$), sich aber bei leicht variierender Zusammenhangsstärke durch fast alle Untergruppen der Schülerstichprobe zieht. Demzufolge ist die durchschnittliche *on task*-Rate einer Stunde tendenziell höher, wenn mehr unterschiedliche Geräte und Materialien eingesetzt und über den Unterrichtsverlauf hinweg oftmals variiert werden. Nennenswerte Korrelationsunterschiede zwischen Untergruppen lassen sich dabei nur im familiären Sprachhintergrund erkennen: Bezogen auf den Indikator Variationsbreite trifft der gefundene Zusammenhang nicht auf Schüler mit deutschsprachigem Hintergrund zu, dafür aber umso stärker auf Schüler mit nicht-deutschsprachigem Hintergrund ($r_s^{\text{Variationsbreite}} = -0,04/0,31$; $q = 0,37$; s. Tab. 23 und 24). Dieser Befund schließt an die Ergebnisse auf Verlaufsebene des Unterrichts an (Kap. 7.4.1), wo bedeutend niedrigere *on task*-Raten beim Einsatz des Lehrbuchs belegt werden und sich auch dieser Effekt besonders stark bei eben jener Schülergruppe zeigt (vgl. Tab. 22).

Abgesehen davon deuten die insgesamt kaum nennenswerten Koeffizienten auf Seiten des Indikators der Wechselhaftigkeit darauf hin, dass nicht das häufige Wechseln des Mediums innerhalb einer Unterrichtsstunde von Bedeutung ist für das Beteiligungsverhalten der Schüler, sondern – mit Blick auf die Koeffizienten des Indikators Gleichförmigkeit – eher die Vermeidung eines allzu dominanten Einsatzes desselben Geräts oder Materials über die Stunde hinweg. Abschließend ist zu den Aspekten des Medieneinsatzes festzustellen, dass für den Umfang der Medieninhalte ein auf allgemeiner Ebene schwacher ($r_s^{\text{Gleichförmigkeit}} = -0,20$; Tab. 23) und mit Blick auf Schüler mit deutschsprachigem Familienhintergrund ($r_s^{\text{Gleichförmigkeit}} = -0,34$) und Schüler im Leistungsmittelfeld ($r_s^{\text{Gleichförmigkeit}} = -0,28$) etwas stärkerer Zusammenhang zum Beteiligungsverhalten erkennbar ist.

Im Bereich der Merkmale zur Beschreibung der Aufgabeninhalte ist bezüglich der Variationsbreite des Lebensweltbezugs ein geringer Zusammenhang zur stundenspezifischen *on task*-Rate auf allgemeiner Ebene zu verzeichnen ($r_s^{\text{Gleichförmigkeit}} = 0,20$). Dieser Zusammenhang setzt sich in ähnlichem Ausmaß über die Untergruppen hinweg fort, wobei der Effekt bei den Mädchen etwas stärker ausfällt als bei den Jungen und dabei annähernd ein mittleres Niveau erreicht ($r_s^{\text{Variationsbreite}} = 0,29/0,00$; $q = 0,30$). Besonders bei Mädchen ist demnach die *on task*-Rate in einer Stunde tendenziell höher, wenn in dessen Verlauf unterschiedliche Stufen des Lebensweltbezugs hergestellt werden. Dieser Befund kann als Bestätigung der Erkenntnisse auf Ebene des Unterrichtsverlaufs (s. Kap. 7.4.1) gesehen werden, wo aufgezeigt wurde, dass die *on task*-Raten bei Mädchen während Aufgaben ohne Lebensweltbezug signifikant höher ausfallen als bei Aufgaben mit konstruiertem Lebensweltbezug. Letzteres ist die Ausprägungsstufe mit der größten Auftretenshäufigkeit des Merkmals und damit wahrscheinlich der Grund für den gefundenen Effekt an dieser Stelle, da eine hohe Variationsbreite mit häufigeren Abweichungen von dieser Kategorie einhergeht.

Mit Blick auf die Abstraktheit der Aufgabeninhalte scheint die Variationsbreite hingegen kaum relevant zu sein für das Beteiligungsverhalten der Schüler in der Stunde, eine geringere Gleichförmigkeit und größere Abwechslung hingegen schon, zumindest in mäßigem Umfang und auf allgemeiner Ebene ($r_s^{\text{Gleichförmigkeit}} = -0,26$; $r_s^{\text{Wechselhaftigkeit}} = 0,27$). Dieser Zusammenhang differenziert sich jedoch bei Unterscheidung nach Leistungsstand der Schüler aus, das heißt, Schüler im Leistungsmittelfeld zeigen deutlich stärkere Zusammenhänge als ihre Mitschüler, wie die Korrelationsdifferenzen

zeigen ($q = 0,49 - 0,72$; s. Tab. 24). Ähnliche differenzielle Effekte sind auch beim ansonsten auf allgemeiner Ebene unauffälligen Indikator der Variationsbreite zu erkennen.

Die zur Bearbeitung der Aufgaben notwendigen kognitiven Prozesse zeigen für die gesamte Schülerstichprobe mäßige Zusammenhänge zur *on task*-Rate, und zwar hinsichtlich ihrer Variationsbreite ($r_s^{\text{Variationsbreite}} = 0,24$, s. Tab. 23) und der Wechselhaftigkeit ($r_s^{\text{Wechselhaftigkeit}} = 0,25$). Eine größere Vielfalt und häufigere Abwechslung der kognitiven Prozesse hängt demnach tendenziell mit einer höheren *on task*-Rate zusammen. Trotz variierender Stärke der Ausprägung (z.B. Schüler auf mittlerem Leistungsniveau: $r_s^{\text{Variationsbreite}} = 0,41$; Jungen: $r_s^{\text{Variationsbreite}} = 0,15$) lässt sich dieser Zusammenhang als allgemeiner Effekt für alle Personengruppen bezeichnen. Der Gleichförmigkeitsindikator zeigt an dieser Stelle keine Korrelationen nennenswerten Ausmaßes. Das deutet darauf hin, dass die Abwechslung und Vielfalt der zur Bearbeitung notwendigen kognitiven Prozessen im Stundenverlauf eine größere Rolle spielt für das Schülerbeteiligungsverhalten als die besondere Gewichtung eines bestimmten kognitiven Prozesses im Unterrichtsverlauf.

Die zur Bearbeitung notwendigen Wissensdimensionen geben hinsichtlich der Korrelationswerte der unterschiedlichen Variabilitätsindikatoren ein ähnliches Bild ab wie die kognitiven Prozesse: Auch hier sind auf allgemeiner Ebene geringe bis mittlere Zusammenhänge zwischen der Variationsbreite ($r_s^{\text{Variationsbreite}} = 0,24$) und der Wechselhaftigkeit ($r_s^{\text{Wechselhaftigkeit}} = 0,37$) des Merkmals und dem Schülerbeteiligungsverhalten zu finden; ebenso wie bei den kognitiven Prozessen hängt demzufolge eine größere Vielfalt und größere Anzahl an Wechseln der Ausprägungen mit höheren stundenspezifischen *on task*-Raten zusammen. Im Unterschied zu den kognitiven Prozessen, wo die Schüler im Leistungsmittelfeld die größten Zusammenhänge aufweisen, sind es hier allerdings die Schüler auf hohem Leistungsniveau, die deutlich höhere Korrelationen erkennen lassen als die anderen beiden Leistungsgruppen – fachlich stärkere Schüler sprechen somit scheinbar etwas stärker auf die Variabilität der geforderten Wissensarten in einer Unterrichtsstunde an. Daneben sind bei diesem Aufgabenmerkmal weitere differenzielle Effekte in Abhängigkeit vom familiären Sprachhintergrund zu verzeichnen: Schüler mit deutschsprachigem Familienhintergrund zeigen größere Zusammenhänge zur *on task*-Rate als jene Schüler mit nicht-deutscher Familiensprache.

Auch bei dem Aufgabenmerkmal „Format des geforderten Produkts“ (d.h. der Verarbeitung von Vorgaben, eingeschränkter oder erweiterter Eigenkonstruktion) sind mäßige Zusammenhänge auf Ebene der gesamten Schülerstichprobe ($r_s^{\text{Variationsbreite}} = 0,27$; $r_s^{\text{Gleichförmigkeit}} = 0,27$; $r_s^{\text{Wechselhaftigkeit}} = 0,27$) zu konstatieren. Das heißt, insbesondere eine größere Variationsbreite und häufigere Wechsel des geforderten Produktformats hängen leicht mit höheren *on task*-Raten zusammen. Während sich dieser Effekt auf Seiten der Variationsbreite durch alle Personengruppen zieht, spaltet sich der allgemeine Effekt des Indikators der Wechselhaftigkeit auf, wenn man die unterschiedlichen Leistungsgruppen innerhalb der Klasse betrachtet: Schüler mit niedrigem Leistungsstand lassen den beschriebenen Zusammenhang in deutlich geringerem Maße erkennen als Schüler auf mittlerem oder hohem Leistungsniveau ($r_s^{\text{Wechselhaftigkeit}} = 0,12/0,40/0,44$). Diese Gruppendifferenz der Korrelationen ist auch festzustellen im Bereich des Gleichförmigkeitsindikators ($r_s^{\text{Gleichförmigkeit}} = -0,04/-0,24/-0,34$). Ein deutlicher differenzieller Effekt ist überdies bei der Personenunterscheidung nach familiärem Sprachhintergrund zu registrieren ($r_s^{\text{Gleichförmigkeit}} = -0,53/-0,09$; $q = 0,50$). Demnach zeigen besonders Schüler im mittleren und oberen Leistungsbereich und mit deutschsprachigem Familienhintergrund höhere *on task*-Raten in Unterrichtsstunden mit größerer Variabilität des geforderten Produktformats. An dieser Stelle fällt auf, dass die drei zuletzt beschriebenen Aufgabenmerkmale (kognitive Prozesse, Wissensdimensionen und gefordertes Produktformat) insofern ähnliche Ergebnismuster zeigen, dass jeweils auf Ebene der gesamten Schülerstichprobe die stundenspezifische Variationsbreite und Wechselhaftigkeit des Merkmals mäßig bis mittelstark mit höheren *on task*-Raten zusammenhängt, wobei sich diese Zusammenhänge bei einer Unterscheidung nach Leistungsgruppen und Sprachhintergrund der Schüler teilweise deutlich ausdifferenzieren.

Ein anderes Bild bietet sich bei der Art des geforderten Produkts (Sprache, Schrift, Grafik, anderer Handlungsvollzug). Für dieses Merkmal ist den Korrelationskoeffizienten zufolge kein allgemeiner Effekt erkennbar ($r_s = -0,12 - 0,04$, Tab. 23), stattdessen sind teilweise starke differenzielle Effekte bei der Personenunterscheidung nach familiärem Sprachhintergrund zu finden ($q = 0,33 - 0,80$, Tab. 24). Schüler mit deutschsprachigem Hintergrund zeigen höhere *on task*-Raten in Unterrichtsstunden mit größerer Variationsbreite und Wechselhaftigkeit und niedrigerer Gleichförmigkeit des geforderten Produkts ($r_s^{\text{Variationsbreite}} = 0,20$;

$r_s^{\text{Gleichförmigkeit}} = -0,66$; $r_s^{\text{Wechselhaftigkeit}} = 0,59$), Schüler mit nicht-deutschsprachigem Hintergrund hingegen kaum ($r_s^{\text{Variationsbreite}} = -0,13$; $r_s^{\text{Gleichförmigkeit}} = -0,01$; $r_s^{\text{Wechselhaftigkeit}} = 0,02$). Unter Berücksichtigung der Verteilung der Ausprägungsstufen dieses Merkmals (vgl. Kap. 7.3.2, Tab. 15) und der gefundenen Effekte auf Ebene der Einzelaufgaben im Unterrichtsverlauf (Kap. 7.4.1, Tab. 22) scheint es für das Beteiligungsverhalten dieser Schülergruppe von besonderer Bedeutung zu sein, dass das inhaltliche Geschehen innerhalb des öffentlichen Unterrichts einer Unterrichtsstunde nicht monoton auf rein verbale Äußerungen der Schüler aufbaut.

Hinsichtlich der Reflexions- und Auswertungsfunktion der Aufgaben sind vereinzelte Zusammenhänge bzw. auch Korrelationsdifferenzen von nennenswertem Ausmaß in den Schüleruntergruppen feststellbar, welche jedoch im Gesamtbild keine Systematik erkennen lassen und von denen zusätzlich anzunehmen ist, dass sie durch ungleiche Häufigkeitsverteilungen der Ausprägungsstufen in den Klassen verzerrt sind (s. Erläuterungen zu den Ergebnissen dieses Merkmals in Kap. 7.4.1).

Um neben den Indikatoren zur Abbildung verschiedener Aspekte der Variabilität der Aufgabenmerkmale in den Unterrichtsstunden mögliche Zeiteffekte des stundenspezifischen Aufgabenangebots identifizieren zu können, wurde auch der Zeitumfang der Aufgaben, bzw. ihrer Teilkomponenten Stellung und Bearbeitung, in Form des Stundenmittelwerts zur Berechnung der Zusammenhänge herangezogen. Dabei sind zwei allgemeine Effekte festzustellen: Wie die Ergebnisse zeigen, ist die stundenspezifische *on task*-Rate tendenziell höher bei längerer Durchschnittsdauer der Aufgabenstellungen ($r_s = 0,27$), aber auch bei kürzerer durchschnittlicher Bearbeitungszeit ($r_s = -0,31$). Durch den hohen Anteil der Aufgabenbearbeitungszeit an der Aufgabengesamtzeit (vgl. Kap. 7.3.2, Tab. 15) zeigen diese beiden Zeitkomponenten ähnliche Werte. Der Befund auf Seiten der durchschnittlichen Aufgabenbearbeitungsdauer bestätigt die Ergebnisse auf Ebene der Einzelaufgaben (Kap. 7.4.1, Tab. 21). Überraschend ist jedoch, dass die zeitliche Länge der Aufgabenstellungen im Stundendurchschnitt einen – immerhin annähernd mittelstarken – Zusammenhang erkennen lässt, auf Ebene der Einzelaufgabe jedoch nicht (vgl. Tab. 21: Rangkorrelation Zeitumfang der Aufgabenstellung mit gleichzeitiger *on task*-Rate auf allgemeiner Ebene: $r_s = 0,06$; keine differenziellen Effekte in den Untergruppen der Schülerstichprobe). Im Hinblick auf die Ergebnisse dieses Abschnitts ist zu beachten, dass sich die auf allgemeiner Ebene gefundenen Zusammenhänge in den

Leistungsgruppen innerhalb der Schülerstichprobe ausdifferenzieren: Der Effekt höherer *on task*-Raten bei längerer Durchschnittsdauer der Aufgabenstellungen in einer Stunde fällt bei Schülern auf niedrigem Leistungsniveau mittelstark aus, bei den anderen beiden Leistungsgruppen hingegen verschwindet der Zusammenhang ($r_s = 0,36/0,01/0,09$, s. Tab. 23). Dies lässt sich dadurch begründen, dass leistungsschwächere Schüler möglicherweise mehr Zeit brauchen, um zu verstehen, was in der Aufgabe von Ihnen gefordert wird. Weiterhin zeigt der Blick auf die Leistungsgruppen Korrelationsunterschiede hinsichtlich der Bearbeitungsdauer der Aufgaben: Hierbei sind es hauptsächlich die Schüler mit mittlerem Leistungsstand, die bei längerer Durchschnittsdauer niedrigere *on task*-Raten erkennen lassen ($r_s = -0,14/-0,53/-0,12$).

Abschließend zeigt die Anzahl an Aufgaben pro Unterrichtsstunde insgesamt nur schwache Zusammenhänge zum Schülerbeteiligungsverhalten in Form der stundenspezifischen *on task*-Rate. Lediglich jene Schüler mit mittlerem und hohem fachlichem Leistungsstand zeigen mäßige bzw. mittlere positive Zusammenhänge ($r_s = 0,28/0,22$; Tab. 23), Schüler mit niedrigem Leistungsniveau hingegen nicht ($r_s = -0,04$; Korrelationsdifferenz zu mittlerem und hohem Leistungsniveau: $q = 0,32/0,26$, Tab. 24). Das bedeutet, Schüler mit mindestens mittlerem Leistungsstand im Fach lassen tendenziell mehr Beteiligungsverhalten erkennen, wenn in der Unterrichtsstunde mehr Aufgaben gestellt werden.

Zusammenfassung der Ergebnisse dieses Abschnitts: In diesem Auswertungsschritt wurde mit Blick auf die gesamte Unterrichtsstunde als Analyseeinheit geprüft, inwiefern unterschiedliche Aspekte der stundenspezifischen Variabilität der Aufgabeneigenschaften mit der *on task*-Rate im Stundendurchschnitt zusammenhängen. Dabei wurden neben der gesamten Schülerstichprobe auch die anhand der Personenmerkmale unterscheidbaren Schüleruntergruppen in den Blick genommen, um allgemeine und differenzielle Effekte feststellen zu können. Die in diesem Zuge erzeugten Ergebnisse lassen unterschiedliche Ergebnismuster erkennen, die aufgrund der Komplexität der Einzelbefunde nach folgender Systematik beschrieben werden: (1) Vergleich der Aufgabenmerkmale und Merkmalsbereiche, (2) Betrachtung der

Differenzen zwischen den Untergruppen der Schülerstichprobe und (3) Ergebnismuster zwischen den Variabilitätsindikatoren.

(1) Aufgabenmerkmale und Merkmalsbereiche

Unter Berücksichtigung der gesamten Schülerstichprobe weisen viele der untersuchten Aufgabenmerkmale in ihrer Variabilität schwache bis mittlere Zusammenhänge ($r_s = 0,20 - 0,30$) zur stundenspezifischen *on task*-Rate auf. Allerdings zeigt kaum ein Aufgabenmerkmal durchgängig das gleiche Korrelationsmuster über alle Teilgruppen der Schüler hinweg, bei allenfalls geringen gruppenspezifischen Schwankungen der Korrelationshöhe. Dies lässt sich lediglich für die kognitiven Prozesse der Bearbeitung und den inhaltlichen Umfang der verwendeten Medien behaupten. Wie an den Korrelationsdifferenzen zwischen den Untergruppen (Tab. 24) und auch den Korrelationskoeffizienten selbst (Tab. 23) abzulesen ist, ist die Mehrzahl der Aufgabenmerkmale, die einen Zusammenhang auf Ebene der gesamten Schülerstichprobe aufweisen, bei Berücksichtigung der Personenuntergruppen von einer Ausdifferenzierung des allgemeinen Effekts betroffen (s. Sprache der Einführung und Präsentation, verbaler Umfang der Aufgabenstellung, Anzahl an Hilfestellungshinweisen in der Aufgabenstellung, Lebensweltbezug und Abstraktheit der Aufgabeninhalte, in der Bearbeitung geforderte Wissensdimension, Format und Art des geforderten Produkts und mittlerer Zeitumfang der Aufgabenstellungen und -bearbeitungen). Daneben sind bei zwei Aufgabenmerkmalen nur in einzelnen Schüleruntergruppen nennenswerte Zusammenhänge zu finden, ohne einen Zusammenhang auf allgemeiner Ebene (Darstellungsform des Mediums, Anzahl an Aufgaben in der Unterrichtsstunde). Mit Blick auf die übergeordneten Merkmalsbereiche lässt sich keine besondere Gewichtung der beschriebenen Ergebnismuster erkennen, ebenso wenig hinsichtlich der Stärke der Zusammenhänge zum Beteiligungsverhalten.

(2) Differenzen zwischen den Untergruppen der Schülerstichprobe

Es wurde deutlich, dass eine Vielzahl an nennenswerten Korrelationsdifferenzen innerhalb der unterschiedlichen Personenunterscheidungskategorien vorliegt und aus diesem Grund kaum einer der gefundenen Zusammenhänge zwischen stundenspezifischer Variabilität des Aufgabenmerkmals und *on task*-Rate als allgemeingültig (im Rahmen der Stichprobe) bezeichnet werden kann. Die

Korrelationsdifferenzen verteilen sich allerdings nicht gleichmäßig auf alle Personenkategorien. Besonders deutlich treten bei einer Reihe von Aufgabenmerkmalen die Korrelationsunterschiede zwischen den Leistungsgruppen hervor, wie z.B. bei der Sprache der Einführung und Präsentation der Aufgabe, dem verbalen Umfang der Aufgabenstellung, den Darstellungsformen in den Unterrichtsmedien, der Abstraktheit der Aufgabeninhalte, dem Format des geforderten Produkts und den Aspekten des Zeitumfangs. Teilweise in Überschneidung dazu gibt es daneben auch einige Differenzen zwischen den unterschiedlichen Sprachhintergründen in der Familie zu verzeichnen, wie z.B. bei der Anzahl an Hilfestellungshinweisen in den Aufgabenstellungen, der Art des geforderten Produkts und auch den bereits zuvor genannten Merkmalen. Nicht zu übersehen ist dabei die Tatsache, dass bei jenen Aufgabenmerkmalen, die keine oder nur sehr schwache Zusammenhänge auf allgemeiner Ebene zeigen, zugleich deutliche Korrelationsdifferenzen zwischen den Schülern unterschiedlichen familiären Sprachhintergrunds vorliegen (Anzahl an Hilfestellungshinweisen in den Aufgabenstellungen, Darstellungsformen in den Unterrichtsmedien, Art des geforderten Produkts). Zudem fällt ins Auge, dass diese beiden Personenkategorien im Gegensatz zu den Leistungsgruppen in den Aspekten des Zeitumfangs keinerlei Korrelationsdifferenzen zeigen. Bezüglich der Unterscheidung der Schüler nach Geschlecht sind hingegen, bis auf die Darstellungsformen der Unterrichtsmedien und den Lebensweltbezug der Aufgabeninhalte, keine nennenswerten Korrelationsunterschiede festzustellen.

(3) Ergebnismuster zwischen den Variabilitätsindikatoren

Wie die Interkorrelationen zwischen den Variabilitätsindikatoren es bereits nahe legen (s. Kap. 7.3.3. und Anhang B), fallen die Werte der drei Indikatoren innerhalb der Aufgabenmerkmale nicht immer gleichsinnig oder in vergleichbarem Umfang aus, was dafür spricht, dass die Indikatoren an diesen Stellen unterschiedliche Qualitäten der stundenspezifischen Variabilität abbilden. In der Zusammenschau lassen sich unterschiedliche Schwerpunkte in den Ergebnissen der Variabilitätsindikatoren in Bezug auf die übergeordneten Aufgabenmerkmalsbereiche feststellen. Bei fast allen Aufgabenmerkmalen aus dem Bereich der kognitiven Anforderungen der Bearbeitung (kognitiver Prozess, angesprochene Wissensdimension und Format des geforderten Produkts) schlagen sich die nennenswerten Zusammenhänge hauptsächlich in den Indikatoren Variationsbreite und Wechselhaftigkeit nieder (vgl. Tab. 23). Bei diesen

Aufgabenmerkmalen ist es für das Schülerbeteiligungsverhalten demnach weniger von Bedeutung, ob ein bestimmter kognitiver Prozess, eine Wissensdimension oder ein Produktformat im Stundenverlauf überrepräsentiert ist, sondern vielmehr, ob eine solche quantitative Dominanz zwischenzeitlich immer wieder durch Ausprägungswechsel unterbrochen wird, und dies mit einer möglichst breiten Spanne an Ausprägungen. Bei anderen Aufgabenmerkmalen hingegen spielt die Gleichförmigkeit des Merkmals im Unterrichtsverlaufs eine gleich geordnete Rolle (Sprache der Einführung und Präsentation, Mediengerät bzw. -material, Abstraktheit der Aufgabeninhalte). Konkret heißt das, je mehr Aufgaben einer Stunde sich beispielsweise auf dasselbe Arbeitsmaterial beziehen bzw. je monotoner das Aufgabenmerkmal in der Unterrichtsstunde gehalten wird, desto niedriger fällt die *on task*-Rate in dieser Stunde aus.

8. Zusammenfassung und Diskussion

Ausgehend von dem übergeordneten Bestreben, Wirksamkeitszusammenhänge innerhalb von Unterrichtsprozessen zu beleuchten (s. Kap. 1), bestand das Ziel der vorliegenden Studie darin, mittels videogestützter Beobachtungen das Beteiligungsverhalten der Schüler im Unterrichtsverlauf zu untersuchen und durch die Eigenschaften der durch die Lehrkraft eingebrachten Aufgaben zu erklären. Im Fokus stand dabei die Realität „normalen Alltagsunterrichts“ an Regelschulen, womit sich die Untersuchung an einem praxisnahen Unterrichtsbegriff orientiert, welcher die Annahme der Beeinflussbarkeit des institutionalisiert organisierten Unterrichtsgeschehens durch das (teilweise geplante) Handeln der Lehrkraft impliziert (vgl. Kap. 2.1). Die effektivitätsorientierte Sichtweise auf Unterricht stellt Bezüge zu einem Forschungsstrang der internationalen Unterrichtsforschung her, dessen gemeinsamer theoretischer Ausgangspunkt in Carrolls Modell schulischen Lernens (1963) zu suchen ist (s. Kap. 2.2). Dieses Modell bot die Grundlage sowohl für einen breiten Bestand an empirischen Arbeiten als auch eine Reihe theoretischer Weiterentwicklungen, von denen das Modell der Unterrichtseffektivität nach Slavin (1984, 1994) den zentralen theoretischen Bezugspunkt dieser Arbeit darstellt. In dem Modell wird ein Zusammenhang formuliert, welcher in der Forschung zur Unterrichtseffektivität bzw. Unterrichtsqualität – und teilweise auch darüber hinaus – als theoretischer Konsens verstanden werden kann (Kap. 2.2) und in zahlreichen empirischen Studien bestätigt werden konnte (Kap. 2.3.1): Demnach ist die Beteiligung der Schüler an bzw. ihre Zuwendung zu den Lerngelegenheiten des Unterrichts eine wichtige Voraussetzung für potenzielle Lern- oder Bildungseffekte des Unterrichts. Damit nimmt dieser Bereich die Position eines prozessimmanenten Indikators für die Effektivität von Unterrichtsprozessen ein. Als bedingende Größen für diesen im vorliegenden Kontext als Beteiligungsverhalten bezeichneten Bereich (Kap. 2.4) werden sowohl verschiedene Aspekte des persönlichen Hintergrunds bzw. der Persönlichkeit der Schüler als auch die situativen Bedingungen im Unterricht angesehen, wobei die bisherige Forschung auf einen stärkeren Einfluss der Unterrichtsbedingungen hindeutet (Kap. 2.3.2). Innerhalb des Unterrichts ist sowohl dem Modell von Slavin (1994) zufolge, aber auch gemäß einer Reihe empirischer Arbeiten aus der Forschung zum *school engagement* (Fredricks, Blumenfeld & Paris, 2004) ein wichtiger Einflussfaktor zu suchen in der Qualität der Aufgaben im Unterricht. Allerdings liegen dazu kaum Forschungsergebnisse vor, die die

Prozessebene des Unterrichts in den Blick nehmen – womit jene Forschungslücke benannt ist, die in der vorliegenden Studie bearbeitet wurde. Angesichts der Tatsache, dass eine Betrachtung der Effektivität von Unterrichtsprozessen nicht ohne die Berücksichtigung der zeitlichen Dimension auskommt (Kap. 3), wurde eine geeignete Operationalisierung des Schülerbeteiligungsverhaltens im Anschluss an Slavin (1994) in der so genannten *time on task* (Bloom, 1974) gefunden. Im Hinblick auf eine hinsichtlich des Auflösungsgrades analoge Erfassung der Aufgaben wurde das zugrundeliegende Aufgabenverständnis dieser Arbeit – nach grundsätzlichen Betrachtungen zu dem Phänomen „Aufgabe“ – hauptsächlich an dem Aufgabenbegriff von Renkl (1991) orientiert (Kap. 4). Die zentrale Frage nach dem Einfluss der Aufgaben auf das Schülerbeteiligungsverhalten wurde anschließend in mehreren Teilfragestellungen ausformuliert (Kap. 5). Die eigene empirische Untersuchung wurde schließlich anhand von audiovisuellen Aufzeichnungen aus 23 Unterrichtsstunden im Fach Englisch in sechs Schulklassen der Sekundarstufe I und einem schriftlichen Test zum Leistungsstand der Schüler im Unterrichtsfach Englisch, kombiniert mit einer Befragung zum Sprachhintergrund in der Familie und dem Geschlecht ($n = 132$), vorgenommen (Kap. 6). Die Daten zum Leistungsstand im Unterrichtsfach, dem Geschlecht und dem familiären Sprachhintergrund der Schüler sollten dazu dienen, Einsicht in die Einflüsse der Personeneigenschaften auf ihr Beteiligungsverhalten im Unterricht zu bekommen und die Beobachtungsdaten nach diesen Merkmalen differenzieren zu können. Die verwendeten Beobachtungsverfahren zur Erfassung und Beschreibung des Schülerverhaltens und der Aufgaben wurden aufgrund des großen Anteils an Eigenentwicklung bzw. Adaption bestehender Instrumentarien ausführlich dargestellt. Die auf dieser Basis gewonnenen Ergebnisse werden im nachfolgenden Abschnitt zusammenfassend dargestellt.

8.1 Zusammenfassung und Einordnung der Ergebnisse

Ergebnisse zu den Personenmerkmalen der Schüler:

Die Ergebnisse zum Sprachhintergrund zeigen einen im Vergleich zu anderen Erhebungen (vgl. Hesse, Göbel und Hartig, 2008; Walter, 2008) geringen Anteil an Schülern mit allein deutscher Umgangssprache in der Familie (22 %; s. Kap. 7.1.1), während 24 % der Schüler zu Hause kein Deutsch und der Großteil (54 %) Deutsch und

zusätzlich eine andere Sprache sprechen. Daneben zeigt sich die Geschlechterverteilung unauffällig, d.h. annähernd gleich verteilt (s. Kap. 7.1.2). Die Ergebnisse des Leistungstests (Kap. 7.1.3) belegen eine große Leistungsheterogenität in der untersuchten Schülerstichprobe. Bei zu erwartenden Unterschieden in den Leistungsbereichen der unterschiedlichen Schulformen wird der Wertebereich des Leistungstests vollständig ausgeschöpft, und auch innerhalb der Klassen sind teilweise beträchtliche Leistungsspannen zu erkennen. Eine Abhängigkeit der Testleistung vom Geschlecht oder dem familiären Sprachhintergrund konnte nicht festgestellt werden.

Ergebnisse zum Beteiligungsverhalten der Schüler im Unterricht:

Anhand von audiovisuellen Aufzeichnungen, welche den Klassenraum aus zwei Kameraperspektiven zeigen und das Klanggeschehen mit mehreren Mikrofonen festhielten (s. Kap. 6.4), wurden des Weiteren die zu untersuchenden Prozesskomponenten des Unterrichts erfasst. Das Beteiligungsverhalten der Schüler im Unterricht wurde mit einem niedrig-inferenten, auf einem Zeitstichprobenplan basierenden Beobachtungsverfahren kodiert, mit dem das Verhalten jedes Schülers in Zeitschritten von einer Minute hinsichtlich des *on* bzw. *off task*-Status eingeschätzt wurde (s. Kap. 6.6.3). Die Kodierungen wurden aufgrund der situativen Spezifität erwartbaren Verhaltens nur in Phasen öffentlichen Unterrichts vorgenommen, die zuvor in einer Basiskodierung identifiziert worden waren (Kap. 6.6.2). Diese ergab mit durchschnittlich ca. 70 % der tatsächlichen Unterrichtszeit ein klares Gewicht auf der Sozialform des öffentlichen Unterrichts (s. Kap. 7.2.1), was in Einklang mit einer Reihe weiterer Befunde der Unterrichtsforschung steht (Bohl, Kleinknecht, Batzel & Richey, 2012; Götz, Lohrmann, Ganser & Haag, 2005; Hage et al., 1985; Kleinknecht, 2010; Seidel et al., 2006). Die Kodierung des Schülerbeteiligungsverhaltens bezog sich damit auf einen Großteil der Zeit des aufgezeichneten Unterrichts. Auf Grundlage dieser Zeitstichprobe und der Anzahl anwesender und sichtbarer Schüler in den Klassen (durchschnittlich 85 %) konnte in den 23 Unterrichtsstunden eine Beobachtungsstichprobe von insgesamt etwa 13.000 einminütigen Beobachtungsintervallen erreicht werden. Die Ergebnisse der Verhaltenskodierung, abgebildet anhand der *on task*-Rate, d.h. des Anteils der als *on task* kodierten Beobachtungsintervalle an allen untersuchten Beobachtungsintervallen im Verlauf einer

Unterrichtsstunde, zeigen in der Stichprobe eine durchschnittliche *on task*-Rate von 83,4 %, mit einer Spanne von 69 % - 98 % mit Blick auf die stundenweise Aggregation und einer Spanne von 30 % - 100 % beim Blick auf die personenweise Aggregation über die Unterrichtsstunden hinweg (s. Kap. 7.2.2). Diese Ergebnisse schließen an den bisherigen Forschungsstand an, welcher in der Mehrzahl *on task*- bzw. *engagement*-Raten in diesem Wertebereich nachweist (vgl. Kap. 2.3.3).

Ergebnisse zum Einfluss der Personenmerkmale auf das Beteiligungsverhalten:

Mit Blick auf die Wirkung von persönlichen Hintergrundmerkmalen der Schüler auf ihr individuelles Beteiligungsverhalten im Unterricht – bei Slavin (1994) bezeichnet als *student inputs* – wurde die Abhängigkeit des individuenbezogenen *on task*-Durchschnitts von den verfügbaren Personenvariablen (s.o.) geprüft (s. Kap. 7.2.3). Der familiäre Sprachhintergrund und der Leistungsstand im Unterrichtsfach lassen keinerlei Einfluss auf die individuelle Verhaltenstendenz erkennen, wohl aber das Geschlecht: Mädchen zeigen eine erkennbar höhere durchschnittliche *on task*-Rate als Jungen ($M_{\text{männlich}} = 78,7\%$; $M_{\text{weiblich}} = 87,3\%$; $t_{(122)} = -3,5$; $p = 0,001$; $d = 0,6$). Dieser Effekt ist in der Forschung zwar durchaus bekannt (vgl. Marks, 2000; Rimm-Kaufman, Baroody, Larsen, Curby & Abry, 2014; (Helmke & Renkl, 1993), gilt aber nicht als durchgängig belegt und fällt in der vorliegenden Form zudem ungewöhnlich stark aus. Neben dem Einfluss des persönlichen Hintergrunds wurde auch geprüft, ob interindividuelle Unterschiede im Beteiligungsverhalten durch die Zugehörigkeit zu einer Schulform oder die Schulklasse bedingt sind. Ein Effekt der Schulform konnte nicht festgestellt werden, die Zugehörigkeit zur Schulklasse zeigt daneben einen starken Effekt ($F_{(1;118)} = 6,14$; $p = 0,000$; $\eta^2 = 0,21$) – ein Befund, der an Ergebnisse von Helmke und Renkl (1993) und Helmke und Weinert (1997) zur Abhängigkeit des Aufmerksamkeitsverhaltens von der Schulklasse bzw. der Lehrkraft anschließt. Der Schulklasseneffekt überwiegt im direkten Vergleich den – wenngleich ebenfalls starken – Einfluss des Geschlechts, was an dieser Stelle als Hinweis auf ein größeres Gewicht der Unterrichtsbedingungen in ihrem Einfluss auf das Schülerbeteiligungsverhalten im Vergleich zu außerhalb des Unterrichts liegenden Größen gewertet wird. Damit findet auch eine in der bisherigen Forschung überwiegend festzustellende Tendenz Bestätigung (s. Kap. 2.3.2). Im Anschluss daran zeigt eine Betrachtung der zeitlichen Stabilität des Beteiligungsverhaltens bedeutsame

Schwankungen der Ausprägungen, und zwar sowohl innerhalb der Unterrichtsstunden als auch zum überwiegenden Teil zwischen den Unterrichtsstunden (s. Kap. 7.2.4). Dies untermauert zusätzlich die Annahme der Abhängigkeit des Merkmals von situativen Bedingungen im Unterrichtsprozess.

Ergebnisse zu den Aufgaben im Unterricht:

Die von der Lehrkraft in den Unterricht eingebrachten Aufgaben bilden jenes unterrichtliche Prozessmerkmal, dessen Einfluss auf das Beteiligungsverhalten der Schüler geprüft wurde. Dazu wurden die Aufgaben in 16 der 23 Unterrichtsstunden mit einem auf Ereignisstichproben basierenden Beobachtungsverfahren erfasst, welches jedem Aufgabenereignis den Start- und Endzeitpunkt des tatsächlichen Auftretens im Unterrichtsverlauf zuweist (s. Kap. 6.7.1). Die Kodierungen ergeben insgesamt 165 Aufgaben, 144 davon zu verorten innerhalb von Phasen öffentlichen Unterrichts (s. Kap. 7.3.1). Die daraus resultierende durchschnittliche Anzahl von 10 Aufgaben pro Unterrichtsstunde erscheint in Anbetracht der relativ fein auflösenden Ereignisdefinition vergleichsweise niedrig (s. Renkl, 1991: ca. 53 Aufgaben/Stunde; Schabram, 2007: 32 Aufgaben/Stunde). Allerdings ist anzunehmen, dass hierbei die Kodierregel zum Tragen kommt, welche die Kodierung einer neuen Aufgabe erst bei der Ausprägungsveränderung in mindestens einer Aufgabeneigenschaft vorsah (s. Kap. 6.7.1).

Die auf diese Weise erfassten Aufgaben wurden mithilfe eines Kategoriensystems hinsichtlich ihrer Eigenschaften analysiert. Das Kategoriensystem umfasst 14 verschiedene Analysemerkmale, welche zum großen Teil aus bestehenden Instrumentarien übernommen bzw. adaptiert wurden und nach den drei übergeordneten Bereichen „Eigenschaften der Einführung und Präsentation der Aufgabe“, „Eigenschaften der Aufgabeninhalte“ und „Eigenschaften der Anforderungen in der Bearbeitung der Aufgabe“ gruppiert sind (vgl. Kap. 6.7). Zusammengefasst zeigen die Ergebnisse der Aufgabenanalyse folgendes Bild (vgl. Kap. 7.3.2): Die Aufgaben haben im Durchschnitt eine Dauer von 03:18⁵⁶, allerdings bei großer Spannweite der Werte

⁵⁶ Angaben im Format mm:ss

(Minimum: 00:04, Maximum: 17:47, SD: 03:17). Den größten Zeitanteil der Aufgaben nimmt deren Bearbeitung ein (M: 02:49, Minimum: 00:03, Maximum: 15:34, SD: 02:50), die Aufgabenstellungen einen weitaus geringeren Anteil (M: 00:28, Minimum: 00:01, Maximum: 05:43, SD: 00:48). Im Anschluss daran ergibt die verbale Ausführlichkeit der Aufgabenstellungen einen mittleren Umfang von ca. 37 Wörtern, bei ebenfalls großer Spannweite (Minimum: 2, Maximum: 32, SD: 43,3). Der Großteil der Aufgaben wird in englischer Sprache eingeführt bzw. gestellt (64 %). Insgesamt werden von den Lehrkräften in der Formulierung der Aufgabenstellung kaum Hinweise integriert, die eine Hilfestellungsfunktion für die erfolgreiche Bewältigung der Aufgabe erfüllen: 38 % der Aufgabenstellungen enthalten keine entsprechenden Hinweise und 40 % lediglich einen Hinweis, welcher meistens in einer Wiederholung oder Paraphrasierung der Aufgabenstellung besteht. Der Großteil der Aufgaben (85 %) bezieht ein Unterrichtsmedium ein bzw. stützt sich auf mediengebundene Informationen oder Darstellungen. Mehrheitlich sind dies textförmige Informationen (56 %) oder Kombinationen aus Text und Abbildungen oder Text und tabellarisch bzw. diagrammartig dargestellten Informationen (11 %). Träger der Informationen ist bei 25 % der Aufgaben die Tafel, bei 34 % ein Lehrbuch, bei 9 % ein Arbeitsblatt und in 11 % der Fälle ein „anderes“ Gerät oder Material, welches meist ein Audioabspielgerät oder ein Anschauungsgegenstand ist. Der Umfang der in den Unterrichtsmedien abgebildeten Textinformationen, gemessen an der Wortanzahl, differiert zwischen null (bei reinen Abbildungen) und 432 Wörtern mit einem Mittelwert von ca. 70 Wörtern. Der Bezug zu einem bestimmten mediengebundenen Inhalt wird meistens über mehrere Aufgaben hinweg aufrecht erhalten, d.h. die Einführung eines neuen Mediums erfolgt etwa mit jeder vierten Aufgabe. Die Aufgabeninhalte lassen in annähernd der Hälfte der Fälle (47 %) einen „konstruierten“ Lebensweltbezug erkennen, was meistens einer Bezugnahme auf beispielhafte Geschichten mit fiktiven Charakteren einhergeht. Ein knappes Viertel der Aufgaben weist hingegen keinerlei Bezug zur Lebenswelt der Schüler auf (24 %), und 29 % der Aufgaben verknüpfen die Inhalte mit Aspekten der tatsächlichen Lebenswelt der Schüler. Die behandelten Inhalte repräsentieren zumeist einfach beobachtbare Aspekte des Alltags (57 %), wie bspw. Ereignisse oder Objekte, in geringerem Umfang auch leicht abstrakte Sachverhalte wie z.B. Gefühle und Meinungen (22 %) oder abstrakte Dinge wie Strukturen oder Prinzipien (21 %). Hinsichtlich der Anforderungen, die in der Aufgabenbearbeitung an die Schüler gestellt werden, ist ein starkes Gewicht auf den kognitiven Prozessen des Verstehens (36 %) und Anwendens

(30 %) bei geringeren Anteilen von Prozessen des Erinnerns (19 %) oder des Analysierens (11 %) und marginalen Anteilen von Prozessen des Evaluierens (4 %) und Gestaltens (1 %) festzustellen. Die zur Bearbeitung zu aktivierenden Wissensdimensionen liegen dementsprechend zu 48 % im Bereich der Prozeduren, daneben zu 31 % im Faktenwissen und zu 20 % im Wissen über Konzepte. Metakognitive Strategien bilden mit 1 % einen vernachlässigbaren Anteil. Fast die Hälfte der Aufgaben (47 %) erfordert die Verarbeitung, d.h. die Wiedergabe oder das Umorganisieren von vorgegebenen Informationen, während 34 % der Aufgaben die Eigenkonstruktion kleinerer Produkte (z.B. Antworten mit der Länge eines Wortes oder Satzes) und 19 % der Aufgaben ausgedehnte Produkte (z.B. Erzählungen, Berichte, Diskussionsbeiträge) erfordern. Die geforderten Produkte sind in den überwiegenden Fällen (80 %) alleine in sprachlicher Form, selten auch schriftlich (12 %), kaum aber in grafischer Form (1 %) zu erbringen. Die im Rahmen der Probekodierungen zusätzlich eingeführte Identifizierung von Aufgaben mit einer Reflexions- bzw. Auswertungsfunktion (s. Kap. 6.7.2) ergibt bei 8 % die Auswertung einer vorherigen Unterrichtsaufgabe und bei 6 % einer Hausaufgabe. Die Ausprägungen der Aufgabenmerkmale unterscheiden sich zwischen den Unterrichtsstunden fast durchgängig hoch signifikant (gemäß χ^2 -Unabhängigkeitstests und Varianzanalysen, s. Kap. 7.3.2). Ausnahmen bilden die Anzahl der Hilfestellungsmaßnahmen, deren enger Wertebereich offenbar keinen Spielraum für nennenswerte Wertedifferenzen zwischen den Unterrichtsstunden zulässt, und die verschiedenen Aspekte des Zeitumfangs. Eine Prüfung der wechselseitigen Zusammenhänge zwischen den Aufgabeneigenschaften zeigt überwiegend erwartbare Zusammenhänge zwischen Ausprägungen einzelner Merkmale wie z.B. zwischen dem Zeitumfang der Aufgabenstellungen und dem Umfang der verbalen Ausführungen, Ausprägungskombinationen des Medieneinsatzes (Abbildungen und Grafiken zumeist im Lehrbuch, reiner Text zumeist an der Tafel) oder den beiden der Lernzieltaxonomie von Anderson et al. (2001) entstammenden Variablen.

Neben der Charakterisierung der Einzelaufgaben wurden auf Basis der beschriebenen Ausprägungen Indikatoren gebildet, die auf unterschiedliche Weise die Variabilität des jeweiligen Aufgabenmerkmals auf Stundenebene abbilden (s. Kap. 7.3.3). Die Werte dieser drei Indikatoren (Variationsbreite, Gleichförmigkeit, Wechselhaftigkeit) lassen sich aufgrund mangelnder Vergleichbarkeit nicht einordnen, deuten aber auf

unterschiedlich ausgeprägte Variabilitäten der verschiedenen Aufgabenmerkmale hin. Die Begutachtung der Interkorrelationen zwischen den Indikatorenwerten zeigt zum einen, dass die Variabilitätsindikatoren je nach Aufgabenmerkmal zueinander unterschiedlich trennscharf sind. Zum anderen werden einige Zusammenhänge zwischen Aufgabenmerkmalen sichtbar, die auf Ebene der Einzelaufgaben nicht zu erkennen waren.

Ergebnisse zu den Beziehungen zwischen den Eigenschaften der Aufgaben und dem Beteiligungsverhalten der Schüler im Unterricht:

Die Beantwortung der Kernfrage dieser Studie, inwiefern Zusammenhänge zwischen dem Aufgabenangebot und dem Beteiligungsverhalten der Schüler im Unterricht festzustellen sind, geschah in zwei Schritten. Zuerst wurden Variationen der *on task*-Rate im Unterrichtsverlauf auf ihre Abhängigkeit von Ausprägungen der Merkmale der Einzelaufgaben hin geprüft. Im zweiten Schritt wurden Einflüsse der stundenspezifischen Variabilität der Aufgabenmerkmale auf das allgemeine *on task*-Niveau in der jeweiligen Unterrichtsstunde untersucht. Die Ergebnisse beider Schritte wurden dargestellt sowohl für die gesamte Schülerstichprobe als auch differenziert nach Schüleruntergruppen.

Für die Abbildung der Beziehungen zwischen den Ausprägungen der Aufgabenmerkmale und dem Schülerbeteiligungsverhalten auf Prozessebene wurde für jede Aufgabe innerhalb des öffentlichen Unterrichts die durchschnittliche gleichzeitige *on task*-Rate berechnet (s. Kap. 7.4.1). Eine Prüfung der Zusammenhänge zwischen den Aspekten des Zeitumfangs der Aufgaben und den *on task*-Raten zeigt, dass eine längere Dauer der Aufgabenbearbeitung über alle Schülergruppen hinweg leicht mit niedrigeren *on task*-Raten zusammenhängt ($r_s = -0,20$). Demgegenüber ist für die Dauer der Aufgabenstellung kein Zusammenhang festzustellen. Für alle weiteren Aufgabenmerkmale wurde die Beziehung zum Beteiligungsverhalten über einfaktorielle, univariate Varianzanalysen mit dem jeweiligen Aufgabenmerkmal als mehrstufigem Faktor und der durchschnittlichen *on task*-Rate im jeweiligen Aufgabenzeitraum als abhängige Variable geprüft. Die Ergebnisse (im Überblick: Tab. 22) lassen jedoch nur wenige, auf die Ausprägungen der Aufgabenmerkmale zurückzuführende Unterschiede in den zeitgleichen *on task*-Raten erkennen: Zunächst zeigt die Hälfte der

Aufgabenmerkmale keinerlei Effekte. Einige geben ein lediglich schwaches bzw. unklares Effektbild ab: Bei den Merkmalen „Darstellungsform des Unterrichtsmediums“ und „Abstraktheit des Aufgabeninhalts“ ist nur auf Ebene einer einzelnen Untergruppe der Schülerstichprobe ein Effekt festzustellen⁵⁷, und das Merkmal „Lebensweltbezug der Aufgabeninhalte“ zeigt signifikante Unterschiede für mehrere Untergruppen zugleich, diese ergeben allerdings in Kombination kein plausibles Bild⁵⁸. Daneben lassen andere Aufgabenmerkmale (Sprache der Aufgabeneinführung bzw. -präsentation, Bekanntheit des Mediums, Art des geforderten Produkts) Einflüsse auf das zeitgleiche Beteiligungsverhalten erkennen, wobei diese Ergebnisse jedoch von stark ungleichen Häufigkeitsverteilungen zwischen den Ausprägungsstufen betroffen und an diesen Stellen Verzerrungen nicht auszuschließen sind. Ein relativ stabiler Befund ist jedoch bei Betrachtung der Mediengeräte bzw. -materialien zu verzeichnen: Im überwiegenden Teil der Prüfungen fallen die *on task*-Raten unter Verwendung des Lehrbuchs niedriger aus als beim Einsatz „anderer“ Materialien bzw. Geräte, was im Wesentlichen Audioabspielgeräte und Anschauungsgegenstände umfasst. Dieser Effekt differenziert sich jedoch bei Berücksichtigung des familiären Sprachhintergrunds der Schüler aus und ist dann nur noch bei Schülern mit nicht-deutschsprachigem Hintergrund – dafür umso deutlicher – erkennbar. Insgesamt erweisen sich die Effekte auf Ebene der Einzelaufgaben aber als relativ schwach und zudem teilweise unsicher. Denn auch eine Analyse möglicher Wirkungen von Ausprägungskombinationen zwischen verschiedenen Aufgabenmerkmalen mittels paarweiser Prüfung von Interaktionseffekten für alle möglichen Kombinationen liefert keine zusätzlichen Erkenntnisse.

Ergiebiger stellen sich demgegenüber die Befunde zur Variabilität von Aufgabenmerkmalen auf Ebene der Unterrichtsstunden dar. Die Prüfung von Zusammenhängen zwischen den unterschiedlichen Indikatoren stundenspezifischer Variabilität der Aufgabenmerkmale (Variationsbreite, Gleichförmigkeit und Wechselhaftigkeit, vgl. Kap. 7.3.3) und den stundenweise gemittelten *on task*-Raten der

⁵⁷ (1) Schüler mit nicht-deutschsprachigem Hintergrund zeigen signifikant höhere *on task*-Raten beim Einsatz von Hörspielen als bei Verwendung einer Medienkombination. (2) Schüler im Leistungsmittelfeld zeigen signifikant höhere *on task*-Raten bei Aufgaben mit abstrakten Inhalten als bei konkreten Aufgabeninhalten.

⁵⁸ Bei Schülern mit nicht-deutschsprachigem Hintergrund, Mädchen und Schülern im Leistungsmittelfeld fallen die *on task*-Raten bei Aufgaben ohne Lebensweltbezug signifikant höher aus als bei Aufgaben mit konstruiertem Lebensweltbezug.

unterschiedlichen Untergruppen sowie der gesamten Schülerstichprobe (s. Kap. 7.4.2) zeigt eine Reihe schwacher bis mittlerer Korrelationen ($r_s = 0,20 - 0,30$; s. Tab. 23). Die Mehrzahl der Aufgabenmerkmale, die solch einen Zusammenhang auf Ebene der gesamten Schülerstichprobe aufweisen, sind jedoch bei Berücksichtigung der Personenuntergruppen von einer Ausdifferenzierung des allgemeinen Effekts betroffen⁵⁹. Demnach wirkt sich bei diesen Aufgabenmerkmalen das Ausmaß ihrer Ausprägungsvariabilität im Stundenverlauf unterschiedlich stark auf das Beteiligungsverhalten verschiedener Schülergruppen in einer Klasse aus. Lediglich bei den zur Bearbeitung erforderlichen kognitiven Prozessen und dem inhaltlichen Umfang der Medien lässt sich annähernd das gleiche Korrelationsmuster über alle Teilgruppen der Schüler hinweg feststellen, sodass bei diesen Merkmalen von einem allgemeinen, d.h. in ähnlicher Weise für alle Schüler geltenden Effekt gesprochen werden kann.

Die Unterschiede in der Stärke der Zusammenhänge zwischen den verschiedenen Schüleruntergruppen (s. Korrelationsdifferenzen, Tab. 24), die für eine Ausdifferenzierung der allgemeinen Effekte verantwortlich sind, treten allerdings nicht gleichermaßen in allen Personenkategorien auf. Eine Reihe von Aufgabenmerkmalen zeigt deutliche Korrelationsdifferenzen bei Unterscheidung der Schüler nach Leistungsgruppen⁶⁰ und größtenteils in Überschneidung dazu auch bei Unterscheidung nach dem Sprachhintergrund in der Familie⁶¹. Lediglich die Aspekte des Zeitumfangs der Aufgaben zeigen alleine zwischen den Leistungsgruppen Korrelationsdifferenzen, und dies in deutlicher Ausprägung. Im Gegensatz zu den Leistungsgruppen und den Kategorien des familiären Sprachhintergrunds sind zwischen den beiden Geschlechtern kaum nennenswerte Korrelationsunterschiede festzustellen.

⁵⁹ Dies sind: Sprache der Einführung und Präsentation, verbaler Umfang der Aufgabenstellung, Anzahl an Hilfestellungshinweisen in der Aufgabenstellung, Lebensweltbezug und Abstraktheit der Aufgabeninhalte, Wissensdimension in der Bearbeitung, Format und Art des geforderten Produkts, mittlerer Zeitumfang der Aufgabenstellungen und -bearbeitungen

⁶⁰ Dies sind: Sprache der Einführung und Präsentation, verbaler Umfang der Aufgabenstellung, Darstellungsform des Mediums, Abstraktheit der Aufgabeninhalte, Wissensdimension in der Bearbeitung, Format des geforderten Produkts, Zeitumfang von Aufgabenstellung und -bearbeitung

⁶¹ Dies sind: Sprache der Einführung und Präsentation, verbaler Umfang der Aufgabenstellung, Anzahl Hilfestellungshinweise bei Aufgabenstellung, Darstellungsform des Mediums, Mediengerät bzw. -material, Abstraktheit der Aufgabeninhalte, Wissensdimension in der Bearbeitung, Format und Art des geforderten Produkts

Im Hinblick auf die Abbildung unterschiedlicher Qualitäten der stundenspezifischen Variabilität durch die drei generierten Variabilitätsindikatoren lassen sich unterschiedliche Schwerpunkte in den Zusammenhangsmustern zwischen den übergeordneten Aufgabenmerkmalsbereichen erkennen (s. Tab. 23). So sind beim überwiegenden Teil der Aufgabenmerkmale zur Beschreibung der kognitiven Anforderungen der Aufgabenbearbeitung (kognitiver Prozess, zu aktivierende Wissensdimension und Format des geforderten Produkts) die größten nennenswerten Zusammenhänge im Bereich der beiden Indikatoren „Variationsbreite“ und „Wechselhaftigkeit“ zu konstatieren. Bei diesen Aufgabeneigenschaften sind demnach häufige Ausprägungswechsel bei einem möglichst großen Variantenreichtum hinsichtlich der abgedeckten Ausprägungen im Stundenverlauf von größerer Bedeutung für das Beteiligungsverhalten der Schüler als ein quantitativ gleichmäßiges Auftreten der verschiedenen Ausprägungsstufen, abgebildet durch den Indikator „Gleichförmigkeit“. Andere Aufgabenmerkmale lassen demgegenüber eine mindestens ebenso große Relevanz des Gleichförmigkeitsindikators erkennen (Sprache der Einführung und Präsentation, Mediengerät bzw. -material, Abstraktheit der Aufgabeninhalte).

Zusammenfassung der Ergebnisse in Kurzform:

Die dargestellten Ergebnisse lassen sich unter Bezugnahme auf die in Kap. 5.3 Fragestellungen folgendermaßen zusammenfassen:

1. Das beobachtete Schülerbeteiligungsverhalten zeigt mit einer durchschnittlichen *on task*-Rate von ca. 83 % ein vergleichbar hohes Niveau wie der Großteil der bisherigen Forschungsergebnisse zu diesem Konstrukt. (a) Vor dem Hintergrund einer leistungs- und herkunftsheterogenen Schülerstichprobe sind Unterschiede in der individuellen Verhaltenstendenz weder durch den Leistungsstand im Unterrichtsfach noch durch den familiären Sprachhintergrund zu erklären, wohl aber durch das Geschlecht. (b) Daneben ist ein noch größerer Einfluss der Zugehörigkeit zu einer bestimmten Schulklasse festzustellen. Zusammen mit den Ergebnissen zur zeitlichen Instabilität des auf Klassenebene aggregierten Beteiligungsverhaltens sowohl im Stundenverlauf als auch über die untersuchten Unterrichtsstunden hinweg wird dies als Hinweis auf potenzielle Effekte der Unterrichtsbedingungen gewertet.

2. Die Aufgaben in den untersuchten Unterrichtsstunden liefern, entsprechend der Beschreibung mithilfe des mehrdimensionalen Kategoriensystems, ein vielschichtiges Bild. (a) Bis auf einzelne Ausnahmen (z.B. Aspekte des Zeitumfangs der Aufgaben) unterscheiden sich die untersuchten Unterrichtsstunden hinsichtlich der Ausprägung der einzelnen Aufgabenmerkmale deutlich. (b) Die zusammenfassende Beschreibung der Variabilität des stundenspezifischen Aufgabenangebots mithilfe eigens generierter Indikatoren gibt ein zwischen den verschiedenen Aufgabenmerkmalen ebenfalls differierendes Bild ab.

3. Hinsichtlich der Erklärung von Variationen des Schülerbeteiligungsverhaltens im Unterricht durch die Eigenschaften der in den Unterrichtsprozess eingebrachten Aufgaben sind unterschiedliche Ergebnismuster in Abhängigkeit von der Betrachtungsebene zu verzeichnen. (a) Eine Analyse der Merkmale von Einzelaufgaben im Hinblick auf ihren Einfluss auf das zeitgleiche Beteiligungsverhalten der Schüler ergibt kaum aussagekräftige Befunde. Allenfalls die zeitliche Dauer der Aufgabenbearbeitung und die Art des Medieneinsatzes lassen nennenswerte Effekte erkennen. (b) Auf Ebene der zusammenfassenden Betrachtung der stundenspezifischen Variabilität des Aufgabenangebots hingegen sind für einige der untersuchten Aufgabenmerkmale mäßige bis mittlere Zusammenhänge zum Beteiligungsverhalten erkennbar. (c) Der Großteil dieser Effekte differenziert sich jedoch unter Berücksichtigung von Untergruppen der Schülerstichprobe aus, wobei diese Differenzen überwiegend bei Unterscheidung der Schüler nach Leistungsstand im Unterrichtsfach und familiärem Sprachhintergrund sichtbar werden, während zwischen den Geschlechtern kaum Unterschiede in den Ergebnismustern festzustellen sind.

Einordnung und Bewertung der zentralen Ergebnisse

Die Tatsache, dass sich aus dem Bereich der persönlichen Hintergrundmerkmale der Schüler ausgerechnet das Geschlecht, nicht aber der individuelle fachliche Leistungsstand oder der familiäre Sprachhintergrund als einflussreiche Größe für das Beteiligungsverhalten im Unterricht erwiesen hat, war sowohl mit Blick auf das Modell effektiven Unterrichts nach Slavin (1994), welches auf Seiten der so genannten *student inputs* eher als Fähigkeitsparameter zu verstehende Faktoren nennt, als auch die bisherige Forschungslage (s. Kap. 2.3.2) nicht zu erwarten. Begründungen für diesen

Effekt sind auch an anderen Stellen nicht zu finden, womit eine erste Forderung nach weiter gehenden Untersuchungen zu benennen wäre.

Im Hinblick auf die Beziehungen zwischen den Eigenschaften der Aufgaben und dem Beteiligungsverhalten der Schüler im Unterricht sind zunächst einige der Ergebnisse in Anknüpfung an Slavin (ebd.) als Wirkung einer Interessantheit der Unterrichtsinhalte und -aktivitäten und damit des so genannten unterrichtlichen Anreizwerts (*incentive*) zu verstehen. Dazu können etwa die Effekte im Bereich des Medieneinsatzes gezählt werden oder die positiven Einflüsse der Variabilität einer Reihe von Aufgabenmerkmalen im Laufe des Unterrichtsprozesses, im Sinne eines abwechslungsreichen und „spannenden“ Unterrichts (z.B. Variabilität der Bearbeitungsanforderungen oder der Abstraktheit der Inhalte). Darüber hinaus ist ein Teil der Befunde an der von Slavin (ebd.) benannten Schnittstelle zwischen dem Anreizwert des Unterrichts und der Qualität der Darstellung und Präsentation der Inhalte (*quality of instruction*) einzuordnen. Dies betrifft etwa die auf Stundenebene zu erkennenden, positiven Effekte eines höheren durchschnittlichen Zeitumfangs der Aufgabenstellungen und der Variabilität des verbalen Umfangs in den Ausführungen der Lehrkraft während jener Phasen (s. Kap. 7.4.2, Tab. 23) sowie auch hier die Art des Medieneinsatzes.

Eine Wirkung des in diesem Zusammenhang ebenfalls als bedeutsam erachteten Bezugs der Inhalte zur Lebenswelt der Schüler⁶² lässt sich an dieser Stelle hingegen kaum belegen: Neben dem Befund, dass eine höhere Variationsbreite hinsichtlich der hergestellten Lebensweltbezüge im Stundenverlauf leicht mit höheren *on task*-Raten korreliert (s. Kap. 7.4.2, Tab. 23), steht das auf Mikroprozessebene erzielte Ergebnis, welches eher sogar eine Tendenz zu höheren *on task*-Raten während Aufgaben ohne jeglichen Lebensweltbezug belegt (s. Kap. 7.4.1). Der in anderen Studien nachgewiesene, positive Effekt einer *real life significance* des Unterrichts (Lam, Wong, Yang & Liu, 2012) bzw. der *authentic instructional work* (Marks, 2000) auf das Beteiligungsverhalten der Schüler kann somit zumindest in dieser Studie nicht bestätigt werden. Der ausbleibende Effekt in der vorliegenden Stichprobe ist möglicherweise darauf zurückzuführen, dass in den inhaltlichen Darstellungen, d.h. den Beispielszenarien, welche in den

⁶² „relating topics to students' personal lives“ (Slavin, 1994, S. 147)

Lehrmaterialien oder den Äußerungen der Lehrkraft entworfen werden, zwar versucht wird, lebensweltliche Bezüge herzustellen, dies aus Sicht der Schüler jedoch misslingt, z.B. aufgrund mangelnder Authentizität der konstruierten Szenarien. In der Konsequenz würde dies aber auch ein Validitätsdefizit bei der Einschätzung dieses Aspekts im Rahmen der vorliegenden Aufgabenanalyse bedeuten.

Ein Teil der vorliegenden Befunde lässt sich überdies mit Blick auf die von den Schülern subjektiv wahrgenommenen Bewältigungschancen bei der Bearbeitung der Aufgabe erklären, einem bei Slavin (1994, S. 147) an der Schnittstelle zwischen dem Anreizwert des Unterrichts und der Angemessenheit der Lerngelegenheiten angesiedelten Faktor. Wie gezeigt, korrelieren eine hohe Vielfalt und ein hoher Grad an Abwechslung verschiedener Aspekte des kognitiven Anspruchs der Aufgabenbearbeitung in einer Unterrichtsstunde positiv und in mittlerer Höhe mit den *on task*-Raten (s. Kap. 7.4.2, Tab. 23). Eine größere Variation der Anforderungen hängt demnach tendenziell mit einer höheren durchschnittlichen Beteiligung unter den Schülern einer Klasse zusammen. Dies kann durchaus als Beleg des bei Slavin (ebd.) beschriebenen Effekts verstanden werden, da sich bei steigender Variation der Anforderungen im Unterrichtsverlauf aus Sicht des Einzelschülers auch die Wahrscheinlichkeit von Aufgaben mit einem für ihn angemessenen Anforderungsprofil erhöht. Und dies führe – so Slavin (ebd.) – wiederum zu einem Anstieg der individuellen Beteiligungsmotivation. Dieser Effekt einer „gesteigerten individuellen Angemessenheit“ mittels Herstellung einer Variabilität des Aufgabenangebots lässt sich allerdings auch mit Blick auf Aufgabenmerkmale erweitern, die sich nicht alleine auf die kognitiven Anforderungen der Bearbeitung beziehen und deren stundenspezifische Variabilität ebenfalls an verschiedenen Stellen nennenswerte Zusammenhänge zum Beteiligungsverhalten erkennen lässt. Demzufolge ist auch eine höhere Variation von Ausprägungen der Aufgabenpräsentation, des Medieneinsatzes oder der inhaltlichen Abstraktheit im Unterrichtsverlauf mit höheren durchschnittlichen *on task*-Raten verbunden (s. Kap. 7.4.2, Tab. 23). Die Ergebnisse zur Variabilität des Aufgabenangebots lassen sich somit auch in ihrer Gesamtheit in den Zusammenhang von Wirkungen der individuellen Angemessenheit von Lerngelegenheiten im Unterricht stellen⁶³, woraus sich Anschlüsse

⁶³ Der auf Prozess- sowie auf Stundenebene festzustellende negative Zusammenhang zwischen der zeitlichen Dauer der Aufgabenbearbeitung und dem Beteiligungsverhalten der Schüler steht nach Slavin

zu dem in der Unterrichtsforschung bekannten Konzept der Adaptivität von Unterricht, als „*optimale Nutzung von Lerngelegenheiten*“ (Hardy et al., 2011) unter der Bedingung heterogener Lernvoraussetzungen der Schüler, ergeben. Damit verbunden ist auch die Tatsache, dass die Ergebnisse zu den Beziehungen zwischen Aufgabeneigenschaften und Beteiligungsverhalten an vielen Stellen (in besonderer Weise aber auf Stundenebene) unterschiedlich stark ausgeprägte Effekte in Abhängigkeit vom fachlichen Leistungsstand oder dem Sprachhintergrund in der Familie zeigen. Darin findet sich eine Grundfigur wieder, welche bereits im Rahmen der so genannten ATI-Forschung (*Aptitude-Treatment-Interaction*, Corno & Snow, 1986) der 1970er und 1980er Jahre zu finden ist, in deren Rahmen Wechselwirkungen (*interactions*) zwischen individuellen Voraussetzungen der Schüler (*aptitudes*) und den Unterrichtsbedingungen (*treatment*) untersucht wurden. Allerdings lassen sich aus den vorliegenden Ergebnissen über die grundsätzliche Feststellung hinaus, dass bestimmte Personenkategorien unter verschiedenen Bedingungen des Aufgabenangebots unterschiedliche Tendenzen hinsichtlich ihres Beteiligungsverhaltens zeigen, kaum stabile Wechselwirkungsmuster benennen.

In der Zusammenschau der Ergebnisse bleibt abschließend festzuhalten, dass das Schülerbeteiligungsverhalten im Unterricht durchaus zu gewissen Teilen durch die Eigenschaften der von der Lehrkraft in den Unterrichtsprozess eingebrachten Aufgaben zu erklären ist. Dies jedoch mit der Einschränkung, dass der Hauptteil der in der vorliegenden Untersuchung gefundenen Einflüsse bei einer zusammenfassenden Betrachtung des Aufgabenangebots auf Ebene der Unterrichtsstunden, basierend auf einer Betrachtung verschiedener Aspekte der Variabilität der Aufgabeneigenschaften im Unterrichtsverlauf, sichtbar wurde. Die Herstellung von Beziehungen zwischen den Eigenschaften der einzelnen Aufgaben und der zeitgleichen Verhaltenssituation unter den Schülern der Klasse, sozusagen auf Mikroprozessebene des Unterrichts, führte hingegen kaum zu aussagekräftigen Ergebnissen.

(1994) für die Angemessenheit des Unterrichtstempos und ist am Übergang zwischen Angemessenheit des Unterrichtsangebots und Instruktionsqualität zu sehen.

8.2 Grenzen der Untersuchung

Die Ergebnisse sind vor dem Hintergrund einiger Einschränkungen hinsichtlich der forschungsmethodischen Anlage und Durchführung der Untersuchung zu betrachten. Zunächst sind die Befunde mit Blick auf die Eigenschaften der Stichprobe zu interpretieren. Die Ziehung der teilnehmenden Klassen erfolgte nicht nach systematischen bzw. repräsentativen Kriterien, sondern war in erster Linie durch die Teilnahmebereitschaft von Lehrkräften, Schulleitungen und Eltern bedingt (vgl. Kap. 6.3). Die vorliegende Stichprobe ist daher im engeren Sinne als Gelegenheitsstichprobe zu bezeichnen (Bortz & Döring, 2006) und lässt damit weder Aussagen auf Populationsebene zu, noch lassen sich die Befunde ohne Weiteres auf andere Kontexte, wie z.B. andere Schulstufen, Unterrichtsfächer oder gar Regionen⁶⁴, übertragen.

Mit der Zusammensetzung dieser Stichprobe sind methodische Herausforderungen verbunden, welche sich aus der geschachtelten (Schulformen, Klassen, Untergruppen innerhalb der Klassen und Individuen) und über mehrere Messzeitpunkte erstreckenden (Abfolge der Unterrichtsstunden und Verlauf des Unterrichts) Datenstruktur ergeben. Mit Blick auf ersteres zeigen die Ergebnisse, dass die Zugehörigkeit zur speziellen Schulklasse eine große Rolle für interindividuelle Unterschiede des Beteiligungsverhaltens spielt (s. Kap. 7.2.3), was im gegebenen Kontext als Hinweis auf Einflüsse der Unterrichtsbedingungen gewertet wurde. Da in den untersuchten Klassen über die Erhebungszeitpunkte hinweg jeweils die selbe Lehrkraft den Unterricht durchführte, ist es allerdings nicht auszuschließen, dass der Schulklasseneffekt durch Eigenschaften der Lehrpersonen begründet werden kann (z.B. Aspekte der professionellen Kompetenz, vgl. Kunter et al., 2011, oder Handlungsroutinen und Persönlichkeitsmerkmale, vgl. Bromme & Haag, 2008), welche an dieser Stelle aufgrund des Fehlens entsprechender Daten, aber auch der geringen Anzahl an Lehrkräften ($n = 6$) nicht berücksichtigt werden könnten. Bei einer größeren Anzahl an Elementen auf dieser Datenebene könnten entsprechende Effekte mithilfe mehrebenenanalytischer Verfahren (Ditton, 1998) adäquat mit einbezogen werden. Um die Unterrichtsprozesse angemessen beleuchten und ihren Verlaufscharakter berücksichtigen zu können, kam überdies der Zeitebene im Rahmen dieser Studie eine

⁶⁴ Zu den Spezifika der Stichprobe in Bezug auf die Region (Ausprägung des Sprachhintergrunds in der Familie, s. Kap. 7.1.1).

grundlegende Bedeutung zu (vgl. Kap. 3). Allerdings sind damit auswertungsmethodische Herausforderungen verbunden, bei denen die üblichen statistischen Verfahren ab einer höheren Anzahl an Messwiederholungen – wie sie hier mit Blick auf die Unterrichtsverläufe gegeben sind – an ihre Grenzen stoßen (Rietz & Rudinger, 2007, S. 642)⁶⁵. Mögliche Einflüsse von Reihenfolgeeffekten, episodenspezifischen Trends oder sich wiederholenden Verhaltenszyklen (vgl. Vannest & Parker, 2010) sind somit an dieser Stelle nicht auszuschließen, konnten aber nicht systematisch berücksichtigt werden.

Zudem ist zu bemerken, dass sich der Stichprobenumfang für verschiedene Teilkomponenten dieser Untersuchung in unterschiedlicher Weise als geeignet erwiesen hat: Während sich die Analyse des Beteiligungsverhaltens auf einen zufriedenstellenden Umfang an Beobachtungsdaten stützen konnte und eine ausreichende Anzahl an Schülern vorlag, um interindividuelle Unterschiede ausreichend berücksichtigen zu können, wäre im Hinblick auf die statistische Qualität der Auswertungen auf Stundenebene eine höhere Anzahl an Unterrichtsstunden wünschenswert gewesen, ebenso wie eine höhere Anzahl an Aufgaben, um die geringe quantitative Besetzung einzelner Ausprägungsstufen einiger Aufgabenmerkmale (s. Kap. 7.3.2) aufzuwerten.

Eine Limitierung der vorliegenden Ergebnisse – speziell der Befunde zur Beziehung zwischen Aufgabenangebot und Beteiligungsverhalten der Schüler (Kap. 7.4) – besteht darin, dass durch die wiederholte Durchführung von Signifikanztests an der selben Stichprobe ein Teil der gefundenen Effekte lediglich auf Zufall zurückzuführen sein könnte. Dieses in der Statistik als Alpha-Fehlerkumulierung bzw. -inflation bezeichnete Problem (Bortz, 2010; Rasch, Friese, Hofmann & Naumann, 2006) muss an dieser Stelle jedoch mit Blick auf die Ausrichtung der Studie gesehen werden: Vor dem Hintergrund des überwiegend explorativen Untersuchungsinteresses waren vielfältige Merkmalsbeziehungen zu prüfen, die sich als eine Reihe von separaten Hypothesen zwar einer übergeordneten Fragestellung subsumieren, auf Basis des gegenwärtigen wissenschaftlichen Kenntnisstandes aber nicht im Rahmen eines statistischen Gesamtmodells systematisch modellieren lassen. Die Studie ist vielmehr als eine Vorarbeit für Untersuchungen zu verstehen, die die systematische Prüfung eines solchen

⁶⁵ s. mangelnde Differenzierbarkeit zwischen Schwankungen der *on task*-Raten im Unterrichtsverlauf in Kap. 7.2.4

Gesamtmodells zum Ziel haben. Aus diesem Grund wären Maßnahmen der Alpha-Adjustierung in diesem Kontext nicht als zielführend anzusehen, da durch die dafür notwendige Reduktion der Alpha-Niveaus der Einzelprüfungen in den vorliegenden Daten möglicherweise manche Befunde – und damit Anschlussstellen für die weitere Forschung – unsichtbar würden (vgl. Stelzl, 2005). Wenn auch der Verzicht auf eine multivariate Modellierung der Auswertung in gleicher Weise durch den Untersuchungscharakter (und zusätzlich durch die Eigenschaften der erhobenen Daten) zu begründen ist (Bortz, 2012), ist dennoch zu betonen, dass in Folgeuntersuchungen durch den Aufbau einer geeigneten Stichprobe und eine Schärfung der Instrumente der Einsatz multivariater Analysen ermöglicht werden sollte, um mögliche wechselseitige Relationen zwischen den Aufgabenmerkmalen sichtbar machen zu können⁶⁶.

Über diese auswertungsmethodischen Erwägungen hinaus besteht anschließend an die methodischen Erfordernisse, die sich aus der Berücksichtigung der Zeitebene ergeben (s.o.), eine Lücke der vorliegenden Untersuchung in der fehlenden Berücksichtigung sozialer Interaktionen innerhalb des Unterrichtsgeschehens. Basierend auf der empirischen Herangehensweise einer Prozess-Prozess-Studie (Doyle, 1977) vor dem Hintergrund einer effektivitätsorientierten Sichtweise auf Unterricht (s. Kap. 2.1/2.2) folgt diese Untersuchung der Grundannahme einer kausalen Wirkrichtung vom Handeln der Lehrperson auf die (Re-)Aktionen der Schüler. Es muss jedoch angenommen werden, dass auch die Handlungen der Lehrkraft unter bestimmten Bedingungen eine Reaktion auf das Verhalten der Schüler darstellen: *„Auf der Mikroebene des täglichen Unterrichtsgeschehens dürfte ein bidirektionales Modell unterrichtlicher Auswirkungen (auf die Aufmerksamkeit) und deren Rückkopplung (auf den Unterricht) die Wirklichkeit am angemessensten beschreiben“* formulieren Helmke und Renkl (1993, S. 198, Unterstreichung ist im Original kursiv gedruckt) diesbezüglich treffend. Mit Blick darauf wäre es denkbar, dass die Lehrkräfte bestimmte Aufgabeneigenschaften in Abhängigkeit von ihrer Wahrnehmung der momentanen Beteiligungssituation in der Klasse spontan variieren (z.B. die zeitliche Dauer der Aufgabenbearbeitung, die Art und Weise der Formulierung der Aufgabenstellung, die Art des Medieneinsatzes oder auch Aspekte der kognitiven Anforderungen).

⁶⁶ Zu den Grenzen des vorliegenden Untersuchungs- und Auswertungsansatzes hinsichtlich der Identifizierung von Wechselwirkungen zwischen Aufgabenmerkmalen in ihren Einflüssen auf das Schülerverhalten, s. S. 169.

Eine Schwäche der vorliegenden Studie besteht im Übrigen darin, dass neben den Aufgaben kein weiteres Prozessmerkmal des Unterrichts in seinem Einfluss auf das Beteiligungsverhalten der Schüler untersucht wurde. Eine solche kontrastierende Einflussgröße würde eine Gegenüberstellung der Effekte und damit eine bessere Einschätzung der Wichtigkeit des Einflusses der Aufgaben erlauben. In Betracht zu ziehen wäre diesbezüglich etwa der Bereich der Klassenführung bzw. des *classroom management*, welcher seit jeher als ein wesentlicher Einflussfaktor für das Schülerbeteiligungsverhalten diskutiert wird (Anderson, 1984b; Brophy, 2006; Doyle, 1986; Ophardt & Thiel, 2008; s. auch Kap. 2.2/2.3). Weitere Schwierigkeiten in der Einordnung der gefundenen Effekte sind überdies dem explorativen Charakter dieser Studie, in Kombination mit der Fokussierung auf „natürlichen“ Unterricht, zuzuschreiben. Wie deutlich wurde, können Vergleiche zu Ergebnissen anderer Studien daher nur stellenweise und in eingeschränktem Umfang gezogen werden.

Neben den genannten Einschränkungen ist jedoch zu betonen, dass die vorliegenden Ergebnisse durch die Einnahme einer speziellen empirischen Perspektive umfassende Einblicke in die Zusammenhänge zwischen zwei zentralen Geschehensbestandteilen des Unterrichts gewähren, die vielfältige Anknüpfungspunkte bieten, um Anschlussforschung zu generieren und um mit Akteuren der Unterrichtspraxis ins Gespräch zu kommen. Aus Sicht der Forschung könnte dies als Chance zur konkreten Schärfung von Vorgehensweisen und methodischen Konzeptionen, aber auch zur allgemeinen Reflexion ihrer Zugangsweisen gesehen werden.

8.3 Ausblick

Im Anschluss an die Befunde dieser Studie ergeben sich einige Ansatzpunkte für die weitere Forschung. Die Tatsache, dass auf Ebene der zusammenfassenden Betrachtung der stundenspezifischen Aufgabenvariabilität deutlichere Effekte vorliegen als bei der Perspektive auf die Einflüsse der Einzelaufgaben im Unterrichtsverlauf, deutet auf Wirkungen der Aufgaben in ihrem gegenseitigen Zusammenhang bzw. ihrer „Vernetzung“ hin (Neubrand, 2002; vgl. auch den Begriff des „Aufgabensystems“ bei Bromme, Seeger & Steinbring, 1990). Demnach könnte in zukünftigen Untersuchungen die Art der wechselseitigen Verknüpfung von Aufgaben in ihrer Auftretenssequenz berücksichtigt werden, beispielsweise über die Betrachtung der Abfolge von Ausprägungen oder Ausprägungskombinationen von Aufgabeneigenschaften oder der Zusammenfassung von Aufgaben zu Episoden. Für konkrete Vorgehensweisen bei der

Untersuchung solcher Phänomene am Gegenstand des Mathematikunterrichts sei verwiesen auf die Arbeit von Neubrand (2002).

Die vorliegenden Ergebnisse zum Beteiligungsverhalten der Schüler (s. Kap. 7.2.2.) deuten auf ein in der Forschung bereits bekanntes (vgl. Kap. 2.3.3 und 2.4.1), methodisches Grundproblem bzw. genau genommen ein Validitätsproblem bei der beobachtungs-basierten Erfassung des Merkmals hin: Aus der Verhaltensbeobachtung wird nicht ersichtlich, ob der jeweilige Schüler in diesem Moment tatsächlich mental bzw. kognitiv der Lerngelegenheit zugewandt ist, oder ob er lediglich Aufmerksamkeit und Beteiligung „inszeniert“. In der vorliegenden Studie wurde diesem Aspekt dadurch Rechnung getragen, dass sich in dem Kodierverfahren ausschließlich auf zuverlässig beobachtbare Verhaltenskomponenten beschränkt wurde und zusätzlich eine so genannte „in dubio pro discipulo“-Kodierregel zur Anwendung kam (s. Kap. 6.6.3). In zukünftigen Untersuchungen könnte dieser Aspekt aber stärker Berücksichtigung finden, etwa durch flankierende Befragungen der Schüler zu ihrer Wahrnehmung und ihrem Erleben im Unterrichtsverlauf, oder – um im Unterschied zu älteren Studien (Peterson & Swing, 1982; Peterson, Swing, Stark & Waas, 1984) eine Bezugnahme auf bestimmte Zeitpunkte oder Episoden im Unterrichtsverlauf zu erlauben – durch den Einsatz prozesssensibler Erhebungsmethoden wie etwa der *experience sampling method* (Yair, 2000, s. auch Kap. 2.3.2).

In diesem Zusammenhang wären hinausgehend über Aspekte der tatsächlichen inneren Beteiligung der Schüler auch weitere Faktoren von Interesse, insbesondere die aktuelle Motivation der Schüler. In der Forschung zum *school* bzw. *student engagement* spielt Motivation seit jeher eine große Rolle, indem die Beteiligung (in ihren verschiedene Facetten) als Folge einer grundlegenden oder momentanen Motivation der Schüler angesehen wird (Fredricks, Blumenfeld & Paris, 2004; Reschly & Christenson, 2012; Stipek, 2002). Dem im Rahmen des QAIT-Modells effektiven Unterrichts formulierten Zusammenhang zwischen dem Anreizwert des Unterrichts und dem Beteiligungsverhalten der Schüler, welcher den zentralen theoretischen Bezugspunkt der vorliegenden Studie bildet, liegt ebenfalls die Annahme einer Motivationswirkung zugrunde (Slavin, 1994, S. 146f.). Angesichts dessen scheinen Erkenntnisse darüber von Interesse, wie bestimmte Unterrichtsbedingungen mit der individuellen Schülermotivation interagieren, besonders vor dem Hintergrund unterschiedlicher persönlicher Ausgangslagen der Schüler. Mit Blick auf die Aufgaben im Unterricht wäre

in diesem Zusammenhang die „*subjektive Aufgabenwahrnehmung*“ (Bromme, Seeger & Steinbring, 1990, S. 7) durch die Schüler von besonderer Bedeutung. Demnach ist anzunehmen, dass die Auffassung und Bewertung der gestellten Anforderungen interindividuell differiert, sodass ein und die selbe Aufgabe möglicherweise unterschiedliche Motivierungs- und Aktivierungspotenziale entfaltet. Die Erfassung der Aufgabenwahrnehmung durch die Schüler wäre beispielsweise mittels nachträglicher schriftlicher Befragungen der Schüler zu bestimmten Unterrichtsabschnitten bzw. Aufgabenbündelungen (Neubrand, 2002)⁶⁷ denkbar.

In Anknüpfung an den vorhergehenden Abschnitt eröffnet sich eine weitere Forschungsperspektive darin, die wechselseitige Beeinflussung zwischen der Gestaltung und Steuerung des Aufgabengeschehens durch die Lehrkraft einerseits und der Beteiligungsverhalten der Schüler andererseits sichtbar zu machen. Dazu könnten die Lehrkräfte im Nachgang des Unterrichts mittels *stimulated recall*-Interviews unter Einbezug von Videoausschnitten (Calderhead, 1981) zu spontanen Anpassungen der Aufgaben an die Gegebenheiten der jeweiligen Situation (z.B. eine erkennbar niedrige Beteiligung von Schülern mit bestimmten Eigenschaften) befragt werden.

Wie bereits erwähnt, besteht jedoch einer der wichtigsten Ansatzpunkte zukünftiger Forschung in Anknüpfung an die vorliegenden Ergebnisse darin, kontrastierend zur Art und Weise des Aufgabeneinsatzes die Rolle der Klassenführung in ihrem Einfluss auf das Schülerbeteiligungsverhalten in den Blick zu nehmen (vgl. Ende des vorherigen Abschnitts). Dies wäre einerseits mit der Herausforderung verbunden, beide Faktoren auf einer in methodischer Hinsicht analogen Spezifikationsebene abzubilden, birgt aber die Chance in sich, zum einen die (Wechsel-)Wirkungen auf das Beteiligungsverhalten der Schüler zu vergleichen und zum anderen, mögliche Überschneidungsbereiche zwischen diesen beiden Prozesskomponenten aufzuzeigen (s. Konzept des didaktisch-methodisches Klassenmanagements, Wellenreuther, 2007).

Abschließend bleibt zu erwähnen, dass mit Blick auf den explorativen Charakter dieser Untersuchung grundsätzlich weitere Studien wünschenswert sind, die zur Vervollständigung des in den Ergebnissen aufgezeigten Bildes beitragen können.

⁶⁷ s. auch weiter oben in diesem Abschnitt

Literaturverzeichnis

- Adams, R. S. & Biddle, B. J. (1970). *Realities of teaching*. New York: Holt Rinehart Winston.
- Alderson, J. C., Figueras, N., Kuijper, H., Nold, G., Takala, S. & Tardieu, C. (2004). *The development of specifications for item development and classification within The Common European Framework of Reference for Languages: Learning, Teaching, Assessment: Reading and Listening: Final report of The Dutch CEF Construct Project*.
Online verfügbar unter
http://eprints.lancs.ac.uk/44/1/final_report.pdf?origin=publication_detail (Letzter Abruf am 23.11.2014)
- Alderson, J. C., Figueras, N., Kuijper, H., Nold, G., Takala, S. & Tardieu, C. (2006). Analysing Tests of Reading and Listening in Relation to the Common European Framework of Reference: The Experience of The Dutch CEFR Construct Project. *Language Assessment Quarterly*, 3(1), 3-30. doi:10.1207/s15434311laq0301_2
- Anderson, L. W. (1984a). Attention, Tasks and Time. In L. W. Anderson (Hrsg.), *Time and School Learning: Theory, Research, and Practice* (S. 46-68). New York: St. Martin's Press.
- Anderson, L. W. (1984b). Instruction and Time on Task: A Review. In L. W. Anderson (Hrsg.), *Time and School Learning: Theory, Research, and Practice* (S. 143-163). New York: St. Martin's Press.
- Anderson, L. W. (1984c). An Introduction to Time and School Learning. In L. W. Anderson (Hrsg.), *Time and School Learning: Theory, Research, and Practice* (S. 1-12). New York: St. Martin's Press.
- Anderson, L. W. (2004). *Increasing teacher effectiveness* (2nd. Hrsg.) Paris: UNESCO International Institute for Educational Planning.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., ... Wittrock, M. C. (Hrsg.). (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Addison Wesley Longman.

- Appel, J. & Rauin, U. (2015). Methoden videogestützter Beobachtungsverfahren in der Lehr-Lern-Forschung. In D. Elsner & B. Viebrock (Hrsg.), *Triangulation in der Fremdsprachenforschung* (S. 59-79). Frankfurt am Main: Peter Lang Edition.
- Aufschnaiter, S. v. & Welzel, M. (2001). Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen: Eine Einführung. In S. v. Aufschnaiter & M. Welzel (Hrsg.), *Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen. Aktuelle Methoden empirischer pädagogischer Forschung* (S. 7-15). Münster: Waxmann.
- Autorengruppe Bildungsberichterstattung (2010): *Bildung in Deutschland 2010. Ein indikatorengestützter Bericht mit einer Analyse zu Perspektiven des Bildungswesens im demografischen Wandel*. Bielefeld: Bertelsmann. Online verfügbar unter www.bildungsbericht.de/daten2010/bb_2010.pdf (Letzter Abruf am 31.12.2014)
- Bachman, L. F. & Palmer, A. S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: University Press.
- Bachman, L. F. & Palmer, A. S. (2010). *Language Assessment in Practice*. Oxford: University Press.
- Bader, U. & Schaer, U. (2005). *Evaluation Englisch in den 6. Klassen Appenzell Innerrhoden 2005*. Fachhochschule Nordwestschweiz.
- Bakeman, R. & Quera, V. (2011). *Sequential Analysis and Observational Methods for the Behavioral Sciences*. Cambridge: University Press.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., ... Weiß, M (o.J.). *Erfassung fächerübergreifender Problemlösekompetenzen in PISA*. OECD PISA Deutschland. Online verfügbar unter <https://www.mpib-berlin.mpg.de/Pisa/Problemloesen.pdf> (Letzter Abruf am 24.11.2014)
- Beckerman, T. M. & Good, T. L. (1978). Time on Task: A Naturalistic Study in Sixth-Grade Classrooms. *The Elementary School Journal*, 78(3), 192-201.
- Berliner, D. C. (1990). What's all the fuss about instructional time? In M. Ben-Peretz & R. Bromme (Hrsg.), *The nature of time in schools: Theoretical concepts, practitioner perceptions* (S. 3-35). New York: Teachers College Press.
- Bessoth, R. (2010). Im Schatten der Lernzeit. *Pädagogische Führung*, 21(3), 101-105.

- Biddle, B. J. (1967). Methods and concepts in classroom research. *Review of Educational Research*, 37(3), 337-357.
- Blatchford, P., Bassett, P. & Brown, P. (2005). Teachers' and Pupils' Behavior in Large and Small Classes: A Systematic Observation Study of Pupils Aged 10 and 11 Years. *Journal of Educational Psychology*, 97(3), 454-467. doi:10.1037/0022-0663.97.3.454
- Blatchford, P., Bassett, P. & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction*, 21(6), 715-730. doi:10.1016/j.learninstruc.2011.04.001
- Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment*, 1(2), 1-12.
- Bloom, B. S. (1974). Time and learning. *American Psychologist*, 29(9), 682-688.
- Blömeke, S. & Müller, C. (2008). Zum Zusammenhang von Allgemeiner Didaktik und Lehr-Lernforschung im Unterrichtsgeschehen. *Zeitschrift für Erziehungswissenschaft, Sonderheft 9*, 239-258.
- Blömeke, S., Risse, J., Müller, C., Eichler, D. & Schulz, W. (2006). Analyse der Qualität von Aufgaben aus didaktischer und fachlicher Sicht. *Unterrichtswissenschaft*, 34(4), 330-357.
- Bohl, T., Kleinknecht, M., Batzel, A. & Richey, P. (2012). *Aufgabenkultur in der Schule*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Böhme, J. & Kramer, R.-T. (Hrsg.).(2001). *Partizipation in der Schule. Theoretische Perspektiven und empirische Analysen*. Opladen: Leske + Budrich.
- Böhm-Kasper, O. & Weishaupt, H. (2008). Quantitative Ansätze und Methoden in der Schulforschung. In W. Helsper & J. Böhme, (Hrsg.) *Handbuch der Schulforschung* (2. Aufl.)(S. 91-123). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Borg, W. R. (1980). Time and School Learning. In C. Denham & A. Lieberman (Hrsg.), *Time to Learn* (S. 33-72). Washington, D.C.: Department of Education, National Institute of Education.
- Borich, G. D. (2010). *Effective teaching methods: Research based practice* (7th Hrsg.). New York: Allyn & Bacon.

- Bortz, J. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Heidelberg: Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation* (4. Aufl.). Heidelberg: Springer.
- Breidenstein, G. (2006). *Teilnahme am Unterricht: Ethnographische Studien zum Schülerjob*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bremerich-Vos, A. (2008). Benjamin S. Bloom (und andere) revisited. In A. Bremerich-Vos, D. Granzer & O. Köller, *Lernstandsbestimmungen im Fach Deutsch* (S. 29-49). Weinheim: Beltz.
- Bromme, R. & Haag, L. (2008): Forschung zur Lehrerpersönlichkeit. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (2. Aufl.) (S. 803-820). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bromme, R., Seeger, F. & Steinbring, H. (1990). Aufgaben, Fehler und Aufgabensysteme. In R. Bromme, F. Seeger & H. Steinbring (Hrsg.), *Aufgaben als Anforderungen an Lehrer und Schüler* (S. 1-30). Köln: Aulis Verlag Deubner.
- Brophy, J. E. (2006). History of Research on Classroom Management. In C. M. Evertson & C. S. Weinstein (Hrsg.), *Handbook of Classroom Management. Research, Practice, and Contemporary Issues* (S. 17-43). Mahwah, N.J.: Lawrence Erlbaum Publishers.
- Brophy, J. E. & Evertson, C. M. (1980). *Lernen durch Unterricht*. Bochum: Verlag Ferdinand Kamp.
- Buff, A., Reusser, K. & Pauli, C. (2010). Selbstvertrauen ist wichtig, aber nicht ausreichend. Die Bedeutung von Unterricht, Selbstvertrauen und Qualität der Lernmotivation für Engagement und Leistung im Fach Mathematik. In K. Reusser, C. Pauli & M. Waldis (Hrsg.), *Unterrichtsgestaltung und Unterrichtsqualität. Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht* (S. 279-308). Münster: Waxmann.
- Bühner, M. & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson Education.
- Bundesministerium für Bildung und Forschung [BMBF](Hrsg.)(2001). *TIMSS - Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente*. Bonn.

- Burns, M. K. & Dean, V. J. (2005). Effect of acquisition rates on off-task behavior with children identified as having learning disabilities. *Learning Disability Quarterly*, 28(4), 273-281.
- Büttner, G., & Schmidt-Atzert, L. (2004). Diagnostische Verfahren zur Erfassung von Aufmerksamkeit und Konzentration. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 23-62). Göttingen: Hogrefe.
- Calderhead, J. (1981). Stimulated recall: a method for research on teaching. *British Journal of Educational Psychology*, 51(2), 211-217.
- Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, 64(8), 723-733.
- Carroll, J. B. (1984). The Model of School Learning: Progress of an Idea. In L. W. Anderson (Hrsg.), *Time and School Learning: Theory, Research, and Practice* (S. 15-45). New York: St. Martin's Press.
- Caruso, M. (2011). Lernbezogene Menschenhaltung. (Schul)Unterricht als Kommunikationsform. In W. Meseth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 24-36). Bad Heilbrunn: Klinkhardt.
- Christenson, S. L., Reschly, A. L. & Wylie, C. (Hrsg.). (2012). *Handbook of Research on Student Engagement*. New York: Springer.
- Clausen, M. (2002). *Qualität von Unterricht – Eine Frage der Perspektive?* Münster: Waxmann.
- Clausen, M., Reusser, K. & Klieme, E. (2003). Unterrichtsqualität auf der Basis hochinferenter Unterrichtsbeurteilungen. Ein Vergleich zwischen Deutschland und der deutschsprachigen Schweiz. *Unterrichtswissenschaft*, 31(2), 122-141.
- Corno, L. & Snow, R. E. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Hrsg.), *Handbook of Research on Teaching* (S. 605-629). New York: Macmillan.
- Creemers, B. P. M. & Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness*. New York: Routledge.
- De Sa Maini, R. (2011). *Teacher Training in a Proactive Approach to Classroom Behaviour Management: Teacher and Student Outcomes* (Dissertation). University of Toronto.

- DESI-Konsortium (Hrsg.).(2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- Dinkelaker, J. (2011). Aufmerksamkeit. In J. Kade, W. Helsper, C. Lüders, B. Egloff, F. O. Radtke, & W. Thole, *Pädagogisches Wissen* (S. 175-182). Stuttgart: Kohlhammer.
- Dinkelaker, J. & Herrle, M. (2009). *Erziehungswissenschaftliche Videographie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dishaw, M. M. (1978). *Beginning Teacher Evaluation Study Supplement: Phase III-B Instrumentation Package*. San Francisco: Far West Laboratory for Educational Research and Development.
- Ditton, H. (1998). *Mehrebenenanalyse: Grundlagen und Anwendungen des hierarchisch linearen Modells*. Weinheim: Juventa.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. *Zeitschrift für Pädagogik*, 41. Beiheft, 73-92.
- Ditton, H. (2002). Unterrichtsqualität - Konzeptionen, methodische Überlegungen und Perspektiven. *Unterrichtswissenschaft*, 30(3), 197-212.
- Doyle, W. (1977). Paradigms for research on teacher effectiveness. *Review of Research in Education*, 5, 163-198.
- Doyle, W. (1979). Classroom tasks and students' abilities. In P. Peterson & H. Walberg (Hrsg.), *Research on Teaching: Concepts, Findings, Implications* (S. 183-209). Berkeley: McCutchen Publishing.
- Doyle, W. (1983). Academic work. *Review of Educational Research*, 53(2), 159-199.
- Doyle, W. (1986). Classroom management and organization. In M. C. Wittrock (Hrsg.), *Handbook of Research on Teaching* (S. 392-432). New York: Macmillan.
- Drollinger-Vetter, B. (2006). Kognitiver Anspruchsgehalt der Aufgabenstellungen. In Klieme, E., Pauli, C. & Reusser, K. (Hrsg.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie "Unterrichtsqualität, Lernverhalten und mathematisches Verständnis"* (Teil 3: Hugener, I., Pauli, C. & Reusser, K.: Videoanalysen)(S. 148-164). Frankfurt am Main: GFPP.

- Dubberke, T. & Harks, B. (2008). Zur curricularen Validität der DESI-Aufgaben: Ergebnisse eines Expertenratings. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 26-33). Weinheim: Beltz.
- Eickhorst, A. (2011). Das Unterrichtsverständnis der empirischen Lehr-Lern-Forschung. In W. Meseth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 50-66). Bad Heilbrunn: Klinkhardt.
- Emmer, E. T., Sanford, J. P., Clements, B. S. & Martin, J. (1982). *Improving Classroom Management and Organization in Junior High Schools: An Experimental Investigation*. Austin: University of Texas at Austin.
- Europarat (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Fend, H. (1998). *Qualität im Bildungswesen*. Weinheim: Juventa.
- Finn, J. D., Pannozzo, G. M. & Voelkl, K. E. (1995). Disruptive and inattentive-withdrawn behavior and achievement among fourth graders. *Elementary School Journal*, 95(5), 421-434.
- Finn, J. D., & Zimmer, K. S. (2012). Student Engagement: What Is It? Why Does It Matter? In S. L. Christenson, A. L. Reschly & C. Wylie (Hrsg.), *Handbook of Research on Student Engagement* (S. 97-131). New York: Springer.
- Fisher, C. W., Berliner, D. C., Filby, N. N. & Marliave, R. C. (1980). Teaching behaviors, academic learning time, and student achievement: An overview. In C. Denham & A. Lieberman (Hrsg.), *Time to Learn* (S. 23-39). Washington, D.C.: Department of Education, National Institute of Education.
- Fraser, B. J., Walberg, H. J., Welch, W. W. & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research*, 11, 145-252.
- Fredrick, W. C. & Walberg, H. J. (1980). Learning as a function of time. *Journal of Educational Research*, 73, 183-194.
- Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, 74(1), 59-109. doi:10.3102/00346543074001059

- Gagné, R. M. (1980). *Die Bedingungen des menschlichen Lernens*. Hannover: Schroedel.
- Gettinger, M. & Seibert, J. K. (2002). Best practices in increasing academic learning time. *Best Practices in School Psychology, IV(1)*, 773-787.
- Gickling, E. E. & Armstrong, D. L. (1978). Levels of Instructional Difficulty as Related to On-Task Behavior, Task Completion, and Comprehension. *Journal of Learning Disabilities, 11(9)*, 559-566. doi:10.1177/002221947801100905
- Götz, T., Lohrmann, K., Ganser, B. & Haag, L. (2005). Einsatz von Unterrichtsmethoden: Konstanz oder Wandel? *Empirische Pädagogik, 19(4)*, 342-360.
- Gruschka, A. (2011). Der empirische Blick auf das Unterrichten als pädagogischer Prozess. In W. Meseth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 130-144). Bad Heilbrunn: Klinkhardt.
- Hage, K., Bischoff, H., Dichanz, H., Eubel, K.-D., Oehlschläger, H.-J. & Schwittmann, D. (1985). *Das Methodenrepertoire der Lehrer. Eine Untersuchung zum Schulalltag der Sekundarstufe I*. Opladen: Leske+ Budrich.
- Haertel, G. D., Walberg, H. J. & Weinstein, T. (1983). Psychological models of educational performance: A theoretical synthesis of constructs. *Review of Educational Research, 53(1)*, 75-91.
- Hardy, I., Hertel, S., Kunter, M., Klieme, E., Warwas, J., Büttner, G. & Lühken, A. (2011). Adaptive Lerngelegenheiten in der Grundschule: Merkmale, methodisch-didaktische Schwerpunktsetzungen und erforderliche Lehrerkompetenzen. *Zeitschrift für Pädagogik, 57(6)*, 819-832.
- Hartig, J. (2007). Skalierung und Definition von Kompetenzniveaus. In B. Beck & E. Klieme (Hrsg.), *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International)*(S. 83-99). Weinheim: Beltz.
- Hasselhorn, M. & Gold, A. (2013). *Pädagogische Psychologie. Erfolgreiches Lernen und Lehren* (3. Aufl.). Stuttgart: Kohlhammer.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses on achievement*. London: Routledge.
- Heinze, T. (1980). *Schülertaktiken*. München: Urban & Schwarzenberg.

- Helmke, A. (1988). *Das Münchener Aufmerksamkeitsinventar (MAI): Manual für die Beobachtung des Aufmerksamkeitsverhaltens von Grundschulern während des Unterrichts*. München: Max-Planck-Institut für psychologische Forschung.
- Helmke, A. (1997). Individuelle Bedingungsfaktoren der Schulleistung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter*. München: Psychologie Verlags Union.
- Helmke, A. (2003). *Unterrichtsqualität. Erfassen, Bewerten, Verbessern*. Seelze: Kallmeyer.
- Helmke, A. (2007). Aktive Lernzeit optimieren – Was wissen wir über effiziente Klassenführung? *Pädagogik*, 59(5), 44-49.
- Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität (4. Aufl.)*. Seelze-Velber: Klett/Kallmeyer.
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 345-363). Weinheim: Beltz.
- Helmke, A., & Renkl, A. (1992). Das Münchener Aufmerksamkeitsinventar (MAI): Ein Instrument zur systematischen Verhaltensbeobachtung der Schüleraufmerksamkeit im Unterricht. *Diagnostica*, 38, 130-141.
- Helmke, A., & Renkl, A. (1993). Unaufmerksamkeit in Grundschulklassen: Problem der Klasse oder des Lehrers? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, 185-205.
- Helmke, A., Schneider, W. & Weinert, F. E. (1986). Quality of instruction and classroom learning outcomes: The German contribution to the IEA classroom environment study. *Teaching and Teacher Education*, 2(1), 1-18.
- Helsper, W. & Böhme, J. (2008). Einleitung in das Handbuch der Schulforschung. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung (2. Aufl.)* (S. 11-32). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Helsper, W. & Keuffer, J. (1995): Unterricht. In H. H. Krüger & W. Helsper (Hrsg.), *Einführung in Grundbegriffe und Grundfragen der Erziehungswissenschaft* (S. 81-91). Opladen: Leske + Budrich.

- Hesse, H.-G., Göbel, K. & Hartig, J. (2008). Sprachliche Kompetenzen von mehrsprachigen Jugendlichen und Jugendlichen nicht-deutscher Erstsprache. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 208-230). Weinheim: Beltz.
- Hoge, R.D. (1985). The validity of direct observation measures of pupil classroom behavior. *Review of Educational Research*, 55, 469-483.
- Houtveen, A. A. M., Booij, N., de Jong, R. & van de Grift, W. J. C. M. (1999). Adaptive Instruction and Pupil Achievement. *School Effectiveness and School Improvement*, 10(2), 172-192. doi:10.1076/sesi.10.2.172.3508
- Hugener, I. (2006). Sozialformen und Lektionsdauer. In Klieme, E., Pauli, C. & Reusser, K. (Hrsg.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie "Unterrichtsqualität, Lernverhalten und mathematisches Verständnis"* (Teil 3: Hugener, I., Pauli, C. & Reusser, K.: Videoanalysen)(S. 55-61). Frankfurt am Main: GFPP.
- Jackson, P. W. (1966). *The way teaching is: Report of the seminar on teaching*. Washington D.C.: National Education Association (U.S.).
- Jackson, P. W. (1990). *Life in Classrooms* (Neuaufgabe der Originalarbeit von 1968). New York: Teachers College Press.
- Jacobs, J., Garnier, H., Gallimore, R., Hollingsworth, H., Givvin, K. B., Rust, ... Gonzales, P. (2003). *Third international mathematics and science study 1999 video study technical report: Volume 1: Mathematics*. Washington: National Center for Education Statistics, Institute of Education Statistics, U. S. Department of Education.
- Janik, T. & Seidel, T. (Hrsg.).(2009). *The Power of Video Studies in Investigating Teaching and Learning in the Classroom*. Münster: Waxmann.
- Janssen, J. & Laatz, W. (2010). *Statistische Datenanalyse mit SPSS* (7. Aufl.). Heidelberg: Springer.
- Jatzwauk, P., Rumann, S. & Sandmann, A. (2008). Der Einfluss des Aufgabeneinsatzes im Biologieunterricht auf die Lernleistung der Schüler – Ergebnisse einer Videostudie. *Zeitschrift für Didaktik der Naturwissenschaften*, 14, 263-283.
- Jordan, A., Krauss, S., Löwen, K., Blum, W., Neubrand, M., Brunner, M. ... Baumert, J. (2008). Aufgaben im COACTIV-Projekt: Zeugnisse des kognitiven

- Aktivierungspotentials im deutschen Mathematikunterricht. *Journal für Mathematik-Didaktik*, 29(2), 83-107.
- Jordan, A., Ross, N., Krauss, S., Baumert, J., Blum, W., Neubrand, M. ... Kunter, M. (2006). *Klassifikationsschema für Mathematikaufgaben: Dokumentation der Aufgabenklassifikation im COACTIV-Projekt*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Karweit, N. (1989). Time and Learning: A Review. In R. E. Slavin (Hrsg.), *School And Classroom Organization* (S. 69-95). Hillsdale: Lawrence Erlbaum Associates.
- Karweit, N. & Slavin, R. E. (1981). Measurement and modeling choices in studies of time and learning. *American Educational Research Journal*, 18(2), 157-171.
- Kiel, E. (2010). Unterrichtsforschung. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (S. 773-791). Wiesbaden: VS Verlag.
- Kiper, H. (2010). Der systematische Ort von Aufgaben in Theorien des Unterrichts. In H. Kiper, W. Meints, S. Peters, S. Schlump & S. Schmit (Hrsg.), *Lernaufgaben und Lernmaterialien im kompetenzorientierten Unterricht* (S. 44-59). Stuttgart: Kohlhammer.
- Kleinknecht, M. (2010). *Aufgabenkultur im Unterricht - Eine empirisch-didaktische Video- und Interviewstudie an Hauptschulen*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Kleinknecht, M., Bohl, T., Maier, U., & Metz, K. (2013). Aufgaben analysieren: Eine allgemeindidaktische und fachdidaktische Herausforderung. In M. Kleinknecht, T. Bohl, U. Maier & K. Metz (Hrsg.), *Lern- und Leistungsaufgaben im Unterricht. Fächerübergreifende Kriterien zur Auswahl und Analyse* (S. 207-220). Bad Heilbrunn: Julius Klinkhardt.
- Kleinknecht, M., Maier, U., Metz, K. & Bohl, T. (2011). Analyse des kognitiven Aufgabenpotenzials. *Unterrichtswissenschaft*, 39(4), 328-345.
- Klieme, E. (2006). Empirische Unterrichtsforschung: Aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. *Zeitschrift für Pädagogik*, 52(6), 765-773.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. In M. Prenzel & L. Allolio-Näcke (Hrsg.),

- Untersuchungen zur Bildungsqualität von Schule, Abschlussbericht des DFG-Schwerpunktprogramms* (S. 127-146). Münster: Waxmann.
- Klieme, E., Pauli, C. & Reusser, K. (2006). (Hrsg.). *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie "Unterrichtsqualität, Lernverhalten und mathematisches Verständnis"*. Frankfurt am Main: GFPP.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222-237.
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In BMBF (Hrsg.), *TIMSS - Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43-57). Bonn.
- Kötters, C., Schmidt, R., & Ziegler, C. (2001). Partizipation im Unterricht - Zur Differenz von Erfahrung und Ideal partizipativer Verhältnisse im Unterricht und deren Verarbeitung. In J. Böhme & R. T. Kramer (Hrsg.), *Partizipation in der Schule* (S. 93-122). Opladen: Leske + Budrich.
- Kreuzberger, N. (2002). *Aufmerksamkeitsverhalten in offener Lernsituation in der Sekundarstufe*. Frankfurt am Main: Hänsel-Hohenhausen.
- Kron, F. W. (2008). *Grundwissen Didaktik* (5. Aufl.). München: Reinhardt.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Lam, S.-F., Wong, B., Yang, H. & Liu, Y. (2012). Understanding Student Engagement with a Contextual Model. In S. L. Christenson, A. L. Reschly & C. Wylie (Hrsg.), *Handbook of Research on Student Engagement* (S. 403-420). New York: Springer.
- Leisen, J. (2010). Lernaufgaben als Lernumgebung zur Steuerung von Lernprozessen. In H. Kiper, W. Meints, S. Peters, S. Schlump & S. Schmit (Hrsg.), *Lernaufgaben und Lernmaterialien im kompetenzorientierten Unterricht* (S. 60-67). Stuttgart: Kohlhammer.
- Leonhart, R. (2004). *Lehrbuch Statistik. Einstieg und Vertiefung*. Bern: Verlag Hans

Huber.

- Leucht, M., Harsch, C., Pant, H. A. & Köller, O. (2012). Steuerung zukünftiger Aufgabenentwicklung durch Vorhersage der Schwierigkeiten eines Tests für die erste Fremdsprache Englisch durch Dutch Grid Merkmale. *Diagnostica*, 58(1), 31-44. doi:10.1026/0012-1924/a000063
- Lipowsky, F. (2009). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 73-102). Berlin: Springer.
- Lipowsky, F., Pauli, C. & Rakoczy, K. (2008). Schülerbeteiligung und Unterrichtsqualität. In M. Gläser-Zikuda & J. Seifried (Hrsg.), *Lehrerexpertise – Analyse und Bedeutung unterrichtlichen Handelns* (S. 67-90). Münster: Waxmann.
- Lotz, M., Gabriel, K. & Lipowsky, F. (2013). Niedrig und hoch inferente Verfahren der Unterrichtsbeobachtung: Analysen zu deren gegenseitiger Validierung. *Zeitschrift für Pädagogik*, 59(3), 357-380.
- Lotz, M., Lipowsky, F. & Faust, G. (Hrsg.).(2013). *Dokumentation der Erhebungsinstrumente des Projekts „Persönlichkeits- und Lernentwicklung von Grundschulkindern“ (PERLE) - Teil 3. Technischer Bericht zu den PERLE-Videoanalysen*. Frankfurt am Main: GFFP.
- Lüders, M. (1995). *Zeit, Subjektivität und Bildung. Die Bedeutung des Zeitbegriffs für die Pädagogik*. Weinheim: Deutscher Studien Verlag.
- Lüders, M. (2003). *Unterricht als Sprachspiel. Eine systematische und empirische Studie zum Unterrichtsbegriff und zur Unterrichtssprache*. Bad Heilbrunn: Julius Klinkhardt.
- Lüders, M. (2011). Die Sprachspieltheorie des Unterrichts. In W. Meseth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 175-188). Bad Heilbrunn: Klinkhardt.
- Lüders, M. (2011). Unterricht. In U. Sandfuchs, H. Kemnitz, K.-P. Horn & W. Marotzki, W. (Hrsg.), *Lexikon Erziehungswissenschaft* (S. 342). Bad Heilbrunn: Klinkhardt.
- Lüders, M. (2012). Der Unterrichtsbegriff in pädagogischen Nachschlagewerken. *Zeitschrift für Pädagogik*, 58(1), 109-129.
- Lüders, M. & Rauin, U. (2008). Unterrichts- und Lehr-Lern-Forschung. In W. Helsper & J.

- Böhme (Hrsg.), *Handbuch der Schulforschung* (2. Auflage)(S. 717-745). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Maier, U., Bohl, T., Kleinknecht, M. & Metz, K. (2013). Allgemeindidaktische Kategorien für die Analyse von Aufgaben. In M. Kleinknecht, T. Bohl, U. Maier & K. Metz (Hrsg.), *Lern- und Leistungsaufgaben im Unterricht. Fächerübergreifende Kriterien zur Auswahl und Analyse* (S. 9-46). Bad Heilbrunn: Julius Klinkhardt.
- Maier, U., Kleinknecht, M., Metz, K. & Bohl, T. (2010). Ein allgemeindidaktisches Kategoriensystem zur Analyse des kognitiven Potenzials von Aufgaben. *Beiträge Zur Lehrerbildung*, 28(1), 84-96.
- Marks, H. M. (2000). Student Engagement in Instructional Activity: Patterns in the Elementary, Middle, and High School Years. *American Educational Research Journal*, 37(1), 153-184. doi:10.3102/00028312037001153
- Markowitz, J. (1986). *Verhalten im Systemkontext: Zum Begriff des sozialen Epigramms. Diskutiert am Beispiel des Schulunterrichts*. Frankfurt am Main: Suhrkamp.
- Merkens, H. (2010). *Unterricht: Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Meseth, W., Proske, M. & Radtke, F.-O. (2011). Was leistet eine kommunikationstheoretische Modellierung des Gegenstandes „Unterricht“? In W. Meseth, M. Proske & F.-O. Radtke (Hrsg.), *Unterrichtstheorien in Forschung und Lehre* (S. 223-240). Bad Heilbrunn: Klinkhardt.
- Meyer, M. (2007). Unterricht. In H.E. Tenorth & R. Tippelt (Hrsg.), *Lexikon Pädagogik* (S. 728-731). Weinheim: Beltz.
- Meyer, M. A. (2001). Schülermitbeteiligung im Fachunterricht - Schülerpartizipation im Horizont (fach-)didaktischer Überlegungen. In J. Böhme & R. T. Kramer (Hrsg.), *Partizipation in der Schule* (S. 49-58). Opladen: Leske + Budrich.
- Moosbrugger, H., & Kelava, A. (2008). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Müller, H. J. (2010). Lernaufgaben und der Aufbau des Wissens. In H. Kiper, W. Meints, S. Peters, S. Schlump & S. Schmit (Hrsg.), *Lernaufgaben und Lernmaterialien im kompetenzorientierten Unterricht* (S. 84-98). Stuttgart: Kohlhammer.

- Müller, C., Eichler, D. & Blömeke, S. (2006). Chancen und Grenzen von Videostudien in der Unterrichtsforschung. In S. Rahm, I. Mammes, & M. Schratz (Hrsg.), *Schulpädagogische Forschung - Unterrichtsforschung - Perspektiven innovativer Ansätze* (S. 125-139). Innsbruck: Studien Verlag.
- Myers, S. S. (1990). The Management of Curriculum Time as it Relates to Student Engaged Time. *Educational Review*, 42(1), 13-23. doi:10.1080/0013191900420102
- Neubrand, J. (2002). *Eine Klassifikation mathematischer Aufgaben zur Analyse von Unterrichtssituationen: Selbsttätiges Arbeiten in Schülerarbeitsphasen in den Stunden der TIMSS-Video-Studie*. Hildesheim: Verlag Franzbecker.
- Neubrand, M., Jordan, A., Krauss, S., Blum, W. & Löwen, K. (2011). Aufgaben im COACTIV-Projekt: Einblicke in das Potenzial für kognitive Aktivierung im Mathematikunterricht. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 115-132). Münster: Waxmann.
- Niegemann, H. M., & Stadler, S. (2001). Hat noch jemand eine Frage? Systematische Unterrichtsbeobachtungen zu Häufigkeit und kognitivem Niveau von Fragen im Unterricht. *Unterrichtswissenschaft*, 29(2), 171-192.
- Nold, G. (2007). DESI im Kontext des Gemeinsamen Europäischen Referenzrahmens für Sprachen. In B. Beck & E. Klieme (Hrsg.), *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International)*(S. 299-305). Weinheim: Beltz.
- Nold, G., & Rossa, H. (2007). Leseverstehen. In B. Beck & E. Klieme (Hrsg.), *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International)*(S. 197-211). Weinheim: Beltz.
- Nystrand, M. & Gamoran, A. (1991). Student Engagement: When Recitation Becomes Conversation. In H. C. Waxman & H. J. Walberg (Hrsg.), *Effective Teaching: Current Research* (S. 257-276). Berkeley: McCutchan Publishing.
- Ophardt, D. & Thiel, F. (2008). Klassenmanagement als Basisdimension der Unterrichtsqualität. In M. Schweer (Hrsg.) *Lehrer-Schüler-Interaktion* (S. 258-282). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Pauli, C. (2012). Kodierende Beobachtung. In H. de Boer & S. Reh (Hrsg.), *Beobachtung in der Schule – Beobachten lernen* (S. 45-63). Wiesbaden: VS/Springer.
- Pauli, C. & Lipowsky, F. (2007). Mitmachen oder zuhören? Mündliche Schülerinnen- und Schülerbeteiligung im Mathematikunterricht. *Unterrichtswissenschaft*, 35(2), 101-124.
- Pauli, C. & Reusser, K. (2006). Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und -entwicklung. *Zeitschrift für Pädagogik*, 52(6), 774-798.
- Peterson, P. L. & Swing, S. R. (1982). Beyond time on task: Students' reports of their thought processes during classroom instruction. *Elementary School Journal*, 82(5), 481-491.
- Peterson, P. L. Swing, S. R., Stark, K. D. & Waas, G. A. (1984). Students' Cognitions and Time on Task During Mathematics Instruction. *American Educational Research Journal*, 21(3), 487-515. doi:10.3102/00028312021003487
- Petko, D. (2006). Kameraskript. In Klieme, E., Pauli, C. & Reusser, K. (Hrsg.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie "Unterrichtsqualität, Lernverhalten und mathematisches Verständnis"* (Teil 3: Hugener, I., Pauli, C. & Reusser, K.: Videoanalysen)(S. 15-37). Frankfurt am Main: GPPF.
- 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.
- PISA-Konsortium Deutschland (Hrsg.)(2004). *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Piwovar, V., Thiel, F. & Ophardt, D. (2013). Training inservice teachers' competencies in classroom management. A quasi-experimental study with teachers of secondary schools. *Teaching and Teacher Education*, 30(C), 1-12.
doi:10.1016/j.tate.2012.09.007

- Prenzel, M., Haussler, P., Rost, J., & Senkbeil, M. (2002). Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorhersagen? *Unterrichtswissenschaft*, 30(2), 120-135.
- Rasch, B., Frieze, M., Hofmann, W. & Naumann, E. (2010). *Quantitative Methoden 1 – Einführung in die Statistik für Psychologen und Sozialwissenschaftler* (3.. Aufl.). Heidelberg: Springer.
- Rasch, B., Frieze, M., Hofmann, W., & Naumann, E. (2006). *Quantitative Methoden 2 – Einführung in die Statistik für Psychologen und Sozialwissenschaftler* (2. Aufl.). Heidelberg: Springer.
- Renkl, A. (1991). *Die Bedeutung der Aufgaben- und Rückmeldungsgestaltung für die Leistungsentwicklung im Fach Mathematik* (Dissertation). Universität Heidelberg.
- Renkl, A. (2010). Lehren und Lernen. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (S. 737-751). Wiesbaden: VS Verlag.
- Reschly, A. L. & Christenson, S. L. (2012). Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the Engagement Construct. In S. L. Christenson, A. L. Reschly & C. Wylie (Hrsg.), *Handbook of Research on Student Engagement* (S. 3-20). New York: Springer.
- Reusser, K., Pauli, C. & Waldis, M. (Hrsg.). (2010). *Unterrichtsgestaltung und Unterrichtsqualität – Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht*. Münster: Waxmann.
- Rimm-Kaufman, S. E., Baroody, A. E., Larsen, R. A. A., Curby, T. W. & Abry, T. (2014). To What Extent Do Teacher-Student Interaction Quality and Student Gender Contribute to Fifth Graders' Engagement in Mathematics Learning? *Journal of Educational Psychology*. doi:10.1037/a0037252
- Rietz, C. & Rudinger, G. (2007). Analyse von Längsschnittdaten. In M. Hasselhorn & W. Schneider (Hrsg.), *Handbuch Entwicklungspsychologie* (S. 635-645). Göttingen: Hogrefe.
- Romberg, T. A. (1980). Salient Features of the BTES Framework of Teacher Behaviors. In C. Denham & A. Lieberman (Hrsg.), *Time to Learn* (S. 73-93). Washington, D.C.: Department of Education, National Institute of Education.

- Rosenshine, B. V. (1980). How Time Is Spent in Elementary Classrooms. In C. Denham & A. Lieberman (Hrsg.), *Time to Learn* (S. 107-126). Washington, D.C.: Department of Education, National Institute of Education.
- Rosenshine, B. V. & Berliner, D. C. (1978). Academic engaged time. *British Journal of Teacher Education*, 4(1), 3-16. doi:10.1080/0260747780040102
- Sacher, W. (1995). *Meldungen und Aufrufe im Unterrichtsgespräch*. Augsburg: Wißner-Verlag.
- Schabram, K. (2007). *Lernaufgaben im Unterricht: Instruktionspsychologische Analysen am Beispiel der Physik (Dissertation)*. Universität Duisburg-Essen
- Scheerens, J. & Bosker, R. J. (1997). *The Foundations of Educational Effectiveness*. Oxford: Pergamon.
- Schmidt-Atzert, L., Büttner, G. & Bühner, M. (2004). Theoretische Aspekte von Aufmerksamkeit-/Konzentrationsdiagnostik. In G. Büttner & L. Schmidt-Atzert (Hrsg.), *Diagnostik von Konzentration und Aufmerksamkeit* (S. 3-22). Göttingen: Hogrefe.
- Schneider, G. (2007). Auf dem Weg zu Skalen für die rezeptiven Kompetenzen im Bereich des Englischen. In B. Beck & E. Klieme (Hrsg.), *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International)*(S. 273-289). Weinheim: Beltz.
- Seel, N. M. (1981). *Lernaufgaben und Lernprozesse*. Stuttgart: Kohlhammer.
- Seidel, T. (2003a). Sichtstrukturen – Organisation unterrichtlicher Aktivitäten. In T. Seidel, M. Prenzel, R. Duit & M. Lehrke (Hrsg.), *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“* (S. 113-128). Kiel: Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN).
- Seidel, T. (2003b). Videobasierte Kodierverfahren in der IPN Videostudie Physik – ein methodischer Überblick. In T. Seidel, M. Prenzel, R. Duit & M. Lehrke (Hrsg.), *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“* (S. 99-111). Kiel: Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN).
- Seidel, T., Dalehefte, I. M. & Meyer, L. (2005). Standardized guidelines - How to collect videotapes. In T. Seidel, M. Prenzel & M. Kobarg (Hrsg.), *How to run a video study - Technical report of the IPN Video Study* (S. 29-53). Münster: Waxmann.

- Seidel, T. & Prenzel, M. (2010). Beobachtungsverfahren: Vom Datenmaterial zur Datenanalyse. In H. Holling & B. Schmitz (Hrsg.), *Handbuch Statistik, Methoden und Evaluation* (S. 700-707). Göttingen: Hogrefe-Verlag.
- Seidel, T., Prenzel, M., Rimmel, 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. & 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
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: University Press.
- Slavin, R. E. (1984). Component building: A strategy for research-based instructional improvement. *Elementary School Journal*, 84(3), 255-269.
- Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: A model of instructional effectiveness. *International Journal of Educational Research*, 21(2), 141-157.
- Statistisches Bundesamt (2010). Bildung und Kultur - Allgemeinbildende Schulen Schuljahr 2008/2009. Fachserie 11 Reihe 1. Online verfügbar unter www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00006815/2110100097004.pdf (Letzter Abruf am 31.12.2014)
- Stein, M. K., Grover, B. W. & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, 33(2), 455-488.
- Stelzl, I. (2005). Fehler und Fallen der Statistik für Psychologen, Pädagogen und Sozialwissenschaftler. Münster: Waxmann.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S. & Serrano, A. M. (1999). *The TIMSS-Videotape Classroom Study: Methods and findings from an exploratory research project on eighth grade mathematics instruction in Germany, Japan, and the United States*. Washington D.C.: National Center for Education Statistics.
- Stipek, D. J. (2002). Good instruction is motivating. In A. Wigfield & J. S. Eccles (Hrsg.), *Development of achievement motivation* (S. 309-332). San Diego: Academic Press.

- Terhart, E. (2002). Fremde Schwestern. Zum Verhältnis von Allgemeiner Didaktik und empirischer Lehr-Lern-Forschung. *Zeitschrift für Pädagogische Psychologie*, 16, 77-86.
- Trautmann, T. (2008). Schulzeit – Unterrichtszeit – Lernzeit. Die Hamburger Messung der Unterrichtsaktivität. In A. Durdel, A. von der Groeben & T. Trautmann (Hrsg.), *Schule als Lebenszeit: Lern- und Lebensrhythmen von Kindern, Lehrkräften und Schulen* (S. 14-27). Weinheim: Beltz.
- Treiber, B. (1982). Lehr- und Lernzeiten im Unterricht. In B. Treiber & F.E. Weinert (Hrsg.), *Lehr-Lern-Forschung. Ein Überblick in Einzeldarstellungen*. (S. 12-36). München: Urban & Schwarzberg.
- Treiber, B. & Weinert, F. E. (Hrsg.).(1982). *Lehr-Lern-Forschung. Ein Überblick in Einzeldarstellungen*. München: Urban & Schwarzberg.
- Treiber, B., Weinert, F. E. & Groeben, N. (1982). Unterrichtsqualität, Leistungsniveau von Schulklassen und individueller Lernfortschritt. *Zeitschrift für Pädagogik*, 28, 563-576.
- Vannest, K. J. & Parker, R. I. (2010). Measuring Time: The Stability of Special Education Teacher Time Use. *Journal of Special Education*, 44(2), 94-106.
doi:10.1177/0022466908329826
- Vollrath, H.-J. (2000). Problemorientierung als didaktisches Prinzip. In P. Baptist (Hrsg.), *Mathematikunterricht im Wandel: Bausteine für den Unterricht* (S. 31-45). Bamberg: C.C. Buchners.
- Walberg, H. J. (1986). Syntheses of Research on Teaching. In M. C. Wittrock (Hrsg.), *Handbook of Research on Teaching* (S. 214-229). New York: Macmillan.
- Walberg, H. J. (1988). Syntheses of research on time and learning. *Educational Leadership*, 45(6), 76-85.
- Walberg, H. J. & Lai, J.-S. (1999). Meta-Analytic Effects for Policy. In G. J. Cizek (Hrsg.), *Handbook of Educational Policy* (S. 419-453). London: Academic Press.
- Walberg, H. , Niemiec, R.P. & Frederick, W.C. (1994). Productive Curriculum Time. *Peabody Journal of Education*, 69(3), 86-100.
- Walter, O. (2008). Herkunftsassoziierte Disparitäten im Lesen, in der Mathematik und in

- den Naturwissenschaften: ein Vergleich zwischen PISA 2000, PISA 2003 und PISA 2006. *Zeitschrift für Erziehungswissenschaft, Sonderheft 10*, 149-168.
- Wang, M. C. & Walberg, H. J. (1983). Adaptive Instruction and Classroom Time. *American Educational Research Journal*, 20(4), 601-626.
- Wang, M. C. & Walberg, H. J. (1991). Teaching and Educational Effectiveness: Research Synthesis and Consensus from the Field. In H. C. Waxmann & H. J. Walberg (Hrsg.), *Effective Teaching – Current Research* (S. 81-104). Berkeley: McCutchan.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a Knowledge Base: Why, How, for Whom? *Review of Educational Research*, 63(3), 365-376.
- Weinert, F. E. & Helmke, A. (Hrsg.).(1997). *Entwicklung im Grundschulalter*. München: Psychologie Verlags Union.
- Weinstein, C. S., Romano, M. & Mignano, A. J. (2010). *Elementary classroom management*. Boston: McGraw Hill.
- Wellenreuther, M. (2007). *Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Wild, K. P. (2003). Videoanalysen als neue Impulsgeber für eine praxisnahe, prozessorientierte empirische Unterrichtsforschung. *Unterrichtswissenschaft*, 31(2), 98-101.
- Wiley, D. E. & Harnischfeger, A. (1974). Explosion of a myth: Quantity of schooling and exposure to instruction, major educational vehicles. *Educational Researcher*, 3(4), 7-12.
- Wirtz, M. A. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- Yair, G. (2000). Educational battlefields in America: The tug-of-war over students' engagement with instruction. *Sociology of Education*, 73(4), 247-269.

Abbildungsverzeichnis

Abb. 1: Schematische Darstellung von Carrolls Modell schulischen Lernens (nach Berliner, 1990)	10
Abb. 2: Das QAIT-Modell effektiven Unterrichts (nach Slavin 1994).....	13
Abb. 3: Skizze des Geräteaufbau zur Erhebung der Video- und Audiodaten im Klassenraum.....	94
Abb. 4: Streuungs- und Lagemaße der Englischtestleistungen nach Klassen und Schulformen.....	126
Abb. 5: Häufigkeitsverteilung der durchschnittlichen individuenbezogenen on task-Rate über die untersuchten Unterrichtsstunden hinweg.....	133
Abb. 6: Mittlere prozentuale on task-Rate der Klassen nach Messzeitpunkten (Unterrichtsstunden)	138
Abb. 7: Häufigkeitsverteilung des zeitlichen Umfangs von Aufgabenstellung und -bearbeitung.....	149
Abb. 8: Schematische Darstellung der Vorgehensweise bei der Auswertung der Beziehungen zwischen den Eigenschaften der Einzelaufgaben und der gleichzeitigen on task-Rate anhand beispielhafter Daten einer Unterrichtsstunde	162

Tabellenverzeichnis

Tab. 1: Überblick über die Schülerstichprobe	92
Tab. 2: Aufbau des Kodiersystems zur Erfassung des Schülerbeteiligungsverhaltens ..	102
Tab. 3: Kennwerte der Beobachterübereinstimmung zur Kodierung des Schülerbeteiligungsverhaltens	106
Tab. 4: Überblick über das Kategoriensystem zur Beschreibung der Aufgaben im Unterricht	110
Tab. 5: Kennwerte der Beobachterübereinstimmung für die beurteilten Aufgabenmerkmale	121
Tab. 6: Häufigkeiten der familiären Umgangssprache, differenziert nach Klassen	124
Tab. 7: Deskriptive Kennwerte der Leistungen im Englischtest	125
Tab. 8: Gruppenmittelwerte der Englischtestleistung nach Geschlecht und familiärem Sprachhintergrund	127
Tab. 9: Deskriptive Ergebnisse der Basiskodierung der Sozialformen	129
Tab. 10: Deskriptive Ergebnisse zu den klassenweise aggregierten on task-Raten in den Unterrichtsstunden	131
Tab. 11: Ergebnisse der mehrfaktoriellen, univariaten Varianzanalyse mit der individuellen durchschnittlichen on task-Rate als abhängige Variable und Personenmerkmalen der Schüler als Faktoren.....	135
Tab. 12: Ergebnisse des Friedman-Tests zur zeitlichen Stabilität der on task-Rate im Klassendurchschnitt über die Unterrichtsstunden hinweg.....	138
Tab. 13: Ergebnisse des Friedman-Tests zur zeitlichen Stabilität der on task-Rate im Stundenverlauf.....	140
Tab. 14: Aufgabenanzahl nach Unterrichtsstunden.....	142
Tab. 15: Häufigkeitsverteilungen der kategorial erfassten Aufgabeneigenschaften.....	143
Tab. 16: Deskriptive Kennwerte metrisch skaliertes Aufgabeneigenschaften	145
Tab. 17: Ergebnisse der Prüfung auf Unterschiede in den Aufgabenmerkmalsausprägungen zwischen den Unterrichtsstunden.....	151
Tab. 18: Zusammenhänge zwischen den Aufgabenmerkmalen	154
Tab. 19: Datenbeispiel zur Illustration der Berechnung der Indikatoren der stundenspezifischen Aufgabenvariabilität.....	156

Tab. 20: Deskriptive Kennwerte zur Ausprägung der Indikatoren zur Beschreibung der stundenspezifischen Variabilität.....	157
Tab. 21: Zusammenhänge (Rangkorrelation) zwischen Merkmalen des Zeitumfangs der Aufgaben und der gleichzeitigen on task-Rate (aller Schüler und unterschiedlicher Teilgruppen der Schülerstichprobe)	163
Tab. 22: Übersicht über signifikante Mittelwertunterschiede zwischen Ausprägungsstufen der Aufgabenmerkmale in der gleichzeitigen on task-Rate (aller Schüler und unterschiedlicher Teilgruppen der Schülerstichprobe).....	166
Tab. 23: Zusammenhänge (Rangkorrelation Spearman) zwischen Indikatoren stundenspezifischer Variabilität der verschiedenen Aufgabenmerkmale und der stundenweise durchschnittlichen on task-Rate (aller Schüler und unterschiedlicher Teilgruppen der Schülerstichprobe)	173
Tab. 24: Korrelationsdifferenzen (Betrag der Effektgröße q) zwischen Untergruppen der Schülerstichprobe, bezogen auf Zusammenhänge zwischen Indikatoren stundenspezifischer Variabilität der Aufgabenmerkmale und der stundenweise durchschnittlichen on task-Rate (aller Schüler und unterschiedlicher Teilgruppen der Schülerstichprobe)	175

Anhang A: Schriftlicher Leistungstest im Unterrichtsfach Englisch

Schüler/in N°:

Ich bin ein Mädchen

Ich bin ein Junge

Zuhause spreche ich Deutsch

Ich spreche zu Hause eine andere Sprache

Welche?:.....

Schüler/in N°:

A telephone call

Jetzt hört ihr ein Gespräch zwischen Frida und Rick. Frida wohnt in Schweden und Rick in Holland. Sie haben sich letztes Jahr in den Ferien kennen gelernt und sie sind Freunde geblieben. Sie schreiben und sprechen miteinander auf Englisch.

1. Höre zuerst den ganzen Text.
2. Jetzt lese alle Fragen durch.
3. Das Telefongespräch wird nun noch einmal abgespielt, dann wird eine Pause gemacht. In dieser Pause beantwortest du auf Deutsch alle Fragen. Du schreibst deine Antwort auf die Linie unter der Frage. Gib die Antwort so genau wie möglich. Denk daran die Frage auf Deutsch zu beantworten.

1. Wann ist Frida's letzter Schultag?

1.

2. Wohin geht Frida in die Ferien?

2.

3. Wie reisen sie dorthin?

3.

4. Wo schlafen sie?

4.

5. Welche Sprachen spricht Frida ausser Schwedisch?

5.

Schüler/in N°:

Anna, die englische Brieffreundin von Laura schreibt, dass sie sie in Frankfurt besuchen kann.
 Bitte lies den Brief aufmerksam durch und entscheide dich bei jeder Lücke, welches Wort deiner Meinung nach am besten passt:
 Am besten legst du ein anders Blatt quer unter die jeweilige Zeile, damit du beim Lesen nicht verrutschtst.

Dear Laura,			
How are you?	Thank you	Sorry	Excuse me
Next week I a short holiday and I can come to Frankfurt.	has got	have got	had
My uncle will take me to the airport in car.	your	my	his
I'm flying to Frankfurt airport	me	her	him
Can you meet there?	doesn't	haven't got	don't
I am a bit afraid, because I speak any German.	When	Why	Who
..... do your holidays start?	Do	Does	Can
..... your brother still like to read?	one	some	any
Can I bring him English books?	does	have	has
And your sister still got her rabbit?	smaller	small	the smallest
My sister has a new rabbit. It is of our three rabbits.	can't	doesn't	don't
My dog Snoopy can't come with me. He like travelling.	am	are	is
I so happy to see you all. It will be so exciting.	is there	are there	have there
Can I come to your school and meet all your friends?	any	some	one
How many boys and girls in your class?	a	some	any
I want to buy ... chocolate for my school friends before I go home again.			
We can't buy German chocolate in our village.			

Please write soon.

Yours

Anna

/15

Das Haus von Lucy

Lucy ist ein englisches Mädchen, das ihrer Freundin Irma erzählt, wo sie wohnt und was für ein Haus sie hat.

Lesen zuerst den Text und beantworte dann die Fragen:

I live in Fleetwood, a small town near Blackpool. It's in the Northwest of England, on the Irish Sea.

We've got a nice house. Next to our house is a garage. There's a big tree in front of our house.

We've got nice neighbours. The Jones on one side have a small pond with fish in front of their garden.

Behind our house we've got a big garden. It's a beautiful garden with lots of flowers. My father has made a swing in the middle of it for my little brother.

There is a small garden house in the corner at the end of the garden.

I hope you can visit us one day. Then I will show you everything.

Frage 1

Hier ist eine Karte von England. Wo liegt Fleetwood? Schreibe den richtige Buchstaben in dieses Kästchen.

Frage 2

Hier ist die Strasse, in der Lucy wohnt. In welcher Hausnummer wohnt Lucy? Schreibe die Hausnummer in dieses Kästchen.

Schüler/in N°:

Frage 3

Hier sehen wir Bilder von dem Garten hinter dem Haus. Welcher Garten gehört zu Lucy's Haus? Mache ein Kreuzchen in das richtige Kästchen.

Quelle: Bader, U. & Schaer, U. (2005). *Evaluation Englisch in den 6. Klassen Appenzell Innerrhoden 2005*. Fachhochschule Nordwestschweiz.

Anhang B: Zusammenhänge zwischen Indikatoren der stundenspezifischen Aufgabenvariabilität

Tabellenteil I

Variabilitätsindikatoren für die unterschiedlichen Aufgabenmerkmale	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>Variationsbreite</i>													
1 Sprache Aufgabenstellung													
2 Verbaler Umfang Aufgabenstellung	0,39												
3 Hilfestellungen bei Aufgabenstellung	.	.											
4 Medium: Darstellungsform	0,12	0,45	.										
5 Medium: Gerät/Material	-0,34	0,07	.	0,65**									
6 Verbaler Umfang Medieninhalte	0,01	0,25	.	0,14	0,00								
7 Lebensweltbezug Aufgabeninhalte	0,17	0,27	.	0,45	0,12	0,53*							
8 Abstraktheit Aufgabeninhalte	0,49	0,10	.	-0,11	-0,49	0,27	0,45						
9 Kognitiver Prozess Bearbeitung	0,37	0,22	.	-0,02	-0,01	0,07	0,17	0,40					
10 Wissensdimension Bearbeitung	-0,01	0,75**	.	0,28	0,14	0,29	0,29	0,29	0,13				
11 Format d. Produkts	0,12	0,67**	.	0,20	-0,12	0,28	-0,01	0,14	0,14	0,54*			
12 Art d. Produkts	0,11	0,22	.	0,36	0,03	-0,09	0,16	-0,32	-0,48	-0,05	-0,01		
13 Reflexions-/Auswertungsfunktion	0,01	0,09	.	-0,03	0,33	0,32	0,17	-0,12	0,27	0,05	-0,20	0,12	
<i>Gleichförmigkeit</i>													
14 Sprache Aufgabenstellung	-0,89**	-0,32	.	-0,03	0,45	0,14	-0,12	-0,54*	-0,35	-0,03	-0,06	-0,02	0,17
15 Verbaler Umfang Aufgabenstellung	0,06	-0,69**	.	-0,15	-0,17	-0,30	-0,05	-0,11	-0,31	-0,78**	-0,53*	0,26	-0,15
16 Hilfestellungen bei Aufgabenstellung	0,01	-0,44	.	0,07	0,282	-0,28	-0,26	-0,47	-0,28	-0,65**	-0,52*	0,16	0,02
17 Medium: Darstellungsform	-0,07	-0,21	.	-0,82**	-0,45	0,06	-0,42	0,12	0,21	-0,02	-0,05	-0,51*	0,11
18 Medium: Gerät/Material	0,32	0,12	.	-0,59*	-0,65**	0,28	0,05	0,58*	0,25	0,27	-0,05	-0,31	0,12
19 Verbaler Umfang Medieninhalte	-0,42	-0,50*	.	-0,05	0,37	-0,49	-0,14	-0,24	-0,08	-0,18	-0,62**	-0,04	0,07
20 Lebensweltbezug Aufgabeninhalte	-0,64**	-0,29	.	-0,27	0,15	-0,26	-0,51*	-0,56*	-0,24	-0,03	-0,13	0,13	-0,02
21 Abstraktheit Aufgabeninhalte	-0,48	-0,01	.	0,15	0,35	0,10	-0,18	-0,70**	-0,44	-0,02	0,04	0,32	0,11
22 Kognitiver Prozess Bearbeitung	-0,33	-0,12	.	0,14	0,21	-0,14	-0,18	-0,46	-0,88**	-0,09	-0,07	0,42	-0,24
23 Wissensdimension Bearbeitung	0,08	-0,67**	.	-0,35	0,00	-0,33	-0,28	-0,28	0,04	-0,86**	-0,67**	-0,10	0,18

Koeffizient: Spearman r_s ; n = 16 Unterrichtsstunden; Signifikanzniveaus: * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; keine Kennzeichnung $p > 0,05$ (zweiseitig)

Tabellenteil II

Variabilitätsindikatoren für die unterschiedlichen Aufgabenmerkmale	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>(Fortsetzung Variationsbreite, Tabellenteil I)</i>													
24 Format d. Produkts	-0,17	-0,66**	.	-0,18	0,26	-0,12	-0,08	-0,31	0,05	-0,70**	-0,79**	-0,05	0,45
25 Art d. Produkts	-0,05	0,03	.	-0,35	-0,12	0,19	-0,04	0,26	0,20	0,22	-0,02	-0,66**	-0,07
26 Reflexions-/Auswertungsfunktion	0,05	-0,07	.	0,08	-0,14	-0,28	-0,01	0,10	-0,21	-0,03	0,09	-0,07	-0,80**
<i>Wechselhaftigkeit</i>													
27 Sprache Aufgabenstellung	0,89**	0,49	.	0,11	-0,37	-0,01	0,14	0,44	0,21	0,09	0,16	0,22	0,12
28 Verbaler Umfang Aufgabenstellung	0,32	0,68**	.	0,47	0,20	0,53*	0,61*	0,19	0,07	0,50	0,283	0,17	0,20
29 Hilfestellungen bei Aufgabenstellung	0,38	0,50*	.	0,40	0,21	0,55*	0,49	0,08	0,19	0,20	0,226	-0,06	0,20
30 Medium: Darstellungsform	0,34	0,42	.	0,63**	0,10	0,01	0,46	0,15	0,01	0,09	0,17	0,35	-0,05
31 Medium: Gerät/Material	0,08	0,32	.	0,76**	0,67**	-0,06	0,16	-0,33	0,10	0,05	0,22	0,18	0,06
32 Verbaler Umfang Medieninhalte	0,21	0,22	.	0,24	0,135	0,80**	0,52*	0,15	0,20	0,11	0,252	-0,07	0,23
33 Lebensweltbezug Aufgabeninhalte	0,46	0,43	.	0,51*	0,12	0,65**	0,75**	0,32	0,18	0,25	0,04	0,08	0,12
34 Abstraktheit Aufgabeninhalte	0,42	0,26	.	0,05	-0,27	0,25	0,53*	0,67**	0,44	0,25	0,10	0,01	0,10
35 Kognitiver Prozess Bearbeitung	0,55*	0,63**	.	0,16	-0,23	0,40	0,39	0,50*	0,69**	0,37	,65**	-0,18	0,07
36 Wissensdimension Bearbeitung	0,43	0,82**	.	0,46	0,06	0,38	0,37	0,23	0,43	0,60*	,69**	0,09	0,05
37 Format d. Produkts	0,38	0,87**	.	0,34	-0,05	0,33	0,20	0,30	0,39	0,69**	,86**	0,02	0,01
38 Art d. Produkts	0,37	0,31	.	0,51*	0,01	0,01	0,24	-0,04	-0,32	0,00	0,25	0,82**	-0,04
39 Reflexions-/Auswertungsfunktion	0,25	0,44	.	0,27	0,23	0,41	0,48	0,11	0,28	0,27	0,11	0,28	0,77**
<i>Zeitmerkmale & Aufgabenanzahl</i>													
40 Mittlere Aufgabendauer	-0,28	-0,38	.	-0,30	-0,01	-0,28	-0,54*	-0,36	-0,19	-0,38	-0,18	0,05	0,29
41 Mittlere Dauer Aufgabenstellungen	-0,35	-0,14	.	-0,38	-0,04	-0,03	-0,18	-0,12	-0,14	0,11	-0,07	-0,19	0,13
42 Mittlere Dauer Aufgabenbearbeitung	-0,30	-0,40	.	-0,24	0,03	-0,29	-0,57*	-0,45	-0,16	-0,48	-0,19	0,07	0,29
43 Aufgabenanzahl	0,27	0,51*	.	0,42	0,06	0,36	0,63**	0,27	0,00	0,34	0,22	0,25	0,09

Koeffizient: Spearman r_s ; n = 16 Unterrichtsstunden; Signifikanzniveaus: * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; keine Kennzeichnung $p > 0,05$ (zweiseitig)

Tabellenteil III

Variabilitätsindikatoren für die unterschiedlichen Aufgabenmerkmale	14	15	16	17	18	19	20	21	22	23	24	25	26
<i>Gleichförmigkeit</i>													
14 Sprache Aufgabenstellung													
15 Verbaler Umfang Aufgabenstellung	-0,14												
16 Hilfestellungen bei Aufgabenstellung	0,11	0,41											
17 Medium: Darstellungsform	0,11	-0,18	-0,09										
18 Medium: Gerät/Material	-0,37	-0,25	-0,33	0,72**									
19 Verbaler Umfang Medieninhalte	0,2	0,28	0,37	-0,08	-0,21								
20 Lebensweltbezug Aufgabeninhalte	0,56*	-0,03	0,21	0,28	-0,13	0,53*							
21 Abstraktheit Aufgabeninhalte	0,40	0,21	0,15	-0,10	-0,33	0,17	0,48						
22 Kognitiver Prozess Bearbeitung	0,34	0,17	0,31	-0,37	-0,44	0,17	0,22	0,29					
23 Wissensdimension Bearbeitung	-0,04	0,604*	0,72**	0,12	-0,14	0,30	0,05	-0,06	-0,03				
24 Format d. Produkts	0,20	0,44	0,71**	0,06	-0,15	0,56*	0,21	0,05	-0,01	0,83**			
25 Art d. Produkts	-0,03	-0,46	0,01	0,50*	0,51*	0,13	0,00	-0,34	-0,05	-0,06	0,06		
26 Reflexions-/Auswertungsfunktion	-0,05	0,1	0,12	0,02	-0,07	-0,02	0,01	-0,20	0,24	-0,16	-0,34	0,20	
<i>Wechselhaftigkeit</i>													
27 Sprache Aufgabenstellung	-0,86**	-0,01	-0,09	-0,14	0,35	-0,38	-0,66**	-0,38	-0,20	-0,05	-0,21	-0,03	-0,15
28 Verbaler Umfang Aufgabenstellung	-0,13	-0,55*	-0,09	-0,26	0,12	-0,36	-0,42	-0,23	0,11	-0,41	-0,24	0,33	0,07
29 Hilfestellungen bei Aufgabenstellung	-0,29	-0,21	0,07	-0,16	0,10	-0,29	-0,49	0,11	-0,09	-0,19	-0,05	0,35	-0,01
30 Medium: Darstellungsform	-0,35	0,01	0,04	-0,55*	-0,19	-0,11	-0,53*	-0,07	0,01	-0,29	-0,16	-0,12	0,12
31 Medium: Gerät/Material	0,04	-0,06	0,26	-0,49	-0,58*	0,08	-0,14	0,27	0,07	-0,16	-0,01	-0,15	0,17
32 Verbaler Umfang Medieninhalte	-0,06	-0,05	0	-0,05	0,07	-0,39	-0,26	0,22	-0,22	-0,05	-0,00	0,06	-0,16
33 Lebensweltbezug Aufgabeninhalte	-0,37	-0,1	-0,03	-0,29	0,19	-0,31	-0,54*	0,00	-0,21	-0,21	-0,08	0,10	0,00
34 Abstraktheit Aufgabeninhalte	-0,27	-0,34	-0,13	0,07	0,41	-0,18	-0,29	-0,73**	-0,38	-0,19	-0,08	0,28	0,14
35 Kognitiver Prozess Bearbeitung	-0,47	-0,38	-0,48	0,06	0,28	-0,63**	-0,51*	-0,28	-0,68**	-0,36	-0,47	0,05	-0,06
36 Wissensdimension Bearbeitung	-0,28	-0,51*	-0,52*	-0,17	0,05	-0,67**	-0,37	-0,05	-0,41	-0,56*	-0,66**	-0,22	-0,03
37 Format d. Produkts	-0,29	-0,71**	-0,44	-0,07	0,15	-0,54*	-0,24	-0,16	-0,29	-0,66**	-0,66**	0,13	-0,02
38 Art d. Produkts	-0,19	0,22	0,00	-0,64**	-0,40	-0,37	-0,24	0,03	0,33	-0,18	-0,29	-0,73**	0,05
39 Reflexions-/Auswertungsfunktion	-0,14	-0,29	-0,09	-0,28	0,04	-0,10	-0,28	-0,05	-0,18	-0,06	0,17	-0,07	-0,72**

Koeffizient: Spearman r_s ; n = 16 Unterrichtsstunden; Signifikanzniveaus: * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; keine Kennzeichnung $p > 0,05$ (zweiseitig)

Tabellenteil IV

Variabilitätsindikatoren für die unterschiedlichen Aufgabenmerkmale	14	15	16	17	18	19	20	21	22	23	24	25	26
<i>Zeitmerkmale & Aufgabenanzahl</i>													
40 Mittlere Aufgabendauer	0,21	0,34	-0,04	0,06	-0,21	0,117	0,15	0,29	0,11	0,30	0,25	-0,41	-0,54*
41 Mittlere Dauer Aufgabenstellungen	0,35	-0,07	-0,45	0,25	0,07	-0,177	0,04	0,06	0,05	-0,01	-0,16	-0,18	-0,23
42 Mittlere Dauer Aufgabenbearbeitung	0,20	0,38	0,12	0,03	-0,26	0,2	0,20	0,38	0,09	0,36	0,37	-0,36	-0,55*
43 Aufgabenanzahl	-0,10	-0,36	0,01	-0,20	0,14	-0,252	-0,40	-0,27	0,06	-0,40	-0,20	0,27	0,25

Koeffizient: Spearman r_s ; n = 16 Unterrichtsstunden; Signifikanzniveaus: * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; keine Kennzeichnung $p > 0,05$ (zweiseitig)

Tabellenteil V

Variabilitätsindikatoren für die unterschiedlichen Aufgabenmerkmale	27	28	29	30	31	32	33	34	35	36	37	38	39
<i>Wechselhaftigkeit</i>													
27 Sprache Aufgabenstellung													
28 Verbaler Umfang Aufgabenstellung	0,35												
29 Hilfestellungen bei Aufgabenstellung	0,39	0,73**											
30 Medium: Darstellungsform	0,53*	0,47	0,53*										
31 Medium: Gerät/Material	0,07	0,39	0,53*	0,65**									
32 Verbaler Umfang Medieninhalte	0,03	0,44	0,69**	0,02	0,17								
33 Lebensweltbezug Aufgabeninhalte	0,37	0,69**	0,82**	0,46	0,29	0,75**							
34 Abstraktheit Aufgabeninhalte	0,32	0,54*	0,17	0,25	-0,09	0,16	0,34						
35 Kognitiver Prozess Bearbeitung	0,46	0,39	0,49	0,29	0,17	0,51*	0,47	0,48					
36 Wissensdimension Bearbeitung	0,32	0,48	0,38	0,22	0,27	0,46	0,49	0,28	0,81**				
37 Format d. Produkts	0,41	0,57*	0,43	0,34	0,29	0,33	0,33	0,43	0,79**	0,80**			
38 Art d. Produkts	0,36	0,23	-0,02	0,37	0,23	0,08	0,18	0,11	0,10	0,37	0,22		
39 Reflexions-/Auswertungsfunktion	0,40	0,53*	0,43	0,30	0,15	0,40	0,41	0,36	0,36	0,32	0,40	0,23	
<i>Zeitmerkmale & Aufgabenanzahl</i>													
40 Mittlere Aufgabendauer	-0,09	-0,67**	-0,50	-0,25	-0,20	-0,44	-0,61*	-0,65**	-0,40	-0,38	-0,44	-0,02	-0,05
41 Mittlere Dauer Aufgabenstellungen	-0,33	-0,32	-0,48	-0,59*	-0,52*	-0,25	-0,40	-0,40	-0,22	0,01	-0,31	-0,13	-0,18
42 Mittlere Dauer Aufgabenbearbeitung	-0,08	-0,65**	-0,37	-0,11	-0,04	-0,39	-0,56*	-0,65**	-0,38	-0,44	-0,42	-0,07	-0,02
43 Aufgabenanzahl	0,35	0,86**	0,63**	,69**	0,40	0,27	0,58*	0,65**	0,35	0,30	0,48	0,25	0,39

Koeffizient: Spearman r_s ; n = 16 Unterrichtsstunden; Signifikanzniveaus: * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; keine Kennzeichnung $p > 0,05$ (zweiseitig)

Tabellenteil VI

Variabilitätsindikatoren für die unterschiedlichen Aufgabenmerkmale	40	41	42
<i>Zeitmerkmale & Aufgabenanzahl</i>			
40 Mittlere Aufgabendauer			
41 Mittlere Dauer Aufgabenstellungen	0,51*		
42 Mittlere Dauer Aufgabenbearbeitung	0,96**	0,28	
43 Aufgabenanzahl	-0,68**	-0,50*	-0,61*

Koeffizient: Spearman r_s ; n = 16 Unterrichtsstunden;

Signifikanzniveaus: * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; keine Kennzeichnung $p > 0,05$ (zweiseitig)

Anhang C: Mittlere on task-Raten für alle Aufgabenmerkmalsausprägungen (differenziert nach Gesamtstichprobe und Untergruppen der Schülerstichprobe)

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Sprache der Einführung und Präsentation</i>									
Englisch	M	86,04 % ^A	83,40 % ^A	86,82 %	84,02 %	88,23 %	85,38 %	86,64 %	87,01 % ^A
	N	89	89	89	89	89	89	89	89
	SD	12,12 %	17,28 %	13,71 %	15,59 %	13,34 %	15,16 %	15,24 %	17,14 %
Deutsch	M	87,92 %	83,69 % ^B	88,12 %	85,04 %	92,09 %	87,69 %	87,42 %	89,52 %
	N	38	38	38	38	38	38	38	38
	SD	10,56 %	18,63 %	10,97 %	14,09 %	10,45 %	15,24 %	12,13 %	13,02 %
Kombination	M	93,70 % ^A	96,82 % ^{A, B}	92,54 %	92,18 %	93,93 %	92,90 %	90,70 %	96,59 % ^A
	N	9	9	9	9	9	9	9	9
	SD	5,74 %	5,89 %	6,41 %	10,43 %	7,11 %	12,60 %	10,04 %	7,28 %
Varianzanalyse	-	-	-	-	-	-	-	-	-
Post-Hoc	^A p < 0,05	^A p < 0,001 ^B p < 0,001	-	-	-	-	-	-	^A p < 0,05

- Fortsetzung der Tabelle s. nächste Seite -

Anmerkungen:

Nur signifikante Varianzanalysen berichtet (Signifikanzniveau 5 %; bei Verletzung der Varianzhomogenitätsvoraussetzung gemäß Levene-Test herabgesetzt auf 1 %, s. Bühner & Ziegler, 2009, S. 369f.)

Post-Hoc-Tests: Games Howell bei Varianzheterogenität, Bonferroni bei Varianzhomogenität; signifikante Mittelwertunterschiede zwischen Gruppen sind durch gleiche Hochzeichen markiert

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Verbaler Umfang Einführung & Präsentation</i>									
klein (1. Perzentil)	M	83,51 %	77,90 %	85,08 %	83,32 %	86,00 %	84,28 %	86,09 %	82,87 %
	N	30	30	30	30	30	30	30	30
	SD	9,69 %	18,17 %	11,14 %	12,63 %	12,94 %	12,28 %	11,92 %	19,39 %
eher klein (2. Perzentil)	M	88,39 %	87,47 %	88,03 %	85,86 %	90,92 %	87,19 %	87,29 %	90,32 %
	N	38	38	38	38	38	38	38	38
	SD	12,37 %	17,72 %	13,34 %	14,71 %	12,49 %	15,43 %	16,30 %	14,94 %
eher groß (3. Perzentil)	M	85,81 %	86,15 %	85,83 %	82,54 %	88,43 %	85,28 %	83,92 %	89,86 %
	N	32	32	32	32	32	32	32	32
	SD	13,70 %	15,95 %	15,56 %	19,17 %	14,42 %	19,33 %	14,83 %	16,49 %
groß (4. Perzentil)	M	89,76 %	84,90 %	90,69 %	87,09 %	92,55 %	88,80 %	90,68 %	89,46 %
	N	36	36	36	36	36	36	36	36
	SD	9,06 %	16,97 %	9,68 %	12,73 %	8,92 %	12,48 %	12,25 %	11,49 %
Varianzanalyse	-	-	-	-	-	-	-	-	-
Post-Hoc	-	-	-	-	-	-	-	-	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Anzahl Hilfestellungsmaßnahmen in Aufgabenstellung</i>									
wenig (1. Perzentil)	M	86,88 %	83,77 %	87,39 %	84,72 %	89,83 %	87,11 %	86,65 %	87,99 %
	N	49	49	49	49	49	49	49	49
	SD	10,26 %	16,62 %	11,75 %	12,87 %	12,24 %	13,36 %	13,21 %	16,85 %
mittel (2. Perzentil)	M	87,04 %	84,78 %	87,82 %	84,94 %	89,33 %	86,97 %	87,24 %	88,09 %
	N	60	60	60	60	60	60	60	60
	SD	11,59 %	17,18 %	12,91 %	16,66 %	11,99 %	15,80 %	14,42 %	15,28 %
viel (3. Perzentil)	M	87,42 %	84,60 %	87,18 %	85,14 %	90,00 %	84,25 %	87,50 %	89,99 %
	N	26	26	26	26	26	26	26	26
	SD	13,92 %	20,19 %	14,32 %	15,20 %	14,08 %	16,91 %	15,64 %	15,20 %
Varianzanalyse		-	-	-	-	-	-	-	-
Post-Hoc		-	-	-	-	-	-	-	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Medium: Darstellungsform</i>									
kein Medium	M	86,36 %	83,43 %	87,87 %	83,00 %	92,14 %	84,08 %	87,43 %	90,00 %
	N	23	23	23	23	23	23	23	23
	SD	14,08 %	17,18 %	14,21 %	16,79 %	14,28 %	18,81 %	16,78 %	14,52 %
Text	M	88,06 %	85,05 %	88,43 % ^A	85,27 %	90,48 %	87,20 %	87,97 %	89,51 %
	N	72	72	72	72	72	72	72	72
	SD	9,74 %	15,98 %	11,49 %	14,87 %	9,59 %	14,35 %	12,91 %	13,61 %
Abbildungen	M	86,43 %	82,97 %	86,36 %	87,47 %	86,00 %	85,78 %	87,27 %	85,63 %
	N	16	16	16	16	16	16	16	16
	SD	14,53 %	23,64 %	14,30 %	12,94 %	16,48 %	18,33 %	11,47 %	19,28 %
Hörbeispiel	M	91,08 %	83,92 %	94,91 % ^{A, B}	90,14 %	93,30 %	90,93 %	92,41 %	90,19 %
	N	9	9	9	9	9	9	9	9
	SD	6,04 %	14,16 %	4,83 %	10,35 %	9,16 %	12,13 %	8,30 %	13,03 %
Kombination	M	81,03 %	82,06 %	79,42 % ^B	78,81 %	83,25 %	84,37 %	78,53 %	81,79 %
	N	14	14	14	14	14	14	14	14
	SD	13,28 %	20,81 %	14,34 %	16,68 %	16,75 %	10,32 %	18,95 %	23,82 %
Varianzanalyse		-	-	-	-	-	-	-	-
Post-Hoc		-	-	^A p < 0,05 ^B p < 0,05	-	-	-	-	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Medium: Gerät/Material</i>									
kein Medium	M	86,36 %	83,43 %	87,87 %	83,00 %	92,14 %	84,08 %	87,43 %	90,00 %
	N	23	23	23	23	23	23	23	23
	SD	14,08 %	17,18 %	14,21 %	16,79 %	14,28 %	18,81 %	16,78 %	14,52 %
Tafel/Projektor	M	88,12 %	84,10 %	88,54 %	86,07 %	89,46 %	86,93 %	89,35 %	87,22 %
	N	35	35	35	35	35	35	35	35
	SD	10,19 %	17,60 %	11,67 %	16,69 %	11,17 %	14,75 %	12,22 %	16,13 %
Arbeitsblatt	M	88,44 %	84,38 %	90,25 %	86,88 %	92,22 %	92,50 %	81,14 %	93,75 %
	N	10	10	10	10	10	10	10	10
	SD	11,74 %	12,16 %	14,38 %	14,57 %	6,54 %	16,87 %	10,92 %	15,59 %
Lehrbuch	M	84,07 % ^A	82,51 %	83,37 % ^A	81,44 % ^A	86,13 % ^A	83,49 %	84,31 % ^A	85,41 %
	N	49	49	49	49	49	49	49	49
	SD	11,68 %	19,77 %	12,63 %	13,64 %	13,74 %	13,71 %	15,55 %	17,26 %
andere	M	92,62 % ^A	90,44 %	94,56 % ^A	91,62 % ^A	95,22 % ^A	92,66 %	93,06 % ^A	92,70 %
	N	17	17	17	17	17	17	17	17
	SD	7,67 %	12,97 %	8,04 %	10,21 %	8,00 %	10,92 %	8,60 %	11,48 %
Varianzanalyse	-	-	F (4;129) = 2,96; p < 0,05; η ² = 0,08	-	-	-	-	-	-
Post-Hoc	^A p < 0,05	-	^A p < 0,001	^A p < 0,05	^A p < 0,05	-	^A p < 0,05	-	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale		Untergruppen der Schülerstichprobe							
		Umgangssprache in der Familie			Geschlecht		Leistungsniveau im Klassenvergleich		
		Alle Schüler	Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch
<i>Bekanntheit des Mediums</i>									
kein Medium	M	86,36 %	83,43 %	87,87 %	83,00 %	92,14 %	84,08 %	87,43 %	90,00 %
	N	23	23	23	23	23	23	23	23
	SD	14,08 %	17,18 %	14,21 %	16,79 %	14,28 %	18,81 %	16,78 %	14,52 %
bekannt	M	85,72 %	83,30 %	85,64 %	83,67 %	87,31 % ^A	86,82 %	83,68 % ^{A, B}	87,41 %
	N	76	76	76	76	76	76	76	76
	SD	11,95 %	18,13 %	13,63 %	15,93 %	13,26 %	15,19 %	14,50 %	17,09 %
simultane Einführung	M	90,72 %	88,80 %	91,39 %	88,59 %	93,77 % ^A	87,83 %	94,40 % ^A	90,53 %
	N	29	29	29	29	29	29	29	29
	SD	8,17 %	15,97 %	8,84 %	11,46 %	7,22 %	12,77 %	8,25 %	13,29 %
Einführung angekündigt	M	90,42 %	85,60 %	91,52 %	89,05 %	91,85 %	86,19 %	93,17 % ^B	88,20 %
	N	7	7	7	7	7	7	7	7
	SD	5,19 %	10,94 %	6,26 %	9,07 %	7,77 %	11,35 %	6,68 %	10,96 %
Varianzanalyse		-	-	-	-	-	-	F (3;131) = 4,87; p < 0,01; η ² = 0,10	-
Post-Hoc		-	-	-	-	^A p < 0,05	-	^A p < 0,001 ^B p < 0,05	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Verbaler Umfang Medieninhalte</i>									
klein (1. Perzentil)	M	86,99 %	86,47 %	86,66 %	84,49 %	88,03 %	84,46 %	89,23 %	87,38 %
	N	25	25	25	25	25	25	25	25
	SD	15,27 %	18,27 %	16,03 %	19,26 %	16,24 %	18,42 %	14,31 %	22,12 %
eher klein (2. Perzentil)	M	85,76 %	82,92 %	85,08 %	84,91 %	86,60 %	86,32 %	81,35 %	87,90 %
	N	26	26	26	26	26	26	26	26
	SD	12,18 %	20,85 %	14,19 %	15,92 %	12,76 %	14,06 %	17,10 %	18,65 %
eher groß (3. Perzentil)	M	87,21 %	85,94 %	87,01 %	84,71 %	90,19 %	87,57 %	87,18 %	87,74 %
	N	25	25	25	25	25	25	25	25
	SD	8,94 %	15,63 %	9,40 %	12,87 %	10,28 %	11,18 %	11,91 %	12,72 %
groß (4. Perzentil)	M	89,77 %	82,69 %	92,37 %	88,83 %	91,61 %	90,09 %	90,23 %	92,56 %
	N	26	26	26	26	26	26	26	26
	SD	6,53 %	16,56 %	6,56 %	10,55 %	8,15 %	11,75 %	10,27 %	8,57 %
Varianzanalyse	-	-	-	-	-	-	-	-	-
Post-Hoc	-	-	-	-	-	-	-	-	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale		Untergruppen der Schülerstichprobe							
		Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich		
			Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch
<i>Lebensweltbezug d. Aufgabeninhalte</i>									
ohne	M	89,98 %	81,85 %	92,03 % ^A	86,18 %	93,81 % ^A	91,07 %	91,78 % ^A	87,67 %
	N	34	34	34	34	34	34	34	34
	SD	7,47 %	17,46 %	7,31 %	12,47 %	9,35 %	11,34 %	9,32 %	12,11 %
konstruiert	M	86,24 %	86,24 %	85,95 % ^A	85,59 %	88,21 % ^A	86,12 %	85,08 % ^A	88,70 %
	N	63	63	63	63	63	63	63	63
	SD	9,64 %	15,45 %	11,75 %	12,81 %	10,54 %	12,65 %	13,68 %	16,41 %
authentisch	M	85,84 %	83,39 %	86,28 %	82,31 %	88,51 %	83,12 %	86,25 %	88,56 %
	N	37	37	37	37	37	37	37	37
	SD	16,48 %	20,72 %	16,91 %	20,04 %	16,70 %	20,54 %	17,72 %	17,99 %
Varianzanalyse		-	-	-	-	-	-	-	-
Post-Hoc		-	-	^A p < 0,01	-	^A p < 0,05	-	^A p < 0,05	-
<i>Abstraktheit d. Aufgabeninhalte</i>									
konkret	M	86,82 %	84,70 %	87,38 %	85,94 %	88,99 %	86,05 %	85,61 % ^A	90,18 %
	N	76	76	76	76	76	76	76	76
	SD	11,74 %	17,37 %	12,82 %	13,10 %	13,08 %	15,47 %	14,69 %	15,01 %
leicht abstrakt	M	84,80 %	84,64 %	84,37 %	79,28 %	88,43 %	84,60 %	85,91 %	82,49 %
	N	28	28	28	28	28	28	28	28
	SD	13,69 %	16,98 %	15,96 %	20,23 %	12,97 %	16,87 %	15,83 %	20,22 %
abstrakt	M	89,87 %	83,13 %	91,10 %	87,21 %	92,74 %	89,65 %	92,00 % ^A	89,41 %
	N	30	30	30	30	30	30	30	30
	SD	8,26 %	18,85 %	7,57 %	13,19 %	9,89 %	12,39 %	10,03 %	11,77 %
Varianzanalyse		-	-	-	-	-	-	-	-
Post-Hoc		-	-	-	-	-	-	^A p < 0,05	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Kognitiver Prozess d. Bearbeitung</i>									
erinnern	M	83,52 %	83,29 %	83,91 %	81,90 %	89,08 %	79,56 %	87,59 %	87,29 %
	N	22	22	22	22	22	22	22	22
	SD	14,63 %	17,91 %	14,64 %	16,96 %	14,29 %	20,24 %	12,92 %	16,34 %
verstehen	M	89,68 %	89,45 %	89,32 %	86,20 %	91,43 %	88,13 %	89,25 %	91,28 %
	N	45	45	45	45	45	45	45	45
	SD	9,64 %	14,28 %	11,96 %	16,20 %	10,10 %	14,03 %	12,40 %	12,10 %
anwenden	M	87,41 %	82,17 %	89,12 %	86,35 %	89,61 %	86,92 %	87,87 %	87,41 %
	N	36	36	36	36	36	36	36	36
	SD	11,76 %	19,68 %	12,18 %	12,85 %	13,61 %	14,61 %	14,97 %	16,34 %
analysieren	M	90,10 %	82,75 %	89,87 %	87,79 %	90,44 %	93,09 %	85,55 %	92,82 %
	N	11	11	11	11	11	11	11	11
	SD	5,99 %	18,60 %	8,45 %	11,63 %	10,86 %	9,99 %	12,80 %	7,86 %
evaluieren	M	89,26 %	86,17 %	93,48 %	86,03 %	93,46 %	88,00 %	91,83 %	92,84 %
	N	5,00	5,00	5,00	5	5	5	5	5
	SD	4,72 %	10,60 %	5,31 %	7,60 %	6,09 %	17,89 %	9,87 %	8,77 %
Varianzanalyse		-	-	-	-	-	-	-	-
Post-Hoc		-	-	-	-	-	-	-	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale		Untergruppen der Schülerstichprobe							
		Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich		
			Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch
<i>Wissensdimension i. d. Bearbeitung</i>									
Fakten	M	87,26 %	83,50 %	88,60 %	84,45 %	91,59 %	86,98 %	87,83 %	88,74 %
	N	36	36	36	36	36	36	36	36
	SD	12,11 %	15,15 %	12,83 %	14,28 %	12,24 %	17,20 %	12,44 %	15,58 %
Konzepte	M	85,64 %	90,20 %	83,64 %	80,07 %	90,24 %	82,91 %	86,67 %	87,34 %
	N	22	22	22	22	22	22	22	22
	SD	10,89 %	13,14 %	13,73 %	18,79 %	10,58 %	15,44 %	13,65 %	12,98 %
Prozeduren	M	88,78 %	85,27 %	89,61 %	87,77 %	90,05 %	87,96 %	89,03 %	89,84 %
	N	59	59	59	59	59	59	59	59
	SD	10,92 %	19,25 %	11,03 %	13,11 %	12,45 %	14,89 %	13,61 %	14,43 %
Varianzanalyse		-	-	-	-	-	-	-	-
Post-Hoc		-	-	-	-	-	-	-	-
<i>Format d. geforderten Produkts</i>									
Verarbeitung v. Vorgaben	M	87,86 %	83,31 %	88,65 %	86,39 %	89,28 %	87,46 %	87,49 %	89,60 %
	N	54	54	54	54	54	54	54	54
	SD	10,12 %	17,22 %	11,46 %	13,45 %	11,32 %	13,05 %	13,37 %	15,45 %
eingeschränkte Konstr.	M	83,23 %	83,98 %	83,29 %	80,90 %	89,09 %	82,62 %	84,96 %	84,18 %
	N	41	41	41	41	41	41	41	41
	SD	14,93 %	19,48 %	15,75 %	16,78 %	15,15 %	18,27 %	16,21 %	18,09 %
erweiterte Konstr.	M	88,54 %	82,21 %	90,26 %	83,40 %	90,99 %	86,89 %	90,43 %	88,40 %
	N	23	23	23	23	23	23	23	23
	SD	8,81 %	17,72 %	9,90 %	17,72 %	10,33 %	16,53 %	9,88 %	15,10 %
Varianzanalyse		-	-	-	-	-	-	-	-
Post-Hoc		-	-	-	-	-	-	-	-

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Art d. geforderten Produkts</i>									
Sprache	M	86,33 % ^{A, B}	83,57 % ^A	86,81 % ^A	83,94 % ^A	89,05 %	85,58 % ^{A, B}	86,58 % ^{A, B}	87,94 % ^A
	N	119	119	119	119	119	119	119	119
	SD	11,59 %	17,96 %	12,72 %	15,10 %	12,87 %	15,14 %	14,50 %	16,01 %
Schrift	M	81,41 %	84,88 %	79,73 %	78,03 %	88,71 %	78,30 %	85,08 %	77,82 %
	N	4	4	4	4	4	4	4	4
	SD	15,46 %	14,69 %	17,69 %	19,46 %	7,60 %	20,29 %	10,60 %	20,14 %
Kombination	M	90,97 % ^A	81,25 %	92,71 %	87,50 %	90,69 %	100,00 % ^A	75,00 % ^A	96,88 %
	N	2	2	2	2	2	2	2	2
	SD	0,98 %	8,84 %	4,42 %	5,89 %	2,55 %	0,00 %	0,00 %	4,42 %
andere	M	97,31 % ^B	97,92 % ^A	97,85 % ^A	98,72 % ^A	97,50 %	100,00 % ^B	94,44 % ^B	98,81 % ^A
	N	6	6	6	6	6	6	6	6
	SD	2,79 %	5,10 %	3,34 %	3,14 %	6,12 %	0,00 %	8,86 %	2,92 %
Varianzanalyse	-	-	-	-	-	-	-	-	-
Post-Hoc	^A p < 0,05 ^B p < 0,001	^A p = 0,001	^A p < 0,001	^A p < 0,001	-	^A p < 0,001 ^B p < 0,001	^A p < 0,001 ^B p < 0,05	p < 0,001	

- Fortsetzung der Tabelle s. nächste Seite -

Aufgabenmerkmale	Untergruppen der Schülerstichprobe								
	Alle Schüler	Umgangssprache in der Familie		Geschlecht		Leistungsniveau im Klassenvergleich			
		Deutsch	Andere/ Deutsch + andere	männlich	weiblich	niedrig	mittel	hoch	
<i>Reflexions-/Auswertungsfunktion</i>									
keine	M	87,74 % ^A	85,41 %	88,31 % ^A	85,47 %	90,59 % ^A	86,01 %	88,55 % ^A	89,62 %
	N	114	114	114	114	114	114	114	114
	SD	11,07 %	16,75 %	12,17 %	14,66 %	11,97 %	15,45 %	13,17 %	14,04 %
Refl./Ausw. einer Unterrichtsaufgabe	M	88,94 % ^B	83,89 %	90,36 % ^B	86,90 %	92,21 % ^B	93,84 %	87,05 % ^B	87,10 %
	N	13	13	13	13	13	13	13	13
	SD	10,21 %	18,32 %	9,93 %	12,60 %	8,49 %	11,70 %	11,16 %	15,23 %
Refl./Ausw. einer Hausaufgabe	M	76,39 % ^{A, B}	73,30 %	74,62 % ^{A, B}	74,21 %	76,18 % ^{A, B}	81,77 %	69,28 % ^{A, B}	75,86 %
	N	10	10	10	10	10	10	10	10
	SD	13,11 %	20,53 %	14,52 %	17,63 %	13,37 %	11,53 %	17,00 %	27,05 %
Varianzanalyse		F (2;134) = 4,99; p < 0,01; η ² = 0,07	-	F (2;134) = 6,22; p < 0,01; η ² = 0,09	-	F (2;134) = 7,12; p < 0,01; η ² = 0,10	-	F (2;134) = 9,65; p < 0,001; η ² = 0,13	-
Post-Hoc		^A p < 0,01 ^B p = 0,05	-	^A p < 0,01 ^B p < 0,01	-	^A p < 0,01 ^B p < 0,01	-	^A p < 0,001 ^B p < 0,01	-

- Tabellenende -